



UNIVERSIDADE CATÓLICA PORTUGUESA
CENTRO REGIONAL DAS BEIRAS
DEPARTAMENTO DE CIÊNCIAS DA SAÚDE

OralOme – O Contributo dos Microrganismos

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

Por:

Inês Carlos Marques Colaço da Silveira

Julho de 2012



UNIVERSIDADE CATÓLICA PORTUGUESA
CENTRO REGIONAL DAS BEIRAS
DEPARTAMENTO DE CIÊNCIAS DA SAÚDE

OralOme – O Contributo dos Microrganismos

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

Por:

Inês Carlos Marques Colaço da Silveira

Orientadora: Professora Doutora Maria José Correia

Co-Orientador: Professor Doutor Nuno Rosa

Julho de 2012

“Per ardua ad astra”

Lema da *Royal Air Force*

À minha avó,
a única pessoa verdadeiramente altruísta que conheço. Obrigada.

Agradecimentos

À Professora Doutora Maria José Correia,

Pelo apoio incansável, dedicação e disponibilidade, e por ter sido a docente que mais me marcou durante estes cinco anos. Obrigada por toda a sabedoria que transmitiu desde o primeiro ano de curso.

Ao Professor Doutor Nuno Rosa,

por toda a sua colaboração, apoio e disponibilidade, sem a qual teria sido impossível ultrapassar algumas das barreiras mais complicadas na realização deste trabalho.

Aos meus pais,

por terem dado sempre tudo o que podiam e o que não podiam.

Sam,

for making me laugh when everyone else made me cry.

À Ana,

por ter sido uma grande amiga desde aquela aula de biofísica.

À minha binómia,

por me ter aturado tanto. É de louvar a tua paciência.

Aos meus colegas e amigos Marcelo, Fred, Leonor e Valentina,

por terem tornado este 5º ano inesquecível.

Ao Batman e ao Darth Vader,

por terem conservado a minha imaginação e a criança que há em mim.

Resumo:

Na saliva humana podem ser encontradas moléculas de diversas fontes: secreções das glândulas salivares major e minor, moléculas de origem plasmática, do fluido crevicular, da mucosa oral e do microbiota oral. Todas estas moléculas interagem e conferem à saliva as suas propriedades.

De entre as moléculas existentes, as proteínas desempenham um papel muito importante na definição das propriedades da saliva. A identificação de todas as proteínas presentes na cavidade oral, permitirá definir o proteoma oral humano e a caracterização funcional dessas proteínas nos diversos contextos fisiológicos/patológicos, contribuindo para a exploração do OralOme

Está em desenvolvimento uma ferramenta bioinformática que permite armazenar, integrar e visualizar os dados publicados de proteómica da cavidade oral de forma interpretativa. Esta ferramenta é designada por OralCard, e tem como objectivo contribuir para a clarificação dos mecanismos moleculares da biologia oral humana, para a identificação de marcadores moleculares, úteis no diagnóstico de diversas patologias, orais e sistémicas, e para a identificação de novas moléculas que possam ser alvos terapêuticos.

Partindo de uma revisão bibliográfica exaustiva, este estudo tem como objectivo verificar quais as proteínas de origem microbiana experimentalmente determinadas em amostras da cavidade oral, inseri-las na base de dados OralOme, que suporta a ferramenta OralCard, e anotá-las com informação publicada, contribuindo assim para a completa definição do proteoma microbiano da cavidade oral.

Foram adicionadas ao OralOme e manualmente anotadas 1212 proteínas de origem microbiana experimentalmente determinadas e publicadas até Abril de 2012.

Palavras-chave: proteínas microbianas, microbioma oral, OralOme, OralCard, saliva.

Abstract:

In human saliva there are different molecules from diverse sources; secretions from major and minor salivary glands, molecules of plasmatic origin, crevicular fluid, oral mucosa and oral microbiome. All of these components interact and give saliva its properties.

Of the existing molecules, proteins have a very important role in defining saliva's properties. The identification of all the proteins in the oral cavity will allow the definition of the oral proteome and the functional characterization of said proteins in the different physiological/pathological contexts.

A tool that allows the storage, integration and visualization of published data on the proteomics of the oral cavity in a comprehensive way is being developed. This tool is called OralCard, and has the objective of contributing to the clarification of molecular mechanisms of human oral biology, identification of molecular markers useful in the diagnosis of several oral and systemic diseases, and identification of new molecules that can act as therapeutic targets.

Starting with an extensive bibliographical revision, this study aims to verify which are the experimentally determined microbial proteins identified in samples from the oral cavity, insert them in the OralOme data-base supporting OralCard, and annotate them with published information, hence contributing to the complete definition of the microbial proteome of the oral cavity.

From the bibliography reviewed up to April 2012, 1212 experimentally determined proteins of microbial origin were identified, manually annotated and added to OralOme.

Keywords: microbial proteins, oral microbiota, OralOme, OralCard, saliva.

Nota prévia

Neste trabalho surgem por vezes expressões em língua Inglesa, cuja tradução poderia implicar a adulteração do seu significado original. De entre estas expressões, destacamos a descrição dos termos das ontologias. Como é justificado na revisão bibliográfica, e discutido por Pesquita, *et al*, [1], as classificações ontológicas que surgiram na era das ciências ómicas, resultam precisamente da necessidade de haver termos comuns para descrever conceitos biológicos entre laboratórios, culturas e línguas diferentes. Uma tradução dos termos de classificação ontológica poderia implicar uma reinterpretação desses termos, alterando o conceito biológico inicial que se pretende identificar. Assim, optámos por incluir estes termos sem tradução, mas grafados em itálico para facilitar a sua identificação.

Conteúdo

Nota prévia.....	VI
1. Introdução.....	1
1.1. Microbioma Humano	2
1.1.1. Microbioma Oral	4
1.1.1.1. - Bactérias Comensais da cavidade oral	4
1.1.1.2. - Microbiotas diferentes na cavidade oral	5
1.2. Estratégias para a colonização do hospedeiro.....	7
1.2.1. - Formação do biofilme	7
1.2.2. - Quorum Sensing.....	9
1.2.3. - Moléculas importantes na relação microrganismo/hospedeiro	11
1.2.2.1. - Adesinas	11
1.2.2.2. - Proteases bacterianas	12
1.2.2.3. - Proteínas Efectoras	13
1.3 - Ferramentas Bioinformáticas e o estudo do Microbioma Humano	13
1.3.1. - OralOme e OralCard.....	14
1.3.2. - Gene ontology.....	15
2. Objectivos.....	19
3. Material e Métodos.....	21
3.2. Catalogação das proteínas microbianas.....	21
3.3. Análise da classificação ontológica das proteínas microbianas.....	22
3.4. Análise das proteínas microbianas associadas a patologias	26
4. Resultados.....	27
4.2. Catalogação das proteínas microbianas.....	27
4.3. Análise da classificação ontológica das proteínas microbianas.....	29
4.4. Análise das proteínas microbianas associadas a patologias	37
5. Discussão.....	39
6. Conclusão	47
7. Bibliografia	49
8. Índice de Tabelas.....	53
9. Índice de Figuras.....	55

1. Introdução

As propriedades da saliva humana derivam, em grande medida, do conjunto de moléculas que nela existem e da sua interação. Estas moléculas têm origens muito diversas, incluindo assecções das glândulas salivares major e minor, o plasma sanguíneo, o fluido crevicular, a mucosa oral e o microbiota oral.

A identificação de todas as proteínas presentes na cavidade oral permitirá definir o proteoma oral humano. Por sua vez, a caracterização funcional dessas proteínas nos diversos contextos fisiológicos/patológicos permitirá esclarecer o fisioma oral, ou OralOma.

Com o objetivo de esclarecer o OralOma, foi desenvolvida e está em permanente actualização uma ferramenta que permite armazenar, integrar e visualizar resultados de proteómica de forma interpretativa. Esta ferramenta, designada de OralCard[2] (<http://bioinformatics.ua.pt/OralCard/>) assenta na base de dados OralOme. O OralCard, tem como objectivo contribuir para a clarificação da biologia oral humana, para a identificação de marcadores moleculares úteis no diagnóstico de diversas patologias, orais e sistémicas, e para a identificação de novos alvos terapêuticos.

Em Dezembro de 2011, o OralOme incluía 3397 proteínas de origem Humana, identificadas em estudos de proteómica da saliva revistas manualmente, e uma previsão das proteínas de origem bacteriana potencialmente presentes na saliva, previstas a partir dos genomas publicados dos organismos da cavidade oral. Uma vez que havia alguma informação disponível na literatura, com a identificação experimental de proteínas de origem microbiana em amostras da cavidade oral, este trabalho foi proposto para permitir a actualização da base de dados OralOme com resultados experimentais relativos a proteínas de origem microbiana.

1.1. Microbioma Humano

Espalhadas no nosso organismo, existem dez vezes mais células bacterianas que eucariotas. Além de existirem muitos microrganismos espalhados pelo organismo do ser Humano, esta flora é muito variada, mesmo em indivíduos saudáveis. Os locais onde há maior variabilidade de microrganismos, são a cavidade oral e orofaringe, pele, intestinos e vagina[3]. Esta variabilidade está associada aos diversos factores do hospedeiro, que incluem temperatura, humidade, disponibilidade de oxigénio e nutrientes, osmolaridade, receptores do hospedeiro, competição e resistência a outros microrganismos, bem como o sistema imune do hospedeiro [4].

Ao conjunto das comunidades de microrganismos (incluindo bactérias, vírus e eucariotas microbianos) que povoam o organismo Humano, dá-se o nome de Microbiota ou Microbioma Humano[5]. Estas comunidades microbianas são de vital importância para a nossa saúde, e o seu estudo leva a um melhor conhecimento da sua dinâmica complexa, e pode conduzir ao desenvolvimento de novas formas de diagnóstico e mesmo de tratamento [6].

A necessidade de obter informação sobre o microbioma Humano levou à fundação do Human Microbiome Project (HMP), cujo principal objectivo consiste na identificação do microbioma nuclear em pessoas saudáveis. O conceito de microbioma nuclear, assenta, ele próprio, no facto de ser amplamente reconhecido que existe uma grande variabilidade inter-individual, e que portanto não será possível definir “um” microbioma em condições de saúde, mas haverá muitas espécies/géneros de microrganismos que podem ser encontrados em vários indivíduos. No projecto HMP, são usados e desenvolvidos protocolos de metagenómica (a análise de comunidades inteiras de microrganismos que colonizam um determinado habitat, por estudo dos genes ribossomais 16S bacterianos, que distinguem cada organismo pelo seu filo, género e mesmo por espécie [4]) que, aplicados ao microbioma Humano, proporcionam um método estandardizado para criar, processar e interpretar os diferentes tipos de resultados de metagenómica disponíveis na comunidade científica [7]. Para isso, o projecto HMP recorre a vários campos da biologia celular, microbiologia, genética e imunologia[6].

O HMP promove a exploração das comunidades microbianas, e a sua relação com o hospedeiro Humano, estando a ser anotado e analisado actualmente o microbioma Humano em indivíduos saudáveis, que fornece um suporte a outros estudos, que investigam a associação de determinados microbiomas, e das suas dinâmicas ao aparecimento e progressão de patologias. A identificação e ditas associações, pode levar ao desenvolvimento de ferramentas com implicações terapêuticas[7].

Existe uma grande variabilidade de microrganismos que colonizam o ser Humano, sendo a sua diversidade muitas vezes ligada a determinadas patologias, tanto por falta de variabilidade (por exemplo, no caso da doença inflamatória intestinal) ou por excesso (como é o caso da vaginose bacteriana, por exemplo). No entanto, cada local de colonização microbiano no organismo Humano parece ser dominado por um ou alguns *taxa* de microrganismos, embora não existam *taxa* que estejam presentes em todos os habitats do organismo humano. No que diz respeito à variação inter indivíduos, esta parece estar relacionada com factores funcionalmente relevantes, que podem exercer uma pressão selectiva entre espécies. Os *taxa* menos dominantes parecem ser bastante característicos, quer entre indivíduos, quer entre habitats dentro do mesmo organismo[3].

Apesar da grande variabilidade de microrganismos, as diferentes vias metabólicas presentes em cada um dos microbiomas são constantes, ou seja, existe uma “assinatura” ecológica de todo o microbioma Humano com pequenas especificidades para cada microbioma em particular, sendo as vias mais comuns, por exemplo, a maquinaria tradutora no ribossoma, a síntese de ATP e a glicólise, que consistem em vias fundamentais à sobrevivência do microrganismo, bem como à sua interacção com o hospedeiro. Apesar desta constância de vias metabólicas, existem variações que são consequência de alterações funcionais entre indivíduos, incluindo exposições ambientais e dietéticas diferentes, variabilidade de sistemas imunes e genética do hospedeiro[3]. Essas alterações, ao modificar o microbioma, podem levar a disbioses (alterações da composição de determinada comunidade microbiana, associadas a patologias), como é o caso da diabetes ou da doença inflamatória intestinal [4].

1.1.1. Microbioma Oral

Em 2009, o microbioma melhor descrito era o do intestino. No entanto com o projeto HMP outros microbiomas estão a ser investigados neste momento: o nasal, oral, urogenital e a pele [6]. Em 2012, os resultados do projecto HMP revelam que os microbiomas com maior diversidade de microrganismos, são os do intestino e cavidade oral [3]. O objectivo major da investigação, tanto no microbioma Humano geral como no da cavidade oral, é a determinação das espécies presentes, embora também se estejam a abordar a interacção entre o microbioma e o hospedeiro, e a caracterização do comportamento de espécies microbianas de forma individual [6].

Na cavidade oral saudável existem centenas de microrganismos, incluindo bactérias, vírus e fungos, a maioria coexistindo em biofilmes, o que lhes confere uma resistência adicional a stress mecânico e antibioterapia. O modo de vida em biofilme implica uma alteração das expressão genética nos microrganismos, sendo que essa alteração envolve a expressão de genes que, além de estarem associados à formação do biofilme, podem também modelar factores de interacção com o hospedeiro, nomeadamente factores de virulência [8, 9]. Concomitantemente, uma mudança no ambiente ou outra alteração (como por exemplo, a higiene do indivíduo), podem alterar o estado de equilíbrio num biofilme comensal e “ativar” algumas espécies patogénicas oportunistas [6].

1.1.1.1. - Bactérias Comensais da cavidade oral

Teoricamente, calcula-se existirem cerca de 750 diferentes espécies bacterianas comensais na cavidade oral[6], embora a dificuldade de duplicar este ambiente em laboratório (pois muitas espécies não são cultiváveis), e a semelhança entre os genomas dos microrganismos (o que pode não permite a sua distinção através de porções fragmentadas de ADN), fazem suspeitar que ainda não foram identificadas todas as espécies[6].O conhecimento atual permite afirmar que as bactérias que estão representadas em maior número na cavidade oral incluem os géneros *Streptococcus*, *Actinomyces*, *Veillonella*, *Fusobacterium*, *Porphyromonas*, *Prevotella*, *Treponema*, *Neisseria*, *Haemophilus*, *Eubacteria*, *Lactobacterium*, *Capnocytophaga*, *Eiknella*,

Leptotrichia, *Peptostreptococcus*, *Staphylococcus* e *Propionibacterium*, tendo a maioria destes uma relação de mutualismo com o ser Humano[6].

Estas bactérias comensais, ao existir em maior quantidade, não só competem com as patogênicas por nutrientes, como também as impedem de chegar aos tecidos e provocar patologias, constituindo uma barreira física à adesão de microrganismos patogênicos, através do seu número elevado[6].

Alguns destes colonizadores são benéficos mas outros têm efeitos nefastos na fisiologia do hospedeiro. Exemplos de microrganismos que realizam funções importantes para o hospedeiro são algumas bactérias intestinais como a *E. coli* fundamentais na biossíntese de vitamina K₂ [10], essenciais para o ser Humano). Microrganismos causadores de patologias também existem como é o caso de *Streptococcus pneumoniae*, causador de infecções pulmonares em crianças e idosos [11]. Este microbioma normal é conhecido por microbioma comensal, e nela estão incluídas eubactérias, archaeobactérias e fungos [6]. Estas bactérias, quando colonizam outros locais do organismo Humano, podem ter efeitos nefastos (por exemplo, a *E.coli* é uma dos principais agentes etiológicos das infecções urinárias [12]), tornando-se assim um microrganismo virulento, ou patogênico. Existe, no entanto, outra teoria no que diz respeito à patogenicidade bacteriana: as bactérias que causam doenças estão sempre presentes em estado patogênico, embora não consigam prosperar, já que as bactérias comensais são mais abundantes [6]. Além destes 2 tipos de bactérias, existem ainda bactérias que num estado de saúde do hospedeiro não são virulentas, mas quando este está num estado mais debilitado, estas tornam-se patogênicas (por exemplo, a bactéria *Pseudomonas aeruginosa* afecta os pulmões de pacientes com fibrose cística mas não o de indivíduos saudáveis [13])[14].

1.1.1.2. - Microbiotas diferentes na cavidade oral

Apesar de ser conhecida a predominância de *Streptococci*, no biofilme oral, é também sabido que existe uma grande variedade de microbiotas, pois os nichos e superfícies colonizáveis são bastante diferentes: superfície dentária acima do crevículo gengival, superfície dentária abaixo do crevículo gengival, língua, superfícies das mucosas jugais,

próteses dentárias e materiais restauradores. Em Junho de 2012, foram publicados resultados que indicam quais as bactérias com maior predominância em cada um dos nichos da cavidade oral de indivíduos saudáveis[15]. A dominância do género *Streptococcus*, é secundada pelos géneros *Haemophilus* na mucosa jugal, *Actinomyces* na placa bacteriana supra-gengival e *Prevotella* na placa sub-gengiva[3].

As diferenças entre os géneros de microrganismos característicos de cada habitat podem ser explicadas pelas diferentes superfícies dentárias nas quais se deposita o biofilme. Considerando apenas a exposição ao oxigénio, podemos encontrar vários tipos de microrganismos com requisitos de crescimento muito diferentes. Os aeróbios obrigatórios necessitam a presença de oxigénio molecular para o seu crescimento, como é exemplo a bactéria *Mycobacterium tuberculosis*[14]). Os anaeróbios obrigatórios são microrganismos que requerem a ausência de oxigénio para o seu crescimento, pois não têm enzimas para processar este gás e os seus derivados [14], como é exemplo a bactéria *Porphyromonas gingivalis*[16]. Os anaeróbios facultativos, grupo que corresponde à maioria das bactérias orais e que caso esteja presente oxigénio utilizam-no para gerar energia, mas podem sintetizar ATP através da fermentação, caso este gás esteja ausente [17]. Um exemplo de anaeróbio facultativo é o caso de um dos principais agentes etiológicos da cárie dentária o *Streptococcus mutans*[18]. Existem ainda os microrganismos microaerófilos que, em ambientes nos quais o oxigénio está presente em quantidades muito elevadas (ambientes óxicos), não crescem ou crescem pouco. É um exemplo, a bactéria *Helicobacter pylori*[19]. Finalmente há os capnófilos, espécies bacterianas que têm um crescimento óptimo a elevadas quantidades de dióxido de carbono, como é exemplo *Mannheimia succiniciproducens*[20][6].

Uma vez que a anatomia e fisiologia da cavidade oral permite a existência de microhabitats muito variáveis em termos de concentrações de oxigénio, é possível encontrar espécies microbianas com diferentes requisitos deste gás. Esta é com certeza uma das razões que explica a grande variedade de espécies microbianas neste microbioma em particular.

1.2. Estratégias para a colonização do hospedeiro

Para muitos microrganismos, o corpo humano apresenta-se como um “meio de cultura”, que lhes proporciona vários ambientes propícios à sua prosperidade, já que lhes fornece calor, humidade e nutrientes, garantindo assim o seu crescimento e potencializando a sua reprodução[14]. Neste habitat protegido, à semelhança daquilo que acontece com grande parte das bactérias na Natureza, existem comunidades complexas e dinâmicas associadas a superfícies, os biofilmes. Apesar da ideia antiga de que os organismos procaríotas se comportavam como indivíduos auto-suficientes, sabe-se actualmente que dentro as comunidades do biofilme assumem características semelhantes às de organismos multicelulares.

1.2.1. - Formação do biofilme

A formação do biofilme oral é de elevada importância na medicina dentária, já que pode provocar alterações neste sistema e no equilíbrio dinâmico estabelecido como o hospedeiro. As duas das patologias mais prevalentes neste campo da medicina: a cárie dentária e a patologia periodontal têm como facilitadores etiológicos o biofilme oral.

O biofilme mais importante na cavidade oral, é aquele que se deposita sobre as peças dentárias (supra e infra-gengivalmente), e é caracterizado por uma comunidade muito variada (com mais de 700 espécies), e de elevada densidade celular (10^{11} células por grama de peso molhado), o que, juntamente com as flutuações e hostilidades apresentadas pelo meio ambiente, torna as interações intra e inter-espécies inevitáveis e fundamentais para a sobrevivência bacteriana [21]. Estes complexos microbianos além das células bacterianas, são constituídos por uma matriz de glicoproteínas salivares e polissacarídeos extracelulares [22].

A adesão inicial verifica-se entre as moléculas da superfície dos colonizadores primários, e as moléculas de saliva adsorvidas à superfície dentária (película dentária adquirida). Alguns receptores salivares, como é o caso das proteínas ricas em prolina e a estaterina, têm epítomos reconhecidos por *Streptococci*. Outras moléculas como a α -

amilase, imunoglobulinas, fibronectina e lactoferrina, também se unem a adesinas à superfície dos colonizadores iniciais, conferindo-lhes vantagem sobre outras bactérias com receptores com menos afinidade para as moléculas da película dentária [23].

É de referir, que algumas moléculas desta película inicial só expõem determinados domínios, se estiverem adsorvidas à superfície, pois os locais disponíveis para ligação quando as moléculas estão em solução na saliva são diferentes. Inicialmente, a adesão pode ser conseguida através de vários tipos de ligações, incluindo reconhecimento de receptores oligossacarídeos, reacções do tipo lectina, interacções do tipo proteína-proteína, ou ainda ligações iónicas ou hidrofóbicas, conferindo assim uma união bastante firme à superfície que estão a colonizar [23].

Os receptores à superfície das células microbianas constituem um possível alvo terapêutico, já que um possível agente patogénico pode ser excluído do biofilme por não ter receptores adequados à sua adesão, e consequente inserção na comunidade microbiana. No entanto, devem ser cuidadosamente analisados os novos biofilmes a que estas possíveis terapias iriam dar origem, já que no lugar deixado pelo agente patogénico alvo, pode-se instalar um novo microrganismo, com consequências tão ou mais patogénicas que o anterior [23].

O crescimento do biofilme, vai também depender dos colonizadores iniciais. Assim, por exemplo, quando a cavidade oral é colonizada muito precocemente pela bactéria *Streptococcus sanguis*, a colonização pela espécie *Streptococcus mutans* é mais tardia, diminuindo assim a cariogénese. Desta forma, a colonização inicial vai limitar e seleccionar os colonizadores secundários.

As bactérias colonizadoras da cavidade oral, na sua maioria, limitam-se ao espaço extracelular do hospedeiro, embora existam algumas com capacidade de invadir as células do hospedeiro, e instalar-se intra-celularmente (em células gengivais e da mucosa jugal). Um exemplo são os oportunistas *P. gingivalis*[24], *Prevotella intermedia*[25] e *Aggregatibacter actinomycetemcomitans*[6, 26].

1.2.2. - Quorum Sensing

As bactérias que crescem em biofilme têm a capacidade de comunicar entre si e com o hospedeiro através da libertação, resposta e reconhecimento de moléculas sinalizadoras difusíveis. A este sistema dá-se o nome de *quorum sensing* o que significa a regulação controlada pela resposta à densidade de células dentro de uma população.

Este tipo de comunicação traz grandes vantagens para as células bacterianas, incluindo influência na colonização do hospedeiro, formação do biofilme, defesa contra outros microrganismos competidores, adaptação a mudanças ambientais, entre outras funções essenciais à população [21, 27]. Esta comunicação cooperativa facilita a agregação bacteriana entre microrganismos e à superfície do dente, a sinergia nutricional, o crescimento e desenvolvimento celular na saliva, e a formação de cadeias alimentares através de cooperação metabólica [21].

Todos os passos dados na formação do biofilme estão regulados pelo *quorum sensing*, começando na adesão bacteriana à superfície do dente, agregação inter-celular, produção da matriz exopolissacarídea, crescimento, maturação e, finalmente, degradação ou destacamento do biofilme. A regulação da densidade populacional dentro do biofilme, e a actividade metabólica dentro deste, são funções vitais na manutenção da micro-colónia, de acordo com a disponibilidade nutricional e condições ambientais, e este ajuste é regulado também pelas moléculas do *quorum sensing*[27].

As moléculas envolvidas no *quorum sensing*, como é o caso dos auto-indutores, vão interferir na colonização multi-espécies, contribuindo assim para a selecção na constituição e crescimento do biofilme [23]. Dentro dos biofilmes, as bactérias não têm só relações antagonizantes (especialmente quando há competição por nutrientes), mas também de sinergia [28].)

A comunicação entre microrganismos, utilizando o *quorum sensing* é também um elevado factor de virulência reconhecido, pois coordena as oportunidades para minimizar perdas e maximizar a capacidade de infectar determinado habitat. Um exemplo é a repressão da virulência de determinada espécie quando a sua densidade populacional é baixa, para evitar a estimulação do sistema imune do hospedeiro. Quando o número de bactérias aumenta e ultrapassa determinado limiar, a expressão do viruloma é activada e factores de virulência como resistência a antibióticos, mobilidade,

adesinas, citotoxinas e outras toxinas são expressas, permitindo um aumento da virulência e conseqüente aumento da eficácia da colonização [27].

Esta comunicação através do *quorum sensing*, também pode ser tida com o hospedeiro, regulando a transcrição dos genes Humanos de resposta à invasão bacteriana, ou promovendo a apoptose das células do sistema imune. Pode, no entanto, dar-se o acontecimento contrário, e moléculas produzidas pelo hospedeiro (tais como as citocinas Humanas), podem ser utilizadas como moléculas de sinalização no *quorum sensing* bacteriano [27].

As moléculas de *quorum sensing* são fundamentais para um balanço entre a competição e a coexistência entre bactérias dentro de um biofilme, mantendo assim a diversidade e estabilidade de microrganismos em determinado ecossistema microbiano [21].

A comunicação entre microrganismos tem, no entanto, uma característica desfavorável aos microrganismos: a capacidade de interromper uma via de *quorum sensing* pode resultar numa vantagem, num meio em que várias espécies bacterianas competem por quantidades limitadas de recursos [29].

Assim, a comunicação entre uma determinada espécie bacteriana pode ser recebida por outra, que desse modo vai responder de forma a diminuir ou mesmo neutralizar a resposta às moléculas de *quorum sensing*, limitando assim a eficiência desta forma de comunicação [27]. De forma semelhante, a capacidade do hospedeiro interferir com a comunicação entre microrganismos, pode prevenir a sua colonização usando o *quorum sensing* para coordenar a virulência. Aos mecanismos envolvidos na interferência entre a comunicação microbiana, dá-se o nome de *quorum quenching*[29].

As moléculas naturais que fazem o *quorum quenching*, e impedem a colonização patológica de microrganismos, são alvos actuais de investigação como ferramentas terapêuticas. Alguns fármacos actuais já possuem essa capacidade, como é o caso dos macrólidos, e a sua função inibitória da formação de muco em pacientes que sofrem de infecções pulmonares [27].

Assim, a inibição do *quorum sensing* e a utilização do *quorum quenching*, é uma estratégia terapêutica com potencial de substituição de outros fármacos em utilização actual (como os antibióticos), já que da sua aplicação não resultam resistências microbianas. Outra das suas possíveis aplicações é a potencialização da produção

industrial em massa de recursos resultantes da actividade bacteriana (como é o caso da insulina humana obtida pela tecnologia de ADN recombinante) [29].

1.2.3. -Moléculas importantes na relação microrganismo/hospedeiro

As interacções entre as células microbianas e o hospedeiro, são de importância fundamental para a sua sobrevivência e podem determinar se a relação é de mutualismo ou de parasitismo. Existem várias moléculas fundamentais à interacção entre microrganismo e hospedeiro, sendo algumas consideradas factores de virulência (atributos moleculares de determinado microrganismo patogénico utilizado para uma colonização e infecção bem sucedidas) [30]. Estas moléculas podem ser de origem proteica, lipídica ou de hidratos de carbono [31] e podem ser classificadas em dois grupos considerando a fase da colonização em que atuam: i) as moléculas envolvidas na adesão e reconhecimento inicial, nas quais se incluem as adesinas, incluídas nas estruturas externas das bactérias e ii) as moléculas associadas a uma interacção funcional com o hospedeiro onde se podem inserir as toxinas e protéases bacterianas (que não pressupõem necessariamente colonização) e efetores bacterianos que interferem com os mecanismos de sinalização intracelular do hospedeiro e pressupõem um sistema de injeção no espaço intracelular. Muitos destas moléculas são alvos de investigação com o objectivo de fazer o controlo e modulação do desenvolvimento do biofilme oral.

1.2.2.1. - Adesinas

A adesão é um processo fundamental na interacção entre microrganismo e hospedeiro, e pode estar na origem de várias patologias. Microrganismos comensais e patogénicos, expressam moléculas adesivas nas suas superfícies que promovem a interacção com o hospedeiro. No entanto, estas moléculas podem também ser problemáticas para esses mesmos microrganismos, já que são estas que são reconhecidas por células fagocitárias do sistema imune do hospedeiro, estimulando a diapedese, activação e realização da fagocitose. De forma a contornar este problema, alguns microrganismos desenvolveram mecanismos que lhes permitem expressar as adesinas em estruturas poliméricas, que se

estendem da superfície microbiana, fazendo assim com que as interações entre microrganismo e hospedeiro se realizem a uma distância segura da célula do microrganismo [32]. Existem várias estruturas envolvidas na adesão, algumas delas poliméricas, que são as mais estudadas (ex: pili bacterianos e adesinas fimbriais em microrganismos Gram negativos), mas também podem existir também estruturas mais pequenas e monoméricas, mas de igual importância [32].

1.2.2.2. - *Proteases bacterianas*

A função mais descrita das proteases bacterianas, é a sua capacidade de fornecer nutrientes a uma bactéria. No entanto, estas moléculas têm também a capacidade de ser factores de virulência, através de vários mecanismos. As proteases produzidas por uma bactéria extra celular e libertadas no biofilme podem ter uma acção directa sobre os tecidos do hospedeiro, quer destruindo a matriz extracelular, quer activando proteases do hospedeiro, quer inactivando inibidores naturais das proteases do hospedeiro quer ainda inactivando proteínas fundamentais na resposta imune do hospedeiro. As proteases de *Porphyromonas gingivalis* são um bom exemplo de uma protease bacteriana que comprovadamente exerce um efeito de amplo espectro nos tecidos do hospedeiro [33][30].

A maioria das cascatas de proteases humanas são muito bem reguladas, actuam localmente, e são facilmente inactivadas de forma a evitar danos. Além disso, as proteases são normalmente expressas em forma de zimogénios (precursores inactivos que podem ser activados pela acção de outras proteases [34]), sendo as cascatas de proteases são um grupo de sistemas de zimogénios altamente regulados, minimizando assim ao máximo a possibilidade de dano [30]. No entanto, basta um distúrbio na regulação desta cascata para que o organismo se encontre perante uma patologia. Este distúrbio pode ser resultado de uma falha do hospedeiro, ou de intervenção de moléculas de agentes patogénicos, como é o caso das proteases bacterianas. Estas moléculas são extracelulares ou da superfície celular, e podem actuar fazendo a lise das proteínas do hospedeiro, ou activando proteases do mesmo [30].

1.2.2.3. - Proteínas Efectoras

Algumas das bactérias que colonizam o biofilme oral, produzem moléculas que só conseguem ter efeito no hospedeiro quando entram na célula do mesmo. Assim, as bactérias desenvolveram sistemas de secreção proteica, que injectam esses factores de virulência no meio intracelular, com o fim de modificar internamente a fisiologia e/ou morfologia deste último. Essas moléculas denominadas de efectoras diferem das moléculas efectoras do hospedeiro apenas na sua origem, pois a sua forma de atuação e objectivo são idênticos ou seja atuar nas vias de sinalização interna da célula eucariota. As moléculas efectoras podem ter várias consequências benéficas ou prejudiciais para o hospedeiro: inibir a resposta imune (promovendo assim a sua colonização) ou promovê-la (em casos de mutualismo). Os efectores são diferentes segundo o microrganismo que os secreta, mas as suas estratégias para ultrapassar as barreiras de defesa dos hospedeiros são comuns [35].

Recentemente tem sido também verificado que estas moléculas efectoras além de interferirem com moléculas do hospedeiro podem também ser reguladas e interagir com moléculas da própria ou de outras bactérias. Assim as bactérias regulam estes factores de virulência, através da degradação dos efectores por proteases bacterianas, ou através de outros efectores, sendo este efeito conhecido como “friendly-fire” (ou seja, “fogo amigo”, um eufemismo militar). No entanto, os efectores podem ter também um efeito de “equipa”, regulando-se mutuamente [31]. É um exemplo deste tipo de molécula, a proteína efectora EspF da bactéria *E. coli* enteropatogénica, que tem como alvo o nucléolo, e inibe a expressão de um conjunto de factores nucleolares [36].

Estas moléculas são importantes na interacção com o hospedeiro, já que a promoção ou inibição da resposta imune pode ser a origem de várias patologias.

1.3 - Ferramentas Bioinformáticas e o estudo do Microbioma Humano

Para fazer o estudo do Proteoma oral é necessário o recurso a várias ferramentas bioinformáticas, já que após a realização de estudos de proteómica e de metaproteómica, há como resultado um conjunto massivo de dados, que seria impossível de analisar sem

o recurso a estas ferramentas. Assim, para a compreensão do proteoma oral Humano é necessário estudar não só as proteínas orais de origem Humana, mas também as de origem microbiana, já que contribuem também para o OralOma Humano.

Existem várias ferramentas bioinformáticas associadas ao armazenamento e análise de dados relativos a proteínas, tais como: OMIM (*Online Mendelian Inheritance in Man*[37]), UniProtKB (*UniProt KnowledgeBase*[38]), PANTHER (*Protein ANalysis THrough Evolutionary Relationships*[39]), STRING (*Search Tool for the Retrieval of Interacting Genes*[40]), BRENDA (*the Comprehensive Enzyme Information System*[41]) e VFDB (*Virulence Factor DataBase* [42]).

1.3.1. -OralOme e OralCard

Existem ainda algumas bases de dados dedicadas à acumulação e anotação de toda a informação obtida sobre o microbioma Humano, tais como o Human Microbiome Project (HMP, <http://www.hmpdacc.org/>), e mais especificamente sobre o microbioma oral Humano, HOMD (Human Oral Microbiome Database [43]). No entanto, ainda não existia uma ferramenta semelhante, que permitisse armazenar e anotar os resultados dos estudos de proteómica do microbioma oral, nem uma ferramenta que oferecesse aos investigadores um meio de explorar, de forma restrita e, portanto, mais rápida e eficaz os microrganismos, proteínas e doenças associados à cavidade oral. Foi com estas premissas em mente que foi criado o OralCard e a base de dados OralOme que lhe serviu de suporte[44].

Assim, esta ferramenta bioinformática e a informação nela incluída, vai permitir oferecer um recurso para os investigadores e clínicos interessados no estudo da biologia oral, na procura de marcadores moleculares de patologias, no desenvolvimento de testes de diagnóstico e prognóstico, e na descoberta e desenvolvimento de novos agentes terapêuticos[2].

1.3.2. -Gene ontology

Além da existência de bases de dados e de ferramentas bioinformáticas, essenciais para determinar o significado biológico dos vários estudos de genómica e de proteómica, é necessário também desenvolver uma linguagem universal que seja entendida pelos vários laboratórios, e por investigadores de várias origens. Assim, houve a necessidade de criar uma ontologia que permitisse anotar as funções, localização celular e os processos biológicos em que estavam envolvidas as moléculas identificadas e propostas nos vários estudos de proteómica e genómica [45].

O *Gene Ontology* (literalmente, ontologia génica) é um projecto bioinformático que ultrapassa esta barreira, no qual os seus colaboradores desenvolveram vocabulários (ontologias), que descrevem os atributos de produtos génicos provenientes de vários organismos, no contexto da sua função molecular (*Molecular Function*), processo biológico (*Biological Process*) e localização celular (*Cellular Component*)[35].

Este projecto é uma ajuda fundamental na interpretação dos dados experimentais, já que facilita e promove uma consulta uniforme entre os termos GO, permitindo assim uma síntese sistemática de sequências provenientes de vários estudos. O aumento do envolvimento da comunidade científica, e a adição de novos termos GO tem permitido que ocorram avanços em várias áreas, e que haja uma linguagem universal quando nos referimos a produtos génicos[45].

A ontologia *Cellular Component*, refere-se ao conjunto de estruturas intracelulares e complexos macromoleculares, dos quais um determinado produto génico se pode localizar ou fazer parte (como por exemplo: membrana nuclear interna ou o complexo ubiquitina ligase).

A ontologia *Molecular Function*, indica as actividades elementares de um determinado produto génico a nível molecular, como ter função catalítica ou de adesão.

A ontologia *Biological Process*, corresponde a uma série de eventos ou funções moleculares, ou seja, um conjunto de actividades moleculares relacionadas, com um princípio e um fim definidos. São exemplos de *Biological Process* ciclo celular ou a comunicação celular. Ocorrem, por vezes, alterações no processo biológico, causadas por fenótipos mutantes.

Estas ontologias e os termos utilizados, tornam muitas vezes a interpretação confusa. As ontologias GO são, muitas vezes, pouco intuitivas numa primeira análise, dada a complexidade e variedade de antecessores e descendentes que um termo pode ter. Para aliviar este facto, tem havido na comunidade científica alguns grupos com interesses mais específicos, que se têm dedicado a simplificar as ontologias para formas mais abreviadas.

Um desses grupos é o PANTHER [39], que usa uma versão mais simplificada das ontologias dos GO para fazer a anotação das proteínas de organismos tais como *Homo sapiens* e outros organismos modelo como *Mus musculus*, *Rattus norvegicus* ou *Drosophila melanogaster*. No entanto, esta ferramenta tem um número limitado de microrganismos, e nem todos estão presentes na cavidade oral. Assim, apesar de esta ferramenta ser útil para o estudo da cavidade oral no que se refere às proteínas humanas, e portanto estar incluída nas ferramentas a que o OralCard dá acesso, quando queremos analisar as proteínas de origem microbiana, esta ferramenta é muito limitada. Existem várias ferramentas em que há a anotação das ontologias, de forma mais simplificada, de microrganismos. Neste trabalho vamos focar-nos na AgBase, pois é uma base de dados na qual é feita a anotação de vários microrganismos que interagem com hospedeiros vertebrados. É possível encontrar nesta base de dados, a anotação de várias proteínas que são comuns aos microrganismos da cavidade oral Humana, e portanto úteis no presente estudo.

Quando consideramos o caso particular das interacções hospedeiro/microbioma é importante referir que existe um grupo de estudo, o PAMGO (*Plant-Associated Microbe Gene Ontology Initiative*), que está dedicado a estender os GO a incluir termos que descrevam processos envolvidos na interacção entre microrganismos e vários hospedeiros[45]. Este grupo reúne investigadores de várias áreas, mas a sua investigação inicial debruçou-se na simbiose entre bactérias e plantas, sendo essa a origem do nome do grupo.

Um dos passos mais importantes para a criação do termo GO para simbiose foi a sua definição. O PAMGO define simbiose como “um contínuo que abrange interacções que vão desde o mutualismo, passando pelo comensalismo até ao parasitismo”, sendo por isso desencorajada a utilização do termo “simbiose” como um sinónimo de “mutualismo”.

Assim, foram criadas novos termos ontológicos que descrevem as relações inter-espécies como o GO:0044403 “*symbiosis, encompassing mutualism through parasitism*”, que tem 20 termos descendentes[45]:

- GO:0008150 biological process [447337 gene products]
 - GO:0051704 multi-organism process [17615 gene products]
 - GO:0044419 interspecies interaction between organisms [4950 gene products]
 - ▼ GO:0044403 symbiosis, encompassing mutualism through parasitism [4328 gene products]
 - Ⓟ GO:0051816 acquisition of nutrients from other organism during symbiotic interaction [12 gene products]
 - Ⓟ GO:0051825 adhesion to other organism involved in symbiotic interaction [298 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 [67 gene products]
 - Ⓟ GO:0051821 dissemination or transmission of organism from other organism involved in symbiotic interaction [7 gene products]
 - Ⓟ GO:0044110 growth involved in symbiotic interaction [337 gene products]
 - Ⓟ GO:0051701 interaction with host [3343 gene products]
 - Ⓟ GO:0052047 interaction with other organism via secreted substance involved in symbiotic interaction [244 gene products]
 - Ⓟ GO:0051702 interaction with symbiont [238 gene products]
 - Ⓟ GO:0051708 intracellular protein transport in other organism involved in symbiotic interaction [35 gene products]
 - Ⓟ GO:0052214 metabolism of substance in other organism involved in symbiotic interaction [9 gene products]
 - Ⓟ GO:0051817 modification of morphology or physiology of other organism involved in symbiotic interaction [1223 gene products]
 - Ⓟ GO:0052192 movement in environment of other organism involved in symbiotic interaction [1100 gene products]
 - GO:0044399 multi-species biofilm formation [10 gene products]
 - GO:0085030 mutualism [0 gene products]
 - GO:0009877 nodulation [24 gene products]
 - GO:0072519 parasitism [2 gene products]
 - Ⓟ GO:0051824 recognition of other organism involved in symbiotic interaction [5 gene products]
 - GO:0043903 regulation of symbiosis, encompassing mutualism through parasitism [206 gene products]
 - Ⓟ GO:0052173 response to defenses of other organism involved in symbiotic interaction [1181 gene products]
 - Ⓟ GO:0051836 translocation of molecules into other organism involved in symbiotic interaction [0 gene products]

Figura 1 - Vista em árvore dos termos antecedentes e descendentes do “GO:0044403 symbiosis encompassing mutualism through parasitism”, adaptado da vista em árvore de <http://amigo.geneontology.org/> a 28/06/2012.

Não são apenas os processos biológicos que têm sido alvo da criação de novas ontologias, também as estruturas específicas que se criam durante o processo de simbiose (como caveolas, membranas e vacúolos, por exemplo) foram também anotados com novos códigos GO dentro da classificação *Cellular Component*. O termo GO:0033643 “host cell”, é um exemplo de estruturas importantes na relação entre o hospedeiro e o microbioma, e inclui 16 termos descendentes[45]:

2. Objectivos

O objectivo deste trabalho, é completar a base de dados OralOme com as proteínas produzidas pelos microrganismos, e identificadas experimentalmente em estudos de proteómica. Com este trabalho aumentar-se-á o número e tipo de proteínas na base de dados e portanto a capacidade de utilização da ferramenta OralCard (<http://bioinformatics.ua.pt/OralCard/>).

Após a concretização do objectivo anterior, tornou-se possível a análise da informação recolhida, concretizando a proposta de um conjunto de proteínas de origem microbiana interessantes do ponto de vista da interacção que estabelecem com o hospedeiro.

3. Material e Métodos

3.2. Catalogação das proteínas microbianas

Foi feita uma pesquisa bibliográfica no motor de busca PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>) usando como palavras-chave “saliva, proteomics, bacteria”, “saliva, proteomics, microbial” e “saliva proteomics”.

Foram seleccionados artigos, com resultados experimentais de estudos de proteómica da saliva, em que foram encontradas proteínas microbianas. [46-49]

Usou-se a ferramenta bioinformática UniProtKB (UniProt Knowledgebase, <http://www.uniprot.org>[38]), uma base de dados de proteínas que apresenta a informação estrutural e funcional sobre cada proteína, incluindo os seus diferentes nomes, estados de revisão, organismo onde foram identificadas, domínios, estrutura tridimensional, interacções, entre outros.

A partir do número de acesso de cada uma das proteínas microbianas identificadas na bibliografia, foi usada a interface da base de dados do UniProtKB para verificar o estado de revisão das mesmas: *Reviewed*, que corresponde a proteínas com registos anotados manualmente através de informação obtida por literatura, ou *Unreviewed*, que corresponde a registos analisados computacionalmente por algoritmos, aguardando ainda um registo manual. Com este recurso, foi também possível verificar se já existe uma determinação da estrutura tridimensional das proteínas em questão, informação fundamental para um estudo posterior sobre a interacção molecular destas proteínas com o hospedeiro na cavidade oral.

As proteínas microbianas identificadas nas várias referências, foram adicionadas à base de dados OralOme. Esta base de dados, que reúne informação acerca de proteínas salivares humanas identificadas em estudos de proteómica contínua, em Dezembro de 2011, proteínas de origem microbiana cuja presença na cavidade oral foi prevista teoricamente a partir da sequência do genoma das bactérias da cavidade oral [depositadas na *Human Oral Microbiome Database* (<http://www.homd.org>)] [43]. Simultaneamente, as proteínas determinadas experimentalmente foram anotadas manualmente, relativamente ao tipo de amostra em que foram identificadas, o género, idade e estado de saúde do dador, técnicas usadas

para a identificação das proteínas e existência de modificações pós-tradução das mesmas. A anotação da patologia dos dados de amostras orais, foi feita recorrendo ao código OMIM [*Online Mendelian Inheritance in Man* (<http://www.ncbi.nlm.nih.gov/omim>)[37]], mas como algumas situações de alteração do sistema oral como as cáries e a Gengivite, não estavam incluídas na base do OMIM, foram também usadas as designações das alterações descritas no MeSH (Medical Subject Headings <http://www.nlm.nih.gov/mesh/>). Para cada uma das proteínas identificadas como alteradas, no caso de patologia, foi ainda registado se as proteínas identificadas estavam aumentadas, diminuídas ou ausentes.

3.3. Análise da classificação ontológica das proteínas microbianas

Para se obterem as ontologias referentes a cada uma das 1212 proteínas identificadas, foi utilizada a ferramenta bioinformática UniProtKB, inserindo os códigos de acesso correspondentes às proteínas identificadas. Embora nalguns dos artigos originais as proteínas apareçam listadas com códigos IPI, a ferramenta de “*mapping*” no UniProtKB permite fazer a sua conversão para os códigos próprios desta base de dados. Assim, este procedimento foi adoptado para todas as proteínas, de forma aos identificadores armazenados no OralOme sejam os códigos do UniProtKB.

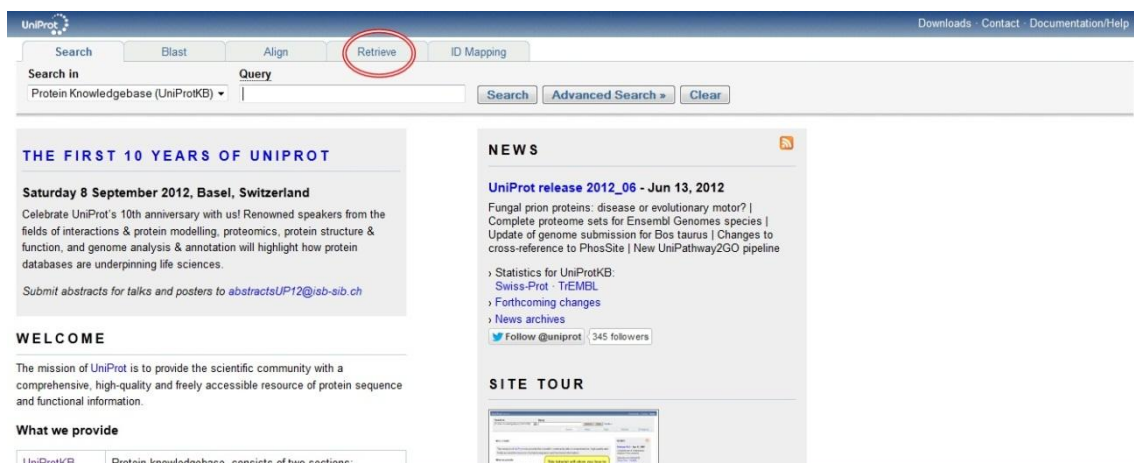


Figura 3 - Obtenção de informação sobre as 1212 proteínas identificadas: utilização da opção “Retrieve” no portal web do UniProtKB. Figura adaptada do portal Web do UniProtKB [29]



Figura 4 - Obtenção de informação sobre as 1212 proteínas identificadas: colocação dos identificadores de todas as proteínas e selecção da opção “Retrieve”. Figura adaptada do portal Web do UniProtKB[29]

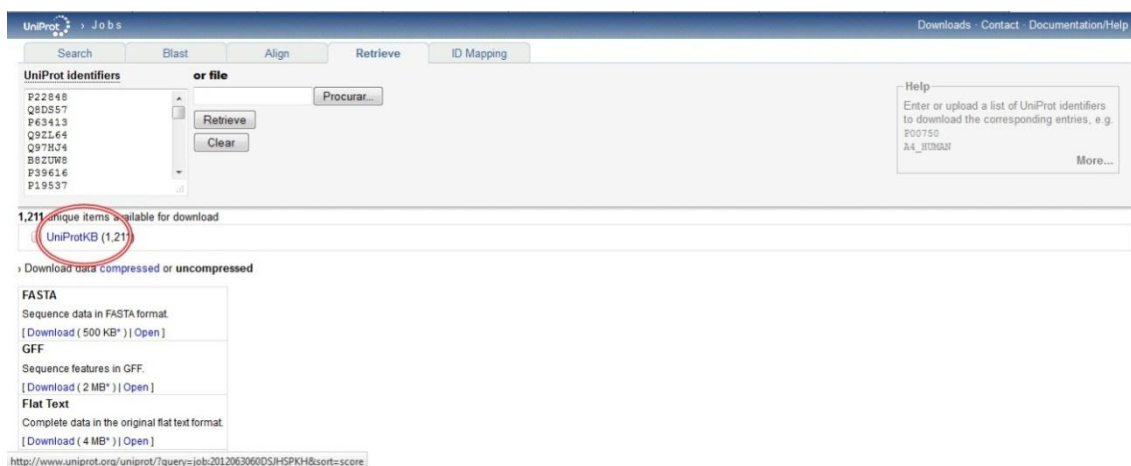


Figura 5 - Obtenção de informação sobre as 1212 proteínas identificadas: selecção da opção “UniProtKB. Figura adaptada do portal Web do UniProtKB[29]

De seguida, foram seleccionadas quais as características de interesse ao estudo das proteínas, tendo sido escolhidos o nome da proteína, função, microrganismo da qual origina, identificação da ontologia e descrição da ontologia.

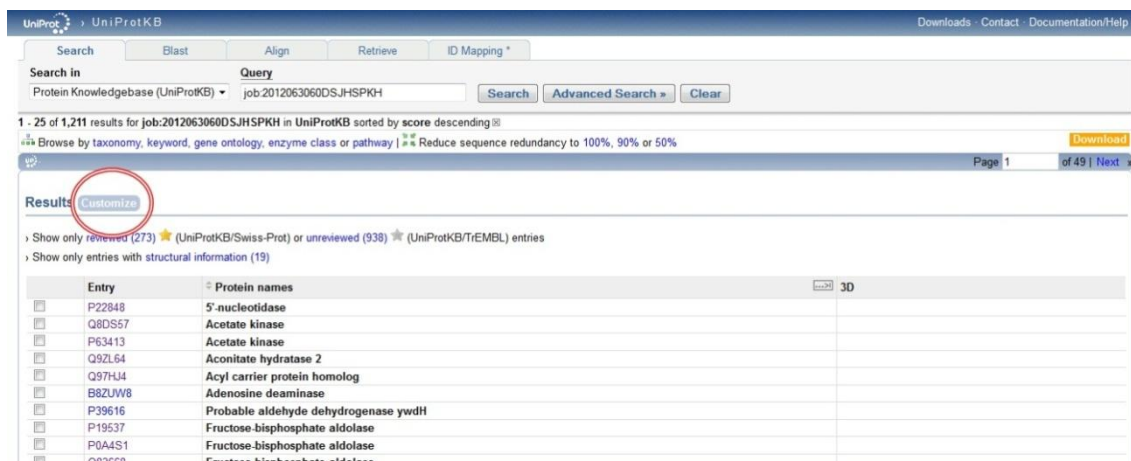


Figura 6 - Obtenção de informação sobre as 1212 proteínas identificadas: selecção da opção "Customize" para escolher as informações mais relevantes das proteínas para este estudo. Figura adaptada do portal Web do UniProtKB[29]

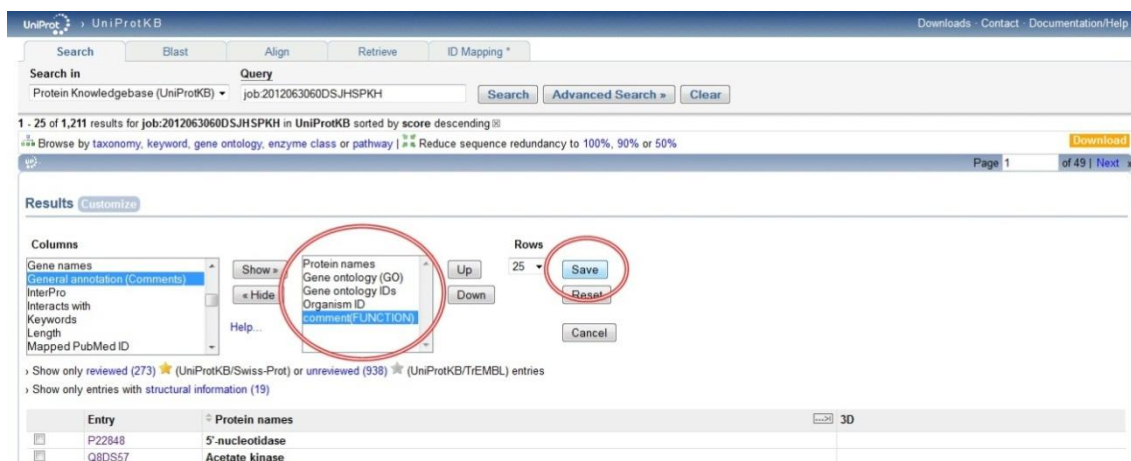


Figura 7 - Obtenção de informação sobre as 1212 proteínas identificadas (nome das proteínas, GOs, identificação das ontologias, organismo de origem e função da proteína) Seleccionou-se de seguida a opção "Save". Figura adaptada do portal Web do UniProtKB[29]

Após as informações de interesse terem sido seleccionadas, foi feito o download de um ficheiro *.xlsx* a tabela foi trabalhada no Microsoft Excel® 2010.

UniProtKB Search results for job:2012063060DSJHSPKH. The interface shows search options (Search, Blast, Align, Retrieve, ID Mapping) and a search bar with the query 'job:2012063060DSJHSPKH'. Below the search bar, it indicates '1 - 25 of 1,211 results for job:2012063060DSJHSPKH in UniProtKB sorted by score descending'. A 'Download' button is circled in red. The results table is as follows:

Entry	Protein names	Gene ontology (GO)	Gene ontology IDs	Organism ID	General annotation (FUNCTION)
P22848	5'-nucleotidase	5-nucleotidase activity; cell outer membrane; metal ion binding; nucleotide binding; nucleotide catabolic process; plasma membrane	GO:0008253; GO:0009279; GO:0046872; GO:0001166; GO:0009166; GO:0005886	670	Degradation of extracellular 5-nucleotides for nutritional needs.
Q8DS57	Acetate kinase	ATP binding; acetate kinase activity; cytoplasm; organic acid metabolic process	GO:0005524; GO:0008776; GO:0005737; GO:0006082	1309	
P63413	Acetate kinase	ATP binding; acetate kinase activity; cytoplasm; organic acid metabolic process	GO:0005524; GO:0008776; GO:0005737; GO:0006082	1313	
Q9ZL64	Aconitate hydratase 2	2-methylisocitrate dehydratase activity; 4 iron, 4 sulfur cluster binding; citrate hydro-lyase (cis-aconitate-forming) activity; isocitrate hydro-lyase (cis-aconitate-forming) activity; metal ion binding; tricarboxylic acid cycle	GO:0047456; GO:0051539; GO:0052632; GO:0052633; GO:0046872; GO:0006099	85963	Catalyzes the isomerization of citrate to isocitrate via cis-aconitate as well as the dehydration of 2-methylisocitrate to cis-2-methylaconitate (By similarity).

Figura 8 - Obtenção de informação sobre as 1212 proteínas identificadas: foi seleccionada a opção “Download” para descarregar o ficheiro. Figura adaptada do portal Web do UniProtKB[29]

UniProtKB search results for job:2012063060DSJHSPKH. The interface shows search options and a search bar with the query 'job:2012063060DSJHSPKH'. Below the search bar, it indicates '1,211 results for job:2012063060DSJHSPKH in UniProtKB sorted by score descending'. A 'Download data compressed or uncompressed' section is visible. The 'Excel' option is circled in red. The download options are as follows:

- Tab-Delimited: Summary information from the result view. [Download] [Open] [Open first 10]
- Excel: Summary information from the result view for MS Excel™. [Download] [Open] [Open first 10]
- FASTA: Canonical sequence data in FASTA format. [Download (600 KB*)] [Open] [Open first 10]
- FASTA: Canonical and isoform sequence data in FASTA format. [Download (600 KB*)] [Open] [Open first 10]
- GFF: Sequence annotation in GFF format. [Download (2 MB*)] [Open] [Open first 10]
- Flat Text: Complete data in the original flat text format.

URL: [http://www.uniprot.org/uniprot/?query=job:2012063060DSJH-id,protein.names,go,go-id,organism-id,comment\(FUNCTION\)](http://www.uniprot.org/uniprot/?query=job:2012063060DSJH-id,protein.names,go,go-id,organism-id,comment(FUNCTION))

Figura 9 - Obtenção de informação sobre as 1212 proteínas identificadas: escolha do formato a descarregar. Foi seleccionado o formato .xlsx para ser possível trabalhar com o Microsoft Excel. Figura adaptada do portal Web do UniProtKB[29]

Para realizar a descrição compreensiva das anotações ontológicas das 1212 proteínas microbianas inseridas no OralOme, essas anotações foram agrupadas em 3 grupos de ontologias: *Cellular Component*, *Molecular Function* e *Biological Process*. Devido à grande variedade de ontologias presentes, foi necessário contornar este problema através do agrupamento das várias ontologias em classes mais abrangentes, sendo para isso utilizada a ferramenta bioinformática AgBase[50], que apesar de ser desenvolvida para a comunidade agro-pecuária contém a anotação de muitos dos microrganismos presentes na cavidade oral.

De forma a encontrar as proteínas de interesse para a interacção com o hospedeiro, foi feita uma pesquisa no ficheiro contendo a informação descrita na Figura 7, com as palavras-chave “*adhesin*”, “*adhesion*”, “*pili*”, “*effector*”, “*bacterial protease*”, “*quorum sensing*”, “*quorum quenching*”, “*virulence*” e “*toxin*”.

3.4. Análise das proteínas microbianas associadas a patologias

O OralCard foi utilizado para fazer a associação entre as proteínas de origem microbiana alteradas em situações de patologia. Das várias patologias inseridas na base de dados OralOme apenas Gengivite, Periodontite Agressiva e Cancro Oral apresentavam proteínas microbianas relacionadas. Para cada uma delas foram identificadas todas as ontologias existentes, e seleccionadas as que teriam maior relevância na interacção com o hospedeiro, de forma idêntica à descrita no parágrafo anterior.

Todas as tabelas e gráficos apresentados foram realizados com o programa Microsoft Excel® 2010.

4. Resultados

4.2. Catalogação das proteínas microbianas

Após pesquisa bibliográfica na base de dados online PubMed (<http://www.ncbi.nlm.nih.gov/pubmed/>), com as palavras-chave “*bacteria*”, “*microbial*”, “*saliva*” e “*proteomics*”, foram encontrados artigos científicos, dentre os quais foram seleccionados cinco, que apresentavam nos seus resultados a identificação de proteínas microbianas em amostras de saliva, fluido crevicular, ou mucosa oral (em pacientes com patologia e/ou pacientes saudáveis).

Assim, os artigos seleccionados foram:

Tabela I - Artigos de estudos de proteómica salivar seleccionados para este estudo.

Artigo	Número de proteínas identificadas	Patologia associada (OMIM/MeSH)	Local de origem da amostra
Bostanci, N. et al., 2010[48]	28 proteínas	Periodontite agressiva generalizada 170650/68010520	Fluido Crevicular
Esser, D. et al., 2008[47]	25 proteínas	NA/	Sobrenadante obtido da centrifugação de Saliva
Grant, M. et al., 2010[49]	16 proteínas	Gengivite NA/68005891	Fluido Crevicular
Xie, H. et al., 2008[46]	136 proteínas	Cancro Oral 275355/68009062	Pellet obtido da centrifugação de Saliva
Jagtap, P. et al., 2012 [51]	2176 proteínas	NA/	Sobrenadante obtido da centrifugação de Saliva

A partir desses artigos, foram reunidas todas as proteínas identificadas nos diferentes estudos e verificado que das 1212 proteínas identificadas, na base de dados UniProtKB, havia 273 destas proteínas revistas e 939 não revistas, ou seja, cuja existência e nível de anotação, não foi ainda revista manualmente pelos curadores da base de dados UniProtKB.

Ainda com recurso ao UniProtKB foi verificado que para 1193 destas proteínas não existia estrutura 3D determinada, e que para apenas 15 proteínas essa estrutura existia determinada por cristalografia de RX, para 2 delas existia estrutura determinada por microscopia electrónica e cristalografia de RX, para uma delas existia estrutura determinada por cristalografia de RX e espectroscopia NMR e finalmente para uma existia estrutura determinada por espectroscopia NMR.

As 1212 proteínas experimentalmente determinadas revelam a presença de 170 organismos, sendo a maioria destes bactérias, uma levedura (*Saccharomyces cerevisiae*, com as proteínas P0CF16 e P0CF17) e dois vírus (*Equine herpesvirus 1*, com a proteína P28966 e *Encephalomyocarditis virus*, com a proteína P17593 e P17594).

A maioria das espécies bacterianas é identificada por poucas proteínas (uma ou duas), sendo a espécie *Rothia mucilaginosa* aquela em que foram identificadas mais proteínas, como se pode observar na Figura 10.

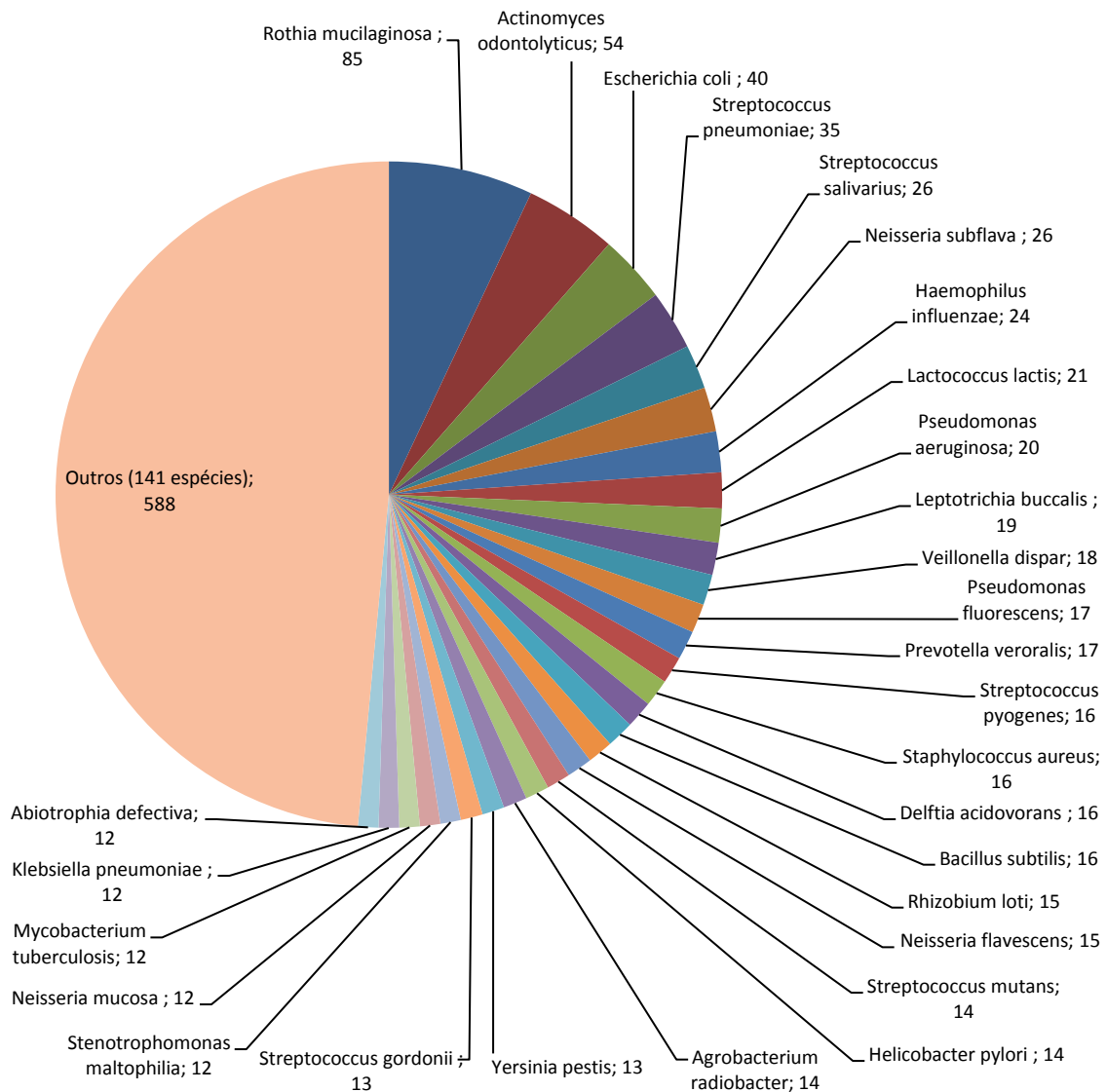


Figura 10- Representação do número de proteínas de origem microbiana presentes no OralCard por espécie de microrganismo. No grupo “Outros” estão incluídas 141 espécies cuja identificação resulta da presença de 11 ou menos proteínas. A correspondência entre a proteína identificada e o microrganismo foi obtida com o UniProtKB.

4.3. Análise da classificação ontológica das proteínas microbianas

No sentido de esclarecer a importância funcional das proteínas microbianas presentes no ecossistema da cavidade oral, foi determinada a anotação ontológica de cada uma das 1212 proteínas. Assim, foram agrupadas de acordo com as três ontologias principais: *Cellular Component*, *Molecular Function* e *Biological Process* representadas nas figuras seguintes:

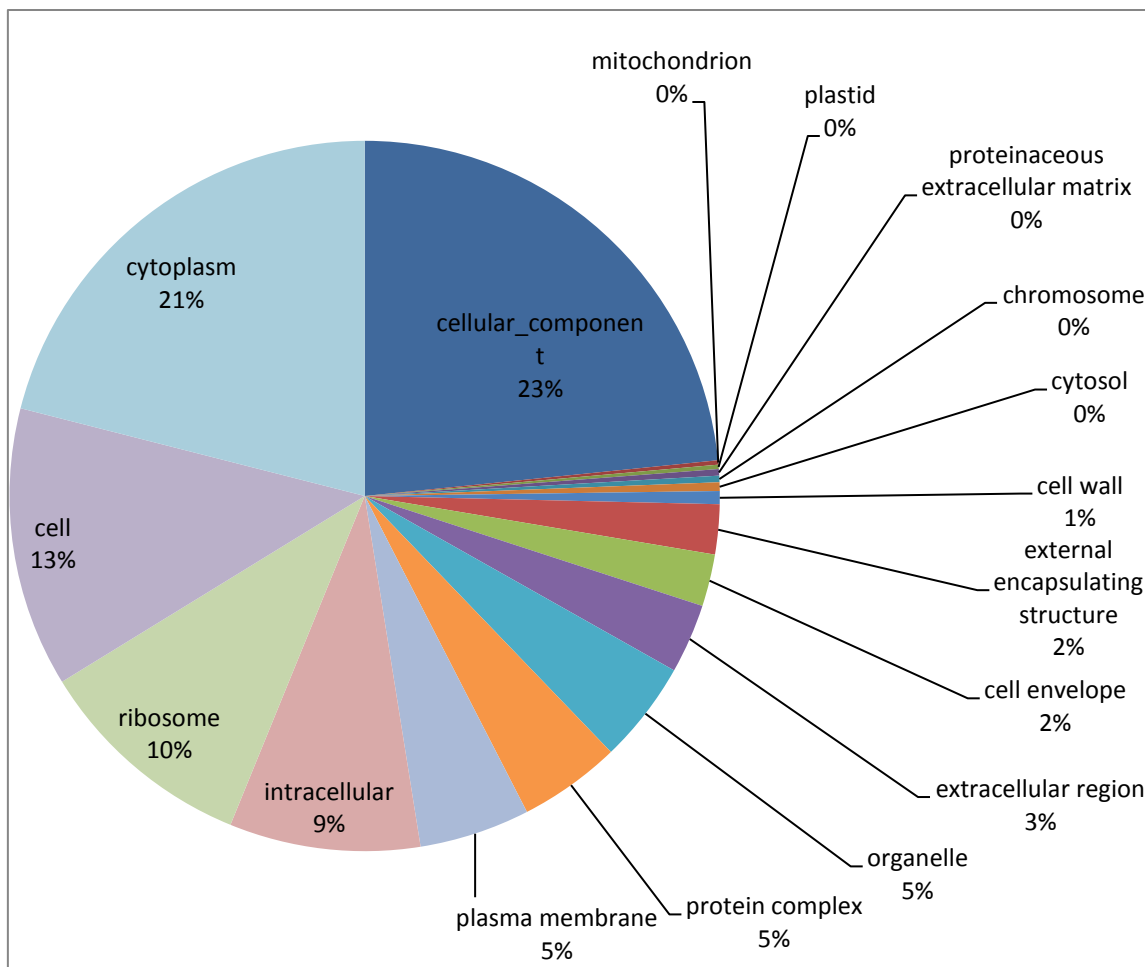


Figura 11 - Representação da classificação ontológica das 1212 proteínas de origem microbiana presentes no OralOme quanto à classe Cellular Component. As classes ontológicas mantêm a designação em inglês pelas razões apresentadas na nota prévia do prefácio.

Através da análise da Figura 11, que ilustra a *Cellular Component*, pode-se concluir que a maioria das ontologias analisadas corresponde a produtos génicos localizados no citoplasma (21% correspondem ao GO:0005737, *cytoplasm*) e no ribossoma (10,1% correspondem ao GO:0005840, *ribosome*). Existe também um número considerável de produtos génicos localizados na membrana plasmática (5% correspondem ao GO:0005886, *plasma membrane*), organelos (4,6% correspondentes ao GO:0043226, *organelle*) e complexos proteicos (4,6% correspondentes ao GO:0043234, *protein complex*), região extra-celular (3,2% correspondentes ao GO:0005576, *extracellular region*), envelope celular (2,4% correspondentes ao GO:0030313, *cell envelope*) e estrutura encapsuladora externa (2,3% correspondentes ao GO:0030312, *external encapsulating structure*). As ontologias representadas em menor número são as de produtos génicos encontrados na parede celular (0,6% correspondentes ao GO:0005618,

cell wall), citosol (0,4% correspondentes ao GO:0005829, *cytosol*), cromossoma (0,3% correspondentes ao GO:0005694, *chromosome*), matriz extra-celular proteínácea (0,3% correspondentes ao GO:0005578, *proteinaceous extracellular matrix*), plastídeo (0,2% correspondentes ao GO:0009536, *plastid*) e mitocôndria (0,2% correspondentes ao GO:0005739, *mitochondrion*).

Para facilitar a compreensão e análise do gráfico que ilustra as ontologias correspondentes a Cellular Component, estas foram agrupadas em categorias mais abrangentes através do recurso à ferramenta bioinformática AgBase, como foi descrito anteriormente na secção de Material e Métodos, obtendo-se assim 3 grupos de ontologias: i) intracelulares, o maior, contendo 50,1% de todos os produtos génicos; ii) extracelulares, que apesar de serem o grupo com menos produtos génicos (13,7%), correspondem ao mais relevante para a interacção com o hospedeiro; iii) componente celular indefinida, com 36,1% dos produtos génicos.

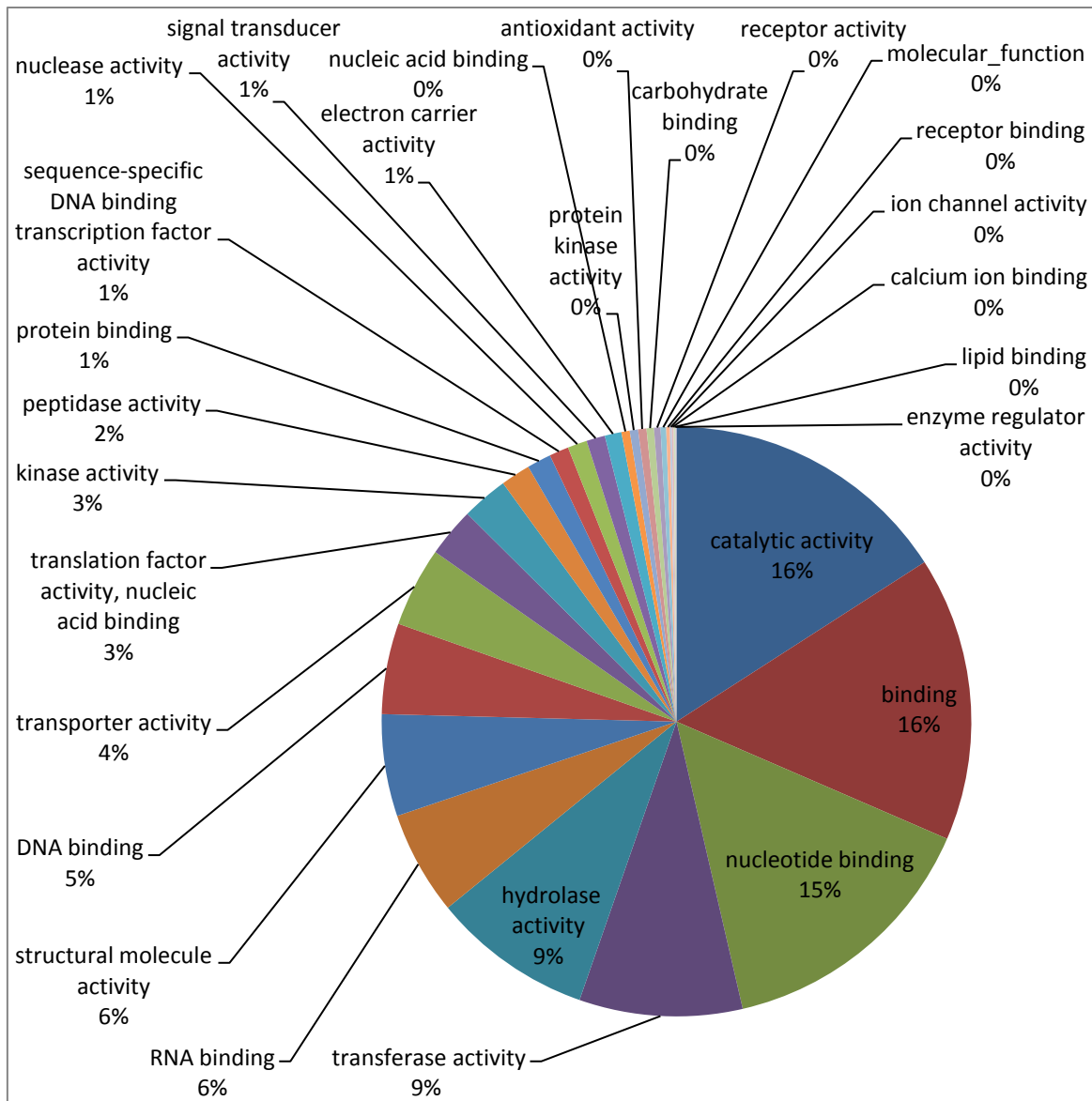


Figura 12 - Representação da classificação ontológica das 1212 proteínas de origem microbiana presentes no OralOme quanto à classe Molecular Function. As classes ontológicas mantêm a designação em inglês pelas razões apresentadas na nota prévia do prefácio.

Na Figura 12, pode verificar-se que os produtos génicos com maior representação na amostra estão envolvidos em actividade catalítica (15,9% correspondentes ao GO:0003824, *catalytic activity*), de ligação (15,6% correspondentes ao GO:0005488, *binding*), ligação de nucleótidos (14,9% correspondentes ao GO:0000166, *nucleotide binding*), actividade de transferase (8,9% correspondentes a GO:0016740, *transferase activity*), actividade de hidrolase (8,8% correspondentes ao GO:0016787, *hydrolase activity*), ligação a ARN (5,7% correspondentes ao GO:0003723, *RNA binding*) e

atividade molecular estrutural (5,6% correspondentes ao GO:0005198, *structural molecule activity*). Existem também uma quantidade relevante de produtos génicos correspondentes a ligação a ADN (5% correspondentes ao GO:0003677, *DNA binding*), atividade de transporte (4,4% correspondentes ao GO:0005215, *transporter activity*), atividade de factor de tradução, ligação a ácido nucleico (2,7% correspondentes ao GO:0008135, *translation factor activity, nucleic acid binding*), atividade de cinase (2,5% correspondentes ao GO:0016301, *kinase activity*), atividade de peptidase (1,6% correspondentes ao GO:0008233, *peptidase activity*), ligação proteica (1,3% correspondentes ao GO:0005515, *protein binding*), atividade de nuclease (1,1% correspondentes ao GO:0004518, *nuclease activity*), ligação a sequências específicas de ADN (1,1% correspondentes ao GO:0003700, *sequence-specific DNA binding*), atividade de transdução de sinal (1% correspondente ao GO:0004871, *signal transducer activity*) e atividade de transporte de electrões (0,9% correspondentes ao GO:0009055, *electron carrier activity*). Em quantidades menores, existem produtos génicos de atividade anti-oxidante (0,5% correspondentes ao GO:0016209, *antioxidant activity*), atividade de cinase de proteínas (0,5% correspondentes ao GO:0004672, *protein kinase activity*), ligação a ácidos nucleicos (0,5% correspondentes ao GO:0003676, *nucleic acid binding*), ligação a hidratos de carbono (0,4% correspondentes ao GO:0030246, *carbohydrate binding*), atividade de receptor (0,4% correspondentes ao GO:0004872, *receptor activity*), ligação a receptores (0,2% correspondentes ao GO:0005102, *receptor binding*), ligação a lípidos (0,1% correspondentes ao GO:0008289, *lipid binding*), ligação ao ião cálcio (0,1% correspondentes ao GO:0005509, *calcium ion binding*), atividade de canal iónico (0,1% correspondentes ao GO:0005216, *ion channel activity*) e atividade de regulação enzimática (0,05% correspondentes ao GO:0030234, *enzyme regulator activity*).

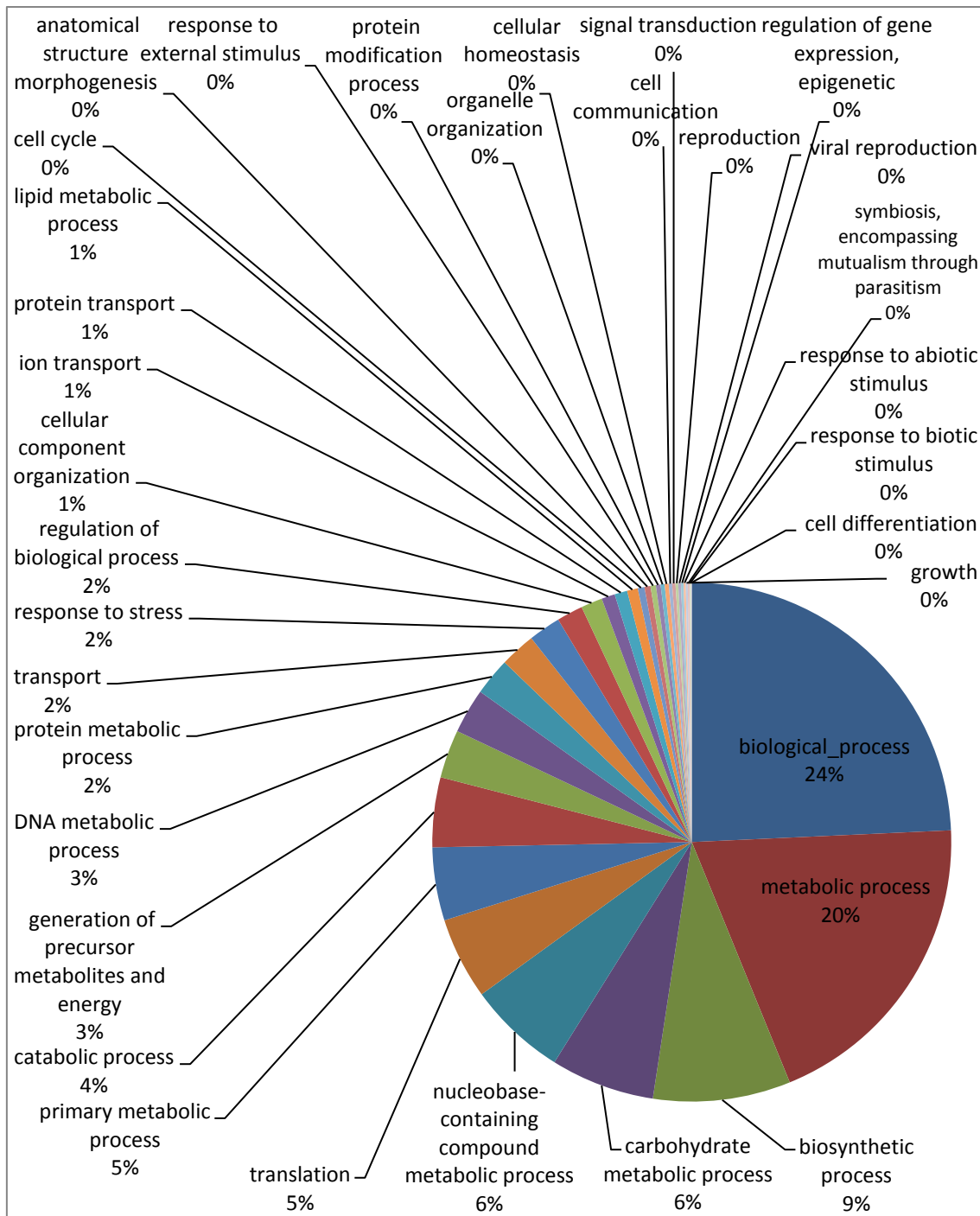


Figura 103 - Representação da classificação ontológica das 1212 proteínas de origem microbiana presentes no OralOme quanto à classe Biological Process. As classes ontológicas mantêm a designação em inglês pelas razões apresentadas na nota prévia do prefácio.

Quanto à classe ontológica *Biological Process* (Figura 13), conclui-se que predominam os processos correspondentes ao processo metabólico (19,6% correspondentes ao GO:0008152, *metabolic process*), processo biossintético (8,6% correspondentes ao GO:0009058, *biosynthetic process*), processo metabólico de hidratos de carbono (6,5% correspondentes ao GO:0005975, *carbohydrate metabolic process*), compostos que contêm uma nucleobase (6,1% correspondentes ao GO:0006139, *nucleobase-containing compound*), tradução (5,1% correspondentes ao GO:0006412, *translation*), processo metabólico primário (4,6% correspondentes ao GO:0044238, *primary metabolic process*) e processo catabólico (4,3% correspondentes ao GO:0009056, *catabolic process*). Em quantidades consideráveis foram encontrados produtos génicos com ontologias correspondentes à geração de metabolitos precursores e energia (3% correspondentes ao GO:0006091, *generation of precursor metabolites and energy*), processo metabólico de ADN (2,8% correspondentes ao GO:0006259, *DNA metabolic process*), processo metabólico de proteínas (2,3% correspondentes ao GO:0019538, *protein metabolic process*), transporte (2,2% correspondentes ao GO:0006810, *transport*), resposta a stress (2% correspondentes ao GO:0006950, *response to stress*), regulação do processo biológico (1,6% correspondentes ao GO:0050789, *regulation of biological process*), organização da componente celular (1,3% correspondentes ao GO:0016043, *cellular component organization*), transporte iónico (0,8% correspondentes ao GO:0006811, *ion transport*), transporte proteico (0,8% correspondentes ao GO:0015031, *protein transport*), processo metabólico de lípidos (0,7% correspondentes ao GO:0006629, *lipid metabolic process*) e ciclo celular (0,4% correspondentes ao GO:0007049, *cell cycle*). Em quantidades menores encontram-se representados os produtos génicos das ontologias de morfogénese da estrutura anatómica (0,4% correspondentes ao GO:0009653, *anatomical structure morphogenesis*), resposta a estímulos externos (0,3% correspondentes ao GO:0009605, *response to external stimulus*), processo de modificação de proteínas (0,3% correspondentes ao GO:0006464, *protein modification process*), homeostase celular (0,3% correspondentes ao GO:0019725, *cellular homeostasis*), organização de organelos (0,3% correspondentes ao GO:0006996, *organelle organization*), transdução de sinal (0,2% correspondentes ao GO:0007165, *signal transduction*), comunicação celular (0,2% correspondentes ao GO:0007154, *cell communication*), regulação da expressão génica, epigenética (0,2% correspondentes ao GO:0040029, *regulation of gene expression, epigenetic*), reprodução viral (0,2% correspondentes ao GO:0016032, *viral*

reproduction), reprodução (0,2% correspondentes ao GO:0000003, *reproduction*), simbiose e mutualismo abrangente através de parasitismo (0,1% correspondentes ao GO:0044403, *symbiosis, encompassing mutualismo through parasitism*), crescimento (0,1% correspondentes ao GO:0040007, *growth*), resposta a estímulos abióticos (0,1% correspondentes ao GO:0009607, *response to biotic stimulus*) e diferenciação celular (0,04% correspondentes ao GO:0030154, *cell differentiation*).

Através da lista que continha as 1212 proteínas, obtidas nas amostras com o procedimento descrito anteriormente na secção de material e métodos, foram seleccionadas aquelas que apresentavam maior potencial de interesse na interacção entre microrganismo e hospedeiro, sendo apresentadas na Tabela I todas as que apresentavam os factores de virulência referidos anteriormente: moléculas de adesão, proteínas efectoras e proteases bacterianas. Das 1212 proteínas microbianas identificadas, existem 6 que se enquadram neste grupo, existindo 3 toxinas, 2 proteínas envolvidas na adesão e 1 proteína efectora.

Tabela II - Proteínas microbianas inseridas no OralOme com propriedades de interesse na interação microrganismo hospedeiro. A informação relativa à anotação ontológica e descrição da função resultam da análise feita com as ferramentas do UniProtKB.

Código	Proteína	Microrganismo	Função	GO's
P35645	<i>Fimbrial protein ecpA (Pilin)</i>	<u>Eikenella corrodens</u>	Adesão celular; pilus	GO:0007155; GO:0009289
Q07564	<i>Fimbrial protein ecpC (Pilin)</i>	<u>Eikenella corrodens</u>	Adesão celular; pilus	GO:0007155; GO:0009289
C8UDG8	<i>T3SS secreted effector NleA-like protein</i>	<u>Escherichia coli</u>	Proteína efectora	Sem GO identificados
C8MGI5	<i>Exotoxin 15</i>	<u>Staphylococcus aureus</u>	Patogénese	GO:0005576; GO:0009405
C2DNR2	<i>RTX-toxin family pore-forming toxin</i>	<u>Escherichia coli</u>	Ligação ao ião cálcio; patogénese	GO:0005509; GO:0005576; GO:0009405
B3WLE9	<i>Putative toxin B</i>	<u>Escherichia coli</u>	Endopeptidase; transferase; patogénese	GO:0004197; GO:0009405; GO:0016757
O34812	<i>Putative NADP-dependent oxidoreductase yfmJ</i>	<u>Bacillus subtilis</u>	Resposta a toxinas	GO:0019439; GO:0000166; GO:0016491; GO:0009636; GO:0008270
P16966	<i>Acetyltransferase (EC 2.3.1.-) (Tabtoxin resistance protein)</i>	<u>Pseudomonas syringae</u>	Actividade N-acetyltransferase	GO:0008080
C2H5F5	<i>Possible virulence protein EssB</i>	<u>Enterococcus faecalis</u>	N/A	Sem códigos GO associados

4.4. Análise das proteínas microbianas associadas a patologias

A ferramenta OralCard foi utilizada para verificar quais as proteínas de origem microbiana às quais estavam associadas patologias. Verificou-se que havia apenas 3 patologias em que foram associadas amostras com proteínas microbianas: Gengivite, Periodontite Agressiva e Cancro Oral. Foram identificados como presentes no OralOme

5 proteínas microbianas associadas a Gengivite (de um total de 95 proteínas), 11 associadas a Periodontite Agressiva (de um total de 123 proteínas) e 53 associadas a Cancro Oral (de um total de 315 proteínas). Após a identificação destas proteínas procedeu-se à análise da anotação das mesmas, seleccionando as proteínas cujas ontologias apresentavam maior potencial de interesse para a interacção com o hospedeiro, nomeadamente as que na ontologia *Cellular Component* (GO:0016043) apresentavam produtos génicos correspondentes a *Extracellular Region* (GO:0005576) e *Plasma Membrane* (GO:0005886). Os resultados desta análise encontram-se na Tabela III

Tabela III - Resumo da informação obtida usando a ferramenta OralCard relativa às proteínas de origem microbiana associadas a patologias orais de interesse na interacção entre hospedeiro e microrganismo

Código da Proteína	Nome da Proteína	Microrganismo	Patologia	GO de interesse	Significado do GO
O07497	<i>Protein translocase subunit SecA</i>	<u><i>Borrelia burgdorferi</i></u>	Periodontite Agressiva	GO:0005886	<i>Plasma membrane</i>
Q02929	<i>Putative sensory transducer protein</i>	<u><i>Clostridium thermocellum</i></u>	Periodontite Agressiva	GO:0005886	<i>Plasma membrane</i>
P07018	<i>Methyl-accepting chemotaxis protein IV</i>	<u><i>Escherichia coli</i></u>	Periodontite Agressiva	GO:0005886	<i>Plasma membrane</i>
Q6MPQ2	<i>Enolase EC=4.2.1.11</i>	<u><i>Bdellovibrio bacteriovorus</i></u>	Cancro Oral	GO:0005576	<i>Extracellular Region</i>
P57216	<i>Electron transport complex protein RnfD</i>	<u><i>Buchnera aphidicola</i></u>	Cancro Oral	GO:0005886	<i>Plasma membrane</i>
Q8NRS1	<i>Enolase EC=4.2.1.11</i>	<u><i>Corynebacterium glutamicum</i></u>	Cancro Oral	GO:0005576	<i>Extracellular Region</i>
P10324	<i>Outer membrane protein P6</i>	<u><i>Haemophilus influenzae</i></u>	Cancro Oral	GO:0005886	<i>Plasma membrane</i>
P44592	<i>UPF0092 membrane protein HI_0241</i>	<u><i>Haemophilus influenzae</i></u>	Cancro Oral	GO:0005886	<i>Plasma membrane</i>
Q06282	<i>Glycerophosphoryl diester phosphodiesterase</i>	<u><i>Haemophilus influenzae</i></u>	Cancro Oral	GO:0005886	<i>Plasma membrane</i>

5. Discussão

A análise dos resultados do proteoma de origem microbiana da cavidade oral, permite concluir que a cobertura dos resultados experimentais de proteómica é inferior à cobertura genómica do microbioma oral. Na realidade, diversos estudos de genómica de vários tipos de amostra da cavidade oral, sugerem que o número de espécies presentes está acima dos 600 [21], enquanto que, pela análise dos resultados de proteómica, apenas são “visíveis” 170 espécies. Este resultado pode ser explicado pelo facto de os estudos de proteómica da saliva, apesar de abundantes, só raramente apresentarem a identificação de proteínas de origem bacteriana.

É cada vez mais consensual na comunidade científica que a identificação de proteínas de origem microbiana em amostras de tecidos provenientes da cavidade oral (saliva, fluido crevicular, raspagens da mucosa ou outros), requer procedimentos específicos quer no processamento inicial da amostra quer na estratégia de identificação dos fragmentos obtidos na espectrometria de massa [51]. Só assim se explica que dos 5 artigos encontrados em que eram referidas proteínas microbianas nos resultados, no caso do artigo de Japtrap *et al* [51], direccionado para a identificação de proteínas microbianas, fossem identificadas mais de 2000 proteínas, enquanto nos artigos anteriores, em que o enfoque eram as proteínas humanas, foram detectadas apenas algumas dezenas de proteínas [47-49] (Tabela I).

É interessante verificar que nos resultados de proteómica, surge a evidência da presença de vírus e leveduras que se sabe que existem no microbioma humano [3] embora os números estejam, mais uma vez, aquém daquilo que os estudos de metagenómica indicam.

Quando se compararam os resultados das proteínas específicas, encontradas em cada um dos estudos da tabela I, é também possível verificar que não há sobreposição, ou seja em cada artigo publicado foram identificadas proteínas diferentes de todas as até então publicadas. Além das razões técnicas descritas acima (técnicas de recolha da amostra, técnicas de processamento e de identificação dos fragmentos), que justificam a ausência de proteínas microbianas dos resultados de análises proteómicas de amostras da cavidade oral, acresce o facto de, tal como indicam os resultados da genómica [3], as amostras de vários locais da cavidade oral representarem um microbioma distinto.

Assim, e dado que os 5 estudos de proteómica incluídos analisam 3 tipos de amostras diferentes, a não sobreposição nos resultados pode também ser justificada.

Após a análise do gráfico correspondentes a *Cellular Component* (GO:0005575), foram identificadas várias ontologias, sendo seleccionados os termos de maior relevância para a interacção com o hospedeiro, nomeadamente os que dizem respeito a ontologias extracelulares (correspondentes a 13,7% do total das ontologias), que incluem: *Plasma Membrane* (5% correspondentes ao GO:00058865 “*The membrane surrounding a cell that separates the cell from its external environment. It consists of a phospholipid bilayer and associated proteins*” [52]); *Extracellular region* (3,2% correspondentes ao GO:0005576 “*The space external to the outmost structure of a cell. For cells without external protective or external encapsulating structures this refers to the space outside of the plasma membrane. This term covers the host cell environment outside an intracellular parasite*” [52]); *Cell Envelope* (2,4% correspondentes ao GO:0030313 “*An envelope that surrounds a bacterial cell and includes the cytoplasmic membrane and everything external, encompassing the periplasmic space, cell wall, and outer membrane if present*” [52]); *External Encapsulating Structure* (2,3% correspondentes ao GO:0030312 “*A structure that lies outside the plasma membrane and surrounds the entire cell*”[52]); *Cell Wall* (0,6% correspondentes ao GO:0005618 “*The rigid or semi-rigid envelope lying outside the cell membrane of plant, fungal, and most prokaryotic cells, maintaining their shape and protecting them from osmotic lysis. In plants it is made of cellulose and, often, lignin; in fungi it is composed largely of polysaccharides; in bacteria it is composed of peptidoglycan*” [52]) e *Proteinaceous Extracellular Matrix* (0,3% correspondentes ao GO:0005578 “*A layer consisting mainly of proteins (especially collagen) and glycosaminoglycans (mostly as proteoglycans) that forms a sheet underlying or overlying cells such as endothelial and epithelial cells. The proteins are secreted by cells in the vicinity. An example of this component is found in *Mus musculus**” [52]). Foram escolhidas estas anotações ontológicas, já que a interacção que poderá existir com o hospedeiro será possivelmente realizada entre este e as proteínas microbianas localizadas no exterior da célula. Por outras palavras, serão estas as primeiras estruturas a entrar em contacto com as superfícies do hospedeiro.

Quanto ao gráfico ilustrativo da *Molecular Function* (GO:0003674), as ontologias de maior relevância para o presente estudo são as que indicam proteínas com actividade de cinase (0,46% correspondentes ao GO:0004672, *Protein Kinase Activity: "Catalysis of the phosphorylation of an amino acid residue in a protein, usually according to the reaction: a protein + ATP = a phosphoprotein + ADP."* [52]), sendo estas enzimas importantes na fosforilação de proteínas envolvidas na transdução de sinal. Este fenómeno ocorre quando uma molécula de sinalização extracelular activa um receptor da superfície celular, sendo por essa razão importante na interacção com o hospedeiro. As proteínas classificadas com esta ontologia, estão envolvidas na resposta da célula microbiana aos estímulos externos. Como foi referido na revisão bibliográfica, estes sistemas de transdução de sinais estão, muitas vezes, associados ao modo de vida em biofilme, e na modulação da expressão de factores de virulência nalgumas espécies oportunistas[53, 54].

A actividade antioxidante (0,46% correspondentes ao GO:0016209, *Antioxidant Activity: "Inhibition of the reactions brought about by dioxygen (O₂) or peroxides. Usually the antioxidant is effective because it can itself be more easily oxidized than the substance protected. The term is often applied to components that can trap free radicals, thereby breaking the chain reaction that normally leads to extensive biological damage."*) é também uma característica importante a ter em conta, já que a inibição das moléculas anti-oxidantes dos microrganismos impede a sua tolerância ao stress oxidativo, provocado pelos radicais livres de oxigénio. O stress oxidativo faz parte dos mecanismos de defesa usados pelo sistema imunitário humano, nomeadamente pelos pela fagocitose dos macrófagos.

A ligação a hidratos de carbono (0,4% correspondentes ao GO:0030246, *Carbohydrate Binding: "Interacting selectively and non-covalently with any carbohydrate, which includes monosaccharides, oligosaccharides and polysaccharides as well as substances derived from monosaccharides by reduction of the carbonyl group (alditols), by oxidation of one or more hydroxy groups to afford the corresponding aldehydes, ketones, or carboxylic acids, or by replacement of one or more hydroxy group(s) by a hydrogen atom. Cyclitols are generally not regarded as carbohydrates."* [52]) é de elevada importância na recolha de nutrientes pelos microrganismos, o que implica que a ausência destas moléculas pode resultar de na perda de capacidade de entrada de nutrientes para o espaço intracelular da bactéria, o que pode interferir com a capacidade

de sobrevivência e virulência. Estes efeitos são teoricamente mais nefastos em espécies fermentadoras como os *Streptococci*. Acresce ainda o facto de muitas interacções dos microrganismos com as superfícies, se darem através de moléculas que sofrem glicosilações, e portanto a capacidade de ligação a epítomos glicosilados pode alterar as propriedades de colonização das espécies do biofilme[55].

Quanto à actividade de receptor (0,36% correspondentes ao GO:0004872, *Receptor Activity*: “Combining with an extracellular or intracellular messenger to initiate a change in cell activity.” [52]), pode estar envolvida na entrada de nutrientes para a célula, bem como comunicações via *quorum sensing*, sendo assim estas proteínas potenciais alvos terapêuticos, e como tal, importantes na interacção com o hospedeiro. A ligação de receptores (0,2%, correspondentes ao GO:0005102, *Receptor Binding*: “Interacting selectively and non-covalently with one or more specific sites on a receptor molecule, a macromolecule that undergoes combination with a hormone, neurotransmitter, drug or intracellular messenger to initiate a change in cell function.” [52]) é, à semelhança das proteínas incluídas no termo *Receptor Activity*, essencial na entrada de nutrientes para a célula e vias de comunicação inter-microrganismo através de *quorum sensing*, sendo uma das proteínas que poderá ser de elevado interesse para o objectivo do desenvolvimento de ferramentas terapêuticas que não causem resistência nos microrganismos.

A ligação ao ião cálcio (0,1% correspondentes ao GO:0005509, *Calcium Ion Binding*: “Interacting selectively and non-covalently with calcium ions (Ca²⁺).” [52]) é de elevada importância na biologia celular, o seu envolvimento é fundamental na transdução de sinal, é um co-factor de várias enzimas, vital para o funcionamento de canais membranares e desempenha um papel essencial na manutenção da diferença de potencial. Assim, a ausência de ligação ao ião cálcio tem várias repercussões a nível celular que podem ser de interesse na colonização do hospedeiro.

Por fim, a ligação a lípidos, (0,1% correspondentes ao GO:0008289, *Lipid Binding*: “Interacting selectively and non-covalently with a lipid” [52]), foi considerada importante pois estas moléculas são componentes das membranas celulares quer do hospedeiro quer das bactérias. Na grande maioria da cavidade oral, os nichos disponíveis para a colonização, não apresentam estas membranas fosfolipídicas disponíveis, como é o caso do esmalte e dentina, materiais restauradores ou mesmo o

epitélio queratinizado. No entanto, existem zonas da mucosa que podem conter células que exibam a sua membrana fosfolipídica, sendo por isso estas proteínas importantes.

Por fim, após a análise da Figura 13, que ilustra os termos descendentes de *Biological Process* (GO:0008150), concluiu-se que as ontologias que poderiam ser relevantes para a interacção com o hospedeiro seriam as correspondentes a *Cell Communication* (0,2% com o GO:0007154) e *Symbiosis, Encompassing Mutualism Through Parasitism* (0,1% correspondentes ao GO:0044403).

Quanto à ontologia *Cell Communication*, nela estão incluídos todos os produtos génicos que “*Any process that mediates interactions between a cell and its surroundings. Encompasses interactions such as signaling or attachment between one cell and another cell, between a cell and an extracellular matrix, or between a cell and any other aspect of its environment*” [52], ou seja, são todos os processos que são intermediários em qualquer interacção entre a célula e o meio ambiente circundante. Abrange interacções como sinalização ou união entre 2 células, entre uma célula e a matriz extracelular ou entre uma célula e outro aspecto do meio ambiente que a rodeia. Assim, tal como os produtos génicos referidos anteriormente para a ontologia *Symbiosis, Encompassing Mutualism Through Parasitism*, os produtos génicos pertencentes à ontologia *Cell Communication* são de elevado interesse para a interacção com o hospedeiro, já que correspondem a proteínas microbianas que podem modelar a interacção com o hospedeiro, e determinar se a relação estabelecida é de comensalismo ou parasitismo[56].

Após a análise da Tabela II, pode-se verificar que existem 6 proteínas com potencial interesse para o estudo das interacções entre o microbioma e o hospedeiro. As moléculas de adesão *Fimbrial protein ecpA (Pilin)* e *Fimbrial protein ecpC (Pilin)* atribuídas à bactéria *Eikenella corrodens*. Esta é uma bactéria comensal gram-negativa anaeróbia facultativa. As proteínas e as estruturas a que correspondem são muito importantes na colonização da cavidade oral, e podem ter influência no estabelecimento destas bactérias no biofilme oral. Existe evidência que noutras bactérias, como a *Porphyromonas gingivalis*, a virulência e capacidade de colonização da cavidade oral estão muito ligadas à capacidade de produzir pili e às proteínas específicas que estão presentes nessas estruturas [57][56].

As proteínas *T3SS secreted effector NleA-like protein* e *Putative toxin B*, pertencentes a *E. coli*, foram também identificadas como interessantes, sendo a primeira um efector e a segunda uma toxina. A toxina B produzida pela bactéria *E. coli*, é uma potente citotoxina, mas a sua resposta não é muito elevada em estudos com modelos animais. No entanto, ao ser uma toxina, poderá ser a responsável pelo aparecimento de patologias ou de respostas imunológicas do hospedeiro [58]. Quanto à proteína efectora secretada por T3SS, esta é usada pela *E. coli* para colonizar vertebrados e promover o parasitismo desta bactéria, através da supressão do sistema imune do hospedeiro e indução da morte celular [59]. Assim, qualquer uma destas duas proteínas poderá ser importante na interacção com o hospedeiro e no surgimento de patologias.

No microrganismo *Staphylococcus aureus*, existem várias exotoxinas associadas a patologias, incluindo a síndrome do choque tóxico. Assim, a produção da Exotoxina 15 pelo *Staphylococcus aureus* pode ter consequências a níveis de interacção com o hospedeiro e patologias [60].

A *Putative NADP-dependent oxidoreductase* presente no microrganismo *Bacillus subtilis*, tem como função a resposta a toxinas. Assim, esta proteína funciona como um mecanismo de resistência, promovendo uma colonização mais eficiente deste microrganismo, que pode ter influência na etiopatogenia de algumas patologias, como doenças pulmonares[61].

Foi seleccionada a proteína *Acetyltransferase (Tabtoxin resistance protein)* do microrganismo *Pseudomonas syringae*. Esta bactéria produz toxinas com efeitos antibióticos, que lhe permite uma vantagem sobre os outros microrganismos do seu meio. No entanto, necessita desta proteína para se proteger dessa mesma toxina. Garantindo então maior capacidade de colonização sobre outras bactérias do meio, que não possuam esta resistência [62].

Por fim, o possível factor de virulência *EssB* identificado no microrganismo *Enterococcus faecalis*, pode ser importante no surgimento de várias patologias perirradiculares e endodônticas, já que um factor de virulência iria conferir maior facilidade de colonização e, conseqüentemente, maior virulência[63].

A análise da Tabela II, indica-nos a presença de moléculas microbianas associadas a patologias orais, que estão anotadas com ontologias que sugerem que possam estar envolvidas na interacção com o hospedeiro. Dessas moléculas destacam-se as duas primeiras uma *Protein translocase subunit SecA*, e a *Putative sensory transducer protein*, pelas razões mencionadas anteriormente, ou seja, o seu envolvimento na produção de estruturas como os pili (*Protein translocase subunit SecA*), e na percepção de sinais externos à bactéria e que modulam a sua resposta metabólica (*Putative sensory transducer protein*). Das outras proteínas identificadas nenhuma delas foi descrita como podendo estar directa ou indirectamente envolvida na etiologia molecular das patologias apresentadas sendo portanto de investigar esta relação em estudos futuros.

6. Conclusão

Neste trabalho foi realizado um levantamento exaustivo dos estudos de proteómica de amostras da cavidade oral, em que houvesse a identificação de proteínas microbianas.

O resultado principal deste trabalho, foi permitir a actualização da base de dados OralOme, no que se refere às proteínas de origem microbiana experimentalmente determinadas, e publicadas até Fevereiro de 2012. Estas proteínas foram não só adicionadas à base de dados, como também manualmente anotadas.

A comparação dos dados de proteómica da cavidade oral existentes, no que se refere aos microrganismos, com os recentes resultados da metagenómica do microbioma humano da cavidade oral em saúde, permite concluir que a cobertura em termos de proteínas microbianas identificadas é muito baixa. Adicionalmente, conclui-se que a sobreposição de resultados entre os vários estudos de proteómica não se verifica, devido à aplicação de técnicas diferentes de processamento de amostras e de identificação dos fragmentos obtidos na análise de espectrometria de massa. Contribuirá também para estas diferenças, o facto de tecidos diferentes na cavidade oral terem espécies diferentes de bactérias, ainda que as vias metabólicas envolvidas sejam idênticas.

Das proteínas microbianas determinadas experimentalmente, em amostras da cavidade oral, foram identificadas aquelas que apresentam maiores probabilidades de ser importantes na interacção com o hospedeiro. Assim, estas moléculas serão os alvos primários em estudos futuros de interacção bactéria/tecidos, presentes na cavidade oral. Estes estudos têm o potencial de revelar mecanismos e moléculas alvo de novas formas de interferência, com alteração do equilíbrio dinâmico estabelecido entre o biofilme oral e o hospedeiro.

7. Bibliografía

1. Pesquita, C., et al., , *Semantic Similarity in Biomedical Ontologies*. PLoS Comput Biol, 2009. **5**(7).
2. Rosa, N., et al., *From the salivary proteome to the OralOme: Comprehensive molecular oral biology*. Arch Oral Biol, 2012.
3. *Structure, function and diversity of the healthy human microbiome*. Nature, 2012. **486**(7402): p. 207-14.
4. Janoff, E.N., C. Gustafson, and D. Frank, *The world within: living with our microbial guests and guides*. Transl Res, 2012.
5. Nelson, K.E., et al., *A catalog of reference genomes from the human microbiome*. Science, 2010. **328**(5981): p. 994-9.
6. Avila, M., D.M. Ojcius, and O. Yilmaz, *The oral microbiota: living with a permanent guest*. DNA Cell Biol, 2009. **28**(8): p. 405-11.
7. *A framework for human microbiome research*. Nature, 2012. **486**(7402): p. 215-21.
8. Lemos, J.A., T.A. Brown, Jr., and R.A. Burne, *Effects of RelA on key virulence properties of planktonic and biofilm populations of Streptococcus mutans*. Infect Immun, 2004. **72**(3): p. 1431-40.
9. Maeda, K., et al., *A Porphyromonas gingivalis tyrosine phosphatase is a multifunctional regulator of virulence attributes*. Mol Microbiol, 2008. **69**(5): p. 1153-64.
10. Sharma, V., et al., *Menaquinone (vitamin K2) biosynthesis: nucleotide sequence and expression of the menB gene from Escherichia coli*. J Bacteriol, 1992. **174**(15): p. 5057-62.
11. Munoz, N., et al., *Mucosal administration of flagellin protects mice from Streptococcus pneumoniae lung infection*. Infect Immun, 2010. **78**(10): p. 4226-33.
12. Hooton, T.M. and W.E. Stamm, *Diagnosis and treatment of uncomplicated urinary tract infection*. Infect Dis Clin North Am, 1997. **11**(3): p. 551-81.
13. Bjarnsholt, T., et al., *Pseudomonas aeruginosa biofilms in the respiratory tract of cystic fibrosis patients*. Pediatr Pulmonol, 2009. **44**(6): p. 547-58.
14. Patrick R. Murray, K.S.R., George S. Kobayashi, Michael A. Pfaller, ed. *Microbiología Médica*. 4ª ed. 2004, Gunabara-Koogan. 762.
15. Peterson, J., et al., *The NIH Human Microbiome Project*. Genome Res, 2009. **19**(12): p. 2317-23.
16. Kuramitsu, H.K., *Proteases of Porphyromonas gingivalis: what don't they do?* Oral Microbiol Immunol, 1998. **13**(5): p. 263-70.
17. Livingstone, C., ed. *Essential Microbiology for Dentistry*. 2005.
18. Busuioc, M., B.A. Buttaro, and P.J. Piggot, *The pdh operon is expressed in a subpopulation of stationary-phase bacteria and is important for survival of sugar-starved Streptococcus mutans*. J Bacteriol, 2010. **192**(17): p. 4395-402.
19. Smith, M.A., et al., *Characteristics of the aerobic respiratory chains of the microaerophiles Campylobacter jejuni and Helicobacter pylori*. Arch Microbiol, 2000. **174**(1-2): p. 1-10.
20. Hong, S.H., et al., *The genome sequence of the capnophilic rumen bacterium Mannheimia succiniciproducens*. Nat Biotechnol, 2004. **22**(10): p. 1275-81.
21. Li, Y.H. and X. Tian, *Quorum sensing and bacterial social interactions in biofilms*. Sensors (Basel), 2012. **12**(3): p. 2519-38.
22. Carranza, N., Takei, Klokkevold, ed. *Periodontia Clínica*. 10ª ed. 2007, Saunders Elsevier.
23. Jenkinson, H.F. and R.J. Lamont, *Oral microbial communities in sickness and in health*. Trends Microbiol, 2005. **13**(12): p. 589-95.

24. Singh, A., et al., *The capsule of Porphyromonas gingivalis leads to a reduction in the host inflammatory response, evasion of phagocytosis, and increase in virulence*. Infect Immun, 2011. **79**(11): p. 4533-42.
25. Fenno, J.C., *Treponema denticola interactions with host proteins*. J Oral Microbiol, 2012. **4**.
26. Arirachakaran, P., et al., *Infection of human gingival fibroblasts with Aggregatibacter actinomycetemcomitans: An in vitro study*. Arch Oral Biol, 2012. **57**(7): p. 964-72.
27. Asad, S. and S.M. Opal, *Bench-to-bedside review: Quorum sensing and the role of cell-to-cell communication during invasive bacterial infection*. Crit Care, 2008. **12**(6): p. 236.
28. Madigan, M., Parker, ed. *Brock's Biology of Microorganisms*. 2000, Prentice Hall.
29. Waters, C.M. and B.L. Bassler, *Quorum sensing: cell-to-cell communication in bacteria*. Annu Rev Cell Dev Biol, 2005. **21**: p. 319-46.
30. Lantz, M.S., *Are bacterial proteases important virulence factors?* J Periodontal Res, 1997. **32**(1 Pt 2): p. 126-32.
31. Shames, S.R. and B.B. Finlay, *Bacterial effector interplay: a new way to view effector function*. Trends Microbiol, 2012. **20**(5): p. 214-9.
32. Kline, K.A., et al., *Bacterial adhesins in host-microbe interactions*. Cell Host Microbe, 2009. **5**(6): p. 580-92.
33. Makela, M., et al., *Matrix metalloproteinases (MMP-2 and MMP-9) of the oral cavity: cellular origin and relationship to periodontal status*. J Dent Res, 1994. **73**(8): p. 1397-406.
34. Nicholas C. Price, L.S., ed. *Fundamentals of Enzymology*. 3rd ed. 1999, Oxford University Press. 478.
35. Torto-Alalibo, T., et al., *Common and contrasting themes in host cell-targeted effectors from bacterial, fungal, oomycete and nematode plant symbionts described using the Gene Ontology*. BMC Microbiol, 2009. **9 Suppl 1**: p. S3.
36. Dean, P., et al., *The enteropathogenic E. coli effector EspF targets and disrupts the nucleolus by a process regulated by mitochondrial dysfunction*. PLoS Pathog, 2010. **6**(6): p. e1000961.
37. McKusick, V.A., *Mendelian Inheritance in Man and its online version, OMIM*. Am J Hum Genet, 2007. **80**(4): p. 588-604.
38. Boutet, E., et al., *UniProtKB/Swiss-Prot*. Methods Mol Biol, 2007. **406**: p. 89-112.
39. Mi, H., et al., *PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium*. Nucleic Acids Res, 2010. **38**(Database issue): p. D204-10.
40. Szklarczyk, D., et al., *The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored*. Nucleic Acids Res, 2011. **39**(Database issue): p. D561-8.
41. Scheer, M., et al., *BRENDA, the enzyme information system in 2011*. Nucleic Acids Res, 2011. **39**(Database issue): p. D670-6.
42. Chen, L., et al., *VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors*. Nucleic Acids Res, 2012. **40**(Database issue): p. D641-5.
43. Chen, T., et al., *The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information*. Database (Oxford), 2010. **2010**: p. baq013.
44. Melo, J.A., JP; Lopes, P; Rosa, N; Correia, MJ; Barros, MT; Oliveira, JL. , *Web information system for the study of oral health*. HEALTHINF, 2012.
45. Torto-Alalibo, T., et al., *Unifying themes in microbial associations with animal and plant hosts described using the gene ontology*. Microbiol Mol Biol Rev, 2010. **74**(4): p. 479-503.

46. Xie, H., 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): p. 486-98.
47. Esser, D., et al., *Sample Stability and Protein Composition of Saliva: Implications for Its Use as a Diagnostic Fluid*. Biomark Insights, 2008. **3**: p. 25-27.
48. Bostanci, N., et al., *Application of label-free absolute quantitative proteomics in human gingival crevicular fluid by LC/MS E (gingival exudatome)*. J Proteome Res, 2010. **9**(5): p. 2191-9.
49. Grant, M.M., et al., *Proteomic analysis of a noninvasive human model of acute inflammation and its resolution: the twenty-one day gingivitis model*. J Proteome Res, 2010. **9**(9): p. 4732-44.
50. McCarthy, F.M., et al., *AgBase: supporting functional modeling in agricultural organisms*. Nucleic Acids Res, 2011. **39**(Database issue): p. D497-506.
51. Jagtap, P., et al., *Deep metaproteomic analysis of human salivary supernatant*. Proteomics, 2012. **12**(7): p. 992-1001.
52. Consortium, T.G.O., *Gene ontology: tool for the unification of biology*. Nat Genet, 2000. **25**(1): p. 25-29.
53. Bandara, H.M., et al., *Microbial chemical signaling: a current perspective*. Crit Rev Microbiol, 2012. **38**(3): p. 217-49.
54. Senadheera, D.B., et al., *Regulation of bacteriocin production and cell death by the VicRK signaling system in Streptococcus mutans*. J Bacteriol, 2012. **194**(6): p. 1307-16.
55. Iwashkiw, J.A., et al., *Identification of a General O-linked Protein Glycosylation System in Acinetobacter baumannii and Its Role in Virulence and Biofilm Formation*. PLoS Pathog, 2012. **8**(6): p. e1002758.
56. Sbordone, L. and C. Bortolaia, *Oral microbial biofilms and plaque-related diseases: microbial communities and their role in the shift from oral health to disease*. Clin Oral Investig, 2003. **7**(4): p. 181-8.
57. Njoroge, T., et al., *A role for fimbriae in Porphyromonas gingivalis invasion of oral epithelial cells*. Infect Immun, 1997. **65**(5): p. 1980-4.
58. Song, K.P., Faust, C., *Molecular analysis of the promoter region of the Clostridium difficile toxin B gene that is functional in Escherichia coli*. J. Med. Microbiol., 1998. **47**: p. 309-316.
59. Preston, G.M., *Metropolitan Microbes: Type III Secretion in Multihost Symbionts*. Cell Host Microbe, 2007. **2**: p. 291-294.
60. Bunikowski, R., et al., *Evidence for a disease-promoting effect of Staphylococcus aureus-derived exotoxins in atopic dermatitis*. J Allergy Clin Immunol, 2000. **105**(4): p. 814-9.
61. Flindt, M.L.H., *Pulmonary Disease due to Inhalation of derivatives of Bacillus subtilis containing proteolytic enzyme*. The Lancet, 2003. **293**(7607): p. 1177-1181.
62. Mourgues, F., M.N. Brisset, and E. Chevreau, *Strategies to improve plant resistance to bacterial diseases through genetic engineering*. Trends Biotechnol, 1998. **16**(5): p. 203-10.
63. Rocas, I.N., J.F. Siqueira, Jr., and K.R. Santos, *Association of Enterococcus faecalis with different forms of periradicular diseases*. J Endod, 2004. **30**(5): p. 315-20.

8. Índice de Tabelas

Tabela I - Artigos de estudos de proteômica salivar seleccionados para este estudo.
27

Tabela III - Resumo da informação obtida usando a ferramenta OralCard relativa às proteínas de origem microbiana associadas a patologias orais de interesse na interação entre hospedeiro e microrganismo 38

9. Índice de Figuras

- Figura 1 – Vista e m árvore dos termos antecedentes e descendentes do “GO:0044403, symbiosis encompassing mutualism through parasitism”, adaptado da vista em árvore de amigo.geneontology.org a 28/06/2012. 17**
- Figura 2 – Vista em árvore dos termos descendentes do “GO:0043675, host cell”, adaptado da vista em árvore de amigo.geneontology.org a 28/06/2012. 18**
- Figura 3 - Obtenção de informação sobre as 1212 proteínas identificadas: utilização da opção “Retrieve” no portal web do UniProtKB. Figura adaptada do portal Web do UniProtKB [29]22**
- Figura 4 - Obtenção de informação sobre as 1212 proteínas identificadas: colocação dos identificadores de todas as proteínas e selecção da opção “Retrieve”. Figura adaptada do portal Web do UniProtKB[29] 23**
- Figura 5 - Obtenção de informação sobre as 1212 proteínas identificadas: selecção da opção “UniProtKB. Figura adaptada do portal Web do UniProtKB[29] 23**
- Figura 7 - Obtenção de informação sobre as 1212 proteínas identificadas (nome das proteínas, GOs, identificação das ontologias, organismo de origem e função da proteína) Seleccionou-se de seguida a opção “save”. Figura adaptada do portal Web do UniProtKB[29] 24**
- Figura 8 - Obtenção de informação sobre as 1212 proteínas identificadas: foi seleccionada a opção “Download” para descarregar o ficheiro. Figura adaptada do portal Web do UniProtKB[29] 25**
- Figura 9 - Obtenção de informação sobre as 1212 proteínas identificadas: escolha do formato a descarregar. Foi seleccionado o formato .xlsx para ser possível trabalhar com o Microsoft Excel. Figura adaptada do portal Web do UniProtKB[29] 25**
- Figura 10 - Representação do número de proteínas de origem microbiana presentes no OralCard por espécie de microrganismo. No grupo “Outros” estão incluídas 141 espécies cuja identificação resulta da presença de 11 ou menos proteínas. A correspondência entre a proteína identificada e o microrganismo foi obtida com o UniProtKB. 29**
- Figura 11 - Representação da classificação ontológica das 1212 proteínas de origem microbiana presentes no OralOme quanto à classe Cellular Component. As classes ontológicas mantêm a designação em inglês pelas razões apresentadas na nota prévia do prefácio. 30**
- Figura 12 - Representação da classificação ontológica das 1212 proteínas de origem microbiana presentes no OralOme quanto à classe Molecular Function. As classes ontológicas mantêm a designação em inglês pelas razões apresentadas na nota prévia do prefácio. 32**
- Figura 13 - Representação da classificação ontológica das 1212 proteínas de origem microbiana presentes no OralOme quanto à classe Biological Process. As classes ontológicas mantêm a designação em inglês pelas razões apresentadas na nota prévia do prefácio 34**
-