

**UNIVERSIDADE FEDERAL DO PAMPA**

**ANNA CAROLINA BOECK**

**PA<sup>3</sup>P: UMA NOVA FERRAMENTA PARA DETERMINAÇÃO *IN SILICO* DA  
ALERGENICIDADE E ANTIGENICIDADE DE PROTEÍNAS**

**São Gabriel**

**2016**

ANNA CAROLINA BOECK

**PAP: UMA NOVA FERRAMENTA PARA DETERMINAÇÃO *IN SILICO* DA ALERGENICIDADE E  
ANTIGENICIDADE DE PROTEÍNAS**

2016

**ANNA CAROLINA BOECK**

**PA<sup>3</sup>P: UMA NOVA FERRAMENTA PARA DETERMINAÇÃO *IN SILICO* DA  
ALERGENICIDADE E ANTIGENICIDADE DE PROTEÍNAS**

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Orientador: Juliano Tomazzoni Boldo

**São Gabriel**

**2016**

Ficha catalográfica elaborada automaticamente com os dados fornecidos  
pelo(a) autor(a) através do Módulo de Biblioteca do  
Sistema GURI (Gestão Unificada de Recursos Institucionais) .

B669p Boeck, Anna Carolina

PA<sup>3</sup>P: uma nova ferramenta para determinação in silico da  
alergenicidade e antigenicidade de proteínas. / Anna Carolina  
Boeck.

153 p.

Dissertação (Mestrado) -- Universidade Federal do Pampa,  
MESTRADO EM CIÊNCIAS BIOLÓGICAS, 2016.

"Orientação: Juliano Boldo".

1. Análise de proteínas. 2. Bioinformática. 3. Predição. 4.  
Transgênicos. I. Título.

ANNA CAROLINA BOECK

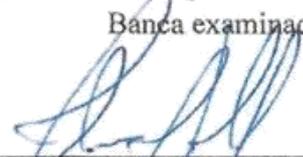
PA³P: UMA NOVA FERRAMENTA PARA DETERMINAÇÃO *IN SILICO* DA  
ALERGENICIDADE E ANTIGENICIDADE DE PROTEÍNAS

Dissertação apresentada ao Programa de Pós-Graduação *Stricto Sensu* em Ciências Biológicas da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Ciências Biológicas.

Área de concentração: Genética.

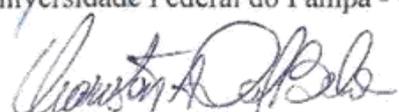
Dissertação defendida e aprovada em: 06 de abril de 2016.

Banca examinadora:



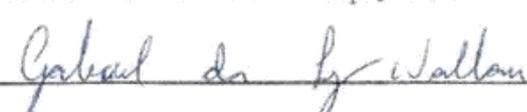
---

Prof. Dr. Juliano Tomazzoni Boldo (Orientador)  
Universidade Federal do Pampa - UNIPAMPA



---

Prof. Dr. Cháriston André dal Belo  
Universidade Federal do Pampa - UNIPAMPA



---

Dr. Gabriel da Luz Walau  
Fundação Oswaldo Cruz - FIOCRUZ



---

Prof. Dr. Jeferson Franco  
Universidade Federal do Pampa - UNIPAMPA

Dedico esta dissertação, bem como todas as minhas demais conquistas, à minha mãe Arcia e ao meu Pequeno Príncipe Gabriel. Amo vocês!

## AGRADECIMENTOS

Sinceros agradecimentos ao meu orientador, Juliano Boldo, pela orientação, dedicação, paciência e, principalmente, pela amizade.

Aos docentes do PPGCB - Unipampa pelo conhecimento proporcionado, colaborando na minha formação tanto profissional quanto pessoal;

Aos membros da Banca Examinadora, por terem atendido ao convite para desempenhar este papel, dispondo de seu tempo e conhecimento para analisar este trabalho;

Ao Bruno Dotto e ao Gabriel Walau por aceitarem o desafio de contribuir com este trabalho e torná-lo possível e pelo empenho dedicado à elaboração deste;

À minha mãe Arcia pelo incentivo e amor. Obrigada por ser meu porto seguro, me ensinar a ser forte e estar presente em todos os momentos e conquistas. Te amo!;

Ao meu irmão Leonardo e minha cunhada Patrícia pelo carinho e apoio incondicional e pelo maior presente desta vida, nosso anjo Gabriel. Serei eternamente grata por tudo que dividimos e, principalmente, por permitirem que eu faça parte da vida do Gabriel. Estaremos juntos em muitas outras conquistas que virão;

À minha segunda família, meu pais afetivos Luiz e Geni, minhas irmãs de vida Leticia e Aline, e minhas pequenas Nicolý e Sarah, agradeço pela presença e compreensão constantes;

À minha avó Elvira e aos demais familiares por apoiarem minhas escolhas e compreenderem minha ausência em muitas ocasiões;

Aos meus afilhados Ana Clara, Cassiano, Gabriel e Nicolý. Vocês me ensinaram o verdadeiro significado do *amor*. O sorriso de vocês ilumina meus dias. Amo vocês!;

Aos amigos que me acompanham desde a graduação, Anderson, Andressa, Jeferson e Priscila. Agradeço, em especial, à Pri pela companhia silenciosa entre tantas sequências a serem analisadas, pelas discussões sobre nossos trabalhos que sempre renderam grandes ideias e, simplesmente, por estar ao meu lado;

À Clarrisa, Raíssa e Renara, nosso quarteto tem inúmeras histórias para contar e ainda terá outras tantas para serem vividas. Levo vocês no coração para a vida. Parafraseando Mário Quintana “A amizade é um amor que nunca morre.”;

Às *Delícias*, Fernanda, Giseli, Janaína e Luani, por acompanharem, mesmo que de longe, mais esta etapa, por cada (re)encontro e por tantos outros que virão. Muito obrigada por animarem meus dias e por me alimentarem sempre;

Aos colegas do Laboratório APIPAMPA, em especial, Carla e Everton. Agradeço pela amizade que construímos e por tornarem esta etapa mais leve e feliz;

Aos colegas da Sala 320 pelas discussões, companhia, almoços no RU e pelos sorrisos nos dias difíceis;

Uma dissertação não se faz sozinha, ou sozinho não se faz uma dissertação. Agradeço à todos que de alguma forma contribuíram na construção e finalização deste trabalho.



## RESUMO

Uma das preocupações envolvendo a expressão de diferentes proteínas heterólogas em alimentos é a possibilidade do desenvolvimento de reações alérgicas ou antigênicas nos consumidores finais. A *Food and Agriculture Organization of the United Nations* juntamente com a *World Health Organization* definiram que uma sequência de aminoácidos seria considerada alergênica/antigênica quando apresentasse ao menos 35 % de identidade em uma janela de 80 aminoácidos ou 6 – 8 aminoácidos contíguos e idênticos quando comparada com sequências de alérgenos conhecidos. Para categorizar proteínas alergênicas ou antigênicas foi construída uma plataforma (chamada Plataforma de Análise da Alergenicidade e Antigenticidade de Proteínas – PA<sup>3</sup>P - e disponível em <http://lpa.saogabriel.unipampa.edu.br:8080/pa3p/index.jsp> ou <http://pluriserver.com.br/pa3p/>) que congrega um conjunto de ferramentas específicas para tais análises. A plataforma fora construída através da linguagem de programação Java e HTML. Desta forma, foram separados grupos de proteínas alergênicas/antigênicas de não alergênicas/não antigênicas com redução dos casos de falsos positivos ou falsos negativos. As análises complementares utilizadas neste trabalho são necessárias, pois a metodologia proposta pela *Food and Agriculture Organization of the United Nations* juntamente com a *World Health Organization* é insuficiente, gerando resultados duvidosos. A plataforma construída neste trabalho obteve valores de 98 % de sensibilidade e 96 % de especificidade, comparada com 89 % e 85 %, respectivamente, dos testes utilizados isoladamente. Esta plataforma pode ser utilizada para embasar a decisão de utilizar determinada proteína na construção de um Organismo Geneticamente Modificado com finalidade nutricional.

**Palavras-chave:** Análise de Proteínas. Bioinformática. Predição. Transgênico.

## ABSTRACT

A major drawback involving heterologous protein expression in organisms used for human consumption is the possibility of allergenic or antigenic reactions. Both Food and Agriculture Organization of the United Nations and World Health Organization determined that a potentially allergenic protein would present at least 35 % identity in a 80 amino acid window or 6 – 8 contiguous and identical amino acids when compared with known allergens. In order to assess the allergenic or antigenic potential of a given protein, the Platform for Analysis of Allergenicity and Antigenicity of Proteins (PA<sup>3</sup>P – available at <http://lpa.saogabriel.unipampa.edu.br:8080/pa3p/index.jsp> or <http://pluriserver.com.br/pa3p/>) was developed. The platform was constructed using Java and HTML programming languages. It was possible to discriminate allergenic/antigenic from non-allergenic/non-antigenic proteins, with reduction of false positive and false negative samples. The additional analyses used in the platform are necessary, since the insufficient FAO and WHO proposed methods, rendering doubtful results. In our tests, PA<sup>3</sup>P presented 98% sensibility and 96% specificity when compared with the web tools used independently (89 and 85%, respectively). This tool has a potential to contribute to decision making regarding the of a given protein in GMO construction with nutritional intent.

**Palavras-chave:** Protein analysis. Bioinformatics. Prediction. Transgenic.

## LISTA DE FIGURAS

Figura 1 - Principais técnicas e ferramentas da engenharia genética. ....	16
Figura 2 - Árvore de decisão da FAO. ....	18
Figura 3 - Diferença entre um OGM e um melhoramento genético clássico. ....	19
Figura 4 - Transformação de plantas via <i>Agrobacterium tumefaciens</i> . ....	21
Figura 5 - Transformação por biobalística. ....	22
Figura 6 – Método de transformação por microinjeção. ....	23
Figura 7 - Eletroporação de protoplastos vegetais. ....	24
Figura 8 - Eventos da avaliação de risco ambiental (ERA) pelo cultivo de um OGM. .	26
Figura 9 - Fluxograma da análise de risco de um OGM. ....	27
Figura 10 - Apresentação gráfica do símbolo transgênico para rotulagem de alimentos embalados. ....	27
Figura 11 - Classificação das Reações Adversas aos Alimentos. ....	28
Figura 12 - Eventos na reação de alergenicidade. ....	29
Figura 13 - Alinhamento de duas sequências de proteínas. ....	33
Figura 14 - Diferença entre alinhamento global e local. ....	33
Figura 15 - Determinação do conjunto de classificadores do SVM. ....	34
Figura 16 - Fluxograma de construção e funcionamento da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA <sup>3</sup> P). ....	44
Figura 17 - Interface da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA <sup>3</sup> P). ....	46
Figura 18 - Especificidade e sensibilidade da PA <sup>3</sup> P. ....	47
Figura 19 - Destaque para o campo destinado à inserção da sequência de consulta para submissão à análise na PA <sup>3</sup> P. ....	50
Figura 20 - Tabela de resultados individuais conforme ferramenta e método. ....	51
Figura 21 - Gráficos para melhor visualização e interpretação dos resultados. ....	51
Figura 22 - <i>Download</i> dos resultados gráficos. ....	52
Figura 23 - Redução dos casos de falsos positivos/negativos pelo uso da PA <sup>3</sup> P. ....	54

## LISTA DE TABELAS

Tabela 1 - Manifestações Clínicas nas Diferentes Reações Alérgicas. ....	30
Tabela 2 - Métodos <i>in vivo</i> , <i>in vitro</i> e <i>in silico</i> para definição do potencial de alergenicidade. ....	31
Tabela 3 - Principais resultados referentes ao cálculo da acurácia das <i>web tools</i> . ....	41
Tabela 4 - Planilha de pontuação da acurácia conforme o método e o resultado da análise. ....	42
Tabela 5 - Valores referentes ao erro amostral. ....	48
Tabela 6 - Matriz binária ( 0 = presença; 1 = ausência) relativa ao processo de determinação da acurácia das <i>web tools</i> e validação da PA <sup>3</sup> P. ....	69
Tabela 7 - Validação da PA <sup>3</sup> P mediante grupo de proteínas com alergenicidade desconhecidas, onde 0 = presença e 1 = ausência. ....	90

## LISTA DE ABREVIATURAS E SIGLAS

aa – aminoácido

DNA – ácido desoxirribonucleico

FA – *Food Allergy*

FAO – *Food and Agriculture Organization of the United Nations*

IgE – Imunoglobulina E

OGM – Organismo Geneticamente Modificado

PA<sup>3</sup>P – Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas

PGM – Plantas Geneticamente Modificadas

RNA – ácido ribonucleico

WHO – *World Health Organization*

## SUMÁRIO

<b>LISTA DE FIGURAS.....</b>	<b>11</b>
<b>LISTA DE TABELAS .....</b>	<b>12</b>
<b>LISTA DE ABREVIATURAS E SIGLAS .....</b>	<b>13</b>
<b>1. INTRODUÇÃO .....</b>	<b>15</b>
<b>2. REVISÃO BIBLIOGRÁFICA.....</b>	<b>19</b>
2.1 <i>ENGENHARIA GENÉTICA (TECNOLOGIA DO DNA RECOMBINANTE)</i> .....	19
2.2 <i>METODOLOGIAS DE TRANSFORMAÇÃO</i> .....	20
2.2.1 <i>Agrotransformação</i> .....	20
2.2.2 <i>Biobalística</i> .....	22
2.2.3 <i>Microinjeção em células vegetais</i> .....	23
2.2.4 <i>Eletroporação de protoplastos vegetais</i> .....	23
2.3 <i>BIOSSEGURANÇA</i> .....	24
2.4 <i>ORGANISMOS GENETICAMENTE MODIFICADOS E TRANSGÊNICOS</i> .....	25
2.5 <i>ALERGENICIDADE E ANTIGENICIDADE</i> .....	28
2.6 <i>BIOINFORMÁTICA</i> .....	32
<b>3. CONTEXTUALIZAÇÃO.....</b>	<b>36</b>
<b>4. OBJETIVOS.....</b>	<b>37</b>
<b>5. CAPÍTULO I - CONSTRUÇÃO DA PLATAFORMA DE ANÁLISE DA ALERGENICIDADE E ANTIGENICIDADE DE PROTEÍNAS .....</b>	<b>38</b>
<b>6. CAPÍTULO II – CATEGORIZAÇÃO <i>IN SILICO</i> DE PROTEÍNAS ALERGÊNICAS OU ANTIGÊNICAS POR MEIO DA PLATAFORMA DE ANÁLISE DA ALERGENICIDADE E ANTIGENICIDADE DE PROTEÍNAS (PA<sup>3</sup>P).....</b>	<b>49</b>
<b>7. DISCUSSÃO .....</b>	<b>53</b>
<b>8. CONCLUSÕES E PERSPECTIVAS .....</b>	<b>57</b>
<b>REFERÊNCIAS .....</b>	<b>58</b>
<b>APÊNDICES .....</b>	<b>64</b>
APÊNDICE 1: MANUSCRITO GERADO A PARTIR DESTE TRABALHO.....	65
APÊNDICE 2: PLANILHA DO CÁLCULO DA ACURÁCIA E VALIDAÇÃO DA PA <sup>3</sup> P VIA DATASET. ....	69
APÊNDICE 3 – PLANILHA DE VALIDAÇÃO. ....	90
APÊNDICE 4 – CÓDIGOS-FONTE DE CONSTRUÇÃO DA PA <sup>3</sup> P.....	92
4.1 <i>CÓDIGOS-FONTE REFERENTES AO ALGPRED</i> .....	92
4.2 <i>CÓDIGOS-FONTE REFERENTES AO ALLERGEN ONLINE</i> .....	95
4.3 <i>CÓDIGOS-FONTE REFERENTES AO PROAP</i> .....	98
4.4 <i>CÓDIGOS-FONTE REFERENTES À ADFS</i> .....	100
4.5 <i>CÓDIGOS-FONTE REFERENTES À SEÇÃO “ABOUT”</i> .....	103

4.6 CÓDIGOS-FONTE REFERENTES À ANÁLISE NA PA <sup>3</sup> P .....	111
4.8 CÓDIGOS-FONTE REFERENTES À SEÇÃO “DATASET” .....	125
4.9 CÓDIGOS-FONTE REFERENTES À “HOME” DA PA <sup>3</sup> P .....	126
4.10 CÓDIGOS-FONTE REFERENTES À SEÇÃO “CONTACT” .....	128
4.11 CÓDIGOS-FONTE REFERENTES À SEÇÃO “USEFULL LINKS” .....	129
4.12 CÓDIGOS-FONTE REFERENTES AO CABEÇALHO DA PA <sup>3</sup> P .....	130
4.13 CÓDIGOS-FONTE REFERENTES AO MENU DAS SEÇÕES DA PA <sup>3</sup> P .....	133
4.14 CÓDIGOS-FONTE REFERENTES AOS RESULTADOS DA ADFS IMPRESSOS NA PA <sup>3</sup> P ...	134
4.15 CÓDIGOS-FONTE REFERENTES AOS RESULTADOS DO ALGPRED IMPRESSOS NA PA <sup>3</sup> P .....	135
4.16 CÓDIGOS-FONTE REFERENTES AOS RESULTADOS DO ALLERGEN ONLINE IMPRESSOS NA PA <sup>3</sup> P .....	137
4.17 CÓDIGOS-FONTE REFERENTES AOS GRÁFICOS IMPRESSOS PELA PA <sup>3</sup> P .....	139
4.18 CÓDIGOS-FONTE REFERENTES À IMPRESSÃO DOS RESULTADOS NA PA <sup>3</sup> P .....	142
4.19 CÓDIGOS-FONTE REFERENTES AOS RESULTADOS DO PROAP IMPRESSOS NA PA <sup>3</sup> P..	153



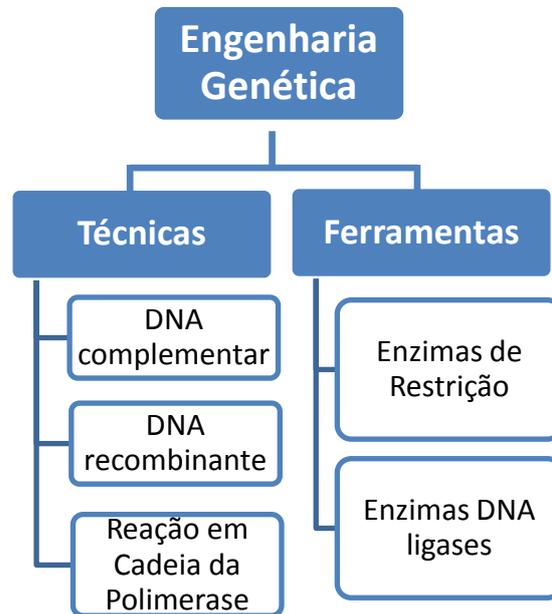
## 1. Introdução

O termo alergia pode ser definido como uma resposta imunológica anormal (SICHERER; SAMPSON, 2010). Na alergia alimentar (do inglês *food allergy* – FA) esta resposta anormal ocorre pela exposição a uma proteína dietética (WASERMAN; WATSON, 2011) uma vez que algumas famílias de proteínas possuem capacidade de gerar alergias, ou seja, possuem potencial alergênico (ROTH-WALTER, 2014). Inúmeros alimentos podem desencadear uma resposta imunológica adversa, destacando-se como principais alérgenos proteínas do leite, ovo, amendoim, nozes, mariscos, peixe, trigo e soja (SICHERER; SAMPSON, 2010).

De acordo com Solé et al. (2008) cerca de 3,5 % da população apresenta alergia alimentar. A FA atinge de 3 a 6 % das crianças e por este motivo é considerada um problema de saúde pública em diversos países (TOIT et al., 2016). Dados dos Centros de Controle e Prevenção de Doenças indicam um aumento de 18 % nos casos de alergia alimentar em crianças entre os anos de 1997 e 2007 (SICHERER; SAMPSON, 2010) e estes valores estão aumentando (SOLÉ et al., 2008). A introdução de novas proteínas em alimentos, medicamentos e outros produtos podem expor os consumidores a novos riscos (MARI et al., 2009), como, por exemplo, o risco de reação alérgica e toxicidade (COSTA et al., 2011).

A inserção de genes que codificarão novas proteínas ocorre através de técnicas de engenharia genética. O termo engenharia genética refere-se ao uso de técnicas capazes de inserir fragmentos de ácido desoxirribonucleico (DNA), geralmente genes, de um organismo a outro, mudando, portanto, a constituição genética do receptor (SHRADER-FRECHETTE, 2005). As principais técnicas e ferramentas utilizadas na engenharia genética são apresentadas na Figura 1. Dentro deste panorama, a utilização de tais técnicas permite a construção de transgênicos, um tipo de Organismo Geneticamente Modificado (OGM). Diferentemente das técnicas tradicionais de melhoramento genético, a engenharia genética permite um maior controle sobre a alteração pretendida. Apenas o gene de interesse é inserido no organismo (HUG, 2008). Assim, podem ser construídos OGM expressando proteínas que confirmam vantagens econômicas, como vegetais mais resistentes a pragas ou a herbicidas (VON KRIES; WINTER, 2011).

Figura 1 - Principais técnicas e ferramentas da engenharia genética.



Fonte: adaptado de <http://oportaldomundobiologico.blogspot.com.br/2011/04/engenharia-genetica-tecnicas.html>. Acesso em 04/01/2016.

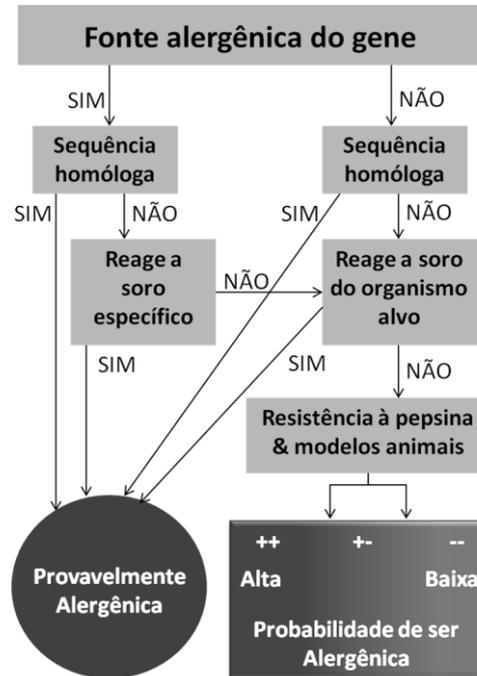
Diferente de um OGM que tem seu genoma modificado, um transgênico é submetido à alguma metodologia de transformação para que um trecho do DNA de outra espécie seja introduzido. Em suma, todo transgênico pode ser considerado um OGM, mas nem todo OGM constitui um transgênico.

O maior impasse para a ampla comercialização de alimentos transgênicos é garantir a sua segurança ambiental e a confiabilidade do consumidor nos produtos finais, uma vez que a modificação genética de tais organismos leva a expressão de proteínas antes ausentes. Para minimizar um possível impacto negativo devem ser analisados os possíveis efeitos sobre a segurança alimentar bem como aqueles relacionados com o meio ambiente (*Food and Agriculture Organization of the United Nations – FAO/ World Health Organization - WHO*, 2001). Os principais perigos potenciais dos OGM podem estar associados com toxicidade, alergenicidade e antigenicidade, alterações nutricionais e efeitos antinutrientes e possibilidade remota de transferência horizontal de genes (COSTA et al., 2011) sendo que todos estes perigos devem ser testados antes da liberação do organismo transgênico e monitorados após a liberação para comercialização. No Brasil, a Lei de Biossegurança (Lei nº 11.105 de 24 de

março de 2005) foi aprovada para normatizar a pesquisa, o manuseio e o uso de transgênicos (BRASIL, 2005).

Para minimizar um possível impacto negativo devem ser analisados os possíveis efeitos sobre a segurança alimentar bem como aqueles relacionados com o meio ambiente (*Food and Agriculture Organization of the United Nations – FAO/ World Health Organization - WHO*, 2001). Assim, a avaliação do potencial de alergenicidade e antigenicidade das proteínas codificadas pelos genes inseridos é imprescindível, uma vez que não é facilmente previsível, sem estudos específicos. Para resolver esta questão, a FAO e a WHO, em 2001, determinaram testes que devem ser realizados para identificar características que definem o potencial alergênico e antigênico de uma nova proteína introduzida em alimentos geneticamente modificados. Também foi criada, pela FAO (2001) uma árvore de decisão para a avaliação do potencial de alergenicidade de proteínas introduzidas no OGM (KLETER; KUIPER, 2002). De acordo com a árvore de decisão da FAO (Figura 2), primeiramente, deve-se comparar a estrutura da nova proteína com as estruturas de alérgenos conhecidos através do alinhamento das sequências. O conhecimento acerca da origem do gene é de extrema importância, pois caso esta seja de um alérgeno conhecido a sua reação com soros de pacientes que são alérgicos à fonte específica deve ser testada. Testes adicionais como, por exemplo, a digestão *in vitro* com pepsina, podem ser necessários. A FAO/WHO (2001) indica que testes adicionais devem ser realizados quando não houver nenhuma homologia de sequência com um alérgeno conhecido. O teste de resistência à pepsina é realizado quando o resultado para o teste de reação a soro específico é negativo (FAO/WHO, 2001). O teste de digestão *in vitro* com pepsina baseia-se na premissa de que a maioria dos alérgenos alimentares são estáveis à digestão (KLETER; KUIPER, 2002). Entretanto, deve-se considerar que a promoção ou a redução da degradação de proteínas podem estar relacionadas à presença de inibidores de proteases ou outras substâncias. Se qualquer uma destas etapas produzir um resultado positivo, a proteína deve ser considerada passível de ser alergênica (FAO/WHO, 2001).

Figura 2 - Árvore de decisão da FAO.



Fonte: adaptado de Kleter & Kuiper (2002).

A verificação destas características isoladas não é suficiente e por vezes apresenta resultados duvidosos. Assim, a utilização de vários algoritmos de avaliação torna-se necessária (KLETER; KUIPER, 2002). Neste contexto, alguns algoritmos foram desenvolvidos e hoje são amplamente utilizados, o que torna as avaliações mais robustas. Contudo, a análise por meio destes algoritmos é dificultada por dois problemas principais: (i) dificuldades como, por exemplo, a exigência de registro para ter acesso aos *sites* que hospedam tais algoritmos, muitas vezes com interface pouco amigável; e, (ii) resultados em gráficos ou tabelas não autoexplicativos e de qualidade gráfica baixa, comprometendo a análise realizada pelo leigo em bioinformática. O desenvolvimento de uma plataforma que congregue tais análises, somadas aos testes sugeridos pela FAO/WHO, pouparia tempo de análise por proteína avaliada e ofereceria ao usuário um resultado mais compreensível e confiável.

Posto isso, este trabalho objetivou a construção de uma plataforma.

## 2. Revisão Bibliográfica

### 2.1 Engenharia Genética (tecnologia do DNA recombinante)

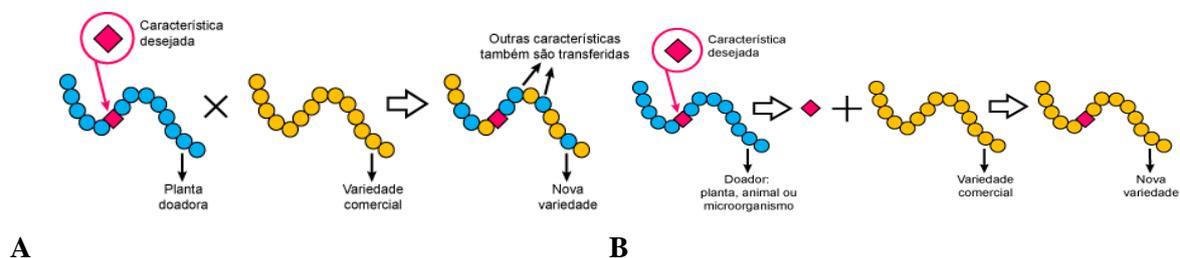
Cordeiro (2003) conceitua a engenharia genética como um conjunto de técnicas que possibilitam avanços em diferentes campos, incluindo modificações do material genético, seja ele DNA ou ácido ribonucleico (RNA). Modificações em sequências genômicas são realizadas através de recombinações homólogas pela engenharia genética (ROCHA-MARTINS et al., 2015). Por meio destas recombinações originam-se DNA recombinantes e por essa razão a engenharia genética é comumente chamada de “tecnologia do DNA recombinante” (ALBERTS et al., 2010).

Na década de 1970 foram descobertas algumas enzimas de restrição e realizados os primeiros ensaios de clonagem. Em 1977 foi desenvolvida a técnica de sequenciamento de DNA (CORDEIRO, 2003).

Dentre as técnicas da tecnologia do DNA recombinante pode-se destacar como técnicas-chave: (i) clivagem do DNA em sítios específicos; (ii) ligação de DNA; (iii) clonagem de DNA via vetores de clonagem; dentre outras (ALBERTS et al., 2010). A utilização de tais técnicas permite a construção de Organismos Geneticamente Modificados (OGM), comumente chamados de transgênicos (FALEIRO; ANDRADE, 2009).

A construção de OGM através de técnicas de engenharia genética torna o processo de manipulação genética mais eficaz (CANHOTO, 2010), ou seja, apenas o gene que possui a característica desejada é transferido (AMARAL; SILVA, 2003). Esta diferença quando comparada a engenharia genética com o melhoramento genético clássico pode ser observada na Figura 3.

Figura 3 - Diferença entre um OGM e um melhoramento genético clássico.



Em (A) tem-se um esquema de como ocorre no melhoramento genético clássico e em (B) como um organismo geneticamente modificado é produzido.

Fonte: adaptado de <http://ciencia.hsw.uol.com.br/transgenicos3.htm>. Acesso em 21/01/2016.

Segundo Faleiro e Andrade (2009) a produção de plantas geneticamente modificadas (PGM) acarretou em novas possibilidades na agricultura beneficiando tanto produtor como consumidor.

Em suma, a engenharia genética permite a seleção de determinada sequência nucleotídica de uma espécie e sua posterior associação ao DNA de outra espécie. A inserção do gene selecionado ocorre por transformação genética (AMARAL; SILVA, 2003).

## 2.2 Metodologias de Transformação

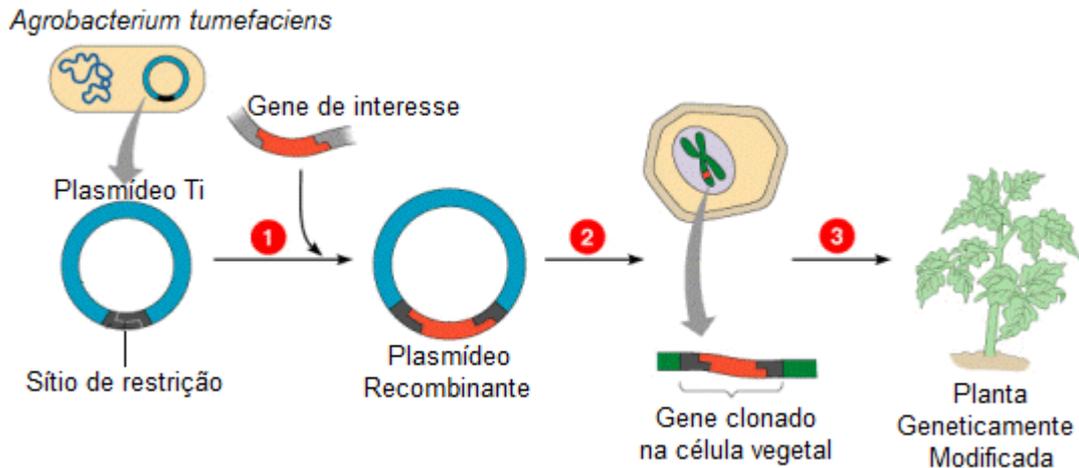
Através das técnicas modernas de transformação genética é possível inserir genes de interesse de outros organismos em plantas (CANHOTO, 2010). A transformação indireta utiliza-se de um vetor para realizar a transferência de DNA e a transformação direta utiliza métodos químicos ou físicos para transferir o DNA (ANDRADE, 2003). A seguir são apresentadas as principais metodologias de transformação utilizadas para obter transgênicos.

### 2.2.1 Agrotransformação

A transformação via *Agrobacterium*, ou agrotransformação, é um método de transferência indireta (ANDRADE, 2003) que realiza a transformação genética em plantas por meio de uma bactéria gram-negativa chamada *Agrobacterium* spp. (CANHOTO, 2010), tipicamente do solo (BRASILEIRO; LACORTE, 2000). As espécies deste gênero induzem a formação de tumores nas plantas, chamados tumores da galha-da-coroa (*crown gall*), reflexo direto da transferência de genes específicos, e por esta razão são utilizadas na engenharia genética (RAVEN; EVERT; EICHHORN, 1996). Nesta doença, as células vegetais infectadas pela bactéria possuem a capacidade de se multiplicarem sem a necessidade de estímulos externos (BRASILEIRO; LACORTE, 2000).

A espécie *Agrobacterium tumefaciens* induz a formação de tumores pela transferência de uma região específica (região T ou T-DNA) de um plasmídeo indutor de tumor (Ti) no DNA nuclear do hospedeiro (RAVEN; EVERT; EICHHORN, 1996) através da região de virulência (região *vir*) (BRASILEIRO; LACORTE, 2000). A presença do plasmídeo Ti está diretamente relacionada à capacidade que as agrobactérias têm de infectar células vegetais (BRASILEIRO; LACORTE, 2000). A Figura 4 apresenta a metodologia proposta na agrotransformação.

Figura 4 - Transformação de plantas via *Agrobacterium tumefaciens*.



Em 1 ocorre a inserção do gene de interesse no plasmídeo por enzimas de restrição e DNA ligase, dando origem a um plasmídeo Ti recombinante. Este plasmídeo Ti recombinante é transferido para uma célula de *A. tumefaciens*. O *A. tumefaciens* contendo o plasmídeo modificado é utilizado para infectar uma planta(2). Para que a infecção ocorra é necessário que haja algum ferimento no tecido vegetal por onde a agrobactéria possa reconhecer a planta e iniciar a transferência do T-DNA do plasmídeo Ti. Desta forma, o T-DNA é transferido da bactéria para a célula vegetal onde será integrado ao genoma da planta. Por fim, ocorre a regeneração da planta que agora expressa novas características (3).

Fonte: adaptado de <http://transgeniaemvegetais.blogspot.com.br/2010/08/tecnicas-de-transformacao-de-dna.html>. Acesso em 21/01/2016.

O gênero *Agrobacterium* possui 5 espécies que diferem entre si pela patogenicidade e o modo de infecção em diferentes plantas (BRASILEIRO; LACORTE, 2000). A espécie *Agrobacterium rhizogenes* é utilizada quando o princípio ativo a ser transformado encontra-se nas partes subterrâneas da planta (AMARAL; SILVA, 2003) uma vez que é o agente etiológico da raiz em cabeleira (*hairy root*) (BRASILEIRO; LACORTE, 2000). Já para princípios ativos localizados nas partes aéreas utiliza-se preferencialmente a espécie *A. tumefaciens* (AMARAL; SILVA, 2003). Cabe ressaltar que esta condição não constitui uma regra, ou seja, os sistemas podem ser utilizados independentes da localização do princípio ativo.

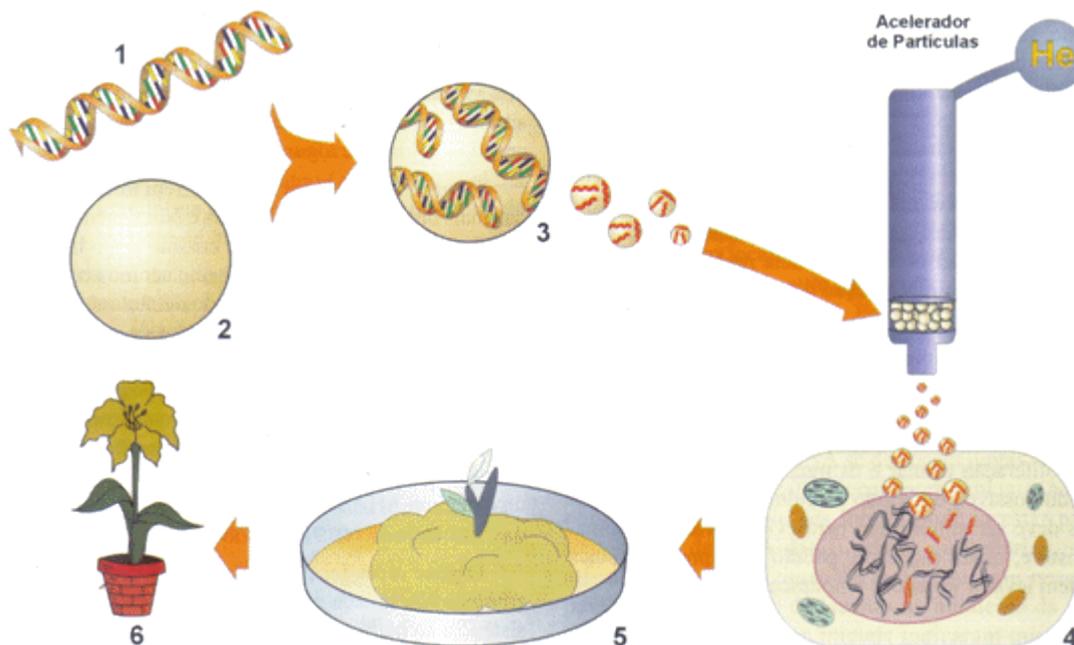
Dentre as vantagens da técnica destaca-se a transmissão dos novos genes às progênes e como desvantagem pode-se citar a eficiência de transformação, que varia drasticamente para diferentes espécies vegetais (ANDRADE, 2003).

### 2.2.2 Biobalística

A biobalística, também conhecida como bombardeamento de (micro) partículas, foi desenvolvida por Klein et al. (1987 apud NONOHAY, 2002) e Sanford et al. (1987 apud SARTORETTO; SALDANHA; CORDER, 2008).

Este método de transformação direta consiste no bombardeamento de micropartículas com DNA exógeno aderido (ANDRADE, 2003). Desta forma, moléculas de DNA são introduzidas em células e tecidos intactos (NONOHAY, 2002); Este mecanismo busca ultrapassar as principais barreiras da célula vegetal, a parede celular e a membrana plasmática (ANDRADE, 2003). O protocolo do método é apresentado na Figura 5.

Figura 5 - Transformação por biobalística.



Nesta técnica, os fragmentos de DNA exógeno (1) são aderidos à micropartículas metálicas (2) e (3). Estas micropartículas são aceleradas (~1.200 km/h) em direção à célula vegetal (4). Quando dentro da célula, o DNA exógeno se desliga das micropartículas e, eventualmente, se integra ao DNA nuclear. Por fim, as células transformadas são selecionadas (5) e regeneradas em plantas geneticamente modificadas (6).

Fonte: adaptado de <http://www.cgm.icb.ufmg.br/oquesao/figura05.gif>. Acesso em 21/01/2016.

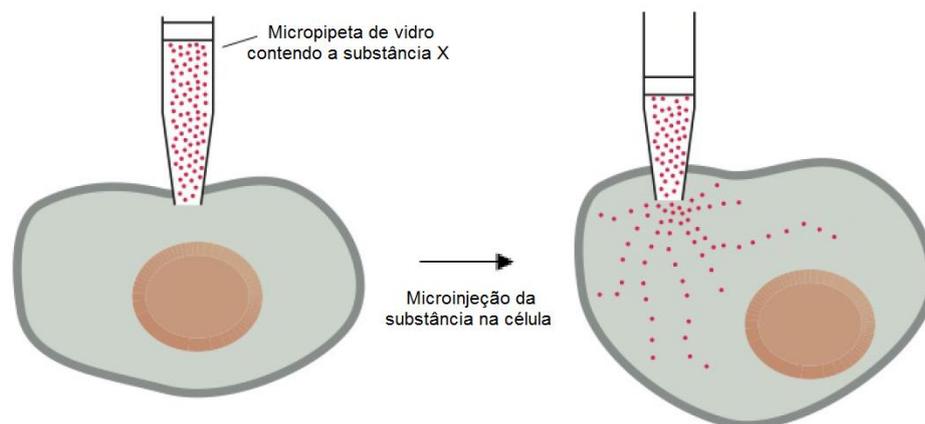
Existem diferentes aparatos para bombardear as células vegetais com as micropartículas. A aceleração pode ocorrer por diferentes mecanismos como, por exemplo, pela aceleração de gás hélio ou por descarga elétrica, dentre outros mecanismos (NONOHAY, 2002). Contudo, é necessário que o tecido a ser bombardeado apresente potencial de regeneração para que seja transformado através desta técnica (SARTORETTO; SALDANHA; CORDER, 2008).

A grande vantagem deste método é a transformação sem limitação de hospedeiro (NONOHAY, 2002) de forma direta, simples e rápida (ANDRADE, 2003). Em contrapartida, este método é limitado por aspectos técnicos e pela resistência à penetração das partículas que alguns tecidos apresentam (NONOHAY, 2002)

### 2.2.3 Microinjeção em células vegetais

A microinjeção é uma metodologia de transformação direta (NONOHAY, 2002). Nesta técnica as células podem estar intactas ou desprovidas total ou parcialmente de suas paredes celulares (ANDRADE, 2003). Contudo, há a necessidade de que cada célula seja injetada individualmente (ALBERTS et al., 2010). Desta forma, macromoléculas são introduzidas de forma precisa dentro da célula (ANDRADE, 2003). A técnica de microinjeção é ilustrada na Figura 6.

Figura 6 – Método de transformação por microinjeção.



Fonte: adaptado de Alberts et al. (2010).

A principal vantagem desta técnica de transformação é que a introdução da macromolécula ocorre de forma precisa dentro da célula e até o volume introduzido pode ser controlado. Contudo, este método é exaustivo e não é indicado para um número grande de células a serem transformadas (ANDRADE, 2003). Este método foi desenvolvido para transformação em animais e, posteriormente foi adaptado a plantas.

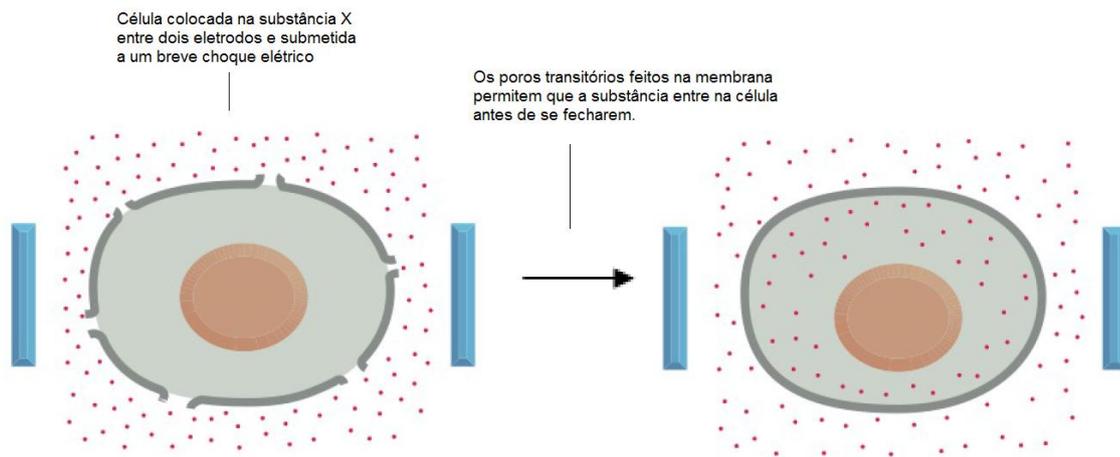
### 2.2.4 Eletroporação de protoplastos vegetais

Este método de transformação direta faz uso do protoplasto vegetal (ANDRADE, 2003). Raven, Evert e Eichhorn (1996) definem protoplasto como uma unidade do

protoplasma constituído por um citoplasma e um núcleo, ou seja, são células vegetais que não possuem parede celular.

Nesta técnica, um choque elétrico, ou um composto químico como, por exemplo, baixas concentrações de detergente, torna a membrana celular temporariamente permeável permitindo que o DNA exógeno entre no citoplasma (ALBERTS et al., 2010). Esta técnica é apresentada na Figura 7.

Figura 7 - Eletroporação de protoplastos vegetais.



Fonte: adaptado de Alberts et al. (2010).

O principal limitante desta técnica refere-se à obtenção dos protoplastos bem como a regeneração do material. Por outro lado, nesta transformação, moléculas de diferentes tamanhos são passíveis de serem introduzidas nos protoplastos (ANDRADE, 2003).

### 2.3 Biossegurança

Juntamente com o advento da biotecnologia moderna surgiu a preocupação com os riscos relacionados à, por exemplo, experimentos laboratoriais e seus possíveis impactos ambientais (SCHOLZE, 1999). A palavra biossegurança está atrelada à ideia de minimizar ou eliminar estes possíveis riscos, não sendo limitada apenas ao campo da biotecnologia moderna.

A FAO define biossegurança como um conjunto de medidas destinadas a evitar os riscos, para a saúde humana e para a conservação ambiental, advindos da prática de diferentes tecnologias aplicadas em pesquisa e práticas comerciais (FAO/WHO, 2001).

No Brasil, em 1995 foi criada a Lei de Biossegurança (Lei nº 8.974) para regulamentar as pesquisas na área da biotecnologia (BRASIL, 1995). A criação desta Lei levou à criação da

Comissão Técnica Nacional de Biossegurança (CTNBio), responsável pelo que tange o uso e a liberação de organismos geneticamente modificados no ambiente (CASTRO, 1998). Em 2005, a Lei de Biossegurança nº 8.974 de 1995 foi revogada pela Lei nº 11.105 de 24 de março de 2005 (BRASIL, 2005). Esta Lei foi aprovada para normatizar a pesquisa, o manuseio e o uso de OGM. Desta forma, uma instituição só poderá desenvolver atividades com OGM se obtiver o Certificado de Qualidade em Biossegurança (CQB) concedido pela CTNBio (SCHOLZE, 1999). Além de emitir os certificados, compete à CTNBio avaliar os requerimentos de liberação planejada no meio ambiente de um OGM (BRASIL, 2005). Uma instituição que utiliza técnicas de engenharia genética ou realiza pesquisas com OGM deve solicitar o CQB, sendo inicialmente necessário a criação de uma Comissão Interna de Biossegurança (CIBio) (BRASIL, 2006<sup>a</sup>).

A Resolução Normativa nº 2, de 27 de novembro de 2006, dispõe sobre a classificação de riscos de OGM e os níveis de biossegurança a serem aplicados. De acordo com o Art. 7º, os OGM serão classificados em quatro classes de risco de acordo com os seguintes critérios: (i) potencial patogênico dos organismos doador e receptor; (ii) a sequência nucleotídica transferida; (iii) a expressão desta no organismo receptor; e, (iv) o OGM resultante e seus efeitos adversos à saúde humana e animal, ao vegetais e ao meio ambiente (BRASIL, 2006<sup>b</sup>). As classes de risco previstas pela Resolução Normativa nº 2/2006 são: (i) Classe de Risco 1: baixo risco individual e baixo risco para a coletividade; (ii) Classe de Risco 2: moderado risco individual e baixo risco para a coletividade; (iii) Classe de Risco 3: alto risco individual e risco moderado para a coletividade; e, (iv) Classe de Risco 4: alto risco individual e alto risco para a coletividade (BRASIL, 2006<sup>b</sup>).

Tais medidas juntamente com a análise de riscos de produtos biotecnológicos, incluindo OGM, e mecanismos de monitoramento são essenciais para garantir a segurança de tais produtos (NODARI; GUERRA, 2003). Responder a estas inquietações é o contexto da biossegurança (ANDRADE; PARROTT; ROCA, 2012).

#### *2.4 Organismos Geneticamente Modificados e Transgênicos*

Um organismo geneticamente modificado (OGM) é aquele que, através da engenharia genética, foi submetido à adição ou subtração de um gene (BROWN, 2010). Para a Lei de Biossegurança vigente no Brasil, um OGM é aquele organismo cujo material genético (DNA ou RNA) tenha sido modificado por qualquer técnica de engenharia genética (BRASIL,

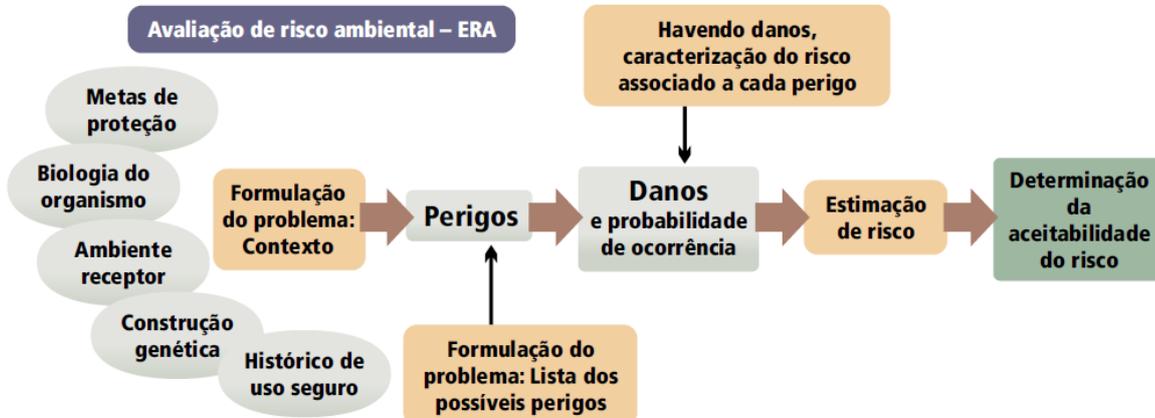
2005). A partir da interpretação da Lei se conclui que a denominação OGM é utilizada como sinônimo de transgênico no Brasil.

Além das metodologias de transformação outras etapas são de suma importância para a construção de um organismo geneticamente modificado como, por exemplo, o isolamento e a caracterização do gene de interesse (FALEIRO; ANDRADE, 2009).

As primeiras plantas transgênicas foram criadas no início da década de 1980 (CANHOTO, 2010), mas só em 1994 foi comercialmente cultivada a primeira variedade geneticamente modificada nos Estados Unidos (BORÉM, 2005).

Uma variedade geneticamente modificada que possivelmente será utilizada como alimento, antes de sua liberação deverá passar por uma análise acerca de seu histórico como cultivar e sua relação com a saúde humana incluindo, por exemplo, informações sobre toxicidade e alergenicidade (FAO/WHO, 2001). Esta avaliação de risco considera os potenciais perigos associados a possíveis impactos negativos (ANDRADE; PARROTT; ROCA, 2012). A Figura 8 ilustra os eventos referentes à avaliação de risco ambiental pelo cultivo de um OGM.

Figura 8 - Eventos da avaliação de risco ambiental (ERA) pelo cultivo de um OGM.

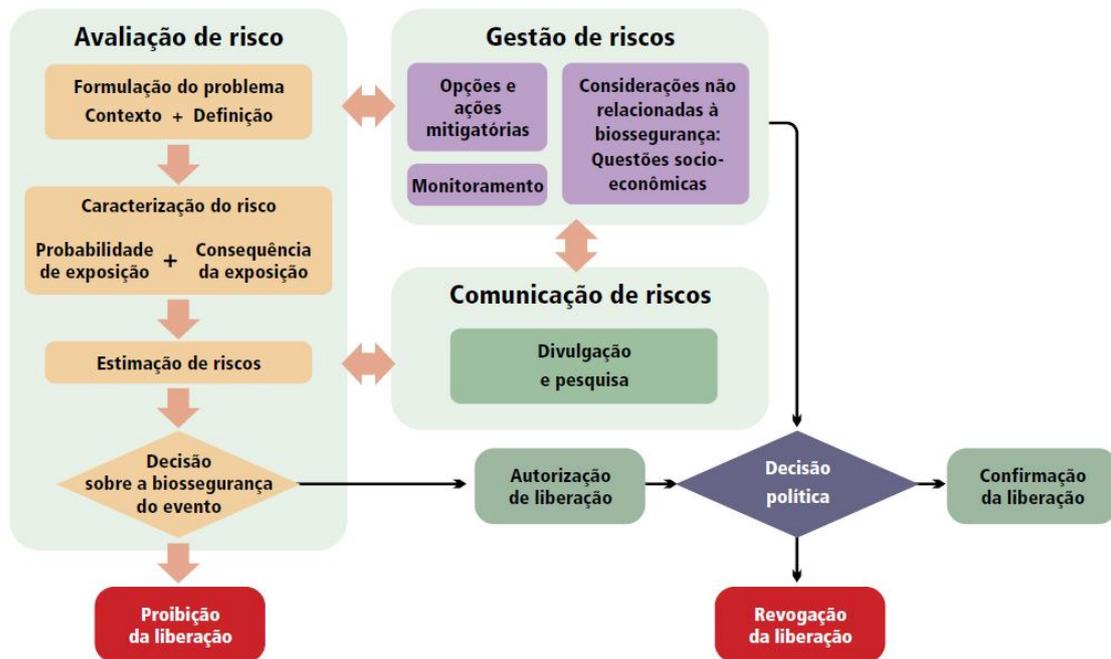


A ERA inicia com a formulação do problema. Posteriormente, são identificados a probabilidade de ocorrência e o dano potencial associado a cada novo perigo. Perigos improváveis ou que não são novos são descartados. Por conseguinte, o risco de cada perigo restante é estimado com base em um algoritmo qualitativo tabular que inclui a dimensão do dano potencial estimado. Havendo risco determina-se se ele é aceitável ou se pode ser manejado para reduzir, evitar, mitigar ou minimizar seus efeitos.

Fonte: adaptado de Andrade, Parrott, Roca (2012).

A avaliação de risco é o primeiro passo no fluxograma da análise de risco. Para a liberação de um OGM além da avaliação de risco é necessária a gestão e a comunicação de risco. Este fluxograma é apresentado na Figura 9.

Figura 9 - Fluxograma da análise de risco de um OGM.



Fonte: adaptado de Andrade, Parrott, Roca (2012).

Uma vez liberado o OGM para cultivo, se utilizado para alimentação, deverá ser rotulado adequadamente. No Brasil, em abril de 2003 foi emitido o Decreto nº 4.680 que regulamenta o direito à informação (BRASIL, 2003<sup>a</sup>), assegurado pela Lei nº 8.078/1990 (Código de Defesa do Consumidor – CDC), quanto aos alimentos e ingredientes alimentares que contenham ou sejam produzidos a partir de OGM. O artigo 2º deste decreto versa sobre a obrigatoriedade de rotulagem informando a presença de OGM (limite de 1 %). A Portaria nº 2.658 de 22 de dezembro de 2003 define o símbolo de que trata o artigo 2º, § 1º, do Decreto 4.680 (BRASIL, 2003<sup>b</sup>). A Figura 10 apresenta a simbologia.

Figura 10 - Apresentação gráfica do símbolo transgênico para rotulagem de alimentos embalados.



Fonte: adaptado de Brasil (2003<sup>b</sup>).

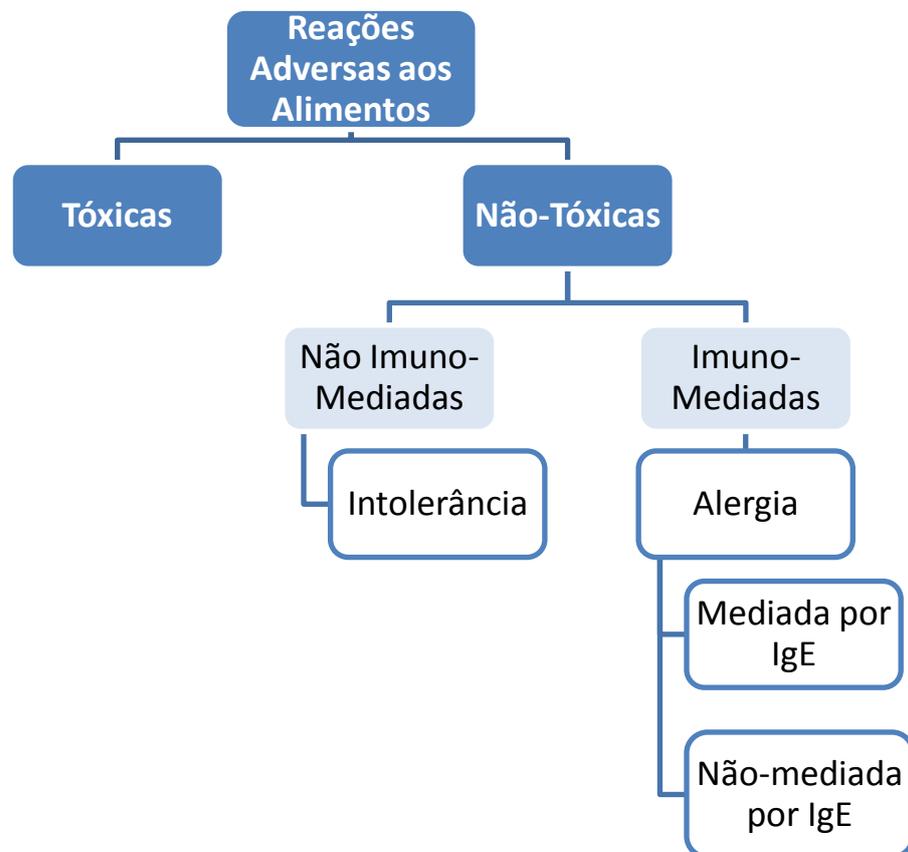
A avaliação de risco supracitada considera não só riscos ambientais como também riscos à saúde humana no que tange, por exemplo, a alergenicidade e antigenicidade.

### 2.5 Alergenicidade e Antigenicidade

O termo alergia refere-se a uma reação imunológica anormal em resposta a um material estranho, ou seja, um alérgeno (TAYLOR, 1992). Na alergia alimentar (FA) esta reação anormal ocorre em resposta a um alimento (LONGO et al., 2013), mais especificamente a uma proteína dietética (WASERMAN; WATSON, 2011). Para Sánchez-Monge e Salcedo (2005) os protagonistas da alergia alimentar, além dos alérgenos alimentares de natureza proteica, são os anticorpos de Imunoglobulinas E (IgE) específicos.

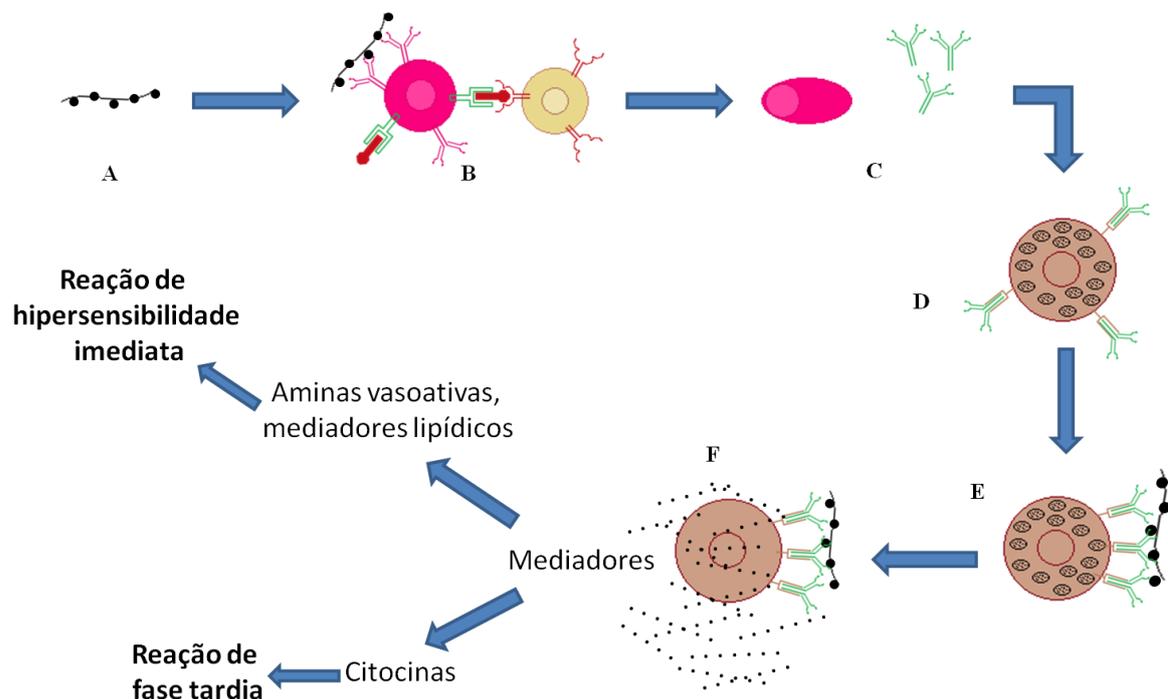
A alergia alimentar é um dos tipos de reações adversas decorrentes da exposição a determinados alimentos (LONGO et al., 2013). A figura 11 apresenta uma breve classificação quanto as reações adversas aos alimentos.

Figura 11 - Classificação das Reações Adversas aos Alimentos.



Quando uma proteína estranha é reconhecida pelo sistema imunológico ocorre a secreção de Imunoglobulinas A (IgA) que bloquearão a absorção de tais proteínas (SOUZA JÚNIOR; MARTINS, 2003). Caso esta proteína atinja o sistema circulatório, de acordo com Souza Júnior e Martins (2003), serão gerados anticorpos da classe IgA, IgM (Imunoglobulinas M) e IgG (Imunoglobulinas G) em pequenas quantidades para eliminar este antígeno alimentar. A alergia alimentar do tipo mediada por IgE ocorre pela sensibilização a alérgenos alimentares com formação de anticorpos específicos da classe IgE (SOLÉ et al., 2008) e liberação de histaminas, prostaglandinas e leucotrienos resultando em uma reação alérgica imediata (SOUZA JÚNIOR; MARTINS, 2003). Por outro lado, as reações alérgicas não-mediadas por IgE referem-se à reações citotóxicas, reações por imunocomplexos e hipersensibilidade mediada por células (SOLÉ et al., 2008). Solé et al. (2008) caracterizam ainda as reações mistas que são aquelas que são mediadas por IgE e células onde ocorrem manifestações decorrentes de mecanismos mediados por IgE e com participação de linfócitos T. Os eventos que ocorrem durante uma reação alérgica são apresentados na Figura 12.

Figura 12 - Eventos na reação de alergicidade.



Em A tem-se a primeira exposição ao alérgeno. Esta exposição leva à ativação dos linfócitos  $T_H2$  e à estimulação da troca para classe IgE nos linfócitos B (B). Posteriormente, (C) ocorre a produção de IgE. Este IgE produzido se liga aos mastócitos (D). Uma nova exposição ao alérgeno ocorre (E) e com a ativação do mastócito (F) ocorre a liberação dos mediadores. Quando liberados mediadores do tipo citocinas há uma reação de fase tardia (6-8 horas após a exposição ao alérgeno). Por outro lado, quando liberados mediadores lipídicos há uma reação de hipersensibilidade imediata.

Fonte: adaptado de Abbas, Lichtman e Pillai, 2008.

O processo de sensibilização, ou seja, a ligação de anticorpos IgE a receptores Fc dos mastócitos, (ABBAS; LICHTMAN; PILLAI, 2008) leva a liberação de mediadores químicos ou farmacológicos da inflamação (SOUZA JÚNIOR; MARTINS, 2003). As histaminas (aminas vasoativas), prostaglandinas e os leucotrienos (metabólitos do ácido araquidônico) e as citocinas, entre outros, são mediadores químicos que são liberados pelos mastócitos (CAMPILLO-NAVARRO et al., 2014).

As manifestações clínicas são apresentadas na Tabela 1.

Tabela 1 - Manifestações Clínicas nas Diferentes Reações Alérgicas.

Reações Mediadas por IgE	Reações Mistas	Reações Não-Mediadas por IgE
Angiodema	Dermatite Atópica	Proctocolite Induzida por Proteína
Urticária	Doenças Digestivas Eosinofílicas	Enterocolite
Asma	Gastrite	Doença Celíaca
Rinite		Dermatite de Contato e Herpetiforme
Anafilaxia		Síndrome de Heiner

Fonte: adaptado de Sampson (1999).

Alergias, de modo geral, afetam de 20 – 30 % da população a nível mundial (GALLI et al., 2016). Dados apresentados por Toit et al. (2016) demonstram que a alergia alimentar afeta de 3 a 6 % das crianças. Solé et al. (2008) indica que para a população mundial este valor é de 3,5 %. Já para Sicherer (2011) este valor é maior que 1 – 2% mas menor que 10%. Sicherer e Sampson (2010) identificaram um aumento de 18 % nos casos de alergia alimentar em crianças entre os anos de 1997 e 2007. Para Longo et al. (2013) a alergia alimentar influencia substancialmente a qualidade de vida das crianças afetadas e suas famílias. A ocorrência da alergia alimentar depende de fatores genéticos, idade, hábitos alimentares, posição geográfica e procedimentos de diagnóstico (SOLÉ et al., 2008).

De acordo com o exposto anteriormente e com Roth-Walter et al. (2014), algumas famílias de proteínas possuem capacidade de gerar alergias, ou seja, possuem potencial alergênico. As características daquele que é alérgeno definem a alergenicidade (MOHABATKAR et al. 2013). Estes alérgenos podem induzir uma sensibilização pela

indução de respostas imunes e produção de IgE (TSAI et al., 2015). Já a antigenicidade é conceituada como a capacidade de uma substância ou fração da mesma de ligar-se a anticorpos ou células do sistema imune (ABBAS; LICHTMAN; PILLAI, 2008), ou seja, capacidade de atuar como antígeno. Antígenos são substâncias que além de estimular a produção de anticorpos são capazes de reagir especificamente com eles (TSAI et al., 2015). A menor parte deste antígeno que possui capacidade de ligar-se a anticorpos é conhecida como epítopo (ABBAS; LICHTMAN; PILLAI, 2008).

A partir da caracterização de diferentes alérgenos foi possível identificar as características, como, por exemplo, a similaridade com famílias de alérgenos conhecidas, responsáveis pelo potencial de alergenicidade de determinada proteína (SÁNCHEZ-MONGE; SALCEDO, 2005). Contudo, apenas a análise isolada destas características não é suficiente para predizer que determinada sequência de aminoácidos possui potencial alergênico. Métodos *in vivo*, *in vitro* e *in silico* foram desenvolvidos para definir o potencial alergênico de uma proteína, conforme Tabela 2.

Tabela 2 - Métodos *in vivo*, *in vitro* e *in silico* para definição do potencial de alergenicidade.

Métodos	Fundamentação	Processo de Avaliação
<b><i>In vivo</i></b>		
- DBPCFC	Humano, ingestão	Sintomas clínicos
- SPT	Humano, punção da pele	Reatividade da pele
- Modelos Animais	Camundongos, cães, porcos, coelhos, oral ou intraperitoneal	Sensibilização, sintomas de desenvolvimento
<b><i>In vitro</i></b>		
- Imunoensaios	Soro humano	Ligação à IgE específica
- <i>Immunoblotting</i>	Soro humano	Ligação à IgE específica
- Teste de Ativação de Basófilos	Sangue humano, preparação de leucócitos	Desgranulação de células efectoras
- Ensaio de resposta das células T	Células mononucleares de sangue, clones de célula T	Proliferação de linfócitos T e perfil das citocinas
<b><i>In silico</i></b>		
- Busca por homologias	Bancos de dados de alérgenos	Homologia com sequências de alérgenos conhecidos
- Baseados em Algoritmos	<i>Support Vector Machine</i>	Identificação de padrões nas sequências

Fonte: adaptado de Sánchez-Monge e Salcedo (2005).

Para satisfazer o método de busca por homologias para identificação do potencial de alergenicidade de proteínas alguns algoritmos foram desenvolvidos e bancos de dados de alérgenos foram elaborados. A ciência da bioinformática é a responsável por tais ferramentas e metodologias.

## 2.6 Bioinformática

Um dos conceitos mais aceitos para esta ciência é a de Luscombe, Greenbaum e Gerstein (2001) que definem a bioinformática como uma união entre a biologia e a informática.

Xiong (2006) apresenta um breve histórico referente à bioinformática:

- Década de 1960: Dayhoff desenvolveu o “*Atlas of Protein Sequence and Structure*”;
- Década de 1970: Needleman e Wunsch criaram o primeiro algoritmo para o alinhamento de sequências e Chou e Fasman foram responsáveis pelo primeiro algoritmo de predição de estruturas proteicas;
- Década de 1980: Altschul desenvolveu o BLAST, Pearson o FASTA e teve início o Projeto Genoma Humano; e,
- Década de 1990: advento da internet.

Com o advento da internet, a bioinformática assumiu sua posição na ciência. Para Xiong (2006), o destaque da bioinformática ocorreu com a explosão de informação gerada pelo sequenciamento genômico. Esta informação necessitava, e ainda necessita, de ferramentas computacionais eficientes para gerenciar e analisar os dados gerados (XIONG, 2006).

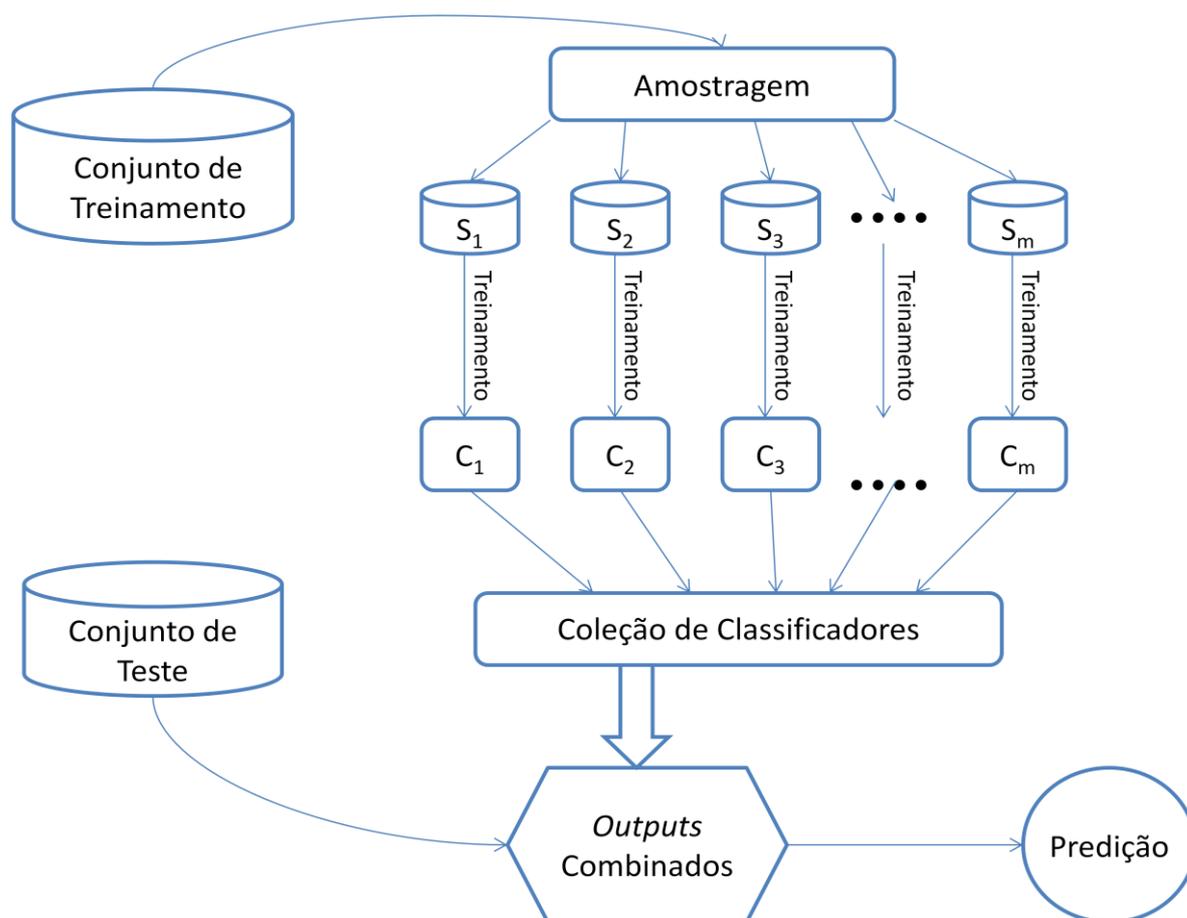
Para melhor compreender a função de um gene, diversos algoritmos foram desenvolvidos. Estes algoritmos podem ser sumarizados em; (i) comparação de sequências; (ii) árvores filogenéticas; (iii) detecção de padrões em sequências; (iv) determinação de estruturas tridimensionais; (v) inferência da regulação celular; (vi) determinação de funções proteicas e vias metabólicas; e, (vii) montagem de fragmentos de DNA, dentre outros algoritmos (COHEN, 2004). No que tange a comparação de sequências destaca-se o alinhamento de sequências.

O alinhamento de sequências é considerada uma das ferramentas mais importantes da bioinformática (PROSDOCIMI et al., 2002). A Figura 13 apresenta um alinhamento de duas sequências de proteínas.



Outro método utilizado na análise do potencial de alergenicidade de proteínas é o método baseado em algoritmos, mais especificamente as Máquinas de Vetores de Suporte (do inglês *Support Vector Machine* - SVM). As SVM constituem uma técnica de aprendizado por vezes superior a modelos como as Redes Neuras Artificiais (LORENA; CARVALHO, 2007). A Figura 15 apresenta uma breve descrição sobre o método.

Figura 15 - Determinação do conjunto de classificadores do SVM.



Cada classificador ( $C_m$ ) é treinado por uma subamostra ( $S_m$ ) a partir de um conjunto de treinamento (conjunto de dados que possui um número equilibrado entre casos negativos e positivos). O conjunto de classificadores é treinado e avaliado pelo conjunto de dados de teste. Por fim, a predição é calculada a partir da predição dos classificadores individuais.

Fonte: adaptado de Caragea et al. (2007).

O SVM analisa dados a fim de reconhecer padrões para classificar determinada amostra (LORENA; CARVALHO, 2007). O conjunto de treinamento e o conjunto de teste, poderá ser constituído por aminoácidos bem como dipeptídeos originando atributos e rótulos (classificadores).

O objetivo final da bioinformática é entender melhor uma célula e suas funcionalidades a nível molecular (XIONG, 2006). A bioinformática revolucionou a avaliação

do potencial de alergenicidade de proteínas por fornecer uma base de dados e ferramentas de predição *in silico* para tal (SÁNCHEZ-MONGE; SALCEDO, 2005). Neste ponto, a análise das sequências é essencial para detectar e prever estruturas, reações adversas e o potencial de alergenicidade (SOUZA JÚNIOR; MARTINS, 2003).

### 3. Contextualização

Este trabalho descreve a construção de uma plataforma, chamada Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas – PA<sup>3</sup>P, utilizada para categorizar proteínas alergênicas ou antigênicas. A categorização foi realizada de acordo com as características da sequência putativa de aminoácidos e pelo uso de um conjunto de ferramentas específicas que contribuirá para embasar a decisão de utilizar ou não determinada proteína na construção de um Organismo Geneticamente Modificado (OGM) com finalidade nutricional. Com o intuito de facilitar a compreensão tanto da metodologia utilizada quanto dos resultados apresentados, esta dissertação foi dividida em dois capítulos:

Capítulo I, intitulado “Construção da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas” dispõe sobre a concatenação das principais ferramentas disponíveis na *web* para predição do potencial de alergenicidade e antigenicidade de proteínas;

e

Capítulo II, intitulado “Categorização *in silico* de proteínas alergênicas ou antigênicas por meio da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P)” apresenta a funcionalidade da PA<sup>3</sup>P para a identificação de proteínas potencialmente alergênicas.

A dissertação também apresenta apêndices contendo informações suplementares para melhor descrição do trabalho:

- Apêndice 1: manuscrito gerado a partir deste trabalho;
- Apêndice 2: planilha do cálculo da acurácia e validação da PA<sup>3</sup>P via *dataset*;
- Apêndice 3: planilha de validação; e,
- Apêndice 4: códigos-fonte de construção da PA<sup>3</sup>P.

## 4. Objetivos

### 4.1 *Objetivo Geral*

Construir uma plataforma – chamada PA<sup>3</sup>P (Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas) de análise integrada.

### 4.2 *Objetivos Específicos*

- Estimar a acurácia das ferramentas *web* disponíveis para análise do potencial de alergenidade e antigenidade;
- Compilar as informações e tecnologias dos diferentes *sites* de pesquisa e bancos de dados públicos em um único local;
- Categorizar proteínas alergênicas ou antigênicas, de acordo com as características biológicas da expressão da sequência de aminoácidos, pelo uso de um conjunto de ferramentas específicas; e,
- Analisar grupos de proteínas *in silico*, definindo seu potencial alergênico e antigênico.

## 5. Capítulo I - Construção da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas

### 5.1 Introdução

Para que novas proteínas, oriundas de modificações genéticas, possam estar presentes em alimentos é necessário que seu potencial alergênico seja avaliado (FIERS et al., 2004). A predição e avaliação do potencial alergênico é essencial para a avaliação da segurança do consumidor final (WANG et al., 2013). Contudo, tal potencial não é facilmente previsível uma vez que depende de diversos fatores como, por exemplo, a variabilidade da resposta de IgE específicas (SOUZA JUNIOR; MARTINS, 2003).

A FAO juntamente com a WHO, em 2001, definiram que uma proteína seria possivelmente alergênica quando apresentasse: (i) seis ou mais aminoácidos contíguos e idênticos quando comparada com uma proteína sabidamente alergênica; ou, (ii) mais de 35 % de identidade em uma janela de 80 aminoácidos quando comparada com uma proteína sabidamente alergênica (FAO/WHO, 2001). Tais características isoladas não são suficientes para determinar um possível potencial alergênico (STADLER; STADLER, 2003).

Para cumprir a exigência de avaliação do potencial de alergenidade de novas proteínas, diversas abordagens computacionais foram desenvolvidas (WANG et al., 2013). Tais abordagens investigam: (i) a fonte do gene; (ii) a homologia da sequência com alérgenos conhecidos; (iii) a associação com IgEs de fonte sorológica de indivíduos alérgicos; e, (iv) propriedades físico-químicas da proteína codificada pelo gene introduzido (SOUZA JUNIOR; MARTINS, 2003). Para identificar homologias entre a sequência de consulta (*query sequence*) e sequências de aminoácidos de alérgenos conhecidos novas ferramentas *web* foram desenvolvidas como, por exemplo, *Prediction of Allergenic Proteins and Mapping of IgE Epitopes* (AlgPred), *Allergen Online*, *Allergen Database for Food Safety* (ADFS), dentre outras.

A ferramenta AlgPred (descrita por SAHA; RAGHAVA, 2006; e, disponível em <http://www.imtech.res.in/raghava/algpred/>) realiza predições de proteínas alergênicas e mapeamento de epítomos elicitores de IgE através da busca por similaridades. De acordo com este método uma proteína é predita como alergênica quando possuir um ou mais epítomos IgE. A busca por *Allergen Representative Peptides* (ARPs) é uma abordagem com alta sensibilidade e precisão. Consiste em um banco de 2.890 ARPs e quando um *hit* é encontrado a sequência de consulta é categorizada como alérgeno (SAHA; RAGHAVA, 2006). AlgPred

não realiza apenas predições baseadas em similaridades como também realiza predições *in silico* baseadas em algoritmos, como a predição baseada em *Support Vector Machine* (SVM). SVM pode ser conceituado como um conjunto de métodos do aprendizado supervisionado que analisam os dados a fim de reconhecer padrões e, desta forma, classificar determinada amostra (LORENA; CARVALHO, 2007).

Outra ferramenta que possibilita a predição *in silico* é o banco de dados *Allergen Online* (disponível em <http://allergenonline.com/>). Esta possibilita a comparação entre a sequência de aminoácidos da proteína de interesse com sequências de proteínas alergênicas. A predição ocorre seguindo as recomendações da FAO/WHO, conforme segue: (i) *Allergen Online* com janelas de 80 aminoácidos: uma proteína será considerada alergênica quando apresentar uma similaridade igual ou maior que 35 % - em uma janela de 80 aminoácidos; e, (ii) *Allergen Online* com *wordmatch* de 8 aminoácidos: uma proteína será considerada alergênica quando apresentar 8 aminoácidos contíguos comparada a uma proteína comprovadamente alergênica.

Um método baseado na busca por similaridades entre motivos de proteínas consiste na principal ferramenta do banco de dados *Allergen Database for Food Safety* [ADFS; disponível em <http://allergen.nihs.go.jp/ADFS/> e descrito por NAKAMURA; NAKAMURA; TESHIMA (2009)]. O ADFS possui um banco de dados próprio com mais de 1.300 alérgenos (NAKAMURA; NAKAMURA; TESHIMA, 2009).

Por fim, a ferramenta *web* chamada *Protein Allergenicity Prediction* [proAP; disponível em <http://gmobl.sjtu.edu.cn/proAP/main.html> e descrito por WANG et al. (2012)] realiza análises por meio de algoritmo de SVM baseado nos aminoácidos que compõe a sequência de interesse. Dentro do contexto das novas abordagens computacionais que avaliam o potencial de alergenicidade e antigenicidade, a Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P – disponível em <http://lpa.saogabriel.unipampa.edu.br:8080/pa3p/index.jsp> ou <http://pluriserver.com.br/pa3p/>) foi desenvolvida concatenando as ferramentas supracitadas.

A construção da plataforma PA<sup>3</sup>P (Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas) visa contribuir para embasar a decisão de utilizar ou não determinada proteína na construção de OGM com finalidade nutricional. A plataforma conta com três seções: (i) as análises sugeridas pela FAO/WHO; (ii) alinhamentos em diferentes bancos de dados que possuem sequências de proteínas comprovadamente ou potencialmente

alergênicas e antigênicas; e, (iii) análises por diferentes algoritmos de predição por busca de motivos e epítomos.

## 5.2 Metodologia

### 5.2.1 Dataset

Para definir o  $n$  amostral que constituiria o *dataset* utilizou-se o cálculo de amostragem para populações finitas, uma vez que até o momento foram descritas 1.897 sequências de proteínas sabidamente alergênicas, conforme equação:

$$n = \frac{Z^2 * P * Q * N}{e^2 * (N - 1) + Z^2 * P * Q}, \text{ onde:}$$

Z – nível de confiança de 99 %;

P – acerto esperado de 60 %;

Q – erro esperado de 40 %;

N – população total de 1.897 indivíduos; e,

e – nível de precisão de 5 %.

Através deste cálculo foi fixado um valor amostral de  $n = 475$ . Então, o *dataset* foi constituído de 475 sequências de proteínas comprovadamente alergênicas obtidas da listagem de alérgenos do banco de dados *Allergen Online* (disponível em <http://allergenonline.com/> versão de Janeiro de 2015) por meio de amostragem aleatória. Para sequências de proteínas não-alergênicas realizou-se uma busca no banco de dados de proteínas do *National Center for Biotechnology Information* (NCBI – disponível em <http://www.ncbi.nlm.nih.gov>). O valor de  $n$  deveria ser igual para o *dataset* das sequências de proteínas alergênicas e não-alergênicas. Contudo, a busca junto ao NCBI (em junho de 2015) retornou apenas 466 sequências de proteínas relacionadas às palavras-chave “*immunogenic*” e “*Homo sapiens*”.

O *dataset* completo encontra-se em <http://pluriserver.com.br/pa3p/database/database.jsp> ou <http://lpa.saogabriel.unipampa.edu.br:8080/pa3p/database/database.jsp>.

### 5.2.2 Acurácia

A acurácia das ferramentas compiladas na plataforma foi definida de acordo com a seguinte equação:

$$Acurácia = \frac{VP+VN}{VP+FP+VN+FN}, \text{ onde:}$$

	Alergênica	Não-Alergênica
Alergênica	Verdadeiro Positivo (VP)	Falso Positivo (FP)
Não-alergênica	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Para chegar aos valores finais de acurácia, as 475 sequências de proteínas alergênicas e as 466 sequências de proteínas não-alergênicas foram submetidas às ferramentas individualmente. Uma matriz binária (0 para ausência; 1 para presença) foi construída com o resultado total de todas as análises (ver Tabela 6 - Apêndice 2) e um resultado sumarizado é apresentado na Tabela 3.

Tabela 3 - Principais resultados referentes ao cálculo da acurácia das *web tools*.

	1	2	3	4	5*	6	7
Verdadeiros Positivos (VPs)	103	439	378	416	98	462	472
Verdadeiros Negativos (VNs)	462	337	459	388	94	437	453
Falsos Positivos (FPs)	372	36	97	59	2	13	3
Falsos Negativos (FNs)	4	129	7	78	6	29	13
Acurácia	60,04%	82,47%	88,95%	85,44%	96,00%	95,54%	98,30%

*Web tools*: (1) *AlgPred – IgE*; (2) *AlgPred - SVM dipeptide*; (3) *AlgPred – ARP*; (4) *proAP - SVM aa*; (5) *ADFS - Motif-based*; (6) *Allergen Online - 80 aa window*; e, (7) *Allergen Online - 8 aa wordmatch*.

\* Foram analisadas 100 sequências de proteínas alergênicas e 100 de proteínas não-alergênicas devido à indisponibilidade do banco de dados *ADFS*.

Fonte: dados produzidos pelo autor (2016).

### 5.2.3 Construção da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas

As ferramentas *web* selecionadas foram concatenadas na PA<sup>3</sup>P através das linguagens de programação *HyperText Markup Language* (HTML), Java SE 1.7 e JavaScript em um servidor *web* (Apache Tomcat 7 Server). Programas em Java podem ser incorporados a HTML (DEITEL;DEITEL, 2005), ou seja, páginas da *web* como a PA<sup>3</sup>P.

Os *plugins* utilizados na construção da PA<sup>3</sup>P foram: (i) jQuery, uma biblioteca de funções de JavaScript; (ii) jQuery User Interface (UI), responsável pela interface amigável e

interativa da PA<sup>3</sup>P; e, (iii) Highcharts JS, uma biblioteca de gráficos baseada em JavaScript responsável pelos resultados gráficos da PA<sup>3</sup>P. As ferramentas utilizadas para construção da PA<sup>3</sup>P são gratuitas para projetos pessoais ou sem fins lucrativos.

A plataforma foi alocada no servidor *web* da Universidade Federal do Pampa. Além disso, será mantida no <http://pluriserver.com.br> como alternativa para o caso de o endereço principal estar *offline*. Todos os códigos-fonte são apresentados no Apêndice 4.

A sequência de consulta quando submetida na PA<sup>3</sup>P é processada em todas as *web tools* implementadas. Conforme cada ferramenta retorna resultados como *allergen* ou *non-allergen* uma planilha (ver Tabela 4) é preenchida de acordo com o método e respectiva acurácia para que ao final obtenha-se o resultado sumarizado.

Tabela 4 - Planilha de pontuação da acurácia conforme o método e o resultado da análise.

	1	2	3	4	5	6	7	Média
Alergênica	96%	88,95%	60,04%	82,47%	95,54%	98,30%	85,44%	85,12%
Não-Alergênica	0%	0%	0%	0%	0%	0%	0%	0%

Conforme a *query* é processada na plataforma, uma planilha de pontuação é preenchida. Quando ocorre um erro no processamento, o valor de acurácia da ferramenta é desconsiderado. Ao final da análise uma média é calculada de acordo com os resultados (alergênica/não-alergênica) de cada *web tool*. *Web tools*: (1) ADFS – Motif-based; (2) AlgPred – ARP; (3) AlgPred – IgE; (4) AlgPred – SVM dipeptide; (5) Allergen Online – 80 aa window; (6) Allergen Online – 8 aa wordmatch; e, (7) proAP – SVM aa.

Fonte: dados produzidos pelo autor (2016).

#### 5.2.4 Validação

O processo de validação da PA<sup>3</sup>P consistiu na submissão do *dataset*, de 475 sequências de aminoácidos referentes à proteínas alergênicas bem como 466 sequências de aminoácidos referentes à proteínas não-alergênicas, à plataforma. Esta etapa é de extrema importância para analisar a confiabilidade da nova ferramenta.

O resultado do processo de validação também possibilita a delimitação da sensibilidade e especificidade da PA<sup>3</sup>P com as seguintes equações:

$$\text{Sensibilidade} = \frac{VP}{VP+FN} \text{ e,}$$

$$\text{Especificidade} = \frac{VN}{VN+FP}, \text{ onde:}$$

	Alergênica	Não-Alergênica
Alergênica	Verdadeiro Positivo (VP)	Falso Positivo (FP)
Não-alergênica	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Concomitantemente ao processo de validação ocorreu a delimitação do valor do erro de acordo com a seguinte equação:

$$\text{Erro Padrão} = \frac{\text{Desvio Padrão}}{\sqrt{\text{Tamanho da Amostra}}}$$

### 5.2.5 Categorização de Proteínas

Um grupo composto de 100 sequências de aminoácidos de proteínas com índices de alergenidade e antigenidade desconhecidos foi submetido à análise na PA<sup>3</sup>P. Tais proteínas foram selecionadas aleatoriamente do banco de dados *Protein Data Bank* (disponível em <http://www.rcsb.org/pdb/home/home.do>). Este teste foi realizado para comprovar a funcionalidade da PA<sup>3</sup>P e seu resultado completo pode ser consultado na Tabela 7 no Apêndice 3.

### 5.3 Resultados

A plataforma conta com três seções: (i) as análises sugeridas pela FAO/WHO; (ii) alinhamentos em diferentes bancos de dados que possuem sequências de proteínas comprovadamente ou potencialmente alergênicas e antigênicas; e (iii) análises por diferentes algoritmos de predição por busca de motivos e epítomos (Figura 16).



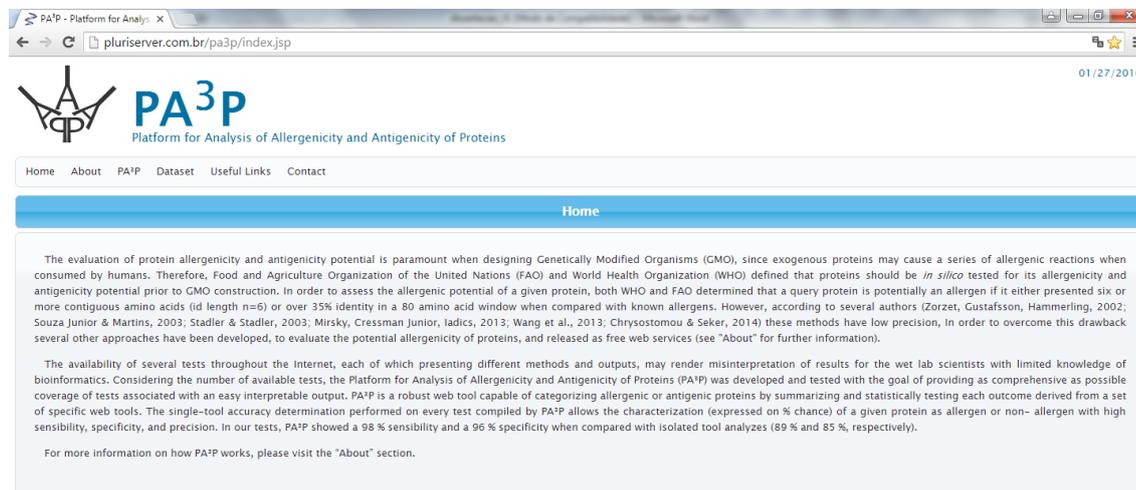
Figura 16 – Fluxograma de construção e funcionamento da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P).

**A.** *Dataset* da PA<sup>3</sup>P. O cálculo de amostragem para população finita de allergen consiste em: 99 % de nível de confiança (*Z*); acerto esperado ( $P = 60\%$ ) complementar ao erro esperado ( $Q = 40\%$ ); população total de  $N = 1.897$ ; e, um nível de precisão  $e = 5\%$ ; resultando em  $n = 475$  sequencias allergen por meio de sorteio aleatório. Para o dataset de sequências non-allergen o *n* foi fixado em 466 sequências uma vez que estas são as únicas sequências relacionadas a imunnogenic e *Homo sapiens* depositadas no National Center for Biotechnology Information (NCBI). **B.** O dataset foi submetido as web tools isoladamente para que por meio do cálculo da acurácia fosse possível identificar quais métodos apresentavam menos casos de falsos positivos e falsos negativos (**C**). Após a identificação dos métodos que melhor se adequam ao modelo, estes foram compilados pela linguagem Java dando origem a PA<sup>3</sup>P (**D**). Ao inserir sua sequencia query em **D**, esta é processada nos diferentes métodos (**C**) presentes na plataforma. Em **E** temos a tabela de pontuação baseada na acurácia, conforme o método retorna o resultado alergênico/não alergênico a pontuação é distribuída e por fim há o cálculo da média adicionado/subtraído o valor do erro populacional. Quando uma das ferramentas retorna 'Error' este valor não é atribuído como, por exemplo, a ADFS. Baseado nesta media, os gráficos são gerados (**F**).

Fonte: dados produzidos pelo autor (2016).

A Plataforma de Análise da Alergenicidade e Antigenicidade pode ser acessada em <http://lpa.saogabriel.unipampa.edu.br:8080/pa3p> ou <http://www.pluriserver.com.br/pa3p>. A Figura 17 apresenta a interface da nova ferramenta.

Figura 17 - Interface da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P).

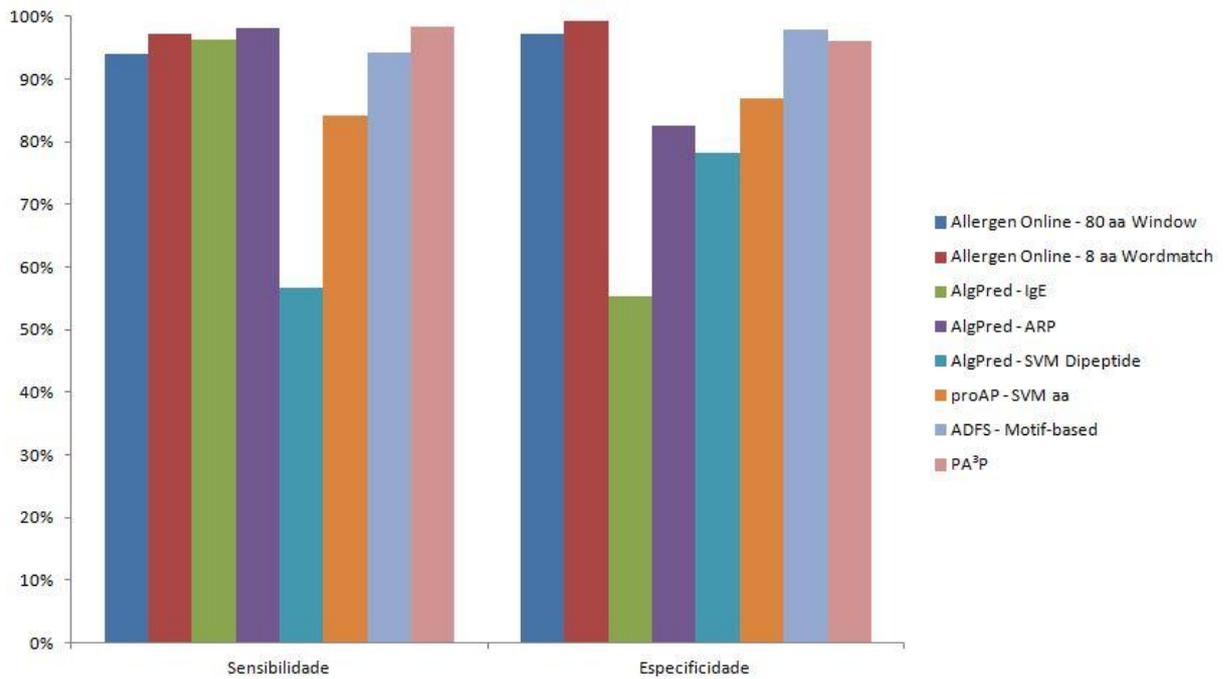


Fonte: dados produzidos pelo autor (2016).

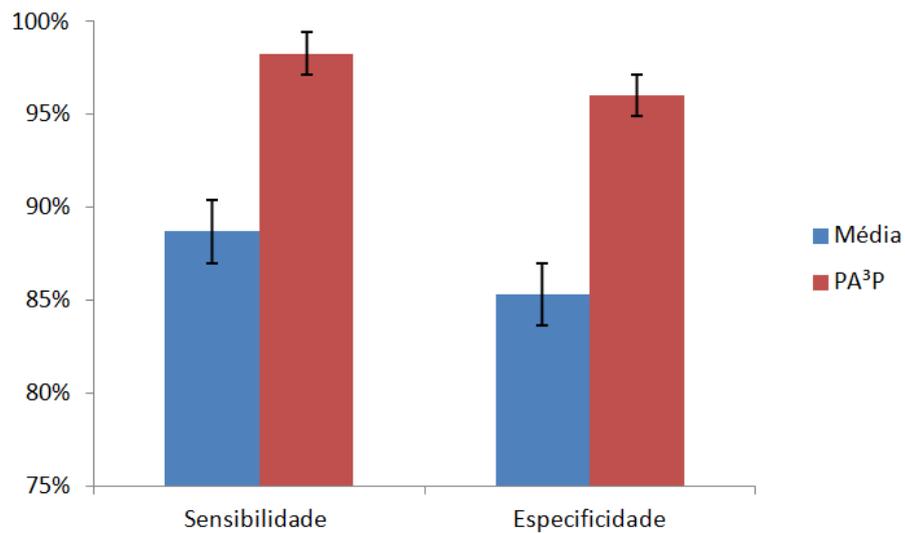
O quadrante superior esquerdo da PA<sup>3</sup>P é composto de diferentes seções. A aba “Home” apresenta uma breve descrição sobre a plataforma. A seção “About” apresenta informações acerca da funcionalidade da plataforma bem como descreve as ferramentas utilizadas e sua respectiva metodologia. A seção “PA<sup>3</sup>P” é destinada ao processamento da análise de interesse. “Dataset” refere-se às sequências de aminoácidos utilizadas para calcular tanto a acurácia das *web tools* como validar a plataforma. Em “Useful Links” estão disponíveis os acessos às ferramentas individuais e outros bancos de dados de interesse. E, por fim, na seção “Contact” são disponibilizados os e-mails de contato.

A validação da PA<sup>3</sup>P revelou um aumento substancial tanto na sensibilidade quanto na especificidade quando se compara aos valores obtidos pelas ferramentas individualmente, conforme Figura 18 (A e B). Resultado semelhante foi encontrado tanto para a análise do *dataset* (Apêndice 2 – Tabela 6) quanto para o grupo de sequências de aminoácidos com índices de alergenidade e antigenidade desconhecidos (Apêndice 3 – Tabela 7).

Figura 18 - Especificidade e sensibilidade da PA<sup>3</sup>P.



A Sensibilidade e Especificidade das ferramentas *web* individualmente e da PA<sup>3</sup>P.



**B** Média (em azul) refere-se a média dos resultados obtidos nas *web tools*. PA<sup>3</sup>P (em vermelho) refere-se ao resultado obtido quando utilizada a plataforma como ferramenta de categorização de proteínas.

Fonte: dados produzidos pelo autor (2016).

Concomitante ao processo de validação da PA<sup>3</sup>P ocorreu a delimitação dos valores referentes ao erro amostral observáveis na Tabela 5.

Tabela 5 - Valores referentes ao erro amostral.

	Erro Amostral
Resultado baseado em algoritmo	$\pm 1,20 \%$
Resultado baseado em similaridade	$\pm 1,02 \%$
Resultado sumarizado	$\pm 0,97 \%$

Fonte: dados produzidos pelo autor (2016).

Novas ferramentas podem ser adicionadas à PA<sup>3</sup>P. Outras informações podem ser obtidas mediante acesso à plataforma.

## 6. Capítulo II – Categorização *in silico* de proteínas alergênicas ou antigênicas por meio da Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P)

### 6.1 Introdução

Um OGM é aquele que teve seu material genético modificado por qualquer técnica de engenharia genética (BRASIL, 2005). Esta modificação levará a novas características biológicas originárias da expressão da sequência de aminoácidos (HJÄLTÉN et al., 2013). Contudo, nem sempre estas novas características são benéficas, ou seja, existem alguns riscos. Para os OGM os principais perigos potenciais incluem a alergenicidade e antigenicidade, alterações nutricionais e efeitos antinutrientes, toxicidade e possibilidade remota de transferência horizontal de genes (COSTA et al., 2011).

O potencial de alergenicidade e antigenicidade de proteínas pode ser predito e avaliado. O potencial de alergenicidade de uma proteína pode ser previsível por meio de testes *in silico* (MOHABATKAR et al., 2013), *in vivo* e *in vitro*. Diversas ferramentas foram criadas para realizar predições *in silico* do potencial alergênico e antigênico de proteínas, dentre estas a Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P).

O método de predição baseado em similaridades realiza comparações entre sequências. Estas comparações, conhecidas como alinhamentos, buscam identificar características comuns entre os pares. Por outro lado, o método de predição *in silico* baseado em algoritmos utiliza de modelos que toma como padrão para determinada classe (SOUZA JUNIOR; MARTINS, 2003). O *Support Vector Machine* (SVM) pode ser conceituado como um conjunto de métodos do aprendizado supervisionado que analisam os dados a fim de reconhecer padrões e, desta forma, classificar determinada amostra (LORENA; CARVALHO, 2007). Neste tipo de aprendizado, o indutor (algoritmo de aprendizado) recebe exemplos (padrões) de treinamento. Estes padrões são descritos por atributos e pelo rótulo da classe (neste caso, os alérgenos).

Neste contexto, a categorização de proteínas é um passo essencial na construção de OGM, pois são os resultados das análises prévias que definirão a possibilidade do uso seguro de determinada proteína.

## 6.2 Metodologia

A categorização (em alergênica ou não-alergênica) da sequência de consulta (*query*) foi realizada pela inserção da sequência de aminoácidos correspondente na Plataforma de Análise da Alergenicidade e Antigenicidade de Proteínas (PA<sup>3</sup>P) (<http://pluriserver.com.br/pa3p/> ou <http://lpa.saogabriel.unipampa.edu.br:8080/pa3p>), conforme ilustrado na Figura 19. Exemplos de sequências de aminoácidos são apresentados a seguir.

- Exemplo de sequência de uma proteína alergênica:

```
>gi|14423645/sp|P79085.1|ALTA1_ALTAL RecName: Full=Major allergen Alt a 1;
AltName: Allergen=Alt a 1; Flags: Precursor
MQFTTIASLFAAAGLAAAAPLESRQDTASCPVTTEGDYVWKISEFYGRKPEGTYYNS
LGFNIKATNGGTLDFTCQAQADKLEDHKWYSCGENSFMDFSFSDRSGLLLKQKVSD
DITYVATATLPNYCRAGGNGPKDFVCQGVADAYITLVTLPKSS
```

- Exemplo de sequência de uma proteína não-alergênica:

```
>gi|68161511/ref|NP_001020131.1| actin-related protein 2/3 complex subunit 4 isoform b
[Homo sapiens]
MRFMMRAENFFILRRKPVEGYDISFLITNFHTEQMYKHKLVDFVIHFMEEIDKEISE
MKLSVNARARIVAEFLKNF
```

Figura 19 - Destaque para o campo destinado à inserção da sequência de consulta para submissão à análise na PA<sup>3</sup>P.

A sequência é analisada pela PA<sup>3</sup>P ao clicar-se em “Run”. Uma nova busca pode ser conduzida clicando em “New Query”.

Fonte: dados produzidos pelo autor (2016).

A sequência analisada foi o alérgeno Alt a 1 pertencente ao fungo *Alternaria alternata*.

### 6.3 Resultados

Após o processamento da sequência de consulta, uma tabela apresenta os resultados para ferramenta conforme a categorização da sequência em alergênico/não alergênico (Figura 20).

Figura 20 - Tabela de resultados individuais conforme ferramenta e método.

Individual Results		
WebTool	Method	Result
Algpred - IgE	Similarity (60.04%)	Allergen
Algpred - SVM Dipeptide	Algorithm (82.47%)	Allergen
Algpred - ARP	Similarity (88.95%)	Allergen
Proap - SVM aa	Algorithm (85.44%)	Allergen
ADFS - Motif-based	Similarity (96.0%)	Error
Allergen Online - 80 aa Window	Similarity (95.54%)	Allergen
Allergen Online - 8 aa Wordmatch	Similarity (98.3%)	Allergen

No centro da tabela observa-se a categorização das diferentes *web tools* conforme seu método de predição *in silico*. No canto superior esquerdo há a possibilidade de visualizar o resultado de cada ferramenta isoladamente. Quando um erro em uma das *web tools* ocorre, este resultado não é contabilizado no processamento final e o resultado impresso no *output* é “Error”.

Fonte: dados produzidos pelo autor (2016).

Para facilitar a interpretação dos resultados expostos pela PA<sup>3</sup>P são gerados gráficos (Figura 21).

Figura 21 - Gráficos para melhor visualização e interpretação dos resultados.



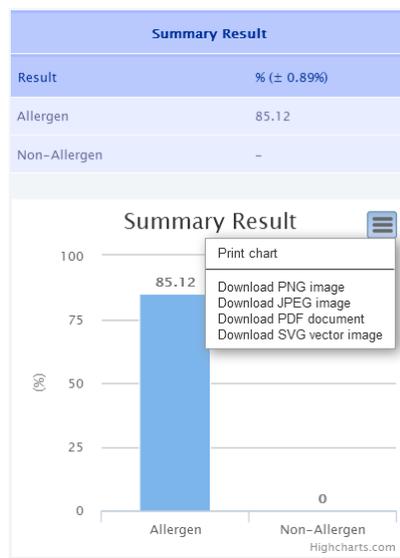
Da esquerda para direita: resultado dos métodos baseados em algoritmos, resultado dos métodos baseados em similaridade e um resultado sumarizado. Acima dos gráficos há os resultados tabelados e juntamente o erro padrão para cada método. O erro padrão ( $\text{erro padrão} = \text{desvio padrão} / \sqrt{\text{tamanho da amostra}}$ ) é necessário uma vez que trabalhamos com uma amostra representativa de uma população. Este erro pode ser

subtraído ou adicionado ao resultado, desta forma infere-se que uma população possui comportamento igual ao apresentado nos resultados.

Fonte: dados produzidos pelo autor (2016).

A diferença observada entre os resultados baseados em similaridade e os baseados em algoritmos é exclusivamente devido à diferença na acurácia de cada ferramenta isoladamente. Há a possibilidade de realizar o *download* dos resultados gráficos conforme ilustrado na Figura 22.

Figura 22 - *Download* dos resultados gráficos.



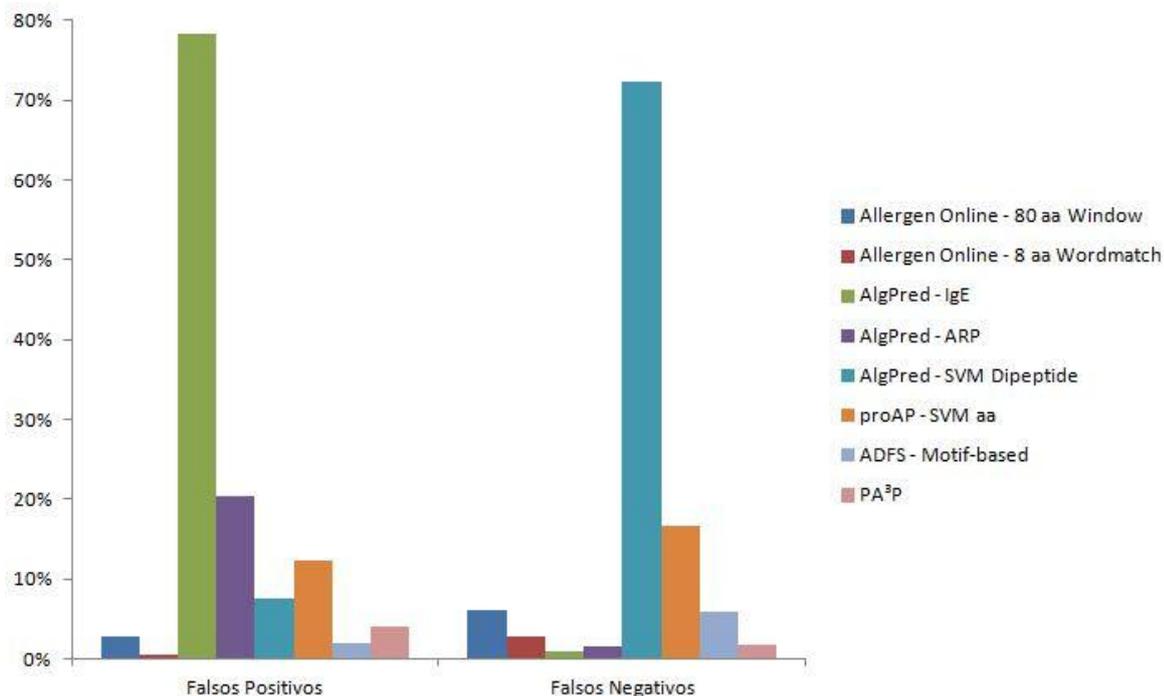
Clicando-se nas barras do quadrante superior direito do gráfico é possível realizar o download do mesmo em diferentes formatos.

Fonte: dados produzidos pelo autor (2016).

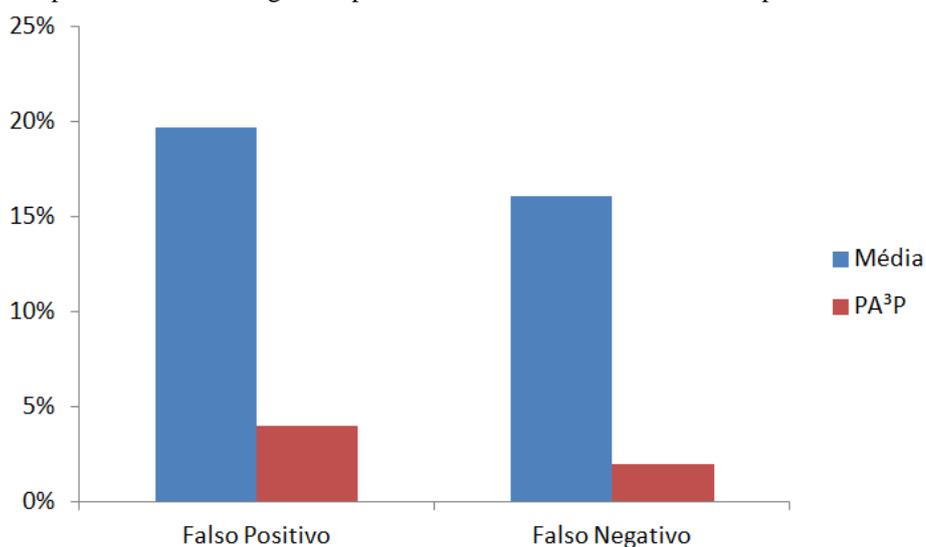
A análise demonstra um índice de 85,12 % ( $\pm 0,89$  %) para categorizar esta sequência de aminoácidos como alergênica e um índice de 0 % ( $\pm 0,89$  %) como não-alergênica. Desta forma, pode-se constatar que se trata de uma proteína potencialmente alergênica.

## 7. Discussão

O critério proposto pela FAO/WHO de 6 ou mais aminoácidos contíguos e idênticos (FAO/WHO, 2001) é criticado por produzir um grande número de falsos positivos (STADLER; STADLER, 2003; MIRSKY; CRESSMAN JUNIOR; LADICS, 2013) sendo sugerido que este número seja fixado em 8 aminoácidos contíguos (SÁNCHEZ-MONGE; SALCEDO, 2005). Stadler e Stadler (2003) indicam que utilizando o método proposto pela FAO/WHO, um a cada 200 resultados positivos são verdadeiramente positivos, ou seja, representam um alérgeno de fato. Estes resultados não condizem com os resultados obtidos pela PA<sup>3</sup>P onde se identificou uma redução de 24 % nos casos de falsos negativos e 15,06 % nos casos de falsos positivos quando submetido o *dataset* à análise, conforme Figura 23 (A e B).

Figura 23 - Redução dos casos de falsos positivos/negativos pelo uso da PA<sup>3</sup>P.

A Casos de falsos positivos e falsos negativos para as ferramentas *web* individuais e para PA<sup>3</sup>P.



B Média (em azul) refere-se a média dos resultados obtidos nas *web tools*. PA<sup>3</sup>P (em vermelho) refere-se ao resultado obtido quando utilizada a plataforma como ferramenta de categorização de proteínas.

Fonte: dados produzidos pelo autor (2016).

Wang et al. (2013) relata que quando alterado o critério de 6 para 8 aminoácidos contíguos ocorre um aumento da acurácia e da especificidade e uma redução da sensibilidade. Estes dados não são sustentados por nossas análises.

A ferramenta *Allermatch* (disponível em <http://www.allermatch.org>, descrito por descrito por FIERS et al., 2004) apresenta baixa sensibilidade (60 %) e alta especificidade (95 %) para o método de 6 aminoácidos contíguos. Por outro lado, utilizando como critério 8

aminoácidos contíguos na ferramenta *Allergen Online* obtém-se alta especificidade (99 %) e sensibilidade (97 %). Por esta razão, optamos por utilizar a ferramenta *Allergen Online* para satisfazer a metodologia proposta pela FAO/WHO para predição do potencial de alergenicidade de proteínas. Dimitrov et al. (2014) questiona o grande número de falsos alérgenos produzidos pelas ferramentas SDAP, *AllerTool* e *Allermatch*.

Durante o processo de validação da nova ferramenta não foi identificado nenhum caso de falso positivo, ou seja, todos os resultados positivos representavam um alérgeno. Este decréscimo nos casos de falsos positivos corrobora com o aumento substancial da sensibilidade e especificidade.

Em nossos testes, a PA<sup>3</sup>P apresentou 98 % de sensibilidade e 96 % de especificidade quando comparada com as análises das ferramentas isoladas (89 % e 85 %, respectivamente). Outro ponto favorável da plataforma, no que tange à confiabilidade da mesma, é o desvio padrão de 19 % para os resultados obtidos a partir do seu uso contra um desvio padrão de 35 % para as análises isoladamente.

Para Mirsky, Cressman Junior e Ladics (2013) há negligência nos resultados da predição *in silico* por homologia do potencial de alergenicidade de proteínas. Desta forma, é de suma importância que sejam combinados diferentes métodos para categorizar proteínas como alergênicas/não-alergênicas. Para sanar a carência dos resultados obtidos nos métodos baseados em similaridade, compilamos métodos baseados em algoritmos. Ferramentas com métodos baseados em algoritmos funcionam através da identificação de padrões nas famílias de amostras.

Existem diversas ferramentas disponíveis *online* para predição *in silico* do potencial de alergenicidade e antigenicidade de proteínas. Apesar de possuírem métodos parcialmente satisfatórios de predição, suas interfaces não são amigáveis e seus resultados são confusos. A PA<sup>3</sup>P foi construída para apresentar resultados sumarizados e de fácil entendimento até mesmo para leigos no assunto. Os resultados gráficos podem ser utilizados para justificar o uso ou não de determinada proteína uma vez que podem ser salvos em diferentes formatos diretamente da plataforma.

Uma alternativa ao uso desta metodologia de predição *in silico* é o acompanhamento das recomendações da FAO/WHO. O pesquisador, primeiramente, deve obter a sequência de aminoácidos de todas as proteínas comprovadamente ou supostamente alergênicas a partir de bancos de dados como, por exemplo, o SwissProt e o GenBank para construir um banco de dados em formato FASTA. Outro banco de dados deveria ser construído contendo sequências

de 80 aminoácidos cada, derivados da sequência da proteína que está sendo analisada. Por fim, deve comparar cada uma das sequências do banco de dados (que contém as sequências de 80 aminoácidos) com todas as sequências do banco de dados composto de todas as proteínas comprovadamente ou supostamente alergênicas. Com base nos resultados dos alinhamentos, uma proteína seria considerada possivelmente alergênica quando houvesse mais de 35% de identidade em uma janela de 80 aminoácidos ou houvesse identidade de 6 aminoácidos contíguos (FAO/WHO, 2001). Contudo, apenas a identidade entre as sequências não é suficiente para atestar o potencial alergênico de determinada proteína. Conforme descrito por Stadler & Stadler (2003), quando utilizada a metodologia proposta pela FAO/WHO, 1 a cada 200 resultados positivos são verdadeiramente positivos, ou seja, representam um alérgeno. Visando agregar valor as predições *in silico* é que as ferramentas que foram compiladas neste trabalho foram criadas.

## 8. Conclusões e Perspectivas

A PA<sup>3</sup>P consiste em uma ferramenta que foi capaz de categorizar proteínas alergênicas ou antigênicas, de acordo com as características biológicas originárias da expressão da sequência de aminoácidos. A PA<sup>3</sup>P permite que análises *in silico* fossem realizadas e que potenciais alergênicos ou antigênicos fossem definidos concomitantemente a uma redução dos casos de falsos positivos/falsos negativos bem como uma elevada sensibilidade e especificidade. Esta ferramenta poderá contribuir para embasar a decisão de utilizar ou não determinada proteínas na construção de um Organismo Geneticamente Modificado com finalidade nutricional.

**Referências**

- ABBAS, A.K.; LICHTMAN, A.H.; PILLAI, S.; **Imunologia Celular e Molecular**. 6ª Edição. Rio de Janeiro: Elsevier. 2008.
- ALBERTS, B.; et al.; **Biologia Molecular da Célula**. 5ª Edição. Porto Alegre: Artmed. 2010.
- AMARAL, CLF; SILVA, AB. **Melhoramento Biotecnológico de Plantas Medicinais**. Rev Biotecnologia Ciência e Desenvolvimento, n. 30, 55- 59. 2003.
- ANDRADE, P.P.; PARROTT, W.; ROCA, M.M.; **Guia para a Avaliação do Risco Ambiental de Organismos Geneticamente Modificados**. 1ª Edição. São Paulo: International Life Sciences Institute do Brasil. 2012.
- ANDRADE, S.R.M.; **Transformação de Plantas**. 1ª Edição. Planaltina: Embrapa Cerrados. 2003.
- BORÉM, A.; **Variedades Transgênicas e Meio Ambiente – Segurança Ambiental das Variedades Comerciais**. Rev Biotecnologia Ciência e Desenvolvimento, n. 34, 91-99. 2005.
- BRASIL. Decreto nº 4.680, de 24 de abril de 2003. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, v. 140, n. 79, 24 de abril de 2003. Seção 1, p. 2. 2003<sup>a</sup>.
- BRASIL. Lei nº 11.105, de 24 de março de 2005. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, v. 142, n. 58, 28 de março de 2005. Seção 1, p. 1-5. 2005.
- BRASIL. Lei nº 8.974, de 5 de janeiro de 1995. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, v. 133, n. 5, 6 de janeiro de 1995. Seção 1, p. 1-3. 1995.
- BRASIL. Portaria nº 2.658, de 22 de dezembro de 2003. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, v. 140, n. 251, 26 de dezembro de 2003. Seção 1, p. 13-14. 2003<sup>b</sup>.
- BRASIL. Resolução Normativa nº 1, de 20 de junho de 2006. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, v. 142, n. 117, 21 de junho de 2006. Seção 1, p. 7-8. 2006<sup>a</sup>.

- BRASIL. Resolução Normativa nº 2, de 27 de novembro de 2006. **Diário Oficial da República Federativa do Brasil**, Brasília, DF, v. 142, n. 227, 28 de novembro de 2006. Seção 1, p. 90-93. 2006<sup>b</sup>.
- BRASILEIRO, ANA C.M.; LACORTE, CRISTIANO; **Agrobacterium: um sistema natural de transferência de genes para plantas**. Rev Biotecnologia Ciência e Desenvolvimento, n. 15, 12-15. 2000.
- BROWN, TERENCE A.; **Gene cloning and DNA analysis: an introduction**. Edição nº 6. New York: Wiley-Blackwell. 2010.
- BRUIJNZEEL-KOOMEN, C; et al. **Adverse Reactions to Food: Position Paper**. Allergy, v. 50, 623 – 635. 1995.
- CAMPILLO-NAVARRO, M; et al. **Mast Cells in Lung Homeostasis: Beyond Type I Hypersensitivity**. Curr Respir Med Rev, v. 10, n. 2, 115 – 123. 2014.
- CANHOTO, JORGE M; **Biotecnologia Vegetal – da Clonagem de Plantas à Transformação Genética**. Coimbra: Imprensa da Universidade de Coimbra. 1ª Edição. 2010.
- CARAGEA, C; et al. **Glycosylation Site Prediction Using Ensembles of Support Vector Machine Classifiers**. BMC Bioinformatics, v. 8, n. 438, 1 – 13. 2007.
- CHAPMAN, M.D.; et al.; **Nomenclature And Structural Biology Of Allergens**. Allergy Clinical Immunology, v.199, 414-420. 2007.
- COHEN, JACQUES. **Bioinformatics – an introduction for computer scientists**. ACM Computing Surveys, v.36, n.2, 122-158. 2004.
- CORDEIRO, MARIA C.R.; **Engenharia Genética: conceitos básicos, ferramentas e aplicações**. Planaltina: Embrapa Cerrados. 2003.
- COSTA, T. E. M. M.; et al.; **Avaliação de Risco dos Organismos Geneticamente Modificados**. Ciência e Saúde coletiva, v.16, 327-336. 2011.
- DEITEL, H.M.; DEITEL, P.; **Java: como programar**. 6ª Edição. São Paulo: Pearson Education do Brasil. 2005.

- DIMITROV, I.; et al.; **Allergen FP: allergenicity prediction descriptor fingerprints.** Bioinformatics, v.30, n.6, 846-851. 2014.
- FALEIRO, FÁBIO G.; ANDRADE, SOLANGE R.M.; **Biotechnologia, Transgênicos e Biossegurança.** Planaltina: Embrapa Cerrados. 2009.
- FAO/WHO. **Evaluation of Allergenicity of Genetically Modified Foods.** Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology. 2001.
- FIERS, M.W.; et al.; **Allermatch™, a Webtool for the Prediction of Potential Allergenicity According to Current FAO/WHO Codex Alimentarius Guidelines.** BMC Bioinformatics, v.5, 1-6. 2004.
- GALLI, SJ; et al. **Mast Cells and IgE in Defense Against Venoms: Possible “Good Side” of Allergy?** Allergology International, v. 65, 3 – 15. 2016.
- HJÄLTÉN, J; et al.; **Innate and Introduced Resistance Traits in Genetically Modified Aspen Trees and Their Effect on Leaf Beetle Feeding.** Plos One, v.8, 1-7. 2013.
- HUG, K.; **Genetically Modified Organisms: do the Benefits Outweigh the Risks?** Medicina, v. 44, n.2, 87-99. 2008.
- KLETER, G. A.; KUIPER, H. J.; **Considerations for the Assessment of the Safety of Genetically Modified Animals Used for Human Food or Animal Feed.** Livestock Production Science, v. 74, 275-285. 2002.
- LONGO, G; et al. **IgE-Mediated Food Allergy in Children.** Lancet, v. 382, 1656 - 64. 2013.
- LORENA, A.C.; CARVALHO, A.C.P.L.F.; **Uma introdução às Support Vector Machines.** Rev de Informática Teórica e Aplicada, v.14, n.2, 43-67. 2007.
- LUSCOMBE, N.M.; GREENBAUM,D.; GERSTEIN, M.; **What is Bioinformatics? A proposed definition and overview of the field.** Method Inform Med, v.40, n.4, 346-358. 2001.
- MARI, A. et al.; **Allergen Databases: Current Status and Perspectives.** Current allergy and asthma reports, v.5, 376-383. 2009.

- MIRSKY, HENRY P.; CRESSMAN JUNIOR, ROBERT F.; LADICS, GREGORY S.; **Comparative assessment of multiple criteria for the *in silico* prediction of cross-reactivity of proteins to known allergens.** Regul Toxicol Pharmacol, v. 67, n.2, 232-239. 2013.
- MOHABATKAR, H. et al.; **Prediction of Allergenic Proteins by Means of the Concept of Chou's Pseudo Amino Acid Composition and a Machine Learning Approach.** Medicinal Chemistry, v.9, 133-137. 2013.
- NAKAMURA, R.; NAKAMURA, R.; TESHIMA,R.; **Major revision of the allergen database for food safety (ADFS) and validation of the motif- based allergenicity prediction tool.** Kokuritsu Iyakuhin Shokuhin Eisei Kenkyusho Hokoku, v.127, 44-49, 2009.
- NELSON, D.L.; COX, M.M.; **Lehninger Princípios de Bioquímica.** 3ª Edição. São Paulo: Sarvier. 2002.
- NODARI, R.O.; GUERRA, M.P.; **Avaliação de riscos ambientais de plantas transgênicas.** Cadernos de Ciência e Tecnologia, v.18, 81-116. 2001.
- NONOHAY, JULIANA S.; **Transformação Genética em Cevada por Bombardeamento de Partículas.** Tese apresentada ao Programa de Pós-graduação em Genética e Biologia Molecular da Universidade Federal do Rio Grande do Sul para obtenção do título de Doutor em Ciências. 2002.
- PROSDOCIMI, F; et al. **Bioinformática: Manual do Usuário.** Rev Biotecnologia Ciência e Desenvolvimento, v.29, 12 - 25. 2002.
- RAVEN, P.H. et al.; **Biologia Vegetal.** 5ª Edição. Rio de Janeiro: Guanabara Koogan. 1996.
- ROTH-WALTER, F.; et al.; **Bet v 1 from birch pollen is a lipocalin-like protein acting as allergen only when devoid of iron by promoting Th2 lymphocytes.** J Biol Chem, v.289, n.25, 17416 - 17421. 2014.
- ROTH-WALTER, F; et al. **Bet V 1 from Birch Pollen is a Lipocalin-Like Protein Acting as Allergen Only When Devoid of Iron by Promoting Th2 Lymphocytes.** J Biol Chem, v. 289, n. 25, 17416 – 21. 2014.
- SAHA, S.; RAGHAVA, G.P.S.; **AlgPred: Prediction of Allergenic Proteins and Mapping of IgE Epitopes.** Nucleic Acids Research, v.34, p.202-209. 2006.

- SAMPSON, HA. **Food Allergy. Part 1: Immunopathogenesis and Clinical Disorders.** J Allergy Clin Immunol, v. 103, n. 5, 717 – 728. 1999.
- SÁNCHEZ-MONGE, ROSA; SALCEDO, GABRIEL; **Analytical methodology for assessment of food allergens: opportunities and challenges.** Biotechnology Advances, v.23, 415-422. 2005.
- SARTORETTO, LAUDETE M.; SALDANHA, CLEBER W.; CORDER, MAISA P.M.; **Transformação Genética: estratégias e aplicações para o melhoramento genético de espécies florestais.** Revista Ciência Rural, v.38, n.3, 861-871. 2008.
- SCHOLZE, SIMONE H.C.; **Biossegurança e Alimentos Transgênicos – o papel da CTNBio.** Rev Biotecnologia Ciência e Desenvolvimento, n.9, 32-34. 1999.
- SHRADER-FRECHETTE, K.; **Property Rights and Genetic Engineering: developing nations at risk.** Science and Engineering Ethics, v.1, 137-149. 2005.
- SICHERER, SH; SAMPSON, HA. **Food Allergy.** J Allergy Clin Immunol, v. 126, n. 2, 116 – 125. 2010.
- SOLÉ, D; et al. **Consenso Brasileiro sobre Alergia Alimentar: 2007.** Rev Brasi Alerg Immunopatol, v. 31, n. 2, 64 – 89. 2008.
- SOUZA JUNIOR, M.T.; MARTINS, N.F.; **Predição do Potencial de Alergenicidade em OGMs – Estudo de Caso.** Rev Biotecnologia Ciência e Desenvolvimento, v.30, 10-15. 2003.
- STADLER, M.B.; STADLER, B.M.; **Alergenicidade Prediction by Protein Sequence.** The FASEB Journal, v.17, 1141-1143. 2003.
- TAYLOR, S. L. **Chemistry and detection of food allergens.** Food Technol, v.46, n.5, 146-152. 1992.
- TOIT, G; et al. **Prevention of Food Allergy.** J Allergy Clin Immunol, v. 137, n. 4, 998 – 1010. 2016.
- TSAI, M; et al. **Testing the ‘Toxin Hypothesis of Allergy’: Mast Cells, IgE, and Innate and Acquired Immune Responses to Venoms.** Curr Opin Immunol, v. 36, 80 – 87. 2015.

VON KRIES, C.; WINTER, G. **The Structuring of GMO Release and Evaluation in EU law.** Journal Biotech, v.4, 569-581. 2011.

WANG, J.; et al.; **Evaluation and Integration of Existing Methods for Computational Prediction of Allergens.** BMC Bioinformatics, v.14, 1-9. 2013.

WASERMAN, S; WATSON, W. **Food Allergy.** Allergy Asthma Clin Immunol, v. 10, n. 7, 1 – 7. 2011.

XIONG, J.; **Essential Bioinformatics.** 1<sup>a</sup> Edição. New York: Cambridge University Press. 2006.

## **APÊNDICES**

**Apêndice 1: manuscrito gerado a partir deste trabalho.**

ScholarOne Manuscripts

Página 1 de 2



Bioinformatics



---

## Submission Confirmation

[Print](#)

---

Thank you for your submission

---

**Submitted to**

Bioinformatics

**Manuscript ID**

BIOINF-2016-0139

**Title**PA<sup>3</sup>P – Platform for Analysis of Allergenicity and Antigenicity of Proteins**Authors**

Boeck, Anna Carolina

Dotto, Bruno

Wallau, Gabriel

Pinto, Paulo

Schenkel Homrich, Milena

Boldo, Juliano

**Date Submitted**

29-Jan-2016

---

---

[Author Dashboard](#)

**PA<sup>3</sup>P – Platform for Analysis of Allergenicity and Antigenicity of Proteins**

Journal:	<i>Bioinformatics</i>
Manuscript ID	Draft
Category:	Applications Note
Date Submitted by the Author:	n/a
Complete List of Authors:	Boeck, Anna Carolina; Universidade Federal do Pampa - Campus Sao Gabriel, Programa de Pós-Graduação em Ciências Biológicas Dotto, Bruno; Universidade Federal do Pampa - Campus Sao Gabriel Wallau, Gabriel; Centro de Pesquisas Aggeu Magalhaes, Departamento de Entomologia Pinto, Paulo; Universidade Federal do Pampa - Campus Sao Gabriel, Programa de Pós-Graduação em Ciências Biológicas Schenkel Homrich, Milena; Universidade do Vale do Rio dos Sinos Boldo, Juliano; Universidade Federal do Pampa - Campus Sao Gabriel, Programa de Pós-Graduação em Ciências Biológicas
Keywords:	Web services, Protein sequence analysis, Allergenicity

Database and ontologies

## PA<sup>3</sup>P – Platform for Analysis of Allergenicity and Antigenicity of Proteins

Anna Carolina Boeck<sup>1</sup>, Bruno Reis Dotto<sup>2</sup>, Gabriel Luz Wallau<sup>3</sup>, Paulo Marcos Pinto<sup>1</sup>, Milena Schenkel Homrich<sup>4</sup>, Juliano Tomazzoni Bolo<sup>1\*</sup>

<sup>1</sup> Programa de Pós-Graduação em Ciências Biológicas, Universidade Federal do Pampa, São Gabriel, 97300-000, RS, Brazil

<sup>2</sup> Universidade Federal do Pampa, São Gabriel, 97300-000, RS, Brazil

<sup>3</sup> Departamento de Entomologia, Centro de Pesquisas Aggeu Magalhães—FIOCRUZ-CPqAM, Recife, 50740-465, PE, Brazil

<sup>4</sup> Universidade do Vale dos Sinos, São Leopoldo, 93022-000, RS, Brazil.

\*julianoboldo@unipampa.edu.br

Received on XXXXX, revised on XXXXX, accepted on XXXXX

Associate Editor: XXXXXXXX

### ABSTRACT

**Summary:** The Platform for Analysis of Allergenicity and Antigenicity of Proteins (PA<sup>3</sup>P) was developed with the goal of providing a comprehensive set of tests assembled into one platform which has a clear output for bench scientists with limited knowledge of bioinformatics. PA<sup>3</sup>P is a robust web tool capable of categorizing allergenic or antigenic proteins by summarizing and statistically testing each output derived from a set of specific web tools prior to genetically modified organisms construction.

**Availability and Implementation:** This platform is freely available on the web at <http://pa.saogabriel.unipampa.edu.br:8080/pa3p> or <http://pluriserver.com.br/pa3p/>. PA<sup>3</sup>P was implemented in HTML and Java with all major browsers supported. Tools and software packages used are free for personal or non-profit projects.

**Contact:** julianoboldo@unipampa.edu.br

**Supplementary Information:** For more information on how PA<sup>3</sup>P works, please visit the "About" section in <http://pa.saogabriel.unipampa.edu.br:8080/pa3p/about/about.jsp> or <http://pluriserver.com.br/pa3p/about/about.jsp>. PA<sup>3</sup>P design and analyses pipeline is displayed in "About" section.

### 1 INTRODUCTION

Genomes of plants and animals have been altered by traditional breeding techniques for many years. Although artificial selection was proven efficient, undesirable traits and time-consuming experiments drove researchers to develop more direct and controlled modification methods. Today, it is possible to introduce one specific foreign gene to a recipient genome, rendering new characteristics to the genetically modified organism (GMO) (Hjältén et al. 2013). The major drawbacks when designing GMO are the potential development of toxicity, allergenicity and antigen-

icity, nutritional alterations, and the possibility of horizontal gene transfer (Costa et al. 2011). Thus, such drawbacks must be assessed prior to GMO construction and environmental and commercial releases.

New transgenic proteins allergenicity/antigenicity must be investigated before GMO construction and introduction in human diet (Fiers et al. 2004). The allergenicity potential of a given protein can be assessed by *in silico* (Mohabatkar et al. 2013), *in vivo*, and *in vitro* tests. However, there is not a single-testing method for allergenic potential evaluation that is fully reliable (Zorzet et al. 2002).

Concerned about the safe use of transgenic proteins, both Food and Agriculture Organization of the United Nations (FAO) and World Health Organization (WHO) defined, in 2001, that proteins should be tested *in silico* for its allergenicity and antigenicity potential prior to GMO construction. In order to assess the allergenic potential of a given protein, both WHO and FAO determined that a query protein is potentially an allergen if it either presented six or more contiguous amino acids (id length  $n=6$ ) or over 35% identity in a 80 amino acid window when compared with known allergens (FAO/WHO 2001).

In order to meet the allergenic potential assessment requirement, several bioinformatics approaches were developed (Wang et al. 2013). Within the context of new approaches to determine the allergenic and antigenic potential of proteins more accurately, the Platform for Analysis of Allergenicity and Antigenicity of Proteins (PA<sup>3</sup>P) was developed. PA<sup>3</sup>P unifies the FAO/WHO suggested tests, analysis by aligning the query sequence against several databases composed by proven or potential allergenic and antigenic amino acid sequences, and analysis using different epitopes and motifs prediction algorithms.

\*To whom correspondence should be addressed.

Boeck et al.

In addition, PA<sup>3</sup>P summarizes the results obtained from several web tools, allowing accurate and simpler interpretation for the end user.

## 2 DATA MINING AND METHODS

### 2.1 Dataset and Accuracy

Considering the sampling effort calculation for a finite population,  $n$  was fixed in 475 allergenic proteins amino acid sequences and 466 non-allergenic proteins sequences. All allergenic protein sequences were obtained from the Allergen Online database (<http://allergenonline.com/>). Non-allergenic protein sequences were obtained from the National Center for Biotechnology Information (NCBI) database using “immunogenic” and “*Homo sapiens*” as keywords. The complete dataset ( $n=941$  sequences) was subjected to every web tool allergenic evaluation in order to determine each web tool accuracy. For this version, we used Allergen Online – 8 aa wordmatch; Allergen Online – 80 aa window; AlgPred – IgE (Saha and Raghava, 2006); AlgPred – ARP (Saha and Raghava, 2006); and ADFS – Motif-based (Nakamura et al. 2009); SVM-amino acid composition (Wang et al. 2012); and (ii) SVM dipeptide (Saha and Raghava 2006) web tools.

### 2.2 Platform Development

PA<sup>3</sup>P was implemented in HTML and Java, running into a Web Server (Apache Tomcat 7 Server). We used Eclipse Java EE IDE as the developing tool. As programming languages, we used Java SE - 1.7, HTML, JavaScript, and as plugins we used JQuery, JQuery UI and Highcharts JS. Tools and software packages used are free for personal or non-profit projects. The “About” section presents the PA<sup>3</sup>P workflow in detail.

### 2.3 Validation

After the platform development, the  $n=941$  dataset was subjected to PA<sup>3</sup>P analysis, allowing sensibility and specificity determination of our platform compared with the web tools used independently. After, a dataset composed by 100 proteins with unknown allergenicity and antigenicity potential was subjected to PA<sup>3</sup>P and web tools for further platform validation.

## 3 USING PA<sup>3</sup>P

The query sequence is input in the proper space, in the context of the “PA<sup>3</sup>P” tab. After hitting “Run”, the sequence is analyzed by the implemented tools. Returning outputs are organized in a table and presented as “allergen” or “non-allergen”. A chart summarizing the result is presented. The end user is able to export the chart in different file formats.

## 4 RESULTS AND DISCUSSION

The methods proposed by FAO/WHO for allergenicity and antigenicity potential are highly criticized by several authors. Stadler and Stadler (2003) demonstrated when the methodology is employed, one out of 200 positive results is a true positive, in other words, FAO/WHO methodology has a high false positive rate.

PA<sup>3</sup>P reduced 24 and 15.06 % of false positives and false negatives cases, respectively, when compared with the web tools used independently. In the PA<sup>3</sup>P validation process, no false negative results were obtained and all true positives represented a known allergen. This reduction in the false positive cases is consistent with the increase in the sensibility and specificity percentages. In our tests, PA<sup>3</sup>P presented 98% sensibility and 96% specificity when compared with the web tools used independently (89 and 85%, respectively).

Concluding, PA<sup>3</sup>P consists in a free, web based platform, that is highly reliable in categorizing proteins as allergenic or antigenic based on its amino acid sequence with the reduction in false positive and false negative cases as well as with high sensibility and specificity. This tool will contribute to support the decision whether to use or discard a given protein in GMO construction with nutritional intent.

## 5 AVAILABILITY AND LONG-TERM MAINTENANCE

The platform is available at the web server of the Universidade Federal do Pampa and will be maintained by this university for an unlimited time. Also, PA<sup>3</sup>P will be maintained at <http://pluriserver.com.br> as an alternative link in order to keep providing access if the main address becomes offline. PA<sup>3</sup>P is a free platform, which is publicly available at (<http://pa.saogabriel.unipampa.edu.br:8080/pa3p> or <http://pluriserver.com.br/pa3p/>). Other tools and methods may be added in the future.

Funding information: This work was supported by the funding agencies Coordenação de Aperfeiçoamento de Pessoal de Nível Superior; and Conselho Nacional de Desenvolvimento Científico e Tecnológico [473528/2013-4].

## REFERENCES

- Costa, T.E.M.M. et al. (2011) Avaliação de Risco dos Organismos Geneticamente Modificados. *Cienc Saude Coletiva*, 16, 327-336.
- FAO/WHO (2001) Evaluation of Allergenicity of Genetically Modified Foods. Report of a joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology.
- Fiers, M.W. et al. (2004) Allermatch™, a Webtool for the Prediction of Potential Allergenicity According to Current FAO/WHO Codex Alimentarius Guidelines. *BMC Bioinformatics*, 5, 1-6.
- Hjältén, J. et al. (2013) Innate and Introduced Resistance Traits in Genetically Modified Aspen Trees and Their Effect on Leaf Beetle Feeding. *PLoS ONE*, 8, e73819.
- Mohabatkar, H. et al. (2013) Prediction of Allergenic Proteins by Means of the Concept of Chou's Pseudo Amino Acid Composition and a Machine Learning Approach. *Med Chem*, 9, 133-137.
- Saha, S. and Raghava, G.P.S. (2006) AlgPred: Prediction of Allergenic Proteins and Mapping of IgE Epitopes. *Nucleic Acids Res*, 34, 202-209.
- Stadler, M.B. and Stadler, B.M. (2003) Allergenicity Prediction by Protein Sequence. *EASEB J*, 17, 1141-1143.
- Wang, J. et al. (2013) Evaluation and Integration of Existing Methods for Computational Prediction of Allergens. *BMC Bioinformatics*, 14, 1-9.
- Zorzet, A. et al. (2002) Prediction of Food Protein Allergenicity: a Bioinformatics Learning Systems Approach. *In Silico Biol*, 2, 525-534.

**Apêndice 2: Planilha do cálculo da acurácia e validação da PA<sup>3</sup>P via *dataset*.**

Tabela 6 - Matriz binária ( 0 = presença; 1 = ausência) relativa ao processo de determinação da acurácia das *web tools* e validação da PA<sup>3</sup>P.

Sequência de Aminoácido <sup>1</sup>	1 <sup>2</sup>	2 <sup>3</sup>	3 <sup>4</sup>	4 <sup>5</sup>	5 <sup>6</sup>	6 <sup>7</sup>	7 <sup>8</sup>	8 <sup>9</sup>
1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1
4	1	1	1	1	1	1	1	1
5	1	1	1	1	1	1	1	1
6	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	1
9	0	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1
12	0	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1
18	1	1	0	1	1	1	1	1
19	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1

<sup>1</sup> Sequências de aminoácidos referentes ao *dataset*. De 1- 475 são sequências de aminoácidos de proteínas alergênicas; de 476 – 941 são sequências de aminoácidos de proteínas não-alergênicas.

<sup>2</sup> AlgPred - IgE

<sup>3</sup> AlgPred - SVM dipeptide

<sup>4</sup> AlgPred - ARPs

<sup>5</sup> proAP – SVM aa

<sup>6</sup> ADFS – Motif-based

<sup>7</sup> Allergen Online – 80 aa window

<sup>8</sup> Allergen Online – 8 aa wordmatch

<sup>9</sup> PA<sup>3</sup>P

31	1	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1	1
33	0	1	0	0	1	0	0	0
34	1	1	1	1	1	1	1	1
35	0	1	0	0	1	0	1	0
36	0	1	0	0	1	1	1	1
37	0	1	1	1	1	1	1	1
38	0	1	1	1	1	1	1	1
39	1	1	1	0	1	1	1	1
40	0	1	1	1	1	1	1	1
41	0	1	1	1	1	1	1	1
42	0	1	1	1	1	1	1	1
43	0	1	1	1	1	1	1	1
44	0	1	1	1	1	1	1	1
45	0	1	1	1	1	1	1	1
46	0	1	1	1	1	1	1	1
47	1	1	1	1	1	1	1	1
48	0	1	1	1	1	1	1	1
49	0	1	1	1	1	1	1	1
50	1	1	1	1	1	1	1	1
51	0	1	1	1	1	1	1	1
52	0	1	1	1	1	1	1	1
53	0	1	0	1	1	1	1	1
54	0	1	1	1	1	1	1	1
55	0	1	1	1	1	1	1	1
56	0	1	1	1	1	1	1	1
57	0	0	1	1	1	1	1	1
58	0	1	0	1	1	1	1	1
59	0	1	1	1	1	1	1	1
60	0	1	1	1	1	1	1	1
61	0	1	1	1	1	1	1	1
62	0	1	0	1	1	1	1	1
63	0	1	1	1	1	1	1	1
64	0	1	1	1	1	1	1	1
65	0	1	1	1	1	1	1	1
66	0	1	1	0	1	1	1	1
67	0	1	1	1	1	1	1	1
68	0	1	1	1	1	1	1	1
69	0	1	1	1	1	1	1	1
70	0	1	1	1	1	1	1	1
71	0	1	0	1	1	1	1	1
72	0	1	1	1	1	1	1	1
73	0	0	1	1	1	1	1	1
74	0	1	1	0	1	1	1	1
75	0	1	1	1	1	1	0	1
76	0	1	1	1	1	1	1	1

77	1	1	0	1	1	1	1	1
78	0	1	1	1	1	1	1	1
79	0	1	1	1	1	1	1	1
80	0	1	1	1	1	1	1	1
81	0	1	1	1	1	1	1	1
82	1	1	1	1	1	1	1	1
83	0	1	1	1	1	1	1	1
84	0	1	1	1	1	1	1	1
85	0	1	1	1	1	1	1	1
86	0	0	0	1	1	1	1	1
87	0	1	0	1	0	1	1	1
88	1	1	1	1	1	1	1	1
89	0	1	1	1	1	1	1	1
90	0	1	1	1	1	1	1	1
91	0	1	0	1	1	1	1	1
92	0	1	1	1	1	1	1	1
93	0	1	1	1	1	1	1	1
94	1	1	1	1	1	1	1	1
95	0	1	1	1	1	1	1	1
96	0	1	1	1	1	1	1	1
97	0	1	1	1	1	1	1	1
98	0	1	1	1	1	1	1	1
99	0	1	1	1	1	1	1	1
100	0	0	0	0	0	1	1	0
101	0	1	1	1		1	1	1
102	0	1	1	1		1	1	1
103	0	1	1	1		1	1	1
104	0	1	1	1		1	1	1
105	0	1	1	1		1	1	1
106	0	1	1	1		1	1	1
107	0	1	0	1		1	1	1
108	0	1	1	1		1	1	1
109	0	1	0	1		1	1	1
110	0	1	1	1		1	1	1
111	0	1	0	1		1	1	1
112	0	1	1	1		1	1	1
113	0	1	1	1		1	1	1
114	0	1	1	1		1	1	1
115	0	1	1	1		1	1	1
116	0	1	1	1		1	1	1
117	0	1	0	1		1	1	1
118	1	1	1	1		1	1	1
119	1	1	1	1		1	1	1
120	0	1	1	0		1	1	1
121	1	1	1	1		1	1	1
122	0	1	1	1		1	1	1

123	0	1	1	1	1	1	1
124	0	1	1	1	1	1	1
125	0	1	1	1	1	1	1
126	1	1	1	1	1	1	1
127	0	1	1	1	1	1	1
128	0	1	1	1	1	1	1
129	0	1	1	1	1	1	1
130	0	1	1	1	1	1	1
131	0	1	1	1	1	1	1
132	0	1	1	1	1	1	1
133	0	1	0	1	1	1	1
134	0	1	0	1	1	1	1
135	0	1	1	1	1	1	1
136	0	1	0	1	1	1	1
137	0	1	1	1	1	1	1
138	0	1	1	1	1	1	1
139	0	1	1	1	1	1	1
140	0	1	0	1	1	1	1
141	1	1	1	1	1	1	1
142	0	1	1	1	1	1	1
143	0	1	1	1	1	1	1
144	0	1	1	0	1	1	1
145	0	0	0	1	1	1	1
146	0	1	0	0	1	1	1
147	1	1	1	1	1	1	1
148	0	1	1	1	1	1	1
149	0	0	0	1	1	1	1
150	0	1	1	1	1	1	1
151	1	1	1	1	1	1	1
152	0	1	1	1	1	1	1
153	0	0	1	1	1	1	1
154	0	0	1	0	1	1	1
155	0	0	1	1	1	1	1
156	0	1	1	0	1	1	1
157	0	1	1	1	1	1	1
158	0	1	0	1	1	1	1
159	1	1	1	1	1	1	1
160	0	1	1	1	1	1	1
161	0	1	1	1	1	1	1
162	0	1	1	1	1	1	1
163	0	0	0	1	1	1	1
164	0	1	1	1	1	1	1
165	0	1	1	1	1	1	1
166	0	1	0	1	1	1	1
167	0	0	0	1	1	1	1
168	1	1	1	1	1	1	1

169	0	1	1	1	1	1	1
170	0	1	1	1	1	1	1
171	0	1	1	1	1	1	1
172	0	1	1	1	1	1	1
173	0	1	1	1	1	1	1
174	0	1	1	1	1	1	1
175	0	1	1	1	1	1	1
176	0	1	1	1	1	1	1
177	0	1	0	1	1	1	1
178	0	1	1	1	1	1	1
179	0	1	0	1	1	1	1
180	0	1	1	1	1	1	1
181	0	0	0	0	1	1	0
182	0	1	0	1	1	1	1
183	0	1	1	1	1	1	1
184	1	1	1	1	1	1	1
185	0	1	0	1	1	1	1
186	0	1	1	1	1	1	1
187	0	1	1	1	1	1	1
188	0	1	0	1	1	1	1
189	0	1	1	1	1	1	1
190	0	1	1	1	1	1	1
191	0	1	1	1	1	1	1
192	1	1	1	1	1	1	1
193	0	1	1	1	1	1	1
194	0	1	0	1	1	1	1
195	1	1	0	0	1	1	1
196	0	1	1	1	1	1	1
197	0	0	0	1	0	1	0
198	1	1	1	1	1	1	1
199	0	1	0	0	0	1	0
200	0	1	1	1	1	1	1
201	1	1	1	1	1	1	1
202	0	1	1	0	1	1	1
203	0	1	1	1	1	1	1
204	0	0	0	0	1	1	0
205	0	1	1	1	1	1	1
206	1	1	1	1	1	1	1
207	0	1	1	1	1	1	1
208	0	1	1	1	1	1	1
209	0	1	1	1	1	1	1
210	0	1	1	1	1	1	1
211	0	1	1	1	1	1	1
212	0	1	1	1	1	1	1
213	0	1	1	1	1	1	1
214	0	1	1	0	1	1	1

215	0	1	1	1	1	1	1
216	0	1	1	1	1	1	1
217	0	1	1	1	1	1	1
218	0	0	0	1	1	1	1
219	0	1	1	1	1	1	1
220	0	1	0	1	1	1	1
221	0	1	1	1	1	1	1
222	0	1	1	1	1	1	1
223	1	1	1	1	1	1	1
224	0	0	1	0	1	1	1
225	0	1	1	1	1	1	1
226	0	1	1	1	1	1	1
227	0	1	1	1	1	1	1
228	0	1	0	1	1	1	1
229	0	1	1	0	1	1	1
230	1	1	1	1	1	1	1
231	0	1	1	1	1	1	1
232	0	1	1	1	1	1	1
233	0	1	1	1	1	1	1
234	0	1	1	1	1	1	1
235	0	1	0	1	1	1	1
236	0	1	0	1	1	1	1
237	0	1	1	1	1	1	1
238	0	1	0	0	1	1	1
239	0	1	1	1	1	1	1
240	1	1	1	1	1	1	1
241	1	1	1	1	1	1	1
242	1	0	1	1	1	1	1
243	0	1	1	1	1	1	1
244	0	1	0	1	1	1	1
245	0	1	1	1	1	1	1
246	0	1	1	1	1	1	1
247	1	1	1	1	1	1	1
248	1	1	0	1	1	1	1
249	0	1	1	1	1	1	1
250	0	1	1	1	1	1	1
251	0	1	0	0	1	1	1
252	1	1	1	1	1	1	1
253	0	0	1	0	1	1	1
254	1	1	1	1	1	1	1
255	0	1	1	0	1	1	1
256	0	1	0	0	1	1	1
257	0	1	1	1	1	1	1
258	1	1	1	1	1	1	1
259	1	1	1	1	1	1	1
260	0	0	1	1	1	1	1

261	0	1	1	1	1	1	1
262	0	1	0	0	1	1	1
263	0	1	0	1	1	1	1
264	0	1	1	1	1	1	1
265	0	1	1	1	1	1	1
266	0	1	0	1	1	1	1
267	0	1	1	1	1	1	1
268	0	1	1	1	1	1	1
269	1	1	1	1	1	1	1
270	0	1	0	1	1	1	1
271	0	1	1	1	1	1	1
272	0	1	1	1	1	1	1
273	0	1	1	1	1	1	1
274	0	1	1	1	1	1	1
275	0	1	1	1	1	1	1
276	0	1	1	1	1	1	1
277	0	1	0	1	1	1	1
278	1	1	1	1	1	1	1
279	1	0	0	1	1	1	1
280	0	1	1	1	1	1	1
281	0	1	1	1	1	1	1
282	0	1	1	1	1	1	1
283	0	1	1	0	1	1	1
284	0	1	0	1	1	1	1
285	0	1	0	1	1	1	1
286	0	1	1	1	1	1	1
287	0	1	1	1	1	1	1
288	0	1	1	0	1	1	1
289	0	1	1	0	1	1	1
290	0	1	1	1	1	1	1
291	0	0	1	1	1	1	1
292	0	0	1	1	1	1	1
293	0	1	1	1	1	1	1
294	0	1	1	1	1	1	1
295	0	1	0	0	0	1	0
296	1	1	1	1	1	1	1
297	0	1	1	1	1	1	1
298	0	1	1	1	1	1	1
299	1	1	0	1	1	1	1
300	0	0	1	0	1	1	1
301	0	1	1	1	1	1	1
302	0	1	1	1	1	1	1
303	1	1	1	1	1	1	1
304	0	1	1	0	0	1	0
305	0	1	0	1	1	1	1
306	0	1	0	1	1	1	1

307	1	1	0	1	1	1	1
308	0	1	1	0	1	1	1
309	0	1	1	1	1	1	1
310	0	0	1	0	1	1	1
311	1	1	1	1	1	1	1
312	0	1	1	1	1	1	1
313	0	1	1	1	1	1	1
314	0	1	1	1	1	1	1
315	0	1	1	1	1	1	1
316	0	1	1	1	1	1	1
317	1	1	1	1	1	1	1
318	0	1	1	1	1	1	1
319	0	1	1	1	1	1	1
320	0	1	1	0	1	1	1
321	0	1	1	1	1	1	1
322	0	1	1	1	1	1	1
323	0	1	1	1	1	1	1
324	0	1	1	0	1	1	1
325	0	1	1	1	1	1	1
326	1	1	1	1	1	1	1
327	0	1	1	1	1	1	1
328	0	1	1	0	1	1	1
329	0	1	1	1	1	1	1
330	0	0	1	1	1	1	1
331	0	1	1	1	1	1	1
332	0	1	1	1	1	1	1
333	0	0	1	0	0	1	0
334	0	1	0	1	1	1	1
335	1	1	0	1	1	1	1
336	0	1	1	0	1	1	1
337	0	1	1	1	1	1	1
338	0	0	0	1	0	1	0
339	0	1	1	1	1	1	1
340	1	1	1	1	1	1	1
341	1	1	1	1	1	1	1
342	0	1	0	1	1	1	1
343	0	1	1	1	1	1	1
344	0	1	1	1	1	1	1
345	1	0	1	1	1	1	1
346	0	1	0	1	1	1	1
347	0	1	0	1	1	1	1
348	0	1	1	1	1	1	1
349	0	1	1	1	1	1	1
350	1	1	1	1	1	1	1
351	0	1	0	0	1	1	1
352	1	1	1	1	1	1	1

353	0	1	0	1	1	1	1
354	0	1	1	1	1	1	1
355	1	0	1	1	1	1	1
356	0	1	1	1	1	1	1
357	0	1	0	1	1	1	1
358	0	1	1	1	1	1	1
359	0	1	1	1	1	1	1
360	0	1	1	1	1	1	1
361	0	1	1	0	1	1	1
362	0	1	1	1	1	1	1
363	0	1	0	1	1	1	1
364	1	1	1	1	1	1	1
365	0	1	1	1	1	1	1
366	0	1	0	1	1	1	1
367	0	1	0	0	0	1	0
368	0	0	0	0	1	1	0
369	0	1	1	1	1	1	1
370	0	1	1	1	1	1	1
371	0	1	1	1	1	1	1
372	0	1	1	1	1	1	1
373	0	0	1	0	1	1	1
374	0	1	1	1	1	1	1
375	0	1	1	1	1	1	1
376	0	1	1	1	1	1	1
377	1	1	1	1	1	1	1
378	0	1	1	1	1	1	1
379	0	1	1	1	1	1	1
380	0	1	0	1	1	1	1
381	0	1	0	0	0	1	0
382	0	1	1	0	1	1	1
383	0	1	1	1	1	1	1
384	0	1	1	1	1	1	1
385	1	1	1	1	1	1	1
386	0	1	1	1	1	1	1
387	0	0	0	0	1	1	0
388	0	1	1	1	1	1	1
389	1	1	1	1	1	1	1
390	0	1	0	0	1	1	0
391	0	1	0	0	0	1	0
392	0	1	0	1	1	1	1
393	0	1	1	1	1	1	1
394	0	1	1	1	1	1	1
395	0	0	0	1	1	1	1
396	1	1	1	1	1	1	1
397	0	1	1	1	1	1	1
398	0	1	1	1	1	1	1

399	0	1	0	1	1	1	1
400	0	1	1	1	1	1	1
401	1	1	1	1	1	1	1
402	1	1	1	1	1	1	1
403	0	1	1	1	1	1	1
404	0	1	1	1	1	1	1
405	0	1	0	0	1	1	1
406	0	1	1	1	1	1	1
407	0	1	0	1	1	1	1
408	0	1	1	1	1	1	1
409	0	1	1	1	1	1	1
410	0	1	1	0	1	1	1
411	0	1	0	1	1	1	1
412	0	1	1	1	1	1	1
413	0	1	1	1	1	1	1
414	1	1	1	1	1	1	1
415	0	1	1	0	1	1	1
416	0	1	0	1	1	1	1
417	0	1	1	1	1	1	1
418	0	1	1	1	1	1	1
419	0	1	0	1	1	1	1
420	0	1	1	1	1	1	1
421	0	1	1	1	1	1	1
422	1	1	1	1	1	1	1
423	1	1	1	1	1	1	1
424	0	1	1	0	1	1	1
425	0	1	0	1	1	1	1
426	0	1	1	1	1	1	1
427	0	1	1	0	1	1	1
428	1	1	1	1	1	1	1
429	0	1	1	1	1	1	1
430	0	1	1	1	1	1	1
431	0	1	1	1	1	1	1
432	0	1	1	1	1	1	1
433	0	1	1	1	1	1	1
434	0	1	1	1	1	1	1
435	0	1	0	1	1	1	1
436	0	1	1	1	1	1	1
437	0	1	1	1	1	1	1
438	0	1	0	0	0	0	0
439	0	1	1	1	1	1	1
440	1	1	1	1	1	1	1
441	1	1	1	1	1	1	1
442	0	1	1	1	1	1	1
443	1	1	1	1	1	1	1
444	0	1	1	1	1	1	1

445	1	1	1	1		1	1	1
446	0	1	1	1		1	1	1
447	0	1	1	1		1	1	1
448	0	1	0	1		1	1	1
449	0	1	0	1		1	1	1
450	0	1	1	1		1	1	1
451	0	1	1	1		1	1	1
452	0	1	1	1		1	1	1
453	0	1	0	1		1	1	1
454	0	1	1	1		1	1	1
455	0	1	1	1		1	1	1
456	0	1	1	1		1	1	1
457	0	1	0	1		1	1	1
458	1	1	1	1		1	1	1
459	0	1	1	0		1	1	1
460	0	1	1	1		1	1	1
461	0	1	1	0		1	1	1
462	0	0	1	1		1	1	1
463	0	1	1	1		1	1	1
464	0	1	1	1		1	1	1
465	0	1	0	1		1	1	1
466	0	0	1	1		1	1	1
467	1	1	1	1		1	1	1
468	1	1	1	1		1	1	1
469	0	0	0	0		0	1	0
470	1	1	1	1		1	1	1
471	0	1	1	1		1	1	1
472	0	1	1	0		1	1	1
473	1	1	1	1		1	1	1
474	1	1	1	1		1	1	1
475	1	1	1	1		1	1	1
476	1	0	1	1	1	1	1	1
477	1	1	1	1	1	1	1	1
478	0	1	1	1	1	1	1	1
479	1	1	1	1	1	1	0	1
480	1	1	1	1	1	1	1	1
481	1	1	1	1	1	1	0	1
482	1	1	1	1	1	1	1	1
483	1	1	1	0	1	1	1	1
484	1	1	1	0	1	1	1	1
485	1	1	1	0	0	1	1	1
486	1	1	1	0	1	1	1	1
487	1	1	1	0	1	1	1	1
488	1	1	1	0	1	1	1	1
489	1	1	1	0	1	1	1	1
490	1	1	1	0	1	1	1	1

491	1	0	1	1	1	1	1	1
492	1	1	1	1	1	1	1	1
493	0	0	0	0	1	1	1	1
494	1	0	1	1	1	1	1	1
495	1	1	1	1	1	1	1	1
496	1	1	1	1	1	1	1	1
497	1	1	1	1	1	1	1	1
498	1	0	1	1	1	1	1	1
499	1	1	1	1	1	1	1	1
500	1	1	1	1	1	1	1	1
501	1	1	1	1	1	1	1	1
502	1	1	1	1	1	1	1	1
503	1	0	1	1	1	1	1	1
504	1	1	1	1	1	1	1	1
505	1	0	1	1	1	1	1	1
506	1	1	1	1	1	1	1	1
507	1	0	1	1	0	1	1	1
508	1	1	1	1	1	1	1	1
509	1	1	1	1	1	1	1	1
510	1	0	1	1	1	1	1	1
511	1	0	1	1	1	1	1	1
512	1	0	1	1	1	1	1	1
513	1	0	1	1	1	1	1	1
514	1	1	1	1	1	1	1	1
515	1	1	1	1	1	1	1	1
516	1	1	1	1	1	1	1	1
517	1	1	1	1	1	1	1	1
518	0	0	1	1	1	0	0	0
519	1	1	1	1	1	1	1	1
520	1	1	1	1	1	1	1	1
521	1	1	1	1	1	1	1	1
522	1	1	1	1	0	1	1	1
523	1	0	1	1	1	1	1	1
524	1	1	1	1	1	0	1	1
525	1	1	1	1	1	1	1	1
526	1	1	1	1	1	1	1	1
527	1	0	1	1	1	1	1	1
528	1	1	1	1	1	1	1	1
529	1	1	1	1	1	1	1	1
530	1	0	1	1	1	0	0	1
531	1	1	1	1	1	0	1	1
532	1	1	1	1	1	1	1	1
533	1	1	1	1	1	1	1	1
534	1	1	1	1	1	1	1	1
535	1	1	1	1	1	1	1	1
536	1	1	1	1	1	1	1	1

537	1	1	1	1	1	1	1	1
538	1	1	1	1	1	1	1	1
539	1	1	1	1	1	1	1	1
540	1	1	1	0	1	1	1	1
541	1	1	1	0	1	1	1	1
542	1	1	1	1	1	1	1	1
543	1	1	1	1	0	1	1	1
544	1	1	1	1	1	1	1	1
545	1	1	1	1	1	1	1	1
546	1	1	1	1	1	1	1	1
547	1	1	1	1	1	1	1	1
548	1	1	1	1	1	1	1	1
549	1	1	1	1	1	1	1	1
550	1	1	1	1	1	1	1	1
551	1	1	1	1	1	1	1	1
552	1	1	1	1	1	1	1	1
553	1	1	1	1	1	1	1	1
554	1	1	1	0	1	1	1	1
555	1	1	1	0	1	1	1	1
556	1	1	1	1	1	1	1	1
557	1	1	1	1	1	1	1	1
558	1	1	1	1	1	1	1	1
559	1	1	1	1	0	1	1	1
560	1	1	1	1	1	1	1	1
561	1	1	1	1	1	1	1	1
562	1	1	1	1	1	1	1	1
563	1	1	1	1	1	1	1	1
564	1	1	1	1	1	1	1	1
565	1	1	1	1	1	1	1	1
566	1	0	1	1	1	1	1	1
567	1	0	1	1	1	1	1	1
568	1	1	1	1	1	1	1	1
569	1	0	1	1	1	1	1	1
570	1	1	1	1	1	1	1	1
571	1	1	1	1	1	1	1	1
572	1	1	1	1	0	0	1	1
573	1	1	1	1	1	0	1	1
574	1	1	1	1	1	0	1	1
575	1	1	1	1	1	1	1	1
576	1	1	1	1		1	1	1
577	1	1	1	1		1	1	1
578	1	1	1	1		1	1	1
579	1	1	1	1		1	1	1
580	1	1	1	1		1	1	1
581	1	1	1	1		0	1	1
582	1	1	1	1		0	0	1

583	1	1	1	1	1	1	1
584	1	1	1	1	1	1	1
585	1	1	1	1	1	1	1
586	1	1	1	1	1	1	1
587	1	1	1	1	1	1	1
588	1	1	1	1	1	1	1
589	1	1	1	1	1	1	1
590	1	1	1	1	1	1	1
591	1	1	1	1	1	0	1
592	1	1	1	1	1	1	1
593	1	1	1	1	1	1	1
594	1	1	1	1	1	1	1
595	1	0	1	1	0	1	1
596	1	1	1	1	1	1	1
597	1	1	1	1	1	1	1
598	1	0	1	1	0	1	1
599	1	1	1	1	1	1	1
600	1	1	1	1	1	1	1
601	1	1	1	1	1	1	1
602	1	1	1	1	1	1	1
603	1	1	1	1	1	1	1
604	1	1	1	1	1	1	1
605	1	1	1	1	0	1	1
606	1	1	1	1	0	1	1
607	1	1	1	1	0	1	1
608	1	1	1	1	0	1	1
609	1	0	1	1	1	1	1
610	1	1	1	1	1	1	1
611	1	1	1	1	1	1	1
612	1	1	1	1	1	1	1
613	1	1	1	1	0	1	1
614	1	1	1	1	1	1	1
615	1	1	1	1	1	1	1
616	1	1	1	1	1	1	1
617	1	1	1	1	1	1	1
618	1	1	1	1	1	1	1
619	1	0	1	1	0	1	1
620	1	1	1	1	0	0	1
621	1	0	1	0	1	1	1
622	1	1	1	1	1	1	1
623	1	0	1	1	1	1	1
624	1	1	1	1	1	1	1
625	1	1	1	1	1	1	1
626	1	1	1	1	1	1	1
627	1	1	1	1	1	1	1
628	1	1	1	1	1	1	1

629	1	1	1	1	1	1	1
630	1	1	1	1	1	1	1
631	1	1	1	1	1	1	1
632	1	1	1	1	1	1	1
633	1	0	1	1	1	1	1
634	1	0	1	1	1	1	1
635	1	1	1	1	1	1	1
636	1	1	1	1	1	1	1
637	1	1	1	1	1	1	1
638	1	1	1	1	1	1	1
639	1	1	1	1	1	1	1
640	1	1	1	1	1	1	1
641	1	0	1	1	1	1	1
642	1	1	1	1	1	1	1
643	1	0	1	1	1	1	1
644	1	1	1	1	1	1	1
645	1	1	1	1	1	1	1
646	1	1	1	0	1	1	1
647	1	1	1	1	1	1	1
648	1	1	1	0	1	1	1
649	1	1	1	1	1	1	1
650	1	1	1	0	1	1	1
651	1	1	1	1	1	1	1
652	1	1	1	0	1	1	1
653	1	1	1	1	1	1	1
654	1	1	1	0	1	1	1
655	1	1	1	1	1	1	1
656	1	1	1	0	1	1	1
657	1	1	1	1	1	1	1
658	1	1	1	0	1	1	1
659	1	1	1	1	1	1	1
660	1	1	1	0	1	1	1
661	1	1	1	1	1	1	1
662	1	1	1	0	1	1	1
663	1	0	1	1	1	1	1
664	1	1	1	1	1	1	1
665	1	1	1	0	1	1	1
666	1	1	1	1	1	1	1
667	1	1	1	0	1	1	1
668	1	1	1	1	1	1	1
669	1	1	1	0	1	1	1
670	1	1	1	1	1	1	1
671	1	1	1	0	1	1	1
672	1	1	1	1	1	1	1
673	1	1	1	0	1	1	1
674	1	0	1	1	1	1	1

675	1	1	1	1	1	1	1
676	1	1	1	1	1	1	1
677	1	0	1	0	1	1	1
678	1	1	1	1	1	1	1
679	1	1	1	1	1	1	1
680	1	1	1	1	1	1	1
681	1	0	1	1	1	1	1
682	1	1	1	1	1	1	1
683	1	1	1	1	1	1	1
684	1	1	1	1	1	1	1
685	1	1	1	1	1	1	1
686	1	1	1	1	1	1	1
687	1	1	1	1	1	1	1
688	1	1	1	1	1	1	1
689	1	1	1	1	1	1	1
690	1	1	1	1	1	1	1
691	1	1	1	1	1	1	1
692	1	1	1	1	1	1	1
693	1	1	1	1	1	1	1
694	1	1	1	1	1	1	1
695	1	1	1	1	1	1	1
696	1	1	1	1	1	1	1
697	1	1	1	1	1	1	1
698	1	1	1	1	1	1	1
699	1	1	1	1	1	1	1
700	1	1	1	1	1	1	1
701	1	1	1	1	1	1	1
702	1	1	1	1	1	1	1
703	1	1	1	1	1	1	1
704	1	1	1	1	1	1	1
705	1	1	1	1	1	1	1
706	1	1	1	1	1	1	1
707	1	1	1	1	1	1	1
708	1	1	1	1	1	1	1
709	1	1	1	1	1	1	1
710	1	1	1	1	1	1	1
711	1	0	1	1	1	1	1
712	1	1	1	1	1	1	1
713	1	0	1	1	1	1	1
714	1	1	1	1	1	1	1
715	1	0	1	1	1	1	1
716	1	1	1	1	1	1	1
717	1	0	1	1	1	1	1
718	1	1	1	1	1	1	1
719	1	0	1	1	1	1	1
720	1	1	1	1	1	1	1

721	1	0	1	1	1	1	1
722	1	1	1	1	1	1	1
723	1	0	1	1	1	1	1
724	1	1	1	1	1	1	1
725	1	0	1	1	1	1	1
726	1	1	1	1	1	1	1
727	1	0	1	1	1	1	1
728	1	1	1	1	1	1	1
729	1	0	1	1	1	1	1
730	1	1	1	1	1	1	1
731	1	0	1	1	1	1	1
732	1	1	1	1	1	1	1
733	1	0	1	1	1	1	1
734	1	1	1	1	1	1	1
735	1	1	1	1	1	1	1
736	1	1	1	1	1	1	1
737	1	1	1	1	1	1	1
738	1	0	1	1	1	1	1
739	1	1	1	1	1	1	1
740	1	0	1	1	1	1	1
741	1	1	1	1	1	1	1
742	0	0	0	1	1	1	1
743	1	0	1	1	1	1	1
744	1	1	1	1	1	1	1
745	1	1	1	1	1	1	1
746	1	0	1	1	1	1	1
747	1	1	1	1	1	1	1
748	1	0	1	1	1	1	1
749	1	0	1	1	1	1	1
750	1	1	1	1	1	1	1
751	1	0	1	0	1	1	1
752	1	1	1	1	1	1	1
753	1	0	1	0	1	1	1
754	1	1	1	1	1	1	1
755	1	0	1	0	1	1	1
756	1	1	1	1	1	1	1
757	1	0	1	1	1	1	1
758	1	1	1	1	1	1	1
759	1	0	1	1	1	1	1
760	1	1	1	1	1	1	1
761	1	0	1	1	1	1	1
762	1	1	1	1	1	1	1
763	1	0	1	1	1	1	1
764	1	1	1	1	1	1	1
765	1	1	1	0	1	1	1
766	1	0	1	0	1	1	1

767	1	1	1	0	1	1	1
768	1	0	1	0	1	1	1
769	1	1	1	0	1	1	1
770	1	0	1	0	1	1	1
771	1	0	1	0	1	1	1
772	1	1	1	0	1	1	1
773	1	0	1	0	1	1	1
774	1	1	1	0	1	1	1
775	1	1	1	0	1	1	1
776	1	0	1	0	1	1	1
777	1	0	1	1	1	1	1
778	1	1	1	1	1	1	1
779	1	0	1	1	1	1	1
780	1	1	1	1	1	1	1
781	1	0	1	1	1	1	1
782	1	1	1	1	1	1	1
783	1	0	1	1	1	1	1
784	1	1	1	1	1	1	1
785	1	0	1	1	1	1	1
786	1	1	1	1	1	1	1
787	1	0	1	1	1	1	1
788	1	1	1	1	1	1	1
789	1	0	1	1	1	1	1
790	1	1	1	1	1	1	1
791	1	0	1	1	1	1	1
792	1	1	1	1	1	1	1
793	1	0	1	1	1	1	1
794	1	1	1	1	1	1	1
795	1	0	1	1	1	1	1
796	1	1	1	1	1	1	1
797	1	0	1	1	1	1	1
798	1	1	1	1	1	1	1
799	1	1	1	1	1	1	1
800	1	1	1	1	1	1	1
801	1	1	1	1	1	1	1
802	1	1	1	1	1	1	1
803	1	1	1	1	1	1	1
804	1	1	1	1	1	1	1
805	1	1	1	1	1	1	1
806	1	1	1	1	1	1	1
807	1	0	1	1	1	1	1
808	1	0	1	1	1	1	1
809	1	1	1	1	1	1	1
810	1	1	1	1	1	1	1
811	1	0	1	1	1	1	1
812	1	0	1	1	1	1	1

813	1	1	1	1	1	1	1
814	1	1	1	1	1	1	1
815	1	1	1	1	1	1	1
816	1	1	1	1	1	1	1
817	1	1	1	1	1	1	1
818	1	1	1	1	1	1	1
819	1	1	1	1	1	1	1
820	1	1	1	1	1	1	1
821	1	0	1	1	1	1	1
822	1	1	1	1	1	1	1
823	1	0	1	1	1	1	1
824	1	1	1	1	1	1	1
825	1	0	1	1	1	1	1
826	1	1	1	1	1	1	1
827	1	0	1	1	1	1	1
828	1	1	1	1	1	1	1
829	1	0	1	0	1	1	1
830	1	0	0	0	0	0	0
831	1	0	0	0	0	0	0
832	1	0	0	0	0	0	0
833	1	0	0	0	0	0	0
834	1	0	1	0	1	1	1
835	1	1	1	1	1	1	1
836	1	0	1	0	1	1	1
837	1	1	1	1	1	1	1
838	1	0	1	0	1	1	1
839	1	0	1	0	1	1	1
840	1	0	1	0	1	1	1
841	1	1	1	1	1	1	1
842	1	0	1	0	1	1	1
843	1	1	1	1	1	1	1
844	1	0	1	1	1	1	1
845	1	1	1	1	1	1	1
846	1	0	1	1	1	1	1
847	1	1	1	1	1	1	1
848	1	0	1	1	1	1	1
849	1	1	1	1	1	1	1
850	1	0	1	1	1	1	1
851	1	1	1	1	1	1	1
852	1	0	1	1	1	1	1
853	1	1	1	1	1	1	1
854	1	0	1	1	1	1	1
855	1	1	1	1	1	1	1
856	1	0	1	1	1	1	1
857	1	1	1	1	1	1	1
858	1	0	1	0	1	1	1

859	1	0	1	1	1	1	1
860	1	1	1	1	1	1	1
861	1	0	1	1	1	1	1
862	1	1	1	1	1	1	1
863	1	0	1	1	1	1	1
864	1	1	1	1	1	1	1
865	1	0	1	1	1	1	1
866	1	1	1	1	1	1	1
867	1	1	1	1	1	1	1
868	1	1	1	0	1	1	1
869	1	1	1	0	1	1	1
870	1	1	1	0	1	1	1
871	1	1	1	0	1	1	1
872	1	1	1	0	1	1	1
873	1	1	1	0	1	1	1
874	1	1	1	0	1	1	1
875	1	1	1	0	1	1	1
876	1	0	1	0	1	1	1
877	1	1	1	1	1	1	1
878	1	1	1	1	1	1	1
879	1	1	1	1	1	1	1
880	1	0	1	1	1	1	1
881	1	1	1	0	1	1	1
882	1	1	1	0	1	1	1
883	1	1	1	0	1	1	1
884	1	1	1	0	1	1	1
885	1	1	1	0	1	1	1
886	1	1	1	0	1	1	1
887	1	1	1	0	1	1	1
888	1	1	1	1	1	1	1
889	1	1	1	1	1	1	1
890	1	1	1	1	0	1	1
891	1	1	1	1	1	1	1
892	1	0	1	1	1	1	1
893	1	1	1	1	1	1	1
894	1	0	1	1	1	1	1
895	1	0	1	1	1	1	1
896	1	0	1	1	1	1	1
897	1	0	1	1	1	1	1
898	1	0	1	0	0	1	0
899	1	1	1	1	1	1	1
900	1	1	1	0	1	1	1
901	1	1	1	1	1	1	1
902	1	1	1	1	1	1	1
903	1	1	1	1	0	1	1
904	1	0	1	0	0	1	0

905	1	0	1	1	1	1	1
906	1	1	1	1	1	0	1
907	1	1	1	1	1	1	1
908	1	1	1	1	1	1	1
909	1	0	1	1	1	1	1
910	1	0	1	1	1	1	1
911	1	1	1	1	1	1	1
912	1	0	0	1	0	0	0
913	1	1	1	1	1	1	1
914	1	1	1	1	1	1	1
915	1	1	1	1	0	1	1
916	1	1	1	1	0	1	1
917	1	1	1	1	1	1	1
918	1	0	1	1	1	1	1
919	1	1	1	1	1	1	1
920	1	1	1	1	1	1	1
921	1	0	1	0	1	1	1
922	1	1	1	1	1	1	1
923	1	1	1	1	1	1	1
924	1	0	1	1	1	1	1
925	1	1	1	1	1	1	1
926	1	0	1	1	1	1	1
927	1	1	1	1	1	1	1
928	1	1	1	1	1	1	1
929	1	1	1	1	1	1	1
930	1	0	1	1	1	1	1
931	1	0	1	1	1	1	1
932	1	1	1	1	1	1	1
933	1	0	1	1	1	1	1
934	1	1	1	0	1	1	1
935	1	1	1	1	1	1	1
936	1	1	1	0	1	1	1
937	1	1	1	1	1	1	1
938	1	1	1	1	1	1	1
939	1	0	1	1	1	1	1
940	1	0	1	1	1	1	1
941	1	1	1	1	1	1	1

### Apêndice 3 – Planilha de Validação.

Tabela 7 - Validação da PA<sup>3</sup>P mediante grupo de proteínas com alergenicidade desconhecidas, onde 0 = presença e 1 = ausência.

ID	PA <sup>3</sup> P	ID	PA <sup>3</sup> P		
1	102m_A	1	38	1t26_A	1
2	1a5n_C	1	39	1ukl_B	1
3	1ai0_C	1	40	1vf8_A	1
4	1ahz_A	1	41	1vfc_A	1
5	1aqm_A	0	42	1vqo_A	1
6	1aqu_A	1	43	1vqo_E	1
7	1aqp_A	1	44	1vz3_A	1
8	1aux_B	1	45	1vz6_A	1
9	1av1_B	1	46	1w6f_C	1
10	1b3z_A	1	47	1w6g_A	1
11	1b41_B	1	48	1wrq_B	1
12	1b44_D	1	49	1wrt_S	1
13	1bjp_E	1	50	1xmh_C	1
14	1bjq_A	1	51	1xmi_C	1
15	1bv1_D	1	52	1y77_K	1
16	1bvn_T	1	53	1yj9_M	1
17	1bvp_3	1	54	1z63_A	1
18	1bvs_H	1	55	1zem_A	1
19	1bvw_A	1	56	1zsb_A	1
20	1bwj_A	0	57	1zsf_B	1
21	1bwk_A	1	58	1zsh_A	1
22	1bwy_A	0	59	1zsn_A	1
23	1bwz_A	1	60	1zsp_B	0
24	1bx1_A	1	61	1zsq_A	1
25	1bxq_A	0	62	1zsv_A	1
26	1e2m_A	1	63	1zsz_B	1
27	1fj1_F	1	64	1zt2_A	1
28	1gk8_I	1	65	1zt3_A	1
29	1jra_B	1	66	1zvt_B	1
30	1jrb_B	1	67	20gs_B	1
31	1kyi_M	1	68	21bi_A	1
32	1nnj_A	1	69	2a03_B	0
33	1nno_A	1	70	2a0q_D	0
34	1ox6_B	1	71	2ald_H	1
35	1ox9_C	1	72	2d5f_B	0
36	1px4_C	1	73	2hu0_H	1
37	1rzm_B	1	74	2mvz_A	0

	<b>ID</b>	<b>PA<sup>3</sup>P</b>
75	2mwp_A	1
76	2mxi_A	1
77	2qci_B	1
78	2vv5_G	1
79	2xrx_B	1
80	3aam_A	1
81	3ab1_B	1
82	3ab2_M	1
83	3gj1_D	1
84	3gje_L	1
85	3l2e_D	1
86	3nub_A	1
87	3rla_B	1
88	4bts_DE	1
89	4bts_D7	1
90	4bud_A	1
91	4dmr_A	1
92	4fca_A	1
93	4fff_B	1
94	4mnf_B	1
95	4msf_A	1
96	4q9h_A	1
97	4qa1_A	1
98	4u56_S7	1
99	4u5b_A	1
100	1b43_A	1

## Apêndice 4 – Códigos-fonte de construção da PA<sup>3</sup>P

### 4.1 Códigos-fonte referentes ao AlgPred

```

package org.unipampa.lpa.classe;
import java.io.IOException;
import java.net.URI;
import org.apache.http.HttpEntity;
import org.apache.http.HttpResponse;
import org.apache.http.client.ClientProtocolException;
import org.apache.http.client.ResponseHandler;
import org.apache.http.client.config.RequestConfig;
import org.apache.http.client.methods.HttpUriRequest;
import org.apache.http.client.methods.RequestBuilder;
import org.apache.http.impl.client.BasicCookieStore;
import org.apache.http.impl.client.CloseableHttpClient;
import org.apache.http.impl.client.HttpClients;
import org.apache.http.util.EntityUtils;
public class Algpred {
    private String sequence;
    private String htmlresult;
    private double[] peso;
    private String resultado;
    public Algpred(){
        /*
        0 - Algpred - IgE - Similarity
        3 - Algpred - SVM Dipeptide - Predicted
        4 - Algpred - ARP - Similarity
        */
        peso = new double[6];
        peso[0] = 60.04;
        peso[3] = 82.47;
        peso[4] = 88.95;
        peso[5] = 0;
    }
    public String getSequence() {
        return sequence;
    }
    public void setSequence(String sequence) {
        this.sequence = sequence;
    }
    public String getHtmlResult(){
        return htmlresult;
    }
    public int execute(int approach){
        int timeout = 60;
        RequestConfig config = RequestConfig.custom()
            .setConnectTimeout(timeout * 1000)
            .setConnectionRequestTimeout(timeout * 1000)
            .setSocketTimeout(timeout * 1000).build();

        BasicCookieStore cookieStore = new BasicCookieStore();
        CloseableHttpClient httpclient = HttpClients.custom()
            .setDefaultCookieStore(cookieStore)
            .setDefaultRequestConfig(config)
            .build();

        try {

```

```

HttpUriRequest consulta = null;

consulta = RequestBuilder.post()
    .setUri(new
URI("http://www.imtech.res.in/cgibin/algpred/algpred_main_temp.pl"))
    .addParameter("SEQ", sequence)
    .addParameter("SEQNAME", "")
    .addParameter("format", "nformat")
    .addParameter("approach", Integer.toString(approach))
    .build();

// Create a custom response handler
ResponseHandler<String> responseHandler = new ResponseHandler<String>() {

    @Override
    public String handleResponse(
        final HttpResponse response) throws
ClientProtocolException, IOException {
        int status = response.getStatusLine().getStatusCode();
        if (status >= 200 && status < 300) {
            HttpEntity entity = response.getEntity();
            return entity != null ? EntityUtils.toString(entity) : null;
        } else {
            throw new ClientProtocolException("Unexpected response
status: " + status);
        }
    }
};

String strret = httpclient.execute(consulta, responseHandler);
httpclient.close();
//Retira somente a parte a exibir
int pos = 0, posini = 0, posfim = 0;
if (approach > 1)
    pos = strret.toLowerCase().indexOf("<h3>", 0);
if (pos > 0){
    posini = pos;
    if (approach > 1)
        posfim = strret.toLowerCase().indexOf("</h3>", pos);
    if (posfim > 1)
        strret = strret.substring(pos, posfim + 5);
    strret = strret.replaceAll("<font color=[\\"]yellow[\\"]>", "");
}
if (approach == 1){
    strret = strret.replaceAll("<h3>", "<h2>");
    strret = strret.concat("</tr></td></center></TBODY></TABLE>");
}
if (strret.toLowerCase().indexOf("non allergen", 0) > 0)
    resultado = "Non-Allergen";
else if (strret.toLowerCase().indexOf("allergen", 0) > 0)
    resultado = "Allergen";
else
    resultado = "-";
htmlresult = strret;
return 1;
} catch (Exception e) {
    System.out.println(e.getMessage());
    return -1;
}
}

```

```
public double getPeso(int approach) {  
    return peso[approach - 1];  
}  
public String getResultado() {  
    return resultado;  
}  
}
```

#### 4.2 Códigos-fonte referentes ao Allergen Online

```

package org.unipampa.lpa.classe;
import java.io.File;
import java.io.IOException;
import java.net.URI;
import org.apache.commons.io.FileUtils;
import org.apache.http.Consts;
import org.apache.http.HttpEntity;
import org.apache.http.HttpResponse;
import org.apache.http.client.ClientProtocolException;
import org.apache.http.client.ResponseHandler;
import org.apache.http.client.config.RequestConfig;
import org.apache.http.client.methods.HttpUriRequest;
import org.apache.http.client.methods.RequestBuilder;
import org.apache.http.impl.client.BasicCookieStore;
import org.apache.http.impl.client.CloseableHttpClient;
import org.apache.http.impl.client.HttpClients;
import org.apache.http.util.EntityUtils;

public class Allergenonline {
    private String sequence;
    private String htmlresult;
    private double[] peso;
    private String resultado;
    private String sessionid;

    public Allergenonline(){
        peso = new double[2];
        peso[0] = 95.54;
        peso[1] = 98.3;
    }

    public String getSequence() {
        return sequence;
    }

    public void setSequence(String sequence) {
        this.sequence = sequence.toUpperCase().trim();
    }

    public void setSessionid(String sessionid) {

        this.sessionid = sessionid;
    }

    public String getHtmlResult(){
        return htmlresult;
    }

    public int execute(int approach){

        int timeout = 90;

        RequestConfig config = RequestConfig.custom()
            .setConnectTimeout(timeout * 1000)
            .setConnectionRequestTimeout(timeout * 1000)
            .setSocketTimeout(timeout * 1000).build();

        BasicCookieStore cookieStore = new BasicCookieStore();
        CloseableHttpClient httpclient = HttpClients.custom()

```

```

        .setDefaultCookieStore(cookieStore)
        .setDefaultRequestConfig(config)
        .build();

    try {

        HttpRequest consulta = null;

        if (approach == 1){
            consulta = RequestBuilder.post()
                .setUri(new URI("http://allergenonline.com/search80-
15.cgi"))
                .addParameter("SEARCH_SEQ", sequence)
                .addParameter("searchtype", "sliding")
                .addParameter("submit", "Submit")
                .setHeader("Content-Type", "application/x-www-form-
urlencoded")
                .build();
        }
        else{
            consulta = RequestBuilder.post()
                .setUri(new
URI("http://allergenonline.com/shortmer15.php"))
                .addParameter("SEARCH_SEQ", sequence)
                .addParameter("searchtype", "shortmer")
                .addParameter("submit", "Submit")
                .build();
        }

        // Create a custom response handler
        ResponseHandler<String> responseHandler = new ResponseHandler<String>() {

            @Override
            public String handleResponse(
                final HttpResponse response) throws
ClientProtocolException, IOException {

                int status = response.getStatusLine().getStatusCode();
                if (status >= 200 && status < 300) {
                    HttpEntity entity = response.getEntity();
                    return entity != null ? EntityUtils.toString(entity) : null;
                } else {
                    throw new ClientProtocolException("Unexpected response
status: " + status);
                }
            }
        };

        String strret = httpClient.execute(consulta, responseHandler);

        strret = strret.replaceAll("<a href='#1'>&gt;query</a><br><hr><a name='1'>", "");
        if (approach == 1){
            if (strret.toLowerCase().indexOf("no matches of greater", 0) > 0)
                resultado = "Non-Allergen";
            else
                resultado = "Allergen";
        }
        else{
            if (strret.toLowerCase().indexOf("no sequences found with an exact", 0) > 0)
                resultado = "Non-Allergen";
            else

```

```
                resultado = "Alergen";
            }
            htmlresult = strret;
            httpclient.close();
            return 1;
        } catch (Exception e) {
            System.out.println(e.getMessage());
            return -1;
        }
    }

    public double getPeso(int approach) {
        return peso[approach - 1];
    }

    public String getResultado() {
        return resultado;
    }
}
```

### 4.3 Códigos-fonte referentes ao proAP

```

package org.unipampa.lpa.classe;
import java.io.IOException;
import java.net.URI;
import org.apache.http.HttpEntity;
import org.apache.http.HttpResponse;
import org.apache.http.client.ClientProtocolException;
import org.apache.http.client.ResponseHandler;
import org.apache.http.client.config.RequestConfig;
import org.apache.http.client.methods.HttpUriRequest;
import org.apache.http.client.methods.RequestBuilder;
import org.apache.http.impl.client.BasicCookieStore;
import org.apache.http.impl.client.CloseableHttpClient;
import org.apache.http.impl.client.HttpClients;
import org.apache.http.util.EntityUtils;
public class Proap {
    private String sequence;
    private String htmlresult;
    private double peso[];
    private String resultado;
    public Proap(){
        peso = new double[1];
        peso[0] = 85.44;
    }
    public String getSequence() {
        return sequence;
    }
    public void setSequence(String sequence) {
        this.sequence = sequence;
    }
    public String getHtmlResult(){
        return htmlresult;
    }
    public int execute(int approach){
        int timeout = 60;
        RequestConfig config = RequestConfig.custom()
            .setConnectTimeout(timeout * 1000)
            .setConnectionRequestTimeout(timeout * 1000)
            .setSocketTimeout(timeout * 1000).build();
        BasicCookieStore cookieStore = new BasicCookieStore();
        CloseableHttpClient httpclient = HttpClients.custom()
            .setDefaultCookieStore(cookieStore)
            .setDefaultRequestConfig(config)
            .build();
        try {
            HttpUriRequest consulta = null;
            consulta = RequestBuilder.post()
                .setUri(new URI("http://gmobl.sjtu.edu.cn/proAP/blast.php"))
                .addParameter("sequence", sequence)
                .addParameter("Sdegree", "35")
                .addParameter("Mlength", "6")
                .addParameter("rules[]", "svm-aac")
                .addParameter("category", "AeroAnimal.fasta")
                .addParameter("extent", "All.fasta")
                .addParameter("B1", "Predict")
                .build();
            // Create a custom response handler
            ResponseHandler<String> responseHandler = new ResponseHandler<String>() {
                @Override
                public String handleResponse(

```

```

        final HttpResponse response) throws
ClientProtocolException, IOException {
    int status = response.getStatusLine().getStatusCode();
    if (status >= 200 && status < 300) {
        HttpEntity entity = response.getEntity();
        return entity != null ? EntityUtils.toString(entity) : null;
    } else {
        throw new ClientProtocolException("Unexpected response
status: " + status);
    }
}

String strret = httpClient.execute(consulta, responseHandler);

httpClient.close();

//Retira somente a parte a exibir
int pos = 0, posini = 0, posfim = 0;

pos = strret.indexOf("<div class=\"content\">", 0);
pos = strret.indexOf("<div class=\"content\">", pos + 1);
pos = strret.indexOf("<div class=\"content\">", pos + 1);

if (pos > 0){
    posini = pos;
    posfim = strret.indexOf("</div>", pos);

    if (posfim > 0)
        strret = strret.substring(pos, posfim + 8);
}

if (strret.indexOf("non-allergen", 0) > 0)
    resultado = "Non-Allergen";
else if (strret.indexOf("allergen", 0) > 0)
    resultado = "Allergen";
else
    resultado = "-";

htmlresult = strret;

return 1;

} catch (Exception e) {
    return -1;
}
}
public double getPeso(int approach) {
    return peso[approach];
}
public String getResultado() {
    return resultado;
}
}

```

#### 4.4 Códigos-fonte referentes à ADFS

```
package org.unipampa.lpa.classe;

import java.io.IOException;
import java.net.URI;

import org.apache.http.HttpEntity;
import org.apache.http.HttpResponse;
import org.apache.http.client.ClientProtocolException;
import org.apache.http.client.ResponseHandler;
import org.apache.http.client.config.RequestConfig;
import org.apache.http.client.methods.HttpUriRequest;
import org.apache.http.client.methods.RequestBuilder;
import org.apache.http.impl.client.BasicCookieStore;
import org.apache.http.impl.client.CloseableHttpClient;
import org.apache.http.impl.client.HttpClients;
import org.apache.http.util.EntityUtils;

public class ADFS {
    private String sequence;
    private String htmlresult;
    private double peso[];
    private String resultado;

    public ADFS(){
        peso = new double[1];
        peso[0] = 96;
    }

    public String getSequence() {
        return sequence;
    }

    public void setSequence(String sequence) {

        this.sequence = sequence;
    }

    public String getHtmlResult(){
        return htmlresult;
    }

    public int execute(int approach){

        int timeout = 30;

        RequestConfig config = RequestConfig.custom()
            .setConnectTimeout(timeout * 1000)
            .setConnectionRequestTimeout(timeout * 1000)
            .setSocketTimeout(timeout * 1000).build();

        BasicCookieStore cookieStore = new BasicCookieStore();
        CloseableHttpClient httpclient = HttpClients.custom()
            .setDefaultCookieStore(cookieStore)
            .setDefaultRequestConfig(config)
            .build();

        try {
```

```

    HttpRequest login = RequestBuilder.post()
        .setUri(new URI("http://allergen.nihs.go.jp/ADFS/login.do"))
        .addParameter("uname", "bdotto82@gmail.com")
        .addParameter("passwd", "1234")
        .build();

    httpClient.execute(login);

    HttpRequest consulta = RequestBuilder.post()
        .setUri(new URI("http://allergen.nihs.go.jp/ADFS/motif.do"))
        .addParameter("seq", sequence)
        .addParameter("expect", "0.0010")
        .build();

    // Create a custom response handler
    ResponseHandler<String> responseHandler = new ResponseHandler<String>() {

        @Override
        public String handleResponse(
            final HttpResponse response) throws ClientProtocolException, IOException {
            int status = response.getStatusLine().getStatusCode();
            if (status >= 200 && status < 300) {
                HttpEntity entity = response.getEntity();
                return entity != null ? EntityUtils.toString(entity) : null;
            } else {
                throw new ClientProtocolException("Unexpected response status: " + status);
            }
        }
    };

    String strret = httpClient.execute(consulta, responseHandler);

    //Retira somente a parte a exibir
    int pos = 0, posini = 0, posfim = 0;

    pos = strret.indexOf("<body>", 0);
    pos = strret.indexOf("<body>", pos + 1);

    if (pos > 0){

        posini = pos;
        posfim = strret.indexOf("</body>", pos);

        if (posfim > 0)
            strret = strret.substring(pos, posfim + 8);

    }

    //Procura so no campo "Total"
    String strres = strret;

    pos = strres.indexOf("<b>Total</b>", 0);

    if (pos > 0){

        posini = pos;
        posfim = strres.indexOf("</tr>", pos);

        if (posfim > 0)
            strres = strres.substring(pos, posfim + 8);
    }

```

```
        }  
        if (strres.indexOf("Positive", 0) > 0)  
            resultado = "Alergen";  
        else if (strres.indexOf("Negative", 0) > 0)  
            resultado = "Non-Alergen";  
        else  
            resultado = "-";  
        htmlresult = strret;  
        httpclient.close();  
        return 1;  
    } catch (Exception e) {  
        System.out.println(e.getMessage());  
        return -1;  
    }  
}  
  
public double getPeso(int approach) {  
    return peso[approach];  
}  
  
public String getResultado() {  
    return resultado;  
}  
}
```



```

<table class="table-imagem">
  <tr>
    <td>
      
    </td>
  </tr>
  <tr><td>
    <div style="padding: 10px;">

```

**Figure 1 – Platform for Analysis of Allergenicity and Antigenicity of Proteins (PA<sup>3</sup>P) design and analyses pipeline.** **A.** PA<sup>3</sup>P dataset. The sample calculus for allergen finite population consists of: confidence level at 99 % (Z); expected hit (P= 60 %) additional to the expected error (Q= 40 %); total sequence population (N= 1,897), and an accuracy level at 5 %, resulting n= 475 allergen sequences set by random drawing. The non-allergen sequences dataset is composed by a fixed n= 466 sequences, since those were the only sequences that returned after a National Center for Biotechnology Information (NCBI) search using "immunogenic" and "Homo sapiens" as keywords. **B.** Both datasets were entirely subjected to the web tools testing in order to estimate the respective web tool accuracy, sorting by false positive and false negative results **(C)**. After defining which tests best fit our model, they were compiled using Java programming language leading to PA<sup>3</sup>P **(D)**. When submitting the query sequence in D, it is processed by the different methods **(C)** in the platform. In **E** it is presented the score table based on accuracy. Each method returns a given result (allergen/non-allergen), and the scores are distributed as they become available. An average is calculated for each result and the mean population error is attributed. If a web tool is offline, and therefore no result is sent back to the platform, the term "Error" will be the output and the specific test is discarded from the analyses. Charts are generated based on the result averages **(F)**. 

\* Unavailable ADFS webservice since February, 2015. When the server comes online, PA<sup>3</sup>P will connect automatically.

```

    </div>
  </td>
</tr>
</table>

```

<p>

PA<sup>3</sup>P was built for categorizing proteins as allergens or antigens according to the biological characteristics embedded within their amino acids chain by using a specific set of web based tools and statistical analyzes of results. The PA<sup>3</sup>P output can be used to support or to discard the employment of a given protein in Genetically Modified Organism (GMO) construction. The platform allows allergen and non-allergen protein sorting with significant reduction on false positive or false negative cases and a substantial increase in sensibility and specificity (Table 1 and Figure 2).

</p>

<p>

Table 1. Possible outcomes and sensibility, specificity, and accuracy formula.

</p>

<p>

&nbsp;

</p>

<p>

&nbsp;Where:

</p>

<p>

&nbsp;

</p>

<p>

The equations are applied based on a binary matrix (0 = presence; 1 = absence) which summarizes all the performed tests with the datasets, so each sequence in the datasets (allergen and non-allergen sequence datasets are available at the "Dataset" section) was submitted to analysis both on web tools independently and on PA<sup>3</sup>P.

</p>



```
<h3 class="titulo-borda-janela borda-janela" style="width: 100%; margin-bottom:5px; ">SUBMITTING
YOUR SEQUENCE TO PA<sup>3</sup>P ANALYSIS</h3>
```

```
<div class="borda-janela" style="padding-top: 8px; padding-bottom: 8px;width: 100%;" >
```

```
<table style="width: 100%; padding-right: 20px; padding-left: 20px;" >
```

```
<tr>
```

```
<td>
```

```
<p>
```

&nbsp;The first step to PA<sup>3</sup>P analysis is to insert a query sequence within the text box ("PA<sup>3</sup>P" tab) (Figure 4).

```
<br>
```

Allergenic protein amino acid sequence example:

```
</p>
```

```
<p style="font-style: italic;">
```

```
>gi|14423645|sp|P79085.1|ALTA1_ALTAL RecName: Full=Major
```

allergen Alt a 1;

```
AltName: Allergen=Alt a 1; Flags: Precursor
```

```
<br>
```

```
MQFTTIASLFAAAGLAAAAPLESRQDTASCPVTTEGDYVWKISEFYGRKPEGTYYNLSLGFNIKA
TNGGTL<br>
```

```
DFTCSAQADKLEDHKWYSCGENSFMDFSFSDRSGLLLKQKVSDDITYVATATLPNYCRAGG
NGPKDFVC<br>
```

```
QGVADAYITLVTLPKSS
```

```
</p>
```

```
<p>
```

Non-allergenic protein amino acid sequence example:

```
</p>
```

```
<p style="font-style: italic;">
```

```
>gi|68161511|ref|NP_001020131.1| actin-related protein 2/3
```

complex subunit 4 isoform b [Homo sapiens]

```
<br>
```

```
MRFMMMRAENFFILRRKPVEGYDISFLITNFHTEQMYKHKLVDFVIHFMEIIDKEISEMKLSVN
ARARIV<br>
```

```
AEEFLKNF
```

```
</p>
```

```
<table class="table-imagem">
```

```
<tr>
```

```
<td>
```

```

```

```
</td>
```

```
</tr>
```

```
<tr>
```

```
<td>
```

```
<div style="padding: 10px;">
```

```
<b>Figure 4 – Text box for
```

query sequence insertion and submission to PA<sup>3</sup>P analysis.

```
</b> <br>
```

The sequence is inserted in the blank space and sent for analysis by hitting "Run". A new search can be made by hitting "New Query".

```
</div>
```

```
</td>
```

```
</tr>
```

```
</table>
```

```
<p>
```

The query sequence is subjected to several prediction methods concomitantly, depending on the web tool: the guidelines proposed by the FAO/WHO, algorithm prediction and sequence alignment to known allergens datasets. The results (allergen or non-allergen) are compiled and a score sheet (Figure 5) is filled. The average accuracy is then determined. If a given result is not informed (e.g., if a webserver is offline) the correspondent web tool accuracy value is discarded from the analysis.

```

</p>
<table class="table-imagem">
  <tr>
    <td>
      
    </td>
  </tr>
  <tr>
    <td>
      <div style="padding: 10px;">
        <b>Figure 5 – Accuracy score

```

sheet.

```

</b> <br>

```

As the query is analyzed on each web tool, a score sheet is filled with each result (allergen or non-allergen). If a web tool error occur and no result is obtained, the correspondent web tool accuracy is discarded from analysis. A final average is calculated according to each approach accuracy. <p>\* Unavailable ADFS webservice since February, 2015. When the server comes online, PA<sup>3</sup>P will connect automatically. </p>

```

</div>
</td>
</tr>
</table>

```

```

<p>

```

After the query analysis, a table presents the results of each web tool according to the sequence category (allergen or non-allergen) (Figure 6).

```

</p>

```

```

<table class="table-imagem">
  <tr>
    <td>
      
    </td>
  </tr>
  <tr>
    <td>
      <div style="padding: 10px;">
        <b>Figure 6 – Individual results

```

table.

```

</b> <br>

```

Left row informs the web tool denomination, center row informs the approach used by each web tool, and right row shows the result from each web tool. On the top, the user is able to visualize the web tools outputs by clicking the respective tab. If no output is obtained, an "Error" message is informed.

```

</div>
</td>
</tr>
</table>

```

```

<p>

```

The similarity based prediction approach compares two or more amino acid sequences in order to assess the query sequence homology and biological information. Such

comparisons, known as alignments, aim identifying shared sequence characteristics. Five similarity/homology based web tools are compiled on PA<sup>3</sup>P: Allergen Online – 8 aa wordmatch; Allergen Online – 80 aa window; AlgPred – IgE (Saha & Raghava, 2006); AlgPred – ARP (Saha & Raghava, 2006); and ADFS – Motif-based (Nakamura, Nakamura & Teshima, 2009).

</p>

<div class="accordion">

<h3 style="text-align: left">About the similarity based prediction approaches</h3>

<div>

<p>

&nbsp;The web tool Allergen Online (available at <http://allergenonline.com/>) allows the comparison of the query sequence with an allergenic protein sequences dataset. The prediction follows the FAO/WHO guidelines:

</p>

<p>

&nbsp;- Allergen Online 80 aa window: a query protein is potentially an allergen if it presents over 35% identity in a 80 amino acid window when compared with known allergens.

</p>

<p>

&nbsp;- Allergen Online 8 aa wordmatch: a query protein is potentially an allergen if it presents six or more contiguous amino acids (id length  $n \geq 6$ ) when compared with known allergens.

</p>

<p>

&nbsp;The Allergen Online web tool has its own dataset composed by 1,897 allergenic protein sequences.

</p>

<p>

&nbsp;The Prediction of Allergenic Proteins and Mapping of IgE Epitopes (AlgPred; Saha & Raghava, 2006), available at <http://www.imtech.res.in/raghava/algpred/> is also a similarity based prediction web tool. This tool maps eliciting epitopes of immunoglobulin E (IgE) in the query sequence. The method considers a protein as a potential allergen when the sequence possesses one or more IgE epitopes. The AlgPred database has 183 IgE epitopes. The method also performs a search for allergen representative peptides (ARPs). This approach is highly sensitive and accurate. It consists in a 2,890 ARPs, and when a hit is significant, the query is promptly characterized as potential allergen.

</p>

<p>

&nbsp;The Allergen Database for Food Safety (ADFS; Nakamura, Nakamura & Teshima, 2009), available at <http://allergen.nihs.go.jp/ADFS/> has a database consisted of more than 1,300 allergen sequences. The tool performs a search for known allergen motifs in the query sequence.

</p>

</div>

</div>

<p>

&nbsp;Another approach is the in silico allergen prediction based on algorithms using models derived from a given class of sequences. The Support Vector Machine (SVM) is a set of methods of supervised learning that analyses data in order to identify patterns, and therefore classify a query sequence. In this kind of learning, the inducer (learning algorithm) receives examples (patterns) for training. These patterns are described by certain attributes and by the class label (in this case, allergens). Two SVMs are implemented on PA<sup>3</sup>P: (i) SVM-amino acid composition (Wang et al., 2012); and (ii) SVM dipeptide (Saha & Raghava, 2006).

</p>

<div class="accordion">

<h3 style="text-align: left">About the algorithm based prediction approaches</h3>

<div>

The SVM Amino Acid Composition (SVM-AA) built on PA<sup>3</sup>P is derived from Protein Allergenicity Prediction (proAP; Wang et al., 2012) available at <http://gmobl.sjtu.edu.cn/proAP/main.html>.

The prediction based on Support Vector Machine – Dipeptide is conducted by the Prediction of Allergenic Proteins and Mapping of IgE Epitopes web tool (AlgPred; Saha & Raghava, 2006), available at <http://www.imtech.res.in/raghava/algpred/>.

In order to ease results interpretation generated by PA<sup>3</sup>P, a series of charts are generated (Figure 7).

interpretation charts.

result. **B** – Similarity based result. **C** – Summary result. Chart heading shows tabbed results alongside the standard error calculated for each method. The standard error ((standard error=standard deviation)<sup>1/2</sup>(sample)) is required since PA<sup>3</sup>P works with a representative sample of a specific population. This error can be added to or subtracted from the result, allowing the inference that the sampled population behaves similarly to the obtained result.

charts (Figure 8).

download.

By clicking the horizontal parallel bars at the upper right corner, user can download the result carts in different file formats.

```

                </td>
            </tr>
        </table>

        <p>
        Other tools and methods may be added in the future.
        </p>
        <p>
        For publishing, please cite.
        </p>

        <p>
        <b>References:</b><br>
        - Allergen Online, available at http://allergenonline.com/. Access in february
of 2015. <br>
        - CHRYSOSTOMOU, C.; SEKER, H. <b>Prediction of protein allergenicity
based on signal-processing bioinformatics approach.</b> Conf Proc IEEE Eng Med Biol Soc., 808-811, 2014;
        <br>
        - FAO/WHO. <b>Evaluation of Allergenicity of Genetically Modified
Foods.</b> Joint FAO/WHO Expert Consultation on Allergenicity of Foods Derived from Biotechnology, 27 pg,
2001;

        <br>

        - MIRSKY, H.P.; CRESSMAN JR., R.F.; LADICS, G.S.; <b>Comparative
assessment of multiple criteria for the in silico prediction of cross-
reactivity of proteins to known allergens.</b> Regulatory Toxicology and
Pharmacology, 67, 232–239, 2013;
        <br>

        - NAKAMURA, R.; NAKAMURA, R.; TESHIMA,R.; <b>Major revision of
the allergen database for food safety (ADFS) and validation of the motif-based allergenicity prediction tool.</b>
Kokuritsu Iyakuhiin Shokuhin Eisei Kenkyusho Hokoku, 127, 44-49, 2009.
        <br>

        - SAHA, S.; RAGHAVA, G.P.S.; <b>AlgPred: Prediction of Allergenic
Proteins and Mapping of IgE Epitopes.</b> Nucleic Acids Research, 34, 202-209, 2006;
        <br>

        - SOUZA JUNIOR, M.T.; MARTINS, N.F.; <b>Predição do Potencial de
Alergenicidade em OGMs – Estudo de Caso.</b> Biotecnologia, Ciência e Desenvolvimento, 30, 10-15, 2003;
        <br>

        - STADLER, M.B.; STADLER, B.M.; <b>Alergenicity Prediction by Protein
Sequence.</b> The FASEB Journal, 17, 1141-1143, 2003;

        <br>

        - WANG, J.; et al.; <b>Evaluation and Integration of Existing Methods for
Computational Prediction of Allergens.</b> BMC Bioinformatics, 14, 1-9, 2013;
        <br>
        - ZORZET, A.; GUSTAFSSON, M.; HAMMERLING, U.; <b>Prediction of
Food Protein Allergenicity: a Bioinformatic Learning Systems Approach.</b> In Silico Biology, 2, 525-534,
2002;

        <br>
    </td>
</tr>
</table>
</div>
</body>
</html>

```

#### 4.6 Códigos-fonte referentes à análise na PA<sup>3</sup>P

```

<% @page import="org.unipampa.lpa.classe.Allermatch"%>

<% @ page language="java" contentType="text/html; charset=UTF-8" pageEncoding="UTF-8"%>

<% @ include file="/menu.jsp" %>

<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>

<head>
    <meta http-equiv="Content-type" content="text/html;charset=UTF-8">
    <meta HTTP-EQUIV="Pragma" CONTENT="no-cache" />
    <meta HTTP-EQUIV="Cache-control" CONTENT="no-cache" />
    <meta HTTP-EQUIV="Expires" CONTENT="0" />

    <script type="text/javascript" src="<%=request.getContextPath()%>/js/mask.js" charset="utf-
8"></script>
    <script type="text/javascript" src="<%=request.getContextPath()%>/pa3p/grafico.js" charset="utf-
8"></script>
</head>

</head>

<script type="text/javascript">
var total = 0;
var totalalergenico = 0.00;
var totalnaoalergenico = 0.00;
var ct = 0;
var ctpred = 0;
var ctsim = 0;
var qtproc = 0;

var totalsim = 0, totalpred = 0;
var totalalergenicosisim = 0.00, totalalergenicopred = 0.00;
var totalnaoalergenicosisim = 0.00, totalnaoalergenicopred = 0.00;

var WAITGIF = '<div style="text-align: center;margin: 0 auto;"></div>';

function newquery(){
    $("#tabs").tabs("option", "active", 0);
    $("#tabs").tabs("option", "disabled", [ 1, 2, 3, 4 ] );

    document.getElementById("allermatch").innerHTML = "";

    document.getElementById("resultadoalergenico").innerHTML = "";
    document.getElementById("resultadonaoalergenico").innerHTML = "";

    document.getElementById("resultadoalergenicosisim").innerHTML = "";
    document.getElementById("resultadonaoalergenicosisim").innerHTML = "";

    document.getElementById("resultadoalergenicopred").innerHTML = "";
    document.getElementById("resultadonaoalergenicopred").innerHTML = "";

    for (var i=0; i<2;i++){

        document.getElementById("idv1allergenonline" + (i)).innerHTML = "";
        document.getElementById("idresultallergenonline" + (i)).innerHTML = "";
    }
}

```

```

}

for (var i=0; i<5;i++){
    if (i != 1 && i !=2){
        document.getElementById("idvialgpred" + (i + 1)).innerHTML = "";
        document.getElementById("idresultalgpred" + (i + 1)).innerHTML = "";
    }
}

document.getElementById("idvlproap0").innerHTML = "";
document.getElementById("idresultproap0").innerHTML = "";

document.getElementById("idvladfs0").innerHTML = "";
document.getElementById("idresultadfs0").innerHTML = "";

$('#container').highcharts().series[0].setData([0,0]);
$('#containersim').highcharts().series[0].setData([0,0]);
$('#containerpred').highcharts().series[0].setData([0,0]);

document.forms['formpesq']['ds_sequence'].value="";
document.forms['formpesq']['ds_sequence'].focus();
}

$(function() {
    $("#tabs").tabs();
//    $("#tabs").tabs("option", "disabled", [ 1, 2, 3 ] );
    inicializacontroles();
    newquery();
});

function run(){

    totalalergenico = 0.00;
    totalnaoalergenico = 0.00;
    qtproc = 0;
    ct = 0;
    ctsim = 0;
    ctpred = 0;

    totalsim = 0;
    totalpred = 0;
    totalalergenicosim = 0.00;
    totalalgenicopred = 0.00;
    totalnaoalergenicosim = 0.00;
    totalnaoalgenicopred = 0.00;

    $("#tabs").tabs("option", "active", 0 );

    $("#tabs").tabs("option", "disabled", [ 1, 2, 3, 4 ] );

    document.getElementById("resultadoalergenico").innerHTML = WAITGIF; ;
    document.getElementById("resultadonaoalergenico").innerHTML = WAITGIF; ;

    document.getElementById("resultadoalgenicosim").innerHTML = WAITGIF; ;
    document.getElementById("resultadonaoalgenicosim").innerHTML = WAITGIF; ;

    document.getElementById("resultadoalgenicopred").innerHTML = WAITGIF; ;
    document.getElementById("resultadonaoalgenicopred").innerHTML = WAITGIF; ;
}

```

```

for (var i=0; i<2;i++){
document.getElementById("idvllallergenonline" + (i)).innerHTML = WAITGIF;
document.getElementById("idresultallergenonline" + (i)).innerHTML = WAITGIF;
}

for (var i=0; i<5;i++){
if (i != 1 && i !=2){
document.getElementById("idvlgpred" + (i + 1)).innerHTML = WAITGIF;
document.getElementById("idresultalgpred" + (i + 1)).innerHTML = WAITGIF;
}
}

document.getElementById("idviproap0").innerHTML = WAITGIF;
document.getElementById("idresultproap0").innerHTML = WAITGIF;

document.getElementById("idvldfs0").innerHTML = WAITGIF;
document.getElementById("idresultadfs0").innerHTML = WAITGIF;

$('#container').highcharts().series[0].setData([0,0]);
$('#containersim').highcharts().series[0].setData([0,0]);
$('#containerpred').highcharts().series[0].setData([0,0]);

//ALGPRED - INICIO
$.ajax({
type: "POST",
url: "algpredresult.jsp",
data: $('#formpesq').serialize(),
cache: true,
success: function(response) {
document.getElementById("algpred").innerHTML = response;

$('# #accordionalgpred" ).accordion({collapsible: true, heightStyle: "content", active: false });
$('# #tabs" ).tabs( "enable", 1 );

var vlinput = document.getElementById('algpredresultform').elements["vlgpred[]"];
var vlresult = document.getElementById('algpredresultform').elements["resultalgpred[]"];

for (i = 0; i < vlinput.length; i++) {

/*
0 - Algpred - IgE - Similarity
3 - Algpred - SVM Dipeptide - Predicted
4 - Algpred - ARP - Similarity
*/
if (i != 1 && i != 2){

if (i == 0 || i == 4)
document.getElementById("idvlgpred" + (i + 1)).innerHTML = "Similarity
(" + vlinput[i].value + "%)";
else
document.getElementById("idvlgpred" + (i + 1)).innerHTML = "Algorithm
(" + vlinput[i].value + "%)";

document.getElementById("idresultalgpred" + (i + 1)).innerHTML =
vlresult[i].value;

if (vlresult[i].value == "Allergen"){
ct = ct + 1;
totalalergenico = parseFloat(totalalergenico) + parseFloat(vlinput[i].value);
}
}
}
}

```

```

        if (i == 0 || i == 4){
            ctsim = ctsim + 1;
            totalalergicosim = parseFloat(totalalergicosim) +
parseFloat(vlininput[i].value);
        }
        else{
            ctpred = ctpred + 1;
            totalalergicopred = parseFloat(totalalergicopred) +
parseFloat(vlininput[i].value);
        }
    }
    else if (vlresult[i].value == "Non-Allergen"){
        ct = ct+ 1;
        totalnaoalergenico = parseFloat(totalnaoalergenico) +
parseFloat(vlininput[i].value);
    }
    if (i == 0 || i == 4){
        ctsim = ctsim + 1;
        totalnaoalergicosim = parseFloat(totalnaoalergicosim) +
parseFloat(vlininput[i].value);
    }
    else{
        ctpred = ctpred + 1;
        totalnaoalergicopred = parseFloat(totalnaoalergicopred) +
parseFloat(vlininput[i].value);
    }
}
}

qtproc = qtproc + 1;
finaliza();
}
});
//Fim Algpred

//PROAP - INICIO
$.ajax({
    type: "POST",
    url: "proapresult.jsp",
    data: $('#formpesq').serialize(),
    cache: true,
    success: function(response) {
        document.getElementById("proap").innerHTML = response;

        $( "#tabs" ).tabs( "enable", 2 );

        var vlininput = document.forms['proapresultform']['vlproap'].value;
        var vlresult = document.forms['proapresultform']['resultproap'].value;

        document.getElementById("idvlproap0").innerHTML = "Algorithm (" + vlininput + "%)";
        document.getElementById("idresultproap0").innerHTML = vlresult;

        if (vlresult == "Allergen"){
            ct = ct+ 1;
            totalalergenico = parseFloat(totalalergenico) + parseFloat(vlininput);
            ctpred = ctpred + 1;

```

```

        totalalergenicopred = parseFloat(totalalergenicopred) + parseFloat(vlininput);
    }
    else if (vlresult == "Non-Allergen"){
        ct = ct+ 1;
        totalnaoalergenicico = parseFloat(totalnaoalergenicico) + parseFloat(vlininput);
        ctpred = ctpred + 1;
        totalnaoalergenicopred = parseFloat(totalnaoalergenicopred) + parseFloat(vlininput);
    }

    qtproc = qtproc + 1;
    finaliza();

}
});
//Fim proap

//ADFS - INICIO
$.ajax({
    type: "POST",
    url: "adsresult.jsp",
    data: $('#formpesq').serialize(),
    cache: true,
    beforeSend: function(){
        //document.getElementById("algpred").innerHTML = '<div style="text-align: center;margin: 0
auto;padding-top: 50px;padding-bottom: 50px;"></div>';
    },
    success: function(response) {
        document.getElementById("ads").innerHTML = response;

        $("#tabs").tabs("enable", 3);

        var vlinput = document.forms['adsresultform']['vladfs'].value;
        var vlresult = document.forms['adsresultform']['resultadfs'].value;

        document.getElementById("idvladfs0").innerHTML = "Similarity (" + vlinput + "%)";
        document.getElementById("idresultadfs0").innerHTML = vlresult;

        if (vlresult == "Allergen"){
            ct = ct+ 1;
            totalalergenicico = parseFloat(totalalergenicico) + parseFloat(vlinput);
            ctsim = ctsim + 1;
            totalalergenicocosim = parseFloat(totalalergenicocosim) + parseFloat(vlinput);
        }
        else if (vlresult == "Non-Allergen"){
            ct = ct+ 1;
            totalnaoalergenicico = parseFloat(totalnaoalergenicico) + parseFloat(vlinput);
            ctsim = ctsim + 1;
            totalnaoalergenicocosim = parseFloat(totalnaoalergenicocosim) + parseFloat(vlinput);
        }

        qtproc = qtproc + 1;
        finaliza();

    }
}); //Fim ADFS

//Allergen Online - INICIO
$.ajax({
    type: "POST",

```

```

url: "allergenonlineresult.jsp",
data: $('#formpesq').serialize(),
cache: true,
beforeSend: function(){
    //document.getElementById("algpred").innerHTML = '<div style="text-align: center;margin: 0
auto;padding-top: 50px;padding-bottom: 50px;"></div>';

    },
success: function(response) {
    document.getElementById("allergenonline").innerHTML = response;

    $('#accordionallergenonline').accordion({collapsible: true, heightStyle: "content", active: false
});

    $('#tabs').tabs("enable", 4);

    var vlininput =
document.getElementById('allergenonlineresultform').elements["vlallergenonline[]"];
    var vlresult =
document.getElementById('allergenonlineresultform').elements["resultallergenonline[]"];

    for (i = 0; i < 2; i++) {

        /*
            0 - allergenonline - Sliding 80mer Window
            1 - allergenonline - 8mer Exact Match
        */
        document.getElementById("idvlallergenonline" + (i)).innerHTML = "Similarity (" +
vlininput[i].value + "%)";
        document.getElementById("idresultallergenonline" + (i)).innerHTML = vlresult[i].value;

        if (vlresult[i].value == "Allergen"){
            ct = ct + 1;
            totalalergenico = parseFloat(totalalergenico) + parseFloat(vlininput[i].value);
            ctsim = ctsim + 1;
            totalalergenicosim = parseFloat(totalalergenicosim) +
parseFloat(vlininput[i].value);
        }
        else if (vlresult[i].value == "Non-Allergen"){
            ct = ct+ 1;
            totalnaoalergenico = parseFloat(totalnaoalergenico) +
parseFloat(vlininput[i].value);
            ctsim = ctsim + 1;
            totalnaoalergenicosim = parseFloat(totalnaoalergenicosim) +
parseFloat(vlininput[i].value);
        }

    }

    qtproc = qtproc + 1;
    finaliza();

    }
}); //Fim AllergenOnline

}

function finaliza(){

    //Só finaliza se rodaram todos os algoritmos
    if (qtproc < 4) return;

```

```

        if (ct > 0 && totalalergenico > 0)
            document.getElementById("resultadoalergenico").innerHTML = (totalalergenico /
ct).toFixed(2);
        else
            document.getElementById("resultadoalergenico").innerHTML = "-";

        if (ct > 0 && totalnaoalergenico > 0)
            document.getElementById("resultadonaoalergenico").innerHTML = (totalnaoalergenico /
ct).toFixed(2);
        else
            document.getElementById("resultadonaoalergenico").innerHTML = "-";

        $('#container').highcharts().series[0].setData([Number((totalalergenico / ct).toFixed(2)),
Number((totalnaoalergenico / ct).toFixed(2))]);

        //SIMILARIDADE
        if (ctsim > 0 && totalalergenicosim > 0)
            document.getElementById("resultadoalergenicosim").innerHTML = (totalalergenicosim /
ctsim).toFixed(2);
        else
            document.getElementById("resultadoalergenicosim").innerHTML = "-";

        if (ctsim > 0 && totalnaoalergenicosim > 0)
            document.getElementById("resultadonaoalergenicosim").innerHTML =
(totalnaoalergenicosim / ctsim).toFixed(2);
        else
            document.getElementById("resultadonaoalergenicosim").innerHTML = "-";

        $('#containersim').highcharts().series[0].setData([Number((totalalergenicosim / ctsim).toFixed(2)),
Number((totalnaoalergenicosim / ctsim).toFixed(2))]);

        //PREDIÇÃO
        if (ctpred > 0 && totalalergenicipred > 0)
            document.getElementById("resultadoalergenicipred").innerHTML = (totalalergenicipred /
ctpred).toFixed(2);
        else
            document.getElementById("resultadoalergenicipred").innerHTML = "-";

        if (ctsim > 0 && totalnaoalergenicipred > 0)
            document.getElementById("resultadonaoalergenicipred").innerHTML =
(totalnaoalergenicipred / ctpred).toFixed(2);
        else
            document.getElementById("resultadonaoalergenicipred").innerHTML = "-";

        $('#containerpred').highcharts().series[0].setData([Number((totalalergenicipred / ctpred).toFixed(2)),
Number((totalnaoalergenicipred / ctpred).toFixed(2))]);
    }
</script>

<body>

<h3 class="titulo-borda-janela borda-janela" style="width: 800px; margin-bottom:5px;">PA³P Search</h3>
<div id="cabpesq" class="borda-janela" style="padding-top: 8px; padding-bottom: 8px; width: 800px;" >
    <fieldset>
        <form action="" name="formpesq" id="formpesq" >

            <label style="width: auto;">Sequence: </label>
            <br>
            <textarea name="ds_sequence" rows="5" style="width: 780px;"></textarea>

```

```

        <br>
        <button class="botao" id="bpesquisar" onClick="run()" style="float: right; margin-
left: 5px; margin-top: 10px;">Run</button>
        <button class="botao" id="blimpar" onClick="newquery()" style="float: right;
margin-top: 10px;">New Query</button>
    </form>
</fieldset>
</div>

```

```

<div id="tabs" class="borda-janela" style="margin-top: 10px; width: 1010px;">

```

```

    <ul>
    <li><a href="#pa3p">PA3P</a></li>
    <li><a href="#algpred">Algpred</a></li>
    <li><a href="#proap">Proap</a></li>
    <li><a href="#adfs">ADFS</a></li>
    <li><a href="#allergenonline">Allergen Online</a></li>
    </ul>

```

```

<div id="pa3p" class="borda-janela" style="margin-top: 10px; width: 1000px; margin: 0 auto;">

```

```

    <table style="margin: 0 auto; padding-top: 20px; padding-bottom: 20px;">

```

```

        <tr>
        <td style="vertical-align: super;">

```

```

            <table width="600px" style="margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >

```

```

                <thead>
                <tr style="text-align: center;">
                <th width="100%" scope="col"><b>Individual Results</b></th>
                </tr>
                </thead>
            </table>

```

```

            <table id="pa3presume" width="600px" style="margin-left: auto; margin-right: auto; text-
align: left;" class='tabela' >

```

```

                <thead>
                <tr>
                <th width="40%" scope="col">WebTool</th>
                <th width="30%" scope="col">Method</th>
                <th width="30%" scope="col">Result</th>
                </tr>
                </thead>
                <tbody>

```

```

                    <tr>
                    <td>Algpred - IgE</td>
                    <td id="idvlgpred1"></td>
                    <td id="idresultalgpred1"></td>
                    </tr>
                    <tr>
                    <td>Algpred - SVM Dipeptide</td>
                    <td id="idvlgpred4"></td>
                    <td id="idresultalgpred4"></td>
                    </tr>
                    <tr>
                    <td>Algpred - ARP</td>
                    <td id="idvlgpred5"></td>
                    <td id="idresultalgpred5"></td>

```



```

                <td id="resultadonoalergenicopred"></td>
            </tr>
        </tbody>
    </table>
    <br>
    <div id="containerpred" style="min-width: 310px; height: 300px; margin: 0
auto"></div>

</td>

<td style="vertical-align:super; padding-left: 0px;">

    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
        <thead>
            <tr style="text-align:center;">
                <th width="100%" scope="col"><b>Similarity Based
Result</b></th>
            </tr>
        </thead>
    </table>

    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
        <thead>
            <tr>
                <th width="60%" scope="col">Result</th>
                <th width="40%" scope="col">% (&#177; 0.87%)</th>
            </tr>
        </thead>
        <tbody>
            <tr>
                <td>Allergen</td>
                <td id="resultadoalergenicosisim"></td>
            </tr>
            <tr>
                <td>Non-Allergen</td>
                <td id="resultadonoalergenicosisim"></td>
            </tr>
        </tbody>
    </table>
    <br>
    <div id="containersim" style="min-width: 310px; height: 300px; margin: 0
auto"></div>

</td>

<td style="vertical-align:super; padding-left: 0px;">

    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
        <thead>
            <tr style="text-align:center;">
                <th width="100%" scope="col"><b>Summary
Result</b></th>
            </tr>
        </thead>
    </table>

```

```

class='tabela' >
    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
        <thead>
            <tr>
                <th width="60%" scope="col">Result</th>
                <th width="40%" scope="col">% (&#177; 0.89%)</th>
            </tr>
        </thead>
        <tbody>
            <tr>
                <td>Allergen</td>
                <td id="resultadoalergenico"></td>
            </tr>
            <tr>
                <td>Non-Allergen</td>
                <td id="resultadonaoalergenico"></td>
            </tr>
        </tbody>
    </table>
    <br>
    <div id="container" style="min-width: 310px; height: 300px; margin: 0 auto"></div>
</td>
</tr>
</table>
</div>
<div id="allermatch" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<div id="algpred" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<div id="proap" class="borda-janela" style="margin-top: 10px;width: 990px; word-wrap: break-word; ">
</div>
<div id="adfs" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<div id="allergenonline" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<!-- <div id="blastp" class="borda-janela" style="margin-top: 10px;width: 990px; text-align: left;"> -->
<!-- </div> -->
</div>
</body>
</html>
4.7 Códigos-fonte referentes aos gráficos gerados pela PA3P
$(function () {
    $('#container').highcharts({

```

```

chart: {
  type: 'column'
},
legend: {
  enabled: false
},
title: {
  text: 'Summary Result'
},
xAxis: {
  categories: ['Allergen', 'Non-Allergen']
},
yAxis: {
  min: 0,
  title: {
    text: '%'
  },
  stackLabels: {
    enabled: true,
    style: {
      fontWeight: 'bold',
      color: (Highcharts.theme && Highcharts.theme.textColor) || 'gray'
    }
  }
},
plotOptions: {
  column: {
    stacking: 'normal',
    dataLabels: {
      enabled: false,
      color: (Highcharts.theme && Highcharts.theme.dataLabelsColor) || 'white',
      style: {
        textShadow: '0 0 3px black'
      }
    }
  }
},
series: [{
  data: [0, 0]
}]
});

```

//SIMILARIDADE

```

$('#containersim').highcharts({
  chart: {
    type: 'column'
  },
  legend: {
    enabled: false
  },
  title: {
    text: 'Similarity Based Result'
  },
  xAxis: {
    categories: ['Allergen', 'Non-Allergen']
  },
  yAxis: {
    min: 0,
    title: {

```

```

        text: '(%)'
    },
    stackLabels: {
        enabled: true,
        style: {
            fontWeight: 'bold',
            color: (Highcharts.theme && Highcharts.theme.textColor) || 'gray'
        }
    }
},

plotOptions: {
    column: {
        stacking: 'normal',
        dataLabels: {
            enabled: false,
            color: (Highcharts.theme && Highcharts.theme.dataLabelsColor) || 'white',
            style: {
                textShadow: '0 0 3px black'
            }
        }
    }
},
series: [{
    data: [0, 0]
}]
});
//PREDIÇÃO
$('#containerpred').highcharts({
    chart: {
        type: 'column'
    },
    legend: {
        enabled: false
    },
    title: {
        text: 'Algorithm Based Result'
    },
    xAxis: {
        categories: ['Allergen', 'Non-Allergen']
    },
    yAxis: {
        min: 0,
        title: {
            text: '(%)'
        },
    },
    stackLabels: {
        enabled: true,
        style: {
            fontWeight: 'bold',
            color: (Highcharts.theme && Highcharts.theme.textColor) || 'gray'
        }
    }
},

plotOptions: {
    column: {
        stacking: 'normal',
        dataLabels: {
            enabled: false,

```

```
        color: (Highcharts.theme && Highcharts.theme.dataLabelsColor) || 'white',
        style: {
            textShadow: '0 0 3px black'
        }
    },
    series: [{
        data: [0, 0]
    }]
});
});
```





```
“About” section.      <p style="text-align:justify;">
                        &nbsp;For more information on how PA3P works, please visit the
                        </p>
                        </td>
                    </tr>
                </table>
                <br>
</div>
</body>
</html>
```

#### 4.10 Códigos-fonte referentes à seção “Contact”

```
<%@ include file="/menu.jsp" %>

<style>

    a {
        font-size: small;
    }

</style>

<body onload="inicializacontroles()">

<!-- <h3 class="titulo-borda-janela borda-janela" style="width: 800px; margin-bottom:5px;">Contact</h3> -->

<h3 class="titulo-borda-janela borda-janela" style="width: 800px; margin-bottom:5px; ">Contact</h3>

<div class="borda-janela" style="padding: 40px;width: 720px; text-align: center;" >

Anna Carolina Boeck - <a href="mailto:anna.boeck@gmail.com" target="_blank">Email:
anna.boeck@gmail.com</a>
<br><br>
Juliano Tomazzoni Boldo - <a href="mailto:jtboldo@gmail.com" target="_blank">Email:
jtboldo@gmail.com</a>
<br>
</div>
</body>
```

### 4.11 Códigos-fonte referentes à seção “Usefull Links”

```

<% @ include file="/menu.jsp" %>

<style>

    a {
        font-size: small;
    }

</style>

<body onload="inicializacontroles()">

<!-- <h3 class="titulo-borda-janela borda-janela" style="width: 800px; margin-bottom:5px;">Usefull
Links</h3> -->

<h3 class="titulo-borda-janela borda-janela" style="width: 800px; margin-bottom:5px; ">Useful Links</h3>

<div class="borda-janela" style="padding: 40px;width: 720px; text-align: left;" >

Universidade Federal do Pampa - <a href="http://www.unipampa.edu.br"
target="_blank">(www.unipampa.edu.br)</a>
<br><br>Allermatch - <a href="http://www.allermatch.org/" target="_blank">(www.allermatch.org)</a>
<br><br>AlgPred - <a href="http://www.imtech.res.in/raghava/algpred/"
target="_blank">(www.imtech.res.in/raghava/algpred)</a>
<br><br>ADFS - <a href="http://allergen.nihs.go.jp/ADFS/" target="_blank">(allergen.nihs.go.jp/ADFS)</a>
<br><br>ProAP - <a href="http://gmobl.sjtu.edu.cn/proAP/" target="_blank">(gmobl.sjtu.edu.cn/proAP)</a>
<br><br>Allergen Online - <a href="http://allergenonline.com/"
target="_blank">(www.allergenonline.com)</a>
<br><br>NCBI - <a href="http://ncbi.nlm.nih.gov/" target="_blank">(ncbi.nlm.nih.gov)</a>
<br><br>Uniprot - <a href="http://www.uniprot.org/" target="_blank">(www.uniprot.org)</a>
<br><br>BLAST - <a href="http://blast.ncbi.nlm.nih.gov/Blast.cgi"
target="_blank">(blast.ncbi.nlm.nih.gov/Blast.cgi)</a>
<br><br>Peptide Match - <a href="http://research.bioinformatics.udel.edu/peptidematch/index.jsp"
target="_blank">(research.bioinformatics.udel.edu/peptidematch/index.jsp)</a>

</div>
</body>

```

### 4.12 Códigos-fonte referentes ao cabeçalho da PA<sup>3</sup>P

```

<!DOCTYPE html>
<% @page import="org.dotto.util.Util"%>
<% @page import="java.util.Date"%>

<% @ page language="java" contentType="text/html; charset=UTF-8" pageEncoding="UTF-8"%>

<head>
    <meta http-equiv="Content-type" content="text/html;charset=UTF-8">
    <meta HTTP-EQUIV="Pragma" CONTENT="no-cache" />
    <meta HTTP-EQUIV="Cache-control" CONTENT="no-cache" />
    <meta HTTP-EQUIV="Expires" CONTENT="0" />

    <link href="<%=request.getContextPath()%>/images/favicon.ico" rel="icon" type="image/x-icon" />
    <link href="<%=request.getContextPath()%>/images/favicon.ico" rel="shortcut icon" type="image/x-icon"
/>

    <link href="<%=request.getContextPath()%>/jquery/css/cupertino/jquery-ui-
1.10.3.custom.min.css" rel="stylesheet">
    <link href="<%=request.getContextPath()%>/css/tabelas.css" rel="stylesheet">
    <link href="<%=request.getContextPath()%>/css/plurigestao.css" rel="stylesheet">
    <link href="<%=request.getContextPath()%>/css/multiple-select.css" rel="stylesheet"/>

    <script src="<%=request.getContextPath()%>/jquery/js/jquery-1.10.2.min.js"></script>
    <script src="<%=request.getContextPath()%>/jquery/js/jquery-ui-
1.10.3.custom.min.js"></script>
    <script src="<%=request.getContextPath()%>/js/jquery.multiple.select.js"></script>
    <script src="<%=request.getContextPath()%>/js/jquery.tablesorter.min.js"></script>

    <script src="<%=request.getContextPath()%>/js/jquery.highchartTable-min.js"></script>
    <script src="<%=request.getContextPath()%>/js/highcharts.js"></script>
    <script src="<%=request.getContextPath()%>/js/highchart/modules/exporting.js"></script>

    <script type="text/javascript" src="<%=request.getContextPath()%>/js/mask.js" charset="utf-
8"></script>
</head>

<title>PA3P - Platform for Analysis of Allergenicity and Antigenicity of Proteins</title>

<table width=100% style="padding-bottom: 10px;" >
<tr>
<td width=auto style="vertical-align: top;">
    
</td>
<td width=60% style="vertical-align: bottom;">
    <font style="font-size: 60px;"><b>PA<sup>3</sup>P</b></font>
<br>
    <font style="font-size: 15px;">Platform for Analysis of Allergenicity and Antigenicity of
Proteins</font>
</td>

    <td width=33% style="vertical-align: top;" >
        <div style="font-size:small; float: right; vertical-align: top; "><%=Util.format(new Date(),
"MM/dd/yyyy")%></div><br>
    </td>
</tr>
</table>

<script>

```

```

function validarlogin(){

    $.ajax({
    type: "POST",
    url: document.getElementById("context_path").value + "/validarlogin",
    data: $("#formlogin").serialize(),
    cache: true,
    success: function(response) {
        var ret = response.split("|");
        msg = ret[1];
        document.getElementById("msg-dialog-alert").innerHTML = msg;

        if (ret[0] == 0){
            $("#dialog-alert").dialog("open");
            return;
        }

        if (ret[0] == 1){
            location.reload();
            return;
        }

    }
    });

}

$(function() {$( "#dialog-alert" ).dialog({autoOpen: false, resizable: false, height:220, modal: true,
buttons: {"Ok": function() {$( this ).dialog( "close" );}} }); });

$(function() { $( "#dialog-confirm" ).dialog({autoOpen: false, resizable: false, height:200, modal: true
});});

$(function() {$( "#dialog-login" ).dialog({autoOpen: false, resizable: false, height:220, width:270,
modal: true,
    buttons: {
        "Cancel": function() {
            $( this ).dialog( "close" );
        },
        "Login": function() {
            validarlogin();
        }
    }
});
});

function inicializacontroles(){
    $(".beditar").button({icons: {primary: "ui-icon-pencil"}, text: false});
    $(".bexcluir").button({icons: {primary: "ui-icon-trash"}, text: false});
    $(".bsalvar").button({icons: {primary: "ui-icon-disk"}, text: false});
    $(".bseleccionar").button({icons: {primary: "ui-icon-check"}, text: false});
    $(".bimprimir").button({icons: {primary: "ui-icon-print"}, text: false});
    $(".bclose").button({icons: {primary: "ui-icon-closethick"}, text: false});
    $(".blink").button({icons: {primary: "ui-icon-link"}, text: false});
    $("input[type=submit], button").button().click(function( event ) {event.preventDefault();});
    $("#dialog-alert").dialog({autoOpen: false, resizable: false, height:220, modal: true,
buttons: {"Ok": function() {$( this ).dialog( "close" );}} });
}

```

```
$( "#dialog-confirm" ).dialog({autoOpen: false,  resizable: false, height:200, modal: true
});
}

$('body')
.on(
    'keydown',
    'input, select',
    function(e) {
        var self = $(this), form = self.parents('form:eq(0)'), focusable, next;
        if (e.keyCode == 13) {
            focusable = form.find(
                'input,a,select,button,textarea').filter(
                ':visible');
            next = focusable.eq(focusable.index(this) + 1);
            if (next.length) {
                next.focus();
            } else {
            }
            return false;
        }
    });

</script>

<input type="hidden" id="context_path" value="<%=request.getContextPath()%" >
```

### 4.13 Códigos-fonte referentes ao menu das seções da PA<sup>3</sup>P

```

<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<%-- <% @page import="org.dotto.plurifotoestudio.classe.Usuario"%> --%>

<% @ include file="/cabecalho.jsp" %>

<div class="cssmenu" >
  <ul id="nav" >
    <li><a href="<%=request.getContextPath()%>/index.jsp">Home</a></li>
    <li><a href="<%=request.getContextPath()%>/about/about.jsp">About</a></li>
    <li><a href="<%=request.getContextPath()%>/pa3p/pa3p.jsp">PA3P</a></li>
    <li><a href="<%=request.getContextPath()%>/database/database.jsp">Dataset</a></li>
    <li><a href="<%=request.getContextPath()%>/usefull_links.jsp">Useful Links</a></li>
    <li><a href="<%=request.getContextPath()%>/contact.jsp">Contact</a></li>

<%--          <li><a href="<%=request.getContextPath()%>/configuracao/estudio.jsp">Dados do
Estudio</a></li> --%>
  </ul>
</div>
<script >

  $( "#nav" ).menu({position: {at: "left bottom" }});

</script>

<!-- MENSAGEM DE ALERTA -->
<div id="dialog-alert" title="Atenção" style="display: none;">
  <p><span class="ui-icon ui-icon-alert" style="float: left; margin: 0 7px 20px 0; "><div id="msg-dialog-
alert">.</div></span></p>
</div>

<!-- MENSAGEM DE CONFIRMAÇÃO -->
<div id="dialog-confirm" title="Atenção" style="display: none;">
  <p><span class="ui-icon ui-icon-alert" style="float: left; margin: 0 7px 20px 0; "></span><div id="msg-
confirm-alert">Confirma o cancelamento da edição do cliente?</div></p>
</div>

<!-- MENSAGEM DE LOGIN -->
<div id="dialog-login" title="Login" style="display: none;">
  <form action=# id="formlogin" name="formlogin">
    <label style="width: auto;">User:</label><br><input type="text" name="nm_usuario"
id='nm_usuario' style="width: 230px;"><br>
    <label style="width: auto;">Password:</label><br><input type="password"
name="ds_senha" style="width: 230px;"><br>
  </form>
</div>

```

#### 4.14 Códigos-fonte referentes aos resultados da ADFS impressos na PA<sup>3</sup>P

```

<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<%-- <% @page import="org.dotto.plurifotoestudio.classe.Usuario"%> --%>

<% @ include file="/cabecalho.jsp" %>

<div class="cssmenu" >
  <ul id="nav" >
    <li><a href="<%=request.getContextPath()%>/index.jsp">Home</a></li>
    <li><a href="<%=request.getContextPath()%>/about/about.jsp">About</a></li>
    <li><a href="<%=request.getContextPath()%>/pa3p/pa3p.jsp">PA3P</a></li>
    <li><a href="<%=request.getContextPath()%>/database/database.jsp">Dataset</a></li>
    <li><a href="<%=request.getContextPath()%>/usefull_links.jsp">Useful Links</a></li>
    <li><a href="<%=request.getContextPath()%>/contact.jsp">Contact</a></li>

<%--
    <li><a href="<%=request.getContextPath()%>/configuracao/estudio.jsp">Dados do
Estudio</a></li> --%>
  </ul>
</div>
<script >

  $( "#nav" ).menu({ position: { at: "left bottom" } });

</script>

<!-- MENSAGEM DE ALERTA -->
<div id="dialog-alert" title="Atenção" style="display: none;">
  <p><span class="ui-icon ui-icon-alert" style="float: left; margin: 0 7px 20px 0; "><div id="msg-dialog-
alert">.</div></span></p>
</div>

<!-- MENSAGEM DE CONFIRMAÇÃO -->
<div id="dialog-confirm" title="Atenção" style="display: none;">
  <p><span class="ui-icon ui-icon-alert" style="float: left; margin: 0 7px 20px 0; "></span><div id="msg-
confirm-alert">Confirma o cancelamento da edição do cliente?</div></p>
</div>

<!-- MENSAGEM DE LOGIN -->
<div id="dialog-login" title="Login" style="display: none;">
  <form action=# id="formlogin" name="formlogin">
    <label style="width: auto;">User:</label><br><input type="text" name="nm_usuario"
id='nm_usuario' style="width: 230px;"><br>
    <label style="width: auto;">Password:</label><br><input type="password"
name="ds_senha" style="width: 230px;"><br>
  </form>
</div>

```

#### 4.15 Códigos-fonte referentes aos resultados do AlgPred impressos na PA<sup>3</sup>P

```

<% @page import="org.unipampa.lpa.classe.Algpred"%>

<%
    Algpred algpred = new Algpred();
    algpred.setSequence((String)request.getParameter("ds_sequence"));

    String[] result = {"", "", "", "", ""};
    double[] peso = {0,0,0,0,0};
%>

<div id="accordionalgpred">
    <h3>Mapping of IgE epitopes and PID</h3>
    <div>
        <%
            if (algpred.execute(1) == -1){
                peso[0] = algpred.getPeso(0);
                result[0]="Error";
                out.print("An error occurred during script execution. AlgPred
webservice could be offline.");
            }
            else{
                out.print(algpred.getHtmlResult());
                peso[0] = algpred.getPeso(1);
                result[0]=algpred.getResultado();
            }
        %>
    </div>

    <h3>SVM module based on dipeptide composition</h3>
    <div>
        <%
            if (algpred.execute(4) == -1){
                out.print("Ocorreu um erro ao efetuar a consulta");
            }
            else{
                out.print(algpred.getHtmlResult());
                peso[3] = algpred.getPeso(4);
                result[3]=algpred.getResultado();
            }
        %>
    </div>

    <h3>Blast search on allergen representative peptides (ARPs)</h3>
    <div>
        <%
            if (algpred.execute(5) == -1){
                out.print("Ocorreu um erro ao efetuar a consulta");
            }
            else{
                out.print(algpred.getHtmlResult());
                peso[4] = algpred.getPeso(5);
                result[4]=algpred.getResultado();
            }
        %>
    </div>
</div>

<form name="algpredresultform" id="algpredresultform" style="display: none;">

<%

```

```
for (int i=0;i<5;i++){
    out.print("<input type='hidden' name='v1algpred[]' value='" + peso[i] + "'>");
    out.print("<input type='hidden' name='resultalgpred[]' value='" + result[i] + "'>");
}
%>

</form>
```

#### 4.16 Códigos-fonte referentes aos resultados do Allergen Online impressos na PA<sup>3</sup>P

```

<% @page import="org.unipampa.lpa.classe.Algpred"%>

<%
    Algpred algpred = new Algpred();
    algpred.setSequence((String)request.getParameter("ds_sequence"));

    String[] result = {"", "", "", "", ""};
    double[] peso = {0,0,0,0,0};
%>

<div id="accordionalgpred">
    <h3>Mapping of IgE epitopes and PID</h3>
    <div>
        <%
            if (algpred.execute(1) == -1){
                peso[0] = algpred.getPeso(0);
                result[0]="Error";
                out.print("An error occurred during script execution. AlgPred
webservice could be offline.");
            }
            else{
                out.print(algpred.getHtmlResult());
                peso[0] = algpred.getPeso(1);
                result[0]=algpred.getResultado();
            }
        %>
    </div>

    <h3>SVM module based on dipeptide composition</h3>
    <div>
        <%
            if (algpred.execute(4) == -1){
                out.print("Ocorreu um erro ao efetuar a consulta");
            }
            else{
                out.print(algpred.getHtmlResult());
                peso[3] = algpred.getPeso(4);
                result[3]=algpred.getResultado();
            }
        %>
    </div>

    <h3>Blast search on allergen representative peptides (ARPs)</h3>
    <div>
        <%
            if (algpred.execute(5) == -1){
                out.print("Ocorreu um erro ao efetuar a consulta");
            }
            else{
                out.print(algpred.getHtmlResult());
                peso[4] = algpred.getPeso(5);
                result[4]=algpred.getResultado();
            }
        %>
    </div>
</div>

<form name="algpredresultform" id="algpredresultform" style="display: none;">

<%

```

```
for (int i=0;i<5;i++){
    out.print("<input type='hidden' name='vlalgpred[]' value='" + peso[i] + "'>");
    out.print("<input type='hidden' name='resultalgpred[]' value='" + result[i] + "'>");
}
%>

</form>
```

#### 4.17 Códigos-fonte referentes aos gráficos impressos pela PA<sup>3</sup>P

```

$(function () {
  $('#container').highcharts({
    chart: {
      type: 'column'
    },
    legend: {
      enabled: false
    },
    title: {
      text: 'Summary Result'
    },
    xAxis: {
      categories: ['Allergen', 'Non-Allergen']
    },
    yAxis: {
      min: 0,
      title: {
        text: '(%)'
      },
      stackLabels: {
        enabled: true,
        style: {
          fontWeight: 'bold',
          color: (Highcharts.theme && Highcharts.theme.textColor) || 'gray'
        }
      }
    },
    plotOptions: {
      column: {
        stacking: 'normal',
        dataLabels: {
          enabled: false,
          color: (Highcharts.theme && Highcharts.theme.dataLabelsColor) || 'white',
          style: {
            textShadow: '0 0 3px black'
          }
        }
      }
    },
    series: [{
      data: [0, 0]
    }]
  });

```

//SIMILARIDADE

```

$('#containersim').highcharts({
  chart: {
    type: 'column'
  },
  legend: {
    enabled: false
  },
  title: {
    text: 'Similarity Based Result'
  },
  xAxis: {
    categories: ['Allergen', 'Non-Allergen']
  },

```

```

yAxis: {
  min: 0,
  title: {
    text: '(%)'
  },
  stackLabels: {
    enabled: true,
    style: {
      fontWeight: 'bold',
      color: (Highcharts.theme && Highcharts.theme.textColor) || 'gray'
    }
  }
},

plotOptions: {
  column: {
    stacking: 'normal',
    dataLabels: {
      enabled: false,
      color: (Highcharts.theme && Highcharts.theme.dataLabelsColor) || 'white',
      style: {
        textShadow: '0 0 3px black'
      }
    }
  }
},
series: [{
  data: [0, 0]
}]
});

```

//PREDIÇÃO

```

$('#containerpred').highcharts({
  chart: {
    type: 'column'
  },
  legend: {
    enabled: false
  },
  title: {
    text: 'Algorithm Based Result'
  },
  xAxis: {
    categories: ['Allergen', 'Non-Allergen']
  },
  yAxis: {
    min: 0,
    title: {
      text: '(%)'
    },
    stackLabels: {
      enabled: true,
      style: {
        fontWeight: 'bold',
        color: (Highcharts.theme && Highcharts.theme.textColor) || 'gray'
      }
    }
  }
},

```

```
plotOptions: {
  column: {
    stacking: 'normal',
    dataLabels: {
      enabled: false,
      color: (Highcharts.theme && Highcharts.theme.dataLabelsColor) || 'white',
      style: {
        textShadow: '0 0 3px black'
      }
    }
  }
},
series: [{
  data: [0, 0]
}]
});

});
```

#### 4.18 Códigos-fonte referentes à impressão dos resultados na PA<sup>3</sup>P

```

<% @page import="org.unipampa.lpa.classe.Allermatch"%>

<% @ page language="java" contentType="text/html; charset=UTF-8" pageEncoding="UTF-8"%>

<% @ include file="/menu.jsp" %>

<!DOCTYPE html PUBLIC "-//W3C//DTD HTML 4.01 Transitional//EN"
"http://www.w3.org/TR/html4/loose.dtd">
<html>

<head>
    <meta http-equiv="Content-type" content="text/html; charset=UTF-8">
    <meta HTTP-EQUIV="Pragma" CONTENT="no-cache" />
    <meta HTTP-EQUIV="Cache-control" CONTENT="no-cache" />
    <meta HTTP-EQUIV="Expires" CONTENT="0" />

    <script type="text/javascript" src="<%=request.getContextPath()%>/js/mask.js" charset="utf-
8"></script>
    <script type="text/javascript" src="<%=request.getContextPath()%>/pa3p/grafico.js" charset="utf-
8"></script>
</head>

</head>

<script type="text/javascript">
var total = 0;
var totalalergenico = 0.00;
var totalnaoalergenico = 0.00;
var ct = 0;
var ctpred = 0;
var ctsim = 0;
var qtproc = 0;

var totalsim = 0, totalpred = 0;
var totalalergenicosisim = 0.00, totalalergenicopred = 0.00;
var totalnaoalergenicosisim = 0.00, totalnaoalergenicopred = 0.00;

var WAITGIF = '<div style="text-align: center;margin: 0 auto;"></div>';

function newquery(){
    $("#tabs").tabs("option", "active", 0 );
    $("#tabs").tabs("option", "disabled", [ 1, 2, 3, 4 ] );

    document.getElementById("allermatch").innerHTML = "";

    document.getElementById("resultadoalergenico").innerHTML = " ";
    document.getElementById("resultadonaalergenico").innerHTML = " ";

    document.getElementById("resultadoalergenicosisim").innerHTML = " ";
    document.getElementById("resultadonaalergenicosisim").innerHTML = " ";

    document.getElementById("resultadoalergenicopred").innerHTML = " ";
    document.getElementById("resultadonaalergenicopred").innerHTML = " ";

    for (var i=0; i<2;i++){

        document.getElementById("idvlallergenonline" + (i)).innerHTML = " ";
        document.getElementById("idresultallergenonline" + (i)).innerHTML = " ";
    }
}

```

```

}

for (var i=0; i<5;i++){
    if (i != 1 && i !=2){
        document.getElementById("idvialgpred" + (i + 1)).innerHTML = "";
        document.getElementById("idresultalgpred" + (i + 1)).innerHTML = "";
    }
}

document.getElementById("idvlproap0").innerHTML = "";
document.getElementById("idresultproap0").innerHTML = "";

document.getElementById("idvladfs0").innerHTML = "";
document.getElementById("idresultadfs0").innerHTML = "";

$('#container').highcharts().series[0].setData([0,0]);
$('#containersim').highcharts().series[0].setData([0,0]);
$('#containerpred').highcharts().series[0].setData([0,0]);

document.forms['formpesq']['ds_sequence'].value="";
document.forms['formpesq']['ds_sequence'].focus();
}

$(function() {
    $("#tabs").tabs();
    // $("#tabs").tabs("option", "disabled", [ 1, 2, 3 ] );
    inicializacontroles();
    newquery();
});

function run(){

    totalalergenico = 0.00;
    totalnaoalergenico = 0.00;
    qtproc = 0;
    ct = 0;
    ctsim = 0;
    ctpred = 0;

    totalsim = 0;
    totalpred = 0;
    totalalergenicosisim = 0.00;
    totalalergenicopred = 0.00;
    totalnaoalergenicosisim = 0.00;
    totalnaoalergenicopred = 0.00;

    $("#tabs").tabs("option", "active", 0 );

    $("#tabs").tabs("option", "disabled", [ 1, 2, 3, 4 ] );

    document.getElementById("resultadoalergenico").innerHTML = WAITGIF; ;
    document.getElementById("resultadonaoalergenico").innerHTML = WAITGIF; ;

    document.getElementById("resultadoalergenicosisim").innerHTML = WAITGIF; ;
    document.getElementById("resultadonaoalergenicosisim").innerHTML = WAITGIF; ;

    document.getElementById("resultadoalergenicopred").innerHTML = WAITGIF; ;
    document.getElementById("resultadonaoalergenicopred").innerHTML = WAITGIF; ;
}

```

```

for (var i=0; i<2;i++){
document.getElementById("idvlallergenonline" + (i)).innerHTML = WAITGIF;
document.getElementById("idresultallergenonline" + (i)).innerHTML = WAITGIF;
}

for (var i=0; i<5;i++){
    if (i != 1 && i !=2){
        document.getElementById("idvlgpred" + (i + 1)).innerHTML = WAITGIF;
        document.getElementById("idresultalgpred" + (i + 1)).innerHTML = WAITGIF;
    }
}

document.getElementById("idvlproap0").innerHTML = WAITGIF;
document.getElementById("idresultproap0").innerHTML = WAITGIF;

document.getElementById("idvldfs0").innerHTML = WAITGIF;
document.getElementById("idresultadfs0").innerHTML = WAITGIF;

$('#container').highcharts().series[0].setData([0,0]);
$('#containersim').highcharts().series[0].setData([0,0]);
$('#containerpred').highcharts().series[0].setData([0,0]);

//ALGPRED - INICIO
$.ajax({
    type: "POST",
    url: "algpredresult.jsp",
    data: $('#formpesq').serialize(),
    cache: true,
    success: function(response) {
        document.getElementById("algpred").innerHTML = response;

        $("#accordionalgpred" ).accordion({collapsible: true, heightStyle: "content", active: false });
        $("#tabs" ).tabs( "enable", 1 );

        var vlininput = document.getElementById('algpredresultform').elements["vlgpred[]"];
        var vlresult = document.getElementById('algpredresultform').elements["resultalgpred[]"];

        for (i = 0; i < vlininput.length; i++) {

            /*
                0 - Algpred - IgE - Similarity
                3 - Algpred - SVM Dipeptide - Predicted
                4 - Algpred - ARP - Similarity
            */
            if (i != 1 && i != 2){

                if (i == 0 || i == 4)
                    document.getElementById("idvlgpred" + (i + 1)).innerHTML = "Similarity
(" + vlininput[i].value + "%)";
                else
                    document.getElementById("idvlgpred" + (i + 1)).innerHTML = "Algorithm
(" + vlininput[i].value + "%)";

                document.getElementById("idresultalgpred" + (i + 1)).innerHTML =
vlresult[i].value;

                if (vlresult[i].value == "Allergen"){
                    ct = ct + 1;
                    totalalergenico = parseFloat(totalalergenico) + parseFloat(vlininput[i].value);

```

```

        if (i == 0 || i == 4){
            ctsim = ctsim + 1;
            totalalergenicosisim = parseFloat(totalalergenicosisim) +
parseFloat(vlininput[i].value);
        }
        else{
            ctpred = ctpred + 1;
            totalalergenicopred = parseFloat(totalalergenicopred) +
parseFloat(vlininput[i].value);
        }
    }
    else if (vlresult[i].value == "Non-Allergen"){
        ct = ct+ 1;
        totalnaoalergenico = parseFloat(totalnaoalergenico) +
parseFloat(vlininput[i].value);
    }
    if (i == 0 || i == 4){
        ctsim = ctsim + 1;
        totalnaoalergenicosisim = parseFloat(totalnaoalergenicosisim) +
parseFloat(vlininput[i].value);
    }
    else{
        ctpred = ctpred + 1;
        totalnaoalergenicopred = parseFloat(totalnaoalergenicopred) +
parseFloat(vlininput[i].value);
    }
}
}

qtproc = qtproc + 1;
finaliza();
}
});
//Fim Algpred

//PROAP - INICIO
$.ajax({
    type: "POST",
    url: "proapresult.jsp",
    data: $('#formpesq').serialize(),
    cache: true,
    success: function(response) {
        document.getElementById("proap").innerHTML = response;

        $('#tabs').tabs( "enable", 2 );

        var vlininput = document.forms['proapresultform']['vlproap'].value;
        var vlresult = document.forms['proapresultform']['resultproap'].value;

        document.getElementById("idvlproap0").innerHTML = "Algorithm (" + vlininput + "%)";
        document.getElementById("idresultproap0").innerHTML = vlresult;

        if (vlresult == "Allergen"){
            ct = ct+ 1;
            totalalergenico = parseFloat(totalalergenico) + parseFloat(vlininput);
            ctpred = ctpred + 1;

```

```

                totalalergenicopred = parseFloat(totalalergenicopred) + parseFloat(vlinput);
            }
            else if (vlresult == "Non-Allergen"){
                ct = ct+ 1;
                totalnaoalergenico = parseFloat(totalnaoalergenico) + parseFloat(vlinput);
                ctpred = ctpred + 1;
                totalnaoalergenicopred = parseFloat(totalnaoalergenicopred) + parseFloat(vlinput);
            }

            qtproc = qtproc + 1;
            finaliza();

        }
    });
//Fim proap

//ADFS - INICIO
$.ajax({
    type: "POST",
    url: "adsresult.jsp",
    data: $('#formpesq').serialize(),
    cache: true,
    beforeSend: function(){
        //document.getElementById("algpred").innerHTML = '<div style="text-align: center;margin: 0
auto;padding-top: 50px;padding-bottom: 50px;"></div>';
    },
    success: function(response) {
        document.getElementById("ads").innerHTML = response;

        $('#tabs').tabs("enable", 3);

        var vlinput = document.forms['adsresultform']['vladsf'].value;
        var vlresult = document.forms['adsresultform']['resultadsf'].value;

        document.getElementById("idvladsf0").innerHTML = "Similarity (" + vlinput + "%)";
        document.getElementById("idresultadsf0").innerHTML = vlresult;

        if (vlresult == "Allergen"){
            ct = ct+ 1;
            totalalergenico = parseFloat(totalalergenico) + parseFloat(vlinput);
            ctsim = ctsim + 1;
            totalalergenicosisim = parseFloat(totalalergenicosisim) + parseFloat(vlinput);
        }
        else if (vlresult == "Non-Allergen"){
            ct = ct+ 1;
            totalnaoalergenico = parseFloat(totalnaoalergenico) + parseFloat(vlinput);
            ctsim = ctsim + 1;
            totalnaoalergenicosisim = parseFloat(totalnaoalergenicosisim) + parseFloat(vlinput);
        }

        qtproc = qtproc + 1;
        finaliza();
    }
}); //Fim ADFS

//Allergen Online - INICIO
$.ajax({
    type: "POST",

```

```

url: "allergenonlineresult.jsp",
data: $('#formpesq').serialize(),
cache: true,
beforeSend: function(){
    //document.getElementById("algpred").innerHTML = '<div style="text-align: center;margin: 0
auto;padding-top: 50px;padding-bottom: 50px;"></div>';

    },
success: function(response) {
    document.getElementById("allergenonline").innerHTML = response;

    $('#accordionallergenonline').accordion({collapsible: true, heightStyle: "content", active: false
});

    $('#tabs').tabs("enable", 4);

    var vlinput =
document.getElementById('allergenonlineresultform').elements["vlallergenonline[]"];
    var vlresult =
document.getElementById('allergenonlineresultform').elements["resultallergenonline[]"];

    for (i = 0; i < 2; i++) {

        /*
            0 - allergenonline - Sliding 80mer Window
            1 - allergenonline - 8mer Exact Match
        */
        document.getElementById("idvlallergenonline" + (i)).innerHTML = "Similarity (" +
vlinput[i].value + "%)";
        document.getElementById("idresultallergenonline" + (i)).innerHTML = vlresult[i].value;

        if (vlresult[i].value == "Allergen"){
            ct = ct + 1;
            totalalergenico = parseFloat(totalalergenico) + parseFloat(vlinput[i].value);
            ctsim = ctsim + 1;
            totalalergenicosisim = parseFloat(totalalergenicosisim) +
parseFloat(vlinput[i].value);
        }
        else if (vlresult[i].value == "Non-Allergen"){
            ct = ct+ 1;
            totalnaoalergenico = parseFloat(totalnaoalergenico) +
parseFloat(vlinput[i].value);
            ctsim = ctsim + 1;
            totalnaoalergenicosisim = parseFloat(totalnaoalergenicosisim) +
parseFloat(vlinput[i].value);
        }

    }

    qtproc = qtproc + 1;
    finaliza();

}
}); //Fim AllergenOnline

}

function finaliza(){

    //Só finaliza se rodaram todos os algoritmos
    if (qtproc < 4) return;

```

```

        if (ct > 0 && totalalergenico > 0)
            document.getElementById("resultadoalergenico").innerHTML = (totalalergenico /
ct).toFixed(2);
        else
            document.getElementById("resultadoalergenico").innerHTML = "-";

        if (ct > 0 && totalnaoalergenico > 0)
            document.getElementById("resultadonaoalergenico").innerHTML = (totalnaoalergenico /
ct).toFixed(2);
        else
            document.getElementById("resultadonaoalergenico").innerHTML = "-";

        $('#container').highcharts().series[0].setData([Number((totalalergenico / ct).toFixed(2)),
Number((totalnaoalergenico / ct).toFixed(2))]);

        //SIMILARIDADE
        if (ctsim > 0 && totalalergenicosim > 0)
            document.getElementById("resultadoalergenicosim").innerHTML = (totalalergenicosim /
ctsim).toFixed(2);
        else
            document.getElementById("resultadoalergenicosim").innerHTML = "-";

        if (ctsim > 0 && totalnaoalergenicosim > 0)
            document.getElementById("resultadonaoalergenicosim").innerHTML =
(totalnaoalergenicosim / ctsim).toFixed(2);
        else
            document.getElementById("resultadonaoalergenicosim").innerHTML = "-";

        $('#containersim').highcharts().series[0].setData([Number((totalalergenicosim / ctsim).toFixed(2)),
Number((totalnaoalergenicosim / ctsim).toFixed(2))]);

        //PREDIÇÃO
        if (ctpred > 0 && totalalergenicipred > 0)
            document.getElementById("resultadoalergenicipred").innerHTML = (totalalergenicipred /
ctpred).toFixed(2);
        else
            document.getElementById("resultadoalergenicipred").innerHTML = "-";

        if (ctsim > 0 && totalnaoalergenicipred > 0)
            document.getElementById("resultadonaoalergenicipred").innerHTML =
(totalnaoalergenicipred / ctpred).toFixed(2);
        else
            document.getElementById("resultadonaoalergenicipred").innerHTML = "-";

        $('#containerpred').highcharts().series[0].setData([Number((totalalergenicipred / ctpred).toFixed(2)),
Number((totalnaoalergenicipred / ctpred).toFixed(2))]);
    }
</script>

<body>

<h3 class="titulo-borda-janela borda-janela" style="width: 800px; margin-bottom:5px;">PA³P Search</h3>
<div id="cabpesq" class="borda-janela" style="padding-top: 8px; padding-bottom: 8px; width: 800px;" >
    <fieldset>
        <form action="" name="formpesq" id="formpesq" >

            <label style="width: auto;">Sequence: </label>
            <br>
            <textarea name="ds_sequence" rows="5" style="width: 780px;"></textarea>

```

```

        <br>
        <button class="botao" id="bpesquisar" onClick="run()" style="float: right; margin-
left: 5px; margin-top: 10px;">Run</button>
        <button class="botao" id="blimpar" onClick="newquery()" style="float: right;
margin-top: 10px;">New Query</button>
    </form>
</fieldset>
</div>

<div id="tabs" class="borda-janela" style="margin-top: 10px; width: 1010px;">
<ul>
<li><a href="#pa3p">PA3P</a></li>
<li><a href="#algpred">Algpred</a></li>
<li><a href="#proap">Proap</a></li>
<li><a href="#ads">ADFS</a></li>
<li><a href="#allergenonline">Allergen Online</a></li>
</ul>

<div id="pa3p" class="borda-janela" style="margin-top: 10px; width: 1000px; margin: 0 auto;">

    <table style="margin: 0 auto; padding-top: 20px; padding-bottom: 20px;">

        <tr>
            <td style="vertical-align: super;">

                <table width="600px" style="margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
                    <thead>
                        <tr style="text-align: center;">
                            <th width="100%" scope="col"><b>Individual Results</b></th>
                        </tr>
                    </thead>
                </table>

                <table id="pa3presume" width="600px" style="margin-left: auto; margin-right: auto; text-
align: left;" class='tabela' >
                    <thead>
                        <tr>
                            <th width="40%" scope="col">WebTool</th>
                            <th width="30%" scope="col">Method</th>
                            <th width="30%" scope="col">Result</th>
                        </tr>
                    </thead>
                    <tbody>

                        <tr>
                            <td>Algpred - IgE</td>
                            <td id="idv1algpred1"></td>
                            <td id="idresultalgpred1"></td>
                        </tr>

                        <tr>
                            <td>Algpred - SVM Dipeptide</td>
                            <td id="idv1algpred4"></td>
                            <td id="idresultalgpred4"></td>
                        </tr>

                        <tr>
                            <td>Algpred - ARP</td>
                            <td id="idv1algpred5"></td>
                            <td id="idresultalgpred5"></td>
                    </tbody>
                </table>

```



```

                <td id="resultadonaoalergenicopred"></td>
            </tr>
        </tbody>
    </table>
    <br>
    <div id="containerpred" style="min-width: 310px; height: 300px; margin: 0
auto"></div>

</td>

<td style="vertical-align:super; padding-left: 0px;">

    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
        <thead>
            <tr style="text-align:center;">
                <th width="100%" scope="col"><b>Similarity Based
Result</b></th>
            </tr>
        </thead>
    </table>

    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
        <thead>
            <tr>
                <th width="60%" scope="col">Result</th>
                <th width="40%" scope="col">% (&#177; 0.87%)</th>
            </tr>
        </thead>
        <tbody>
            <tr>
                <td>Allergen</td>
                <td id="resultadoalergenicosisim"></td>
            </tr>
            <tr>
                <td>Non-Allergen</td>
                <td id="resultadonaoalergenicosisim"></td>
            </tr>
        </tbody>
    </table>
    <br>
    <div id="containersim" style="min-width: 310px; height: 300px; margin: 0
auto"></div>

</td>

<td style="vertical-align:super; padding-left: 0px;">

    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
class='tabela' >
        <thead>
            <tr style="text-align:center;">
                <th width="100%" scope="col"><b>Summary
Result</b></th>
            </tr>
        </thead>
    </table>

```

```

class='tabela' >
    <table width="330px" style=" margin-left: auto; margin-right: auto; text-align: left;"
        <thead>
            <tr>
                <th width="60%" scope="col">Result</th>
                <th width="40%" scope="col">% (&#177; 0.89%)</th>
            </tr>
        </thead>
        <tbody>
            <tr>
                <td>Allergen</td>
                <td id="resultadoalergenico"></td>
            </tr>
            <tr>
                <td>Non-Allergen</td>
                <td id="resultadonaoalergenico"></td>
            </tr>
        </tbody>
    </table>
    <br>
    <div id="container" style="min-width: 310px; height: 300px; margin: 0 auto"></div>
</td>
</tr>
</table>
</div>
<div id="allermatch" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<div id="algpred" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<div id="proap" class="borda-janela" style="margin-top: 10px;width: 990px; word-wrap: break-word;">
</div>
<div id="adfs" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<div id="allergenonline" class="borda-janela" style="margin-top: 10px;width: 990px;">
</div>
<!-- <div id="blastp" class="borda-janela" style="margin-top: 10px;width: 990px; text-align: left;"> -->
<!-- </div> -->
</div>
</body>
</html>

```

#### 4.19 Códigos-fonte referentes aos resultados do proAP impressos na PA<sup>3</sup>P

```

<% @page import="org.unipampa.lpa.classe.Proap"%>
<%
    Proap proap = new Proap();
    proap.setSequence((String)request.getParameter("ds_sequence"));

    String[] result = {"-"};
    double[] peso = {0};

    if (proap.execute(0) == -1){
        peso[0] = proap.getPeso(0);
        result[0]="Error";
        out.print("An error occurred during script execution. Proap webserver could be offline.");
    }
    else{
        out.print(proap.getHtmlResult());
        peso[0] = proap.getPeso(0);
        result[0]=proap.getResultado();
    }
%>

<form name="proapresultform" id="proapresultform" style="display: none;">

<%
    out.print("<input type='hidden' name='vlproap' value='" + peso[0] + "'>");
    out.print("<input type='hidden' name='resultproap' value='" + result[0] + "'>");
%>

</form>

```