



UNIVERSIDADE
CATÓLICA | INSTITUTO DE
PORTUGUESA | CIÊNCIAS DA SAÚDE

BIOMARCADORES NA SALIVA – PRESENTE E FUTURO

Dissertação apresentada à Universidade Católica Portuguesa

Para obtenção do grau de Mestre em Medicina Dentária

Por:

Jorge Filipe Raposo Fonseca

Viseu, 2014



UNIVERSIDADE
CATÓLICA | INSTITUTO DE
PORTUGUESA | CIÊNCIAS DA SAÚDE

BIOMARCADORES NA SALIVA – PRESENTE E FUTURO

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

Por:

Jorge Filipe Raposo Fonseca

Sob orientação da Professora Doutora Marlene Barros

Viseu, 2014

“Every path is the right path. Everything could've been anything else. And it would have just as much meaning.”

Mr.Nobody

Agradecimentos

À minha orientadora, Professora Doutora Marlene Barros,

Pela sua capacidade excepcional de coordenar trabalhos.

"Se vi mais longe foi por estar de pé sobre ombros de gigantes.", Robert Hooke.

À Professora Doutora Maria José Correia,

Pela sua frontalidade, simpatia e disponibilidade ao longo de todo este percurso, estando sempre de sorriso aberto e pronta a ajudar.

Ao Professor Doutor Nuno Rosa,

Pela sua dinâmica, gosto e capacidade de transmitir conhecimentos de domínio complexo a quem não pertence à sua área científica. Muito Obrigado.

Ao meu Pai e à minha Mãe,

Por tudo.

Aos meus avós,

Pela sabedoria e valores transmitidos ao longo da minha vida.

À minha tia Carolina,

Que infelizmente e de um modo súbito não me conseguiu acompanhar até ao fim.

Aos meus irmãos,

Por terem tido a paciência para me aturar em criança, adolescente e como jovem adulto.

À Lili,

Por ser o meu pilar.

Ao Ricardo,

Pelo companheirismo ao longo do curso.

Nota prévia

Ao longo deste trabalho irão ser apresentados termos na língua Inglesa cuja tradução não foi realizada com vista a manutenção do seu significado original. A classificação de proteínas de acordo com um sistema de classificação de ontologias pressupõe a sua utilização em língua Inglesa, por forma a evitar ambiguidades e perdas de significados do ponto de vista biológico (1).

Índice

Nota prévia	V
1. Introdução e enquadramento teórico	1
1.1. A cavidade oral e o proteoma salivar	1
1.2. A saliva como secreção ativa glandular.....	2
1.3. A função fisiológica salivar	4
1.4. A saliva como fluído de diagnóstico	4
1.5. A importância de diagnósticos salivares dirigidos	5
1.5.1. <i>Diagnóstico salivar em doenças locais e sistémicas</i>	5
1.6. Biomarcadores na saliva	5
1.7. Resposta imune e proteínas salivares.....	6
1.8. Biologia de sistemas e proteínas salivares.....	6
1.9. Ferramentas bioinformáticas	7
2. Objetivos	9
3. Materiais e métodos	11
3.1. Catalogação das Proteínas Identificadas em cada patologia	12
3.1.1. <i>Proteínas identificadas recolhidas em Whole Saliva</i>	12
3.1.2. <i>Seleção das proteínas de Homo sapiens</i>	13
3.1.3. <i>Seleção das patologias com dados de quantificação</i>	13
3.2. Proteínas Exclusivas de cada patologia	14
3.2.1. <i>Representação gráfica das proteínas exclusivas de cada patologia</i>	15
3.3. Classificação das proteínas por Ontologias Génicas (GO's).....	16
3.3.1. <i>Ontologias génicas derivadas do AgBase</i>	17
3.3.1.1. Ferramenta <i>GORetriever</i>	17
3.3.1.2. Ferramenta <i>GOSlimViewer</i>	19
3.4. Caracterização funcional das proteínas de cada patologia	22
3.4.1. <i>Importação de ontologias a partir de GOSlimViewer</i>	22
3.4.2. <i>Tratamento dos dados provenientes de ontologias</i>	25
3.4.3. <i>Obtenção de atributos para as proteínas exclusivas</i>	29
3.4.4. <i>Tabelas funcionais por patologia</i>	32
3.5. As proteínas propostas como biomarcadores	36

4. Resultados e discussão	39
4.1. Proteínas identificadas por patologia.....	39
4.2. Proteínas exclusivas por patologia	41
4.3. Proteínas exclusivas classificadas por ontologias	43
4.4. Biomarcadores salivares sugeridos	45
4.4.1. Biomarcadores sugeridos com quantificação ≤ -3 ou ≥ 3	45
4.4.2. Biomarcadores sugeridos com quantificação > -3 ou < 3	46
4.4.3. Biomarcadores sugeridos com quantificação qualitativa (aumentados/diminuídos)	47
4.5. Novos biomarcadores sugeridos	49
5. Conclusão	57
6. Bibliografia	59
7. Anexos	67

Índice de figuras

- Figura 1 – Gráfico ilustrativo da contribuição das diferentes glândulas em termos de proteínas presentes na saliva e provenientes de secreção glandular salivar adaptado de (5). 2
- Figura 2 – Gráfico ilustrativo dos grupos proteicos presentes na saliva de *Homo sapiens* adaptado de (6). 3
- Figura 3 – Fluxograma da metodologia seguida neste trabalho. 11
- Figura 4 – Metodologia para seleção das proteínas com recolha na “*Whole Saliva*”, foi desmarcado o visto na opção “(Em branco)” de modo a serem mostradas apenas as patologias com recolha na saliva. 12
- Figura 5 – Metodologia para seleção das proteínas com recolha na “*Whole Saliva*” na espécie *Homo sapiens* removendo-se a selecção de todas as espécies restantes. 13
- Figura 6 – Metodologia para seleccionar as patologias que apresentam proteínas com quantificação. Nesta imagem apenas está representada uma porção das quantificações que na realidade foram seleccionadas. Remoção de “(Em branco)” de modo a remover deste estudo patologias que não possuem proteínas com dados de quantificação. 14
- Figura 7 – Metodologia para seleccionar as proteínas exclusivas em *Breast Neoplasm (MesH ID 68001943)*. É possível verificar a presença de dois gráficos circulares cruzados, sendo que o gráfico a azul possui a contagem de proteínas exclusivas de *Breast Neoplasm (MesH ID 68001943)* e o gráfico a amarelo a contagem de proteínas no OralOma em *Whole Saliva* proveniente de *Homo sapiens*. No cruzamento dos dois gráficos visualiza-se a contagem de proteínas em comum aos dois grupos. Na caixa inferior encontra-se a lista de proteínas exclusivas de *Breast Neoplasm (MesH ID 68001943)*. 16
- Figura 8 – Seleção da ferramenta “*GORetriever*” na página inicial do *AgBase* no separador “*TOOLS*”. 18

Figura 9 – “*Accession Type*” selecionado como “*UniProt Accession*” e seleção do botão “Escolher ficheiro” de modo a fazer *upload* do ficheiro de texto possui as entradas de proteínas de acordo com anotação *UniProtKBAC*.

18

Figura 10 – Página de resultados da ferramenta *GORetriever* com seleção da opção “*GO Summary file (for use in GOSlim Viewer)*” que fornece um ficheiro de texto (MS-DOS) para utilização na ferramenta *GOSlim Viewer*.

19

Figura 11 – Seleção, na página inicial do *AgBase*, da ferramenta *GOSlim Viewer* no separador “*TOOLS*”.

20

Figura 12 – *Upload* do ficheiro proveniente de *GORetriever* e seleção da opção “*Generic*” em “*GOSlim Set*”.

20

Figura 13 – Página de resultados da ferramenta *GOSlimViewer* em que se fez a seleção da opção “*View accessions for each slim id*”.

21

Figura 14 – Página com apresentação detalhada das ontologias génicas das proteínas de interesse e seleção da opção “*View accessions for each slim id text only*” para obtenção do ficheiro em formato texto (MS-DOS).

21

Figura 15 – Primeiro passo para a importação dos dados presentes na ficheiro de texto (MS-DOS) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013* em que se escolheu a opção “Delimitado” para a organização dos dados.

22

Figura 16 – Segundo passo para a importação dos dados presentes na ficheiro de texto (MS-DOS) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013* optando-se por uma delimitação por “Tabulação”.

23

Figura 17 – Terceiro passo para a importação dos dados presentes na ficheiro de texto (MS-DOS) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013*.

23

Figura 18 – Quarto passo para a importação dos dados presentes na ficheiro de texto (*MS-DOS*) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013*. 24

Figura 19 – Porção da tabela obtida com os níveis mais elevados das ontologias génicas (*Parent*) das proteínas de interesse na coluna “*Slim_GO_Name*” e ontologias génicas derivadas (*Child*) na coluna “*Input_GO_Name*”.

Figura 20 – Porção da tabela filtrada em “*Slim_GO_Name*” para “*immune system process*” e “*cell death*” e com apresentação apenas das colunas “*Slim_GO_Name*” (ontologia *Parent*), “*Input_Accession*” (código UniProtKBAC da proteína) e “*Input_GO_Name*” (ontologia *Child* com função biológica específica da proteína). 26

Figura 21 – Porção da tabela dinâmica para “*immune system process*” com apresentação das ontologias *Child* / funções biológicas específicas (“*Input_GO_Name*”) organizadas por grupos e as proteínas correspondentes (“*Input_Accession*”). 27

Figura 22 – Porção da tabela com a ontologia “*immune system process*” na primeira célula e todas as proteínas dessa ontologia *Parent* na mesma linha. As ontologias *Child* / funções biológicas específicas não repetidas encontram-se imediatamente abaixo. 28

Figura 23 – Porção da tabela da ontologia “*immune system process*” e funções biológicas específicas / ontologias *Child* de si derivadas com todas as proteínas relacionadas imediatamente à frente na mesma linha. 29

Figura 24 – Do lado esquerdo visualiza-se porção da tabela derivada do OralOma em *Whole Saliva* proveniente de *Homo sapiens* filtrada de modo a mostrar apenas as colunas que possuem os atributos de interesse “*UniProtKBAC*”, “*Name*”, “*Regulation*” e “*Biomarker*” . Do lado direito, visualiza-se a porção da tabela com a listagem de todas as proteínas exclusivas das patologias em estudo e que permitiu aplicar o filtro avançado do *Microsoft Excel 2013* para essas entradas na tabela do lado esquerdo. 30

- Figura 25 – Do lado esquerdo visualiza-se a porção da tabela derivada do OralOma em *Whole Saliva* proveniente de *Homo sapiens* já filtrada para as proteínas exclusivas das patologias em estudo. Do lado direito visualiza-se a tabela dinâmica derivada da tabela do lado esquerdo com cálculo da média de “*Regulation*” das proteínas com diferentes valores de quantificação referenciados na literatura. 31
- Figura 26 – Porção da tabela com os atributos relevantes das proteínas exclusivas das patologias que servirão de base para a construção das tabelas funcionais de cada patologia. 32
- Figura 27 – Porção da tabela com funções biológicas específicas para proteínas exclusivas de *Breast Neoplasm (MeSH ID 68001943)*. Nesta tabela verifica-se o resultado da aplicação da fórmula supra-citada, estando já assinaladas as proteínas intervenientes em cada ontologia *Parent*, ou seja, “*immune system process*” e “*cell death*”. 33
- Figura 28 – Porção da tabela com os atributos referentes às proteínas exclusivas de *Breast Neoplasm (MeSH ID 68001943)* em “*immune system process*” e “*cell death*”. 34
- Figura 29 – Porção da tabela funcional inicial de *Breast Neoplasm (MeSH ID 68001943)* onde se visualizam as proteínas exclusivas desta patologia em “*immune system process*” e “*cell death*” associadas as suas funções biológicas específicas. Verifica-se também a inclusão dos cabeçalhos que foram obtidos ao longo desta metodologia. 35
- Figura 30 – Porção da tabela funcional final de *Breast Neoplasm (MeSH ID 68001943)* em “*immune system process*” e “*cell death*”. Verifica-se a formatação da tabela funcional e a organização das ontologias *Child* em grupos comuns referentes a cada uma destas ontologias / funções biológicas específicas. 35
- Figura 31 – Proteínas identificadas das patologias do OralOma em *Whole Saliva* proveniente de *Homo sapiens* possuem proteínas com valores de quantificação. 40

Figura 32 – Proteínas exclusivas das patologias do OralOma em *Whole Saliva* proveniente de *Homo sapiens* que possuem proteínas com valores de quantificação. 42

Figura 33 – Neste gráfico estão representadas as proteínas exclusivas em “*immune system process*” e “*cell death*” para cada patologia e respectivos biomarcadores sugeridos até ao momento. 44

Figura 34 – *Clusters* de patologias com proteínas exclusivas ligadas ao nó central (que representa a patologia). Os nós a azul são proteínas sem sugestão como biomarcador, enquanto que os nós a vermelho são as proteínas já sugeridas como biomarcador. 44

Índice de tabelas

Tabela 1 – Proteínas identificadas nas patologias selecionadas. 39

Tabela 2 – Proteínas exclusivas nas patologias selecionadas. 41

Tabela 3 – Tabela resumo com proteínas com sugestão para estudos de validação de biomarcadores para a patologia *Chronic Periodontitis* (MesH ID 68055113), *Breast Neoplasm* (MesH ID 68001943) e *Diabetes Mellitus type 2* (MesH ID 68003924). 56

Lista de abreviaturas

GO	Gene ontology
GTP	guanosine triphosphate
GTPase	guanosine triphosphate hydrolase
HDL	High-density lipoprotein
HIV	human immunodeficiency virus
Hsp10	Heat shock protein 10
IL-1F6	Interleukin-1 family member 6
IL-1β	Interleukin-1 beta
IRS-1	Insulin receptor substrate 1

LCP-1	Lymphocyte cytosolic protein 1
LDL	Low-density lipoprotein
MAC	membrane attack complex
MCP-1	Monocyte chemotactic protein 1
MesH	Medical subjects headings
MHC	histocompatibility complex
MIP-1-alpha	Macrophage inflammatory protein 1-alpha
PCB	polychlorinated biphenyl
QR1	Quinone reductase 1
RNA	Ribonucleic acid
TNF-α	Tumor necrosis factor alpha
UniProtKBAC	Universal Protein Knowledgebase Accession Code
VLDL	Very low-density lipoprotein

Resumo

A detecção de biomarcadores na saliva, proteínas que estão associadas a fenótipos de patologias específicas, apresenta-se como uma área de intervenção diagnóstica pouco estudada na atualidade.

Foi objetivo deste estudo de acordo com os dados de proteômica/ estudos dirigidos depositados no *OralCard* referentes a patologias orais e sistêmicas:

- Realizar uma análise crítica dos biomarcadores até agora propostos;
- Identificar que proteínas podem vir a constituir potenciais novos biomarcadores.

A organização das proteínas exclusivas e identificadas como estando presentes e/ou alteradas em doenças orais e sistêmicas de acordo com a ontologia “*immune system process*” e “*cell death*” foram submetidas a um estudo de acordo com as suas funções biológicas específicas nestas ontologias, realizando-se uma análise tanto aos biomarcadores sugeridos como a novas propostas de biomarcadores.

Foram sugeridas neste estudo proteínas que apresentam valores de quantificação em saliva e que são passíveis de serem integradas em estudos para validação de biomarcadores, pois são proteínas com significado funcional em patologias específicas e que contribuem para processos moleculares alterados subjacentes a cada patologia. Foram também sugeridos novos estudos de quantificação para biomarcadores já sugeridos e com falta de dados de quantificação.

Palavras-chave: Biomarcador, proteoma, *OralCard*, saliva, análise funcional.

Abstract

The detection of biomarkers in saliva, proteins that are associated with specific phenotypes of pathologies, presents itself as an area of diagnostic intervention without much studies.

Using the proteomic data / directed studies deposited in *OralCard* on oral and systemic diseases the primary goals of this study were::

- To conduct a critical analysis of biomarkers already proposed;
- To identify proteins that may prove to be potential new biomarkers.

The organization of the proteins exclusively identified as being present and / or altered in oral and systemic diseases according to the ontologies “immune system process” and “cell death” were submitted to the study of their specific biological functions in these ontologies, in order to make an analysis of both previously suggested biomarkers and biomarkers to be proposed.

This study allowed to suggest proteins which contain regulation values in *Whole Saliva* and which are likely to become part of biomarker validation's studies, because they have functional meaning in specific pathologies and contribute to a better characterization of the underlying molecular processes for each pathology. This study also allowed the suggestion of more regulation studies for the previously suggested biomarkers without regulation data.

Key-words: Biomarker, proteome, *OralCard*, saliva, functional analysis.

1. Introdução e enquadramento teórico

1.1. A cavidade oral e o proteoma salivar

A cavidade oral, como ecossistema complexo que é, combina vários fatores que se encontram interligados, sendo que nestes está integrado todo o universo de proteínas presentes no conteúdo salivar. A compreensão da biologia oral é dependente da integração dos estudos de proteómica numa base de dados e sua consequente interpretação. A inexistência de tal ferramenta suscitou a necessidade de criar a base de dados OralOma, que compila as proteínas referidas em estudos de proteómica salivar em indivíduos saudáveis e indivíduos com patologias orais e/ou sistémicas. Sendo que, na última década o desenvolvimento das mais variadas técnicas de proteómica permitiu o desenvolvimento exponencial do repositório de estudos que envolvem identificação de proteínas em amostras salivares, tanto em amostras salivares de indivíduos saudáveis como em amostras de indivíduos representativos de determinadas condições patológicas orais e/ou sistémicas (2).

Tendo em conta a necessidade de facilitar à comunidade científica o acesso a estudos de proteómica na saliva, uma tarefa que se torna bastante morosa devido à falta de ferramentas que compilem todas as informações de interesse relativamente a este tipo de estudos na saliva, tornou-se imperativa a necessidade de converter a base de dados OralOma numa plataforma que permitisse o acesso à informação de proteómica contida na base de dados, permitindo oferecer à comunidade científica uma ferramenta base para o entendimento dos processos fisiológicos e patológicos da cavidade oral, tornando-os alvos de investigação para a identificação de biomarcadores tanto para doenças orais como sistémicas. Assim surgiu o *OralCard*, uma ferramenta *online* que permite o acesso às informações contidas na base de dados OralOma (3).

1.2. A saliva como secreção ativa glandular

A saliva é constituída ativamente pela junção das secreções das glândulas salivar *major* (parótida, submandibular e sublingual) e das glândulas salivar *minor* distribuídas em grande número na mucosa oral. Porém, a saliva é também constituída por constituintes do fluído crevicular, produtos de atividade bacteriana e células descamadas do epitélio oral. A secreção salivar por parte das glândulas salivares é altamente variável e dependente de fatores como a sua estimulação, grau de estimulação e até mesmo dependente da altura do dia em que se dá secreção, se de dia ou de noite (4).

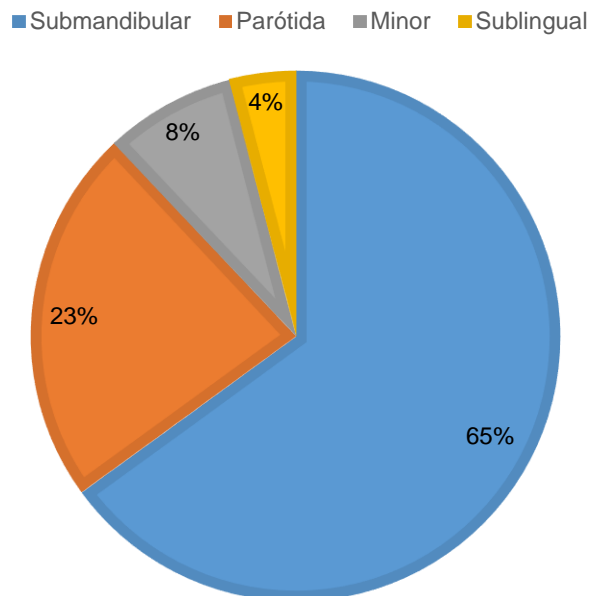


Figura 1 – Gráfico ilustrativo da contribuição das diferentes glândulas em termos de proteínas presentes na saliva e provenientes de secreção glandular salivar adaptado de (5).

É também importante referir que as proteínas salivares estão representadas por quatro grandes grupos de proteínas que são as proteínas ricas em prolina, as estaterinas, as cistatinas e as histatinas. (Figura 2) (6).

Proteínas presentes na saliva

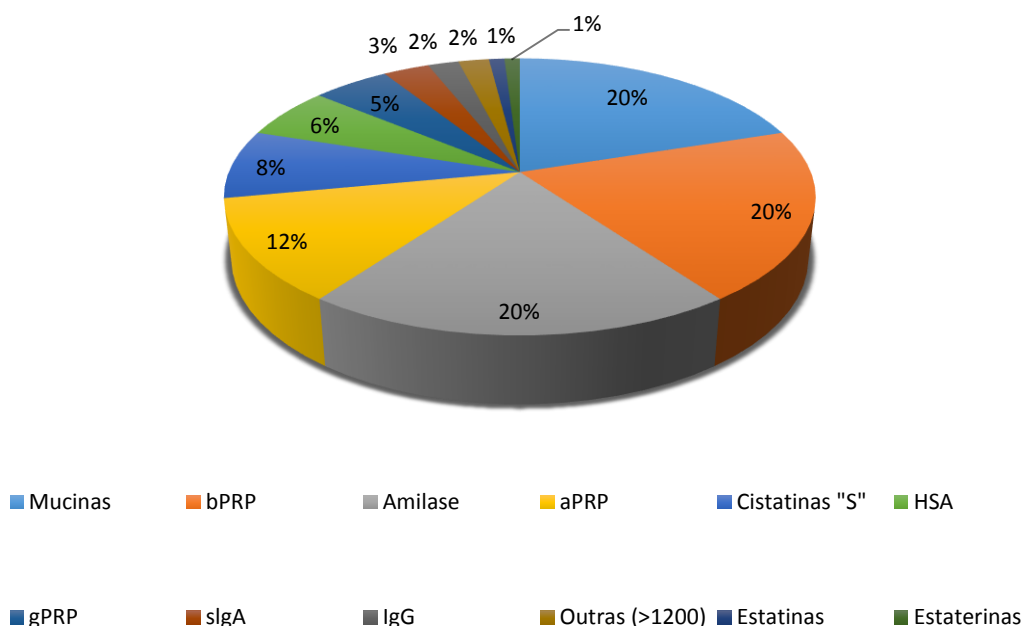


Figura 2 – Gráfico ilustrativo dos grupos proteicos presentes na saliva de Homo sapiens adaptado de (6).

A produção diária de saliva ronda os 500 a 1500mL, sendo que diversas condições fisiopatológicas podem modificar a produção de saliva tanto quantitativamente como qualitativamente. As glândulas produtoras de saliva, constituídas por células epiteliais especializadas secretoras (células acinares), que formam conjuntos denominados ácinos, produzem um fluido que contém água, eletrólitos, secreção mucosa e proteínas (5, 7).

Cerca de 80% das proteínas secretadas para a saliva pelas glândulas salivares têm origem nas células acinares sendo que 20% têm origem na secreção proveniente das células dos ductos. Porém, a saliva que é recolhida do meio oral, designada de *Whole Saliva*, não é apenas constituída pela soma do conteúdo das secreções das glândulas salivares mas também por outros constituintes do ambiente oral que se juntam a essas secreções como já foi referido, como células resultantes da descamação epitelial oral, produtos microbianos ou outras proteínas que também estão presentes no sangue e que por difusão, filtração ou transporte ativo encontraram o seu caminho até chegar à saliva possibilitando assim a sua recolha através da saliva. No entanto há

que referir que das proteínas que possam estar presentes na secreção salivar de uma glândula, não sejam encontradas na *Whole Saliva* devido a, por exemplo, atividade enzimática de degradação dessas mesmas proteínas. Pode-se então dizer que a recolha de saliva como fluido de diagnóstico se pode dar tanto ao nível das glândulas de um modo individual ou através da recolha da saliva presente no meio oral, ou *Whole Saliva* (8, 9).

Em suma, pode-se afirmar que as secreções das glândulas salivares major perfazem cerca de 90% das proteínas endógenas, sendo que os restantes 10% de proteínas endógenas correspondem à secreção das glândulas salivares minor, fluido crevicular, língua e mucosa oral. Ao aliar as proteínas endógenas às proteínas exógenas (produzidas pelos microrganismos em meio oral), obtemos o universo total de proteínas presentes em saliva ou *Whole Saliva* (3).

1.3. A função fisiológica salivar

O papel básico, conhecido, da saliva consiste tanto na proteção como na manutenção da porção superior do trato digestivo, através de funções como a da lubrificação, mecanismo de tampão, atividade antibacteriana e antiviral, manutenção da integridade da mucosa e peças dentárias assim como papel no paladar e digestão (6, 10). No entanto, pouco se sabe acerca do papel que desempenham os muitos milhares de proteínas humanas e microbianas que nela podemos encontrar.

1.4. A saliva como fluido de diagnóstico

A saliva como meio de diagnóstico toma particular importância tendo em conta a facilidade e rapidez na colheita para diagnóstico, o seu carácter não-invasivo e a presença de inúmeras proteínas que podem ter valor em diagnóstico (5, 11). O risco de contração de infeções através da colheita salivar é mínimo, sendo que este tipo de colheita apresenta vantagens em situações clínicas particulares verificadas em crianças pouco colaborantes ou pacientes inimputáveis, demonstrando-se uma alternativa viável à colheita sanguínea, tendo em conta que grande parte das proteínas encontradas no sangue,

através de diferentes meios de transporte, encontram passagem para o meio salivar (5, 12).

1.5. A importância de diagnósticos salivares dirigidos

No entanto, tendo em conta a grande variabilidade de processos de degradação e modificação de proteínas que pode ocorrer no meio oral, temos a limitação de utilizar a saliva como meio de diagnóstico direcionado a perguntas concretas, como por exemplo, a presença de uma determinada proteína que nos possa remeter para uma determinada patologia, ao invés da devolução do estado geral de saúde do indivíduo (8). Deste modo, torna-se relevante focar a atenção para proteínas que se encontram na saliva e que podem estar associadas a fenótipos de patologia, sendo que o estudo dessas mesmas proteínas se demonstra relevante para a determinação do seu potencial no diagnóstico específico da patologia (11).

1.5.1. Diagnóstico salivar em doenças locais e sistémicas

Historicamente, o diagnóstico de patologias realiza-se a partir da revisão da história clínica, dos sintomas relatados pelo paciente, através do exame clínico, com recurso a exames auxiliares de diagnóstico tais como recolhas de sangue ou urina, de modo a comprovar o estabelecimento de patologia. Atualmente, e de um modo geral apenas se recorre à recolha de amostras orais caso haja evidência de patologia oral, no entanto cada vez mais se torna relevante a recolha de amostras orais, nomeadamente a saliva que tem vindo a ser citada na bibliografia como um método de diagnóstico fiável para patologias sistémicas (13).

1.6. Biomarcadores na saliva

Um grande número de biomarcadores pode ser identificado na saliva, desde metais pesados a proteínas (14). Sendo que, ultimamente a análise de “proteínas chave” ou biomarcadores proteicos que permitam o diagnóstico de determinadas patologias tem vindo a despertar cada vez mais a atenção da comunidade científica. O estudo da proteómica salivar tem a potencialidade de

permitir a identificação de proteínas específicas para a doença sendo que a caracterização destas proteínas do ponto de vista fisiopatológico permite obter informação sobre o estabelecimento, progressão e prognóstico da patologia (11).

1.7. Resposta imune e proteínas salivares

A resposta imune, um processo biológico inerente a patologias tão diversas como infecções, traumas, doenças do foro asmático, cancro, diabetes, periodontites, síndrome de Sjögren, todas comportam mecanismos da resposta imunológica. Esta resposta imune ao *stress* biológico desencadeia inúmeras cascatas de eventos mediadas por um variado número de células específicas da resposta imune, desde macrófagos, dendríticas, linfócitos T, B e células *natural killer*.

Envolvidas na resposta imune estão também identificadas moléculas características designadas por citocinas. A associação entre as células características da resposta imune aliadas às proteínas de resposta imune confere ao processo biológico da resposta imunológica a capacidade de identificar possíveis patogénicos, tecido sob *stress* ou danificado e recrutar por sua vez outras células e moléculas que eliminam os agentes agressores e possibilitam o restabelecimento do equilíbrio do organismo (15).

É importante referir que a ação tanto da resposta imune como das suas proteínas, despoleta inúmeros processos na fisiopatologia das doenças que induzem alterações nas quantidades das proteínas encontradas em saúde, pelo que as alterações dessas proteínas em doença são passíveis de ser identificadas e quantificadas em saliva, pois tendo em conta e como já foi referido a capacidade destas proteínas conseguirem atingir o meio salivar.

1.8. Biologia de sistemas e proteínas salivares

A biologia de sistemas é definida como uma abordagem global de dados biológicos nas suas diversas escalas de organização de modo a caracterizar sistemas ou processos biológicos específicos no qual se insere a resposta

imune, com o objetivo primordial de encontrar assinaturas moleculares que possibilitem o diagnóstico (16).

1.9. Ferramentas bioinformáticas

As ferramentas bioinformáticas são instrumentos que permitem a análise integrada de dados produzidos em grande escala e que se encontram armazenados em múltiplas bases de dados. A base concetual para o entendimento da biologia do sistema pressupõe a capacidade instrumental de integrar toda a informação existente por forma a ser possível retirar significado biológico a partir da evidência existente.

O *OralCard*, como já foi referido, é uma ferramenta bioinformática que possibilita aceder aos estudos de proteómica e de estudos dirigidos que foram selecionados da literatura e depositados numa base de dados restrita a estudos na saliva, o OralOma (2, 3).

Os dados obtidos através do *OralCard* podem ser então utilizados na ferramenta bioinformática *AgBase* que disponibiliza vários recursos que facilitam a manipulação e gestão de dados de genómica e de anotações estruturais e funcionais de genomas de diversos animais, plantas, bactérias e parasitas (17).

2. Objetivos

Através do estudo *in silico* dos dados de proteômica e de estudos dirigidos que atualmente estão disponíveis no *OralCard* este estudo pretende alcançar os seguintes objetivos:

- Analisar e selecionar as patologias presentes no *OralCard* com proteínas quantificadas e identificar as proteínas destas patologias como estando presentes e alteradas na saliva;

- Analisar e selecionar as proteínas identificadas das patologias como estando presentes apenas numa patologia, ou seja, selecionar as proteínas exclusivas de cada patologia;

- Classificar as proteínas exclusivas de acordo com o código de ontologias selecionando apenas as proteínas exclusivas presentes em ontologias de interesse;

- Rever a literatura referente às proteínas selecionadas de modo a se proceder à caracterização funcional dessas mesmas proteínas nas ontologias de interesse;

- Revisão e proposta das proteínas que de acordo com a sua função e os dados de quantificação na patologia, melhor cumprem os requisitos para estudos de validação como biomarcadores salivares.

3. Materiais e métodos

No sentido de melhor elucidar a metodologia seguida para a consulta e integração da informação depositada no *OralCard* é apresentada no fluxograma da figura 3 a estratégia seguida (Figura 3):

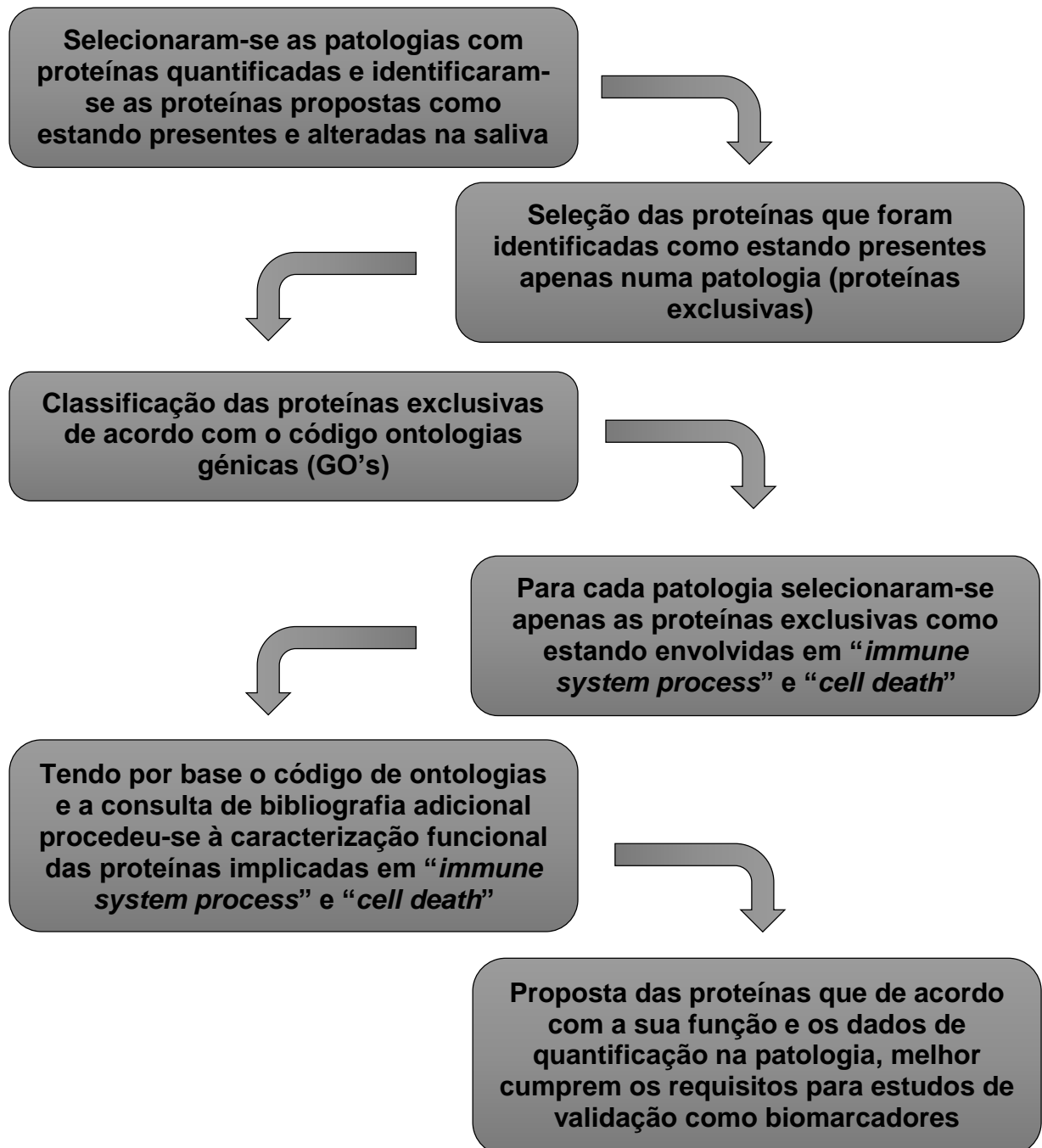


Figura 3 – Fluxograma da metodologia seguida neste trabalho.

3.1. Catalogação das Proteínas Identificadas em cada patologia

A base de dados OralOma (2) atualizada a 15 de Dezembro de 2013, permitiu a identificação das proteínas presentes em cada patologia. Para consulta da base de dados recorreu-se à ferramenta bioinformática *OralCard* (<http://bioinformatics.ua.pt/OralCard/>) (3).

3.1.1. Proteínas identificadas recolhidas em *Whole Saliva*

Uma vez que o interesse estava apenas nas proteínas identificadas na saliva utilizou-se o *software* informático *Microsoft Office Excel* 2013 que por aplicação do filtro para proteínas provenientes de *Whole Saliva* permitiu listar todas as proteínas já identificadas na saliva (Figura 4).

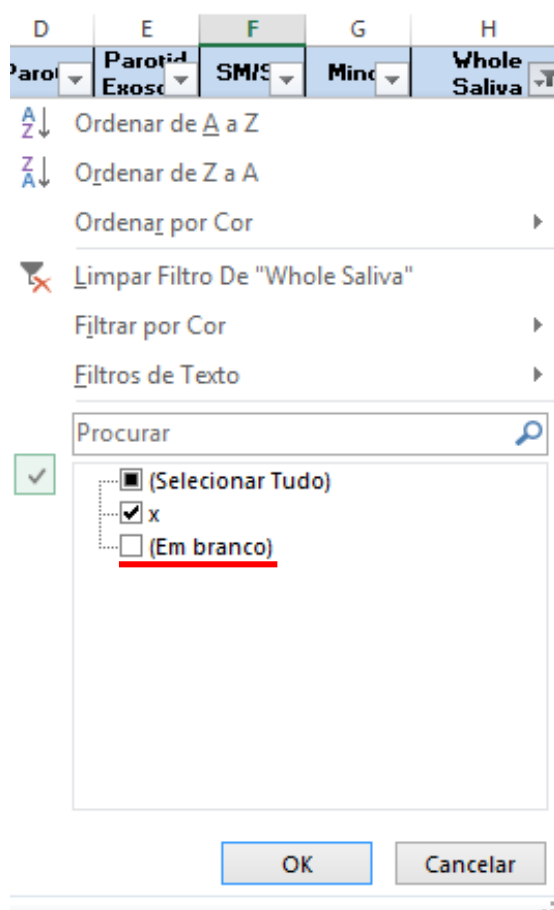


Figura 4 – Metodologia para seleção das proteínas com recolha na “*Whole Saliva*”, foi desmarcado o visto na opção “(Em branco)” de modo a serem mostradas apenas as patologias com recolha na saliva.

3.1.2. Seleção das proteínas de *Homo sapiens*

Uma vez que o *OralCard* também contém a informação correspondente às proteínas do microbioma oral procedeu-se à seleção por filtragem das proteínas provenientes da espécie *Homo sapiens* (Figura 5).

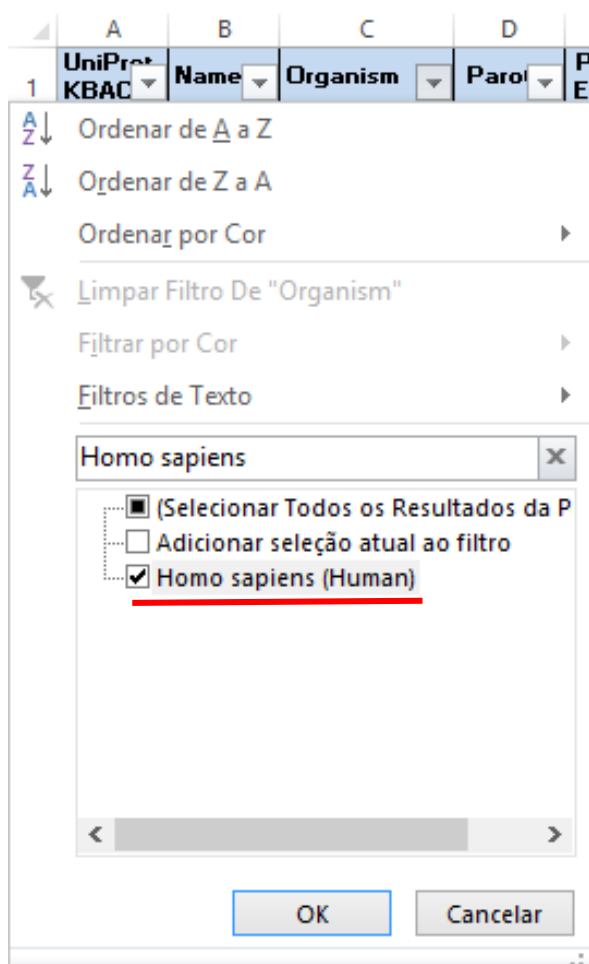


Figura 5 – Metodologia para seleção das proteínas com recolha na “*Whole Saliva*” na espécie *Homo sapiens* removendo-se a seleção de todas as espécies restantes.

3.1.3. Seleção das patologias com dados de quantificação

Como o interesse era selecionar as patologias que possuem proteínas com dados de quantificação de modo a estudar as respetivas proteínas que estão identificadas no *OralCard*, foram selecionadas para estudo apenas as patologias, com base no seu código MesH que cumprissem o requisito mencionado (Figura 6).

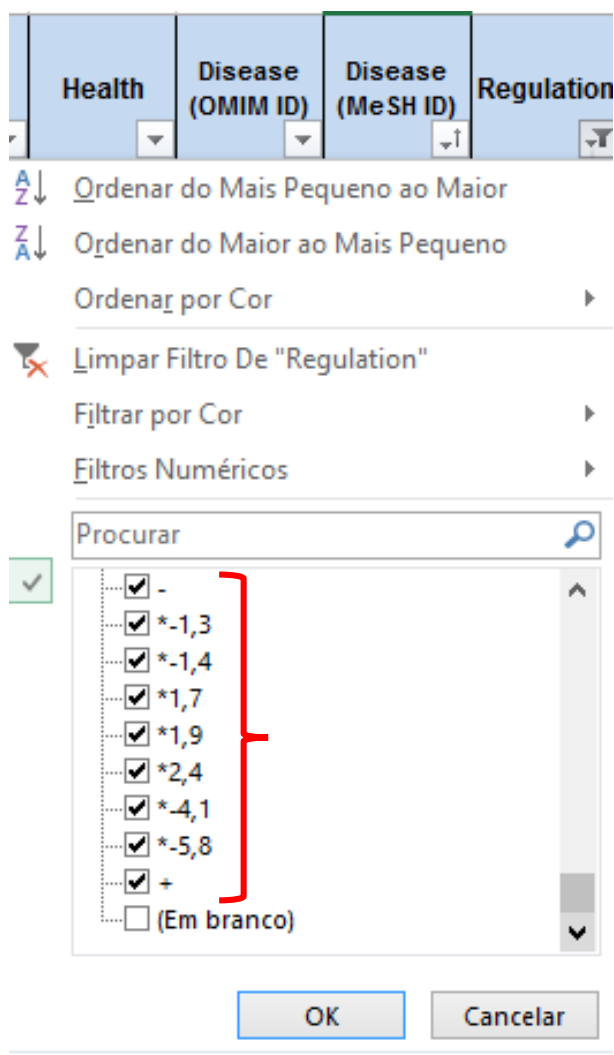


Figura 6 – Metodologia para selecionar as patologias que apresentam proteínas com quantificação. Nesta imagem apenas está representada uma porção das quantificações que na realidade foram selecionadas. Remoção de “(Em branco)” de modo a remover deste estudo patologias que não possuem proteínas com dados de quantificação.

3.2. Proteínas Exclusivas de cada patologia

De modo a ser possível selecionar as proteínas a serem discutidas em cada patologia, procedeu-se à seleção das proteínas exclusivas para cada patologia, ou seja, para cada patologia foram selecionadas apenas as proteínas que não foram identificadas noutras patologias, de acordo com a base de dados OralOma. Listaram-se os códigos UniProtKBAC (18) das proteínas identificadas em cada patologia.

3.2.1. Representação gráfica das proteínas exclusivas de cada patologia

Com o objetivo de facilitar a seleção e a importação da lista dos códigos das proteínas com interesse (proteínas exclusivas) recorreu-se à ferramenta Venny (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>) (19).

Esta ferramenta permitiu gerar gráficos circulares de Venn usando os códigos UniProtKBAC correspondente às proteínas exclusivas de cada patologia. Como resultado foi possível agrupar todas as proteínas da patologia numa caixa e as proteínas de todo o OralOma identificadas em *Whole Saliva* proveniente de *Homo sapiens* noutra caixa. É importante referir que ao OralOma introduzido na segunda caixa foram retiradas as proteínas da patologia que estão na primeira caixa e todas as proteínas que estão associadas apenas a saúde, ou seja, que não têm associação a doença. Desta forma, o gráfico de Venn permitiu o cruzamento dos códigos de todas as proteínas de todas as patologias descritas no OralOma com a doença em estudo. A figura 7 demonstra a construção deste tipo de gráficos para a patologia *Breast Neoplasm (MeSH ID 68001943)*. Idêntico procedimento foi aplicado a todas as patologias.

Após a identificação das proteínas exclusivas de cada patologia foi possível construir tabelas resumo do OralOma para cada uma delas. Estas tabelas são extratos menores do OralOma em *Whole Saliva* proveniente de *Homo sapiens*, obtido anteriormente. Estas tabelas individualizadas e com as proteínas exclusivas de cada patologia possibilitaram tanto uma leitura como um acesso facilitado aos atributos de cada proteína nas patologias em estudo.

A união destas tabelas possibilitou catalogar todas as proteínas exclusivas de cada patologia e que possuem proteínas com quantificação de acordo com o OralOma em *Whole Saliva* proveniente de *Homo sapiens*.

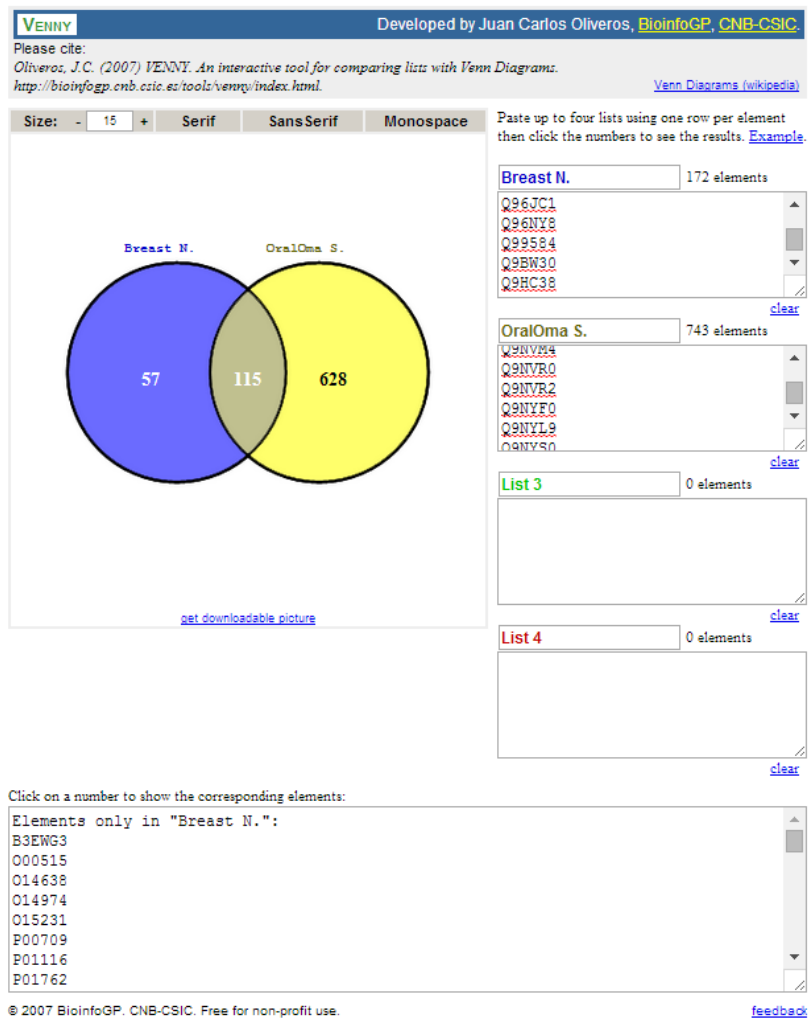


Figura 7 – Metodologia para selecionar as proteínas exclusivas em *Breast Neoplasm* (MesH ID 68001943). É possível verificar a presença de dois gráficos circulares cruzados, sendo que o gráfico a azul possui a contagem de proteínas exclusivas de *Breast Neoplasm* (MesH ID 68001943) e o gráfico a amarelo a contagem de proteínas no OralOma em *Whole Saliva* proveniente de *Homo sapiens*. No cruzamento dos dois gráficos visualiza-se a contagem de proteínas em comum aos dois grupos. Na caixa inferior encontra-se a lista de proteínas exclusivas de *Breast Neoplasm* (MesH ID 68001943).

3.3. Classificação das proteínas por Ontologias Génicas (GO's)

De modo a organizar as proteínas de cada patologia com vista à sua análise funcional, foi necessário obter as ontologias génicas de cada uma das proteínas exclusivas. Posteriormente foram escolhidas as ontologias génicas de processos biológicos que descrevem funções específicas para cada patologia.

3.3.1. Ontologias génicas derivadas do *AgBase*

O recurso à ferramenta *AgBase* (<http://www.AgBase.msstate.edu/>) (17, 20, 21) que nos permite obter as ontologias génicas *Parent* das proteínas e respetivas ontologias derivadas, designadas por ontologias *Child*, que dependendo do tipo de ontologia *Parent* da qual derivam, podem ser ontologias de funções biológicas em processos biológicos específicos, sendo que a sua recolha facilita a análise funcional e consequente interpretação funcional de cada proteína (22). Todas as proteínas exclusivas anteriormente selecionadas foram juntas numa única coluna de um livro do *Microsoft Excel*. Este livro foi então guardado como formato de Texto (*MS-DOS*) o que possibilita a sua utilização na ferramenta *AgBase*.

3.3.1.1. Ferramenta *GORetriever*

Aquando da entrada no sítio da internet do *AgBase*, selecionou-se a opção *GORetriever* no separador *TOOLS*, sendo que esta ferramenta nos permite recolher todas as anotações de ontologias génicas associadas às proteínas a serem introduzidas (Figura 8). Dentro desta ferramenta selecionou-se no campo “*Select Accession Type*” a opção “*UniProt Acession*”, tendo em conta que o ficheiro com formato de Texto (*MS-DOS*) possui as entradas de proteínas de acordo com anotação *UniProtKBAC* (18), e fez-se o *upload* do ficheiro de Texto (*MS-DOS*) para a ferramenta o analisar (Figura 9).

Assim que a ferramenta *GORetriever* processou as proteínas introduzidas, selecionou-se a opção “*GO Summary file (for use in GOSlim Viewer)*” que forneceu um ficheiro de Texto no formato (*MS-DOS*) com as anotações das ontologias génicas associadas a cada uma das proteínas introduzidas para posterior utilização na ferramenta *GOSlim Viewer* (Figura 10).



Figura 8 – Seleção da ferramenta “GORetriever” na página inicial do AgBase no separador “TOOLS”.

GORetriever is used to rapidly retrieve existing GO annotations for a list of proteins. The input file is a text file containing protein accessions separated by a return/line break. Note: Help with Affymetrix probe mapping is available [here](#).

For help using this tool click [here](#).
An online tutorial for this tool is available [here](#).
More information about the tool is [here](#).

Select Accession Type*:

File to Upload**
(Please upload a plain text file): Nenhum ficheiro selecionado

Type of Evidence to Return:

- Check all/Uncheck all
- Experimental Evidence Codes (EXP,IDA,IPI,IMP,IGI,IEP)
- Computational Analysis Evidence Codes (IGC,IBA,IBD,IKR,IRD,ISS,ISO,ISA,ISM,RCA)
- Author Statement Evidence Codes (TAS,NAS)
- Curator Statement Evidence Codes (IC,ND)
- Automatically-assigned Evidence Codes (IEA)
- Obsolete Evidence Codes (NR)
 - EXP IBA IBD IC IDA
 - IEA IEP IGC IGI IKR
 - IMP IPI IRD ISA ISM
 - ISO ISS NAS ND NR
 - RCA TAS

Figura 9 – “Accession Type” selecionado como “UniProt Accession” e seleção do botão “Escolher ficheiro” de modo a fazer upload do ficheiro de texto possui as entradas de proteínas de acordo com anotação UniProtKBAC.

AgBase [Version: 2.00]

MISSISSIPPI STATE UNIVERSITY

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

HOME | SEARCHES | TOOLS | ANIMALS | PLANTS | MICROBES | PARASITES | HELP | CONTACT

GORetriever Results

Number of Input IDs: 563

Input filename: UniProts_ALL_Ex.txt

Evidence Codes to return: IC,IDA,IEA,IEP,IGI,IMP,IPI,ISS,NAS,ND,RCA,TAS,NR,EXP,ISO,ISA,ISM,IGC,IBA,IBD,IKR,IMR,IRD

[View GO annotation file in tab-delimited format \(can be opened in Excel\)](#)

[View ID List without GO annotation](#)

[GO Summary file \(for use in GOSlim Viewer\)](#)

GORetriever Results

Number of Input IDs: 563

Input filename: UniProts_ALL_Ex.txt

Evidence Codes to return: IC,IDA,IEA,IEP,IGI,IMP,IPI,ISS,NAS,ND,RCA,TAS,NR,EXP,ISO,ISA,ISM,IGC,IBA,IBD,IKR,IMR,IRD

[View GO annotation file in tab-delimited format \(can be opened in Excel\)](#)

[View ID List without GO annotation](#)

[GO Summary file \(for use in GOSlim Viewer\)](#)

Figura 10 – Página de resultados da ferramenta *GORetriever* com seleção da opção “*GO Summary file (for use in GOSlim Viewer)*” que fornece um ficheiro de texto (*MS-DOS*) para utilização na ferramenta *GOSlim Viewer*.

3.3.1.2. Ferramenta *GOSlim Viewer*

Após se ter guardado o respetivo ficheiro de texto (*MS-DOS*), recorreu-se então à ferramenta *GOSlim Viewer*, ferramenta esta que permite obter um resumo de nível elevado (*GOSlim*) das funções das proteínas de interesse, funções estas que são retiradas a partir das anotações das ontologias génicas de cada proteína provenientes da ferramenta *GORetriever*. Para aceder a esta ferramenta, é necessário voltar à página inicial do sítio do *AgBase* e no separador “*TOOLS*” escolher a opção “*GOSlim Viewer*” (Figura 11). Já na página da ferramenta deve ser feito o *upload* do ficheiro de texto proveniente da ferramenta *GORetriever* no campo “*Upload GO Summary File*” enquanto no campo “*Select GOSlim Set*” se deve colocar a opção “*Generic*” (Figura 12). Na página de resultados deve-se selecionar a opção “*View accessions for each slim id*” (Figura 13). Esta opção reencaminha para uma nova página que apresenta em detalhe todas as ontologias génicas das proteínas de interesse e respetivas ontologias *Child* (nas quais também se inserem funções específicas das proteínas), sendo então selecionada a opção “*View accessions for each slim id text only*” (Figura 14) que permite o *download* de um ficheiro de texto (*MS-DOS*) que contém a mesma informação da página.



Figura 11 – Seleção, na página inicial do AgBase, da ferramenta GOSlim Viewer no separador “TOOLS”.

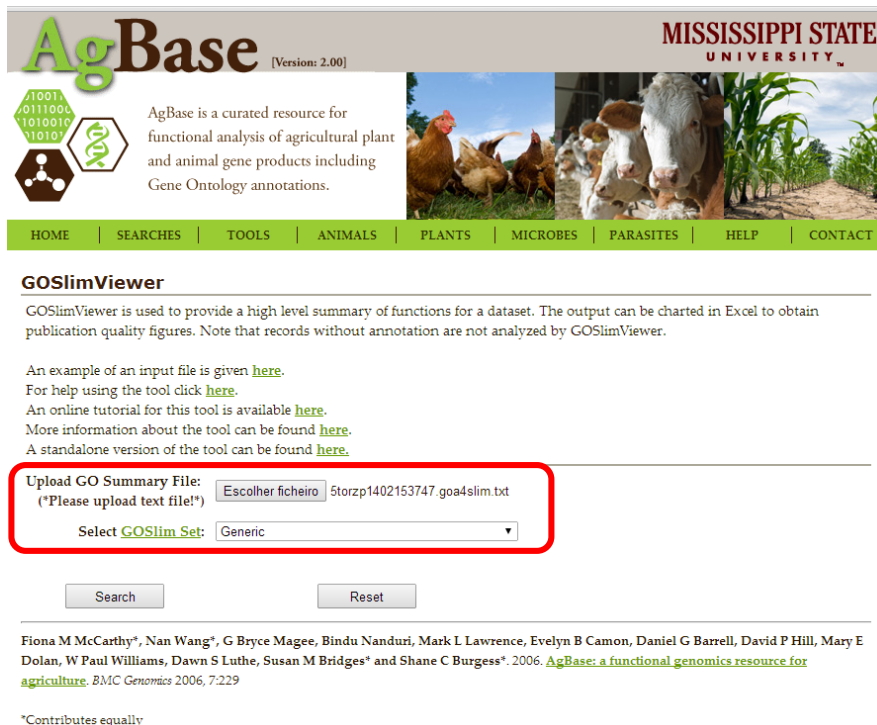


Figura 12 – Upload do ficheiro proveniente de GORetriver e seleção da opção “Generic” em “GOSlim Set”.

AgBase [Version: 2.00]

MISSISSIPPI STATE UNIVERSITY

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

HOME | SEARCHES | TOOLS | ANIMALS | PLANTS | MICROBES | PARASITES | HELP | CONTACT

GOSlimViewer Results

To download, right click a file and select Save As

To view the detail for cellular components, molecular components, and biological processes click the View accessions for each slim id link

[View Cellular Component](#)

[View Molecular Function](#)

[View Biological Process](#)

[View accessions for each slim id](#)

Figura 13 – Página de resultados da ferramenta *GOSlimViewer* em que se fez a seleção da opção “View accessions for each slim id”.

GOSlimViewer accession detail for each slim id

[View accessions for each slim id text only](#)

Index

1 [Cellular Component Accessions](#)

2 [Molecular Function Accessions](#)

3 [Biological Process Accessions](#)

Cellular Component Accessions [Back to top](#)

Slim id:	GO:0000228	nuclear chromosome	8
O75367	GO:0001740		Barr body
Q13185	GO:0005720		nuclear heterochromatin
O75367	GO:0000790		nuclear chromatin
P07305	GO:0000790		nuclear chromatin
P60709	GO:0000790		nuclear chromatin
P61978	GO:0000790		nuclear chromatin
O75367	GO:0001739		sex chromatin
P07305	GO:0005719		nuclear euchromatin
P16403	GO:0005719		nuclear euchromatin
Q13185	GO:0005719		nuclear euchromatin
P13010	GO:0000784		nuclear chromosome, telomeric region
P13010	GO:0000783		nuclear telomere cap complex
Q9H160	GO:0016580		Sin3 complex
Q13185	GO:0031618		nuclear centromeric heterochromatin

Figura 14 – Página com apresentação detalhada das ontologias gênicas das proteínas de interesse e seleção da opção “View accessions for each slim id text only” para obtenção do ficheiro em formato texto (*MS-DOS*).

3.4. Caracterização funcional das proteínas de cada patologia

Com a obtenção do ficheiro proveniente dos resultados da ferramenta *GOSlimViewer* procedeu-se à metodologia que permite a construção das tabelas funcionais de cada patologia, tabelas estas que permitem a análise de funções específicas, e inseridas em ontologias de interesse, de cada proteína alterada na patologia respetiva.

3.4.1. Importação de ontologias a partir de *GOSlimViewer*

Inicialmente procedeu-se à importação dos dados presentes no ficheiro de texto (*MS-DOS*) proveniente da ferramenta *GOSlimViewer* para uma folha num livro do *Microsoft Excel 2013*, obtendo-se assim uma tabela detalhada com os níveis mais elevados das ontologias génicas das proteínas de interesse e respetivas ontologias *Child*, que como já foi referido anteriormente, podem ser de funções específicas das proteínas dependendo do tipo de ontologia *Parent* que derivam (Figura 15-19).

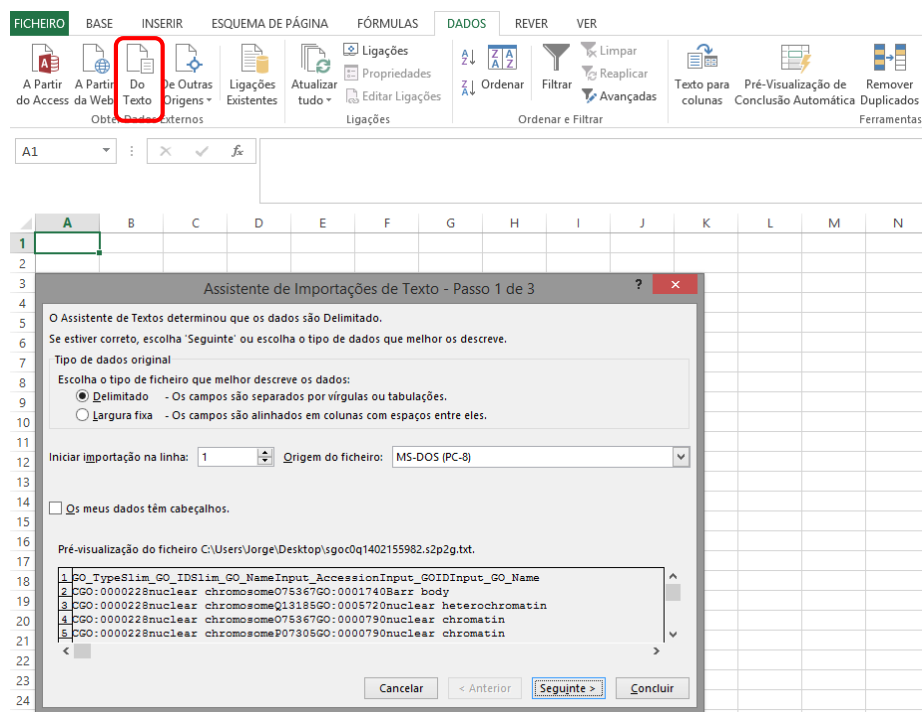


Figura 15 – Primeiro passo para a importação dos dados presentes na ficheiro de texto (*MS-DOS*) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013* em que se escolheu a opção “Delimitado” para a organização dos dados.

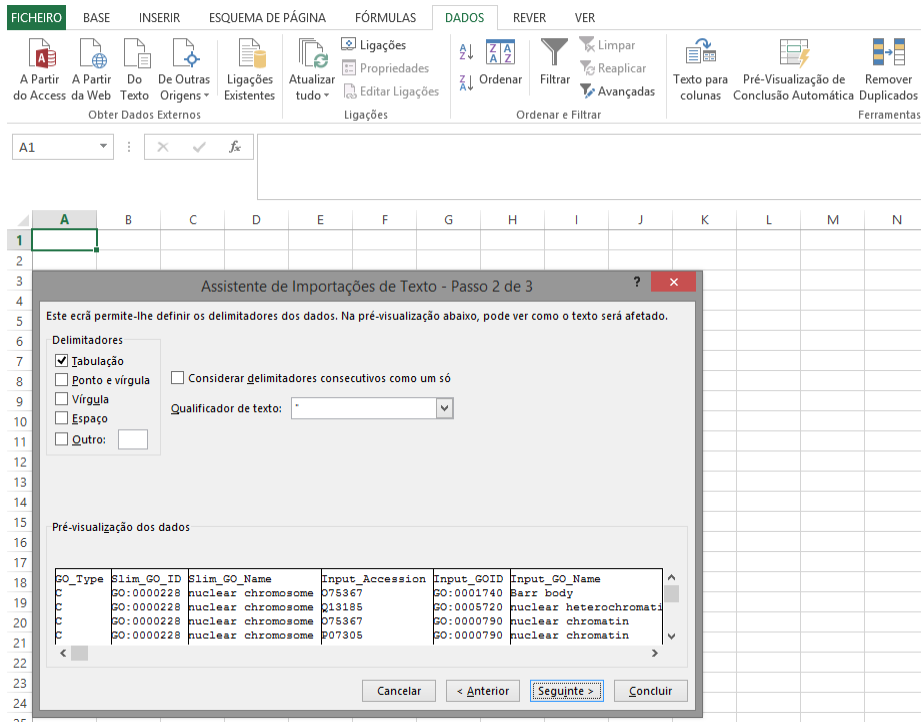


Figura 16 – Segundo passo para a importação dos dados presentes na ficheiro de texto (*MS-DOS*) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013* optando-se por uma delimitação por “Tabulação”.

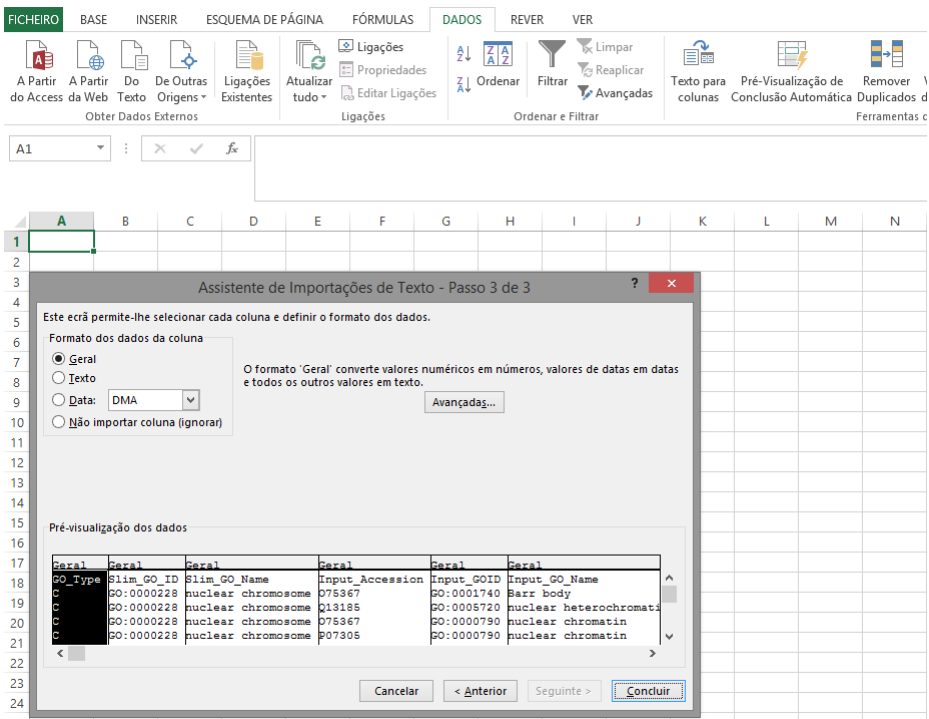


Figura 17 – Terceiro passo para a importação dos dados presentes na ficheiro de texto (*MS-DOS*) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013*.

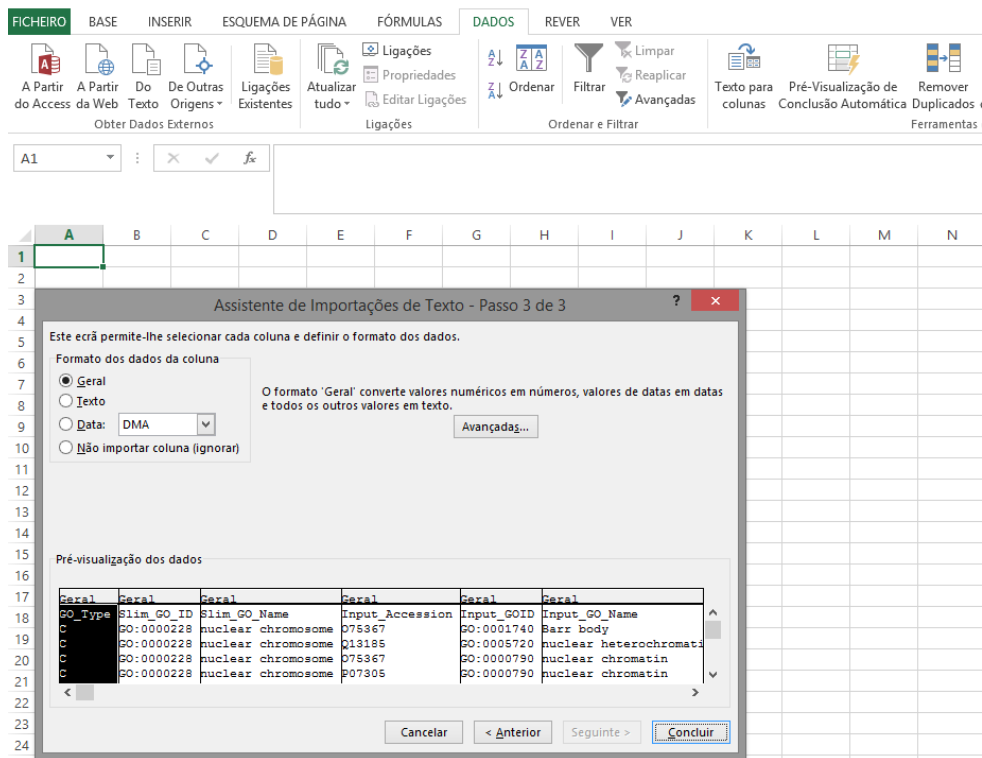


Figura 18 – Quarto passo para a importação dos dados presentes na ficheiro de texto (*MS-DOS*) proveniente da ferramenta *GOSlimViewer* para livro de *Microsoft Excel 2013*.

GO_Type	Slim_GO_ID	Slim_GO_Name	Input_Accession	Input_GOID	Input_GO_Name
C	GO:000228	nuclear chromosome	O75367	GO:0001740	Barr body
C	GO:000228	nuclear chromosome	Q13185	GO:0005720	nuclear heterochromatin
C	GO:000228	nuclear chromosome	O75367	GO:0000790	nuclear chromatin
C	GO:000228	nuclear chromosome	P07305	GO:0000790	nuclear chromatin
C	GO:000228	nuclear chromosome	P60709	GO:0000790	nuclear chromatin
C	GO:000228	nuclear chromosome	P61978	GO:0000790	nuclear chromatin
C	GO:000228	nuclear chromosome	O75367	GO:0001739	sex chromatin
C	GO:000228	nuclear chromosome	P07305	GO:0005719	nuclear euchromatin
C	GO:000228	nuclear chromosome	P16403	GO:0005719	nuclear euchromatin
C	GO:000228	nuclear chromosome	Q13185	GO:0005719	nuclear euchromatin
C	GO:000228	nuclear chromosome	P13010	GO:0000784	nuclear chromosome, telomeric region
C	GO:000228	nuclear chromosome	P13010	GO:0000783	nuclear telomere cap complex
C	GO:000228	nuclear chromosome	Q9H160	GO:0016580	Sin3 complex
C	GO:000228	nuclear chromosome	Q13185	GO:0031618	nuclear centromeric heterochromatin
C	GO:0005575	cellular_component	P02679	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	P02786	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	P11215	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	P27797	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	P50895	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	Q16206	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	Q6UQ28	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	Q8NHL6	GO:0009897	external side of plasma membrane
C	GO:0005575	cellular_component	P27824	GO:0032839	dendrite cytoplasm
C	GO:0005575	cellular_component	P28482	GO:0032839	dendrite cytoplasm
C	GO:0005575	cellular_component	P09104	GO:0043204	perikaryon
C	GO:0005575	cellular_component	P28482	GO:0043204	perikaryon
C	GO:0005575	cellular_component	Q92876	GO:0045171	intercellular bridge

Figura 19 – Porção da tabela obtida com os níveis mais elevados das ontologias génicas (*Parent*) das proteínas de interesse na coluna “*Slim_GO_Name*” e ontologias génicas derivadas (*Child*) na coluna “*Input_GO_Name*”.

3.4.2. Tratamento dos dados provenientes de ontologias

Em seguida, a coluna “*Slim_GO_Name*”, coluna esta que possui como já foi dito anteriormente as ontologias *Parent* das proteínas de interesse, foi filtrada de modo a que a tabela mostrasse apenas as ontologias *Parent* “*immune system process*” e “*cell death*”. A escolha destas ontologias para a construção das tabelas funcionais baseia-se no fato de estarmos a estudar proteínas alteradas em patologias, pelo que, a sua atividade irá afetar a resposta imune a vários níveis, assim como os mecanismos que despoletam na morte celular nas patologias em estudo. A atividade destas proteínas de um ponto de vista imunológico poderá apresentar um significado no desenvolvimento e estabelecimento de cada uma destas patologias.

Estas ontologias são ontologias pertencentes ao grande grupo ontológico “*Biological process (BP)*”, sendo um grupo em que todas as suas ontologias *Parent* nos fornecem ontologias *Child* que possuem informação relativamente a funções e processos biológicos específicos relativamente a cada uma das proteínas (22).

De seguida, foram apenas mantidas na tabela apenas as colunas “*Slim_GO_Name*” referente às ontologias *Parent* “*immune system process*” e “*cell death*”, a coluna “*Input_Accession*” referente ao código *UniProtKBAC* de cada uma das proteínas e a coluna “*Input_GO_Name*” referente à ontologia *Child* derivada da ontologia principal (*Parent*) e que possui informação relativa à função biológica específica de cada proteína (Figura 20).

É importante referir que esta tabela apresenta entradas repetidas da mesma proteína, isto deve-se ao fato de algumas proteínas estarem associadas a ambas as ontologias *Parent* selecionadas e também porque uma mesma proteína pode possuir mais do que uma função biológica específica, pelo que o seu aparecimento em diversas linhas desta tabela é explicado desta forma.

	A	B	C
1	Slim_GO_Name	Input_Accession	Input_GO_Name
9319	immune system process	P13500	positive regulation of macrophage chemotaxis
9320	immune system process	O14788	osteoclast differentiation
9321	immune system process	P02786	osteoclast differentiation
9322	immune system process	P13671	positive regulation of activation of membrane attack complex
9323	immune system process	O14788	monocyte chemotaxis
9324	immune system process	P10147	monocyte chemotaxis
9325	immune system process	P13500	monocyte chemotaxis
9326	immune system process	P17931	monocyte chemotaxis
9327	immune system process	P08571	toll-like receptor 2 signaling pathway
9328	immune system process	P0CG48	toll-like receptor 2 signaling pathway
9329	immune system process	P28482	toll-like receptor 2 signaling pathway
9330	immune system process	P61088	toll-like receptor 2 signaling pathway
17229	cell death	P62258	intrinsic apoptotic signaling pathway
17230	cell death	P99999	intrinsic apoptotic signaling pathway
17231	cell death	Q04917	intrinsic apoptotic signaling pathway
17232	cell death	O14818	negative regulation of apoptotic process
17233	cell death	O95865	negative regulation of apoptotic process
17234	cell death	P01033	negative regulation of apoptotic process
17235	cell death	P04035	negative regulation of apoptotic process
17236	cell death	P04179	negative regulation of apoptotic process
17237	cell death	P06731	negative regulation of apoptotic process
17238	cell death	P09668	negative regulation of apoptotic process
17239	cell death	P0CG48	negative regulation of apoptotic process
17240	cell death	P13693	negative regulation of apoptotic process
17241	cell death	P14174	negative regulation of apoptotic process
17242	cell death	P22392	negative regulation of apoptotic process
17243	cell death	P25786	negative regulation of apoptotic process

Figura 20 – Porção da tabela filtrada em “Slim_GO_Name” para “immune system process” e “cell death” e com apresentação apenas das colunas “Slim_GO_Name” (ontologia Parent), “Input_Accession” (código UniProtKBAC da proteína) e “Input_GO_Name” (ontologia Child com função biológica específica da proteína).

Com vista o tratamento dos dados obtidos da tabela anterior de modo a ser possível construir as tabelas funcionais de cada patologia, foram criadas duas tabelas dinâmicas com os dados desta tabela, uma para “immune system process” e outra para “cell death” em que do lado esquerdo se encontram os “Input_GO_Name” (ontologias Child com funções biológicas específicas das proteínas) e à frente de cada uma o “Input_Accession”, ou seja, a proteína correspondente a essa mesma função. O processo será exemplificado com a metodologia tomada para “immune system process”, tendo sido seguida a mesma metodologia para “cell death” (Figura 21).

O objetivo da construção de ambas as tabelas dinâmicas foi o de organizar a tabela anterior de modo a que todas as proteínas pertencentes a uma mesma ontologia Child e intervenientes numa mesma função biológica específica ficassem corretamente organizadas por grupos de ontologias Child na coluna “Input_GO_Name”, juntando-se portanto todas as ontologias Child no mesmo grupo.

Input_GO_Name	Input_Accession
activated T cell proliferation	P11215
adapptive immune response	P09668
adapptive immune response	P25774
antibacterial humoral response	P12724
antibacterial humoral response	P12838
antigen processing and presentation	P09668
antigen processing and presentation	P25774
antigen processing and presentation	P28062
antigen processing and presentation	Q92930
antigen processing and presentation of endogenous peptide antigen via MHC class I	Q9NZ08
antigen processing and presentation of exogenous antigen	Q06323
antigen processing and presentation of exogenous peptide antigen via MHC class I	O14818
antigen processing and presentation of exogenous peptide antigen via MHC class I	P04839
antigen processing and presentation of exogenous peptide antigen via MHC class I	P0CG48
antigen processing and presentation of exogenous peptide antigen via MHC class I	P14598
antigen processing and presentation of exogenous peptide antigen via MHC class I	P25774
antigen processing and presentation of exogenous peptide antigen via MHC class I	P25786
antigen processing and presentation of exogenous peptide antigen via MHC class I	P25787
antigen processing and presentation of exogenous peptide antigen via MHC class I	P25788
antigen processing and presentation of exogenous peptide antigen via MHC class I	P25789
antigen processing and presentation of exogenous peptide antigen via MHC class I	P27797
antigen processing and presentation of exogenous peptide antigen via MHC class I	P28062
antigen processing and presentation of exogenous peptide antigen via MHC class I	P28066
antigen processing and presentation of exogenous peptide antigen via MHC class I	P49720

Figura 21 – Porção da tabela dinâmica para “immune system process” com apresentação das ontologias *Child* / funções biológicas específicas (“Input_GO_Name”) organizadas por grupos e as proteínas correspondentes (“Input_Accession”).

No entanto e com vista a construção das tabelas funcionais foi de interesse a construção de uma nova tabela ao lado da tabela dinâmica, esta nova tabela permitiu ter grupos organizados com as mesmas ontologias *Child* repetidas e cada uma delas com uma proteína associada, eliminar as repetições das ontologias *Child* e organizar à frente de cada ontologia todas as proteínas intervenientes nessa mesma ontologia, ou seja, função biológica específica. De modo a obter essa tabela colocou-se a ontologia “*immune system process*” em primeiro lugar e na mesma linha foram coladas todas as proteínas correspondentes a “*immune system process*” presentes na tabela da figura 21, este passo permitiu mais tarde identificar quais as proteínas intervenientes em ambas as ontologias *Parent* “*immune system process*” e “*cell death*” e quais as que intervêm em apenas uma destas ontologias. Posteriormente e sob a ontologia “*immune system process*” foram copiados todos os “*Input_GO_Name*” da tabela dinâmica para a nova folha, e usou-se a função “Remove Duplicados” do *Microsoft Excel* 2013 (Figura 22).

	A	B	C	D	E	F
1	immune system process	P11215	P09668	P25774	P12724	P128:
2	activated T cell proliferation					
3	adaptive immune response					
4	antibacterial humoral response					
5	antigen processing and presentation					
6	antigen processing and presentation of endogenous peptide antigen via MHC class I					
7	antigen processing and presentation of exogenous antigen					
8	antigen processing and presentation of exogenous peptide antigen via MHC class I					
9	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent					
10	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent					
11	antigen processing and presentation of exogenous peptide antigen via MHC class II					
12	antigen processing and presentation of peptide antigen via MHC class I					
13	B cell receptor signaling pathway					
14	cellular extravasation					
15	cellular response to interferon-gamma					
16	complement activation					
17	complement activation, alternative pathway					
18	complement activation, classical pathway					
19	complement activation, lectin pathway					

Figura 22 – Porção da tabela com a ontologia “*immune system process*” na primeira célula e todas as proteínas dessa ontologia *Parent* na mesma linha. As ontologias *Child* / funções biológicas específicas não repetidas encontram-se imediatamente abaixo.

De modo a ser possível o preenchimento da presente folha, foi necessário recorrer à fórmula do tipo:

“=SE(DESLOCAMENTO(\$A\$3;CORRESP(\$J2;\$A\$5:\$A\$3000;0)+COL(A1);0)=\$J2;DESLOCAMENTO(\$A\$3;CORRESP(\$J2;\$A\$5:\$A\$3000;0)+COL(A1);1);”

Cuja função respetiva é ir recolher à tabela dinâmica todas as proteínas intervenientes nas ontologias *Child* / funções biológicas específicas “X” repetidas e colocá-las imediatamente à frente dessa mesma ontologia *Child* não repetida na presente folha. A função foi expandida para todas as células desta folha de modo a permitir o seu correto preenchimento (Figura 23), sendo que todo este procedimento foi realizado em paralelo para a ontologia “cell death”.

	J	K	L	M	N	O	P	Q
1	immune system process	P11215	P09668	P25774	P12724	P12838	P09668	P25774
2	activated T cell proliferation	P11215						
3	adaptive immune response	P09668	P25774					
4	antibacterial humoral response	P12724	P12838					
5	antigen processing and presentation	P09668	P25774	P28062	Q92930			
6	antigen processing and presentation of endogenous peptide antigen via MHC class I	Q9NZ08						
7	antigen processing and presentation of exogenous antigen	Q06323						
8	antigen processing and presentation of exogenous peptide antigen via MHC class I	O14818	P04839	P0CG48	P14598	P25774	P25786	P25787
9	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent	O14818	P04839	P0CG48	P14598	P25786	P25787	P25788
10	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent	P25774						
11	antigen processing and presentation of exogenous peptide antigen via MHC class II	P25774	P27824	Q00610	Q10567	Q13409		
12	antigen processing and presentation of peptide antigen via MHC class I	O14818	P04839	P0CG48	P14598	P25774	P25786	P25787
13	B cell receptor signaling pathway	P28482	P41218					
14	cellular extravasation	P11215						
15	cellular response to interferon-gamma	O43293	P10147	P13500				
16	complement activation	O00602	O75636	P00736	P01031	P01609	P01613	P01617
17	complement activation, alternative pathway	P01031	P02748	P07357	P07358	P07360	P08603	P10643
18	complement activation, classical pathway	P00736	P01031	P01609	P01613	P01617	P01625	P01700
19	complement activation, lectin pathway	O00602	O75636					
20	defense response to virus	Q12906	Q70J99	Q8NHL6				
21	dendritic cell differentiation	Q8NHL6						
22	eosinophil chemotaxis	P10147	P17931					
23	eosinophil degranulation	P10147						
24	erythrocyte differentiation	P14317						
25	erythrocyte maturation	P01033	P11413					
26	Fc receptor mediated inhibitory signaling pathway	Q8NHL6						
27	Fc-epsilon receptor signaling pathway	P01116	P01609	P01613	P01617	P01625	P01700	P01714
28	Fc-gamma receptor signaling pathway involved in phagocytosis	O15511	P01609	P01613	P01617	P01625	P01700	P01714

Figura 23 – Porção da tabela da ontologia “*immune system process*” e funções biológicas específicas / ontologias *Child* de si derivadas com todas as proteínas relacionadas imediatamente à frente na mesma linha.

Após a construção de ambas as tabelas para as duas ontologias de interesse, juntaram-se ambas as tabelas, coladas apenas com valores nas células (excluindo assim as fórmulas nelas inseridas), colocando-se a tabela de “*cell death*” imediatamente abaixo da última célula de “*immune system process*”.

3.4.3. Obtenção de atributos para as proteínas exclusivas

Entretanto, e de modo a ser possível o preenchimento das tabelas funcionais para cada patologia, procedeu-se à recolha dos atributos relevantes das proteínas exclusivas. Para isso, retornou-se ao OralOma em *Whole Saliva* proveniente de *Homo sapiens*, e eliminou-se todas as colunas com atributos do OralOma à exceção das colunas com atributos de interesse para a construção das tabelas funcionais, sendo elas a coluna “*UniProtKBAC*” que possui os códigos das proteínas), a coluna “*Name*” com o nome da proteína de acordo com o seu código *UniProtKBAC*, a coluna “*Regulation*” com informação relativa à quantificação da proteína e indicativa na sua alteração em doença comparativamente a saúde, e coluna “*Biomarker*” que contém

informação relativamente às proteínas que já apresentam sugestão como biomarcador de acordo com a literatura.

Estas colunas foram sujeitas à aplicação de filtros avançados do *Microsoft Excel* 2013 de modo a apresentarem apenas os atributos de interesse das proteínas exclusivas das patologias em estudo. Obteve-se com base nesta tabela uma tabela dinâmica que permitiu calcular as médias das “*Regulation*” de proteínas com entradas repetidas devido a diferentes valores de “*Regulation*” referenciados na literatura e registados no OralOma. (Figura 24-25).

	A	B	C	D	E	F	G	H	I	J	L	M
1	UniProtKB AC	Name	Regulation	Biomarker			UniProtKB AC	Exclusivas identificadas todas as patologias				
2	A0N5G3	Rheumatoi					B3EWG3					
3	A0N5G5	Rheumatoi					O00515					
4	A1XBS5	Protein FAM	-91,53				O14638					
5	A2J1M4	Rheumatoi					O14974					
6	A2J1N6	Rheumatoi					O15231					
7	A2RTY6	Inter-alpha	3,20				P00709					
8	A6MUU8	Macrophag					P01116					
9	A6MUU8	Macrophag					P01762					
10	A6NCD1	Tripartite m	-1,23				P02656					
11	A6NJ16	Putative V-ε					P02753					
12	A6NJ16	Putative V-ε	-1,01				P05114					
13	A6NNF4	Zinc finger					P06309					
14	A6QL64	Ankyrin rep					P08779					
15	A8K2U0	Alpha-2-ma					P10619					
16	A8K2U0	Alpha-2-ma					P12532					
17	A8K2U0	Alpha-2-ma					P13646					
18	A8K2U0	Alpha-2-ma	-1,06				P13647					
19	A8MQC9	Calcium-ac					P16083					
20	A8MQC9	Calcium-ac					P16949					
21	A8TX70	Collagen a					P17036					
22	B2R577	cDNA, FLJ3					P17213					
23	B3EWG3	Protein FAM	+				P18135					
24	B3EWG4	Protein FAM					P19013					
25	B3KS75	cDNA FLJ3					P22307					

Figura 24 – Do lado esquerdo visualiza-se porção da tabela derivada do OralOma em *Whole Saliva* proveniente de *Homo sapiens* filtrada de modo a mostrar apenas as colunas que possuem os atributos de interesse “*UniProtKBAC*”, “*Name*”, “*Regulation*” e “*Biomarker*” . Do lado direito, visualiza-se a porção da tabela com a listagem de todas as proteínas exclusivas das patologias em estudo e que permitiu aplicar o filtro avançado do *Microsoft Excel* 2013 para essas entradas na tabela do lado esquerdo.

	A	B	C	D	F	G	I	J	K
1	UniProtKB AC	Name	Regulation	Biomarker			Rótulos de Linha	Média de Regulation	Contagem de Biomarker
2	A0N5G3	Rheumatoid factor					A0N5G3		
3	A0N5G5	Rheumatoid factor					A0N5G5		
4	A1XB55	Protein FAM92A1	-91,53				A1XB55	-91,5272	
5	A2J1M4	Rheumatoid factor					A2J1M4		
6	A2J1N6	Rheumatoid factor					A2J1N6		
7	A2RTY6	Inter-alpha (Globul	3,20				A2RTY6	3,2	
8	A6NCD1	Tripartite motif-con	-1,23				A6NCD1	-1,23	
9	A6NJ16	Putative V-set and					A6NJ16	-1,01	
10	A6NJ16	Putative V-set and	-1,01				B3EWG3	#DIV/0!	
11	B3EWG3	Protein FAM25A	+				B4DUH8		
12	B4DUH8	cDNA FLJ60163, h					F6KPG5	-0,3	
13	F6KPG5	AlbuminAlbumin	-0,30				O00299		
14	O00299	Chloride intracellu					O00391		
15	O00299	Chloride intracellu					O00515	-5,6	
16	O00299	Chloride intracellu					O00560		
17	O00299	Chloride intracellu					O00602		
18	O00391	Sulfhydryl oxidase					O00748	-1,1	
19	O00391	Sulfhydryl oxidase					O00764		
20	O00391	Sulfhydryl oxidase					O14638	-2,1	
21	O00515	Ladinin-1 (Lad-1) (-5,60				O14745		
22	O00515	Ladinin-1 (Lad-1) (O14773		
23	O00515	Ladinin-1 (Lad-1) (O14788		
24	O00560	Syntenin-1 (Melanc					O14818		
25	O00560	Syntenin-1 (Melanc					O14974	-3,4	
26	O00560	Syntenin-1 (Melanc					O15143		
27	O00602	Ficolin-1 (Collager					O15226	-1,01	

Figura 25 – Do lado esquerdo visualiza-se a porção da tabela derivada do OralOma em *Whole Saliva* proveniente de *Homo sapiens* já filtrada para as proteínas exclusivas das patologias em estudo. Do lado direito visualiza-se a tabela dinâmica derivada da tabela do lado esquerdo com cálculo da média de “Regulation” das proteínas com diferentes valores de quantificação referenciados na literatura.

Seguidamente a este passo, voltou-se a converter a tabela dinâmica numa tabela de vista clássica, sendo acrescentada posteriormente a coluna “Name” e “Biomarker” e alterando-se o nome da coluna “Rótulo de linha” para “UniProtKBAC”. Para que a adição das duas novas colunas fosse possível recorreu-se a uma fórmula do tipo:

=PROCV(\$A2;Folha1!\$A\$2:\$D\$20000;2;FALSO)

Fórmula esta que se apresenta com a função de através do código da coluna “UniProtKBAC” ser capaz de ir recolher à tabela filtrada, do OralOma em *Whole Saliva* proveniente de *Homo sapiens* para as proteínas exclusivas das patologias em estudo, os dados necessários para preencher as colunas “Name” e “Biomarker”. Deste modo, obteve-se uma tabela com os atributos das proteínas exclusivas das patologias no que se refere ao seu código “UniProtKBAC”, “Name”, “Regulation” e “Biomarker” (Figura 26). Esta tabela constituirá o cabeçalho das tabelas funcionais finais.

	A	B	C	D	F	G	H
	UniProtKBAC	Name	Regulation	Biomarker			
1							
2	A0N5G3	Rheumatoid factor G9 light chain (Fragment)					
3	A0N5G5	Rheumatoid factor D5 light chain (Fragment)					
4	A1XB55	Protein FAM92A1Protein FAM92A1	-91,53				
5	A2J1M4	Rheumatoid factor RF-ET7 (Fragment)					
6	A2J1N6	Rheumatoid factor RF-ET9 (Fragment)					
7	A2RTY6	Inter-alpha (Globulin) inhibitor H2 (Inter-alpha)	3,20				
8	A6NCD1	Tripartite motif-containing protein ENSP00001	-1,23				
9	A6NJ16	Putative V-set and immunoglobulin domain-c	-1,01				
10	B3EWG3	Protein FAM25AProtein FAM25A	#DIV/0!				
11	B4DUH8	cDNA FLJ60163, highly similar to Carbonic anhydrase 3					
12	F6KPG5	AlbuminAlbuminAlbuminAlbuminAlbumin	-0,30				
13	O00299	Chloride intracellular channel protein 1 (Chloride channel)					
14	O00391	Sulfhydryl oxidase 1 (hQSOX) (EC 1.8.3.2) (Quinone oxidoreductase)					
15	O00515	Lad-1 (Lad-1) (Linear IgA disease antigen)	-5,60				
16	O00560	Syntenin-1 (Melanoma differentiation-associated protein 1)					
17	O00602	Ficolin-1 (Collagen/fibrinogen domain-containing protein)					
18	O00748	Cocaine esterase (EC 3.1.1.84) (Carboxylesterase)	-1,10				
19	O00764	Pyridoxal kinase (EC 2.7.1.35) (Pyridoxine kinase)					
20	O14638	Ectonucleotide pyrophosphatase/phosphodiesterase 1	-2,10				
21	O14745	Na(+)/H(+) exchange regulatory cofactor NHE-1					
22	O14773	Tripeptidyl-peptidase 1 (TPP-1) (EC 3.4.14.9) (Tripeptidase)					
23	O14788	Tumor necrosis factor receptor superfamily member 1B					
24	O14818	Proteasome subunit alpha type-7 (EC 3.4.25.1)					
25	O14974	Protein phosphatase 1 regulatory subunit 12A	-3,40				
26	O15143	Actin-related protein 2/3 complex subunit 1B					
27	O15226	NEK-kappa-B-repressing factor (NEK-repressing factor)	-1,01				

Figura 26 – Porção da tabela com os atributos relevantes das proteínas exclusivas das patologias que servirão de base para a construção das tabelas funcionais de cada patologia.

3.4.4. Tabelas funcionais por patologia

Com a recolha dos dados da tabela anterior e da tabela com a informação das ontologias *Child* / funções biológicas específicas para cada proteína, procedeu-se à construção das tabelas funcionais para cada patologia. O procedimento que foi seguido na construção das tabelas funcionais foi o mesmo para cada patologia, deste modo é descrito apenas vez utilizando como modelo a patologia *Breast Neoplasm (MesH ID 68001943)*.

Numa fase inicial recolheu-se a listagem das funções biológicas específicas da tabela resultante da união das duas tabelas de funções biológicas específicas de “*immune system process*” e “*cell death*” sendo estas colocadas numa nova folha juntamente com as proteínas exclusivas. De seguida, aplicou-se uma fórmula do tipo:

$$=SE(É.ERRO(PROCH(B\$1;Folha9!$B1:$CMM1;1;0));"";"")$$

Cuja função é verificar na tabela de funções biológicas específicas de “*immune system process*” e “*cell death*”, se as proteínas exclusivas de *Breast*

Neoplasm (MesH ID 68001943) estão nas funções biológicas específicas da nova folha e se estiverem será atribuído o valor “x” à célula que corresponde a determinada proteína e correspondente função biológica específica (Figura 27).

	A	B	C	D	E	F	G
1		O14638	P00709	P01116	P01762	P02753	P06309
2	immune system process	+		+	+	+	+
3	cell death		+				
4	activated T cell proliferation						
5	activation of innate immune response						
6	active induction of host immune response by virus						
7	acute inflammatory response to antigenic stimulus						
8	adaptive immune response						
9	alpha-beta T cell differentiation						
10	antibacterial humoral response						
11	antifungal humoral response						
12	antigen processing and presentation						
13	antigen processing and presentation of endogenous peptide antigen via MHC class I						
14	antigen processing and presentation of exogenous antigen						
15	antigen processing and presentation of exogenous peptide antigen via MHC class I						
16	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent						
17	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent						
18	antigen processing and presentation of exogenous peptide antigen via MHC class II						
19	antigen processing and presentation of exogenous protein antigen via MHC class Ib, TAP-dependent						
20	antigen processing and presentation of peptide antigen via MHC class I						
21	antigen processing and presentation of peptide or polysaccharide antigen via MHC class II						
22	antigen processing and presentation, exogenous lipid antigen via MHC class Ib						
23	antigen receptor-mediated signaling pathway						
24	B cell activation						
25	B cell cytokine production						
26	B cell differentiation						
27	B cell proliferation						
28	B cell receptor signaling pathway						

Figura 27 – Porção da tabela com funções biológicas específicas para proteínas exclusivas de *Breast Neoplasm (MesH ID 68001943)*. Nesta tabela verifica-se o resultado da aplicação da fórmula supra-citada, estando já assinaladas as proteínas intervenientes em cada ontologia *Parent*, ou seja, “*immune system process*” e “*cell death*”.

Foram filtradas as tabelas funcionais para mostrarem apenas as proteínas exclusivas da patologia *Breast Neoplasm (MesH ID 68001943)* presentes nas ontologias *Parent* de interesse, que neste caso se trata de “*immune system process*” e “*cell death*”.

Desta forma, obtiveram-se o que se pode apelidar de proteínas exclusivas de *Breast Neoplasm (MesH ID 68001943)* em “*immune system process*” e “*cell death*”. De seguida, foram-se obter os cabeçalhos das tabelas funcionais, e para esse efeito tomou-se como base a tabela da Figura 26 para a recolha dos dados referentes a *Breast Neoplasm (MesH ID 68001943)*. Foram então aplicados filtros avançados para filtrar a tabela de modo a que fossem devolvidos os atributos para as proteínas exclusivas da patologia em “*immune system process*” e “*cell death*” (Figura 28).

	A	B	C	D	E	F	G	I	J	K	M
1	UniProtKE	Name	Regulatio	Biomarker							
20	O14638	Ectonucle	-2,1								
67	P00709	Alpha-lact	2,7								
76	P01116	GTPase KF	-4								
86	P01762	Ig heavy c	#DIV/0!								
103	P02753	Retinol-bi	2,5								
128	P06309	Ig kappa c	#DIV/0!								
210	P17036	Zinc finge	8,1								
211	P17213	Bactericid	2,4								
214	P18135	Ig kappa c	#DIV/0!								
292	P47929	Galectin-7	2,5								
313	P55055	Oxysterol	17,6								
324	P61604	10 kDa he	96								
334	P62987	Ubiquitin-	#DIV/0!								
364	Q13409	Cytoplasm	-50								
472	Q8NHL6	Leukocyte	-2								
509	Q99584	Protein S1	-2								
565											
566											
567											
568											
569											
570											
571											
572											
573											
574											
575											

Figura 28 – Porção da tabela com os atributos referentes às proteínas exclusivas de *Breast Neoplasm (MesH ID 68001943)* em “immune system process” e “cell death”.

Estes dados foram então colocados na tabela da Figura 27, obtendo-se uma tabela funcional para *Breast Neoplasm (MesH ID 68001943)* (Figura 29). Esta tabela foi posteriormente formatada e organizada em grupos comuns referentes a cada uma das ontologias *Child* de modo a facilitar a interpretação funcional de cada proteína na patologia em que intervém (Figura 30). Como já foi dito anteriormente, foram repetidos os mesmos passos para a construção das tabelas funcionais referentes às restantes patologias em “immune system process” e “cell death”.

Após serem construídas todas as tabelas funcionais das patologias em estudo, passou-se à fase seguinte do estudo que consistiu na determinação de critérios para discutir biomarcadores salivares, tanto os que já foram sugeridos na literatura como aqueles que serão sugeridos com este trabalho.

	A	B	C	D
1	UniProtKBAC	O14638	P00709	P01116
2	Name	Ectonucle	Alpha-lact	GTPase KF
3	Regulation	-2,1	2,7	-4
4	Biomarker			
5	Immune system process	+		+
6	cell death		+	
7	antigen processing and presentation of exogenous peptide antigen via MHC class I			
8	antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent			
9	antigen processing and presentation of exogenous peptide antigen via MHC class II			
10	antigen processing and presentation of peptide antigen via MHC class I			
11	complement activation			
12	complement activation, classical pathway			
13	defense response to virus			
14	dendritic cell differentiation			
15	Fc receptor mediated inhibitory signaling pathway			
16	Fc-epsilon receptor signaling pathway			+
17	Fc-gamma receptor signaling pathway involved in phagocytosis			
18	immune response	+		
19	immune response-inhibiting cell surface receptor signaling pathway			
20	Immune system process			
21	innate immune response			+
22	leukocyte activation			
23	leukocyte migration			+
24	mast cell degranulation			

Figura 29 – Porção da tabela funcional inicial de *Breast Neoplasm (MesH ID 68001943)* onde se visualizam as proteínas exclusivas desta patologia em “immune system process” e “cell death” associadas as suas funções biológicas específicas. Verifica-se também a inclusão dos cabeçalhos que foram obtidos ao longo desta metodologia.

	A	B	C	D	E	F	
1			UniProtKBAC	P616	P550	P170	
2			Name	10 kDa h	Oxysterol	Zinc fing	
3			Regulation	36	17,6	8,1	
4			Biomarker				
5		Immune Response	immune system process		+	+	
6	Inate	Antigen	antigen processing and presentation of exogenous peptide antigen via MHC class I, T				
7			antigen processing and presentation of exogenous peptide antigen via MHC class II				
8			negative regulation of T cell activation via T cell receptor contact with antigen bound				
9		Complement	complement activation, classical pathway				
10		Dendritic	negative regulation of dendritic cell differentiation				
11		Macrophage	negative regulation of macrophage activation				
12		Mast Cell	mast cell degranulation				
13		NK Cell	negative regulation of natural killer cell mediated cytotoxicity				
14		Toll-like		toll-like receptor 2 signaling pathway			
15				toll-like receptor 3 signaling pathway			
16				toll-like receptor 4 signaling pathway			
17				toll-like receptor 5 signaling pathway			
18				toll-like receptor 9 signaling pathway			
19				toll-like receptor 10 signaling pathway			
20				toll-like receptor TLR1:TLR2 signaling pathway			
21				toll-like receptor TLR6:TLR2 signaling pathway			
22				TRIF-dependent toll-like receptor signaling pathway			
23				MyD88-dependent toll-like receptor signaling pathway			
24		MyD88-independent toll-like receptor signaling pathway					
25		Immune Response		immune response			
26				innate immune response			
27				regulation of immune response			
28				positive regulation of defense response to virus by host			
29				leukocyte activation			+
30				leukocyte migration			
31				negative regulation of interferon-gamma-mediated signaling pathway		+	
32				immune response-inhibiting cell surface receptor signaling pathway			
33				negative regulation of cytokine secretion involved in immune response			

Figura 30 – Porção da tabela funcional final de *Breast Neoplasm (MesH ID 68001943)* em “immune system process” e “cell death”. Verifica-se a formatação da tabela funcional e a organização das ontologias *Child* em grupos comuns referentes a cada uma destas ontologias / funções biológicas específicas.

3.5. As proteínas propostas como biomarcadores

Como já foi referido, e tendo como base a análise crítica dos biomarcadores sugeridos, procedeu-se à identificação dos mesmos para cada patologia. Estes foram divididos em três categorias: Biomarcadores sugeridos e com dados de quantificação ≤ -3 ou ≥ 3 , biomarcadores sugeridos e com dados de quantificação > -3 ou < 3 e biomarcadores sugeridos com quantificação qualitativa (aumentados/diminuídos). A sua discussão será baseada no fato de serem exclusivos de cada patologia, na sua quantificação e nas suas funções biológicas específicas em “*immune system process*” e “*cell death*” e enquadramento dessas funções na fisiopatologia da patologia.

Relativamente aos biomarcadores a serem sugeridos neste estudo, estes têm de preencher o requisito de serem exclusivos da patologia e possuírem valores de quantificação aumentados (≥ 3) ou diminuídos (≤ -3) (23). De notar, que o valor da quantificação se baseia na variação da quantidade da proteína presente no estado patológico e no controlo saudável, podendo ser calculada recorrendo à fórmula:

$$\frac{\text{Quantificação na doença}}{\text{Quantificação no controlo}} = \text{Valor}$$

Hipótese a) valor ≥ 1 : quantificação positiva (+)

Hipótese b) $0 < \text{valor} < 1$: quantificação negativa (-)

Em **b)**, é necessário realizar o quociente abaixo descrito para que se obtenha uma quantificação com valor negativo:

$$\frac{1}{\text{Valor}} = Y \text{ (a este é atribuído o valor negativo = -Y)}$$

É necessário converter quantificações negativas em valores negativos, pois torna a análise possível, em que, por exemplo, 0,5 significa que há uma

diminuição para a metade de uma proteína em estado de doença, sendo este valor facilmente convertido em valor de quantificação -2.

Após a seleção das proteínas que cumprem os requisitos de potenciais biomarcadores estas serão sujeitas a discussão com base na sua função biológica específica e correlação com a patologia na qual se encontram alterados.

4. Resultados e discussão

4.1. Proteínas identificadas por patologia

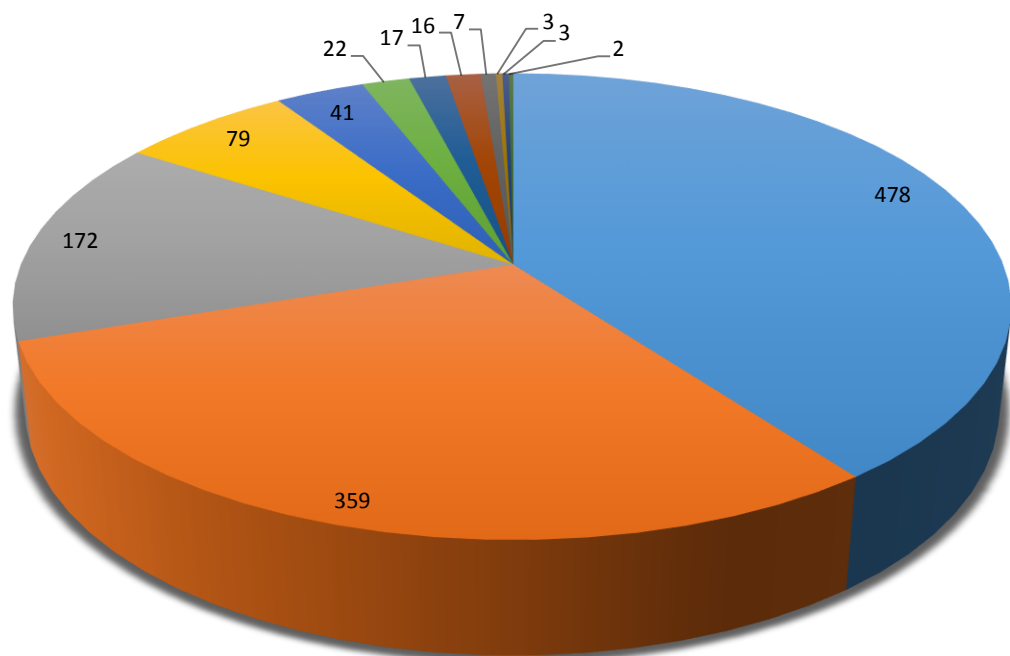
De acordo com a metodologia descrita na secção dos Materiais e Métodos, foram seleccionadas a partir da base de dados OralOma as proteínas que foram identificadas em amostras de *Whole Saliva* de indivíduos com patologias orais e sistémicas e que apresentam proteínas com valores de quantificação. A tabela 1 apresenta as patologias com estes requisitos e as suas respectivas proteínas identificadas perfazendo um total de 1199 proteínas no universo de 2215 proteínas registadas no OralOma em *Whole Saliva*.

Tabela 1 – Proteínas identificadas nas patologias seleccionadas.

Patologia	Proteínas identificadas
<i>Breast Neoplasm (MesH ID 68001943)</i>	172
<i>Diabetes Mellitus type 1 (MesH ID 68003922)</i>	22
<i>Diabetes Mellitus type 2 (MesH ID 68003924)</i>	478
<i>Gingivitis (MesH ID 68005891)</i>	17
<i>Mouth Neoplasms (MesH ID 68009062)</i>	79
<i>Parkinson Disease (MesH ID 68010300)</i>	3
<i>Agressive Periodontitis (MesH ID 68010520)</i>	16
<i>Precancerous Conditions (MesH ID 68011230)</i>	7
<i>Sjogrën Syndrome (MesH ID 68012859)</i>	41
<i>Lichen Planus, Oral (MesH ID 68017676)</i>	2
<i>Chronic Periodontitis (MesH ID 68055113)</i>	359
<i>Peri-Implantitis (MesH ID 68057873)</i>	3
Total	1199

A figura 31 apresenta os dados da tabela 1 mas sobre a forma de gráfico, elucidando a distribuição do número de proteínas identificadas em cada patologia selecionada. Este gráfico realça três patologias com o maior número de proteínas identificadas sendo elas *Diabetes Mellitus type 2* (MesH ID 68003924), *Chronic Periodontitis* (MesH ID 68055113), *Breast Neoplasm* (MesH ID 68001943). Este aspeto deve-se sobretudo à existência de um maior número de estudos de proteómica dirigidos para estas três patologias.

Proteínas Identificadas por patologia



- Diabetes Mellitus type 2 (MesH ID 68003924)
- Chronic Periodontitis (MesH ID 68055113)
- Breast Neoplasm (MesH ID 68001943)
- Mouth Neoplasms (MesH ID 68009062)
- Sjogrën Syndrome (MesH ID 68012859)
- Diabetes Mellitus type 1 (MesH ID 68003922)
- Gingivitis (MesH ID 68005891)
- Agressive Periodontitis (MesH ID 68010520)
- Precancerous Conditions (MesH ID 68011230)
- Parkinson Disease (MesH ID 68010300)
- Peri-Implantitis (MesH ID 68057873)
- Lichen Planus, Oral (MesH ID 68017676)

Figura 31 – Proteínas identificadas das patologias do OralOma em *Whole Saliva* proveniente de *Homo sapiens* que possuem proteínas com valores de quantificação.

4.2. Proteínas exclusivas por patologia

Com a seleção das proteínas exclusivas foi possível verificar a inexistência de proteínas exclusivas em *Agressiva Periodontitis* (MesH ID 68011230) e em *Lichen Planus, Oral* (MesH ID 68017676), sendo portanto retiradas do estudo. A tabela 2 apresenta as proteínas exclusivas das patologias em estudo perfazendo um total de 563 proteínas no universo de 1199 proteínas identificadas nas patologias selecionadas.

Tabela 2 – Proteínas exclusivas nas patologias selecionadas.

Patologia	Proteínas exclusivas
<i>Breast Neoplasm (MesH ID 68001943)</i>	57
<i>Diabetes Mellitus type 1 (MesH ID 68003922)</i>	7
<i>Diabetes Mellitus type 2 (MesH ID 68003924)</i>	287
<i>Gingivitis (MesH ID 68005891)</i>	1
<i>Mouth Neoplasms (MesH ID 68009062)</i>	27
<i>Parkinson Disease (MesH ID 68010300)</i>	2
<i>Agressiva Periodontitis (MesH ID 68010520)</i>	0
<i>Precancerous Conditions (MesH ID 68011230)</i>	4
<i>Sjogrén Syndrome (MesH ID 68012859)</i>	9
<i>Lichen Planus, Oral (MesH ID 68017676)</i>	0
<i>Chronic Periodontitis (MesH ID 68055113)</i>	168
<i>Peri-Implantitis (MesH ID 68057873)</i>	1
Total	563

A figura 32 apresenta os dados da tabela 2 mas sobre a forma de gráfico, elucidando a distribuição do número de proteínas exclusivas em cada patologia selecionada. Este gráfico realça três patologias com o maior número de proteínas identificadas sendo elas *Diabetes Mellitus type 2* (MesH ID 68003924), *Chronic Periodontitis* (MesH ID 68055113), *Breast Neoplasm* (MesH ID 68001943). Tendo em conta que as proteínas exclusivas derivam das proteínas identificadas, é portanto compreensível o destaque destas mesmas patologias.

Proteínas Exclusivas por patologia

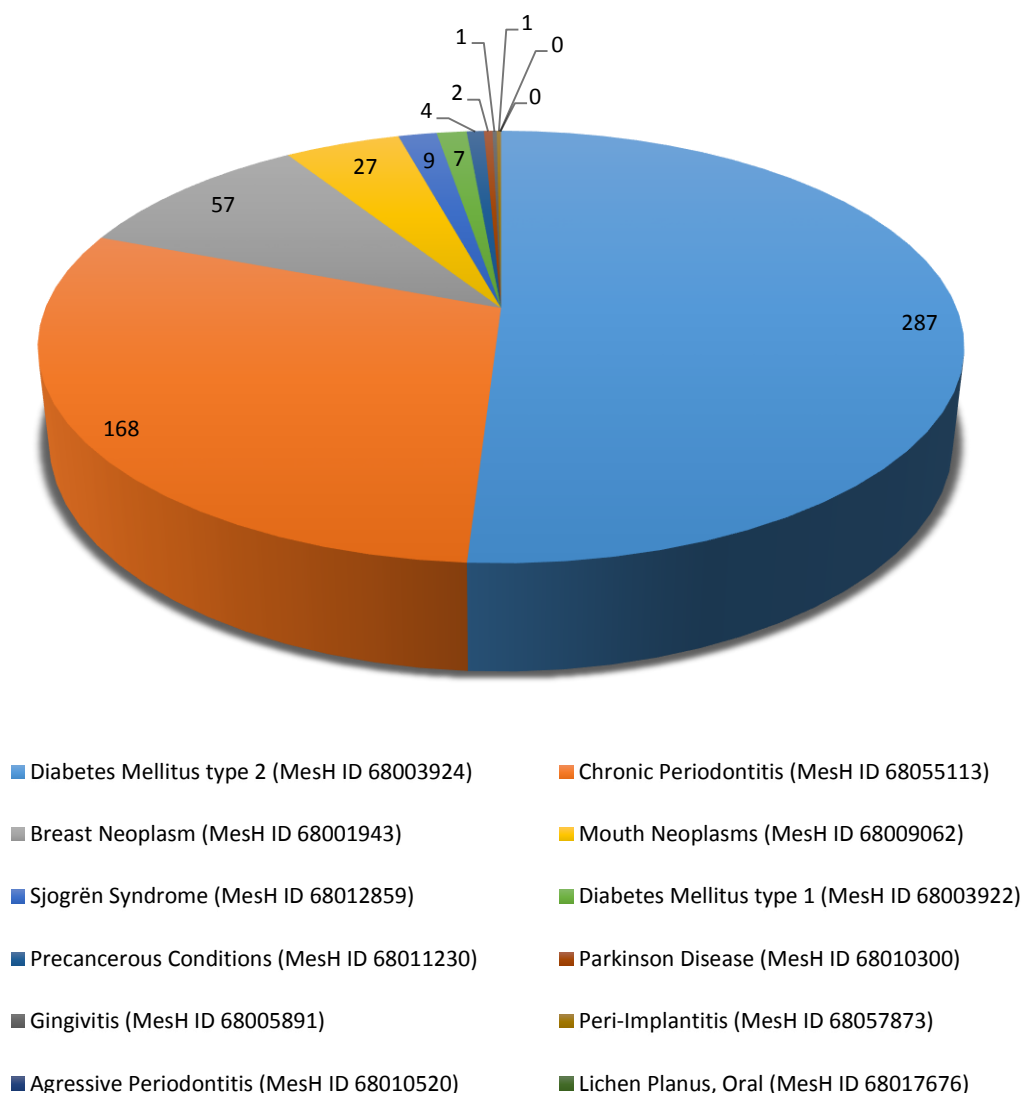


Figura 32 – Proteínas exclusivas das patologias do OralOma em *Whole Saliva* proveniente de *Homo sapiens* que possuem proteínas com valores de quantificação.

4.3. Proteínas exclusivas classificadas por ontologias

A utilização de tabelas funcionais que representam a atribuição de funções biológicas específicas para as proteínas envolvidas em cada uma das patologias em estudo permite a simplificação da análise de caráter funcional de cada uma das proteínas de interesse ao organizar e integrar cada um destes atributos numa tabela de fácil consulta.

A partir da construção das tabelas funcionais (Figura 30) descrita anteriormente foi possível obter proteínas exclusivas para cada uma das patologias ainda em estudo para “*immune system process*” e “*cell death*”, à exceção de *Gingivitis* (MesH ID 68005891), que não apresenta proteínas exclusivas nestas ontologias, pelo que foi retirada de estudo. Quanto às restantes patologias foi possível recolher 16 proteínas para *Breast Neoplasm* (MesH ID 68001943) sem sugestão de biomarcadores; 2 proteínas para *Diabetes Mellitus type 1* (MesH ID 68003922) sem sugestão de biomarcadores; 104 proteínas exclusivas para *Diabetes Mellitus type 2* (MesH ID 68003924) sem sugestão de biomarcadores; 4 proteínas exclusivas para *Mouth Neoplasms* (MesH ID 68009062) com uma sugestão de biomarcador; 2 proteínas exclusivas para *Parkinson Disease* (MesH ID 68010300) sem sugestão de biomarcadores; 3 proteínas exclusivas para *Precancerous Conditions* (MesH ID 68011230) e sugeridas como biomarcadores; 3 proteínas exclusivas para *Sjogren Syndrome* (MesH ID 68012859) sem sugestão de biomarcadores; 63 proteínas exclusivas para *Chronic Periodontitis* (MesH ID 68055113) com três proteínas sugeridas como biomarcadores; 1 proteína exclusiva para *Peri-Implantitis* (MesH ID 68057873) sem sugestão como biomarcador. A partir destes resultados é possível inferir que embora haja um grande painel de 563 proteínas exclusivas, poucos biomarcadores foram sugeridos até ao momento para cada patologia nas ontologias génicas em estudo. No gráfico da figura 33 está representada a relação das proteínas exclusivas de cada patologia com os biomarcadores sugeridos até ao momento para cada patologia.

Proteínas exclusivas em ontologias

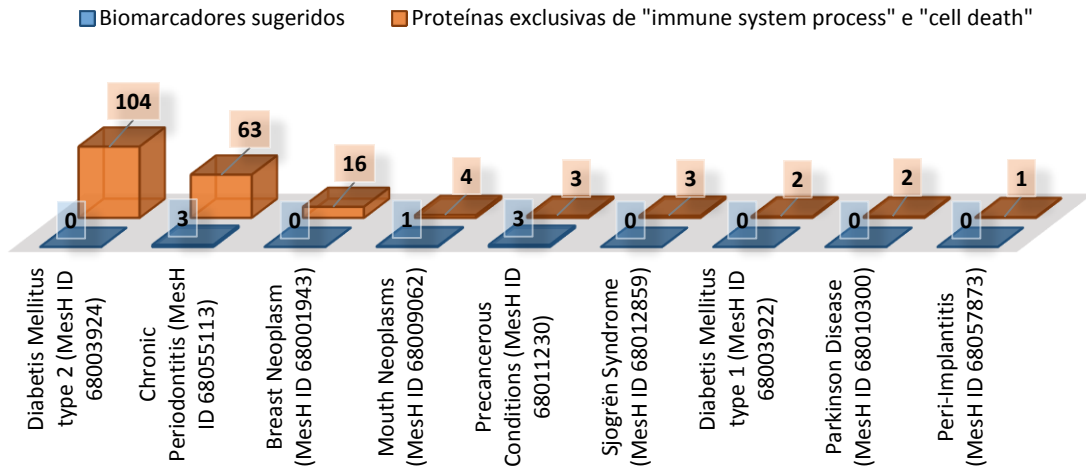


Figura 33 – Neste gráfico estão representadas as proteínas exclusivas em “immune system process” e “cell death” para cada patologia e respectivos biomarcadores sugeridos até ao momento.

A figura 34 gerada pela ferramenta Cytoscape, permite visualizar elucidativamente a disparidade entre proteínas exclusivas e biomarcadores já sugeridos.

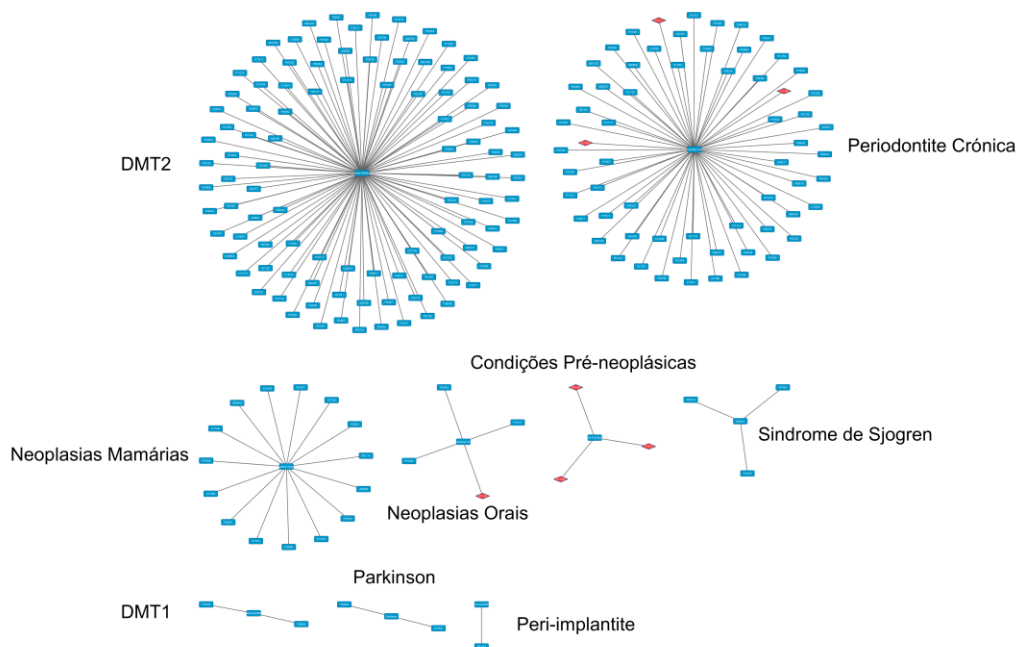


Figura 34 – Clusters de patologias com proteínas exclusivas ligadas ao nó central (que representa a patologia). Os nós a azul são proteínas sem sugestão como biomarcador, enquanto que os nós a vermelho são as proteínas já sugeridas como biomarcador.

Da análise das figuras 33 e 34 é possível concluir que, excluindo *Precancerous Conditions* (MesH ID 68011230) em que todas as proteínas exclusivas já foram sugeridas como biomarcadores, na maioria das patologias grande parte das proteínas identificadas como exclusivas não foram ainda exploradas como biomarcadores.

4.4. Biomarcadores salivares sugeridos

Uma das razões pelas quais uma proteína, apesar de exclusiva, ainda não ter sido proposta como biomarcador poderá ter a ver com a sua quantificação nos vários tipos de amostra. De modo a ser possível discutir os biomarcadores já sugeridos, foi feita a divisão daqueles que se apresentam quantificados daqueles que não se apresentam quantificados. Relembrando o que já foi descrito nos Materiais e Métodos, proteínas que possuam valores de quantificação em diferentes estudos foram sujeitas ao cálculo da média dessa mesma quantificação.

4.4.1. Biomarcadores sugeridos com quantificação ≤ -3 ou ≥ 3

Os biomarcadores já sugeridos das patologias cuja quantificação fosse ≥ 3 (aumento igual ou superior a três vezes da proteína, em doença em relação a saúde) e ≤ -3 (diminuição igual ou inferior a três vezes da proteína, em doença em relação a saúde) foram considerados como proteínas a serem discutidas primariamente por serem as que face a informação atualmente disponível já cumprem os requisitos como marcadores a propor para estudos de validação (23).

Deste modo em *Chronic Periodontitis* (MesH ID 68055113), 3 biomarcadores encontram-se sugeridos até ao momento, sendo que apenas uma cumpre o requisito anterior relativamente à quantificação, a *C-C motif chemokine 3* (P10147) que se encontra com uma quantificação de 18x em doença, e portanto cumpre dois requisitos fundamentais, é exclusiva da patologia e apresenta uma quantificação significativa em doença com valores

passíveis de avaliar pelas atuais técnicas disponíveis e que poderá suportar a provável variabilidade inter-individual prevista.

A *C-C motif chemokine 3* também designada por *MIP-1-alpha* é uma monocina extracelular com propriedades quimiocinéticas envolvida na quimiotaxia de macrófagos, neutrófilos, eosinófilos (e sua desgranulação), granulócitos, linfócitos T e células *natural killer*. Para além do mais possui uma função de regulação negativa da diferenciação osteoclástica e do processo apoptótico neuronal. De notar que esta proteína é um dos principais fatores de supressão produzido pelas células T contra o HIV (24). Em *Chronic Periodontitis* (Mesh ID 68055113), esta função de quimiotaxia das células da resposta inflamatória para as conduzir para os tecidos periodontais é conhecida e é comum a outras patologias inflamatórias crónicas dependentes das interações bactéria-hospedeiro. A chamada ao local da defesa celular é relevante para o estabelecimento do processo inflamatório e constitui o mecanismo responsável pela posterior progressão/destruição de tecidos periodontais característica da periodontite crónica (25) pelo que a sua sugestão como biomarcador apresenta-se pertinente do ponto de vista fisiopatológico da patologia em questão.

4.4.2. Biomarcadores sugeridos com quantificação > -3 ou < 3

Os biomarcadores das patologias cuja quantificação atribuída se encontra no intervalo > -3 e < 3 em doença relativamente a saúde foram classificados como proteínas sem quantificações significativas, pelo que embora sejam exclusivas da patologia, os dados de quantificação existentes não oferecem garantias de que poderão ser obtidas quantificações credíveis face à variabilidade inter-individual expectável pelo que deverão ser realizados mais estudos que comprovem de fato esta falta de quantificação significativa.

Por conseguinte, e novamente em *Chronic Periodontitis* (Mesh ID 68055113) dos 3 biomarcadores já sugeridos, os 2 biomarcadores que ainda não foram especificados, a *Plastin-2* (P13796) apresenta-se com uma quantificação de 2,18x e a *C-C motif chemokine 2* (P13500) com uma

quantificação de 1,8x em relação a saúde, sendo a sua quantificação suportada apenas por um estudo em cada uma das proteínas.

A *Plastin-2* também denominada por *LCP-1* é uma proteína intracelular com ligação a actina que desempenha um papel fundamental na ativação de linfócitos T em resposta a co-estimulação (26, 27), esta proteína modula a expressão por parte da célula T ativada de “*Early activation antigen CD69*” que vai despoletar a sinalização entre células inflamatórias e a proliferação de linfócitos. Esta proteína constitutiva embora com valores de quantificação baixos, desempenha um papel primordial na proliferação linfocitária e sinalização entre células da resposta inflamatória em tecidos periodontais. Estudos posteriores de quantificação desta proteína seriam recomendáveis.

A *C-C motif chemokine 2* também designada por *MCP-1* é um fator quimiotático extracelular que atrai monócitos sendo considerada um dos fatores quimiotáticos mais potentes no que se refere à quimiotaxia de monócitos(25), e está associada a patologias nas quais existe infiltrado de monócitos(28), condição esta que se verifica em *Chronic Periodontitis* (*MesH ID 68055113*). No entanto, esta proteína encontra-se com uma quantificação pouco significativa, sendo que o seu papel funcional em *Chronic Periodontitis* (*MesH ID 68055113*) já é desempenhado pela *C-C motif chemokine 3* (P10147), pelo que a sua sugestão como biomarcador não faz sentido enquanto não forem realizados mais estudos relativos à sua quantificação.

4.4.3. Biomarcadores sugeridos com quantificação qualitativa (aumentados/diminuídos)

Relativamente aos biomarcadores das patologias para os quais só existe quantificação qualitativa, aumentada (+) ou diminuída (-) foram classificados como biomarcadores não quantificados. Tal como sucedeu nos biomarcadores supracitados, estes biomarcadores, por não possuírem dados de quantificação significativa a sua discussão como biomarcador dependerá muito da análise funcional específica de cada proteína no contexto da patologia, sendo nos casos aplicáveis, exigidos estudos para quantificação dos mesmos de modo a validar a sua sugestão como biomarcadores viáveis.

Em *Mouth Neoplasms (MesH ID 68009062)*, apenas um biomarcador se encontra sugerido, a Interleukin 8 (P10145) com uma quantificação aumentada (+).

A *Interleukin 8* também designada por *Chemokine (C-X-C motif) ligand 8* é um fator envolvido na ativação e quimiotaxia de neutrófilos, células T, basófilos, mas não de monócitos. É libertada por vários tipos de células em resposta a estímulos inflamatórios (29, 30). A presença de concentrações elevadas desta citocina está associada ao crescimento e metástase tumoral (31), o que pode contribuir para patogénese em *Mouth Neoplasms (MesH ID 68009062)*, porém existe a necessidade de estudos para obter mais dados de quantificação para esta proteína, de modo a que a sua possível aplicação como biomarcador possa ser considerada.

Em *Precancerous Conditions (MesH ID 68011230)*, 3 biomarcadores encontram-se sugeridos, a *DNA repair protein XRCC5 (P13010)*, a *Quinone reductase 1 (P15559)* e *Interleukin-36 alpha (Q9UHA7)*, todas com apresentação de quantificação aumentada (+).

A *DNA repair protein XRCC5* é uma helicase dependente de cadeia simples de ADN e dependente de ATP, desempenha papel na translocação cromossómica e na diferenciação de células estaminais hematopoiéticas. Está envolvida na estabilização e união de terminações de ADN quebradas, através da eliminação de nucleótidos sem bases nas terminações de ADN resultantes de quebras nas cadeias, sendo que a sua atividade é importante de modo a evitar erros que despoletem em morte celular (32). Esta proteína apresenta-se também com a função de regular negativamente a transcrição (33, 34). É conhecida a ativação desta proteína em *Precancerous Conditions (MesH ID 68011230)* (35) pelo que estudos para quantificação da mesma na saliva seriam favoráveis para a validação da sua sugestão como biomarcador.

A *Quinone reductase 1* ou *QR1* é uma proteína citoplasmática induzida em condições inflamatórias (35) com ligação a FAD e que reduz quinonas a hidroquinonas sendo que a sua atividade enzimática previne a ocorrência de redução de um eletrão das quinonas que tem como produto formas radicais. Esta proteína também tem a função de regular positivamente o processo

apoptótico neuronal. Encontra-se envolvida em processos de biossíntese de protrombina. A sua expressão aumentada poderá estar relacionada com mecanismos de resposta anti-tumoral, no entanto são necessários mais dados de quantificação na saliva de modo a ser passível a sua correta apreciação como biomarcador.

A *Interleukin-36 alpha* também designada por *IL-1F6* apresenta-se com função ao nível da resposta imunitária, no entanto não existem estudos na literatura acerca da sua função específica, pelo que o fato de estar aumentada (35) em *Precancerous Conditions* (*MesH ID 68011230*) não é suficiente para a sugerir como biomarcador, tendo também em conta a falta de dados de quantificação da mesma na saliva.

4.5. Novos biomarcadores sugeridos

De modo a serem sugeridos novos biomarcadores para as patologias em estudo foram definidos dois critérios que devem estar associados a um biomarcador com viabilidade como já foi referido anteriormente neste trabalho. O primeiro diz respeito à quantificação, sendo consideradas como proteínas boas candidatas a biomarcador todas as que apresentem quantificações ≥ 3 (aumento igual ou superior a três vezes da proteína, em doença em relação a saúde) e ≤ -3 (diminuição igual ou inferior a três vezes da proteína, em doença em relação a saúde). O segundo critério diz respeito à proteína ter sido identificada unicamente numa patologia.

Nas diferentes patologias foram selecionadas aquelas cujas proteínas respeitavam estes critérios, sendo que em *Breast Neoplasm* (*MesH ID 68001943*) foram selecionadas 5 proteínas de interesse e em *Diabetes Mellitus type 2* (*MesH ID 68003924*) foram selecionadas 11 proteínas de interesse, perfazendo o total de 16 proteínas sujeitas a análise como possíveis biomarcadores.

Em *Breast Neoplasm* (*MesH ID 68001943*), como já foi referido anteriormente, 5 proteínas foram selecionadas, a *10 kDa heat shock protein* (P61604) com uma quantificação de 96x; a *Oxysterols receptor LXR-beta* (P55055) com uma quantificação de 17,6x; a *Zinc finger protein 3* (P17036)

com uma quantificação de 8,1x; a GTPase KRas (P01116) com uma quantificação de -4x e a *Cytoplasmic dynein 1 intermediate chain 2* (Q13409) que se encontra com uma quantificação de -50x (36).

A 10 kDa *heat shock protein* ou *Hsp10* é um *chaperone*, que está envolvido no *foldings*, montagem e desmontagem de polipéptidos e moléculas de RNA em processos apoptóticos pela ativação de endopeptidases do tipo cisteína. Esta proteína embora presente na matriz mitocondrial, está descrita como presente e expressa em grande quantidade no citosol de células malignas e pré-malignas e no meio extracelular destas, tendo sido assumido que esta proteína se acumula no citoplasma celular e a sua expressão aumenta na transição de displasia para neoplasia. A secreção desta proteína induz resposta imune anti tumoral, associada à patogénese pré-maligna (37). Desta forma, tendo em conta que a sua expressão aumentada está quantificada na saliva (96x) (36), nas fases iniciais de carcinogénese em *Breast Neoplasm* (*MesH ID 68001943*) podemos usá-la como biomarcador para esta patologia.

A *Oxysterols receptor LXR-beta* é um recetor de núcleo e está envolvido na regulação negativa das vias de sinalização mediadas por interferão gama (38). A sua ativação dá-se pela ligação de *oxysterol* (derivado oxidado do colesterol) e culmina na regulação da homeostasia do colesterol através do controlo do seu transporte e neo-síntese. No entanto, a literatura descreve que o aumento da expressão e consequente ativação desta proteína interfere com a função inflamatória das células imunes, nomeadamente e como já foi referido através da diminuição de interferão gama, dando-se então o crescimento tumoral (39). Desta forma, e tendo em conta a sua quantificação significativamente diferente do normal (17,6x) (36) a sua sugestão como biomarcador pode ser considerada viável.

A *Zinc finger protein 3* é um possível fator de transcrição e tem como potencial localização o núcleo, em que a sua função é desempenhada ao nível da diferenciação e proliferação celular, mais especificamente ao nível dos eosinófilos em que esta proteína intervém na sua ativação (40). A sua sugestão como biomarcador não é considerada viável devido à falta de estudos na saliva

relativamente a esta proteína, conhecendo-se pouco da sua ação no contexto de *Breast Neoplasm* (MesH ID 68001943).

A *GTPase KRas* é uma proteína que se liga a *GDP/GTP* e que possui atividade de *GTPase*. Mutações ao nível do gene que codifica esta proteína (nomeadamente nos codões 12, 13, 61) levam a que esta proteína, devido a substituição de aminoácidos, diminua ou perca a sua atividade *GTPase* (conversão de *GTP* a *GDP*), como consequência vamos ter estimulação descontrolada da proliferação celular, sendo também importante referir que esta molécula também intervém na migração leucocitária e nas vias de sinalização com recetor *Fc-epsilon*, pelo que a sua diminuição vai ter consequências severas na capacidade de resposta imune (41). Tendo em conta o que foi dito anteriormente, pode-se sugerir esta proteína como possível biomarcador para *Breast Neoplasm* (MesH ID 68001943).

A *Cytoplasmic dynein 1 intermediate chain 2* é uma proteína citoplasmática presente no citoesqueleto com a função de servir de motor de transporte retrógrado para vesículas e organelos através de microtúbulos estando também envolvida no transporte de antigénios exógenos que despoleta no processamento e apresentação desses antigénios via *MHC* classe II (42). Esta proteína está envolvida na mitose, e é responsável pela inserção do fuso mitótico (constituído por microtúbulos) através da constrição citoplasmática durante a anafase pelo que a sua ausência inviabiliza a correta realização da mitose (43). Tendo em conta este aspeto, e visto que em *Breast Neoplasm* (MesH ID 68001943) esta proteína se encontra bastante diminuída (-50x) (36), é possível presumir que a divisão celular está comprometida e sujeita a erros, pelo que a sugestão desta proteína como biomarcador pode ser viável para esta patologia pois estes processos são característicos das células com fenótipo cancerígeno.

Em *Diabetes Mellitus type 2* (MesH ID 68003924), foram selecionadas, como já foi referido, 11 proteínas de interesse, a *Uteroglobin* (P11684) com uma quantificação de 10,4x; a *Actin-related protein 2/3 complex subunit 5* (O15511) com uma quantificação de 6,1x; a *Complement component C6* (P13671) com uma quantificação de 4,8x; a *Cell division control protein 42*

homolog (P60953) com uma quantificação de 4,8x; a *Pyruvate kinase PKM* (P14618) com uma quantificação de 3,5x; a *Actin-related protein 2* (P61160) com uma quantificação de 3,4x; a *14-3-3 protein eta* (Q04917) com uma quantificação de -3x; a *Apolipoprotein B-100* (P04114) com uma quantificação de -4,1x; a *Desmoplakin* (P15924) com uma quantificação de -5,9x e a *Calnexin* (P27824) com uma quantificação de -7,7x (44) e a *Serum amyloid A-1 protein* (P0DJI8) com uma quantificação de -5,0024 (45).

A *Uteroglobulin* é uma proteína extracelular que se liga e sequestra ligandos hidrofóbicos como a fosfatidilcolina, o fosfatidilinositol, os bifenilpoliclorados (PCB), a progesterona e prostaglandinas (mediadores lipídicos da inflamação). É também uma potente inibidora da atividade da fosfolipase A2 (desencadeia a via pró-inflamatória do ácido araquidónico), regula negativamente a proliferação das células T, inibe a aderência e migração de neutrófilos e monócitos e desempenha um potente papel inibitório na agregação plaquetária característica em diabéticos (46). Em suma, esta proteína apresenta-se com grande atividade anti-inflamatória. De notar que esta proteína devido à sua função fortemente correlacionada com o estabelecimento de hiperlipidémia (47), sinal associado a *Diabetes Mellitus type 2* (MeSH ID 68003924) (48), acaba por ter significado funcional para ser sugerida como biomarcador nesta patologia.

A *Actin-related protein 2/3 complex subunit 5* é uma subunidade de um complexo proteico que é citoplasmático e está presente no citoesqueleto mas também tem localização extracelular em projeções celulares designadas por exossomas, está presente na via de sinalização por recetor *Fc-gamma* envolvida na fagocitose, sendo a sua principal função a de regular a polimerização da actina, mediando a formação de redes de actina do citoesqueleto (49). A evidência sugere que proteínas envolvidas na organização do citoesqueleto de actina são os alvos principais da glicosilação não enzimática verificada nas fases iniciais de *Diabetes Mellitus type 2* (MeSH ID 68003924), sendo que esta glicosilação em condições hiperglicémicas está relacionada, em *Diabetes Mellitus type 2* (MeSH ID 68003924), com a incapacidade das células endoteliais regularem o tónus vascular (50) pelo que

a sua sugestão como biomarcador pode ser considerada viável tendo em conta a sua quantificação significativa.

A *Complement component C6* é uma proteína extracelular envolvida na cascata do complemento envolvida na sua ativação pela via clássica, sendo uma proteína constituinte do complexo de ataque a membrana (MAC) do complemento (51). A evidência sugere que o aumento desta proteína, e consequente montagem do MAC, está relacionado com o aumento leucocitário ao nível do epitélio vascular, assim como ao desenvolvimento de alterações da estrutura vascular que despoleta em lesões ateroscleróticas iniciais verificadas em *Diabetes Mellitus type 2 (MeSH ID 68003924)* (52). Desta forma, é conveniente sugerir esta proteína como possível biomarcador para esta patologia uma vez que está envolvida em processos relacionados com as alterações fisiopatológicas da *Diabetes Mellitus type 2 (MeSH ID 68003924)*.

A *Cell division control protein 42 homolog* é uma proteína localizada na membrana citoplasmática do lado intracelular e está também presente no citoesqueleto, mais propriamente no centróssoma (também designado por centríolo). É uma *GTPase* com atividade na ativação de macrófagos e da fagocitose, co-estimulação de células T e regulação positiva do processo apoptótico neuronal, sendo que sua principal função reside na regulação do arranjo bipolar do fuso mitótico aos cinetocoros (estrutura dos cromátídeos no centrómero que permite a ligação dos microtúbulos do fuso mitótico) antes do alinhamento dos cromossomas na metáfase (53). A literatura sugere também que esta proteína ao ser estimulada por glicose torna-se parte integrante no processo de exocitose de insulina (54), pelo que o seu aumento verificado em *Diabetes Mellitus type 2 (MeSH ID 68003924)* é compatível com os sinais clínicos desta patologia no que diz respeito à produção aumentada de insulina por parte das células β pancreáticas, sendo que a sua sugestão como biomarcador nesta patologia é por isso pertinente do ponto de vista fisiopatológico.

A *Pyruvate kinase PKM* é uma proteína envolvida no metabolismo da glicose, mais propriamente na glicólise, convertendo o ADP + fosfoenolpiruvato em ATP + piruvato, através da transferência do grupo fosforil (55, 56). Esta

proteína está também envolvida em processos de morte celular programada (57). O aumento desta proteína desencadeia um aumento da glicólise, será portanto pertinente sugerir esta proteína como biomarcador com base na sua função no metabolismo de açúcares na fisiopatologia de *Diabetes Mellitus type 2* (MesH ID 68003924).

A *Actin-related protein 2* é a proteína ao qual se liga o ATP na *Actin-related protein 2/3 complex* (já referida anteriormente) (58), pelo que a sua sugestão como biomarcador para *Diabetes Mellitus type 2* (MesH ID 68003924) não se torna muito relevante tendo em conta a sugestão de *Actin-related protein 2/3 complex*, mais determinante como biomarcador na patologia em causa.

A *14-3-3 protein eta* é uma proteína comum a várias vias de sinalização, tanto gerais como específicas, porém tem a capacidade de se ligar ao substrato do recetor de insulina (IRS-1), existindo uma aparente regulação ao nível da sensibilidade da insulina, assim como uma possível associação à resistência a insulina verificada em *Diabetes Mellitus type 2* (MesH ID 68003924) (59), porém não existe na literatura dados concretos sobre os resultados desta interação, pelo que não é viável, por enquanto, sugerir esta proteína como biomarcador para a patologia, sendo sugeridos mais estudos sobre o *outcome* da interação referida anteriormente.

A *Apolipoprotein B-100* é uma proteína com função migratória de leucócitos, é constitutiva dos quilomicrons, *LDL* e *VLDL*, e que a sua função primordial consiste na ligação ao recetor *apoB/E* permitindo a internalização das partículas de *LDL* nas células recetoras. A diminuição desta proteína verificada em *Diabetes Mellitus type 2* (MesH ID 68003924) suporta, como já foi dito anteriormente, a presença de hiperlipidémica característica desta patologia (48), pelo que a sua sugestão como biomarcador é pertinente tendo em conta a fisiopatologia da patologia em questão.

A *Desmoplakin* é uma proteína envolvida na desmontagem dos componentes celulares na fase de execução da apoptose e é uma proteína constituinte dos desmossomas que realizam a adesão celular ao nível epitelial e muscular, sendo que a sua diminuição em *Diabetes Mellitus type 2* (MesH ID

68003924) pode levar a fragilidades ao nível epitelial (60, 61). Pelo que a sua sugestão como biomarcador pode ser considerada viável.

A *Calnexin* é uma proteína que interage com glicoproteínas sintetizadas no retículo endoplasmático, desempenhando a função de reter proteínas com enrolamento incorreto (62). A sua diminuição em *Diabetes Mellitus type 2* (MesH ID 68003924) leva à falha na retenção de proteínas alteradas que podem desempenhar função que leve à resistência a insulina (63), no entanto, mais estudos sobre a sua diminuição em saliva podem ser relevantes para a possibilidade de sugestão da mesma como biomarcador viável.

A *Serum amyloid A-1 protein* é uma proteína envolvida na quimiotaxia de macrófagos, linfócitos e neutrófilos. É também uma apolipoproteína do complexo HDL (64). Existe uma correlação entre o aumento dos níveis plasmáticos desta proteína e produção de citocinas inflamatórias que contribuem para citotoxicidade de células β e resistência a insulina, nomeadamente *IL-1 β* e *TNF- α* (65, 66), no entanto nos estudos dirigidos a amostras salivares, verifica-se a sua diminuição *Diabetes Mellitus type 2* (MesH ID 68003924)(45), sendo que não existe na literatura referência à causa desta diminuição em saliva para *Diabetes Mellitus type 2* (MesH ID 68003924) em contraste ao seu aumento plasmático, no entanto esta sua diminuição em saliva estará em principio relacionada com a fisiopatologia de *Diabetes Mellitus type 2* (MesH ID 68003924), porém, e tendo em conta o desconhecimento desta diminuição à luz da literatura, a sua sugestão como biomarcador está comprometida até à realização de mais estudos a esclarecer os mecanismos envolvidos.

Em suma, foram sugeridos neste trabalho estudos posteriores de validação para 1 biomarcador em *Chronic Periodontitis* (MesH ID 68055113), para 4 biomarcadores em *Breast Neoplasm* (MesH ID 68001943) e para 8 biomarcadores em *Diabetes Mellitus type 2* (MesH ID 68003924). A listagem destas proteínas encontra-se na tabela resumo seguinte (Tabela 3):

Tabela 3 – Tabela resumo com proteínas com sugestão para estudos de validação de biomarcadores para a patologia *Chronic Periodontitis* (MesH ID 68055113), *Breast Neoplasm* (MesH ID 68001943) e *Diabetes Mellitus type 2* (MesH ID 68003924).

<p><i>Chronic Periodontitis</i> (MesH ID 68055113)</p>	<p><i>C-C motif chemokine 3</i> (P10147)</p>
<p><i>Breast Neoplasm</i> (MesH ID 68001943)</p>	<p><i>10 kDa heat shock protein</i> (P61604)</p> <p><i>Oxysterols receptor LXR-beta</i> (P55055)</p> <p><i>GTPase KRas</i> (P01116)</p> <p><i>Cytoplasmic dynein 1 intermediate chain 2</i> (Q13409)</p>
<p><i>Diabetes Mellitus type 2</i> (MesH ID 68003924)</p>	<p><i>Uteroglobin</i> (P11684)</p> <p><i>Actin-related protein 2/3 complex subunit 5</i> (O15511)</p> <p><i>Complement component C6</i> (P13671)</p> <p><i>Cell division control protein 42 homolog</i> (P60953)</p> <p><i>Pyruvate kinase PKM</i> (P14618)</p> <p><i>Apolipoprotein B-100</i> (P04114)</p> <p><i>Desmoplakin</i> (P15924)</p> <p><i>Calnexin</i> (P27824)</p>

5. Conclusão

Ao longo da realização deste estudo foi possível obter diversas conclusões.

A primeira conclusão reside no fato de que se considerarmos as proteínas quantificadas, existem apenas 12 patologias com proteínas quantificadas em *Whole Saliva*.

É também fundamental referir a importância no que diz respeito à quantificação das proteínas em estudo, visto que os dados de quantificação das proteínas alteradas numa determinada patologia permitem avaliar o impacto desta proteína como possível biomarcador e como muitos dos estudos não fazem esta quantificação muita da bibliografia publicada não apresenta valor para aplicação.

Mesmo alguns dos estudos que quantificam as proteínas, apresentam apenas valores de quantificação qualitativa, ou seja, apresentam a proteína como estando aumentada ou diminuída em contexto de patologia.

Alguns estudos embora com apresentação de proteínas com valores de quantificação (quer qualitativa quer quantitativa) discutem a validade da proteína como biomarcador com base apenas na sua quantificação, não considerando a análise funcional que por vezes é essencial na sugestão do biomarcador, visto que uma proteína que apresente valores de quantificação definidos pode não ter consequências significativas na fisiopatologia da doença em que se insere, podendo existir outras proteínas que possuam características funcionais com maior impacto na patologia e que melhor a caracterizam.

Verificou-se ainda que embora existam patologias que possuam muitas proteínas alteradas, poucos biomarcadores se encontram sugeridos atualmente para estas, o que traduz como já foi dito anteriormente a falta de estudos dirigidos para a identificação e quantificação de proteínas específicas de modo a obter informação sobre a variação de expressão das proteínas em causa. Também a falta de estudos de carácter funcional destas proteínas nas patologias em que se inserem constitui uma limitação na proposta de novos biomarcadores. A exceção é *Precancerous Conditions (MeSH ID 68011230)* em que todas as proteínas identificadas foram sugeridas como biomarcadores.

Talvez porque nesta patologia a importância de um diagnóstico precoce é essencial para um bom prognóstico, pelo que terá havido mais interesse e motivação para o estudo de proteínas salivares como biomarcadores.

Deste modo, este trabalho permitiu recolher painéis de possíveis biomarcadores para determinadas patologias de acordo com uma análise funcional específica inserida na fisiopatologia de cada uma das patologias estudadas, sugerindo-se o seu estudo para validação como biomarcadores que identificam as patologias em questão de modo a que no futuro possam ser parte integrante de meios de diagnósticos baseados em saliva.

Este trabalho também demonstrou que apesar da nossa análise final se ter focado nas proteínas implicadas na resposta imunológica, encontramos para as diversas patologias proteínas diferentes e que estão envolvidas em diferentes processos da resposta do hospedeiro. Esta observação permite mostrar que é possível encontrar proteínas que documentam processos específicos da resposta imune que se encontram comprometidos de forma diferente em cada uma das patologias, levando-nos a pensar que em pacientes multi-comprometidos é possível rastrear diferentes biomarcadores para as diferentes patologias, como por exemplo num paciente com periodontite crónica, mas que também seja diabético.

6. Bibliografia

1. Pesquita C, Faria D, Falcao AO, Lord P, Couto FM. Semantic similarity in biomedical ontologies. *PLoS computational biology*. 2009;5(7).
2. Rosa N, Correia M, Arrais J, Lopes P, Melo J, Oliveira J, et al. From the salivary proteome to the OralOme: comprehensive molecular oral biology. *Archives of oral biology*. 2012;57(7):853-64.
3. Arrais J, Rosa N, Melo J, Coelho E, Amaral D, Correia M, et al. *OralCard*: a bioinformatic tool for the study of oral proteome. *Archives of oral biology*. 2013;58(7):762-72.
4. Al Kawas S, Rahim Z, Ferguson D. Potential uses of human salivary protein and peptide analysis in the diagnosis of disease. *Archives of oral biology*. 2012;57(1):1-9.
5. Pfafe T, Cooper-White J, Beyerlein P, Kostner K, Punyadeera C. Diagnostic potential of saliva: current state and future applications. *Clinical chemistry*. 2011;57(5):675-87.
6. Scarano E, Fiorita A, Picciotti P, Passali G, Calò L, Cabras T, et al. Proteomics of saliva: personal experience. *Acta otorhinolaryngologica Italica : organo ufficiale della Società italiana di otorinolaringologia e chirurgia cervico-facciale*. 2010;30(3):125-30.
7. Lawrence H. Salivary markers of systemic disease: noninvasive diagnosis of disease and monitoring of general health. *Journal (Canadian Dental Association)*. 2002;68(3):170-4.
8. Baum B, Yates J, Srivastava S, Wong D, Melvin J. Scientific frontiers: emerging technologies for salivary diagnostics. *Advances in dental research*. 2011;23(4):360-8.
9. Lamy E, Mau M. Saliva proteomics as an emerging, non-invasive tool to study livestock physiology, nutrition and diseases. *Journal of proteomics*. 2012;75(14):4251-8.
10. Rathnayake N, Akerman S, Klinge B, Lundegren N, Jansson H, Tryselius Y, et al. Salivary biomarkers for detection of systemic diseases. *PloS one*. 2013;8(4).

11. Zhang A, Sun H, Wang P, Wang X. Salivary proteomics in biomedical research. *Clinica chimica acta; international journal of clinical chemistry*. 2013;415:261-5.
12. Spielmann N, Wong D. Saliva: diagnostics and therapeutic perspectives. *Oral diseases*. 2011;17(4):345-54.
13. Daniel Malamud IRR-C. Saliva as a Diagnostic Fluid. *Dental clinics of North America*. 2011.
14. Soo-Quee Koh D, Choon-Huat Koh G. The use of salivary biomarkers in occupational and environmental medicine. *Occupational and environmental medicine*. 2007;64(3):202-10.
15. Mi Q, Li N, Ziraldo C, Ghuma A, Mikheev M, Squires R, et al. Translational systems biology of inflammation: potential applications to personalized medicine. *Personalized medicine*. 2010;7(5):549-59.
16. Vodovotz Y, Csete M, Bartels J, Chang... S. Translational systems biology of inflammation. *PLoS computational* 2008.
17. McCarthy FM, Gresham CR, Buza TJ, Chouvarine P, Pillai LR, Kumar R, et al. *AgBase*: supporting functional modeling in agricultural organisms. *Nucleic acids research*. 2011;39(Database issue):506.
18. Consortium U. The universal protein resource (UniProt) in 2010. *Nucleic acids research*. 2010.
19. Oliveros JC. VENNY. An interactive tool for comparing lists with Venn Diagrams. 2007.
20. McCarthy FM, Bridges SM, Wang N, Magee GB, Williams WP, Luthe DS, et al. *AgBase*: a unified resource for functional analysis in agriculture. *Nucleic acids research*. 2007;35(Database issue):603.
21. McCarthy FM, Wang N, Magee GB, Nanduri B, Lawrence ML, Camon EB, et al. *AgBase*: a functional genomics resource for agriculture. *BMC genomics*. 2006;7:229.
22. Ontology G. The Gene Ontology (GO) database and informatics resource. *Nucleic acids research*. 2004.

23. Rosa N, Correia MJ, Arrais J, Costa N, Oliveira... JL. The landscape of protein biomarkers proposed for periodontal disease. Markers with functional meaning. The landscape of protein biomarkers proposed for periodontal disease Markers with functional meaning.
24. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC, Lusso P. Identification of RANTES, MIP-1 alpha, and MIP-1 beta as the major HIV-suppressive factors produced by CD8+ T cells. *Science (New York, NY)*. 1995;270(5243):1811-5.
25. Gupta M, Chaturvedi R, Jain A. Role of monocyte chemoattractant protein-1 (MCP-1) as an immune-diagnostic biomarker in the pathogenesis of chronic periodontal disease. *Cytokine*. 2013;61(3):892-7.
26. Janji B, Giganti A, De Corte V, Catillon M, Bruyneel E, Lentz D, et al. Phosphorylation on Ser5 increases the F-actin-binding activity of L-plastin and promotes its targeting to sites of actin assembly in cells. *Journal of cell science*. 2006;119(Pt 9):1947-60.
27. Wabnitz GH, Köcher T, Lohneis... P. Costimulation induced phosphorylation of L - plastin facilitates surface transport of the T cell activation molecules CD69 and CD25. *European journal of* 2007.
28. Deshmane SL, Kremlev S, Amini... S. Monocyte chemoattractant protein-1 (MCP-1): an overview. *Journal of Interferon &* 2009.
29. Hébert CA, Luscinskas FW, Kiely JM, Luis EA, Darbonne WC, Bennett GL, et al. Endothelial and leukocyte forms of IL-8. Conversion by thrombin and interactions with neutrophils. *Journal of immunology (Baltimore, Md : 1950)*. 1990;145(9):3033-40.
30. Van Damme J, Rampart M, Conings R, Decock B, Van Osselaer N, Willems J, et al. The neutrophil-activating proteins interleukin 8 and beta-thromboglobulin: in vitro and in vivo comparison of NH2-terminally processed forms. *European journal of immunology*. 1990;20(9):2113-8.
31. John MAR, Li Y, Zhou X, Denny P, Ho C-M, Montemagno C, et al. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. *Archives of Otolaryngology–Head & Neck Surgery*. 2004;130(8):929-35.

32. Roberts SA, Strande N, Burkhalter MD, Strom C, Havener JM, Hasty P, et al. Ku is a 5'-dRP/AP lyase that excises nucleotide damage near broken ends. *Nature*. 2010;464(7292):1214-7.
33. Chung U, Igarashi T, Nishishita T, Iwanari H, Iwamatsu A, Suwa A, et al. The interaction between Ku antigen and REF1 protein mediates negative gene regulation by extracellular calcium. *The Journal of biological chemistry*. 1996;271(15):8593-8.
34. Willis DM, Loewy AP, Charlton-Kachigian N, Shao J-SS, Ornitz DM, Towler DA. Regulation of osteocalcin gene expression by a novel Ku antigen transcription factor complex. *The Journal of biological chemistry*. 2002;277(40):37280-91.
35. Kooren JA, Rhodus NL, Tang C, Jagtap PD, Horrigan BJ, Griffin TJ. Evaluating the potential of a novel oral lesion exudate collection method coupled with mass spectrometry-based proteomics for oral cancer biomarker discovery. *Clinical proteomics*. 2011;8:13.
36. Bandhakavi S, Van Riper SK, Tawfik PN, Stone MD, Haddad T, Rhodus NL, et al. Hexapeptide libraries for enhanced protein PTM identification and relative abundance profiling in whole human saliva. *Journal of proteome research*. 2011;10(3):1052-61.
37. Jia H, Halilou AI, Hu L, Cai W, Liu J, Huang B. Heat shock protein 10 (Hsp10) in immune-related diseases: one coin, two sides. *International journal of biochemistry and molecular biology*. 2011;2(1):47-57.
38. Li N, Salter RC, Ramji DP. Molecular mechanisms underlying the inhibition of IFN- γ -induced, STAT1-mediated gene transcription in human macrophages by simvastatin and agonists of PPARs and LXRs. *Journal of cellular biochemistry*. 2011;112(2):675-83.
39. Venteclef N, Ferré P. Liver X receptor: from metabolism to cancer. *The Biochemical journal*. 2014;459(2):3.
40. Thiesen HJ. Multiple genes encoding zinc finger domains are expressed in human T cells. *The New biologist*. 1990;2(4):363-74.
41. Finkelstein SD, Sayegh R, Christensen S, Swalsky PA. Genotypic classification of colorectal adenocarcinoma. Biologic behavior correlates with K - ras - 2 mutation type. *Cancer*. 1993.

42. Bremner KH, Scherer J, Yi J, Vershinin M, Gross... SP. Adenovirus transport via direct interaction of cytoplasmic dynein with the viral capsid hexon subunit. *Cell host & ...* 2009.
43. Karki S, Holzbaur ELF. Cytoplasmic dynein and dynactin in cell division and intracellular transport. *Current opinion in cell biology*. 1999.
44. Rao PV, Reddy AP, Lu X, Dasari S, Krishnaprasad A, Biggs E, et al. Proteomic identification of salivary biomarkers of type-2 diabetes. *Journal of proteome research*. 2009;8(1):239-45.
45. Border MB, Schwartz S, Carlson J, Dibble CF, Kohltfarber H, Offenbacher S, et al. Exploring salivary proteomes in edentulous patients with type 2 diabetes. *Molecular bioSystems*. 2012;8(4):1304-10.
46. Mukherjee AB, Zhang Z, Chilton BS. Uteroglobin: a steroid-inducible immunomodulatory protein that founded the Secretoglobin superfamily. *Endocrine reviews*. 2007;28(7):707-25.
47. Nomori H, Horio H, Takagi... M. Clara cell protein correlation with hyperlipidemia. *CHEST ...* 1996.
48. O'Brien T, Nguyen TT, Zimmerman BR. Hyperlipidemia and diabetes mellitus. *Mayo Clinic Proceedings*. 1998.
49. Welch MD, DePace AH, Verma... S. The human Arp2/3 complex is composed of evolutionarily conserved subunits and is localized to cellular regions of dynamic actin filament assembly. *The Journal of cell ...* 1997.
50. Ghitescu LD, Gugliucci A, Dumas F. Actin and annexins I and II are among the main endothelial plasmalemma-associated proteins forming early glucose adducts in experimental diabetes. *Diabetes*. 2001;50(7):1666-74.
51. Aleshin AE, Schraufstatter IU, Stec B, Bankston LA, Liddington RC, DiScipio RG. Structure of complement C6 suggests a mechanism for initiation and unidirectional, sequential assembly of membrane attack complex (MAC). *The Journal of biological chemistry*. 2012;287(13):10210-22.
52. Fischetti F, Candido R, Toffoli B, Durigutto P, Bernardi S, Carretta R, et al. Innate immunity, through late complement components activation, contributes to the

development of early vascular inflammation and morphologic alterations in experimental diabetes. *Atherosclerosis*. 2011;216(1):83-9.

53. Ocegüera-Yanez F, Kimura K, Yasuda S, Higashida C, Kitamura T, Hiraoka Y, et al. Ect2 and MgcRacGAP regulate the activation and function of Cdc42 in mitosis. *The Journal of cell biology*. 2005;168(2):221-32.

54. Nevins AK, Thurmond DC. A direct interaction between Cdc42 and vesicle-associated membrane protein 2 regulates SNARE-dependent insulin exocytosis. *The Journal of biological chemistry*. 2005;280(3):1944-52.

55. Ashizawa K, McPhie P, Lin KH, Cheng SY. An in vitro novel mechanism of regulating the activity of pyruvate kinase M2 by thyroid hormone and fructose 1, 6-bisphosphate. *Biochemistry*. 1991.

56. Dombrauckas JD, Santarsiero BD, Mesecar AD. Structural basis for tumor pyruvate kinase M2 allosteric regulation and catalysis. *Biochemistry*. 2005.

57. Steták A, Veress R, Ovádi J, Csermely P, Kéri... G. Nuclear translocation of the tumor marker pyruvate kinase M2 induces programmed cell death. *Cancer research*. 2007.

58. Welch MD, Iwamatsu A, Mitchison TJ. Actin polymerization is induced by Arp 2/3 protein complex at the surface of *Listeria monocytogenes*. *Nature*. 1997.

59. Ogihara T, Isobe T, Ichimura T, Taoka... M. 14-3-3 protein binds to insulin receptor substrate-1, one of the binding sites of which is in the phosphotyrosine binding domain. *Journal of Biological ...*. 1997.

60. Armstrong DKB, McKenna KE, Purkis... PE. Haploinsufficiency of desmoplakin causes a striate subtype of palmoplantar keratoderma. *Human molecular ...*. 1999.

61. Jonkman MF, Pasmooij AMG, Pasmans... S. Loss of desmoplakin tail causes lethal acantholytic epidermolysis bullosa. *The American Journal of ...*. 2005.

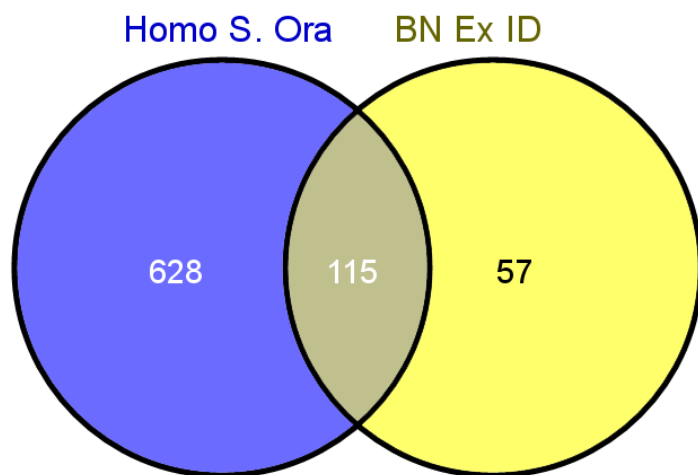
62. David V, Hochstenbach F, Rajagopalan S, Brenner MB. Interaction with newly synthesized and retained proteins in the endoplasmic reticulum suggests a chaperone function for human integral membrane protein IP90 (calnexin). *The Journal of biological chemistry*. 1993;268(13):9585-92.

63. Eizirik DL, Cardozo AK, Cnop M. The role for endoplasmic reticulum stress in diabetes mellitus. *Endocrine reviews*. 2008.

64. Parmelee DC, Titani K, Ericsson LH, Eriksen N, Benditt EP, Walsh KA. Amino acid sequence of amyloid-related apoprotein (apoSAA1) from human high-density lipoprotein. *Biochemistry*. 1982;21(14):3298-303.
65. Serum amyloid A protein in patients with non-insulin-dependent diabetes mellitus. *Clinical Biochemistry*. 1994;27(6):469-473.
66. Interaction between serum amyloid A and leukocytes—A possible role in the progression of vascular complications in diabetes. *Immunology Letters*. 2007;108(2):160-166.

7. Anexos

Figura A. Gráfico de Venn das proteínas exclusivas em *Breast Neoplasm (MesH ID 68001943)*.



B3EWG3	P28838	
O00515	P47929	
O14638	P49189	
O14974	P49221	
O15231	P49411	
P00709	P50895	
P01116	P52790	
P01762	P54727	
P02656	P55055	
P02753	P61604	
P05114	P62834	
P06309	P62987	
P08779	Q08257	
P10619	Q13296	
P12532	Q13409	
P13646	Q13642	
P13647	Q6EEV6	
P16083	Q8IUE6	
P16949	Q8IW75	Q96JC1
P17036	Q8NHL6	Q96NY8
P17213	Q8WWI1	Q99584
P18135	Q92484	Q9BW30
P19013	Q92876	Q9HC38
P22307	Q92930	Q9NY33
P24844	Q96IU4	Q9UNF0

Figura B. Gráfico de Venn das proteínas exclusivas em *Diabetes Mellitus type 1* (MesH ID 68003922).

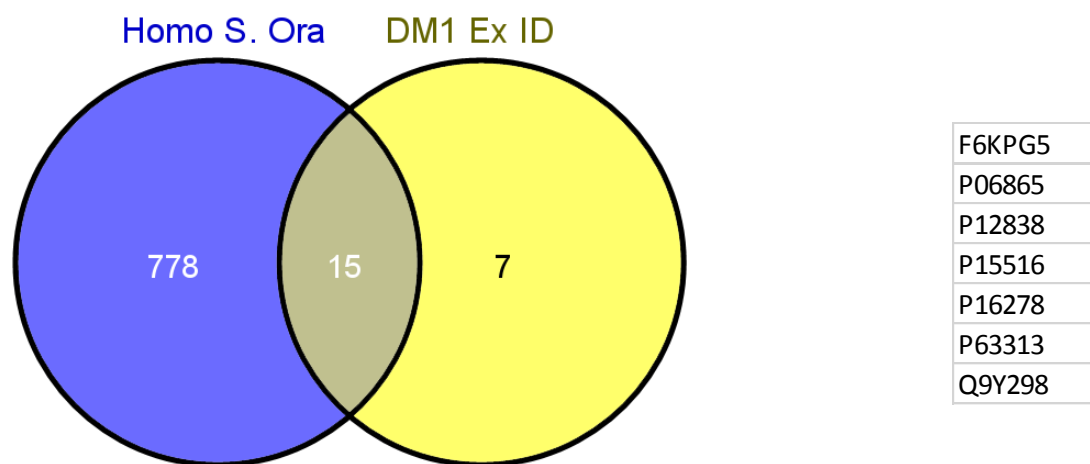
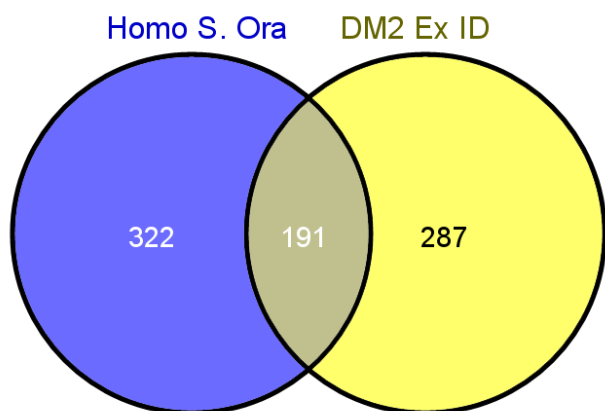


Figura C. Gráfico de Venn das proteínas exclusivas em *Diabetes Mellitus type 2* (MesH ID 68003924).



A0N5G3	P00367	P07384	P16870	P31146	P61160	Q6FHR0	Q15084
A0N5G5	P00390	P07900	P16930	P31943	P61586	Q6F113	Q15149
A1XBS5	P00492	P07988	P18206	P31997	P61916	Q6ICN0	Q15300
A2J1M4	P00505	P07998	P19105	P32320	P61978	Q6P4A8	Q15631
A2J1N6	P00734	P08118	P19827	P34932	P61981	Q6XQN6	Q15782
A2RTY6	P00736	P08238	P19957	P35326	P62258	Q6YHK3	Q16881
O00299	P00747	P08519	P20700	P35754	P62316	Q6ZW64	Q1T7A3
O00391	P01019	P08571	P20933	P36222	P67936	Q70J99	Q4VAX6
O00560	P01031	P08603	P22314	P36871	P68036	Q7M4Q5	Q4VB24
O00602	P01042	P09417	P22735	P36955	P78324	Q7Z3Y5	Q4W4Y1
O00764	P02452	P09467	P23141	P39687	P78417	Q86U62	Q53EM5
O14745	P02549	P09668	P23142	P41218	P98088	Q86UX7	Q53EP0
O14773	P02652	P09758	P23526	P42785	Q00610	Q8N1C8	Q53F10
O14818	P02748	P09871	P25774	P43034	Q03252	Q8N274	Q53F48
O15143	P02760	P09958	P25786	P43490	Q04917	Q8NGY0	Q53G35
O15511	P03952	P0C0L4	P25787	P43652	Q06323	Q8TDL5	Q53G71
O43240	P04114	P0DJ18	P25788	P46926	Q08209	Q8WWA0	Q53H01
O43278	P04179	P0DJ19	P25789	P46940	Q09666	Q8WXI7	Q53H57
O43490	P04196	P10153	P26022	P46976	Q0VDC6	Q92817	Q53XB4
O43866	P04839	P10643	P27169	P47755	Q10567	Q92820	Q59E99
O60361	P04908	P10645	P27487	P48637	Q10588	Q96BQ1	Q59FP5
O60437	P05089	P11215	P27824	P48723	Q12805	Q96PD5	Q59HE3
O60664	P05090	P11413	P27918	P49720	Q12906	Q96RM1	Q5F1R6
O75083	P05107	P11684	P28062	P49721	Q12907	Q99102	Q5NV82
O75367	P05156	P12724	P28066	P50552	Q13126	Q99536	Q5NV91
O75369	P05204	P13489	P28482	P51884	Q13185	Q99877	Q55TX8
O75390	P05387	P13671	P28676	P52907	Q13231	Q9BS26	Q5T949
O75436	P05976	P13693	P28799	P54802	Q14019	Q9BW34	Q5TCI8
O75556	P06576	P13797	P29034	P55042	Q14103	Q9NUQ9	Q5TD07
O75882	P06681	P14550	P29373	P55058	Q14210	Q9NYL9	Q5U071
O95336	P06737	P14598	P30040	P55072	Q14624	Q9NZ08	Q5VXV3
O95445	P07195	P14618	P30043	P58499	Q14914	Q9NZT1	Q5VY30
O95479	P07357	P15586	P30044	P60900	Q14CN2	Q9UBU3	Q6FGD7
O95865	P07358	P15924	P30046	P60953	Q15075	Q9UHL4	Q6FGL0
O95969	P07360	P16035	P30048	P61088	Q15080	Q9UJ68	Q6FGX5
							Q9UJ70
							Q9UL52
							Q9UL88
							Q9UNZ2
							Q9Y3E8
							Q9Y490
							Q9Y5J9

Figura D. Gráfico de Venn das proteínas exclusivas em *Gingivitis* (MesH ID 68005891).

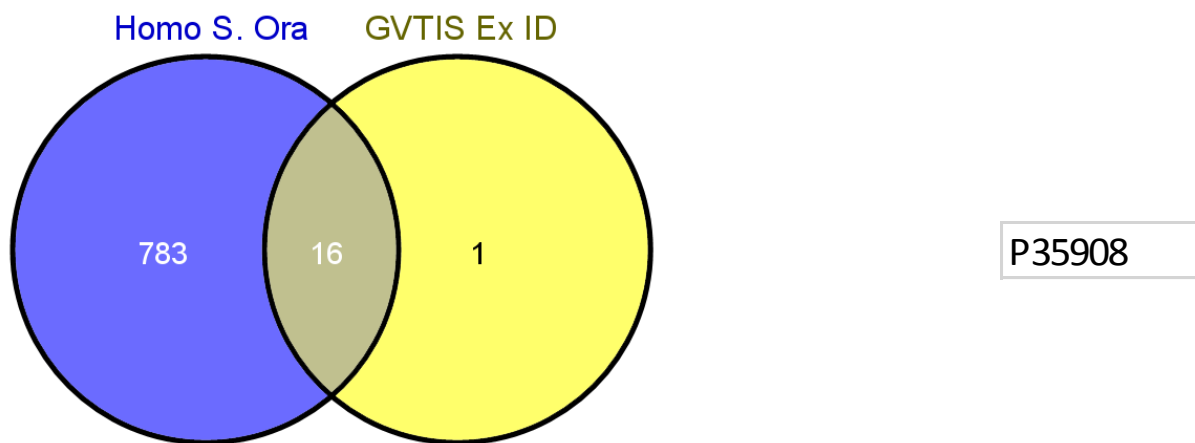
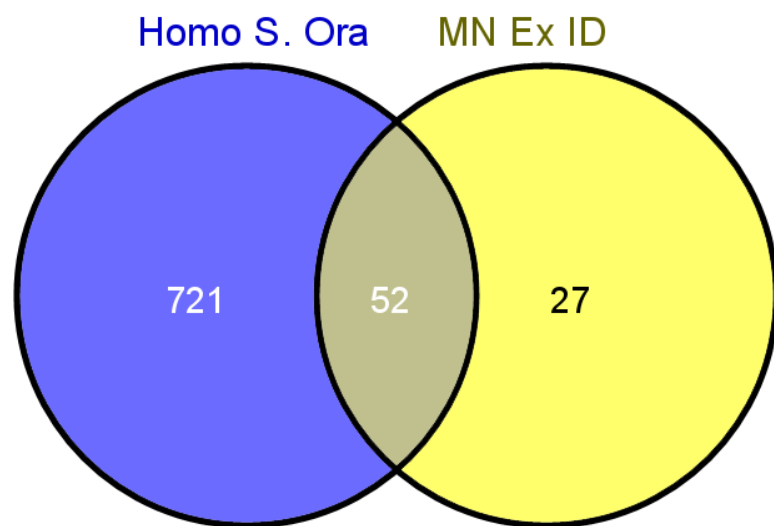


Figura E. Gráfico de Venn das proteínas exclusivas em *Mouth Neoplasms* (MesH ID 68009062).



O43933	
P02100	
P05452	
P07305	
P08253	
P09104	
P10145	
P13929	
P14317	
P16403	
P20309	
P21673	
P28562	
P46013	
P50452	
P54368	
Q53HE2	
Q5T4B6	
Q6ZNH5	
Q8TBR0	
Q8TE27	
Q8TF72	
Q96BU1	
Q9H014	Q9UDX3
Q9NYS0	Q9UPN9

Figura F. Gráfico de Venn das proteínas exclusivas em *Parkinson Disease* (MesH ID 68010300)

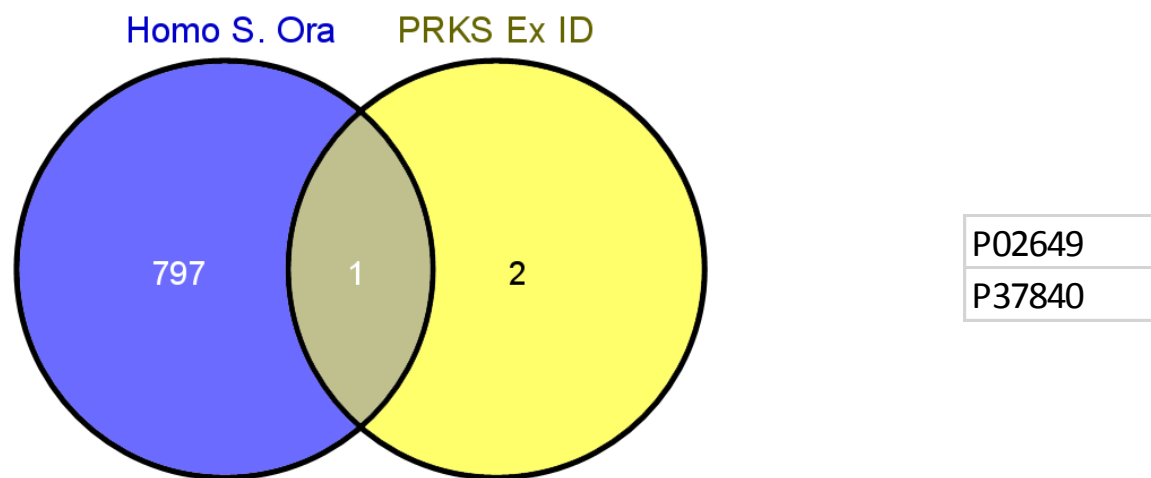


Figura G. Gráfico de Venn das proteínas exclusivas em *Agressive Periodontitis* (MesH ID 68010520)

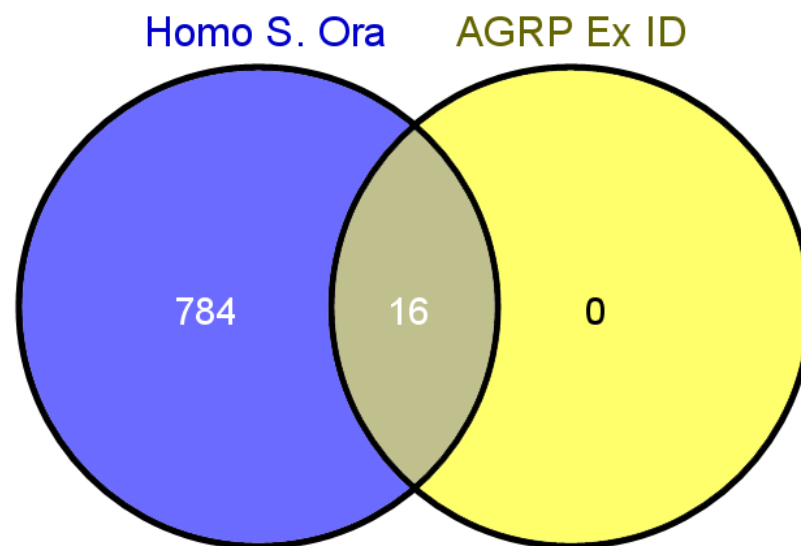


Figura H. Gráfico de Venn das proteínas exclusivas em *Precancerous Conditions* (MesH ID 68011230).

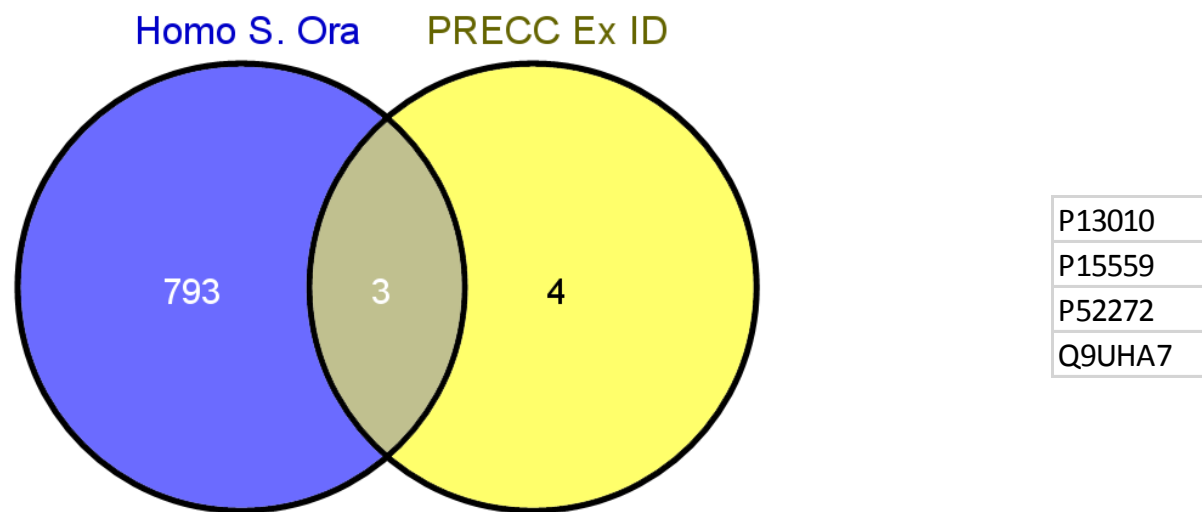


Figura I. Gráfico de Venn das proteínas exclusivas em *Sjogren Syndrome* (MesH ID 68012859).

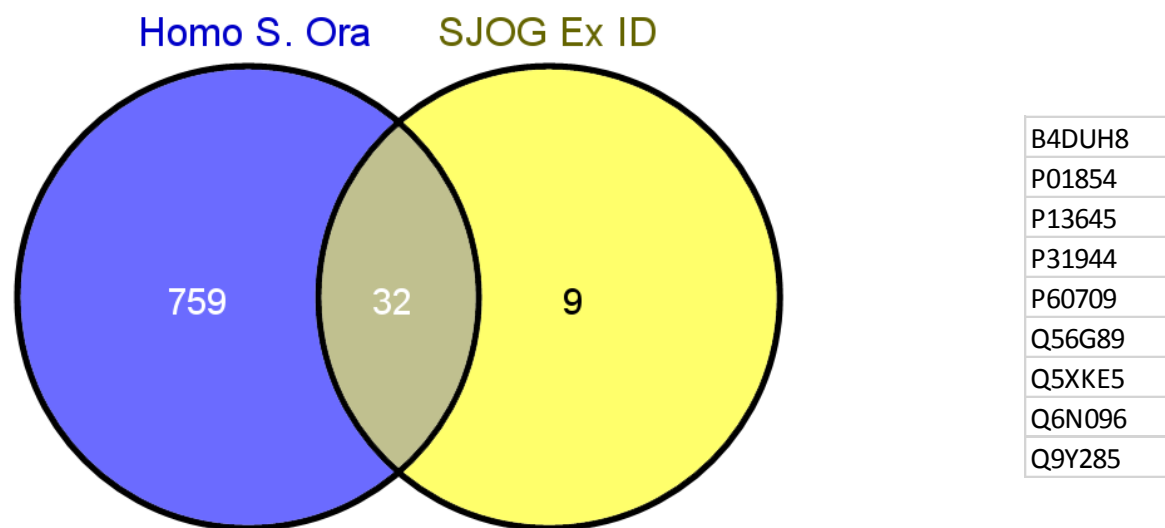


Figura J. Gráfico de Venn das proteínas exclusivas em *Lichen Planus, Oral* (MesH ID 68017676).

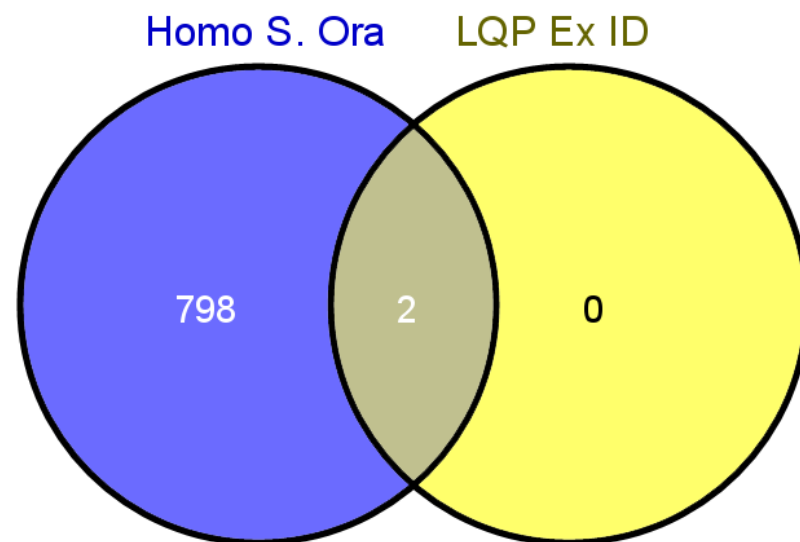
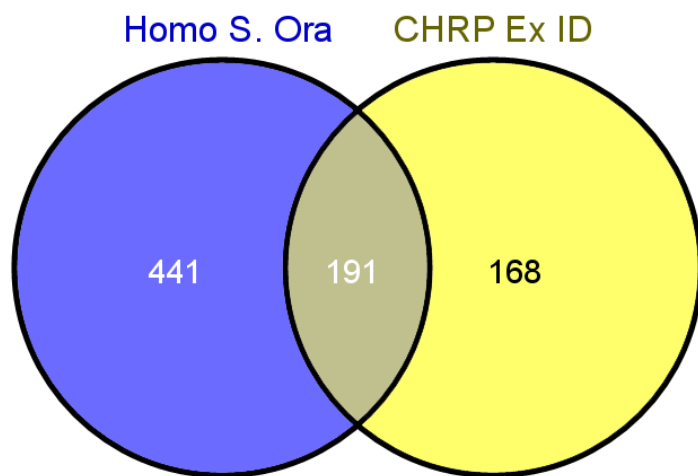


Figura K. Gráfico de Venn das proteínas exclusivas em *Chronic Periodontitis* (MesH ID 68055113).



A6NCD1	P02679	P31751	Q6R327	
A6NJ16	P02786	P35325	Q6UQ28	
O00748	P04035	P36952	Q6UXG3	Q99459
O14788	P04207	P38919	Q6YI46	Q99759
O15226	P04211	P47756	Q6ZN30	Q99856
O15516	P04430	P51530	Q6ZNI0	Q9BQR3
O43293	P04433	P54289	Q6ZRS2	Q9BT49
O43707	P04434	P58062	Q6ZS10	Q9BUH6
O43852	P05787	P61970	Q6ZV73	Q9BWK5
O75037	P06331	P62195	Q86VP1	Q9BYJ1
O75223	P06396	P62826	Q86XP1	Q9BZJ6
O75636	P06731	P62993	Q8IVV2	Q9BZZ5
O75717	P07951	P80419	Q8IWW7	Q9H160
O75829	POC0L5	P80748	Q8IX01	Q9H1A4
O94868	POC221	P85298	Q8IXL6	Q9H799
O95622	POCG05	P99999	Q8IYT8	Q9HBA0
O95625	POCG48	Q02818	Q8IZJ1	Q9HBG4
P00450	P10147	Q13423	Q8N130	Q9HCJ0
P01011	P11142	Q14008	Q8N1A0	Q9NT68
P01033	P12830	Q14264	Q8N7Z2	Q9NUP1
P01609	P12955	Q14517	Q8NB25	Q9NVM4
P01613	P13473	Q14563	Q8TAQ5	Q9NVR0
P01617	P13500	Q14697	Q8TAT5	Q9NVR2
P01625	P13796	Q14978	Q8TD57	Q9NYF0
P01700	P14174	Q15477	Q8TEP8	Q9UL36
P01714	P17931	Q16206	Q8WZ64	Q9ULL4
P01717	P18065	Q2TB90	Q96GX5	Q9UQ16
P01743	P18669	Q49AJ0	Q96HA8	Q9Y2Q0
P01763	P22392	Q4G0P3	Q96N87	Q9Y483
P01764	P22528	Q5M9N0	Q96NL6	Q9Y6R7
P01767	P23083	Q5XG99	Q96NM4	
P01780	P27348	Q6IQ22	Q96NY7	
P01824	P27797	Q6MZM9	Q96PR1	
P01825	P28347	Q6NVV3	Q96PZ7	
P01860	P29400	Q6PIF6	Q96RR4	

Figura L. Gráfico de Venn das proteínas exclusivas em *Peri-Implantitis* (MesH ID 68057873).

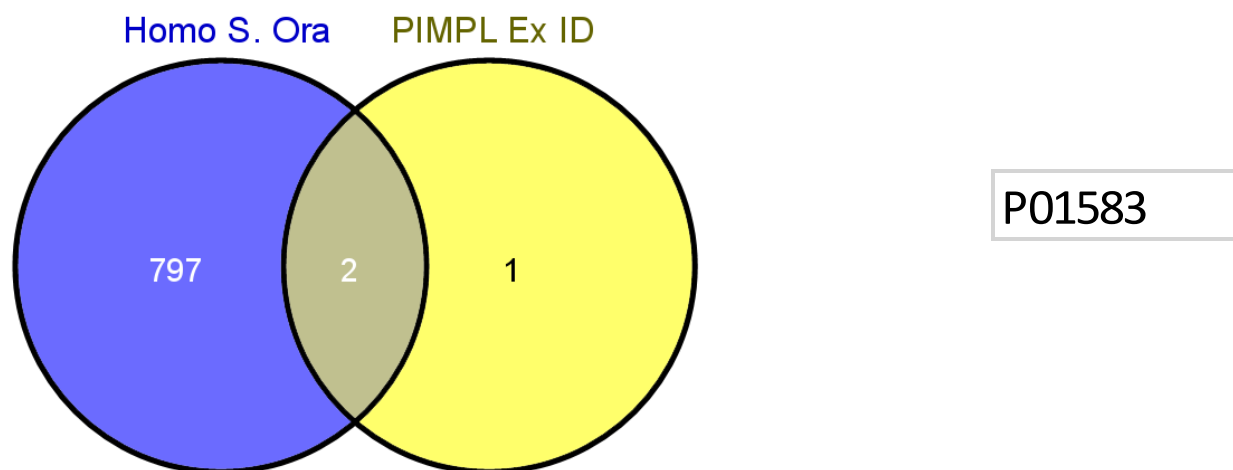


Tabela A. Tabela funcional de *Breast Neoplasm* (MesH ID 68001943) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela B. Tabela funcional de *Diabetes Mellitus type 1* (MesH ID 68003922) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela C. Tabela funcional de *Diabetes Mellitus type 2* (MesH ID 68003924) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela D. Tabela funcional de *Mouth Neoplasms* (MesH ID 68009062) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela E. Tabela funcional de *Parkinson Disease* (MesH ID 68010300) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

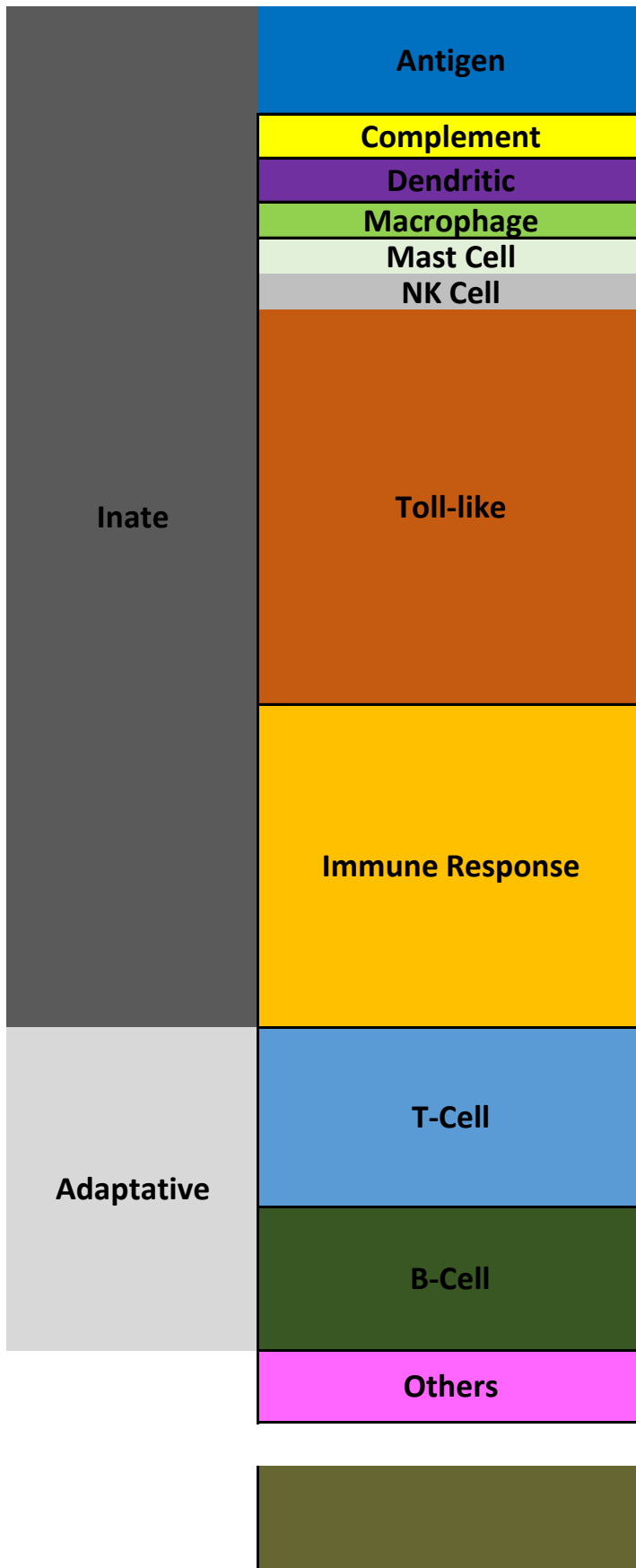
Tabela F. Tabela funcional de *Precancerous Conditions* (MesH ID 68011230) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela G. Tabela funcional de *Sjogren Syndrome* (MesH ID 68012859) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela H. Tabela funcional de *Chronic Periodontitis* (MesH ID 68055113) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Tabela I. Tabela funcional de *Peri-Implantitis* (MesH ID 68057873) com proteínas exclusivas em “immune system process” e “cell death”. Este anexo encontra-se em formato digital (CD-R anexado).

Immune Response



Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent
antigen processing and presentation of exogenous peptide antigen via MHC class II
negative regulation of T cell activation via T cell receptor contact with antigen bound to MHC molecul
complement activation, classical pathway
negative regulation of dendritic cell differentiation
negative regulation of macrophage activation
mast cell degranulation
negative regulation of natural killer cell mediated cytotoxicity
toll-like receptor 2 signaling pathway
toll-like receptor 3 signaling pathway
toll-like receptor 4 signaling pathway
toll-like receptor 5 signaling pathway
toll-like receptor 9 signaling pathway
toll-like receptor 10 signaling pathway
toll-like receptor TLR1:TLR2 signaling pathway
toll-like receptor TLR6:TLR2 signaling pathway
TRIF-dependent toll-like receptor signaling pathway
MyD88-dependent toll-like receptor signaling pathway
MyD88-independent toll-like receptor signaling pathway
immune response
innate immune response
regulation of immune response
positive regulation of defense response to virus by host
leukocyte activation
leukocyte migration
negative regulation of interferon-gamma-mediated signaling pathway
immune response-inhibiting cell surface receptor signaling pathway
negative regulation of cytokine secretion involved in immune response
T cell receptor signaling pathway
positive regulation of gamma-delta T cell activation involved in immune response
negative regulation of CD8-positive, alpha-beta T cell activation
negative regulation of T cell proliferation
negative regulation of T cell mediated cytotoxicity
positive regulation of immunoglobulin secretion
Fc receptor mediated inhibitory signaling pathway
Fc-epsilon receptor signaling pathway
Fc-gamma receptor signaling pathway involved in phagocytosis
negative regulation of osteoclast development
nucleotide-binding oligomerization domain containing signaling pathway
cell death
apoptotic process
positive regulation of apoptotic process
negative regulation of apoptotic process

apoptotic signaling pathway

positive regulation of cytolysis

negative regulation of dendritic cell apoptotic process

activation of cysteine-type endopeptidase activity involved in apoptotic process

							+	
							+	
+								

O14638	P01116	Q13409	P01762	P06309	P18135	P62987
Ectonucleo	GTPase KR	Cytoplasm	Ig heavy ch	Ig kappa ch	Ig kappa ch	Ubiquitin-6
-2.1	-4	-50	+	+	+	+
+	+	+	+	+	+	+
						+
		+				
			+	+	+	
						+
						+
						+
						+
						+
						+
						+
						+
						+
						+
						+
+			+	+	+	
	+		+	+	+	+
			+	+	+	
	+					
	+		+	+	+	+
			+	+	+	
						+
						+
						+
						+
						+

						+

Immune Response

Inate

Immune Response

Cellular death

UniProtKBAC	P12838	P06865
Name	Neutrophil	Beta-hexos
Regulation	1.9	1.8
Biomarker		
immune system process	+	
innate immune response	+	
cell death		+
cell death		+

Immune Response

Antigen

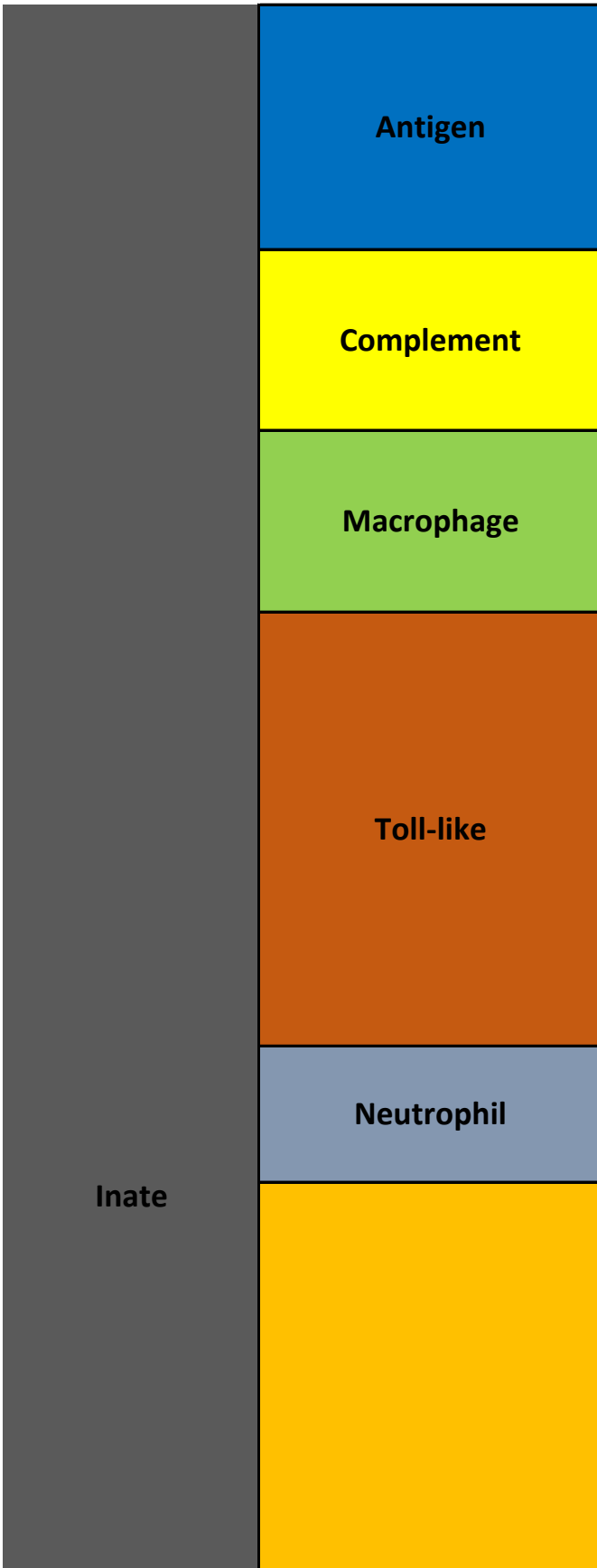
Complement

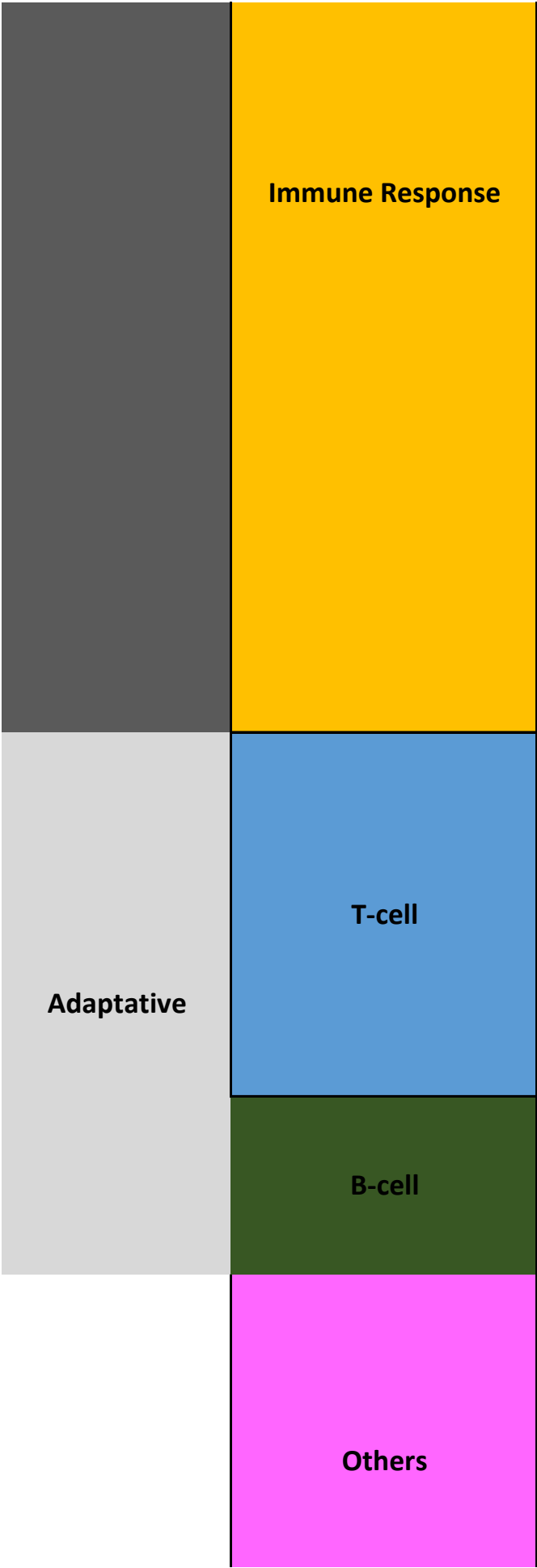
Macrophage

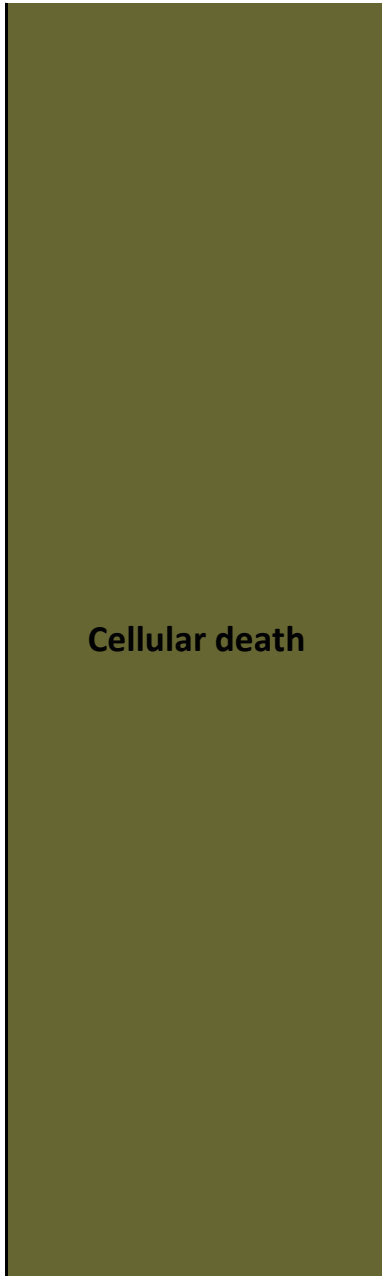
Toll-like

Neutrophil

Inate







Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
antigen processing and presentation
antigen processing and presentation of peptide antigen via MHC class I
antigen processing and presentation of endogenous peptide antigen via MHC class I
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent
antigen processing and presentation of exogenous peptide antigen via MHC class II
complement activation, alternative pathway
complement activation, classical pathway
complement activation, lectin pathway
regulation of complement activation
positive regulation of complement activation
macrophage differentiation
macrophage chemotaxis
negative regulation of macrophage chemotaxis
pattern recognition receptor signaling pathway
toll-like receptor signaling pathway
toll-like receptor 2 signaling pathway
toll-like receptor 3 signaling pathway
toll-like receptor 4 signaling pathway
toll-like receptor 5 signaling pathway
toll-like receptor 9 signaling pathway
toll-like receptor 10 signaling pathway
toll-like receptor TLR1:TLR2 signaling pathway
toll-like receptor TLR6:TLR2 signaling pathway
TRIF-dependent toll-like receptor signaling pathway
MyD88-dependent toll-like receptor signaling pathway
MyD88-independent toll-like receptor signaling pathway
neutrophil chemotaxis
neutrophil mediated killing of fungus
neutrophil mediated killing of gram-positive bacterium
immune system process
immune response
adaptive immune response
humoral immune response
cellular extravasation
regulation of immune response
immune response-regulating signaling pathway
negative regulation of immune response
positive regulation of immune response to tumor cell
negative regulation of cytokine secretion involved in immune response

negative regulation of natural killer cell differentiation involved in immune response
innate immune response
regulation of innate immune response
defense response to virus
regulation of defense response to virus by virus
leukocyte migration involved in inflammatory response
leukocyte chemotaxis
leukocyte mediated cytotoxicity
leukocyte migration
lymphocyte chemotaxis
lymphocyte differentiation
lymphocyte homeostasis
lymphocyte proliferation
respiratory burst involved in defense response
positive regulation of activation of membrane attack complex
regulation of interferon-gamma-mediated signaling pathway
regulation of mast cell degranulation
regulation of type I interferon-mediated signaling pathway
type I interferon signaling pathway
T cell receptor signaling pathway
T cell costimulation
T cell activation
positive regulation of T cell activation
activated T cell proliferation
T cell proliferation
positive regulation of T cell proliferation
negative regulation of T cell proliferation
negative regulation of T cell migration
T cell mediated cytotoxicity
T cell homeostasis
B cell receptor signaling pathway
negative regulation of B cell proliferation
Fc-epsilon receptor signaling pathway
Fc-gamma receptor signaling pathway involved in phagocytosis
opsonization
erythrocyte maturation
germinal center formation
granuloma formation
hemopoiesis
microglia development
mononuclear cell migration
myeloid cell differentiation
natural killer cell degranulation

nucleotide-binding oligomerization domain containing signaling pathway

osteoclast development

cell death

cell death

programmed cell death

apoptotic process

regulation of apoptotic process

positive regulation of apoptotic process

negative regulation of apoptotic process

regulation of cysteine-type endopeptidase activity involved in apoptotic process

negative regulation of cysteine-type endopeptidase activity involved in apoptotic process

activation of cysteine-type endopeptidase activity involved in apoptotic process

apoptotic signaling pathway

positive regulation of apoptotic signaling pathway

intrinsic apoptotic signaling pathway

positive regulation of intrinsic apoptotic signaling pathway

intrinsic apoptotic signaling pathway in response to DNA damage

negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage

negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress

positive regulation of extrinsic apoptotic signaling pathway

negative regulation of extrinsic apoptotic signaling pathway via death domain receptors

positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway

regulation of neuron apoptotic process

positive regulation of neuron apoptotic process

negative regulation of neuron apoptotic process

apoptotic mitochondrial changes

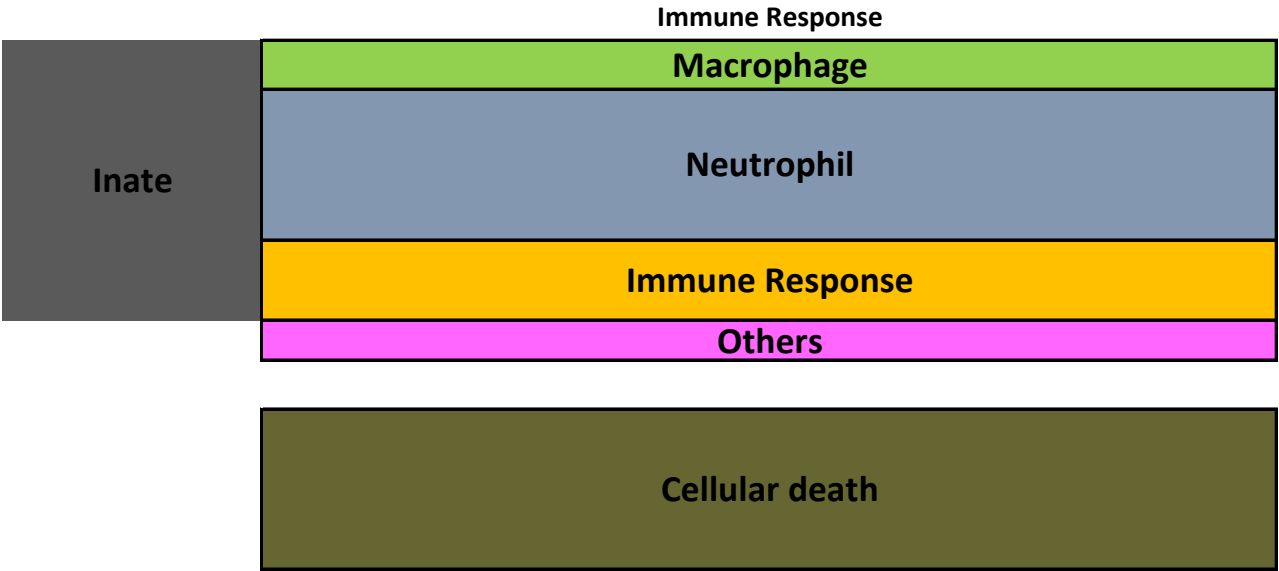
release of cytochrome c from mitochondria

regulation of apoptosis involved in tissue homeostasis

cellular component disassembly involved in execution phase of apoptosis

cardiac muscle cell apoptotic process

+		+	+	+	+			+
								+
					+			
								+
					+			
	+							
						+		
					+			



UniProtKBAC	P07305	P10145
Name	Histone H1	Interleukin
Regulation	+	+
Biomarker		x
immune system process		+
positive regulation of macrophage differentiation		
neutrophil activation		+
neutrophil chemotaxis		+
positive regulation of neutrophil chemotaxis		+
immune response		+
positive regulation of granulocyte differentiation		
erythrocyte differentiation		
cell death	+	
positive regulation of apoptotic process		
negative regulation of apoptotic process		
negative regulation of leukocyte apoptotic process		
apoptotic DNA fragmentation	+	

P14317	P28562
Hematopo	Dual specif
+	
+	
+	
+	
+	+
	+
	+
+	

Inate

Immune Response

Macrophage

Immune Response

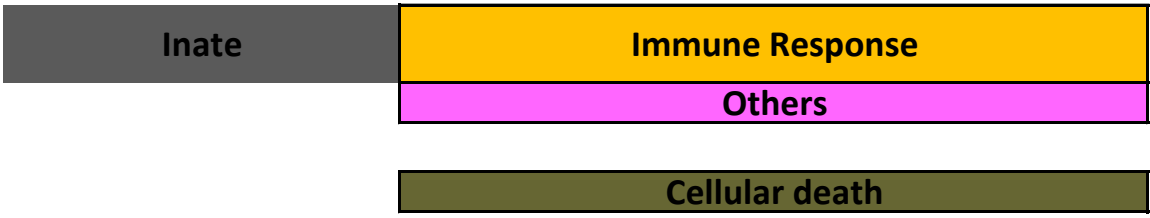
Others

Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
regulation of macrophage activation
response to interferon-gamma
microglial cell activation
cell death
negative regulation of neuron apoptotic process
activation of cysteine-type endopeptidase activity involved in apoptotic process
negative regulation of cysteine-type endopeptidase activity involved in apoptotic process

P37840	P02649
Alpha-synu	Apolipoproc
-	
+	
+	
+	
+	
+	+
+	+
+	
+	

Immune Response



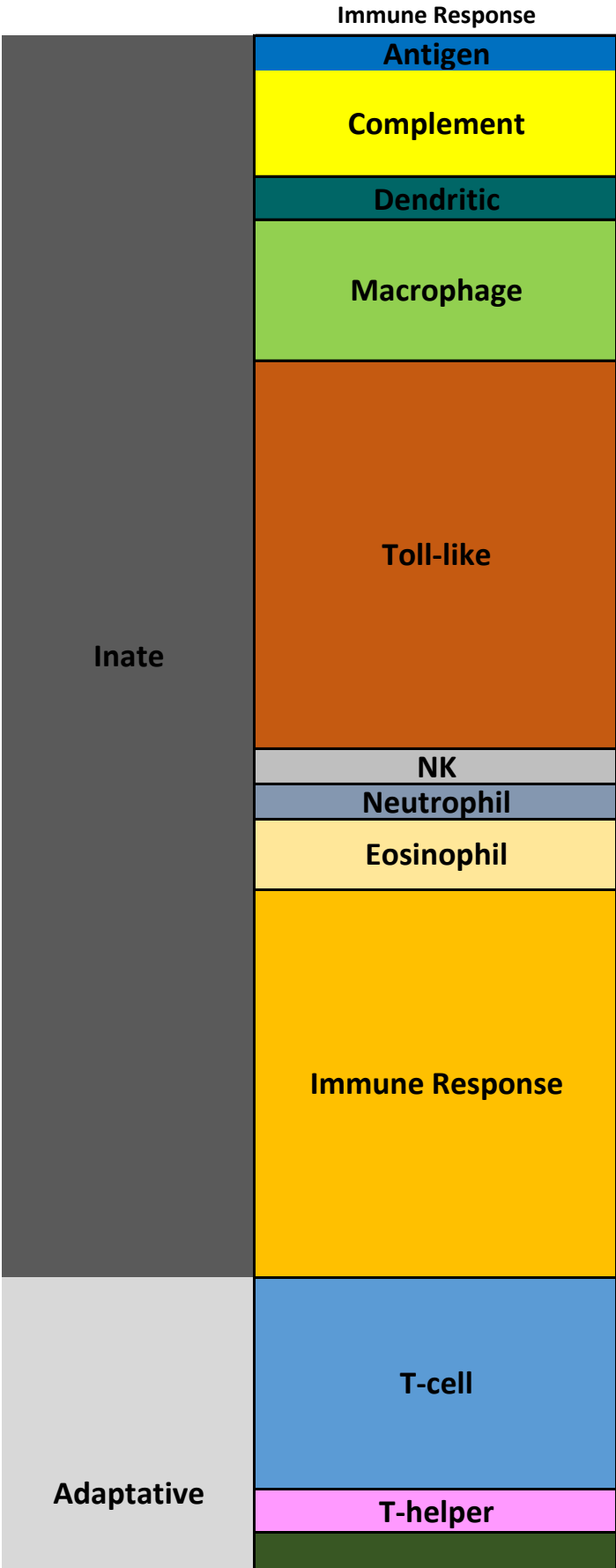
UniProtKBAC	P13010	P15559	Q9UHA7
Name	X-ray repair	NAD(P)H d	Interleukin
Regulation	+	+	+
Biomarker	x	x	x
immune system process	+		+
immune response			+
innate immune response	+		
hematopoietic stem cell differentiation	+		
cell death		+	
positive regulation of neuron apoptotic process		+	

Immune Response



UniProtKBAC	P01854	P31944
Name	Ig epsilon c	Caspase-14
Regulation	+	+
Biomarker		
immune system process	+	
immune response	+	
innate immune response	+	
Fc-epsilon receptor signaling pathway	+	
Fc-gamma receptor signaling pathway involved in phagocytosis		
cell death		+
cornification		+

P60709
Actin, cyto
+
+
+
+





B-cell

Others

Cellular death

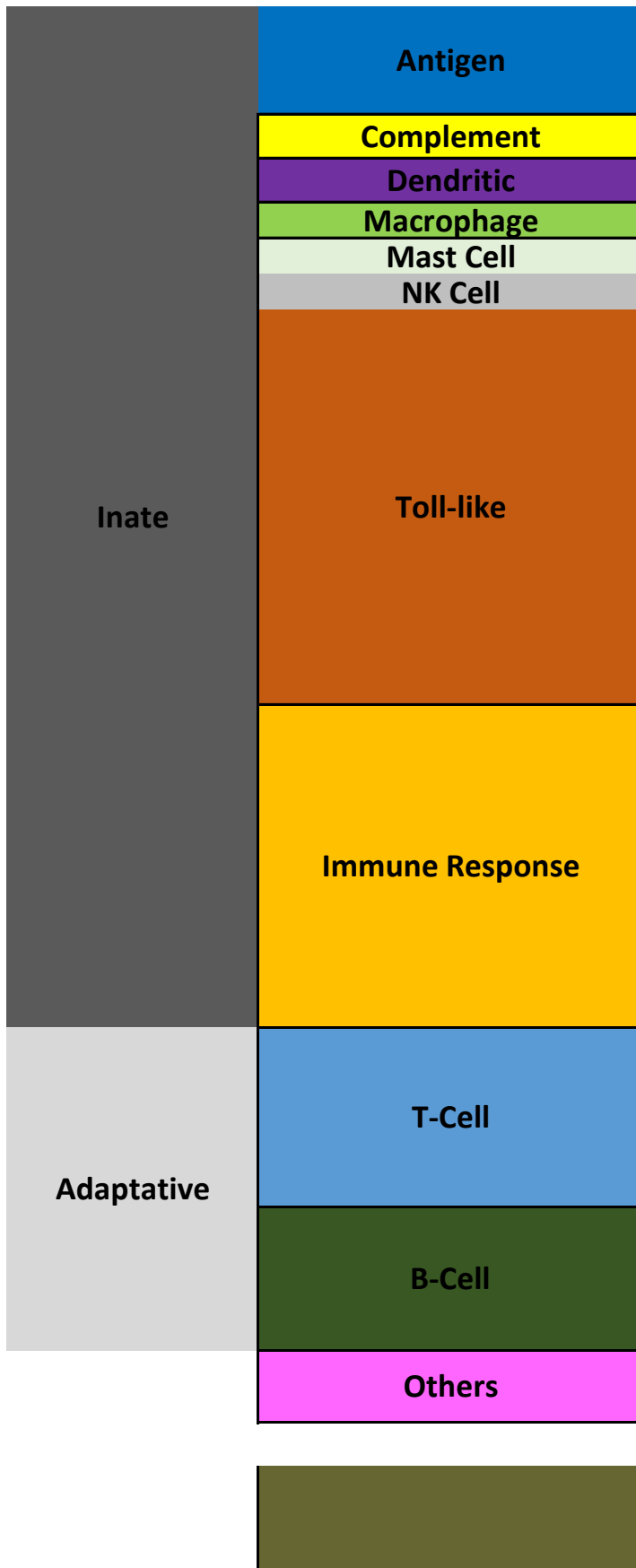
UniProtKBAC
Name
Regulation
Biomarker
immune system process
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent
complement activation, classical pathway
complement activation, lectin pathway
regulation of complement activation
positive regulation of dendritic cell chemotaxis
regulation of macrophage activation
macrophage chemotaxis
positive regulation of macrophage chemotaxis
positive regulation of immune complex clearance by monocytes and macrophages
toll-like receptor 2 signaling pathway
toll-like receptor 3 signaling pathway
toll-like receptor 4 signaling pathway
toll-like receptor 5 signaling pathway
toll-like receptor 9 signaling pathway
toll-like receptor 10 signaling pathway
toll-like receptor TLR1:TLR2 signaling pathway
toll-like receptor TLR6:TLR2 signaling pathway
TRIF-dependent toll-like receptor signaling pathway
MyD88-dependent toll-like receptor signaling pathway
MyD88-independent toll-like receptor signaling pathway
positive regulation of natural killer cell chemotaxis
neutrophil chemotaxis
eosinophil chemotaxis
eosinophil degranulation
immune system process
immune response
regulation of immune response
positive regulation of myeloid leukocyte cytokine production involved in immune response
positive regulation of prostaglandin secretion involved in immune response
innate immune response
cellular response to interferon-gamma
leukocyte migration
lymphocyte chemotaxis
monocyte chemotaxis
negative regulation of natural killer cell chemotaxis
T cell receptor signaling pathway
T cell costimulation
T cell activation involved in immune response
positive regulation of T cell activation
positive regulation of activated T cell proliferation
T cell chemotaxis
helper T cell extravasation
positive regulation of B cell proliferation

negative regulation of mature B cell apoptotic process
Fc-epsilon receptor signaling pathway
Fc-gamma receptor signaling pathway involved in phagocytosis
humoral immune response
erythrocyte maturation
granulocyte chemotaxis
negative regulation of osteoclast differentiation
nucleotide-binding oligomerization domain containing signaling pathway
osteoclast differentiation
osteoclast maturation
positive regulation of osteoclast differentiation
regulation of osteoclast differentiation
cell death
cell death
negative regulation of programmed cell death
apoptotic process
regulation of apoptotic process
positive regulation of apoptotic process
negative regulation of apoptotic process
regulation of cysteine-type endopeptidase activity involved in apoptotic process
activation of cysteine-type endopeptidase activity involved in apoptotic process by cytochrome c
cellular component disassembly involved in execution phase of apoptosis
apoptotic DNA fragmentation
apoptotic signaling pathway
negative regulation of apoptotic signaling pathway
intrinsic apoptotic signaling pathway
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator
extrinsic apoptotic signaling pathway
positive regulation of extrinsic apoptotic signaling pathway in absence of ligand
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand
positive regulation of neuron apoptotic process
negative regulation of neuron apoptotic process
positive regulation of cardiac muscle cell apoptotic process
negative regulation of striated muscle cell apoptotic process
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway
negative regulation of fibroblast apoptotic process
negative regulation of glial cell apoptotic process
negative regulation of mature B cell apoptotic process
negative regulation of myeloid cell apoptotic process
hepatocyte apoptotic process
release of cytochrome c from mitochondria

				+				
		+	+			+		
+		+	+					
				+			+	
							+	
				+				
							+	
				+				
				+				

+	+	+						
+	+	+						
								+
			+					+
								+
				+	+	+		
				+	+			
				+				
						+		
						+		

Immune Response



Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent
antigen processing and presentation of exogenous peptide antigen via MHC class II
negative regulation of T cell activation via T cell receptor contact with antigen bound to MHC molecul
complement activation, classical pathway
negative regulation of dendritic cell differentiation
negative regulation of macrophage activation
mast cell degranulation
negative regulation of natural killer cell mediated cytotoxicity
toll-like receptor 2 signaling pathway
toll-like receptor 3 signaling pathway
toll-like receptor 4 signaling pathway
toll-like receptor 5 signaling pathway
toll-like receptor 9 signaling pathway
toll-like receptor 10 signaling pathway
toll-like receptor TLR1:TLR2 signaling pathway
toll-like receptor TLR6:TLR2 signaling pathway
TRIF-dependent toll-like receptor signaling pathway
MyD88-dependent toll-like receptor signaling pathway
MyD88-independent toll-like receptor signaling pathway
immune response
innate immune response
regulation of immune response
positive regulation of defense response to virus by host
leukocyte activation
leukocyte migration
negative regulation of interferon-gamma-mediated signaling pathway
immune response-inhibiting cell surface receptor signaling pathway
negative regulation of cytokine secretion involved in immune response
T cell receptor signaling pathway
positive regulation of gamma-delta T cell activation involved in immune response
negative regulation of CD8-positive, alpha-beta T cell activation
negative regulation of T cell proliferation
negative regulation of T cell mediated cytotoxicity
positive regulation of immunoglobulin secretion
Fc receptor mediated inhibitory signaling pathway
Fc-epsilon receptor signaling pathway
Fc-gamma receptor signaling pathway involved in phagocytosis
negative regulation of osteoclast development
nucleotide-binding oligomerization domain containing signaling pathway
cell death
apoptotic process
positive regulation of apoptotic process
negative regulation of apoptotic process

apoptotic signaling pathway

positive regulation of cytolysis

negative regulation of dendritic cell apoptotic process

activation of cysteine-type endopeptidase activity involved in apoptotic process

							+	
							+	
+								

O14638	P01116	Q13409	P01762	P06309	P18135	P62987
Ectonucleo	GTPase KR	Cytoplasm	Ig heavy ch	Ig kappa ch	Ig kappa ch	Ubiquitin-6
-2.1	-4	-50	+	+	+	+
+	+	+	+	+	+	+
						+
		+				
			+	+	+	
						+
						+
						+
						+
						+
						+
						+
						+
						+
						+
						+
+			+	+	+	
	+		+	+	+	+
			+	+	+	
	+					
						+
	+		+	+	+	+
			+	+	+	
						+
						+
						+
						+
						+

						+

Immune Response

Inate

Immune Response

Cellular death

UniProtKBAC	P12838	P06865
Name	Neutrophil	Beta-hexos
Regulation	1.9	1.8
Biomarker		
immune system process	+	
innate immune response	+	
cell death		+
cell death		+

Immune Response

Antigen

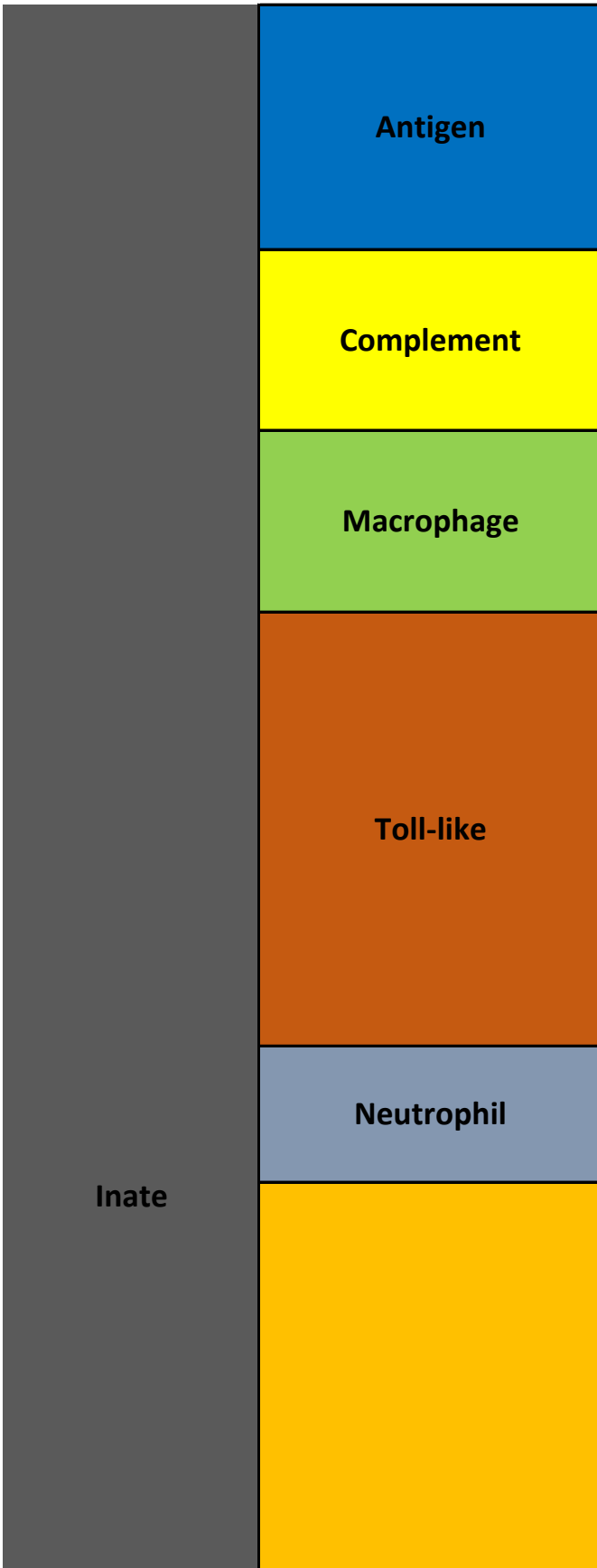
Complement

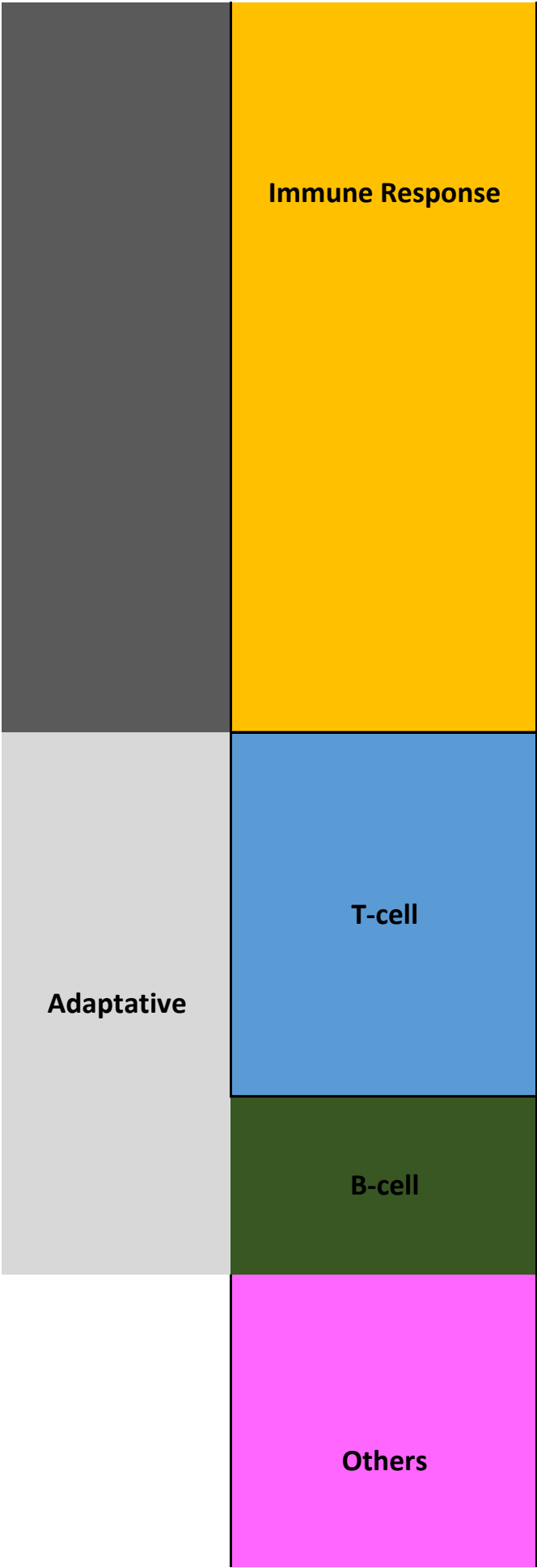
Macrophage

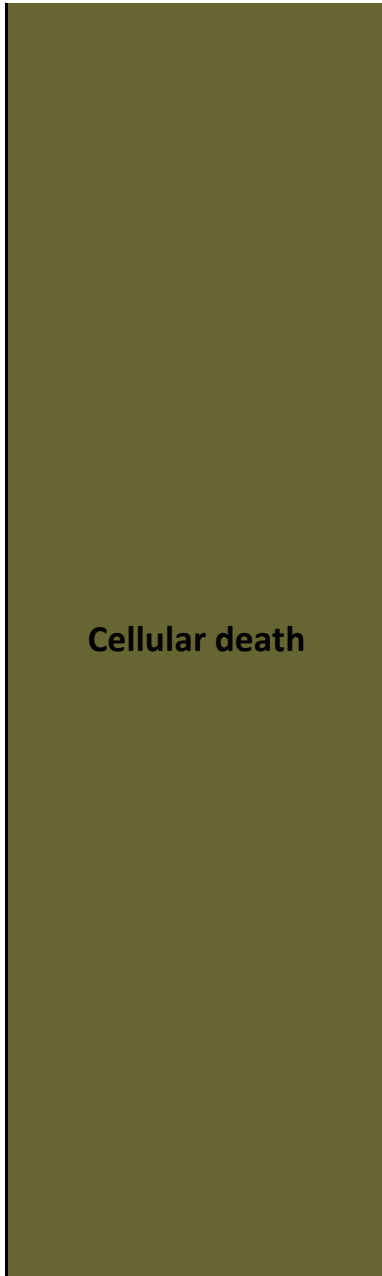
Toll-like

Neutrophil

Inate







Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
antigen processing and presentation
antigen processing and presentation of peptide antigen via MHC class I
antigen processing and presentation of endogenous peptide antigen via MHC class I
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-independent
antigen processing and presentation of exogenous peptide antigen via MHC class II
complement activation, alternative pathway
complement activation, classical pathway
complement activation, lectin pathway
regulation of complement activation
positive regulation of complement activation
macrophage differentiation
macrophage chemotaxis
negative regulation of macrophage chemotaxis
pattern recognition receptor signaling pathway
toll-like receptor signaling pathway
toll-like receptor 2 signaling pathway
toll-like receptor 3 signaling pathway
toll-like receptor 4 signaling pathway
toll-like receptor 5 signaling pathway
toll-like receptor 9 signaling pathway
toll-like receptor 10 signaling pathway
toll-like receptor TLR1:TLR2 signaling pathway
toll-like receptor TLR6:TLR2 signaling pathway
TRIF-dependent toll-like receptor signaling pathway
MyD88-dependent toll-like receptor signaling pathway
MyD88-independent toll-like receptor signaling pathway
neutrophil chemotaxis
neutrophil mediated killing of fungus
neutrophil mediated killing of gram-positive bacterium
immune system process
immune response
adaptive immune response
humoral immune response
cellular extravasation
regulation of immune response
immune response-regulating signaling pathway
negative regulation of immune response
positive regulation of immune response to tumor cell
negative regulation of cytokine secretion involved in immune response

negative regulation of natural killer cell differentiation involved in immune response
innate immune response
regulation of innate immune response
defense response to virus
regulation of defense response to virus by virus
leukocyte migration involved in inflammatory response
leukocyte chemotaxis
leukocyte mediated cytotoxicity
leukocyte migration
lymphocyte chemotaxis
lymphocyte differentiation
lymphocyte homeostasis
lymphocyte proliferation
respiratory burst involved in defense response
positive regulation of activation of membrane attack complex
regulation of interferon-gamma-mediated signaling pathway
regulation of mast cell degranulation
regulation of type I interferon-mediated signaling pathway
type I interferon signaling pathway
T cell receptor signaling pathway
T cell costimulation
T cell activation
positive regulation of T cell activation
activated T cell proliferation
T cell proliferation
positive regulation of T cell proliferation
negative regulation of T cell proliferation
negative regulation of T cell migration
T cell mediated cytotoxicity
T cell homeostasis
B cell receptor signaling pathway
negative regulation of B cell proliferation
Fc-epsilon receptor signaling pathway
Fc-gamma receptor signaling pathway involved in phagocytosis
opsonization
erythrocyte maturation
germinal center formation
granuloma formation
hemopoiesis
microglia development
mononuclear cell migration
myeloid cell differentiation
natural killer cell degranulation

nucleotide-binding oligomerization domain containing signaling pathway

osteoclast development

cell death

cell death

programmed cell death

apoptotic process

regulation of apoptotic process

positive regulation of apoptotic process

negative regulation of apoptotic process

regulation of cysteine-type endopeptidase activity involved in apoptotic process

negative regulation of cysteine-type endopeptidase activity involved in apoptotic process

activation of cysteine-type endopeptidase activity involved in apoptotic process

apoptotic signaling pathway

positive regulation of apoptotic signaling pathway

intrinsic apoptotic signaling pathway

positive regulation of intrinsic apoptotic signaling pathway

intrinsic apoptotic signaling pathway in response to DNA damage

negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage

negative regulation of intrinsic apoptotic signaling pathway in response to oxidative stress

positive regulation of extrinsic apoptotic signaling pathway

negative regulation of extrinsic apoptotic signaling pathway via death domain receptors

positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway

regulation of neuron apoptotic process

positive regulation of neuron apoptotic process

negative regulation of neuron apoptotic process

apoptotic mitochondrial changes

release of cytochrome c from mitochondria

regulation of apoptosis involved in tissue homeostasis

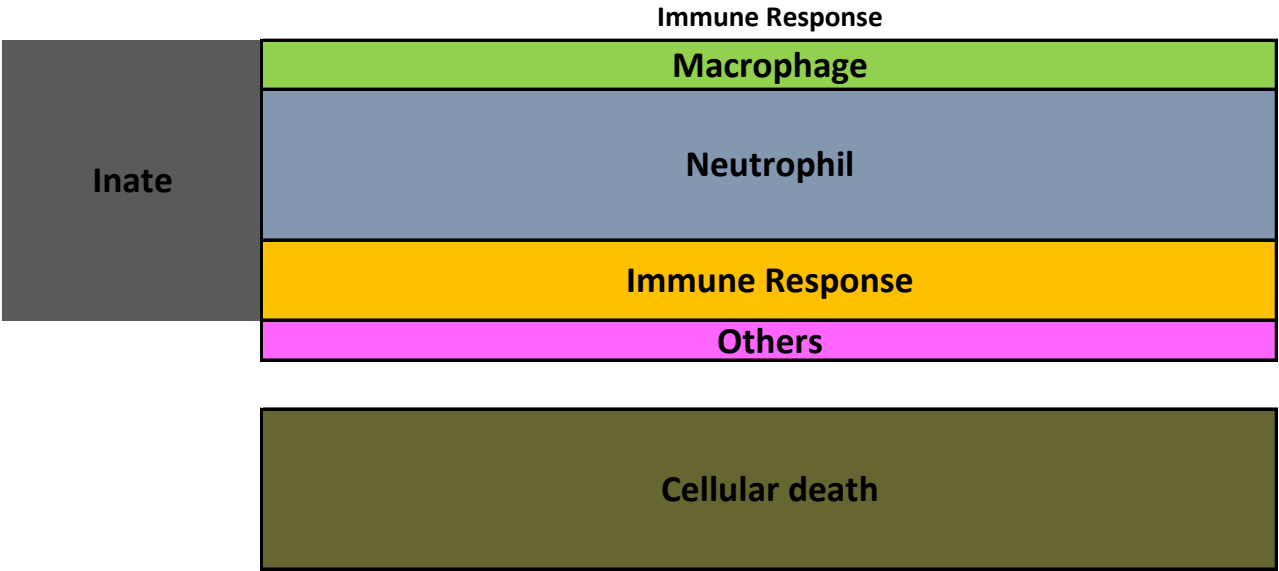
cellular component disassembly involved in execution phase of apoptosis

cardiac muscle cell apoptotic process

					+	+		
						+		
					+			
					+			
					+			
					+			
					+			

+						+	+	+
+								+
							+	
						+		

			+	
			+	
				+
		+		
		+		
			+	
		+		
		+		
		+		



UniProtKBAC	P07305	P10145
Name	Histone H1	Interleukin
Regulation	+	+
Biomarker		x
immune system process		+
positive regulation of macrophage differentiation		
neutrophil activation		+
neutrophil chemotaxis		+
positive regulation of neutrophil chemotaxis		+
immune response		+
positive regulation of granulocyte differentiation		
erythrocyte differentiation		
cell death	+	
positive regulation of apoptotic process		
negative regulation of apoptotic process		
negative regulation of leukocyte apoptotic process		
apoptotic DNA fragmentation	+	

P14317	P28562
Hematopo	Dual specif
+	
+	
+	
+	
+	+
	+
	+
+	

Inate

Immune Response

Macrophage

Immune Response

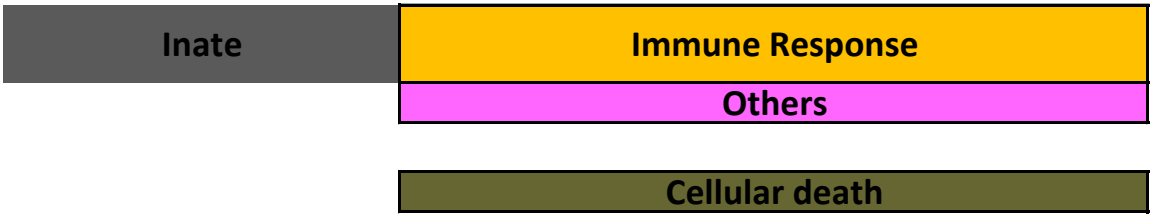
Others

Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
regulation of macrophage activation
response to interferon-gamma
microglial cell activation
cell death
negative regulation of neuron apoptotic process
activation of cysteine-type endopeptidase activity involved in apoptotic process
negative regulation of cysteine-type endopeptidase activity involved in apoptotic process

P37840	P02649
Alpha-synu	Apolipoproc
-	
+	
+	
+	
+	
+	+
+	+
+	
+	

Immune Response



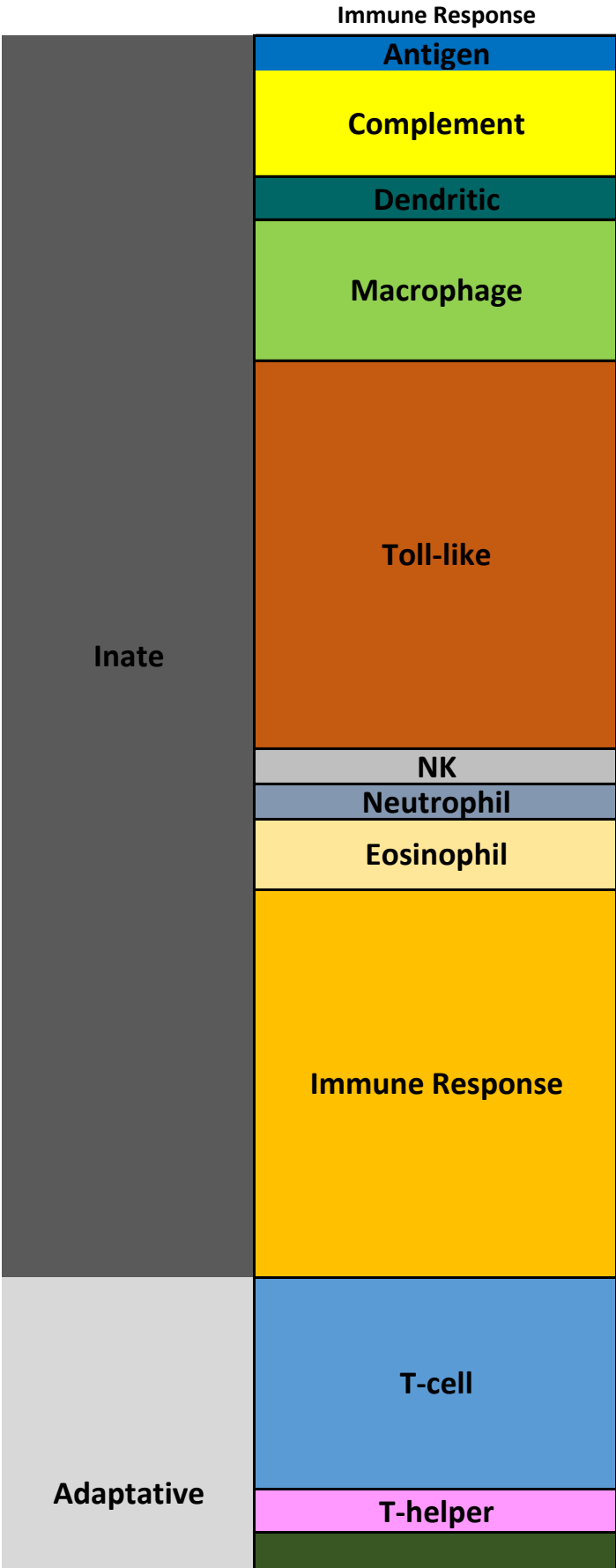
UniProtKBAC	P13010	P15559	Q9UHA7
Name	X-ray repair	NAD(P)H d	Interleukin
Regulation	+	+	+
Biomarker	x	x	x
immune system process	+		+
immune response			+
innate immune response	+		
hematopoietic stem cell differentiation	+		
cell death		+	
positive regulation of neuron apoptotic process		+	

Immune Response



UniProtKBAC	P01854	P31944
Name	Ig epsilon c	Caspase-14
Regulation	+	+
Biomarker		
immune system process	+	
immune response	+	
innate immune response	+	
Fc-epsilon receptor signaling pathway	+	
Fc-gamma receptor signaling pathway involved in phagocytosis		
cell death		+
cornification		+

P60709
Actin, cyto
+
+
+
+





B-cell

Others

Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
antigen processing and presentation of exogenous peptide antigen via MHC class I, TAP-dependent
complement activation, classical pathway
complement activation, lectin pathway
regulation of complement activation
positive regulation of dendritic cell chemotaxis
regulation of macrophage activation
macrophage chemotaxis
positive regulation of macrophage chemotaxis
positive regulation of immune complex clearance by monocytes and macrophages
toll-like receptor 2 signaling pathway
toll-like receptor 3 signaling pathway
toll-like receptor 4 signaling pathway
toll-like receptor 5 signaling pathway
toll-like receptor 9 signaling pathway
toll-like receptor 10 signaling pathway
toll-like receptor TLR1:TLR2 signaling pathway
toll-like receptor TLR6:TLR2 signaling pathway
TRIF-dependent toll-like receptor signaling pathway
MyD88-dependent toll-like receptor signaling pathway
MyD88-independent toll-like receptor signaling pathway
positive regulation of natural killer cell chemotaxis
neutrophil chemotaxis
eosinophil chemotaxis
eosinophil degranulation
immune system process
immune response
regulation of immune response
positive regulation of myeloid leukocyte cytokine production involved in immune response
positive regulation of prostaglandin secretion involved in immune response
innate immune response
cellular response to interferon-gamma
leukocyte migration
lymphocyte chemotaxis
monocyte chemotaxis
negative regulation of natural killer cell chemotaxis
T cell receptor signaling pathway
T cell costimulation
T cell activation involved in immune response
positive regulation of T cell activation
positive regulation of activated T cell proliferation
T cell chemotaxis
helper T cell extravasation
positive regulation of B cell proliferation

negative regulation of mature B cell apoptotic process
Fc-epsilon receptor signaling pathway
Fc-gamma receptor signaling pathway involved in phagocytosis
humoral immune response
erythrocyte maturation
granulocyte chemotaxis
negative regulation of osteoclast differentiation
nucleotide-binding oligomerization domain containing signaling pathway
osteoclast differentiation
osteoclast maturation
positive regulation of osteoclast differentiation
regulation of osteoclast differentiation
cell death
cell death
negative regulation of programmed cell death
apoptotic process
regulation of apoptotic process
positive regulation of apoptotic process
negative regulation of apoptotic process
regulation of cysteine-type endopeptidase activity involved in apoptotic process
activation of cysteine-type endopeptidase activity involved in apoptotic process by cytochrome c
cellular component disassembly involved in execution phase of apoptosis
apoptotic DNA fragmentation
apoptotic signaling pathway
negative regulation of apoptotic signaling pathway
intrinsic apoptotic signaling pathway
negative regulation of intrinsic apoptotic signaling pathway in response to DNA damage by p53 class mediator
extrinsic apoptotic signaling pathway
positive regulation of extrinsic apoptotic signaling pathway in absence of ligand
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand
positive regulation of neuron apoptotic process
negative regulation of neuron apoptotic process
positive regulation of cardiac muscle cell apoptotic process
negative regulation of striated muscle cell apoptotic process
positive regulation of protein insertion into mitochondrial membrane involved in apoptotic signaling pathway
negative regulation of fibroblast apoptotic process
negative regulation of glial cell apoptotic process
negative regulation of mature B cell apoptotic process
negative regulation of myeloid cell apoptotic process
hepatocyte apoptotic process
release of cytochrome c from mitochondria

				+				
		+	+			+		
+		+	+					
				+			+	
							+	
				+				
							+	
							+	
				+				
				+				
				+				

Inate

Immune Response

Immune Response

Cellular death

UniProtKBAC
Name
Regulation
Biomarker
immune system process
immune response
cell death
ectopic germ cell programmed cell death
negative regulation of extrinsic apoptotic signaling pathway in absence of ligand

P01583
Interleukin-1 Alpha
+
+
+
+
+
+