



UNIVERSIDADE CATÓLICA PORTUGUESA | INSTITUTO DE CIÊNCIAS DA SAÚDE

## **INTERATÓMICA DA CAVIDADE ORAL**

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

Por:

André do Nascimento Gonçalves Soares

Sob a orientação da Professora Doutora Maria José Correia e a co-orientação  
do Professor Doutor Joel Arrais

Viseu 2014





UNIVERSIDADE CATÓLICA PORTUGUESA | INSTITUTO DE CIÊNCIAS DA SAÚDE

## **INTERATÓMICA DA CAVIDADE ORAL**

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

Por:

André do Nascimento Gonçalves Soares

Sob a orientação da Professora Doutora Maria José Correia e a co-orientação  
do Professor Doutor Joel Arrais

Viseu 2014



*“Quem não sente a ânsia de ser mais, nunca vai chegar a ser nada”*

**Miguel De Unamuno**



## **AGRADECIMENTOS**

À Professora Doutora Maria José Correia,  
orientadora deste trabalho, pelo apoio, disponibilidade, conhecimentos transmitidos e pelo constante estímulo para ir mais além e fazer sempre melhor. Obrigado pela motivação e calma que transmite.

Ao Professor Doutor Joel Arrais,  
co-orientador deste trabalho, pelos seus comentários e levantamento de questões sempre tão pertinentes.

Ao Professor Doutor Nuno Rosa e à Professora Doutora Marlene Barros,  
Pelo apoio, disponibilidade, troca de impressões e comentários ao longo da realização deste trabalho.

Ao Mestre Edgar Coelho,  
pela sua disponibilidade e ajuda sempre tão atempada em momentos mais complicados durante a elaboração deste trabalho.

Aos meus Pais,  
pelo amor incondicional, pela forma como me inculcaram a alegria de viver, pelo apoio em todas as minhas escolhas e decisões e por estarem sempre comigo. Obrigado por tudo!

Aos meus Avós,  
meus heróis e anjos da guarda, por me providenciarem as fundações éticas e humanas através de exemplos sempre retos que me moldaram na pessoa que sou hoje.

Á Valentina,  
pelo amor e dedicação e por todos os momentos passados, presentes e futuros,  
que com apenas um sorriso ilumina a minha vida, mas também por ser a minha,  
fonte de inestimável apoio, paciência, preocupação e compreensão.

Ao Virgílio China  
meu binómio, pela amizade, apoio e pelos 5 fantásticos anos de convívio repletos  
de bons momentos.

Aos meus amigos André Régua, Bruno Borges e Hugo Ferreira,  
pela amizade e pelos excelentes momentos partilhados.

Aos meus amigos e colegas de curso,  
Por marcarem de forma tão positiva o meu percurso académico.

Ao Chini e ao Thori,  
Por me recordarem a simplicidade da vida todos os dias.

## Resumo

A cavidade oral humana é um ecossistema complexo onde fatores do hospedeiro, microbianos e ambientais interagem num equilíbrio dinâmico. A compreensão da biologia da cavidade oral e dos distúrbios que a afetam depende de ferramentas bioinformáticas que permitam a compilação, integração e aplicação da informação gerada por técnicas de alto rendimento, como as técnicas de proteômica, que se dedica à identificação de todas as proteínas expressas.

A compreensão dos mecanismos moleculares que se desenrolam na cavidade oral assenta na interação entre as moléculas presentes. O conhecimento destes mecanismos tem aplicações na fisiologia e fisiopatologia do sistema cavidade oral.

No sentido de responder à necessidade de estudar os genes e proteínas de forma integrada, dando realce às interações estabelecidas, começaram a ser desenvolvidas bases de dados de interatómica que incluem interações conhecidas e previstas através de algoritmos, associando-lhe um índice de confiança. O nosso trabalho contribuiu para o aperfeiçoamento de um algoritmo que permite a determinação de interações proteína-proteína. É utilizada a informação estrutural e de interações anteriormente determinadas entre os vários domínios das proteínas.

Este estudo tem como objetivo analisar, interpretar e explorar os resultados da aplicação do algoritmo OralInt ao conjunto de proteínas identificado por técnicas de proteômica como presentes na cavidade oral utilizando a base de dados OralOme, assim como verificar se algumas das interações descritas na bibliografia e determinadas experimentalmente na cavidade oral, se encontram presentes nas previsões e validar os resultados através de uma análise da literatura.

O desenvolvimento de ferramentas bioinformáticas que permitam a proposta de interações proteína-proteína, na cavidade oral, pode fornecer uma enorme quantidade de informação que vai assumir implicações na compreensão

dos mecanismos moleculares dos processos que ocorrem na cavidade oral em situações fisiológicas e patológicas.

Só com o conhecimento dos microrganismos que habitam a cavidade oral, das moléculas por eles produzidas e das interações destas com as do hospedeiro, podemos chegar ao entendimento molecular total.

**Palavras-chave:** Proteínas da cavidade oral, OralOme, domínios de interação, interações proteína-proteína, interatômica, interação hospedeiro-patogéneo.

## Abstract

The human oral cavity is a complex ecosystem where host, microbial and environmental factors interact in a dynamic equilibrium. Understanding the biology of oral cavity and disorders affecting it depends on bioinformatics tools to compile, integrate and apply the information generated by high-throughput techniques such as proteomics, which is dedicated to the identification of all the expressed proteins.

The understanding of the molecular mechanisms that take place in the oral cavity is based on the interaction between molecules present. The knowledge of these mechanisms has applications in physiology and pathophysiology of oral cavity system.

In order to address the need for studying genes and proteins in an integrated way, by emphasizing the interactions established, interatomic databases that include known and algorithmically predicted interactions, associated to a confidence index, began to be developed. This work contributes to the improvement of an algorithm allowing the determination of protein-protein interactions based on the structural information and interactions previously described between the various domains of the proteins (OralInt).

This study aims to describe the results of applying the algorithm OralInt to the set of proteins identified by proteomics techniques as in the oral cavity using the OralOme database and to check if some of the interactions described in the literature and determined experimentally in the oral cavity, are present in the predictions.

The development of bioinformatics tools that allow the proposal of protein-protein interactions in the oral cavity may provide enormous information that will have implications on the understanding of the molecular mechanisms occurring in the oral cavity in physiological and pathological situations.

Only with the knowledge of the microorganisms that inhabit the oral cavity, the molecules produced by them and the interactions with the host, can we reach the molecular understanding of the oral cavity.

**Keywords:** Oral cavity proteins, OralOme, interaction domains, protein-protein interactions, interatomic, host-pathogen interaction.

## 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 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.



# ÍNDICE

<b>1. INTRODUÇÃO .....</b>	<b>1</b>
1.1. MICROBIOMA HUMANO .....	3
1.1.1. <i>Microbioma oral</i> .....	5
1.1.1.1. Bactérias comensais da cavidade oral .....	5
1.1.1.2. Heterogeneidade de microbiomas na cavidade oral .....	7
1.2. COLONIZAÇÃO DO HOSPEDEIRO .....	8
1.2.1. <i>Formação do biofilme</i> .....	8
1.2.2. <i>Moléculas importantes na relação microrganismo/hospedeiro</i> .....	10
1.3. INTERATOMA .....	10
1.3.1. <i>Interatômica</i> .....	11
1.3.2. <i>Interatoma bacteriano</i> .....	11
1.3.3. <i>Interatoma Humano - patogéno</i> .....	12
1.3.4. <i>O interatoma oral Humano - bacteriano</i> .....	12
1.3.5. <i>Interações proteína-proteína</i> .....	13
1.4. FERRAMENTAS BIOINFORMÁTICAS E O ESTUDO DO INTERATOMA DA CAVIDADE ORAL .....	15
1.4.1. <i>OralOme e OralCard</i> .....	16
1.4.2. <i>Orallnt</i> .....	17
<b>2. OBJETIVOS.....</b>	<b>19</b>
<b>3. MATERIAIS E MÉTODOS .....</b>	<b>21</b>
3.1. PROTEOMA ORAL.....	21
3.2. ANÁLISE DAS REDES DE INTERATÔMICA.....	22
3.3. ANÁLISE DA CLASSIFICAÇÃO ONTOLÓGICA DAS PROTEÍNAS PRESENTES NOS ORGANISMOS EM ESTUDO .	25
<b>4. RESULTADOS E DISCUSSÃO .....</b>	<b>29</b>
4.1. INTERATOMA DA CAVIDADE ORAL .....	29
4.1.2. <i>Interatoma da cavidade oral – Interações por organismo</i> .....	34
4.1.3. <i>Interatoma da cavidade oral – Interações entre organismos microbianos</i> .....	35
4.2. INTERATOMA HUMANO – MICROBIANO DA CAVIDADE ORAL .....	36
4.2.1. <i>Aggregatibacter</i> .....	36
4.2.2. <i>Porphyromonas</i> .....	37
4.2.3. <i>Rothia</i> .....	42
4.2.3.1. <i>Rothia mucilaginosa</i> atcc 25296.....	43
4.2.3.2. <i>Rothia mucilaginosa (strain dy-18) (stomatococcus mucilaginosus)</i> .....	46
4.2.4. <i>Streptococcus</i> .....	51
4.2.4.1. <i>Streptococcus pneumoniae</i> .....	53
4.2.4.2. <i>Streptococcus salivarius</i> .....	58
4.2.4.3. <i>Streptococcus mutans</i> .....	60
<b>5. DISCUSSÃO.....</b>	<b>63</b>
<b>6. CONCLUSÃO .....</b>	<b>73</b>
<b>7. BIBLIOGRAFIA.....</b>	<b>75</b>



## ÍNDICE DE FIGURAS

<b>Figura 1</b> – Dados importados para o programa Cytoscape, com a definição da fonte e alvo de interação, assim como a classificação do tipo de interação. ...	23
<b>Figura 2</b> – Dados importados para o programa Cytoscape, com a informação referente aos microrganismos. ....	24
<b>Figura 3</b> - Escolha do layout, recorrendo às ferramentas informáticas do programa Cytoscape. ....	25
<b>Figura 4</b> – Página de rosto da base de dados do UniprotKB. ....	26
<b>Figura 5</b> – Secção search da base de dados do UniprotKB. ....	26
<b>Figura 6</b> – Ferramenta retrieve da base de dados do UniprotKB. ....	26
<b>Figura 7</b> – Obtenção da informação sobre as proteínas identificadas, utilizando a ferramenta customize da base de dados do UniprotKB. ....	27
<b>Figura 8</b> – Interações estabelecidas entre as proteínas orais humanas e bacterianas. Os nós a rosa representam as proteínas produzidas pelos organismos de origem microbiana, enquanto os nós a azul representam as proteínas orais de origem humana. Podemos ainda observar as proteínas produzidas pelo organismo <i>Aggregatibacter</i> a púrpura, <i>Porphyromonas</i> a vermelho, <i>Streptococcus</i> a verde, e por último <i>Rothia</i> , a amarelo. As interações estabelecidas entre proteínas de origem humana encontram-se a azul, a verde as realizadas entre proteínas humanas e microbianas, enquanto as interações estabelecidas entre proteínas microbianas encontram-se a amarelo. ....	29
<b>Figura 9</b> – Distribuição do número total de proteínas humanas e bacterianas analisadas pelo algoritmo Orallnt. ....	30
<b>Figura 10</b> – Distribuição do número total dos diferentes tipos de interação estabelecidos entre as proteínas identificadas pelo algoritmo Orallnt. ....	31
<b>Figura 11</b> – Número de proteínas de origem microbiana presentes por espécie de organismo na cavidade oral, analisadas pelo algoritmo Orallnt. ....	32
<b>Figura 12</b> – Género de microrganismos que produzem um maior número de proteínas presentes na cavidade oral, identificadas pelo algoritmo Orallnt. ....	33
<b>Figura 13</b> – Rede de Interações estabelecida entre distintos organismos presentes na cavidade oral. Os nós a rosa representam os organismos de origem	

microbiana, enquanto o nó que se encontra a azul representa o organismo Homo sapiens. As interações estabelecidas entre organismos de origem humana e bacteriana encontram-se a verde, enquanto as interações estabelecidas entre organismos de origem microbiana se encontram a amarelo. .... 34

**Figura 14** – Rede de Interações estabelecida entre os distintos organismos microbianos presentes na cavidade oral. Os nós a azul representam os organismos de origem microbiana. *Aggregatibacter actinomycetemcomitans* a púrpura, *Porphyromonas gingivalis* a vermelho, *Streptococcus mutans* a verde, e *Rothia mucilaginosa* a amarelo. .... 35

**Figura 15** – Rede de interações estabelecidas pelas proteínas bacterianas orais de *Aggregatibacter actinomycetemcomitans* (púrpura) com as proteínas orais humanas (azul) e com as proteínas orais bacterianas interagentes com *Aggregatibacter actinomycetemcomitans* (rosa). .... 37

**Figura 16** – Rede de interações estabelecidas pelas proteínas bacterianas orais de *Porphyromonas* (vermelho) com as proteínas orais humanas (azul) e com as proteínas orais bacterianas interagentes com *Porphyromonas* (rosa). Com um símbolo (triângulo) e cor diferentes (vermelho escuro) está a enolase de *Porphyromonas gingivalis* cujas interações específicas mais interessantes são exploradas. .... 38

**Figura 17** – Rede de interações estabelecida pela proteína Q7MTV8, Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate desidratase) de *Porphyromonas gingivalis* (vermelho) com proteínas orais de origem humana (azul) e com as proteínas orais bacterianas (rosa). .... 39

**Figura 18** – Rede de Interações estabelecida pelas proteínas bacterianas orais do género *Rothia* (amarelo) com as proteínas orais humanas (azul) e com as proteínas orais bacterianas interagentes com *Rothia* (rosa). .... 42

**Figura 19** – Rede de interações estabelecida pelas proteínas C6RE2 Extracellular basic protease (EC 3.4.21.-) e a C6R624 Extracellular solute-binding protein, family 5 de *Rothia mucilaginosa* atcc 25296 (amarelo) com proteínas orais de origem humana (azul). .... 43

**Figura 20** - Rede de interações estabelecida pelas proteínas D2NNZ2 Uncharacterized protein conserved in archaea, D2NPQ4 ABC-type dipeptide transport system, periplasmic component, D2NPZ7 Subtilisin-like serine protease, D2NRY1 Molybdopterin-binding protein e a D2NRY1 Molybdopterin-

binding protein de <i>Rothia mucilaginosa</i> (strain dy-18) ( <i>stomatococcus mucilaginosus</i> ) a (amarelo) com proteínas orais de origem humana (azul). ....	46
<b>Figura 21</b> - Rede de Interações estabelecida pelas proteínas bacterianas orais do género <i>Streptococcus</i> (verde) com as proteínas orais humanas (azul). ....	51
<b>Figura 22</b> - Rede de Interações estabelecida pelas proteínas bacterianas orais do género <i>Streptococcus</i> (verde) com as proteínas orais bacterianas (rosa). .	52
<b>Figura 23</b> – Rede de interações estabelecida pelas proteínas de <i>Streptococcus pneumoniae</i> (verde) com proteínas orais de origem humana (azul) e com as proteínas orais bacterianas (rosa).....	53
<b>Figura 24</b> - Rede de interações estabelecida pela proteína C2LUT1 Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate desidratase) de <i>Streptococcus salivarius</i> SK126 (verde), com as proteínas orais de origem humana (azul). ....	58
<b>Figura 25</b> - Rede de interações estabelecida pela proteína Q8DSC2 Putative PTS system, mannose-specific component IID (EC 2.7.1.69) de <i>Streptococcus mutans</i> serotype c (strain ATCC 700610 / UA159) (verde) com as proteínas orais de origem humana (azul). ....	60



## ÍNDICE DE TABELAS

<b>Tabela 1</b> – Relevância das Interações estabelecidas pela proteína Q7MTV8, Enolase (EC 4.2.1.11) ( <i>2-phospho-D-glycerate hydro-lyase</i> ) ( <i>2-phosphoglycerate dehydratase</i> ) de <i>Porphyromonas gingivalis</i> , já documentadas. ....	40
<b>Tabela 2</b> - Relevância das Interações estabelecidas pela proteína Q7MTV8, Enolase (EC 4.2.1.11) ( <i>2-phospho-D-glycerate hydro-lyase</i> ) ( <i>2-phosphoglycerate dehydratase</i> ) de <i>Porphyromonas gingivalis</i> , propostas pelo algoritmo Orallnt. ....	41
<b>Tabela 3</b> – Relevância da Interação estabelecida pela proteína C6R4E2 <i>Extracellular basic protease</i> (EC 3.4.21.-) de <i>Rothia mucilaginosa</i> atcc 25296, proposta pelo algoritmo Orallnt. ....	44
<b>Tabela 4</b> - Relevância das Interações estabelecidas pela proteína C6R624 <i>Extracellular solute-binding protein, family 5</i> de <i>Rothia mucilaginosa</i> atcc 25296 proposta pelo algoritmo Orallnt. ....	45
<b>Tabela 5</b> – Relevância das Interações estabelecidas pela proteína D2NNZ2 <i>Uncharacterized protein conserved in archaea</i> , de <i>Rothia mucilaginosa</i> (strain dy-18) ( <i>stomatococcus mucilaginosus</i> ). ....	47
<b>Tabela 6</b> – Relevância das Interações estabelecidas pelas proteínas D2NPQ4 <i>ABC-type dipeptide transport system, periplasmic componente</i> e D2NRY1 <i>Molybdopterin-binding protein</i> de <i>Rothia mucilaginosa</i> (strain dy-18) ( <i>stomatococcus mucilaginosus</i> ). ....	48
<b>Tabela 7</b> – Relevância das Interações estabelecidas pela proteína D2NPZ7 <i>Subtilisin-like serine protease</i> de <i>Rothia mucilaginosa</i> (strain dy-18) ( <i>stomatococcus mucilaginosus</i> ). ....	49
<b>Tabela 8</b> - Relevância das Interações estabelecidas pela proteína D2NRY1, <i>Molybdopterin-binding protein</i> de <i>Rothia mucilaginosa</i> (strain dy-18) ( <i>stomatococcus mucilaginosus</i> ). ....	50
<b>Tabela 9</b> - Relevância das Interações estabelecidas pela proteína A5MW80 <i>G5 domain family protein</i> de <i>Streptococcus pneumoniae</i> SP23-BS72. ....	54

<b>Tabela 10</b> - Relevância das Interações estabelecidas pela proteína Q97QS2 Enolase (EC 4.2.1.11) ( <i>2-phospho-D-glycerate hydro-lyase</i> ) ( <i>2-phosphoglycerate dehydratase</i> ) de <i>Streptococcus pneumoniae</i> SP23-BS72. .	55
<b>Tabela 11</b> - Relevância das Interações estabelecidas pela proteína P59213 <i>Maltose/maltodextrin-binding protein</i> , de <i>Streptococcus pneumoniae</i> serotype 4 (strain ATCC BAA-334 / TIGR4).....	56
<b>Tabela 12</b> - Relevância das Interações estabelecidas pelas proteínas Q97NH7 <i>Competence protein CglC</i> de <i>Streptococcus pneumoniae</i> serotype 4 (strain ATCC BAA-334 / TIGR4) e Q8DNE1 <i>Membrane protein insertase YidC 1</i> ( <i>Foldase YidC 1</i> ) ( <i>Membrane integrase YidC 1</i> ) ( <i>Membrane protein YidC 1</i> ) de <i>Streptococcus pneumoniae</i> (strain ATCC BAA-255 / R6) .....	57
<b>Tabela 13</b> - Relevância das Interações estabelecidas pela proteína C2LUT1, Enolase (EC 4.2.1.11) ( <i>2-phospho-D-glycerate hydro-lyase</i> ) ( <i>2-phosphoglycerate dehydratase</i> ) de <i>Streptococcus salivarius</i> SK126, já documentadas.....	59
<b>Tabela 14</b> - Relevância das Interações estabelecidas pela proteína C2LUT1, Enolase (EC 4.2.1.11) ( <i>2-phospho-D-glycerate hydro-lyase</i> ) ( <i>2-phosphoglycerate dehydratase</i> ) de <i>Streptococcus salivarius</i> SK126, propostas pelo algoritmo OralInt. ....	59
<b>Tabela 15</b> - Relevância das Interações estabelecidas pela proteína Q8DSC2 Putative PTS system, mannose-specific component IID (EC 2.7.1.69) de <i>Streptococcus mutans</i> serotype c (strain ATCC 700610 / UA159) propostas pelo algoritmo OralInt. ....	61

# 1. INTRODUÇÃO

O conhecimento do corpo humano começou por ser macroscópico e centrado em entender a sua organização anatómica, uma vez que as técnicas existentes não permitiam ir mais além. Com o avanço de várias técnicas como a microscopia, imagiologia, citoquímica e, posteriormente, de técnicas de genómica e de proteómica o conhecimento do corpo humano passou a ser, primeiro ao nível histológico, depois ao nível celular e, mais tarde, surge o conhecimento a nível molecular com a determinação do genoma humano (conjunto de todos os genes humanos) e de diversos proteomas (conjunto de todas as proteínas) de vários sistemas em diferentes situações fisiológicas e/ou patológicas. Isto permitiu compreender como e porquê ocorrem alterações que promovem diferentes patologias e como algumas das situações podem ser diagnosticadas e resolvidas, de uma forma que nunca antes tinha sido possível. No entanto, depressa se constatou que o conhecimento do genoma não era suficiente para caracterizar um organismo, nem tão pouco o era conhecer o proteoma.

Para caracterizar um dado sistema biológico é necessário conhecer, também, as interações entre todas as moléculas que nele participam, ou seja, conhecer o interatoma desse sistema biológico. À luz da interatómica, qualquer processo biológico é encarado não como resultado de uma ação independente de cada molécula interveniente, mas sim como uma rede de interações moleculares (2). A compreensão dos mecanismos moleculares que se desenrolam na cavidade oral assenta na interação entre as moléculas presentes. O conhecimento destes mecanismos tem aplicações na fisiologia e fisiopatologia do sistema cavidade oral.

No sentido de responder à necessidade de olhar para os genes e proteínas de forma integrada, dando realce às interações entre eles, começaram a ser desenvolvidas bases de dados de interatómica que incluem interações conhecidas e previstas através de algoritmos. Investigadores do Instituto de Engenharia, Eletrónica e Telemática da Universidade de Aveiro e do

Departamento de Informática da Universidade de Coimbra juntamente com investigadores do Departamento de Ciências da Saúde da Universidade Católica Portuguesa desenvolveram um algoritmo que permite a determinação de interações proteína-proteína. Este algoritmo utiliza a informação estrutural e de interações anteriormente determinadas entre os vários domínios que constituem cada proteína (3).

O objetivo desta tese é o de analisar, interpretar e explorar os resultados da aplicação do algoritmo OralInt ao conjunto de proteínas identificado por técnicas de proteómica como presentes na cavidade oral utilizando a base de dados OralOme, assim como verificar a aplicação do mesmo em termos da identificação de interações com significado biológico na cavidade oral.

Os estudos de sistemas biológicos realizados com técnicas de alto rendimento, nomeadamente das ciências “ómicas”, estão a dar origem a uma variedade e quantidade de dados sem precedentes (2). A visualização desses dados é um aspeto crucial para a sua interpretação. A dificuldade está em criar visualizações integradas, claras e com significado biológico, ultrapassando as dificuldades inerentes à complexidade intrínseca dos dados (4). Atualmente existem muitos métodos e ferramentas para a visualização de dados à escolha dos utilizadores.

As redes de interações são universais na biologia. Uma enorme quantidade de dados está a ser reunida para delinear redes de vários tipos a vários níveis, incluindo as redes de interação proteína-proteína (5). Estas redes de dados permitem ampliar e complementar uma grande quantidade de informação disponível nas ciências biomédicas. Apesar de diferentes bases de dados poderem parecer completamente distintas no que concerne à quantidade e qualidade das mesmas, são todas reflexo do sistema biológico que lhe é subjacente. Assim sendo, a inclusão dos dados, juntamente com a sua visualização integrada torna-se fundamental (6).

Na cavidade oral existem várias interações entre os microrganismos e o hospedeiro e o nosso objetivo é verificar que interações são documentadas pelo estudo das proteínas depositadas no OralCard. Do ponto de vista biológico os processos genéricos de adesão, digestão enzimática e sinalização celular, por exemplo, são passíveis de envolver interações entre proteínas que deverão ser evidentes nos nossos resultados. O conhecimento e compreensão molecular

destes processos permite alcançar o conhecimento total dos microrganismos que habitam a cavidade oral, das moléculas por eles produzidas e das interações destas com as do hospedeiro.

### **1.1. Microbioma humano**

No nosso organismo, existem dez vezes mais células bacterianas que eucariotas. Além de existirem muitos microrganismos espalhados pelo organismo do ser Humano, este microbioma é muito variado, mesmo em indivíduos saudáveis (7). Os locais onde há maior variabilidade de microrganismos, são a cavidade oral e orofaringe, pele, intestinos e vagina (7). Esta variabilidade está associada aos diversos fatores do hospedeiro, que incluem: temperatura, humidade, disponibilidade de oxigénio e nutrientes, osmolaridade, recetores do hospedeiro, competição e resistência a outros microrganismos, assim como o sistema imune do hospedeiro (8).

Ao conjunto das comunidades de microrganismos (incluindo bactérias, vírus e eucariotas microbianos) que povoam o organismo Humano, dá-se o nome de Microbioma Humano (9). Estas comunidades microbianas são de vital importância para a nossa saúde, e o seu estudo leva a um melhor conhecimento da dinâmica complexa, que é estabelecida dentro desta comunidade e entra a comunidade e o hospedeiro, e pode conduzir ao desenvolvimento de novas formas de diagnóstico e mesmo de tratamento (10).

A necessidade de obter informação sobre o microbioma Humano levou à criação do Human Microbiome Project (HMP), cujo principal objetivo consiste na identificação do microbioma nuclear em pessoas saudáveis. O conceito de microbioma nuclear assenta, 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 projeto HMP, são utilizados 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) (8), que ao serem

aplicados ao microbioma Humano, proporcionam um método estandardizado para criar, processar e interpretar os diferentes tipos de resultados de metagenômica disponíveis na comunidade científica (11). Para obter tais resultados, o projeto HMP recorre a vários campos da biologia celular, microbiologia, genética e imunologia (10).

O HMP promove a exploração das comunidades microbianas, e a sua relação com o hospedeiro Humano, estando a ser anotado e analisado atualmente 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 das associações pode levar ao desenvolvimento de ferramentas com implicações terapêuticas (11).

Existe uma grande variabilidade de microrganismos que colonizam o ser Humano, sendo a sua diversidade muitas vezes ligada a determinadas patologias, tanto pela sua falta de variabilidade (por exemplo, no caso da doença inflamatória intestinal) como por excesso (como é o caso da vaginose bacteriana). 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 fatores funcionalmente relevantes, que podem exercer uma pressão seletiva entre espécies. Os *taxa* menos dominantes parecem ser bastante característicos, quer entre indivíduos, quer entre habitats dentro do mesmo organismo (7).

Apesar da grande diversidade 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 interação 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

(7). 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 (8).

### **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 (10). Em 2012, os resultados do projeto HMP revelam que os microbiomas com maior diversidade de microrganismos, são os do intestino e cavidade oral (7). O objetivo 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 interação entre o microbioma e o hospedeiro, e a caracterização do comportamento de espécies microbianas de forma individual (10).

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 da 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 fatores de interação com o hospedeiro, nomeadamente fatores de virulência (12, 13). 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 (10).

#### **1.1.1.1. Bactérias comensais da cavidade oral**

Calcula-se que existam cerca de 750 espécies bacterianas comensais diferentes na cavidade oral (10), 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 (10). O conhecimento atual permite afirmar que as bactérias que estão representadas em maior número na cavidade oral incluem os gêneros *Streptococcus*, *Actinomyces*, *Veillonella*, *Fusobacterium*, *Porphyromonas*, *Prevotella*, *Treponema*, *Neisseria*, *Haemophilus*, *Eubacteria*, *Lactobacterium*, *Capnocytophaga*, *Eikenella*, *Leptotrichia*, *Peptostreptococcus*, *Staphylococcus* e *Propionibacterium*, tendo a maioria destes uma relação de mutualismo com o ser Humano (10).

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 (10).

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 K<sub>2</sub> (14), (essenciais para o ser Humano). Microrganismos que dão origem a patologias também existem como é o caso de *Streptococcus pneumoniae*, responsável por infecções pulmonares em crianças e idosos (15). Este microbioma normal é conhecido por microbioma comensal, e nele estão incluídas eubactérias, archaeobacterias e fungos (10). 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) (16), 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 (10). Além destes dois 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* afeta os pulmões de pacientes com fibrose cística mas não o de indivíduos saudáveis (17, 18).

### 1.1.1.2. Heterogeneidade de microbiomas na cavidade oral

Apesar de ser conhecida a predominância de *Streptococcus*, no biofilme oral, é também aceite que existe uma grande variedade de microbiomas, 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 as bactérias com maior predominância em cada um dos nichos da cavidade oral de indivíduos saudáveis (19). 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 (7).

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* (18). 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 (18), como é exemplo a bactéria *Porphyromonas gingivalis* (20). 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 (21). Um exemplo de anaeróbio facultativo é o caso de um dos principais agentes etiológicos da cárie dentária o *Streptococcus mutans* (22). 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* (23). Finalmente há os capnófilos, espécies bacterianas que têm um crescimento ótimo a elevadas quantidades de dióxido de carbono, como é exemplo *Mannheimia succiniciproducens* (10, 24).

Uma vez que a anatomia e fisiologia da cavidade oral permite a existência de microhabitats muito variáveis em termos de concentrações de oxigênio, 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. 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 (18). 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 do antigo conceito de que os organismos procariotas se comportavam como indivíduos autossuficientes, sabe-se atualmente 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 patologias mais prevalentes neste campo da medicina dentária: 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 (25). Estes complexos microbianos além das células bacterianas, são

constituídos por uma matriz de glicoproteínas salivares e polissacarídeos extracelulares (26).

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 recetores salivares, como é o caso das proteínas ricas em prolina e a estaterina, têm epítomos reconhecidos por *Streptococcus*. Outras moléculas como a  $\alpha$  - amilase, imunoglobulinas, fibronectina e lactoferrina, também se unem a adesinas à superfície dos colonizadores iniciais, conferindo-lhes vantagem sobre outras bactérias com recetores com menos afinidade para as moléculas da película dentária (27).

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 recetores oligossacarídeos, reações do tipo lectina, interações do tipo proteína-proteína, ou ainda ligações iónicas ou hidrofóbicas, conferindo assim uma união bastante firme à superfície que estão a colonizar (27).

Os recetores à superfície das células microbianas constituem um possível alvo terapêutico, já que um potencial agente patogénico pode ser excluído do biofilme por não ter recetores adequados à sua adesão, e conseqüente 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 conseqüências tão ou mais patogénicas que o anterior (27).

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 (12).

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

gingivais e da mucosa jugal). Um exemplo são os oportunistas *Porphyromonas gingivalis* (28), *Prevotella intermedia* (29) e *Aggregatibacter actinomycetemcomitans* (10, 30).

### **1.2.2. Moléculas importantes na relação microrganismo/hospedeiro**

As interações 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 interação entre microrganismo e hospedeiro, sendo algumas consideradas fatores de virulência (atributos moleculares de determinado microrganismo patogénico utilizado para uma colonização e infeção bem sucedidas) (31). Estas moléculas podem ser de origem proteica, lipídica ou de hidratos de carbono (32) 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 (32) e ii) as moléculas associadas a uma interação funcional com o hospedeiro onde se podem inserir as toxinas e proteases 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 (33). Muitos destas moléculas são alvos de investigação com o objetivo de fazer o controlo e modulação do desenvolvimento do biofilme oral.

### **1.3. Interatoma**

A palavra interatoma foi originalmente proposta em 1999 por um grupo de cientistas franceses liderados por Bernard Jacq (34). Em biologia molecular, interatoma refere-se a todo o conjunto de interações moleculares que se desenrolam numa determinada célula. Neste trabalho, o termo refere-se especificamente às interações físicas entre as moléculas, contudo, em outro contexto também pode significar as interações indiretas entre os genes, ou

seja, interações genéticas. Matematicamente, os interatomos são geralmente apresentados como gráficos (35).

Para caracterizar um dado sistema biológico é necessário conhecer, as interações entre todas as moléculas que nele participam, ou seja, conhecer o interatoma desse sistema biológico. À luz da interatômica, qualquer processo biológico é encarado não como resultado de uma ação independente de cada molécula interveniente, mas sim como uma rede de interações moleculares. Ainda assim, a interatômica não permite descrever totalmente o funcionamento de um sistema biológico, em que para além de saber que determinada molécula interage com outra molécula, é necessário saber em que contextos anatómicos e fisiológicos é que o fazem (2).

### **1.3.1. Interatômica**

A Interatômica é uma disciplina inserida na bioinformática e biologia que se ocupa do estudo das interações entre proteínas e outras moléculas, assim como das consequências dessas mesmas interações dentro de uma célula (36). Interatômica visa, portanto, comparar as redes de interações (interatomos) entre e dentro das espécies, com o objetivo de descobrir como é que as características dessas redes são conservadas ou alteradas.

### **1.3.2. Interatoma bacteriano**

Sabe-se que relativamente poucas bactérias foram exaustivamente estudadas através das suas interações proteína-proteína, não existindo até ao momento nenhum interatoma totalmente completo (37). Entre os interatomos bacterianos publicados (incluindo os parciais) encontra-se: *Helicobacter pylori* (37), *Campylobacter jejuni* (38), *Treponema pallidum* (39), *Escherichia coli* (40), *Mesorhizobium loti* (41), *Mycobacterium tuberculosis* (42), *Mycoplasma genitalium* (43), *Synechocystis sp.* (44).

### **1.3.3. Interatoma Humano - patogéno**

Apesar da interatômica ser na maioria das vezes considerada numa espécie (ou célula) em particular, neste trabalho o interesse é focado nas interações moleculares entre organismos diferentes. Por um lado entre os organismos microbianos e o hospedeiro humano (em ações simbióticas) e por outro lado as interações entre os vários microrganismos que permitem e determinam a estrutura do biofilme presente.

Os patogénicos bacterianos e virais, afetam a célula eucariota do hospedeiro, em parte, através da interação com as proteínas da célula do hospedeiro. Assim sendo, para melhor compreendermos os mecanismos envolvidos na infeção, é necessário a construção de redes de interação proteína – proteína hospedeiro- patogéno completas e precisas (45).

A nível molecular, a infeção pode ser vista como a interferência das proteínas patogénicas (a maioria das toxinas microbianas são proteínas) na rede de interação do hospedeiro. As proteínas virais utilizam a maquinaria da síntese proteica do hospedeiro, enquanto as bactérias e proteínas dos protozoários são secretadas ou injetadas nas células do hospedeiro (45). Uma estratégia comum dos agentes patogénicos, baseia-se em imitar domínios das interações proteína – proteína eucariotas com o objetivo de remodelar a rede das interações proteína – proteína do hospedeiro (46). Estes mimetismos estruturais são muitas vezes o resultado da transferência horizontal de genes (HGT) (46), em combinação com a rápida evolução dos genomas patogénicos, como parte da coevolução do hospedeiro – patogéno. Estas interações proteína – proteína, podem determinar a virulência do patogéno (47).

### **1.3.4. O interatoma oral Humano - bacteriano**

A compreensão dos mecanismos moleculares que se desenrolam na cavidade oral assenta na interação entre as moléculas presentes.

Praticamente, todos os eventos celulares, tais como a transdução de sinal, transporte intracelular, replicação de ADN, transcrição, tradução, splicing,

secreção, controlo do ciclo celular e intermediários metabólicos são mediados por interações Proteína-proteína (PPIs). No sistema patogéno-hospedeiro as PPIs são essenciais para estabelecer a infeção, logo, o conhecimento sobre estas poderá contribuir para elucidar o papel da flora no desenvolvimento da doença (3). O conhecimento destes mecanismos tem aplicações na fisiologia e fisiopatologia do sistema cavidade oral. Estas interações são importantes no estabelecimento e manutenção da flora oral capaz de desencadear a formação da placa bacteriana, responsável pelo desenvolvimento das cáries e da doença periodontal (48).

A colonização microbiana da cavidade oral é mediada por PPIs entre o hospedeiro e os microrganismos (2). No sistema patógeno-hospedeiro as PPIs são essenciais para estabelecer a infeção, logo, o conhecimento sobre estas poderá contribuir para elucidar o papel da flora no desenvolvimento de determinadas patologias (2).

### **1.3.5. Interações proteína-proteína**

O estudo das interações proteína-proteína, tem vindo a aumentar nos últimos anos, principalmente como resultado do rápido acesso aos dados experimentais obtidos através técnicas de alto rendimento (49).

A maioria dos produtos génicos que constituem uma célula viva, interagem nem que seja transitoriamente com outras moléculas de proteínas. Virtualmente todos os eventos celulares, tais como a transdução de sinal, o transporte intracelular, a replicação do ADN, a transcrição, a tradução, *splicing*, secreção, controlo do ciclo celular e o metabolismo intermediário, são mediados através de interações proteína – proteína (PPIs) (50). O mesmo se aplica no sistema hospedeiro patogénico, onde os sistemas PPIs são essenciais para o estabelecimento da infeção (51).

Uma vez que cada complexo proteína-proteína produz uma resposta no contexto fisiológico, o seu compromisso pode induzir um estado de doença. É por esta razão, que uma adequada identificação e caracterização das PPIs e das suas redes de interação se torna essencial na compreensão dos mecanismos que se encontram por trás de processos biológicos a nível molecular (52).

Existem diferentes classificações dentro das interações proteicas. Dependendo da força, uma interação pode ser permanente ou temporária, em relação à especificidade esta pode ser específica ou não específica, quanto à localização das proteínas intervenientes na interação, estas podem interagir dentro de uma ou em duas cadeias polipeptídicas, e por último as subunidades que interagem podem ser classificadas em homo ou hétero – oligómeros, tendo em conta as semelhanças que partilham (53).

Independentemente dos tipos de interação, para que as proteínas se liguem umas as outras, os pares da interação devem possuir uma elevada afinidade estrutural e físico-química, ou seja, devem ser complementares. Consequentemente, espera-se que os pares de proteínas que interagem partilhem um certo grau de conservação nas regiões de ligação (54, 55).

Vários métodos experimentais foram desenvolvidos para determinar as redes de interação, ou interatoma de um organismo. Os métodos experimentais de alto rendimento foram os mais populares em identificações de PPIs em larga escala, incluindo entre outros, ensaio de duplo híbrido em levedura (56), espectrometria de massa (57), *chips* de proteínas (58), exibição de fagos (59), estudo da letalidade sintética (60), perfil de expressão de mRNA correlacionado (61), *microarrays* de proteínas (62). Contudo, os ensaios experimentais têm as suas desvantagens, sendo que as duas principais são: a grande quantidade de tempo que é necessário para a sua realização, e o seu custo considerável. Além disso, a cobertura destes ensaios a nível da rede das PPIs é mínima. Devemos também ter em conta que as técnicas de alto rendimento nos induzem em erro, e possuem uma baixa especificidade, gerando um elevado número de falsos negativos, e falsos positivos (63). É assim essencial encontrar alternativas a estas técnicas para que se possa prever elevadas porções da rede de interações das proteínas humanas de forma correta e económica. A alternativa mais promissora, é a utilização de uma abordagem computacional que utiliza dados das PPIs obtidos através de métodos experimentais, com o objetivo de prever uma rede mais ampla de PPIs (49).

Com a finalidade de prever com precisão PPIs em larga escala, foram propostas uma série de abordagens computacionais (64, 65). Podemos assim agrupar os métodos computacionais em 5 grandes grupos: métodos baseados

na sequência genómica, na estrutura da proteína, nos dados do domínio, na sequência da proteína, na semelhança semântica (49).

#### **1.4. Ferramentas bioinformáticas e o estudo do interatoma da cavidade oral**

Para fazer o estudo do interatoma 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 sendo, para a compreensão do proteoma e do interatoma oral Humano é necessário estudar não só as proteínas orais de origem Humana, mas também as de origem microbiana.

Existem várias ferramentas bioinformáticas associadas ao armazenamento e análise de dados relativos a proteínas, tais como: UniProtKB (UniProt KnowledgeBase (66), PANTHER (Protein ANalysis THrough Evolutionary Relationships (67), BRENDA (the Comprehensive Enzyme Information System (68) e VFDB (Virulence Factor DataBase (69).

No sentido de responder à necessidade de olhar para os genes e proteínas de forma integrada, dando realce às interações entre eles, começaram também a ser desenvolvidas bases de dados de interatómica. Entre elas, podemos destacar a STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) (70, 71), que inclui interações conhecidas e previstas através de algoritmos, associando-lhes um índice de confiança. Existem outras bases de dados de interatómica porém, de mais difícil utilização, desenhadas com o objetivo primário de serem um repositório de interações, como sejam o UniHI (Unified Human Interactome) (72) BioGRID (Biological General Repository for Interaction Datasets) (73) e a DIP (Database of Interacting Proteins) (74) ou então com fins mais específicos, como por exemplo a PsiBase (75) que apresenta interações baseadas em aspetos estruturais das proteínas.

No que respeita aos dados de interatómica, algumas bases de dados optaram por um sistema de visualização próprio, como o caso do STRING (70, 71). No entanto, a maioria optou por utilizar as potencialidades de um *software* próprio para visualização de dados multifacetado e de utilização transversal,

designado Cytoscape (6). Este *software* é uma plataforma de visualização de redes complexas que permite a integração das redes com diferentes tipos de dados, sendo possível definir para cada rede de interações o índice de confiança nessa interação, o tipo de informação a apresentar relativamente à interação (Evidência da interação, índice de confiança, tipo de ações entre as moléculas que interagem) assim como o número de proteínas que interagem com cada proteína. Como é de código livre, permite que diversos grupos de desenvolvimento de *software* o adaptem às mais diversas situações, criando *plugins* para o efeito (6).

#### **1.4.1. OralOme e OralCard**

Existem 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/](http://www.hmpdacc.org/)), e mais especificamente sobre o microbioma oral Humano, HOMD (Human Oral Microbiome Database (76)). 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 (2). Assim, esta ferramenta bioinformática e a informação nela incluída, vão 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, assim como na descoberta e desenvolvimento de novos agentes terapêuticos (2).

### **1.4.2. OralInt**

O OralInt é um modelo computacional que prevê PPIs inter-espécies na cavidade oral do ser humano, em um ambiente propenso à colonização bacteriana (2, 3). Este algoritmo foi desenvolvido por Investigadores do Instituto de Engenharia, Eletrónica e Telemática da Universidade de Aveiro e do Departamento de Informática da Universidade de Coimbra juntamente com investigadores do Departamento de Ciências da Saúde da Universidade Católica Portuguesa, e utiliza a informação estrutural e de interações anteriormente determinadas entre os vários domínios que constituem cada proteína.

O resultado dos dados de proteómica inseridos no modelo do OralInt é elaborado segundo os níveis de confiança das PPIs. A previsibilidade deste modelo advém do facto de os domínios da ligação das proteínas revelarem elevada afinidade estrutural e físico-química, com um grau de conservação associado e ainda pelas evidências de que os seus homólogos proteicos interagem frequentemente do mesmo modo (3).

O conjunto de dados inseridos no modelo OralInt, podem conter não só as vias mais importantes envolvidas nas doenças infecciosas orais, mas também potenciais alvos farmacológicos e marcadores biológicos relacionados com o processo de estabelecimento e progressão da doença, levando ao desenvolvimento de novos fármacos e permitindo ainda a caracterização das respostas celulares a determinadas drogas, doenças e mudanças ambientais.



## 2. OBJETIVOS

O objetivo deste trabalho, centra-se em testar o algoritmo Orallnt (<http://bioinformatics.ua.pt/software/orallnt>) (3) de previsão de interações proteína - proteína passíveis de ocorrer na cavidade oral, assim como verificar a aplicação do mesmo em termos da identificação de interações com significado biológico.

Este estudo pode fornecer informação com implicações na compreensão dos mecanismos moleculares dos processos que ocorrem na cavidade oral em situações tanto fisiológicas como patológicas. As informações sobre estas interações podem melhorar a nossa compreensão das doenças, bem como fornecer a base para novas abordagens terapêuticas e de diagnóstico.



### 3. MATERIAIS E MÉTODOS

#### 3.1. Proteoma oral

Como ponto de partida para o nosso estudo, foram utilizadas 3572 proteínas, 2747 humanas e 825 bacterianas, disponíveis na base de dados do OralCard (2, 77). Estas proteínas foram identificadas através da análise proteómica da saliva, normalmente realizada através da 2D electroforese/espectrometria de massa ou 2D cromatografia líquida/espectrometria de massa. Para que melhor possam ser compreendidas as premissas subjacentes à criação do algoritmo OralInt, serão brevemente descritos os aspetos principais.

Uma vez que não existe um *gold standard* bem estabelecido para as PPIs, foram recolhidos dados de cinco base de dados que continham interações de elevada confiança determinadas experimentalmente. As bases de dados que serviram de apoio para a recolha de proteínas foram: BIOGRID (Biological General Repository for Interaction Datasets) (78), DIP (Database of Interacting Proteins) (74), HPRD (Human Protein Reference Database) (79), IntAct (80) e MINT (the Molecular INTeraction database) (81). As PPIs extraídas foram assim incorporadas, criando o *gold standard* de interações positivas. O *gold standard* de interações negativas foi obtido através de uma seleção aleatório da lista de proteínas, com a premissa de que todos os pares de proteínas produzidos deveriam ser diferentes dos que se encontravam no conjunto de dados positivo (3).

Simultaneamente, para cada par de proteínas possível, foram construídos cinco grandes grupos de características, baseadas em: Literatura, sequência primária de informação da proteína, perfis de ortólogos, semelhanças no processo biológico, enriquecimento conservado nos domínios dos pares. A realização deste último passo foi possível através do acesso aos bancos de dados públicos, com a posterior extração dos dados e por último o processamento de todos os dados recolhidos. O conjunto de dados padrão foi

utilizado para treinar um classificador Naive Bayes para realizar novas validações no modelo final. O classificador foi então aplicado ao conjunto de todos os pares de interações possíveis. Por último, através da junção de todos os pares individuais de interações previstas, foi obtida a rede (3).

### **3.2. Análise das redes de interatômica**

O Cytoscape (<http://www.cytoscape.org/>) foi a ferramenta utilizada para visualizar e validar as redes de interação entre as proteínas obtidas. As PPIs foram classificadas como “HUMANO-HUMANO”, se as proteínas que interagem são apenas de origem humana, como “MICRO-MICRO”, se as proteínas que interagem são apenas de origem microbiana, ou como “HUMANO-MICRO”, se as proteínas que interagem são de origem tanto humana como bacteriana. Os dados da rede de interações foram importados para o Cytoscape em que foi definido que das duas proteínas que constituem o par de interação, uma seria a fonte da interação (proteína um) e a outra o alvo da interação (proteína dois). O tipo de interação escolhida foi a classificação organismo-organismo mencionado anteriormente.

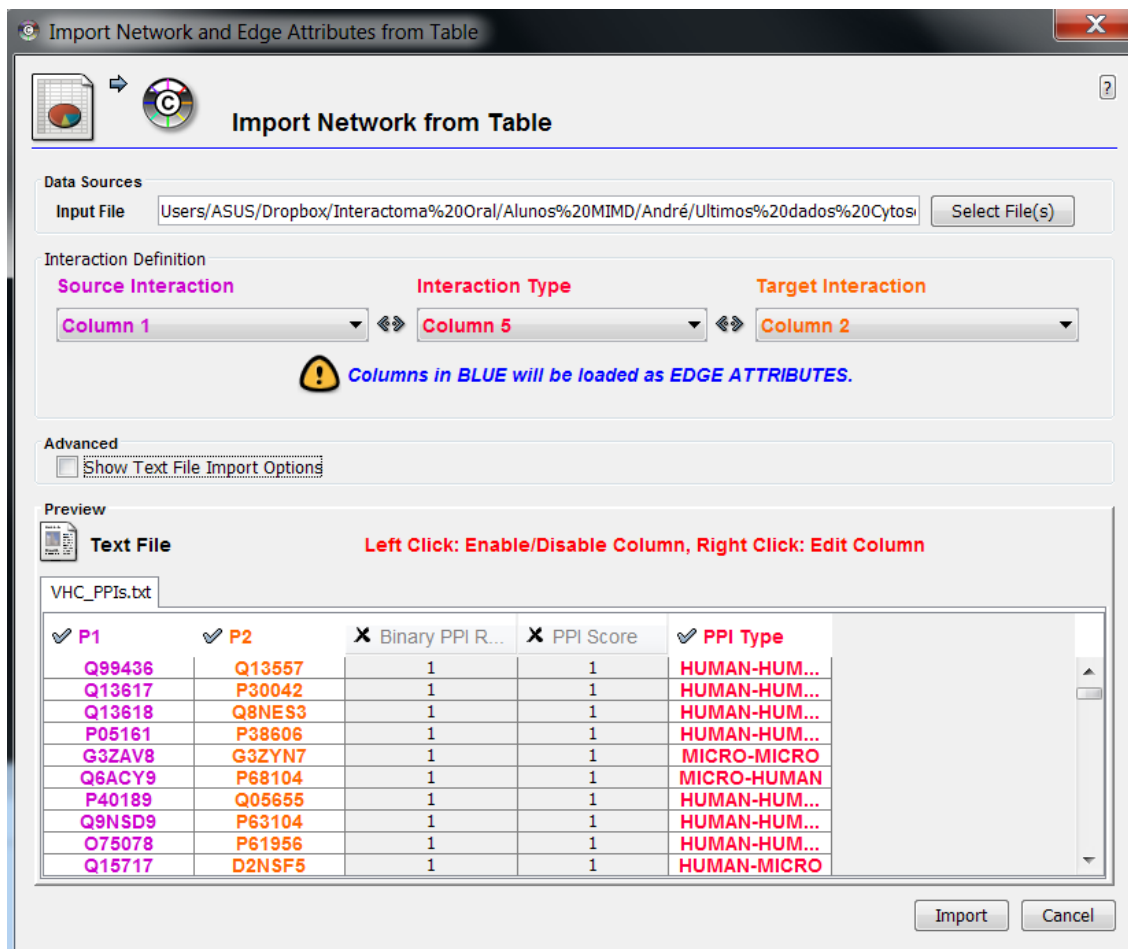


Figura 1 – Dados importados para o programa Cytoscape, com a definição da fonte e alvo de interação, assim como a classificação do tipo de interação.

Um ficheiro com os atributos dos nós foi também importado, onde constava a informação referente aos microrganismos e processos biológicos, extraída da base de dados do UniProt, pertencente a cada proteína presente na nossa rede de interações.

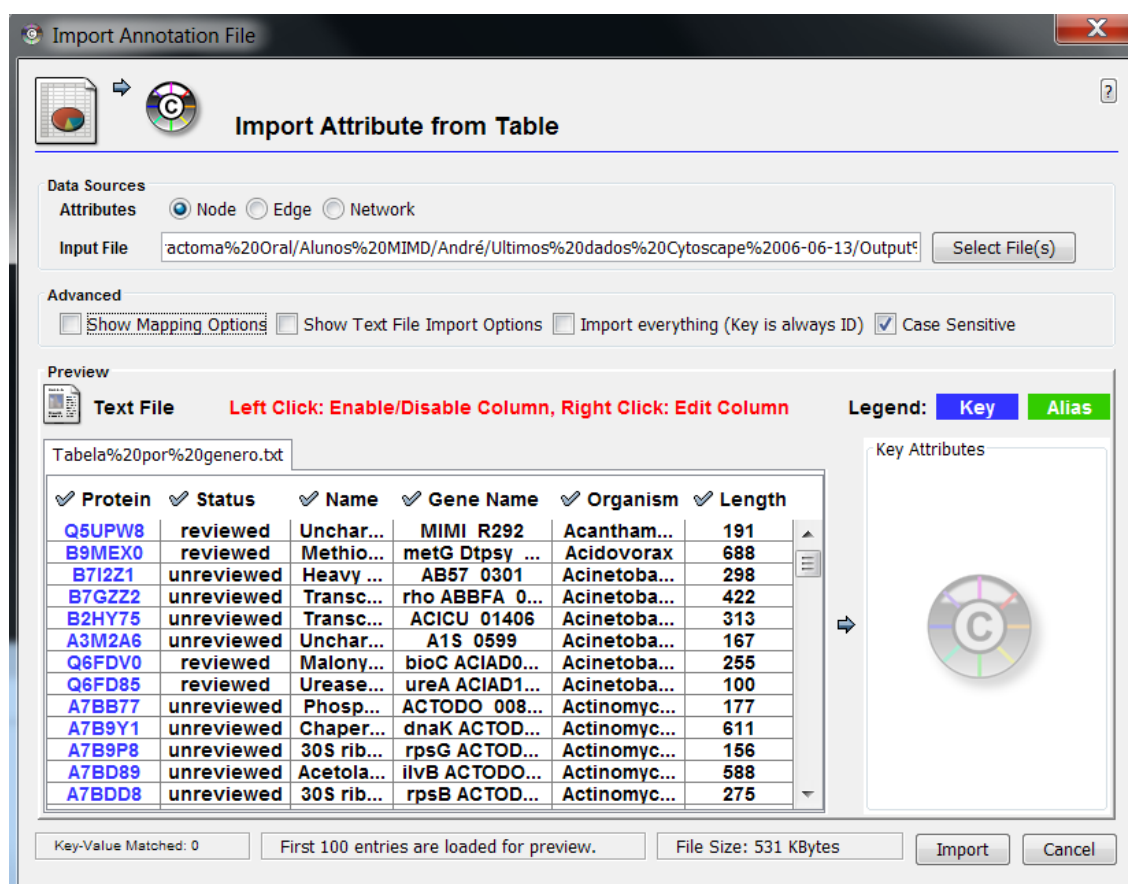
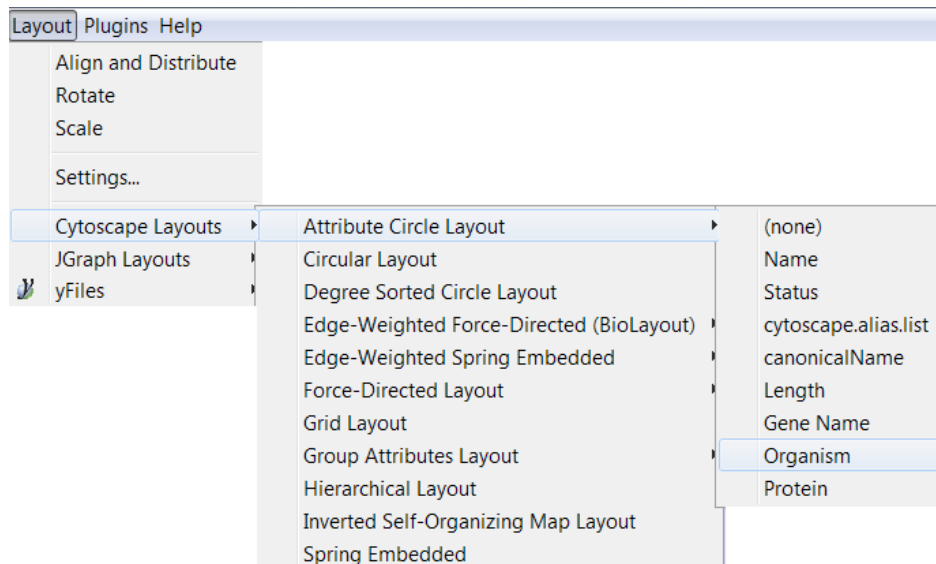


Figura 2 – Dados importados para o programa Cytoscape, com a informação referente aos microrganismos.

O *layout* escolhido para a visualização das redes de interação foi o que colocava as proteínas por organismo, uma vez que seria o de mais fácil compreensão e visualização, devido ao grande número de dados.



**Figura 3 - Escolha do *layout*, recorrendo às ferramentas informáticas do programa Cytoscape.**

### **3.3. Análise da classificação ontológica das proteínas presentes nos organismos em estudo**

Para se obterem as ontologias referentes às proteínas envolvidas nas interações dos géneros *Aggregatibacter*, *Porphyromonas*, *Rothia* e *Streptococcus*, foi utilizada a ferramenta bioinformática UniProtKB, inserindo os códigos de acesso correspondentes às proteínas identificadas. Esta base de dados possui informação estrutural e funcional sobre cada proteína, bem como os seus diferentes nomes, estados de revisão, organismos onde foram identificadas, domínios, interações, estruturas tridimensionais, entre outros.

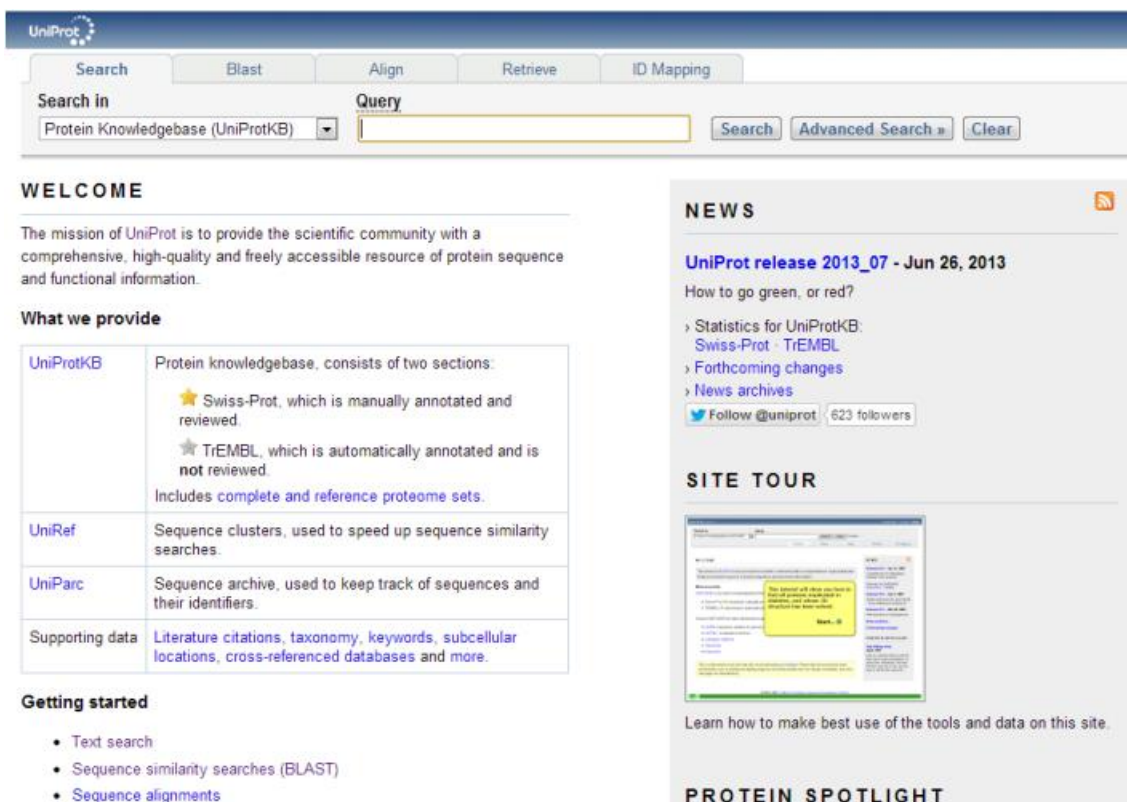


Figura 4 – Página de rosto da base de dados do UniprotKB.

Nesta base de dados, recorreu-se à secção *search*, inserindo-se os respetivos códigos das proteínas identificadas.

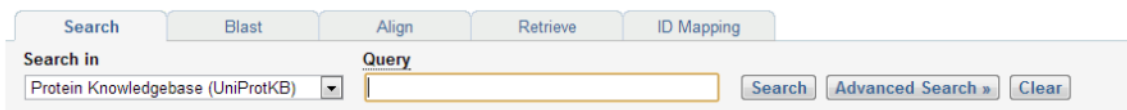


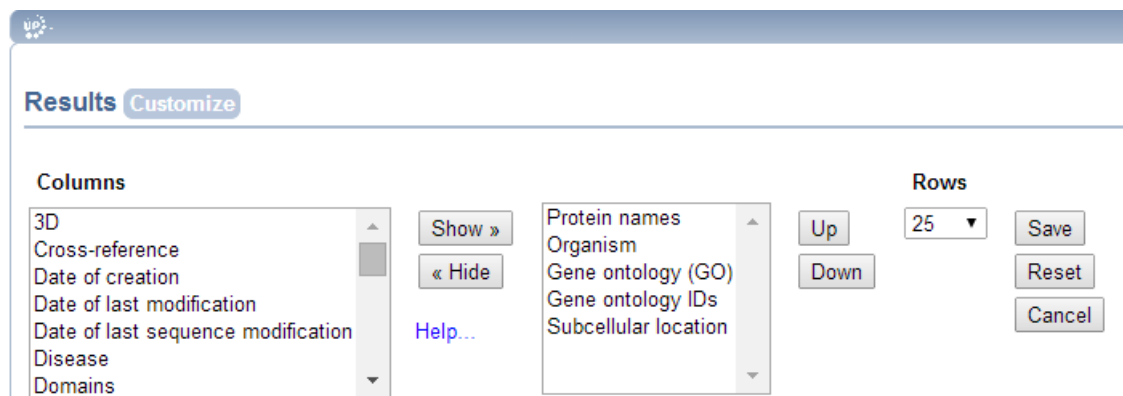
Figura 5 – Secção *search* da base de dados do UniprotKB.

A ferramenta *retrieve* permitiu obter informações específicas e adaptadas às necessidades do utilizador para as proteínas anteriormente obtidas.



Figura 6 – Ferramenta *retrieve* da base de dados do UniprotKB.

De seguida, foram selecionadas as características de interesse ao estudo das proteínas, tendo sido escolhidos o nome da proteína, função, localização celular, microrganismo da qual origina, identificação da ontologia e descrição da ontologia. Após as informações de interesse terem sido selecionadas, foi feito o *download* de um ficheiro *.xls/xw*, e a tabela foi posteriormente trabalhado no Microsoft Excel® 2013.



**Figura 7 – Obtenção da informação sobre as proteínas identificadas, utilizando a ferramenta *customize* da base de dados do UniprotKB.**

Após termos obtido as informações selecionadas de todas as proteínas, foram identificadas aquelas que possuíam ontologias com maior relevância para a interação com o hospedeiro, nomeadamente as que dizem respeito a ontologias extracelulares como sendo *Extracellular Region* (GO:0005576), *Plasma Membrane* (GO:0005886) e *Cell Wall* (GO:0005618).

Identificadas as proteínas que possuíam ontologias extracelulares, foi realizada uma rede de interações individual para cada proteína, para se poder efetuar um estudo mais detalhado sobre as mesmas, assim como uma pesquisa bibliográfica com recuso à ferramenta de pesquisa avançada do portal PubMed, (<http://www.ncbi.nlm.nih.gov/pubmed/>), usando como palavras-chave o nome das proteínas identificadas, assim como os termos MeSH: “interaction domains”, “interatomic”, “host-pathogen interaction”. Com base nos artigos recolhidos, foi possível adicionar informação sobre as interações realizadas pelas proteínas em questão.



## 4. RESULTADOS E DISCUSSÃO

### 4.1. Interatoma da cavidade oral

O algoritmo OralInt foi aplicado no estudo de 3572 proteínas de origem humana e bacteriana que estabelecem entre si exclusivamente interações de elevada confiança (Probabilidade  $\geq 1 \cdot 10^{-7}$ ) (3).

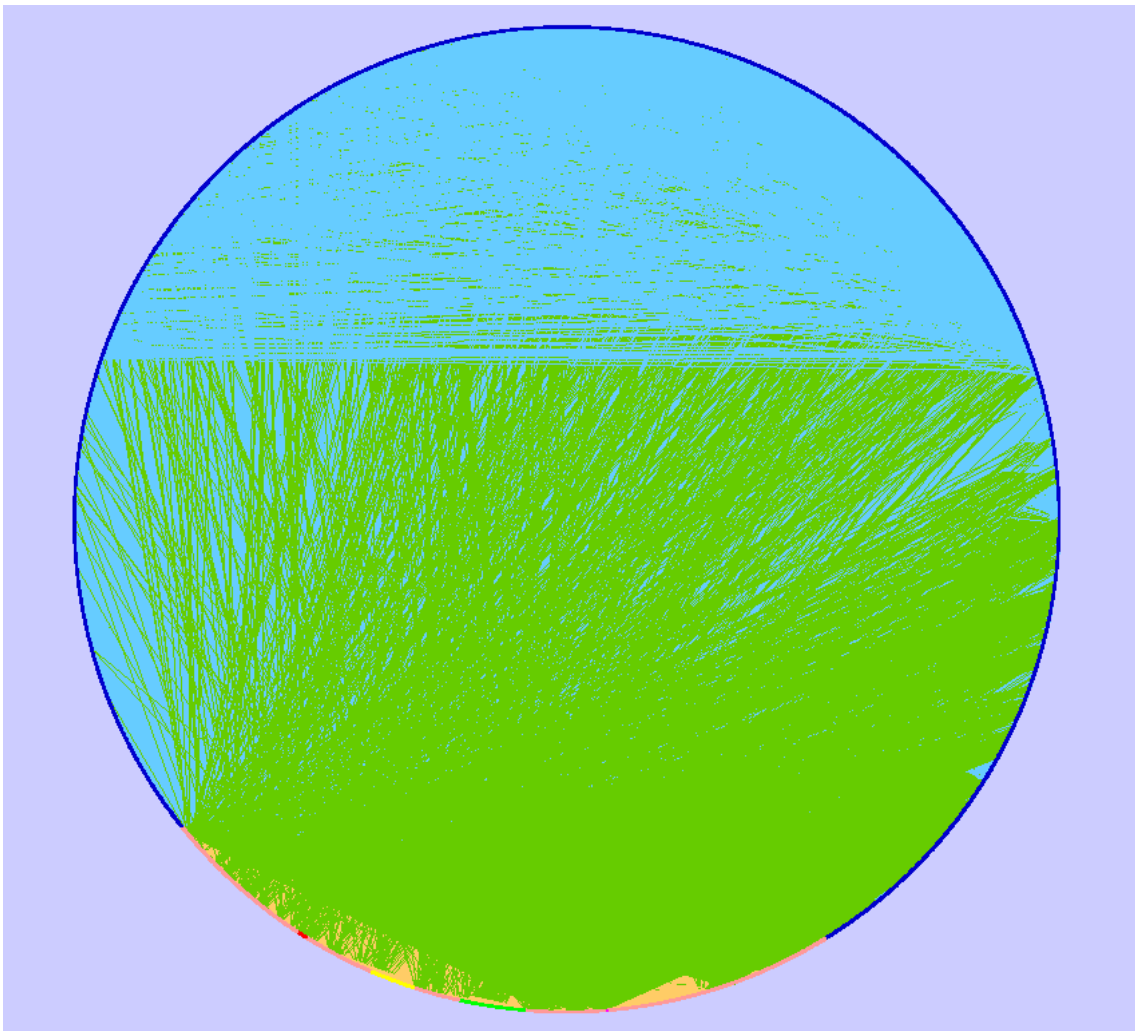
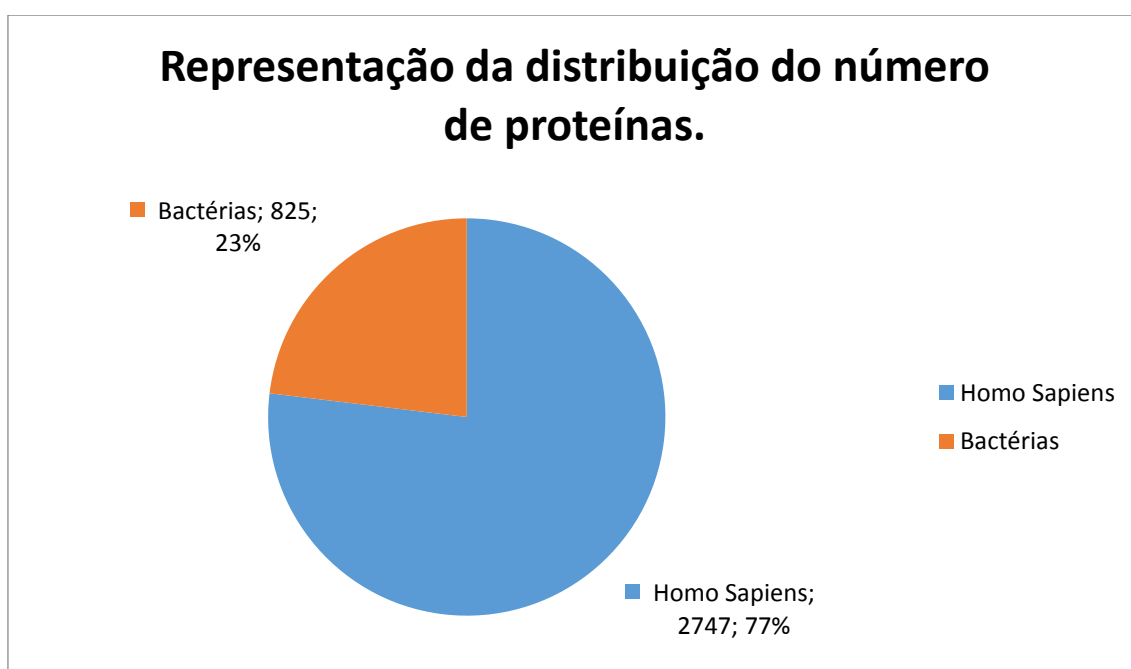


Figura 8 – Interações estabelecidas entre as proteínas orais humanas e bacterianas. Os nós a rosa representam as proteínas produzidas pelos organismos de origem microbiana, enquanto os nós a azul representam as proteínas orais de origem humana. Podemos ainda observar as proteínas produzidas pelo organismo *Aggregatibacter* a púrpura, *Porphyromonas* a vermelho, *Streptococcus* a verde, e por último *Rothia*, a amarelo. As interações estabelecidas entre proteínas de origem humana encontram-se a azul, a verde as realizadas entre proteínas humanas e microbianas, enquanto as interações estabelecidas entre proteínas microbianas encontram-se a amarelo.

Na visualização das redes de interação, cada nó representa uma proteína produzida por um organismo distinto, ligados entre si por interações que foram ponderadas consoante o tipo de interação estabelecida com base na origem do organismo (humano ou bacteriano). Esta abordagem originou uma rede global de interação única com um total de 97032 interações entre as 3572 proteínas, como pode ser observado na figura 8.

Das 3572 proteínas presentes, a maioria (2747 ou 77%) são de origem humana, sendo as bacterianas representadas por apenas 825 proteínas, o que corresponde a 23%, como se pode observar na figura 9.



**Figura 9 – Distribuição do número total de proteínas humanas e bacterianas analisadas pelo algoritmo Orallnt.**

Este resultado é inesperado face ao número de microrganismos diferentes que se sabe colonizarem a cavidade oral. Assim, o facto de 77% das proteínas terem origem num único organismo (*Homo sapiens*), enquanto os restantes 274 organismos dão origem apenas a 23% das proteínas presentes, permite suspeitar que a cobertura em termos de proteínas microbianas identificadas é muito baixa.

As 3572 proteínas identificadas estabelecem entre si 97032 interações, em que 84196 ou 87% do total das interações é estabelecido entre proteínas humanas, 11853 ou 12% entre proteínas bacterianas e humanas, e apenas 983

ou 1% corresponde a interações bacterianas como pode ser observado na figura 10.

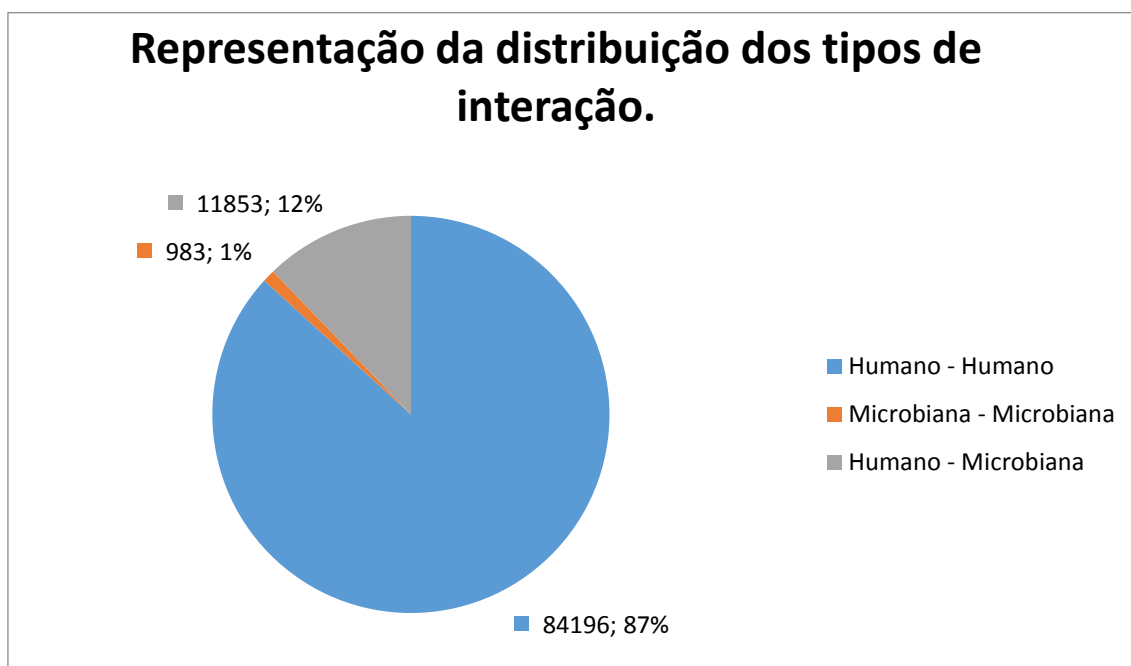


Figura 10 – Distribuição do número total dos diferentes tipos de interação estabelecidos entre as proteínas identificadas pelo algoritmo Oralnt.

O número de interações estabelecidas entre proteínas bacterianas é escasso, podendo ser explicado pela falta de cobertura de proteínas de origem bacteriana identificadas mas também pela falta de anotação que a maioria destas proteínas tem nas bases de dados.

A maioria das espécies bacterianas é identificada por poucas proteínas (uma ou duas), sendo a espécie *Rothia mucilaginosa* aquela em que foi identificada um maior número de proteínas como se pode observar na figura 11.

No que diz respeito ao número de proteínas de origem microbiana presentes por género de microrganismo, o género *Streptococcus* é o que possui um maior número de proteínas como se pode constatar na figura 12, seguido por *Rothia* com 53 proteínas, *Lactobacillus* com 43 proteínas e por último *Neisseria* com 38 proteínas, sendo estes os géneros com maior representação.

## Representação do número de proteínas de origem microbiana presentes por espécie de microrganismo

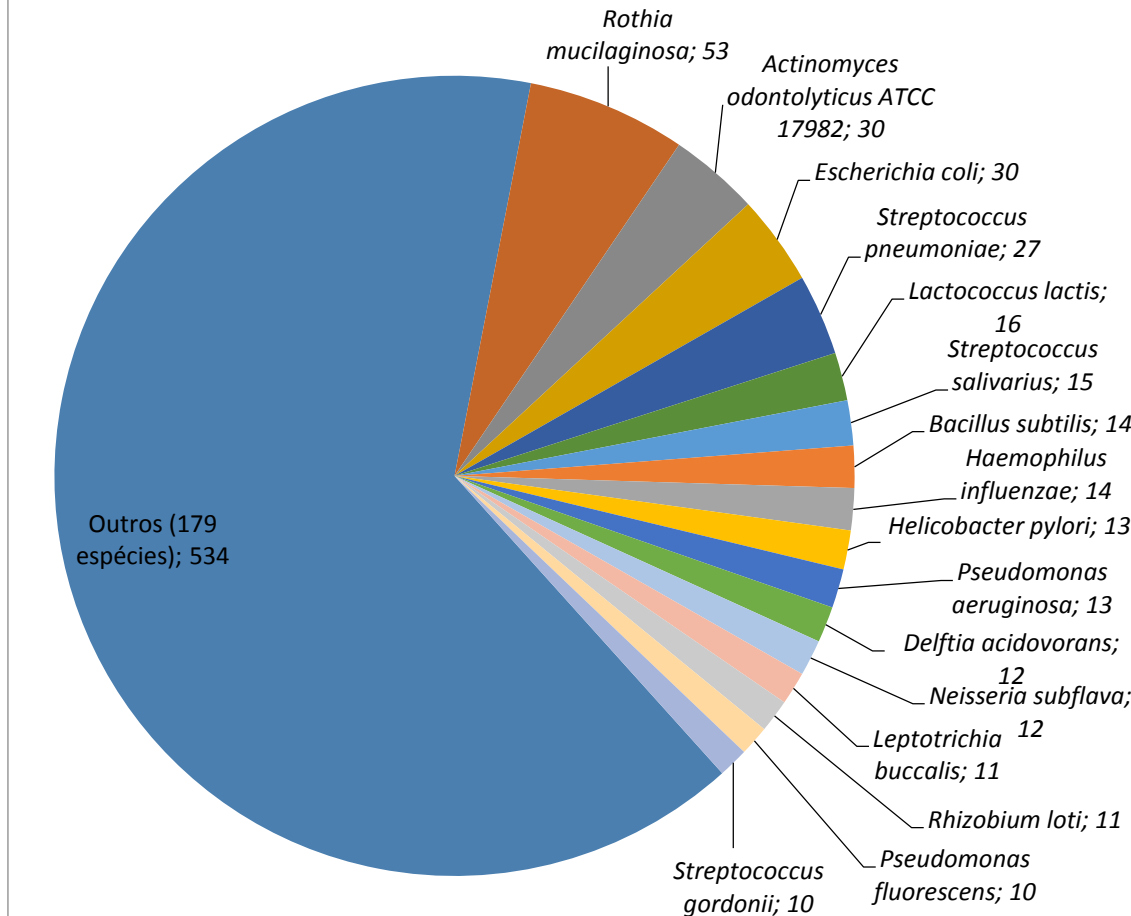


Figura 11 – Número de proteínas de origem microbiana presentes por espécie de organismo na cavidade oral, analisadas pelo algoritmo Orallnt.

Da análise da figura 11 podemos concluir a presença de 195 espécies microbianas, sendo a quase totalidade destas bactérias. Surgem também proteínas de uma levedura *Saccharomyces cerevisiae*, e de dois vírus *Acanthamoeba polyphaga mimivirus* (APMV) e *Carnation etched ring virus* (CERV). A espécie de microrganismos presente na cavidade oral, que possui um maior número de proteínas é *Rothia mucilaginosa*. A barra do gráfico que se encontra legendada com o nome “Outros” representa o conjunto das 179 espécies de microrganismos cuja identificação resulta num número de proteínas compreendidas entre 1 e 10.

Uma vez que a quantidade de proteínas que representa cada espécie é muito heterogênea, e de maneira a podermos incluir mais espécies nesta representação, optou-se por realizar também uma análise por género de organismo, como se pode observar na figura 12.

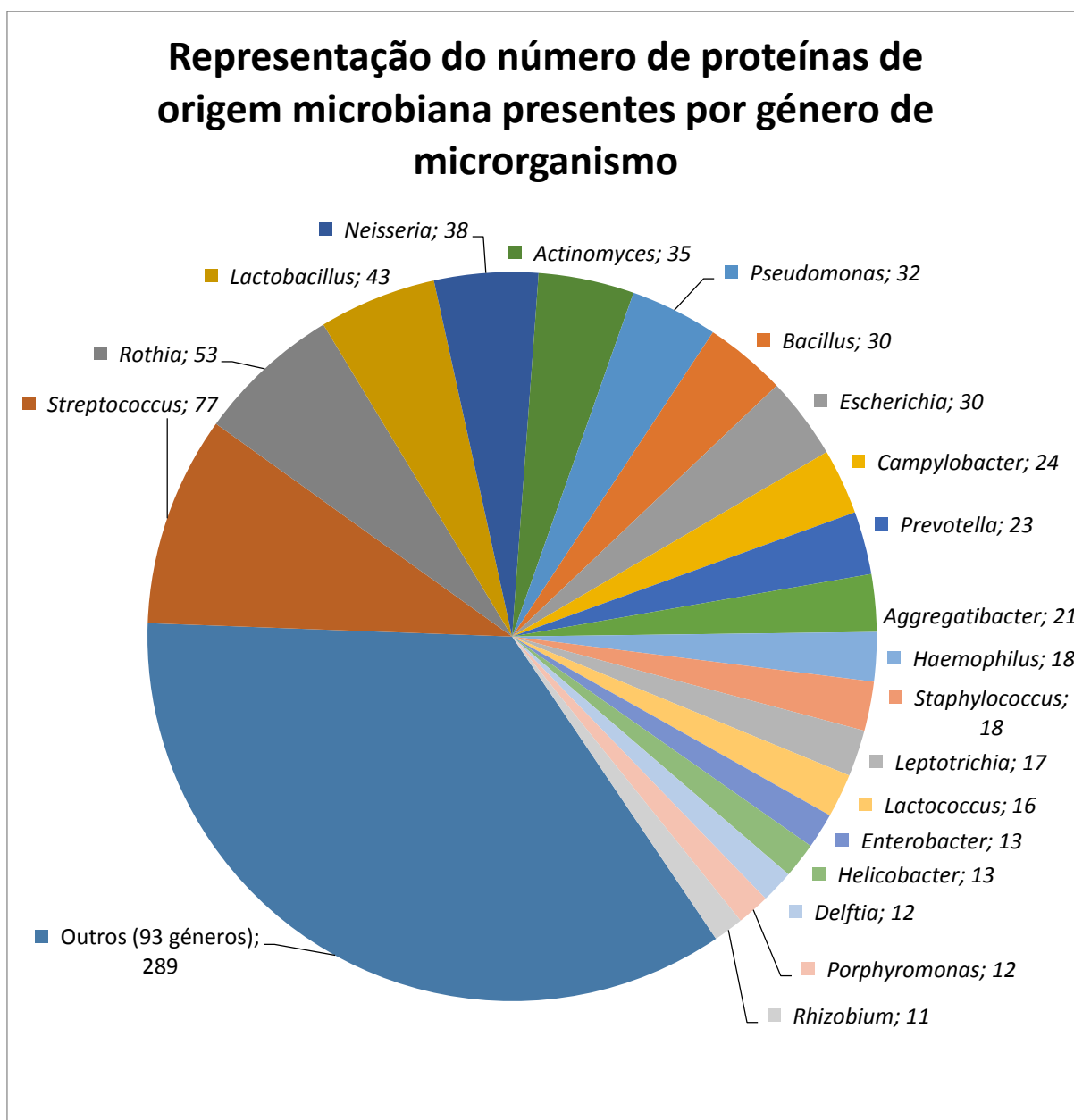
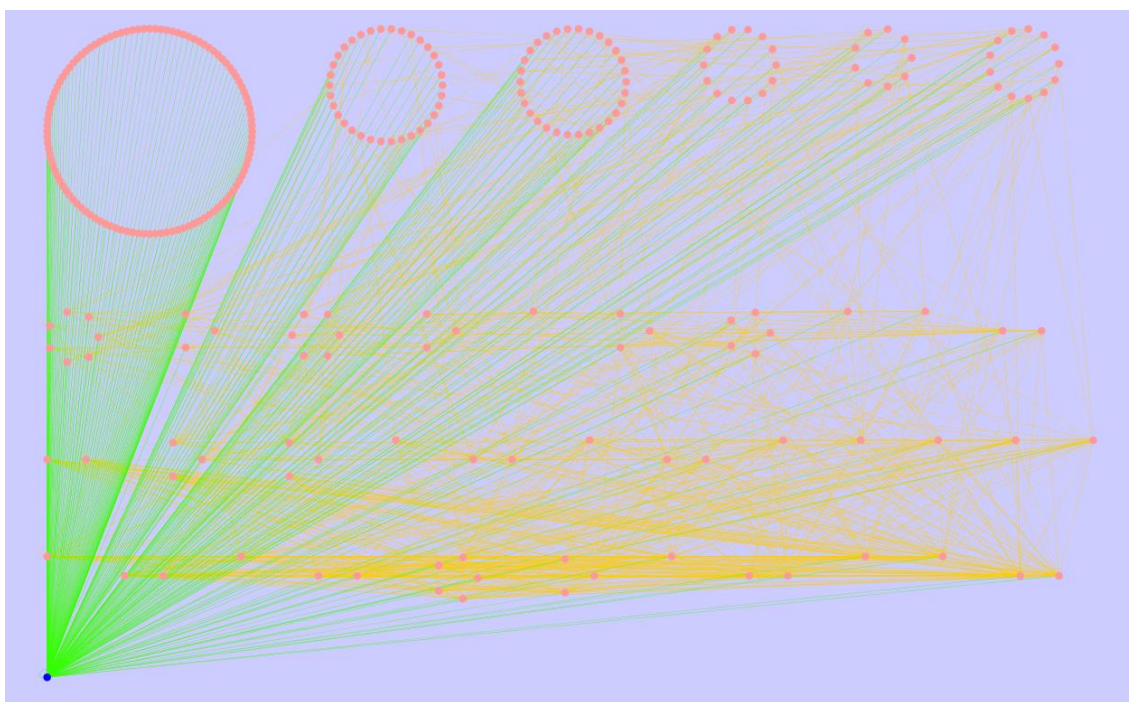


Figura 12 – Género de microrganismos que produzem um maior número de proteínas presentes na cavidade oral, identificadas pelo algoritmo Orallnt.

A análise da figura 12 indica que o género de microrganismos presente na cavidade oral, em que existe uma maior representação em termos de proteínas é *Streptococcus*, seguido de *Rothia* e *Lactobacillus*. A fatia do gráfico

que se encontra legendada com o nome “Outros” representa o conjunto de 93 géneros de microrganismos cuja identificação resulta num número de proteínas compreendidas entre 1 e 10, pelo que se pode concluir que a maioria dos organismos é representada por um número reduzido de proteínas.

#### 4.1.2. Interatoma da cavidade oral – Interações por organismo



**Figura 13 – Rede de Interações estabelecida entre distintos organismos presentes na cavidade oral. Os nós a rosa representam os organismos de origem microbiana, enquanto o nó que se encontra a azul representa o organismo *Homo sapiens*. As interações estabelecidas entre organismos de origem humana e bacteriana encontram-se a verde, enquanto as interações estabelecidas entre organismos de origem microbiana se encontram a amarelo.**

O algoritmo OralInt foi aplicado às proteínas depositadas no OralOme (3) que representam 285 organismos, 1 de origem humana e 284 de origem bacteriana, proporcionando uma rede global de interação única com um total de 1139 interações entre 285 organismos, como pode ser observado na figura 13.

Podemos observar que os organismos dispostos em forma de círculo que se encontram no canto superior esquerdo, realizam um menor número de interações, sendo que os organismos que se situam progressivamente à direita e inferiormente aos mesmos, estabelecem sucessivamente um maior número de

interações, encontrando-se assim nesta ordem de ideias o organismo com o maior número de interações no canto inferior esquerdo.

Se excluirmos as interações realizadas dentro de cada organismo e considerarmos que na figura 13 cada nó representa um organismo distinto, verificamos que o organismo que realiza um maior número de interações é *Homo sapiens*, uma vez que este estabelece 284 interações, interagindo com todos os organismos à exceção da *Serratia proteamaculans* (strain 568). Com 48 interações a *Rothia mucilaginosa* (strain DY-18) (*Stomatococcus mucilaginosus*) e o *Lactococcus lactis subsp. lactis* (strain IL1403) (*Streptococcus lactis*) são os organismos microbianos com o maior número interações estabelecidas. A maioria dos organismos estabelece apenas 1, 2 ou 3 interações, existindo um total de 175 organismos distribuídos apenas nestas três categorias, em que 111 organismos estabelecem apenas uma interação, 34 estabelecem duas interações e 30 organismos realizam três interações.

#### 4.1.3. Interatoma da cavidade oral – Interações entre organismos microbianos

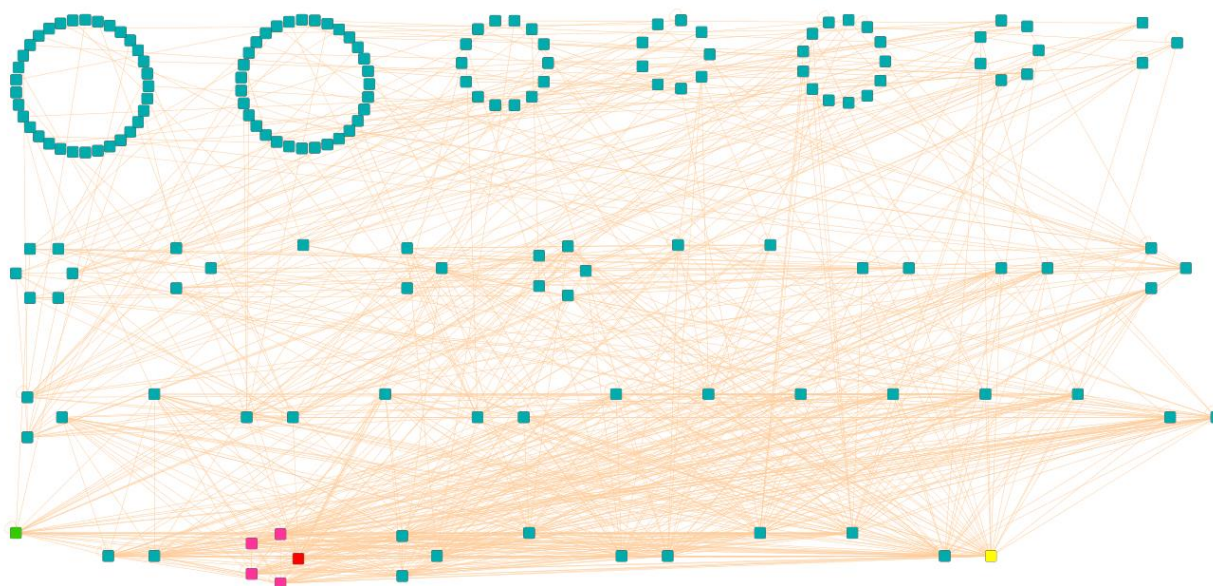


Figura 14 – Rede de Interações estabelecida entre os distintos organismos microbianos presentes na cavidade oral. Os nós a azul representam os organismos de origem microbiana. *Aggregatibacter actinomycetemcomitans* a púrpura, *Porphyromonas gingivalis* a vermelho, *Streptococcus mutans* a verde, e *Rothia mucilaginosa* a amarelo.

Na figura 14, podemos observar que os organismos microbianos dispostos em forma de círculo que se encontram no canto superior esquerdo, realizam um menor número de interações, sendo que os organismos que se situam progressivamente à direita e inferiormente aos mesmos, estabelecem sucessivamente um maior número de interações, encontrando-se assim o organismo com o maior número de interações no canto inferior direito. Obteve-se a representação de 284 organismos de origem microbiana, que estabelecem entres si 855 interações de elevada confiança.

## **4.2. Interatoma Humano – Microbiano da cavidade oral**

Foi feita uma análise mais pormenorizada das interações estabelecidas por quatro microrganismos existentes na cavidade oral que se sabe terem um papel importante na etiologia de patologias orais, *Aggregatibacter*, *Porphyromonas*, *Streptococcus*, e *Rothia*.

### **4.2.1. *Aggregatibacter***

O algoritmo OralInt identifica a participação de 4 proteínas de *Aggregatibacter actinomycetemcomitans* em 188 interações com 10 proteínas humanas e 33 proteínas microbianas. É de salientar que as 4 espécies de *Aggregatibacter actinomycetemcomitans* (*Aggregatibacter actinomycetemcomitans* D17P-3, *Aggregatibacter actinomycetemcomitans* D7S-1, *Aggregatibacter actinomycetemcomitans* serotype a str. H5P1, *Aggregatibacter actinomycetemcomitans* serotype d str. I63B) estabelecem exatamente o mesmo número de interações (47) e com as mesmas proteínas, uma vez que a proteína *TypA protein* é idêntica para as 4 espécies. A *TypA protein* encontra-se ligada a processos de ligação ao GTP, assim como à atividade de GTPase.

O facto de estabelecerem o mesmo número de interações, e de interagirem com as mesmas proteínas deu origem a uma rede de interações interessante apenas do ponto de vista estético devido à sua simetria, como se



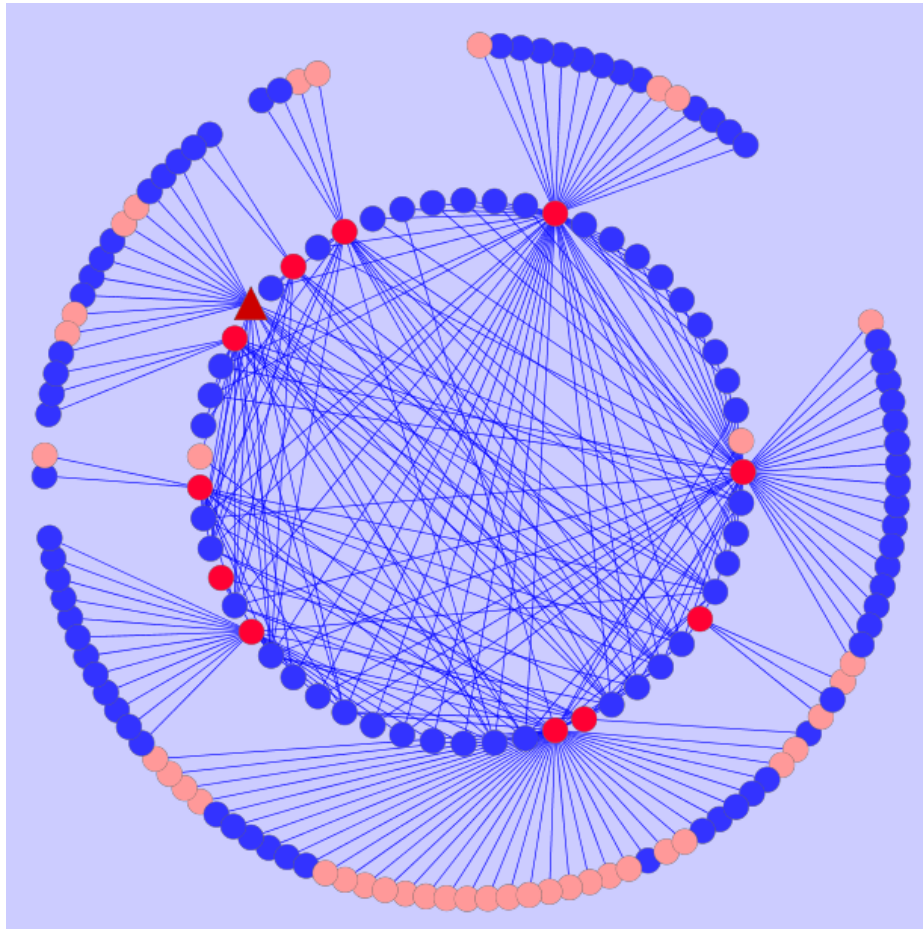


Figura 16 – Rede de interações estabelecidas pelas proteínas bacterianas orais de *Porphyromonas* (vermelho) com as proteínas orais humanas (azul) e com as proteínas orais bacterianas interaguintes com *Porphyromonas* (rosa). Com um símbolo (triângulo) e cor diferentes (vermelho escuro) está a enolase de *Porphyromonas gingivalis* cujas interações específicas mais interessantes são exploradas.

Na figura 16 as proteínas que estabelecem apenas uma interação, encontram-se no círculo exterior, enquanto as proteínas que estabelecem duas ou mais interações como o caso das proteínas pertencentes a *Porphyromonas*, se encontram no círculo interior. A análise da rede de interações permitiu identificar uma proteína de *Porphyromonas gingivalis*, a Q7MTV8, *Enolase* (EC 4.2.1.11) (*2-phospho-D-glycerate hydro-lyase*) (*2-phosphoglycerate dehydratase*), identificada na imagem 16 com forma triangular e a coloração mais escura, cujas ontologias (*cell surface* e *extracellular region*) apresentam maior potencial de interesse para a interação com o hospedeiro. A figura que se segue expõe a rede de interações dessa proteína em detalhe.

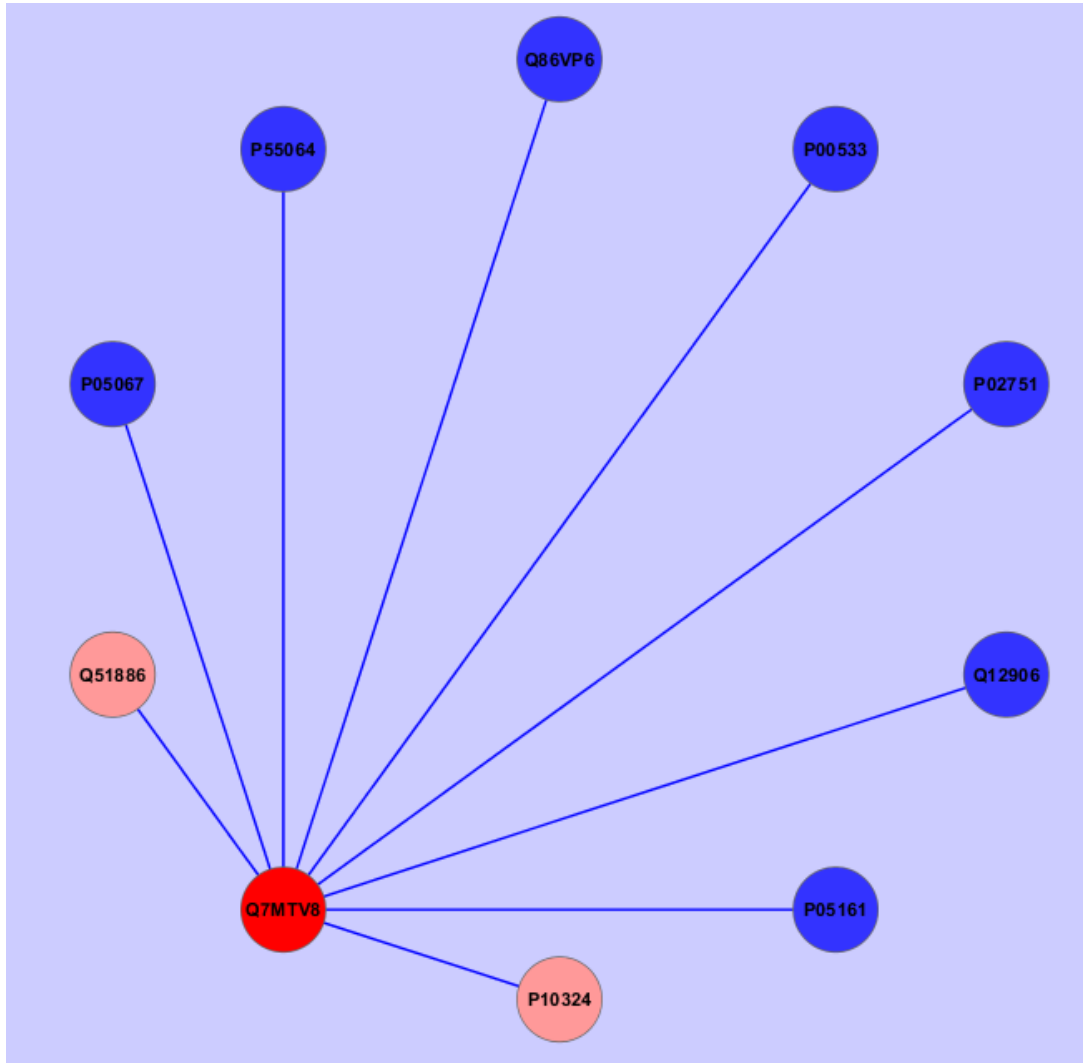


Figura 17 – Rede de interações estabelecida pela proteína Q7MTV8, Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Porphyromonas gingivalis* (vermelho) com proteínas orais de origem humana (azul) e com as proteínas orais bacterianas (rosa)

Através da análise da figura 17, foi possível identificar duas proteínas que interatuam com a *Enolase* (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) já descritas anteriormente em trabalhos publicados (82-85), assim como sete interações novas, propostas pelo algoritmo OralInt que não foram descritas em trabalhos anteriores num total de 9 interações estabelecidas. A anotação das interações com maior potencial de interesse podem ser observadas nas tabelas 1 e 2.

**Tabela 1 – Relevância das Interações estabelecidas pela proteína Q7MTV8, Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Porphyromonas gingivalis*, já documentadas.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Amyloid beta A4 protein	P05067	<i>Homo sapiens</i> (Human)	A enolase prejudica o normal funcionamento da amyloid beta 4, diminuindo a sua expressão (82, 83). No entanto o facto de este recetor membranal estar sobretudo presente nas membranas dos neurónios faz com que a relevância biológica da interação seja nula para a cavidade oral.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	A enolase apresenta-se como uma ligação altamente específica à fibronectina, revelando ainda que a fibronectina é o principal ligando da enolase. Assim poderá haver uma ligação da bactéria à matriz extracelular (84, 85). Promovendo a colonização dos tecidos da cavidade oral.

**Tabela 2 - Relevância das Interações estabelecidas pela proteína Q7MTV8, Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Porphyromonas gingivalis*, propostas pelo algoritmo OralInt.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Aquaporin-5	P55064	<i>Homo sapiens</i> (Human)	Pode interferir em processos como o transporte de dióxido de carbono, excreção, odontogênese, secreção salivar.
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina cinase.
Ubiquitin-like protein ISG15 (Interferon-induced 15 kDa protein)	P05161	<i>Homo sapiens</i> (Human)	Alteração na sinalização célula-célula, resposta da defesa contra bactérias e contra vírus.
Interleukin enhancer-binding factor 3	Q12906	<i>Homo sapiens</i> (Human)	Alteração na resposta à defesa contra vírus.
Cullin-associated NEDD8-dissociated protein 1 (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.
Outer membrane protein P6 (OMP P6)	Q51886	<i>Pasteurella multocida</i> (strain Pm70)	Suspeita de que se podem encontrar em coocorrência.
Outer membrane protein P6 (OMP P6)	P10324	<i>Haemophilus influenzae</i> (strain ATCC 51907 / DSM 11121 / KW20 / Rd)	Suspeita de que se podem encontrar em coocorrência.

### 4.2.3. *Rothia*

Foram analisadas 53 proteínas pertencentes ao género *Rothia*, 10 pertencentes a *Rothia mucilaginosa* atcc 25296 e 43 pertencentes a *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*), que estabelecem um total de 687 interações, com 158 proteínas humanas e 111 proteínas microbianas. A espécie *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*) é aquela que estabelecem o maior número de interações (557), seguindo-se pela *Rothia mucilaginosa* atcc 25296 (112) como se pode observar na figura 18.

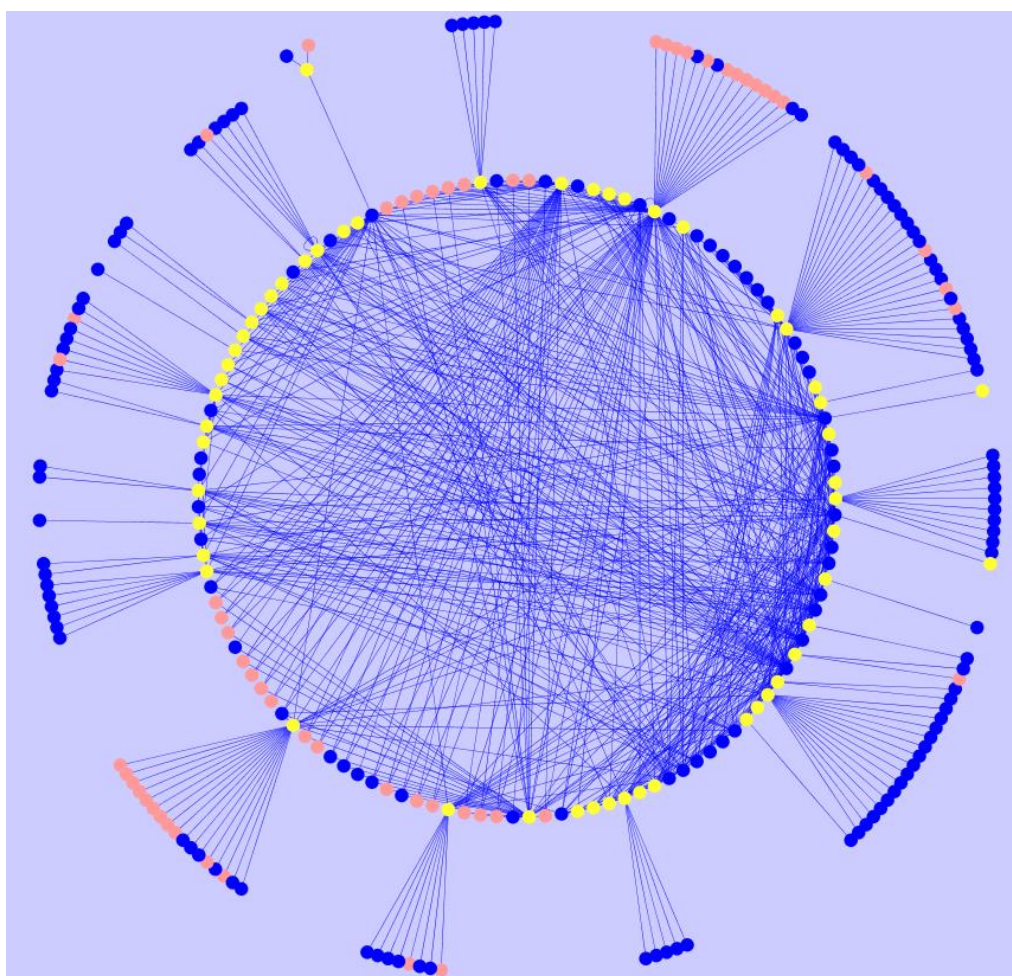


Figura 18 – Rede de Interações estabelecida pelas proteínas bacterianas orais do género *Rothia* (amarelo) com as proteínas orais h.umanas (azul) e com as proteínas orais bacterianas interaguintes com *Rothia* (rosa).

A figura 18 representa o número de interações estabelecidas pelas diversas proteínas. As proteínas que estabelecem apenas uma interação, encontram-se no círculo exterior, enquanto as proteínas que estabelecem duas ou mais interações encontram-se no círculo interior. Como podemos observar, todas as proteínas de *Rothia* encontram-se no círculo interior, com a exceção da *Fructose-1,6-bisphosphatase*, *Endopolygalacturonase* e da *30S ribosomal protein S17*.

#### 4.2.3.1. *Rothia mucilaginosa* atcc 25296

Através da análise da rede de interações da figura 8, foi possível identificar duas proteínas de *Rothia mucilaginosa* atcc 25296, a C6RE2 *Extracellular basic protease* (EC 3.4.21.-) e a C6R624 *Extracellular solute-binding protein, family 5* que devido à sua ontologia (*Extracellular region*) apresentam maior potencial de interesse para a interação com o hospedeiro. A figura que se segue expõe essas mesmas interações ao detalhe.

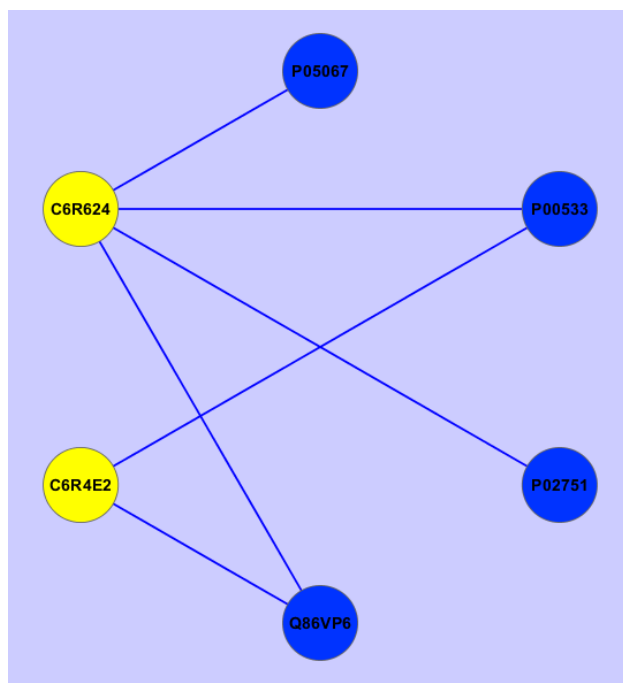


Figura 19 – Rede de interações estabelecida pelas proteínas C6RE2 *Extracellular basic protease* (EC 3.4.21.-) e a C6R624 *Extracellular solute-binding protein, family 5* de *Rothia mucilaginosa* atcc 25296 (amarelo) com proteínas orais de origem humana (azul).

Através da análise da figura 19 foi possível identificar seis interações propostas pelo algoritmo Orallnt, estabelecidas pelas duas proteínas de *Rothia mucilaginosa* atcc 25296 que não foram descritas em trabalhos anteriores. A anotação das interações com maior potencial de interesse podem ser observadas nas tabelas 3 e 4.

**Tabela 3 – Relevância da Interação estabelecida pela proteína C6R4E2 *Extracellular basic protease* (EC 3.4.21.-) de *Rothia mucilaginosa* atcc 25296, proposta pelo algoritmo Orallnt.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.
Cullin-associated NEDD8-dissociated protein 1 (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.

**Tabela 4 - Relevância das Interações estabelecidas pela proteína C6R624 Extracellular solute-binding protein, family 5 de *Rothia mucilaginosa* atcc 25296 proposta pelo algoritmo Orallnt.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Amyloid beta A4 protein	P05067	<i>Homo sapiens</i> (Human)	Pode levar a alterações na resposta imune e inata, assim como no processo de indução da apoptose.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i>	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.
Cullin-associated NEDD8-dissociated protein 1 (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.

#### 4.2.3.2. *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*)

O algoritmo Orallnt foi aplicado ao estudo 4 proteínas de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*), a D2NNZ2 *Uncharacterized protein conserved in archaea*, D2NPQ4 *ABC-type dipeptide transport system, periplasmic component*, a D2NPZ7 *Subtilisin-like serine protease* e a D2NRY1 *Molybdopterin-binding protein*, que devido às suas ontologias (*Cell Wall*) apresentam maior potencial de interesse para a interação com o hospedeiro. A figura que se segue expõe essas mesmas interações ao detalhe.

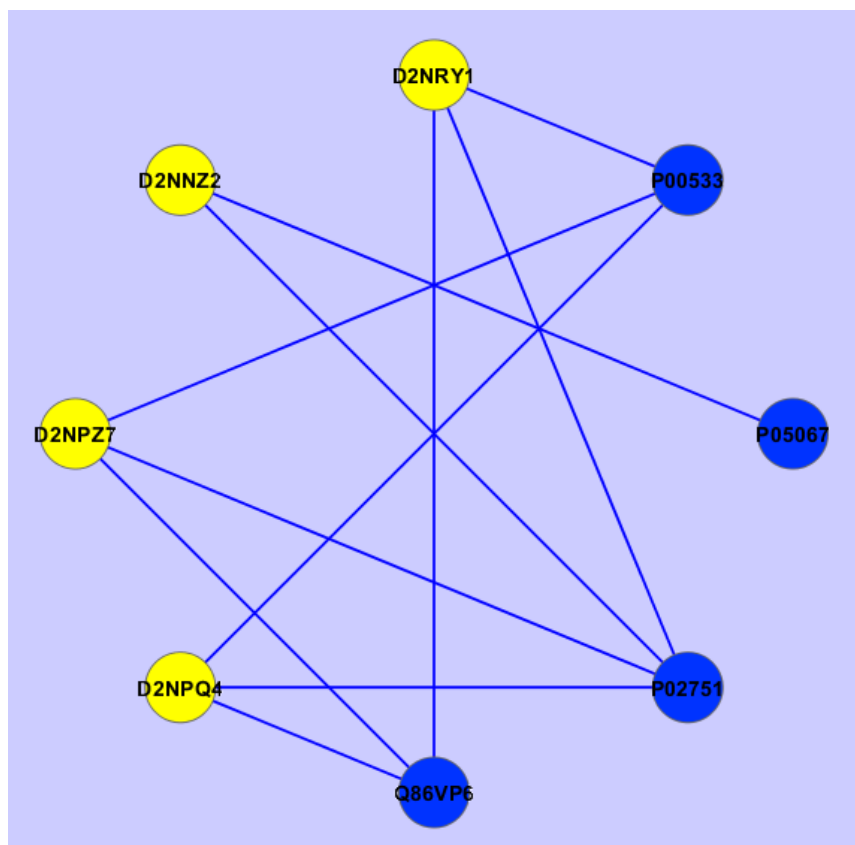


Figura 20 - Rede de interações estabelecida pelas proteínas D2NNZ2 *Uncharacterized protein conserved in archaea*, D2NPQ4 *ABC-type dipeptide transport system, periplasmic component*, D2NPZ7 *Subtilisin-like serine protease*, D2NRY1 *Molybdopterin-binding protein* e a D2NRY1 *Molybdopterin-binding protein* de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*) a (amarelo) com proteínas orais de origem humana (azul).

Através da análise da figura 20 foi possível identificar onze interações estabelecidas pelas proteínas de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*) sendo que duas dessas interações já foram descritas anteriormente em trabalhos publicados, enquanto as restantes seis são interações novas, propostas pelo algoritmo Orallnt e que não foram descritas em trabalhos anteriores. A anotação das interações com maior potencial de interesse podem ser observadas nas tabelas 5, 6, 7 e 8.

**Tabela 5 – Relevância das Interações estabelecidas pela proteína D2NNZ2 *Uncharacterized protein conserved in archaea*, de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*).**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Amyloid beta A4 protein	P05067	<i>Homo sapiens</i> (Human)	Pode levar a alterações na resposta imune e inata, assim como no processo de indução da apoptose.
Fibronectin	P02751	<i>Homo sapiens</i> (human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.

**Tabela 6 – Relevância das Interações estabelecidas pelas proteínas D2NPQ4 ABC-type dipeptide transport system, periplasmic componente e D2NRY1 Molybdopterin-binding protein de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*).**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.
Cullin-associated NEDD8-dissociated protein 1 (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.

**Tabela 7 – Relevância das Interações estabelecidas pela proteína D2NPZ7 *Subtilisin-like serine protease* de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*).**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	O inibidor de protéase serina Kazal tipo 1 (SPINK1) é expresso em células acinares pancreáticas normais humanas, assim como em diversos tumores. Em linhas celulares de cancro de pâncreas esta protéase medeia a proliferação celular através da cascata da proteína cinase ativadora mitogénica, graças à sua ligação com o recetor do fator de crescimento epidermal (EGFR) (86, 87).
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	HTRA1 (protéase serina secretada altamente conservada) induz proteólise da fibronectina, resultando em fragmentos de diversos tamanhos (88).
Cullin (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.

**Tabela 8 - Relevância das Interações estabelecidas pela proteína D2NRY1, *Molybdopterin-binding protein* de *Rothia mucilaginosa* (strain dy-18) (*stomatococcus mucilaginosus*).**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.
Cullin-associated NEDD8-dissociated protein 1 (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.

#### 4.2.4. *Streptococcus*

Das 79 proteínas pertencentes ao género *Streptococcus*, 27 eram de *Streptococcus pneumoniae*, 15 de *Streptococcus salivarius*, 10 de *Streptococcus gordonii*, 9 de *Streptococcus pyogenes*, 8 de *Streptococcus agalactiae*, 7 de *Streptococcus mutans*, 2 de *Streptococcus lactis* e 1 de *Streptococcus dysgalactiae*, que estabelecem um total de 1309 interações, com 277 proteínas humanas, e 78 proteínas microbianas. A espécie *Streptococcus pneumoniae* é aquela que estabelece um maior número de interações, com um total de 468 interações, obtendo-se assim um resultado espectável uma vez que esta é a espécie com maior representação em termos de proteínas.

Na figura 21 podemos observar a rede de interações criada para o organismo *Streptococcus* com proteínas humanas, e na figura 22 as interações com as proteínas bacterianas.

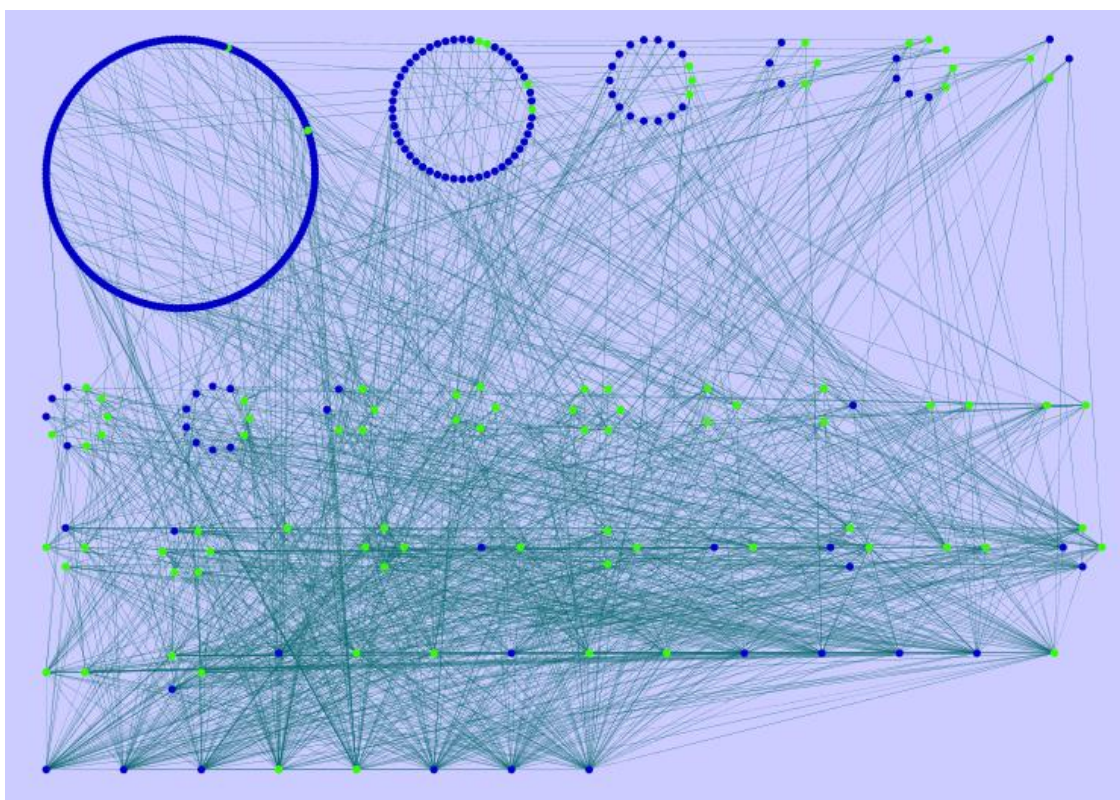
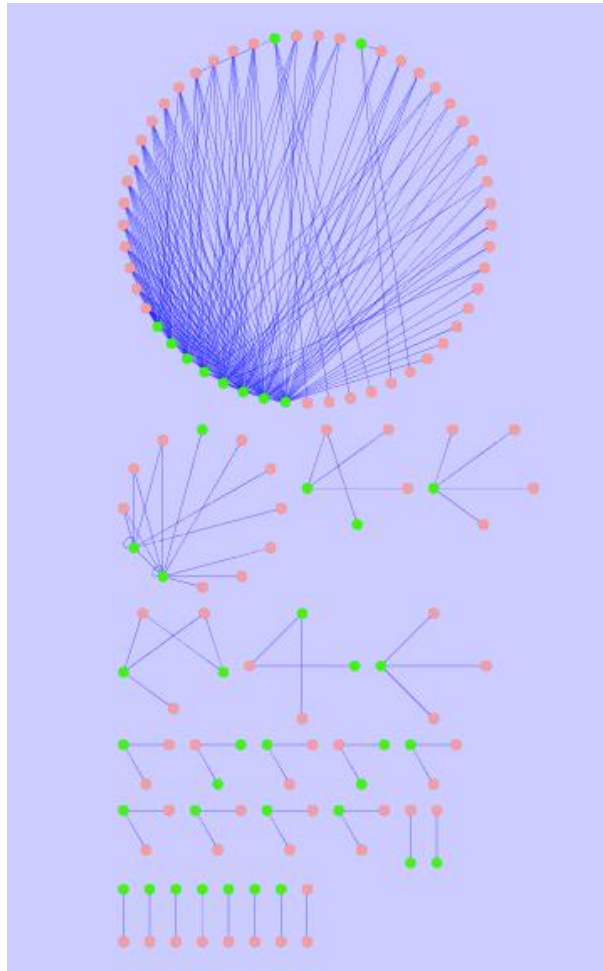


Figura 21 - Rede de Interações estabelecida pelas proteínas bacterianas orais do género *Streptococcus* (verde) com as proteínas orais humanas (azul).

Através da análise da figura 21, podemos observar que as proteínas microbianas e humanas dispostas em forma de círculo que se encontram no canto superior esquerdo realizam um menor número de interações, sendo que as proteínas que se situam progressivamente à direita e inferiormente às

mesmas, estabelecem sucessivamente um maior número de interações, encontrando-se assim nesta ordem de ideias o organismo com o maior número de interações no canto inferior direito, que neste caso é uma proteína humana.



**Figura 22 - Rede de Interações estabelecida pelas proteínas bacterianas orais do género *Streptococcus* (verde) com as proteínas orais bacterianas (rosa).**

A disposição da figura 22 está relacionada com a quantidade de interações que cada proteína estabelece, assim sendo, as proteínas que se encontram no círculo do topo da imagem, apresentam um maior número de interações que as que se encontram na base da mesma.

Identificaram-se 5 proteínas de *Streptococcus pneumoniae*, 1 proteína de *Streptococcus salivarius*, e 1 proteína de *Streptococcus mutans* que devido às suas ontologias (*Cell surface*, *Extracellular region*) apresentam maior potencial de interesse para a interação com o hospedeiro.

#### 4.2.4.1. *Streptococcus pneumoniae*

As 5 proteínas de *Streptococcus pneumoniae* que apresentam um maior potencial de interesse de interação com o hospedeiro são a Q97QS2 *Enolase* (EC 4.2.1.11) (*2-phospho-D-glycerate hydro-lyase*) (*2-phosphoglycerate dehidratase*), A5MW80 *G5 domain family protein*, P59213 *Maltose/maltodextrin-binding protein*, Q97NH7 *Competence protein CglC* e a Q8DNE1 *Membrane protein insertase YidC 1* (*Foldase YidC 1*) (*Membrane integrase YidC 1*) (*Membrane protein YidC 1*).

As interações estabelecidas por estas 5 proteínas podem ser observadas na figura 23.

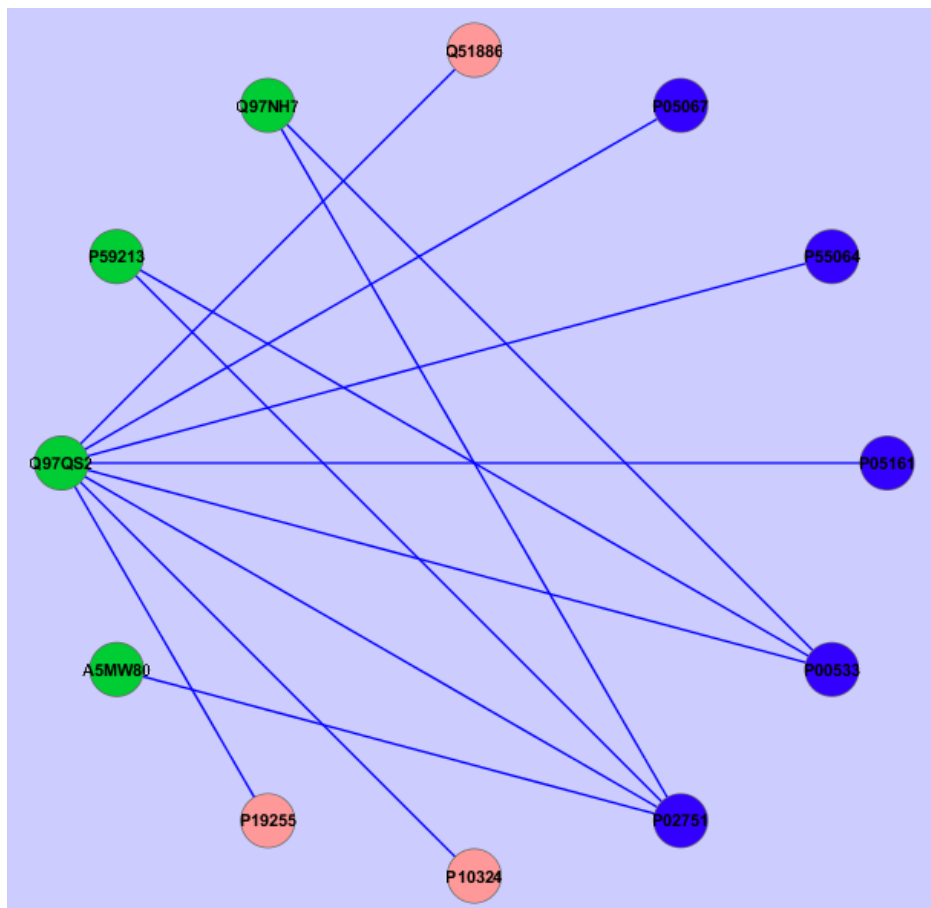


Figura 23 – Rede de interações estabelecida pelas proteínas de *Streptococcus pneumoniae* (verde) com proteínas orais de origem humana (azul) e com as proteínas orais bacterianas (rosa)

Através da análise da figura 23 foi possível identificar dezasseis interações estabelecidas pelas proteínas de *Streptococcus pneumoniae* sendo que duas dessas interações já foram descritas anteriormente em trabalhos publicados, enquanto as restantes doze são interações novas, propostas pelo

algoritmo Orallnt e que não foram descritas em trabalhos anteriores. A anotação das interações com maior potencial de interesse podem ser observadas nas tabelas 9, 10, 11 e 12.

**Tabela 9 - Relevância das Interações estabelecidas pela proteína A5MW80 *G5 domain family protein* de *Streptococcus pneumoniae* SP23-BS72.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.

**Tabela 10 - Relevância das Interações estabelecidas pela proteína Q97QS2 Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Streptococcus pneumoniae* SP23-BS72.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Amyloid beta A4 protein	P05067	<i>Homo Sapiens</i> (Human)	A enolase prejudica o normal funcionamento da amyloid beta 4, diminuindo a sua expressão (82, 83)
Aquaporin-5 (AQP-5)	P55064	<i>Homo sapiens</i> (Human)	Pode interferir em processos como o transporte de dióxido de carbono, excreção, odontogênese, secreção salivar
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	A enolase apresenta-se como uma ligação altamente específica da fibronectina, revelando ainda que a fibronectina é o principal ligando da enolase, podendo haver uma ligação da bactéria à matriz extracelular (84, 85).
Probable glycerol uptake facilitator protein	P19255	<i>Homo sapiens</i> (Human)	Alteração no processo metabólico do glicerol.
Ubiquitin-like protein ISG15 (Interferon-induced 15 kDa protein)	P05161	<i>Homo sapiens</i> (Human)	Alteração na resposta da defesa a bactérias e a vírus.
Outer membrane protein P6 (OMP P6)	P10324	<i>Haemophilus influenzae</i>	Suspeita de que se podem encontrar em coocorrência.
Outer membrane protein P6 (OMP P6)	Q51886	<i>Pasteurella multocida</i> (strain Pm70)	Suspeita de que se podem encontrar em coocorrência.

**Tabela 11 - Relevância das Interações estabelecidas pela proteína P59213 *Maltose/maltodextrin-binding protein*, de *Streptococcus pneumoniae* serotype 4 (strain ATCC BAA-334 / TIGR4).**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.

**Tabela 12 - Relevância das Interações estabelecidas pelas proteínas Q97NH7 *Competence protein CglC* de *Streptococcus pneumoniae* serotype 4 (strain ATCC BAA-334 / TIGR4) e Q8DNE1 *Membrane protein insertase YidC 1 (Foldase YidC 1) (Membrane integrase YidC 1) (Membrane protein YidC 1)* de *Streptococcus pneumoniae* (strain ATCC BAA-255 / R6)**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.
Epidermal growth factor receptor (EC 2.7.10.1)	P00533	<i>Homo sapiens</i> (Human)	Pode prejudicar ou alterar a atividade do recetor de ativação do fator de crescimento epidermal, a via de sinalização do fator de crescimento fibroblástico, a resposta imune e inata, a ativação da atividade da fosfolipase A2 na sinalização mediada pelo cálcio, a proliferação celular, a adesão célula – célula, a regulação negativa do processo apoptótico, a regulação negativa da via de sinalização do fator de crescimento epidermal, a ossificação, a regulação positiva da proliferação fibroblástica, a atividade do recetor de sinalização da proteína tirosina quinase.

#### 4.4.4.2. *Streptococcus salivarius*

Identificou-se uma proteína de *Streptococcus salivarius*, a C2LUT1 Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) em que as suas ontologias (cell surface e extracellular region) apresentam maior potencial de interesse para a interação com o hospedeiro. A figura que se segue expõe essas mesmas interações em detalhe.

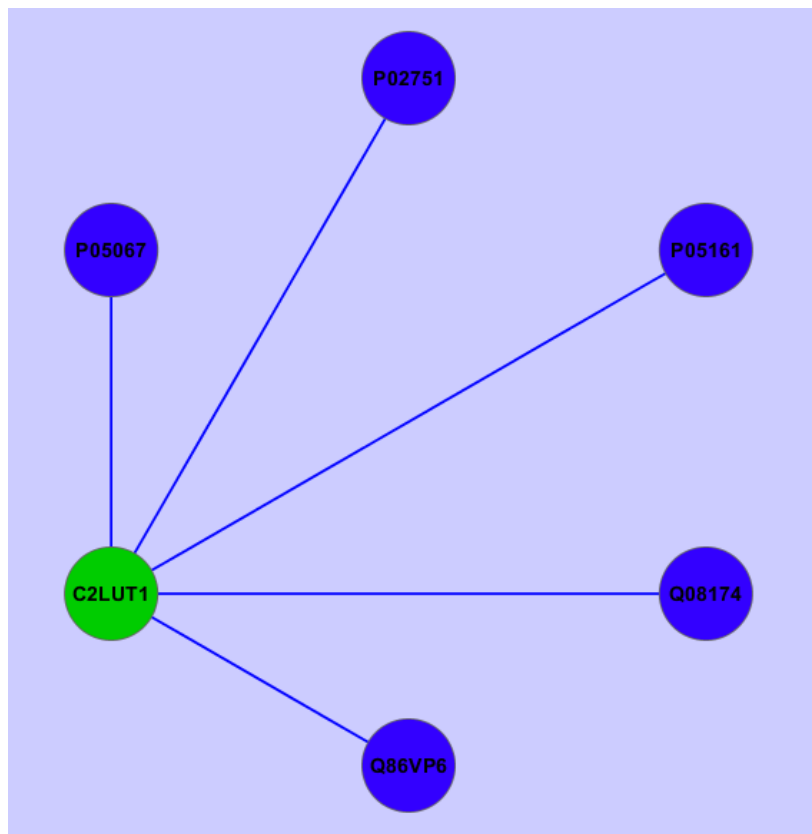


Figura 24 - Rede de interações estabelecida pela proteína C2LUT1 Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Streptococcus salivarius* SK126 (verde), com as proteínas orais de origem humana (azul).

Analisando a figura 24 foi possível identificar cinco interações estabelecidas pela Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Streptococcus salivarius* SK126 sendo que duas dessas interações já foram descritas anteriormente em trabalhos publicados, enquanto as três interações restantes são novas, propostas pelo algoritmo OralInt e que não foram descritas em trabalhos anteriores. A anotação

das interações com maior potencial de interesse podem ser observadas nas tabelas 13 e 14.

**Tabela 13 - Relevância das Interações estabelecidas pela proteína C2LUT1, Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Streptococcus salivarius* SK126, já documentadas.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Amyloid beta A4 protein	P05067	<i>Homo sapiens</i> (Human)	A enolase prejudica o normal funcionamento da amyloid beta 4, diminuindo a sua expressão (82, 83).
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	A enolase apresenta-se como uma ligação altamente específica da fibronectina, revelando ainda que a fibronectina é o principal ligando da enolase, podendo haver uma ligação da bactéria à matriz extracelular (84, 85).

**Tabela 14 - Relevância das Interações estabelecidas pela proteína C2LUT1, Enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (2-phosphoglycerate dehydratase) de *Streptococcus salivarius* SK126, propostas pelo algoritmo Orallnt.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Ubiquitin-like protein ISG15 (Interferon-induced 15 kDa protein)	P05161	<i>Homo sapiens</i> (Human)	Alteração na sinalização célula-célula, resposta da defesa contra bactérias e contra vírus.
Protocadherin-1 (Cadherin-like protein 1) (Protocadherin-42) (PC42)	Q08174	<i>Homo sapiens</i> (Human)	Alteração na junção célula-célula e na sinalização célula - célula.
Cullin-associated NEDD8-dissociated protein 1 (p120 CAND1)	Q86VP6	<i>Homo sapiens</i> (Human)	Alterações na diferenciação celular, assim como, no complexo ubiquitin ligase.

#### 4.4.4.3. *Streptococcus mutans*

Identificou-se uma proteína de *Streptococcus mutans*, a Q8DSC2 *Putative PTS system, mannose-specific component IID (EC 2.7.1.69)* em que a sua ontologia (*Integral to membrane*) e função associada ao transporte de açúcares apresenta maior potencial de interesse para a interação com o hospedeiro. A figura 16 expõe essas mesmas interações em detalhe.

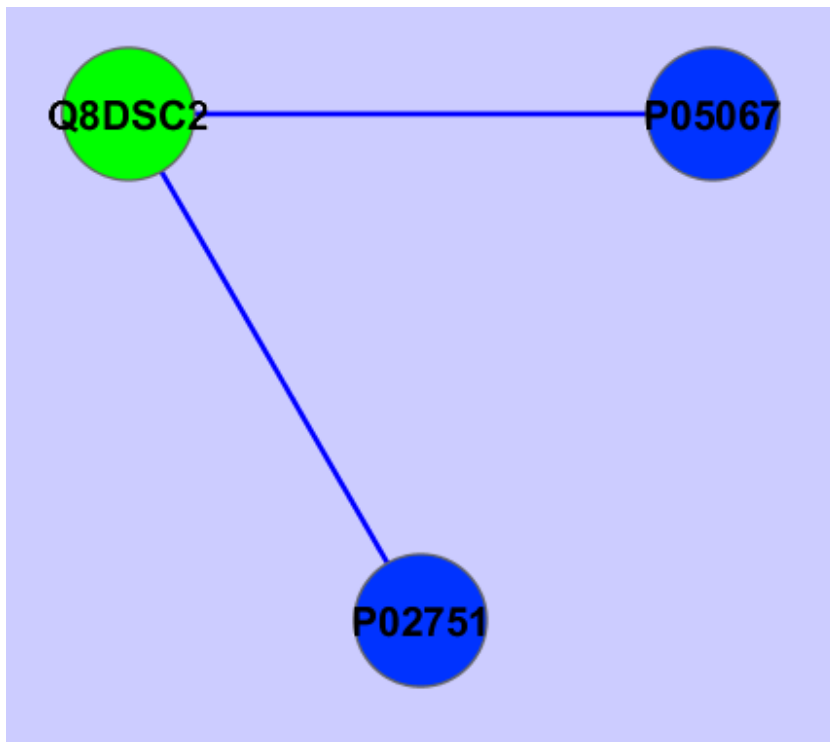


Figura 25 - Rede de interações estabelecida pela proteína Q8DSC2 *Putative PTS system, mannose-specific component IID (EC 2.7.1.69)* de *Streptococcus mutans* serotype c (strain ATCC 700610 / UA159) (verde) com as proteínas orais de origem humana (azul).

Através da análise da figura 25 identificou-se duas interações estabelecidas pela proteína *Putative PTS system, mannose-specific component IID (EC 2.7.1.69)* de *Streptococcus mutans* serotype c (strain ATCC 700610 / UA159) propostas pelo algoritmo OralInt e que não foram descritas em trabalhos anteriores. A anotação das interações com maior potencial de interesse podem ser observadas nas tabelas 15.

**Tabela 15 - Relevância das Interações estabelecidas pela proteína Q8DSC2 Putative PTS system, mannose-specific component IID (EC 2.7.1.69) de *Streptococcus mutans* serotype c (strain ATCC 700610 / UA159) propostas pelo algoritmo Orallnt.**

Proteína com a qual interatua	Código UniProtKB	Organismo	Relevância da interação
Amyloid beta A4 protein	P05067	<i>Homo sapiens</i> (Human)	Pode levar a alterações na resposta imune e inata, assim como no processo de indução da apoptose.
Fibronectin	P02751	<i>Homo sapiens</i> (Human)	Pode alterar a organização e estrutura da matriz extracelular, migração leucocitária, desgranulação plaquetária assim como provocar alteração na regulação da forma da célula.



## 5. DISCUSSÃO

A comparação dos dados de proteômica da cavidade oral usados pelo algoritmo Orallnt no que se refere aos microrganismos, com os resultados da metagenômica do microbioma humano da cavidade oral em saúde, permite concluir que a cobertura em termos de proteínas microbianas identificadas é muito baixa. Quando analisados em detalhe os dados usados nesta análise do ORallnt verifica-se que a sobreposição de resultados entre os vários estudos de proteômica não existe, o que pode ser devido à aplicação de técnicas diferentes de processamento de amostras e de identificação dos fragmentos obtidos na análise de espectrometria de massa (89). 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 (89). Assim, uma das melhorias a implementar em determinações futuras das PPIs presentes na cavidade oral será aumentar a representatividade dos microrganismos em termos de proteínas.

Por outro lado, como há falta de anotação que a maioria das proteínas microbianas nas bases de dados, e porque o Orallnt baseia em parte a previsão das interações em interações já descritas na literatura e na anotação pré-existente das proteínas, menor informação disponível sobre as proteínas implica uma menor capacidade de serem identificadas novas interações. Esta questão tem tendência a ser resolvida pelo aumento da informação existente sobre cada proteína microbiana e portanto o aumento da anotação das proteínas microbianas.

Segundo *The Human Microbiome Consortium*, o meio intra oral é dominado sobretudo por *Streptococcus*, seguido em abundância por *Haemophilus* na mucosa oral, *Actinomyces* na placa supragengival, e *Prevotella* na placa subgengival (7). À exceção do género *Streptococcus*, que é o

dominante em ambos os estudos, os nossos resultados diferem dos resultados obtidos pelo *The Human Microbiome Consortium*, uma vez que nem *Haemophilus* com uma representação de 18 proteínas, *Actinomyces* com 35 proteínas e *Prevotella* com 23 proteínas, se encontram nos géneros mais representados. Esta diferença obtida pode ter como explicação o facto de o *The Human Microbiome Consortium* realizar a identificação das bactérias através do material genético, enquanto o nosso trabalho é focado nas proteínas, o que implica diferenças a nível técnico na identificação e justifica a obtenção de resultados diferentes.

Os microrganismos que se sabe terem um papel importante na etiologia da cárie (*Streptococcus mutans*), e na doença periodontal (*Rothia mucilaginosa*, *Porphyromonas gingivalis* e *Aggregatibacter actinomycetemcomitans*), apresentam um número de interações bastante elevado, sendo mesmo a *Rothia mucilaginosa* o microrganismo que possui um maior número de interações de todos os microrganismos presentes com 47 interações. Os restantes microrganismos associados a patologias orais possuem respetivamente 33 interações no caso de *Porphyromonas gingivalis* e *Aggregatibacter actinomycetemcomitans*, encontrando-se em último lugar *Streptococcus mutans* com 31 interações. É interessante verificar também que os 3 organismos periodontopatogénicos interagem entre si, como é o caso da interação entre: *Rothia mucilaginosa* e *Aggregatibacter actinomycetemcomitans*, *Rothia mucilaginosa* e *Porphyromonas gingivalis*, *Porphyromonas gingivalis* e *Aggregatibacter actinomycetemcomitans*. Identificou-se também possíveis interações entre *Porphyromonas gingivalis* e outros periodontopatogénios como a *Kingella oralis* e *Treponema denticola*, o que demonstra a sinergia existente entre os diferentes periodontopatogénios já anteriormente sugerida (90).

Após a análise da anotação das 4 proteínas de *Aggregatibacter actinomycetemcomitans* e apesar de se saber que este género possui um papel relevante na etiologia da doença periodontal (91), conclui-se que a *TypA protein* não apresenta ontologias com potencial de interesse para a interação com o hospedeiro. Uma vez que não possui anotação ontológica para *Cellular Component* (GO:0016043) o que não permite integrá-la nas categorias *Extracellular Region* (GO:0005576), *Plasma Membrane* (GO:0005886) e *Cell Wall* (GO:0005618). Estas seriam as anotações das proteínas microbianas

localizadas no exterior da célula e portanto as primeiras estruturas a entrar em contacto com as superfícies do hospedeiro. Da sua função nomeadamente em *E.coli* (71) em que interage essencialmente com proteínas ribossomais e de tRNA pode inferir-se que a proteína TypA protein tem uma expressão apenas intracelular e muito provavelmente nunca entra em interação com proteínas de outros organismos. Neste caso as previsões do Orallnt correspondem a interações teóricas mas sem grande relevância biológica.

Nos organismos que possuem um maior número de interações (situados na linha inferior da figura 14) podemos encontrar o *Lactococcus lactis* com 47 interações, um Gram positivo usado extensivamente na produção de soro de leite e queijo, que recentemente se tornou conhecido como o primeiro organismo geneticamente modificado a ser utilizado vivo no tratamento de doenças humanas (92). *Haemophilus influenzae* com 46 interações, um Gram negativo descrito pela primeira vez em 1892 por Richard Pfeiffe durante a gripe pandémica (93), sendo que a maioria das estirpes de *Haemophilus influenzae* têm a característica de serem patógenos oportunistas, ou seja, podem causar problemas apenas na presença de outros fatores (infecção viral, redução da função imune ou alergias) (94). *Streptomyces coelicolor* com 44 interações é um Gram positivo, tendo sido apelidado primeiramente como *Streptothrix coelicolor* em 1908 por R.Muller depois de este o ter encontrado numa batata (95), não causa doenças em humanos, animais ou plantas (96). É notável pela sua produção de um composto único farmacêutico, que inclui agentes anti-tumorais, imunossupressores, dando ainda origem atualmente a um grande número de antibióticos utilizados na medicina humana e veterinária como é o caso do actinorhodina (97), metilenomicina (98) e perimicina (99) assim como a agentes antiparasitários, herbicidas e várias enzimas importantes na indústria da alimentação (100) e mais de dois terços de todos os antibióticos naturais disponíveis atualmente (91). O facto destas bactérias terem sido extensivamente estudadas (ainda que em contextos alheios à cavidade oral) faz com que haja bastante anotação das suas proteínas o que justifica que o Orallnt tenha conseguido prever um número elevado de interações.

Se considerarmos as bactérias mais associadas à doença periodontal, podemos verificar que o algoritmo Orallnt prevê poucas interações para *Aggregatibacter actinomycetemcomitans* e *Porphyromonas gingivalis* com as

proteínas do hospedeiro, pois estes dois organismos, não se encontram bem representados no conjunto de dados de proteínas inicial, o que não deveria ocorrer, uma vez que estes dois organismos pertencem à flora oral comensal (10). Contudo, no caso de *Porphyromonas gingivalis* o algoritmo identificou várias interações interessantes e algumas delas não descritas anteriormente. A interação prevista entre a enolase bacteriana (Q7MTV8) e a aquaporina e um receptor EGF têm a potencialidade de afetar alguns dos mecanismos de homeostase do hospedeiro (101). Por outro lado é também prevista a interação da enolase de *Porphyromonas gingivalis* com duas proteínas do hospedeiro (Ubiquitin-like protein ISG15 e Interleukin enhancer-binding factor 3) envolvidas na defesa contra vírus o que pode fazer com que a resposta imune do hospedeiro possa ficar afetada pela presença da bactéria.

Se considerarmos as interações de *Porphyromonas gingivalis* com outras bactérias, verificamos que a mesma enolase pode interagir com proteínas extra membranares de *Pasteurella multocida* e *Haemophilus influenzae* o que significa que estas espécies podem co-colonizar superfícies do hospedeiro.

A *Rothia mucilaginosa*, um organismo presente na flora microbiana humana oral normal, mas que é considerada um patogénico oportunista (7) apresenta-se no nosso estudo como a espécie associada à doença periodontal que possui mais interações, sendo que algumas delas revelam ser interações importantes e específicas.

Os *Streptococcus* são um género de bactérias Gram positivas em forma de coco, pertencentes ao filo *Firmicutes* e ao grupo de bactérias do ácido lácteo (102). O *Streptococcus pneumoniae* reside de forma assintomática na nasofaringe de portadores sãos. No entanto, em indivíduos suscetíveis como idosos, crianças e imunocomprometidos, a bactéria pode tornar-se patogénica e espalhar-se para outros locais onde irá levar a doença (18). O *Streptococcus pneumoniae* é a principal causa de pneumonia e meningite nas crianças e nos idosos, assim como de septicémia nas pessoas portadoras do Vírus da imunodeficiência humana (VIH) (103). Este organismo pode causar outro tipo de infeções pneumocócicas que incluem a bronquite, sinusite aguda, otite média, conjuntivite, meningite, bacterémia, sépsis, osteomielite, artrite séptica, endocardite, peritonite, pericardite, celulite e abscesso cerebral (103). Se verificarmos as interações da enolase Q97QS2 de *Streptococcus pneumoniae*

podemos verificar que esta pode interagir com proteínas extra membranares de *Pasteurella multocida* e *Haemophilus influenzae* o que significa que a co-colonização das superfícies do hospedeiro por estas espécies é possível.

O *Streptococcus salivarius* é uma das principais bactérias comensais da cavidade oral em humanos saudáveis, podendo ser encontrado no trato respiratório superior, sendo a primeira bactéria a colonizar a placa dentária, antes de ser unida por várias espécies de diferentes géneros. Ao ser o primeiro colonizador da placa dentária, pensa-se que possa criar condições favoráveis para que outras espécies possam iniciar a sua colonização. Esta bactéria pode também ser encontrada em infeções oportunistas, mais comumente após intervenções cirúrgicas (104).

O *Streptococcus mutans* é um dos principais peptógenos da cárie dentária uma vez que é capaz de fermentar uma grande variedade de açúcares, mantendo um metabolismo ativo em pHs baixos, que causam a desmineralização progressiva dos dentes. Este processo é intensificado com o aumento da proporção desta espécie oportunista na placa dentária, a qual utiliza a sacarose da dieta para sintetizar polissacarídeos extracelulares insolúveis, que são o principal componente estrutural da placa dentária (105).

Após uma análise pormenorizada das proteínas presentes nos quatro organismos apresentados anteriormente, podemos verificar que a enolase é a proteína dentro do grupo de proteínas microbianas que apresentam ontologias com potencial de interesse na interação com o hospedeiro mais vezes identificada, uma vez que esta se encontra presente em *Porphyromonas gingivalis*, *Streptococcus pneumoniae* e *Streptococcus salivarius*.

A enolase, também conhecida como fosfopiruvato hidratase, é uma enzima glicolítica importante geralmente encontrada no citoplasma. Esta metaloenzima pertencente à classe das liases, e catalisa a desidratação do 2-fosfoglicerato (2-PG) em fosfo-enol-piruvato (PEP), no penúltimo passo da glicólise, podendo também catalisar a reação inversa, dependendo das concentrações no ambiente de substratos, sendo assim um importante intermediário metabólico (106, 107). O pH mais adequado para esta enzima é 6,5, encontrando-se presente em todos os tecidos e organismos capazes de realizar glicólise ou fermentação . A enzima foi descoberto por Lohmann

e Meyerhof em 1934 (108), tendo sido isolada a partir de uma variedade de origens, em que se inclui o músculo humano e eritrócitos (109).

Sabe-se atualmente que a enolase pode desempenhar uma ou mais funções únicas tanto em procariotas como em eucariotas, sendo este fenómeno conhecido como *protein moonlighting* (109-112). Tem-se descoberto um número crescente de espécies bacterianas que possuem proteínas com atividade *moonlighting*, sabendo-se também que tal processo pode contribuir para um aumento do comportamento de virulência bacteriana.

Um número crescente de bactérias e outros microrganismos têm demonstrado expressar a enolase com atividade *moonlighting* na superfície da célula. Atualmente o maior número de bactérias que possui enolases na sua superfície celular são as Gram positivas, contudo estas existem também Gram negativas (por exemplo, *Borrelia burgdorferi*) (113). A atenção tem-se focado maioritariamente nas enolase das superfícies celulares de *streptococcus* do Grupo A. Pancholi e Fischetti, cuja investigação sobre a *Glyceraldehyde-3-phosphate dehydrogenase* (GAPD) em *Streptococcus pyogenis* foi largamente discutido, contribuiu para muito do trabalho realizado sobre as enolase presentes nas superfícies celulares de *Streptococcus*, e o seu papel na virulência bacteriana. É atualmente reconhecida como estando presente na superfície da maioria dos *Streptococcus* (109).

A enolase é uma proteína que possui várias ações *moonlighting*, tanto em eucariotas como em procariotas (107), tais como a expressão de superfície da enzima como um fator importante na patogénese de *Streptococcus pyogenes* ao contribuir significativamente num aumento da afinidade na ligação a plasminogénios (109). Mais ainda, o efeito bacteriostático do flureto na cárie dentária é considerado ser o resultado da inibição da enolase pelo fluoreto (114, 115). O flúor, ao fazer parte de um complexo constituído por magnésio e fosfato, pode ligar-se ao local ativo da enolase em detrimento do 2-PG (114). Como tal, beber água floretada pode fornecer um nível de flúor suficiente para inibir a enolase das bactérias orais, sem prejudicar os seres humanos. A interrupção da via glicolítica das bactérias, e assim do seu normal funcionamento metabólico, impede a formação da cárie dentária (115, 116).

Partindo destas constatações, Veiga-Malta *et al.* (2004) clonaram e sequenciaram o gene da enolase de *Streptococcus sobrinus* e caracterizaram a expressão da proteína num sistema heterólogo. A enolase recombinante suprime a resposta imune primária contra um antigénio dependente de células – T, mas não estimula as células – B, sugerindo que as atividades imunobiológicas complementares da infeção por *Streptococcus. sobrinus* (estimulação e supressão) possam ser causadas por, pelo menos, duas proteínas. Estes resultados demonstraram que, diferentes proteínas estão envolvidas nos fenómenos imunobiológicos durante a invasão patogénica e subsequente sobrevivência de *Streptococcus sobrinus*, incluindo a enolase como um imunossupressor. No entanto, ter como alvo de neutralização estas conhecidas moléculas pode revelar-se como uma estratégia útil no tratamento da cárie dentária induzida por *Streptococcus sobrinus* (117).

Nesta linha e estudos, Dinis *et al.* (2009) testaram a redução de cárie dentária induzida pela enolase recombinante (enolaseR) de *Streptococcus. sobrinus* na terapêutica de imunidade oral. A enolase é uma proteína que se encontra associada a várias doenças autoimunes por induzir a produção de anticorpos contra ela mesma. Posto isto acharam igualmente necessário averiguar se os anticorpos produzidos pelo sistema imune contra a enolase bacteriana seriam capazes de a distinguir da enolase humana (118).

Este estudo demonstrou que a terapêutica com enolaseR reduziu as lesões por cárie induzidas por *Streptococcus sobrinus* sendo bem tolerada, havendo uma significativa redução de caries no esmalte de ratos imunizados pela enolaseR. Em concordância com a redução do número de lesões por cárie, foi também observado um decréscimo na colonização oral por *Streptococcus sobrinus* em ratos imunizados pela enolaseR. Esta imunidade induziu um elevado nível de IgA na saliva, que reage especificamente à enolaseR, estando este resultado diretamente relacionado com a redução do número de bactérias detetadas, assim como, com uma redução na extensão das cáries observadas neste grupo (118).

Quanto aos efeitos histopatológicos, não parece apresentar efeitos fisiológicos prejudiciais nos ratos, podendo ser um antigénio alvo útil na vacinação humana contra a cárie dentária (118).

Dada a existência de um elevado grau de semelhança entre a enolase de *Streptococcus sobrinus* e de *Streptococcus mutans* (98%), antecipa-se que os anticorpos anti-enolase produzidos contra uma espécie de *Streptococcus* deverão ser capazes de reconhecer a enolase de outra espécie e vice-versa. Desta forma, e ainda que seja necessário uma maior investigação, sai reforçado o interesse da enolaseR como um antigénio candidato à vacina contra a cárie induzida por estas espécies de *Streptococcus* produtoras de ácido lácteo (118).

Da nossa avaliação das interações estabelecidas por estes quatro organismos podemos também verificar que existem três proteínas que são comuns a *Streptococcus*, *Porphyromonas* e *Rothia*, que se caracterizam por serem exploradas pelos patogénicos, numa tentativa de entrarem dentro do hospedeiro e sobreviverem dentro deste. Estas proteínas são: *Epidermal growth factor receptor*, Fibronectina, *Cullin* (CAND1).

O *Epidermal growth factor receptor* (EGFR), é uma proteína transmembranar, produzida principalmente nas glândulas salivares e nos rins (119). A sua associação com a invasão por microrganismos já foi relatada para *Salmonella typhimurium* (120), *Candida albicans* (121), *Reovirus* (122), e *Vaccinia virus* (123). Aparentemente todos estes agentes patogéneos iniciam a invasão celular em certo ponto pela ligação com o EGFR. Este facto sugere a possibilidade de que possam existir outros agentes patogéneos que estejam a usar o EGFR para iniciar a colonização do hospedeiro, como foi sugerido por Buret et al. (124).

De igual modo a Fibronectina parece também ser usada pelos microrganismos na colonização do hospedeiro (3). Esta glicoproteína encontra-se ligada às integrinas  $\beta_1$  na superfície celular, e é considerada uma proteína essencial na aderência de bactérias no interior da cavidade oral (125, 126).

A proteína CAND1, anteriormente designada TIP120A, interage com a maioria das proteínas da família Cullin (127). A família de proteínas Cullin desempenha um papel chave na ubiquitinação das proteínas celulares, isto é, realizam a modificação pós tradução com o objetivo de marcar a proteína alvo com moléculas de ubiquitina. Esta marca na proteína alvo resulta frequentemente no comprometimento da proteína *ubiquitin – linked* à degradação proteosomal (128). Consequentemente foi sugerido que a CAND1 poderia funcionar como um regulador global das *cullin* que possuíam *ubiquitin*

*ligases* (127, 129). Sabe-se que certas bactérias corrompem o mecanismo de ubiquitinação como meio de regular os seus fatores de virulência, ou para desencadear a internalização das bactérias nas células do hospedeiro (130). Este mecanismo melhora a sobrevivência e as hipóteses de replicação das bactérias dentro do hospedeiro (3).

Apesar de o EGFR, a fibronectina e a CAND1 serem as proteínas que possuem um maior número de interações com as proteínas do hospedeiro, foi possível identificar outras proteínas com potencial de interesse na interação com o hospedeiro, como é o caso da proteína *Ubiquitin-like* ISG15 (Interferon-induced 15 kDa protein) (Interferon-induced 17 kDa protein) (IP17) (Ubiquitin cross-reactive protein) (hUCRP) e a *Protocadherin-1* (Cadherin-like protein 1) (Protocadherin-42) (PC42).

A *Ubiquitin-like* ISG15 aparece nos nossos dados a interagir com proteínas de *Porphyromonas gingivalis*, *Streptococcus salivarius* e *Streptococcus pneumoniae*. A ubiquitina é uma proteína encontrada nas células eucariotas e desempenha uma função importante na regulação de proteínas (131), existindo 4 genes no genoma humano que a produzem; UBB, UBC, UBA52, e RPS27A (132). Embora a ubiquitina seja o modificador pós tradução melhor compreendido, existe uma família numerosa de proteínas ubiquitin-like (UBLs), que modificam alvos celulares num percurso que é paralelo ao da ubiquitina, mas distinto (133). É nesta família de UBLs que se encontra a *Ubiquitin-like* ISG15. Esta proteína tem um papel importante na sinalização da resposta do hospedeiro a vírus não apenas intracelularmente mas também como proteína que é secretada para o exterior da célula e atua tanto quanto se sabe em linfócitos T e em células NK de forma a induzir a produção de interferão (IFN)- $\gamma$  (134). É nesta forma secretada que a proteína pode interagir com as várias proteínas microbianas mencionadas diminuindo a sua quantidade disponível para ativar a resposta a vírus pelas células do sistema imunitário do hospedeiro.

A protocadherina-1 é uma proteína membranar localizada nos limites da junção célula-célula, encontrando-se envolvida na adesão das células neurais, o que sugere um possível papel no desenvolvimento neuronal. As células que expressam esta proteína apresentam atividade de agregação celular (135). O facto de a proteína em questão estar sobretudo presente em tecidos nervosos

faz com que a sua relevância para a discussão das interações na cavidade oral seja negligenciável.

## 6. CONCLUSÃO

A comparação dos dados de proteômica da cavidade oral existentes, no que se refere aos microrganismos, com os recentes resultados da metagenômica do microbioma humano da cavidade oral em saúde, permite concluir que a cobertura em termos de proteínas microbianas identificadas é muito baixa.

Das proteínas microbianas dos organismos *Aggregatibacter*, *Porphyromonas*, *Rothia* e *Streptococcus* identificadas pelo algoritmo Orallnt, foram analisadas aquelas que apresentam maiores probabilidades de serem importantes na interação com o hospedeiro. Assim, estas moléculas poderão ser os alvos primários em estudos futuros de interação 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.

A enolase mostrou ser a proteína dentro do grupo de proteínas microbianas que apresentam ontologias com potencial de interesse na interação com o hospedeiro, mais vezes identificada. A capacidade da enolase bacteriana poder desempenhar várias funções, foi reconhecido recentemente, e o algoritmo Orallnt identificou uma possível interação com uma aquaporina do hospedeiro que poderia interferir com os mecanismos de homeostase.

Pelo que foi possível inferir dos nossos resultados, as interações previstas entre proteínas humanas e microbianas, faz prever que a resposta a vírus possa também estar comprometida, devido à interação da *Ubiquitin-like* ISG15 do hospedeiro, com proteínas microbianas, o que pode levar a uma diminuição da sua quantidade disponível para ativar a resposta a vírus pelas células do sistema imunitário do hospedeiro.

Quando testado com exemplos reais, o algoritmo Orallnt identificou interações que já tinham sido descritas anteriormente, mas também novas interações potencialmente interessantes na biologia oral, tanto entre o hospedeiro e os patógenos periodontais, como entre os diferentes patógenos

periodontais, o que demonstra a sinergia existente no biofilme oral na doença periodontal. Ainda no que diz respeito à interação entre bactérias, foram identificadas interações não documentadas para *Streptococcus pneumoniae* SP23-BS72 e *Porphyromonas gingivalis* com *Haemophilus influenzae* e *Pasteurella multocida*, o que nos pode levar a supor que estes microrganismos podem coexistir na colonização das superfícies do hospedeiro e interagir durante a formação do biofilme oral. Contudo, de realçar que não foram encontradas tantas interações como seria esperado para *Aggregatibacter actinomycetemcomitans* e *Porphyromonas gingivalis* e proteínas do hospedeiro, o que não deveria acontecer, uma vez que estas duas bactérias são as mais associadas à doença periodontal e há várias interações descritas entre proteínas destas bactérias e proteínas humanas. No que diz respeito à etiologia da cárie, o algoritmo Orallnt também identificou poucas interações com as proteínas de *Streptococcus mutans*, apesar de muitas destas já se encontrarem, visto que esta bactéria é considerada o principal fator no desenvolvimento da cárie dentária.

Assim sendo, é necessário a incrementação da informação fornecida ao algoritmo Orallnt. Por um lado é necessário aumentar o número de proteínas microbianas incluídas na análise. Por outro lado o aumento gradual da anotação das proteínas microbianas possibilitará um aumento no desempenho do algoritmo. Desta forma as previsões podem tornar-se mais precisas e extensas identificando algumas das interações já descritas e propondo novas interações.

## 7. BIBILOGRAFIA

1. Pesquita C, Faria D, Falcao AO, Lord P, Couto FM. Semantic similarity in biomedical ontologies. *PLoS computational biology*. 2009;5(7):e1000443.
2. Rosa N, Correia MJ, Arrais JP, Lopes P, Melo J, Oliveira JL, et al. From the salivary proteome to the OralOme: comprehensive molecular oral biology. *Archives of oral biology*. 2012;57(7):853-64.
3. Coelho ED, Arrais JP, Matos S, Pereira C, Rosa N, Correia MJ, et al. Computational prediction of the human-microbial oral interactome. *BMC systems biology*. 2014;8:24.
4. Gehlenborg N, O'Donoghue SI, Baliga NS, Goesmann A, Hibbs MA, Kitano H, et al. Visualization of omics data for systems biology. *Nature methods*. 2010;7(3 Suppl):S56-68.
5. Fowler JH, Dawes CT, Christakis NA. Model of genetic variation in human social networks. *Proceedings of the National Academy of Sciences of the United States of America*. 2009;106(6):1720-4.
6. Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics*. 2011;27(3):431-2.
7. Human Microbiome Project C. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486(7402):207-14.
8. Janoff EN, Gustafson C, Frank DN. The world within: living with our microbial guests and guides. *Translational research : the journal of laboratory and clinical medicine*. 2012;160(4):239-45.
9. Human Microbiome Jumpstart Reference Strains C, Nelson KE, Weinstock GM, Highlander SK, Worley KC, Creasy HH, et al. A catalog of reference genomes from the human microbiome. *Science*. 2010;328(5981):994-9.

10. Avila M, Ojcius DM, Yilmaz O. The oral microbiota: living with a permanent guest. *DNA and cell biology*. 2009;28(8):405-11.
11. Human Microbiome Project C. A framework for human microbiome research. *Nature*. 2012;486(7402):215-21.
12. Lemos JA, Brown TA, Jr., Burne RA. Effects of RelA on key virulence properties of planktonic and biofilm populations of *Streptococcus mutans*. *Infection and immunity*. 2004;72(3):1431-40.
13. Maeda K, Tribble GD, Tucker CM, Anaya C, Shizukuishi S, Lewis JP, et al. A *Porphyromonas gingivalis* tyrosine phosphatase is a multifunctional regulator of virulence attributes. *Molecular microbiology*. 2008;69(5):1153-64.
14. Sharma V, Suvarna K, Meganathan R, Hudspeth ME. Menaquinone (vitamin K2) biosynthesis: nucleotide sequence and expression of the *menB* gene from *Escherichia coli*. *Journal of bacteriology*. 1992;174(15):5057-62.
15. Munoz N, Van Maele L, Marques JM, Rial A, Sirard JC, Chabalgoity JA. Mucosal administration of flagellin protects mice from *Streptococcus pneumoniae* lung infection. *Infection and immunity*. 2010;78(10):4226-33.
16. Hooton TM, Stamm WE. Diagnosis and treatment of uncomplicated urinary tract infection. *Infectious disease clinics of North America*. 1997;11(3):551-81.
17. Bjarnsholt T, Jensen PO, Fiandaca MJ, Pedersen J, Hansen CR, Andersen CB, et al. *Pseudomonas aeruginosa* biofilms in the respiratory tract of cystic fibrosis patients. *Pediatric pulmonology*. 2009;44(6):547-58.
18. Murray PR, Rosenthal KS, Pfaller MA. *Medical microbiology*. Philadelphia: Elsevier/Saunders; 2013. p. p.
19. Group NHW, Peterson J, Garges S, Giovanni M, McInnes P, Wang L, et al. The NIH Human Microbiome Project. *Genome research*. 2009;19(12):2317-23.

20. Kuramitsu HK. Proteases of *Porphyromonas gingivalis*: what don't they do? *Oral microbiology and immunology*. 1998;13(5):263-70.
21. Volk WA. *Essentials of medical microbiology*. 5th ed. Philadelphia: Lippincott-Raven Publishers; 1996. viii, 725 p. p.
22. Busuioc M, Buttaro BA, Piggot PJ. The pdh operon is expressed in a subpopulation of stationary-phase bacteria and is important for survival of sugar-starved *Streptococcus mutans*. *Journal of bacteriology*. 2010;192(17):4395-402.
23. Smith MA, Finel M, Korolik V, Mendz GL. Characteristics of the aerobic respiratory chains of the microaerophiles *Campylobacter jejuni* and *Helicobacter pylori*. *Archives of microbiology*. 2000;174(1-2):1-10.
24. Hong SH, Kim JS, Lee SY, In YH, Choi SS, Rih JK, et al. The genome sequence of the capnophilic rumen bacterium *Mannheimia succiniciproducens*. *Nature biotechnology*. 2004;22(10):1275-81.
25. Li YH, Tian X. Quorum sensing and bacterial social interactions in biofilms. *Sensors*. 2012;12(3):2519-38.
26. Newman MG, Takei HH, Carranza FnA. *Carranza's clinical periodontology*. 11th ed. St. Louis, Mo.: Elsevier/Saunders; 2012. xlv, 825 p. p.
27. Jenkinson HF, Lamont RJ. Oral microbial communities in sickness and in health. *Trends in microbiology*. 2005;13(12):589-95.
28. Singh A, Wyant T, Anaya-Bergman C, Aduse-Opoku J, Brunner J, Laine ML, et al. The capsule of *Porphyromonas gingivalis* leads to a reduction in the host inflammatory response, evasion of phagocytosis, and increase in virulence. *Infection and immunity*. 2011;79(11):4533-42.
29. Fenno JC. *Treponema denticola* interactions with host proteins. *Journal of oral microbiology*. 2012;4.

30. Arirachakaran P, Apinhasmit W, Paungmalit P, Jeramethakul P, Rerkyen P, Mahanonda R. Infection of human gingival fibroblasts with *Aggregatibacter actinomycetemcomitans*: An in vitro study. *Archives of oral biology*. 2012;57(7):964-72.
31. Shames SR, Finlay BB. Bacterial effector interplay: a new way to view effector function. *Trends in microbiology*. 2012;20(5):214-9.
32. Kline KA, Falker S, Dahlberg S, Normark S, Henriques-Normark B. Bacterial adhesins in host-microbe interactions. *Cell host & microbe*. 2009;5(6):580-92.
33. Lantz MS. Are bacterial proteases important virulence factors? *Journal of periodontal research*. 1997;32(1 Pt 2):126-32.
34. Sanchez C, Lachaize C, Janody F, Bellon B, Roder L, Euzenat J, et al. Grasping at molecular interactions and genetic networks in *Drosophila melanogaster* using FlyNets, an Internet database. *Nucleic acids research*. 1999;27(1):89-94.
35. Hennah W, Porteous D. The DISC1 pathway modulates expression of neurodevelopmental, synaptogenic and sensory perception genes. *PloS one*. 2009;4(3):e4906.
36. Kiemer L, Cesareni G. Comparative interactomics: comparing apples and pears? *Trends in biotechnology*. 2007;25(10):448-54.
37. Chen YC, Rajagopala SV, Stellberger T, Uetz P. Exhaustive benchmarking of the yeast two-hybrid system. *Nature methods*. 2010;7(9):667-8; author reply 8.
38. Parrish JR, Yu J, Liu G, Hines JA, Chan JE, Mangiola BA, et al. A proteome-wide protein interaction map for *Campylobacter jejuni*. *Genome biology*. 2007;8(7):R130.
39. Titz B, Rajagopala SV, Goll J, Hauser R, McKeivitt MT, Palzkill T, et al. The binary protein interactome of *Treponema pallidum*--the syphilis spirochete. *PloS one*. 2008;3(5):e2292.
40. Hu P, Janga SC, Babu M, Diaz-Mejia JJ, Butland G, Yang W, et al. Global functional atlas of *Escherichia coli* encompassing previously uncharacterized proteins. *PLoS biology*. 2009;7(4):e96.

41. Shimoda Y, Shinpo S, Kohara M, Nakamura Y, Tabata S, Sato S. A large scale analysis of protein-protein interactions in the nitrogen-fixing bacterium *Mesorhizobium loti*. *DNA research : an international journal for rapid publication of reports on genes and genomes*. 2008;15(1):13-23.
42. Wang Y, Cui T, Zhang C, Yang M, Huang Y, Li W, et al. Global protein-protein interaction network in the human pathogen *Mycobacterium tuberculosis* H37Rv. *Journal of proteome research*. 2010;9(12):6665-77.
43. Kuhner S, van Noort V, Betts MJ, Leo-Macias A, Batisse C, Rode M, et al. Proteome organization in a genome-reduced bacterium. *Science*. 2009;326(5957):1235-40.
44. Sato S, Shimoda Y, Muraki A, Kohara M, Nakamura Y, Tabata S. A large-scale protein protein interaction analysis in *Synechocystis* sp. PCC6803. *DNA research : an international journal for rapid publication of reports on genes and genomes*. 2007;14(5):207-16.
45. Arnold R, Boonen K, Sun MG, Kim PM. Computational analysis of interactomes: current and future perspectives for bioinformatics approaches to model the host-pathogen interaction space. *Methods*. 2012;57(4):508-18.
46. Elde NC, Malik HS. The evolutionary conundrum of pathogen mimicry. *Nature reviews Microbiology*. 2009;7(11):787-97.
47. Aguilera G, Refregier G, Yockteng R, Fournier E, Giraud T. Rapidly evolving genes in pathogens: methods for detecting positive selection and examples among fungi, bacteria, viruses and protists. *Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2009;9(4):656-70.
48. Rogers JD, Palmer RJ, Jr., Kolenbrander PE, Scannapieco FA. Role of *Streptococcus gordonii* amylase-binding protein A in adhesion to hydroxyapatite, starch metabolism, and biofilm formation. *Infection and immunity*. 2001;69(11):7046-56.
49. Coelho ED, Arrais JP, Oliveira JL. From protein-protein interactions to rational drug design: are computational methods up to the challenge? *Current topics in medicinal chemistry*. 2013;13(5):602-18.

50. Phizicky EM, Fields S. Protein-protein interactions: methods for detection and analysis. *Microbiological reviews*. 1995;59(1):94-123.
51. Dyer MD, Murali TM, Sobral BW. Computational prediction of host-pathogen protein-protein interactions. *Bioinformatics*. 2007;23(13):i159-66.
52. Shoemaker BA, Panchenko AR. Deciphering protein-protein interactions. Part I. Experimental techniques and databases. *PLoS computational biology*. 2007;3(3):e42.
53. Ofran Y, Rost B. Analysing six types of protein-protein interfaces. *Journal of molecular biology*. 2003;325(2):377-87.
54. Valdar WS, Thornton JM. Protein-protein interfaces: analysis of amino acid conservation in homodimers. *Proteins*. 2001;42(1):108-24.
55. Panchenko AR, Wolf YI, Panchenko LA, Madej T. Evolutionary plasticity of protein families: coupling between sequence and structure variation. *Proteins*. 2005;61(3):535-44.
56. Ito T, Chiba T, Ozawa R, Yoshida M, Hattori M, Sakaki Y. A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences of the United States of America*. 2001;98(8):4569-74.
57. Ho Y, Gruhler A, Heilbut A, Bader GD, Moore L, Adams SL, et al. Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*. 2002;415(6868):180-3.
58. Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, et al. Global analysis of protein activities using proteome chips. *Science*. 2001;293(5537):2101-5.
59. Sidhu SS, Fairbrother WJ, Deshayes K. Exploring protein-protein interactions with phage display. *Chembiochem : a European journal of chemical biology*. 2003;4(1):14-25.
60. Tong AH, Evangelista M, Parsons AB, Xu H, Bader GD, Page N, et al. Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science*. 2001;294(5550):2364-8.

61. Ge H, Liu Z, Church GM, Vidal M. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. *Nature genetics*. 2001;29(4):482-6.
62. Kukar T, Eckenrode S, Gu Y, Lian W, Megginson M, She JX, et al. Protein microarrays to detect protein-protein interactions using red and green fluorescent proteins. *Analytical biochemistry*. 2002;306(1):50-4.
63. Xia JF, Wang SL, Lei YK. Computational methods for the prediction of protein-protein interactions. *Protein and peptide letters*. 2010;17(9):1069-78.
64. Skrabanek L, Saini HK, Bader GD, Enright AJ. Computational prediction of protein-protein interactions. *Molecular biotechnology*. 2008;38(1):1-17.
65. Liu Y, Kim I, Zhao H. Protein interaction predictions from diverse sources. *Drug discovery today*. 2008;13(9-10):409-16.
66. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. UniProtKB/Swiss-Prot. *Methods in molecular biology*. 2007;406:89-112.
67. Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. *Nucleic acids research*. 2010;38(Database issue):D204-10.
68. Scheer M, Grote A, Chang A, Schomburg I, Munaretto C, Rother M, et al. BRENDA, the enzyme information system in 2011. *Nucleic acids research*. 2011;39(Database issue):D670-6.
69. Chen L, Xiong Z, Sun L, Yang J, Jin Q. VFDB 2012 update: toward the genetic diversity and molecular evolution of bacterial virulence factors. *Nucleic acids research*. 2012;40(Database issue):D641-5.
70. Snel B, Lehmann G, Bork P, Huynen MA. STRING: a web-server to retrieve and display the repeatedly occurring neighbourhood of a gene. *Nucleic acids research*. 2000;28(18):3442-4.

71. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguéz P, et al. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic acids research*. 2011;39(Database issue):D561-8.
72. Chaurasia G, Malhotra S, Russ J, Schnoegl S, Hanig C, Wanker EE, et al. UniHI 4: new tools for query, analysis and visualization of the human protein-protein interactome. *Nucleic acids research*. 2009;37(Database issue):D657-60.
73. Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, Livstone MS, et al. The BioGRID Interaction Database: 2011 update. *Nucleic acids research*. 2011;39(Database issue):D698-704.
74. Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D. The Database of Interacting Proteins: 2004 update. *Nucleic acids research*. 2004;32(Database issue):D449-51.
75. Gong S, Yoon G, Jang I, Bolser D, Dafas P, Schroeder M, et al. PSImap: a database of Protein Structural Interactome map (PSIMAP). *Bioinformatics*. 2005;21(10):2541-3.
76. Chen T, Yu WH, Izard J, Baranova OV, Lakshmanan A, Dewhirst FE. The Human Oral Microbiome Database: a web accessible resource for investigating oral microbe taxonomic and genomic information. *Database : the journal of biological databases and curation*. 2010;2010:baq013.
77. Arrais JP, Rosa N, Melo J, Coelho ED, Amaral D, Correia MJ, et al. OralCard: a bioinformatic tool for the study of oral proteome. *Archives of oral biology*. 2013;58(7):762-72.
78. Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M. BioGRID: a general repository for interaction datasets. *Nucleic acids research*. 2006;34(Database issue):D535-9.
79. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, et al. Human Protein Reference Database--2009 update. *Nucleic acids research*. 2009;37(Database issue):D767-72.

80. Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, et al. The IntAct molecular interaction database in 2012. *Nucleic acids research*. 2012;40(Database issue):D841-6.
81. Licata L, Briganti L, Peluso D, Perfetto L, Iannuccelli M, Galeota E, et al. MINT, the molecular interaction database: 2012 update. *Nucleic acids research*. 2012;40(Database issue):D857-61.
82. Pislari AH, Kos J. C-terminal peptide of gamma-enolase impairs amyloid-beta-induced apoptosis through p75(NTR) signaling. *Neuromolecular medicine*. 2013;15(3):623-35.
83. Hafner A, Glavan G, Obermajer N, Zivin M, Schliebs R, Kos J. Neuroprotective role of gamma-enolase in microglia in a mouse model of Alzheimer's disease is regulated by cathepsin X. *Aging cell*. 2013;12(4):604-14.
84. Glenting J, Beck HC, Vrang A, Riemann H, Ravn P, Hansen AM, et al. Anchorless surface associated glycolytic enzymes from *Lactobacillus plantarum* 299v bind to epithelial cells and extracellular matrix proteins. *Microbiological research*. 2013;168(5):245-53.
85. Marcos CM, de Fatima da Silva J, de Oliveira HC, Moraes da Silva RA, Mendes-Giannini MJ, Fusco-Almeida AM. Surface-expressed enolase contributes to the adhesion of *Paracoccidioides brasiliensis* to host cells. *FEMS yeast research*. 2012;12(5):557-70.
86. Ozaki N, Ohmuraya M, Ida S, Hashimoto D, Ikuta Y, Chikamoto A, et al. Serine protease inhibitor Kazal type 1 and epidermal growth factor receptor are expressed in pancreatic tubular adenocarcinoma, intraductal papillary mucinous neoplasm, and pancreatic intraepithelial neoplasia. *Journal of hepato-biliary-pancreatic sciences*. 2013.
87. Chen M, Chen LM, Lin CY, Chai KX. The epidermal growth factor receptor (EGFR) is proteolytically modified by the Matriptase-Prostasin serine protease cascade in cultured epithelial cells. *Biochimica et biophysica acta*. 2008;1783(5):896-903.
88. Tiaden AN, Klawitter M, Lux V, Mirsaidi A, Bahrenberg G, Glanz S, et al. Detrimental role for human high temperature requirement serine protease A1 (HTRA1) in the pathogenesis of

intervertebral disc (IVD) degeneration. *The Journal of biological chemistry*. 2012;287(25):21335-45.

89. Jagtap P, McGowan T, Bandhakavi S, Tu ZJ, Seymour S, Griffin TJ, et al. Deep metaproteomic analysis of human salivary supernatant. *Proteomics*. 2012;12(7):992-1001.

90. Avila-Campos MJ, Velasquez-Melendez G. Prevalence of putative periodontopathogens from periodontal patients and healthy subjects in Sao Paulo, SP, Brazil. *Revista do Instituto de Medicina Tropical de Sao Paulo*. 2002;44(1):1-5.

91. Rico S, Santamaria RI, Yepes A, Rodriguez H, Laing E, Bucca G, et al. Deciphering the Regulon of *Streptomyces coelicolor* AbrC3, a Positive Response Regulator of Antibiotic Production. *Applied and environmental microbiology*. 2014;80(8):2417-28.

92. Braat H, Rottiers P, Hommes DW, Huyghebaert N, Remaut E, Remon JP, et al. A phase I trial with transgenic bacteria expressing interleukin-10 in Crohn's disease. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2006;4(6):754-9.

93. Puig C, Calatayud L, Marti S, Tubau F, Garcia-Vidal C, Carratala J, et al. Molecular epidemiology of nontypeable *Haemophilus influenzae* causing community-acquired pneumonia in adults. *PloS one*. 2013;8(12):e82515.

94. Roberts MC, Soge OO, No DB. Characterization of macrolide resistance genes in *Haemophilus influenzae* isolated from children with cystic fibrosis. *The Journal of antimicrobial chemotherapy*. 2011;66(1):100-4.

95. Conn JE. The Pigment Production of *Actinomyces coelicolor* and *A. violaceus-ruber*. *Journal of bacteriology*. 1943;46(2):133-49.

96. Zhang X, Clark CA, Pettis GS. Interstrain inhibition in the sweet potato pathogen *Streptomyces ipomoeae*: purification and characterization of a highly specific bacteriocin and cloning of its structural gene. *Applied and environmental microbiology*. 2003;69(4):2201-8.

97. Liu CM, McDaniel LE, Schaffner CP. Fungimycin, biogenesis of its aromatic moiety. *The Journal of antibiotics*. 1972;25(3):187-8.
98. Lee CH, Schaffner CP. Perimycin. The structure of some degradation products. *Tetrahedron*. 1969;25(10):2229-32.
99. Brian P, Riggle PJ, Santos RA, Champness WC. Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative signal transduction system. *Journal of bacteriology*. 1996;178(11):3221-31.
100. Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature*. 2002;417(6885):141-7.
101. Ishikawa Y, Cho G, Yuan Z, Inoue N, Nakae Y. Aquaporin-5 water channel in lipid rafts of rat parotid glands. *Biochimica et biophysica acta*. 2006;1758(8):1053-60.
102. Facklam R. What happened to the streptococci: overview of taxonomic and nomenclature changes. *Clinical microbiology reviews*. 2002;15(4):613-30.
103. Siemieniuk RA, Gregson DB, Gill MJ. The persisting burden of invasive pneumococcal disease in HIV patients: an observational cohort study. *BMC infectious diseases*. 2011;11:314.
104. Kaci G, Goudercourt D, Dennin V, Pot B, Dore J, Ehrlich SD, et al. Anti-inflammatory properties of *Streptococcus salivarius*, a commensal bacterium of the oral cavity and digestive tract. *Applied and environmental microbiology*. 2014;80(3):928-34.
105. Nogueira RD, King WF, Gunda G, Culshaw S, Taubman MA, Mattos-Graner RO, et al. Mutans streptococcal infection induces salivary antibody to virulence proteins and associated functional domains. *Infection and immunity*. 2008;76(8):3606-13.
106. Zhang E, Brewer JM, Minor W, Carreira LA, Lebioda L. Mechanism of enolase: the crystal structure of asymmetric dimer enolase-2-phospho-D-glycerate/enolase-phosphoenolpyruvate at 2.0 Å resolution. *Biochemistry*. 1997;36(41):12526-34.

107. Pancholi V. Multifunctional alpha-enolase: its role in diseases. Cellular and molecular life sciences : CMLS. 2001;58(7):902-20.
108. Peshavaria M, Day IN. Molecular structure of the human muscle-specific enolase gene (ENO3). The Biochemical journal. 1991;275 ( Pt 2):427-33.
109. Pancholi V, Fischetti VA. alpha-enolase, a novel strong plasmin(ogen) binding protein on the surface of pathogenic streptococci. The Journal of biological chemistry. 1998;273(23):14503-15.
110. Wistow GJ, Lietman T, Williams LA, Stapel SO, de Jong WW, Horwitz J, et al. Tau-crystallin/alpha-enolase: one gene encodes both an enzyme and a lens structural protein. The Journal of cell biology. 1988;107(6 Pt 2):2729-36.
111. Aaronson RM, Graven KK, Tucci M, McDonald RJ, Farber HW. Non-neuronal enolase is an endothelial hypoxic stress protein. The Journal of biological chemistry. 1995;270(46):27752-7.
112. Breitenbach M, Simon B, Probst G, Oberkofler H, Ferreira F, Briza P, et al. Enolases are highly conserved fungal allergens. International archives of allergy and immunology. 1997;113(1-3):114-7.
113. Nowalk AJ, Nolder C, Clifton DR, Carroll JA. Comparative proteome analysis of subcellular fractions from *Borrelia burgdorferi* by NEPHGE and IPG. Proteomics. 2006;6(7):2121-34.
114. Hamilton IR. Effects of fluoride on enzymatic regulation of bacterial carbohydrate metabolism. Caries research. 1977;11 Suppl 1:262-91.
115. Huther FJ, Psarros N, Duschner H. Isolation, characterization, and inhibition kinetics of enolase from *Streptococcus rattus* FA-1. Infection and immunity. 1990;58(4):1043-7.
116. Centers for Disease C, Prevention. Populations receiving optimally fluoridated public drinking water--United States, 2000. MMWR Morbidity and mortality weekly report. 2002;51(7):144-7.

117. Veiga-Malta I, Duarte M, Dinis M, Tavares D, Videira A, Ferreira P. Enolase from *Streptococcus sobrinus* is an immunosuppressive protein. *Cellular microbiology*. 2004;6(1):79-88.
118. Dinis M, Tavares D, Veiga-Malta I, Fonseca AJ, Andrade EB, Trigo G, et al. Oral therapeutic vaccination with *Streptococcus sobrinus* recombinant enolase confers protection against dental caries in rats. *The Journal of infectious diseases*. 2009;199(1):116-23.
119. Barnard JA, Beauchamp RD, Russell WE, Dubois RN, Coffey RJ. Epidermal growth factor-related peptides and their relevance to gastrointestinal pathophysiology. *Gastroenterology*. 1995;108(2):564-80.
120. Galan JE, Pace J, Hayman MJ. Involvement of the epidermal growth factor receptor in the invasion of cultured mammalian cells by *Salmonella typhimurium*. *Nature*. 1992;357(6379):588-9.
121. Zhu W, Phan QT, Boontheung P, Solis NV, Loo JA, Filler SG. EGFR and HER2 receptor kinase signaling mediate epithelial cell invasion by *Candida albicans* during oropharyngeal infection. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109(35):14194-9.
122. Strong JE, Tang D, Lee PW. Evidence that the epidermal growth factor receptor on host cells confers reovirus infection efficiency. *Virology*. 1993;197(1):405-11.
123. Eppstein DA, Marsh YV, Schreiber AB, Newman SR, Todaro GJ, Nestor JJ, Jr. Epidermal growth factor receptor occupancy inhibits vaccinia virus infection. *Nature*. 1985;318(6047):663-5.
124. Buret A, Gall DG, Olson ME, Hardin JA. The role of the epidermal growth factor receptor in microbial infections of the gastrointestinal tract. *Microbes and infection / Institut Pasteur*. 1999;1(13):1139-44.
125. Llana-Puy MC, Montanana-Llorens C, Forner-Navarro L. Fibronectin levels in stimulated whole-saliva and their relationship with cariogenic oral bacteria. *International dental journal*. 2000;50(1):57-9.

126. Henderson B, Nair S, Pallas J, Williams MA. Fibronectin: a multidomain host adhesin targeted by bacterial fibronectin-binding proteins. *FEMS microbiology reviews*. 2011;35(1):147-200.
127. Min KW, Hwang JW, Lee JS, Park Y, Tamura TA, Yoon JB. TIP120A associates with cullins and modulates ubiquitin ligase activity. *The Journal of biological chemistry*. 2003;278(18):15905-10.
128. Sarikas A, Hartmann T, Pan ZQ. The cullin protein family. *Genome biology*. 2011;12(4):220.
129. Zheng J, Yang X, Harrell JM, Ryzhikov S, Shim EH, Lykke-Andersen K, et al. CAND1 binds to unneddylated CUL1 and regulates the formation of SCF ubiquitin E3 ligase complex. *Molecular cell*. 2002;10(6):1519-26.
130. Munro P, Flatau G, Lemichez E. Bacteria and the ubiquitin pathway. *Current opinion in microbiology*. 2007;10(1):39-46.
131. Goldstein G, Scheid M, Hammerling U, Schlesinger DH, Niall HD, Boyse EA. Isolation of a polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. *Proceedings of the National Academy of Sciences of the United States of America*. 1975;72(1):11-5.
132. Kimura Y, Tanaka K. Regulatory mechanisms involved in the control of ubiquitin homeostasis. *Journal of biochemistry*. 2010;147(6):793-8.
133. Downes BP, Saracco SA, Lee SS, Crowell DN, Vierstra RD. MUBs, a family of ubiquitin-fold proteins that are plasma membrane-anchored by prenylation. *The Journal of biological chemistry*. 2006;281(37):27145-57.
134. Bogunovic D, Boisson-Dupuis S, Casanova JL. ISG15: leading a double life as a secreted molecule. *Experimental & molecular medicine*. 2013;45:e18.
135. Kim SY, Yasuda S, Tanaka H, Yamagata K, Kim H. Non-clustered protocadherin. *Cell adhesion & migration*. 2011;5(2):97-105.

