

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA E FISIOLOGIA

***Análise e caracterização in silico de polimorfismos de base única
dos genes Toll Like Receptor: consequências estruturais e
funcionais associadas ao desenvolvimento do câncer***

Carolina da Rocha Simões

Orientador: Dr. José Luiz de Lima Filho

Coorientador: Dr. Carlos Henrique Madeiros Castelletti

RECIFE
FEVEREIRO 2014

UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOQUÍMICA E FISIOLOGIA

***Análise e caracterização in silico de polimorfismos de base única
dos genes Toll Like Receptor: consequências estruturais e
funcionais associadas ao desenvolvimento do câncer***

Carolina da Rocha Simões

Dissertação apresentada para o
cumprimento parcial das exigências
para obtenção do título de Mestre em
Bioquímica e Fisiologia pela
Universidade Federal de Pernambuco

Orientador: Dr. José Luiz de Lima Filho

Coorientador: Dr. Carlos Henrique Madeiros Castelletti

RECIFE
FEVEREIRO 2014

Catálogo na Fonte:
Bibliotecário Bruno Márcio Gouveia, CRB-4/1788

Simões, Carolina da Rocha

“Caracterização e análise de polimorfismos de base única nos genes das proteínas Tool Like Receptor e sua predisposição ao câncer” / Carolina da Rocha Simões. – Recife: O Autor, 2014.

73 folhas: il.

Orientador: José Luiz de Lima Filho, Carlos Henrique Madeiros Castelletti

Dissertação (mestrado) – Universidade Federal de Pernambuco. Centro de Ciências Biológicas. Programa de Pós-graduação em Bioquímica e Fisiologia, 2014.

Inclui bibliografia e anexos

1. Polimorfismo (Genética) 2. Câncer – Aspectos genéticos I. Lima Filho, José Luiz de (orient.) II. Castelletti, Carlos Henrique Madeiros (coorient.) III. Título.

576.54

CDD (22.ed.)

UFPE/CCB-2014-260

CAROLINA DA ROCHA SIMÕES

Análise e caracterização in silico de polimorfismos de base única dos genes Toll Like Receptor: consequências estruturais e funcionais associadas ao desenvolvimento do câncer

Dissertação apresentada para o cumprimento parcial das exigências para obtenção do título de Mestre em Bioquímica e Fisiologia pela Universidade Federal de Pernambuco

Aprovada em 20/02/2015 por:

José Luiz de Lima Filho
Universidade Federal de Pernambuco
Departamento de Bioquímica

Danyelly Bruneska Gondim Martins
Universidade Federal de Pernambuco
Departamento de Bioquímica

Maria da Paz carvalho da Silva
Universidade Federal de Pernambuco
Departamento de Bioquímica

Roberta Rodrigues de Lemos
Universidade Federal de Pernambuco
Pós-doutoranda do Laboratório de Imunopatologia Keizo Asami

Aos meus pais Marcelo e Cleonice Simões

“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar. Mas o mar seria menor se lhe faltasse uma gota”.

(Madre Teresa de Calcutá)

RESUMO

O aumento da informação proveniente do sequenciamento de alta performance (NGS), e de projetos como o 1000 genomas e HapMap, permitiram a descoberta de milhões de variações. Entretanto, o maior desafio é a identificação da relação entre o genótipo e o fenótipo, proporcionando informações que possam ajudar a definir os polimorfismos que podem ou não causar doenças. Ferramentas computacionais tem auxiliado na predição das modificações estruturais geradas pelos polimorfismos, e as consequentes alterações funcionais sofridas pelas proteínas. Os receptores Toll Like (TLR) são proteínas do sistema imunológico que estão envolvidas na regulação da inflamação e em alguns casos no desenvolvimento do câncer. O objetivo deste projeto foi analisar, através de ferramentas *in silico*, os polimorfismos de base única nos genes das TLRs, buscando por polimorfismos que possam estar relacionados com a predisposição ao câncer e com alterações da via de sinalização das TLRs. Foram encontrados 37 genes que estão envolvidos na via de sinalização e podem ser utilizados como marcadores genéticos (biomarcadores) para o diagnóstico e predição das alterações na expressão dos genes relacionados à esta via. Estes genes, se regulados, podem ser utilizados como inibidores. Em relação aos polimorfismos foram coletados no banco de dados dbSNP/NCBI 5.839 SNPs entre os 10 genes das TLRs. Destes, 1.017 variações foram classificadas como missense e analisadas para avaliar as consequências estruturais pela troca dos aminoácidos. Para isso quatro ferramentas preditoras (SIFT, Polyphen, MutationAssessor e SDM) foram utilizadas gerando informações sobre as modificações e associando-as com possíveis danos nas proteínas. Dos polimorfismos analisados 223 foram classificados como danosos baseados na troca de aminoácido e podem causar uma desregulação funcional na proteína. Entre eles está o rs5743708 (TLR2), rs3775291, (TLR3) e rs11466653 (TLR10) que já foi estudado *in vitro* e tiveram associação com câncer colorectal (TLR2 e 3) e carcinoma da tireóide (TLR10). A predição prévia, *in silico*, das alterações funcionais pode auxiliar na interpretação das variações gênicas, neste caso associadas com o câncer, e também na caracterização precisa dos fatores que levam a estas alterações, contribuindo no diagnóstico, na prevenção e em melhores respostas aos tratamentos oferecidos.

Palavras chaves: Toll Like Receptors, Câncer, Polimorfismo de base única

ABSTRACT

The increase of information from the next generation sequencing (NGS) and projects like 1000 genomes and HapMap permitted the discovery of millions variations. However, the major challenger is the identification of the relationship between phenotype and genotype, proportioning information that could help define polymorphisms that can or not cause disease. Computational tools have supported in the prediction of the structural modifications developed by polymorphisms, and the consequential functional alterations suffered by the protein. TLRs are proteins from the immune system and are involved in the regulation of inflammation and the development of cancer. The objective of this project was analyse through in silico tools the single nucleotide polymorphisms of TLR gene, searching for polymorphisms that can be related with cancer and the changes of the signaling pathway of TLRs. Were founded 37 genes that are involved in the signaling pathway and can be used as genetic markers (biomarkers) to the diagnosis and prediction of alterations in the expression of the genes related with the pathway. These genes if regulated can be used as inhibitors. In relation of polymorphisms were collected in the database dbSNP/NCBI 5.839 SNPs between the 10 genes of TLRs. From that, 1.017 variations were classified as missense and analysed to evaluate the structural consequences of aminoacid changes. For this, four prediction tools (SIFT, Polyphen, MutationAssessor e SDM) were utilized generating information about the modifications and associating them with the possibility to damage the proteins. From polymorphisms, analyzed 223 were classified as damage based in the amino acid change and can cause a functional downregulation in the protein. Between that, are the rs5743708 (TLR2), rs3775291, (TLR3) e rs11466653 (TLR10) that were studied in vitro and had an association with colorectal cancer (TLR2 e 3) and papillary thyroid carcinoma (TLR 10). The previous prediction, in silico, of functional alterations could assist in the interpretation of genetic variation, in this case, associated with cancer, and either in the precision characterization of the factors that lead to these alterations, contributing in the diagnosis, prevention and better responses to the offered treatment.

Keys Woeds: Toll Like Receptors, Cancer, Single Nucleotide Polimorphism.

LISTA DE FIGURAS

Figura 1: Tipos de variações existentes no genoma	14
Figura 2: Polimorfismo de base única (SNP)	15
Figura 3: Esquema estrutural da TLR	18
Figura 4: Esquema do reconhecimento de patógenos e localização das TLRs 1, 2, 4, 5 e 6	19
Figura 5: Esquema do reconhecimento de patógenos e localização das TLRs 3, 7 e 9	20
Figura 6: Esquema da via de sinalização das TLRs	21
Figura 7: Interação de disciplinas que tem contribuído para o desenvolvimento da bioinformática	24

LISTA DE ABREVIATURAS

DAMPs - Damage Associated Molecular Pattern Molecules

DD – Death Domain

HMGB-1 - Mobility Group Box-1

HSP - Heat Shock Protein

IFN – Interferons

IKK - I κ B Kinase

IRAK1 - IL-1 Receptor Associated Kinase

IRF3 - Interferon Regulating Factor3

JNK - c-Jun-N-Terminal Kinase

LBD – Proteína de Ligação do LPS

LPS – Lipopolissacarídeo

LRR – Leucine Rich Repeat

MAL – Myd88 Adaptor Like

MAPKs – Mitogen activated protein kinase

MyD88 - Myeloid Differentiation Antigen 88

NCBI – National Center of Biotechnology Information

NF κ B - Factor Nuclear Kappa B

NGS – Next Generation Sequencing

NLRs - Nucleotide Binding Oligomerization Domain like receptors

nsSNP – nonsynonymous SNP

PCR – Polymerase Chain Reaction

POLYPHEN - Polymorphism Phenotyphin

PRRs - Patter Recognition Receptors

RIP1 - Receptor Interfering Serine/Threonine Protein Kinase-1

RLRs - including retinoic acid-inducible gene-1 (RIG-1)-like receptors

RSV – Respiratory Syncytial Virus

SDM - Site Directed Mutator

SIFT - Sorting Intolerant from Tolerant

SNP - Single Nucleotide Polymorphism

TBK1 - Tank Binding Kinase 1

TIR – Toll IL-1 Receptor

TIRAP - TIR Domain Containing Protein

TLR - Toll Like Receptors

TRAF6 - TNF Receptor Associated Factor 6

TRIF – TIR Domain Containing adapter inducing interferon β

SUMÁRIO

Resumo	V
Abstract	VI
Lista de figuras	VII
Lista de abreviaturas	VIII
1. Introdução	11
2. Fundamentação teórica	13
2.1 Variações Genéticas	13
2.2 Polimorfismo de base única (Single Nucleotide Polymorphism - SNP)	14
2.3 Sistema Imunológico	16
2.3.1 Toll Like Receptor (TLR)	17
2.3.2 Toll Like Receptor e variações genéticas	22
2.4 Bioinformática e servidores computacionais	23
3. Objetivos	28
4. Referências	29
5. Artigo - Functional analysis of Single Nucleotide Polymorphisms in the genes of the Toll Like Receptors (TLR) family and the susceptibility to development of cancer	34
6. Patente – Painel genético na detecção de genes inibidores da via de sinalização das proteínas Toll Like Receptor (TLR) para o tratamento do câncer	55
7. Conclusões	71
ANEXO	73

1. INTRODUÇÃO

Com o término do projeto genoma humano em 2001 e com o desenvolvimento de novas tecnologias de alta capacidade é possível sequenciar um genoma humano a custo significativamente mais baixo, permitindo novos diagnósticos e tratamentos médicos. As plataformas de *Next Generation Sequencing* (NGS) identificaram uma grande quantidade de variações estruturais, permitindo uma associação entre estas e susceptibilidade a doenças. Projetos como o *The SNP Consortium*, *HapMap* e 1000 genomas possibilitaram comparar e caracterizar as variações presentes entre os genomas e associá-las com várias doenças humanas comuns e raras. Bancos de dados como o dbSNP do *National Center of Biotechnology Information* (NCBI) agregam informações acerca de pequenas variações que ocorrem no genoma, tais como SNP, inserção e deleção, repetições curtas e microsátélites.

Os polimorfismos de base única, SNP, do inglês *Single Nucleotide Polimorphisms*, são encontrados em 93% dos genes e são considerados a forma mais comum de variação genética encontrada no DNA humano. Uma abordagem típica para encontrar a causa de uma doença genética envolve a utilização de estudos realizados com marcadores genéticos para a determinação de regiões associadas com a doença. Essas regiões são extraídas de SNPs e um subconjunto destes são genotipados em uma amostra de uma população.

Estudos associativos identificaram várias regiões de genes relacionadas com a suscetibilidade de doenças. Uma vez que uma determinada região é identificada por abrigar um locus de risco, se faz necessária a realização de estudos detalhados de todas as variações genéticas no locus para descobrir a causa desta variação, para quantificar sua contribuição à suscetibilidade da doença e elucidar seus papéis nas vias funcionais.

Na tentativa de reduzir tempo e custo na detecção, tem sido usada a caracterização funcional das variações através de métodos computacionais, que predizem as alterações provocadas pelos SNPs e as consequências funcionais na proteína alvo. A acurada predição dos efeitos dos SNP não sinônimos (do inglês, nonsynonymous SNPs - nsSNPs) será uma ferramenta de grande valor para

analisar e selecionar os SNPs considerados como prejudiciais para serem reavaliados em estudos associados ao genoma.

O estudo dos SNPs nas TLRs pode fornecer informações importantes para o desenvolvimento de estratégias terapêuticas e fármacos especializados e mais efetivos para cada indivíduo, como também para cada tipo de câncer; criação de vacinas; assim como o esclarecimento dos mecanismos de patogenicidade envolvidos nos diversos tipos de câncer.

Informações sobre o impacto de nsSNPs na função e estrutura de proteínas através de ferramentas computacionais permite interpretar e selecionar aqueles que podem ou não estar envolvidos no desenvolvimento do câncer. Com o surgimento da “medicina precisa” o entendimento como a genética pode influenciar no progresso de doenças ajuda no aperfeiçoamento da elaboração de terapias apropriadas e efetivas direcionadas a mutações específicas.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Variações Genéticas

Com o desenvolvimento de novas tecnologias de genotipagem e redução no custo dos sequenciamentos cresceu exponencialmente a quantidade de informação sobre variações genéticas e genômicas. Em paralelo com esses desenvolvimentos, descobertas de genes que contribuem para doenças monogênicas e complexas tem avançado rapidamente, fazendo com que os atuais bancos de dados e softwares relacionados com coleções e análise de dados genéticos cresçam em número e tamanho (JOHNSON, 2009).

Os estudos de variações genéticas surgiram a partir do projeto genoma humano e vários consórcios mundiais identificaram mais do que dois milhões de variações no DNA na população humana (SUNYAEV et al., 2001). Estudos associados ao genoma estão sendo rotineiramente usados para identificar polimorfismos que são fundamentais para a suscetibilidade de doenças em uma população (TORKAMANI; SCHORK, 2008).

Um dos focos da genômica é determinar as variações genéticas existentes entre os indivíduos e entender suas relações com as diferenças fenotípicas. Vários tipos de variações genéticas tem sido identificadas em humanos, tais como o polimorfismo de base única (SNP), pequenas inserções e deleções (INDELs) que variam de 1pb a 10 kb, grandes variações estruturais variando de 10kb a diversas megabases, inversões e translocações (ZHANG et al., 2011). Elas surgem como consequência da formação de diferentes mecanismos incluindo a recombinação do DNA, replicação e processos de reparo (Figura 1) (WEISCHENFELDT et al., 2013). Cada indivíduo possui características únicas que são refletidas em seu genótipo. Cerca de 99.9% dos nucleotídeos são os mesmo para todas as pessoas, o restante 0.1%, é específico de cada indivíduo (TENG et al., 2009).

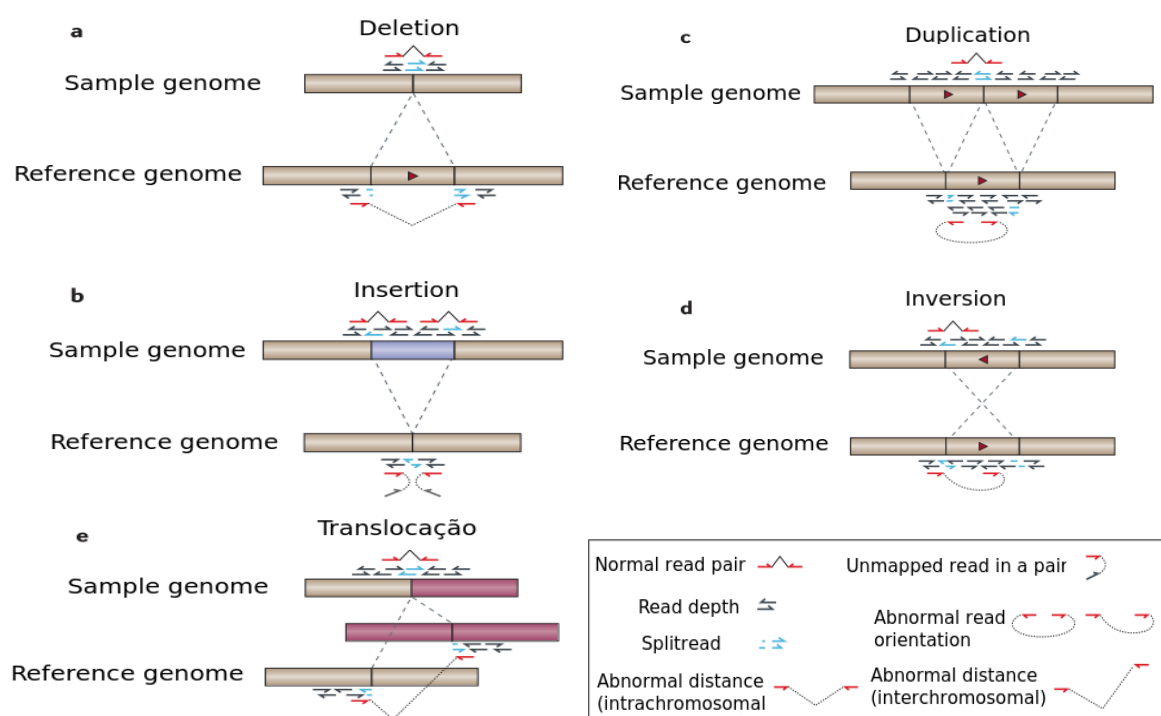


Figura 1 – Tipos de variações existentes no genoma. (a) deleção, (b) Inserção, (c) Duplicação, (d) Inversão e (e) Translocação. (CHMIELECKI; MEYERSON, 2014).

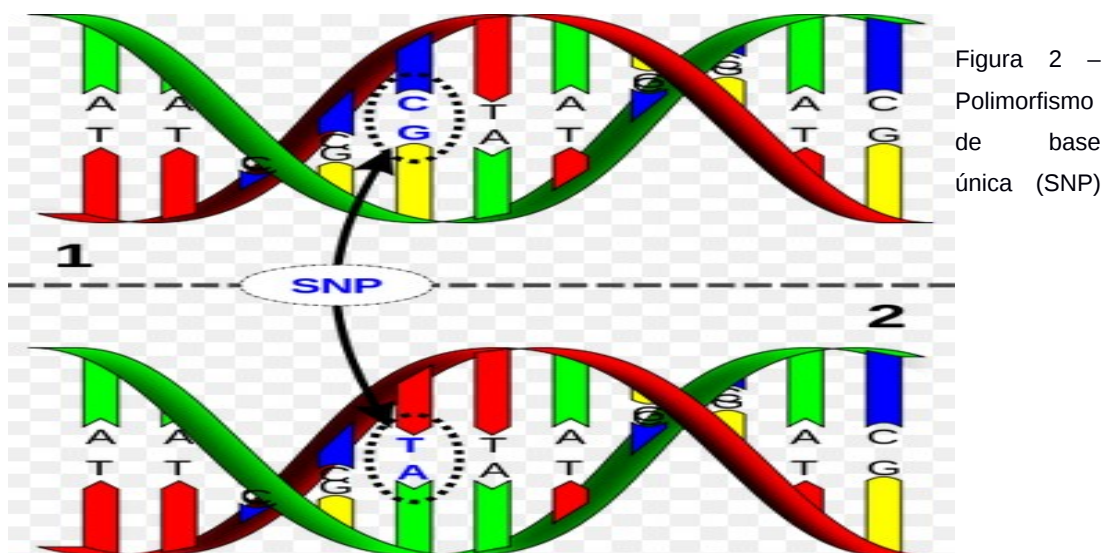
2.2 Polimorfismo de base única (SNP)

Os polimorfismos de base única (SNP do inglês, *Single Nucleotide Polymorphism*), são variantes alélicas do DNA que aparecem como resultado da troca de um nucleotídeo (RAMENSKY; BORK; SUNYAEV, 2002) (Figura 2). Para o genoma humano, composto por mais de três bilhões de pares de base, foi estimada a presença de cerca de dez milhões de SNPs que podem ocorrer aproximadamente a cada trezentos pares de base ao longo dos cromossomos com uma frequência alélica $\geq 1\%$ na população (GEORGE PRIYA DOSS et al., 2008).

Os SNPs são as formas mais simples e frequentes das variações da sequência de DNA e são usados no mapeamento de doenças genéticas complexas (LIU et al., 2011). O polimorfismo em um único nucleotídeo também é utilizado na análise farmacogenética, em genética de populações e em estudos de cunho evolucionário (ZHANG et al., 2005).

A mudança de pares de base individuais em regiões codificadoras de genes funcionais pode resultar na substituição do amino ácido (mutação missense), no truncamento do produto gênico (mutação sem sentido), em mudanças nos sítios de reconhecimento de splicing, ou não haver mudança alguma (mutação silenciosa).

Inserções e deleções podem alterar a função gênica dependendo se eles preservam ou rompem o frame de leitura. A proporção de mutações incluindo tanto variações sinônimas (silenciosas) e não sinônimas (não silenciosa) varia entre tumores de diferentes origens tão bem quanto em regiões (CHMIELECKI; MEYERSON, 2014).



http://en.wikipedia.org/wiki/Single-nucleotide_polymorphism.

Os SNP não sinônimos (do inglês nonsynonymous SNP, nsSNPs) podem ou não levar a alteração da função da proteína. Existem cerca de 24,000 – 40,000 de nsSNP em cada pessoa e um total de 67,000 – 200,000 nsSNPs comuns na população humana (TENG; SRIVASTAVA; WANG, 2010). Em geral, se ocorrem em códons conservados são mais susceptíveis a alterar a função da proteína e serem patogênicos do que os que ocorrem em códons pobremente conservados (CHAN P. A. et al, 2007). Os nsSNPs podem resultar em mais do que 50% das mutações conhecidas envolvidas nas doenças hereditárias tornando-se uma variante de grande interesse (SUO et al., 2013).

A classificação das variantes missense como benignas ou danosas é importante por duas razões. Primeiro, para identificar precisamente as variantes que aumentam a susceptibilidade ao desenvolvimento de doenças e fornece uma avaliação de risco apropriada, como também decisões terapêuticas. Segundo, por realizar pesquisas voltadas à funcionalidade de mutações consideradas importantes por alterarem a função da proteína (LAZARUS et al., 2002). Algumas mutações, SNPs, têm sido associadas com doenças, entretanto a correlação entre o genótipo e o fenótipo não é direta. Para algumas doenças, apenas uma parte das mutações prediz um fenótipo. Doenças hereditárias em humanos podem estar relacionadas

com nsSNPs, pois podem potencialmente afetar a estrutura ou função de proteínas expressas e conseqüentemente causar uma doença (MASOODI et al., 2012).

Mutações podem afetar as características da proteína de diferentes formas: dobramento da proteína, estabilidade, flexibilidade e agregação; altera sítios funcionais, reações cinéticas e dependência de parâmetros do ambiente, tais como pH, concentração de sal e temperatura; altera a expressão proteica e localização subcelular, e também rompe interações proteína-pequenas moléculas, proteína-proteína, proteína-DNA e proteína-membrana (TENG; MICHONNOVA-ALEXOVA; ALEXOV, 2008). As alterações que ocorrem nas proteínas são provocadas pelas diferenças bioquímicas, se o aminoácido modificado é ácido, básico ou hidrofóbico, e pela localização da variação se afeta a estrutura quaternária ou terciária ou o sítio ativo, conduzindo a efeitos deletérios na estrutura e/ou função (JOHNSON; HOUCK; CHEN, 2005).

Existem três tipos de mutação que contribuem com o progresso do câncer: ganho de função, que convertem genes normais em oncogênicos; perda de função que inativam supressores de tumor e as mutações resistentes a drogas que tem efeito inibitório de um fármaco em uma proteína alvo (REVA; ANTIPIN; SANDER, 2011).

Variações no sistema imunológico tem um importante papel no desenvolvimento de várias doenças, pois elas são hereditárias. Elas também são importantes por acionar e sustentar respostas inflamatórias e em providenciar respostas do sistema adaptativo e antígeno específico; como também podem afetar respostas das vias de sinalização que pode ser crítica para o hospedeiro na hora da defesa contra o patógeno (LAZARUS et al., 2003).

2.3 Sistema Imunológico

O sistema imunológico é dividido em duas partes, o sistema imune inato e o adaptativo. O sistema inato é geneticamente programado a detectar características invariantes dos micro-organismos invasores, sendo mediada por fagócitos tais como os macrófagos, células dendríticas e neutrófilos. Já o adaptativo, composto por linfócitos B e T, é caracterizado por sua especificidade por antígenos particulares e depende da seleção clonal de antígenos externos e produção de anticorpos gerados por um mecanismo conhecido como rearranjo gênico, é responsável por matar

patógenos que podem escapar do sistema imune inato (IWASAKI; MEDZHITOV, 2004). O sistema imunológico inato reconhece micro-organismos através de um número limitado de receptores de reconhecimento de padrão (do inglês pattern recognition receptors – PRRs), ao contrário do vasto repertório de receptores rearranjados utilizados pelo sistema adquirido (ABDELSADIK; TRAD, 2011).

Os PRRs possuem algumas características comuns que são o reconhecimento de componentes produzidos por organismos invasores, conhecido por padrões moleculares associados a patógenos (do inglês pathogen associated molecular patterns PAMPs), que são essenciais para a sobrevivência do micro-organismo e não alteradas por eles; são expressos constitutivamente pelo hospedeiro e protegem independentemente de estágio de ciclo de vida; como também independem de memória imunológica (AKIRA; UEMATSU; TAKEUCHI, 2006).

Várias famílias de PRRs tem sido caracterizadas, elucidando o mecanismo básico de infecção. Algumas pesquisas tem mostrado que as proteínas *including retinoic acid-inducible gene-1 (PIG-1)-like receptors* (RLRs) e *nucleotide-binding oligomerization domain (NOD)-like receptors* (NLRs) também desempenham um papel no reconhecimento de padrões (CHANG, 2010).

Os receptores Toll Like (TLRs) por possuírem características semelhantes, pertencem a uma das famílias PRRs e são responsáveis pela efetiva indução do sistema imunológico (TAKEUCHI; AKIRA, 2010).

2.3.1 Toll Like Receptor (TLR)

A proteína Toll, derivada de uma gíria alemã para “fantástico, foi inicialmente proveniente de um gene de *Drosophila* chamado Toll, que desempenhava um papel central no estabelecimento do padrão dorsoventral durante o desenvolvimento embrionário. Ao mesmo tempo que foi descoberto, observou-se que a sua via de sinalização estava relacionada a resistência à infecção de diversos peptídeos pela *Drosophila*. Foi encontrado que o ligante Spz Toll (Spätzle) controlava a expressão de genes antifúngicos em moscas adultas e que mutações poderiam reduzir drasticamente a sobrevivência desses insetos depois de uma infecção. Devido a semelhança com esse receptor ela foi chamada de Receptor Toll Like (TLR) (BASSETT, 2004).

O receptor Toll Like consiste em uma família de dez proteínas que alertam as células imunocompetentes a gerar uma resposta imunológica rápida e eficiente (KIM et al., 2013). As TLRs humanas são receptores de membrana integrais tipo I, com peso molecular variando de 90 a 115 kDa (BELL et al., 2005).

Cada TLR possui um domínio N terminal de reconhecimento, uma hélice simples e um domínio C terminal citoplasmático de sinalização. Os ectodomínios N terminais são glicoproteínas com 500-800 resíduos de aminoácidos (KIRK; BAZAN, 2005). Podem ser encontrados na região extracelular (TLRs 1, 2, 4, 5, and 6) ou em endossomos (TLRs 3, 7, 8, and 9) onde reconhecem e se ligam as moléculas dos patógenos invasores. É constituído por cópias repetidas de uma região conhecida como repetição rica em leucina (LRR, do inglês Leucine Rich Repeats), que é tipicamente composta por 20 a 29 resíduos de leucina (YANG et al., 2010). Os domínios de sinalização são conhecidos como Toll IL-1 (TIR) pois compartilham homologia com os domínios de sinalização da família IL-1R. O domínio transmembranar contém um trecho de aproximadamente 20 resíduos hidrofóbicos não carregados (O'NEILL; BOWIE, 2007) (Figura 3).

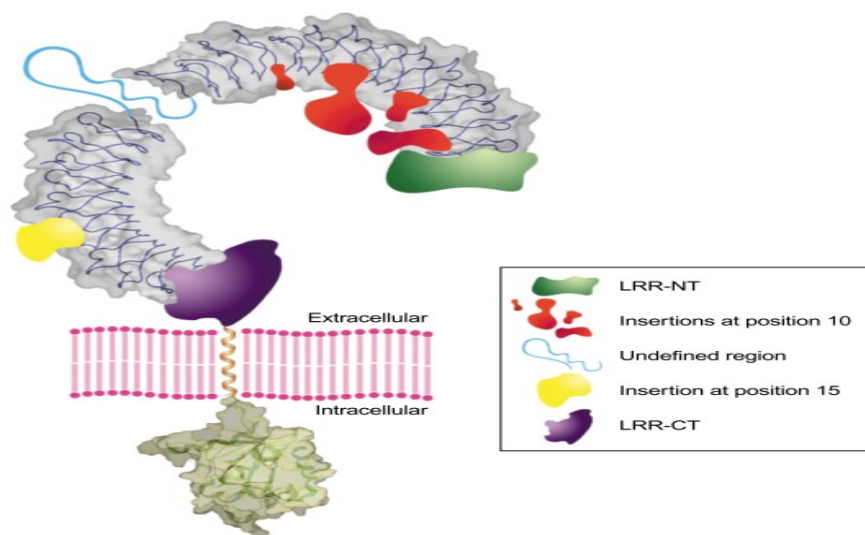


Figura 3 – Esquema estrutural da TLR. Em cinza a região de repetição de leucina, em amarelo a região transmembranar e cinza escuro o domínio de sinalização. (BELL et al., 2005).

A maioria das TLRs são expressas em células imunes tais como neutrófilos polinucleares, monócitos/macrófagos, células dendríticas, células natural killer e mastócitos (NIETERS et al., 2006). Podem também estar presentes em células B e T, células endoteliais, células epiteliais, queratinócitos, fibroblastos e células de

câncer. Isso evidencia que as TLRs regulam diretamente o sistema imune inato e adaptativo, reações de inflamação, doenças inflamatórias, câncer e outras respostas celulares (MEDVEDEV, 2013).

Cada TLR reconhece um padrão molecular associado a patógenos (do inglês *pathogen associated molecular patterns* PAMPs) diferente e a ligação destes desencadeia uma série de eventos bioquímicos. Elas são sensíveis a diversos PAMPs, tais como ácidos nucleicos, proteínas, lipídeos e polissacarídeos. Além disso, reconhecem moléculas autólogas, padrões moleculares associados a dano ou estresse (DAMPs, do inglês *Damage-Associated Molecular Pattern Molecules*), tais como proteínas heat shock (HSP22, HSP60, HSP70, HSP96), produtos da matriz intracelular, DNA genômico, proteína *high mobility group box1* (HMGB1), ácido hialurônico, fibronectina e cristais de ácido úrico (SATO et al., 2009).

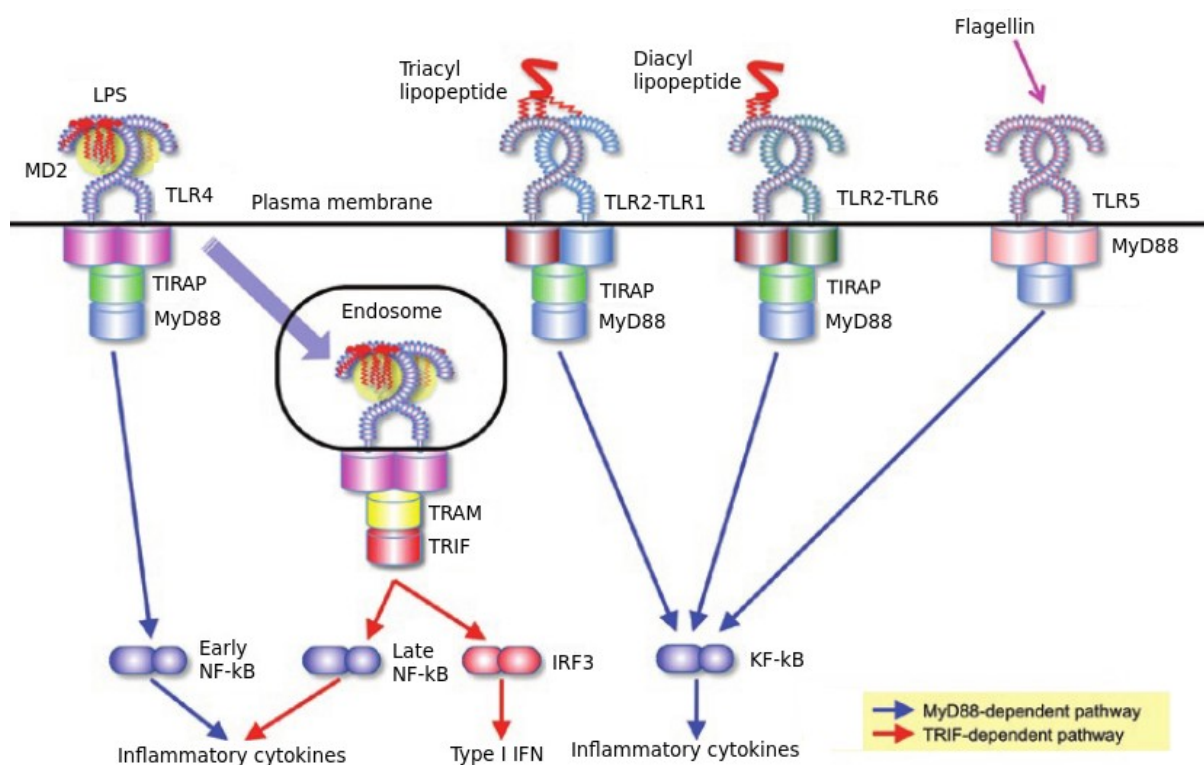


Figura 4 – Esquema do reconhecimento de patógenos e localização das TLRs 1, 2, 4, 5 e 6 (KAWAI; AKIRA, 2010).

A TLR2 detecta lipoproteínas tri ou diacetiladas de bactérias gram positivas e micoplasma em associação com as proteínas TLR1 e TLR6 respectivamente, bem como zymosan, envelope de proteínas do vírus do sarampo, coriomeningite linfocítica, o vírus arena e lipoarabinomannan de micobactéria (KUMAR; KAWAI; AKIRA, 2009).

TLR4 é o principal sensor para o lipopolissacarídeo (LPS) da parede celular de bactérias gram negativas (LPS), mas também detecta outros componentes tais como um polissacarídeo de planta formado por um polímero de manose (mannan), a proteína do vírus sincicial respiratório (RSV) e a proteína clamidial heat shock (Hsp) 60. Ela forma um complexo com a proteína MD2 na superfície celular e juntas servem como o principal componente de ligação do LPS. Outras proteínas como a de ligação do LPS (LBP) e a CD14 estão envolvidas na ligação do LPS (MONTERO VEGA; DE ANDRÉS MARTÍN, 2008).

A TLR5 reconhece a flagelina de bactérias extracelulares (HAJISHENGALLIS; LAMBRIS, 2010) (Figura 4). A TLR9 é sensível a motivos não metilados de DNA (CpG) que são frequentemente presentes em bactérias e vírus, mas são raros em células de mamíferos, enquanto que as TLRs 3, 7 e 8 são sensíveis a dsRNA viral e ssRNA respectivamente (Figura 5) (MANICASSAMY; PULENDRAN, 2009; MUCCIOLI et al., 2012).

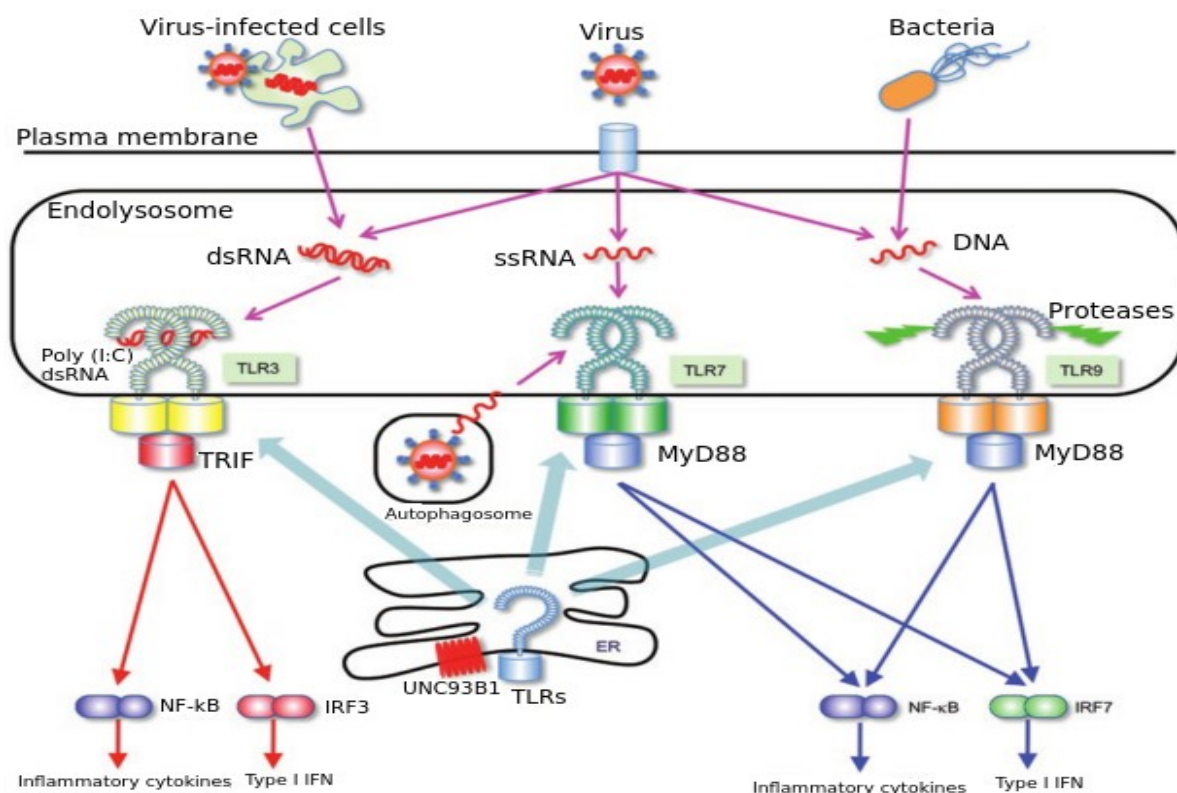


Figura 5 - Esquema do reconhecimento de patógenos e localização das TLRs 3, 7 e 9.(KAWAI; AKIRA, 2010).

O reconhecimento do patógeno inicia uma homodimerização ou heterodimerização (TLR1-TLR2 ou TLR6-TLR2) das TLRs iniciando eventos de

sinalização que conduz a expressão de mediadores inflamatórios incluindo muitas citocinas, quimiocinas e moléculas de adesão celular e promove a aniquilação direta do invasor, ativa a fagocitose e influencia respostas imune adaptativas (LI; JIANG; TAPPING, 2010). Conduz ao aumento da regulação da transcrição de distintos genes, dependendo das TLRs e o tipo de célula envolvido. A via de sinalização é dividida em duas dependendo do uso de diferentes moléculas adaptadoras, MyD88 e TRIF (TAKEUCHI; AKIRA, 2002). O uso seletivo de proteínas adaptadoras é um dos principais mecanismos que diferencia a via de sinalização das TLRs. O Myeloid differentiation antigen 88 (MyD88) é um adaptador central requerido por todas as TLRs, exceto a TLR3. Este necessita de outro adaptador para o recrutamento das TLR2 e 4, que são MyD88 adaptador like (Mal) ou também chamado de TIR domain containing protein (TIRAP) (MANAVALAN; BASITH; CHOI, 2011) (Figura 6).

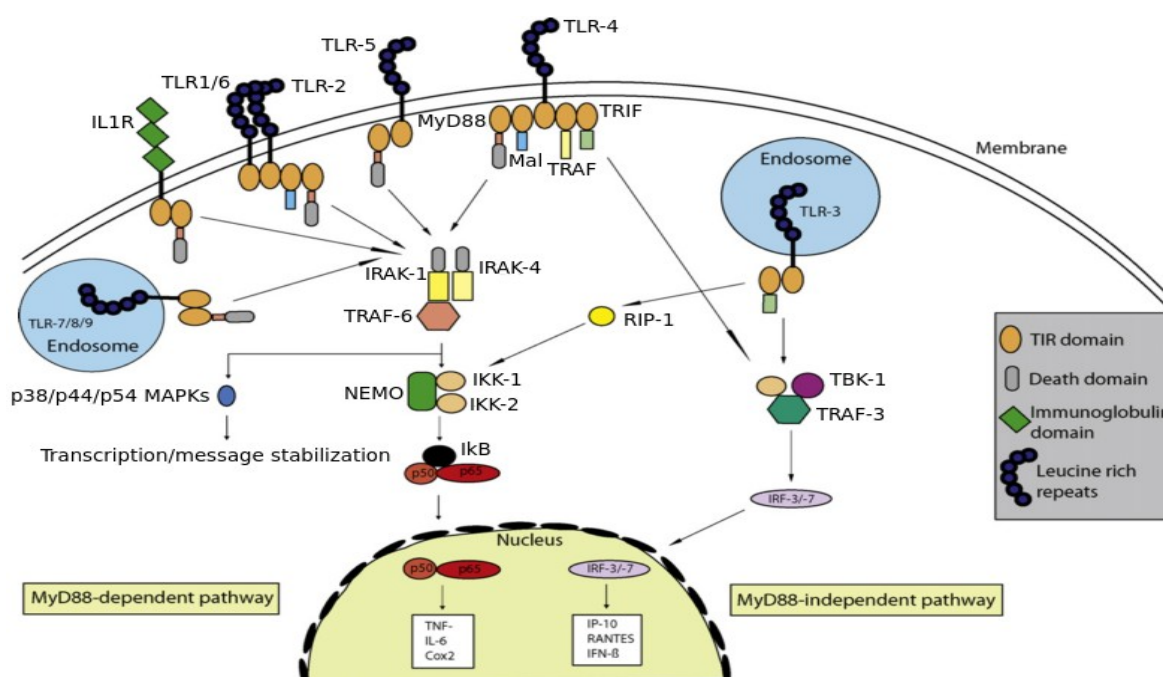


Figura 6 – Esquema da via de sinalização das TLRs. São mostradas as principais proteínas de sinalização intracelular que são ativadas depois do mecanismo de ligação do patógeno a TLR que levam a ativação do NfκB e IFNs (40).

A via do MyD88 irá induzir o fator nuclear kappa B (NfκB) e proteínas quinases associadas a mitose (MAPKs) através das moléculas IRAK (IL-1 receptor associated kinase)-4, IRAK-1 e TRAF-6. Além disso, uma via independente da MyD88 que é utilizada pela TLR3 e TLR4 depende do adaptador TRIF e conduz a indução de IRFs e uma ativação tardia do NFKb para a TLR4 quando esta não é

ativada inicialmente pelo sistema dependente do MyD88. Na via dependente do MyD88 as proteínas IRAK-4 e IRAK-1 são recrutadas pelo domínio DD (do inglês, death domain) do MyD88, onde são fosforiladas e recrutam o TRAF (NF receptor associated factor)-6 (KAWAI; AKIRA, 2011).

A IRAK-1 e TRAF6 dissociadas do receptor ativam o complexo I κ B quinase (IKK) resultando na translocação dos dímeros do Nf κ B (p50/p65) para o núcleo celular e no aumento de genes pro inflamatórios. Além da indução do Nf κ B, a via de sinalização dependente do MyD88 também ativa a p38 uma proteína quinase ativadora de mitose e a JNK (c-Jun-N-terminal kinase) (TAKEDA; KAISHO; AKIRA, 2003). Ademais, a proteína Mal está envolvida também na via de sinalização dependente da MyD88, porém serve como uma ponte para a ligação da TLR2 e TLR4 ao MyD88. Na via independente, as proteínas TLR3 e TLR4 fazem a mediação da ativação da IRF (interferon regulating factor)-3 e dos IFN (interferons). A molécula adaptadora TRIF é essencial para essa via. Ela recruta e ativa a proteína RIP-1 (Receptor-interacting serine/threonine-protein kinase 1) que é responsável pela indução tardia do Nf κ B pela TLR4, bem como TBK-1 (TANK-binding kinase 1) e IKKi, que formam um complexo com a TRAF-3 subsequentemente ativando IRF-3 e IRF-7. A proteína TRAM que está envolvida apenas na via de sinalização independente da MyD88 na TLR4 é necessária para o recrutamento da proteína TRIF (TAKEDA; AKIRA, 2004).

2.3.2 Toll Like Receptor e variações genéticas

O polimorfismo nos genes da família de proteína Toll Like Receptor pode influenciar a susceptibilidade de doenças alterando a resposta a infecção conduzindo a efeitos inflamatórios, onde a ocorrência de inflamações crônicas e recorrentes são implicadas no desenvolvimento e início de vários tipos de câncer (SINGH et al., 2013). SNPs situados na região rica em leucina podem afetar a ligação da TLR aos patógenos, na região transmembranar pode conduzir a defeitos no transporte de receptores intracelulares e na região citoplasmática pode alterar a interação com proteínas adaptadoras ou não ocorre a dimerização das TLRs (KUTIKHIN, 2011).

Devido a importância das TLRs no início e na regulação da resposta inflamatória, a desregulação na ativação dos circuitos de sinalização causada por

variações genéticas pode ter consequências significativas no processo inflamatório e carcinogênico (MEDVEDEV, 2013). A ativação da TLR pode proteger de agentes infecciosos e prevenir, inibir ou bloquear a carcinogênese enquanto que o rompimento da funcionalidade das TLRs pode permitir a infecção dos patógenos invasores ou células tumorais para evitar o reconhecimento do sistema imune e, conseqüentemente, não serem eliminados. Ao mesmo tempo a ativação das TLRs pode promover a carcinogênese, criando um ambiente pro inflamatório que é favorável ao progresso do tumor. Porém, se diminuir a atividade da TLR pode minimizar os efeitos da inflamação crônica e, por conseguinte, decrescer a probabilidade do desenvolvimento do tumor. Da mesma forma, se as proteínas da via de sinalização das TLRs estão pouco expressas, inativadas ou desabilitadas a realizarem sua função, pode resultar nos mesmos efeitos que quando decresce a atividade da TLR (ANTON G KUTIKHIN ARSENIY E YUZHALLIN, 2012).

Em geral, a sinalização da TLR4 promove a progressão e desenvolvimento de tumores, enquanto que as TLR2, 3, 7 e 9 possui um efeito antitumoral. A ativação das TLRs pode induzir tanto efeitos imuno estimulatórios quanto imuno supressores, por conseguinte, a seleção de agonistas deve ser realizada visando o bloqueio desses efeitos (YU; WANG; CHEN, 2013). Ter como alvo a via de sinalização das TLRs pode ser uma abordagem promissora para a prevenção e tratamento do câncer. Portanto, vários agonistas das TLRs foram desenvolvidos e estão em avaliação clínica e pré clínica. A elucidação da função da TLR na biologia do tumor pode levar ao descobrimento de novos alvos terapêuticos no tratamento do câncer (BHATTACHARYA; YUSUF, 2012).

2.4 Bioinformática e servidores computacionais

Desde a descoberta da dupla hélice do DNA feita por Watson & Crick, vários avanços incluindo a técnica de DNA recombinante e a reação em cadeia da polimerase (PCR), levou ao início da análise do sequenciamento e finalmente a formação do projeto genoma humano (PROSDOCIMI; COUTINHO, 2002). Essas descobertas levaram ao desenvolvimento de métodos tais como a genômica e proteômica, que aliadas geram uma complementação e conduzem a correlação entre genótipos e fenótipos (KNAUP et al., 2004). Por causa dos avanços na

genômica, a quantidade de informação a nível molecular e médico tem sido largamente depositado em grandes bancos de dados. Porém, a grande quantidade de informação não estaria disponível sem ferramentas que fossem capazes de minerar os dados, organizá-los e criar hipóteses baseadas em pesquisas laboratoriais para a clínica médica. A partir destas necessidades, surgiu a bioinformática (BUTTE, 2009).

Podemos definir a bioinformática como uma área da ciência que pesquisa, desenvolve e aplica ferramentas computacionais em dados biológicos, médicos e de saúde (ZWART, 2009) (Figura 7). É um campo interdisciplinar, que aborda ciência computacional, matemática, física e biologia (BAYAT, 2002) (Figura 7). Durante o início dos anos 60, a ciência computacional surgiu como uma peça importante no estudo da biologia molecular. O estudo da estrutura e sequência proteica desencadeou o crescimento da bioinformática durante as últimas décadas (CHANG, 2005).

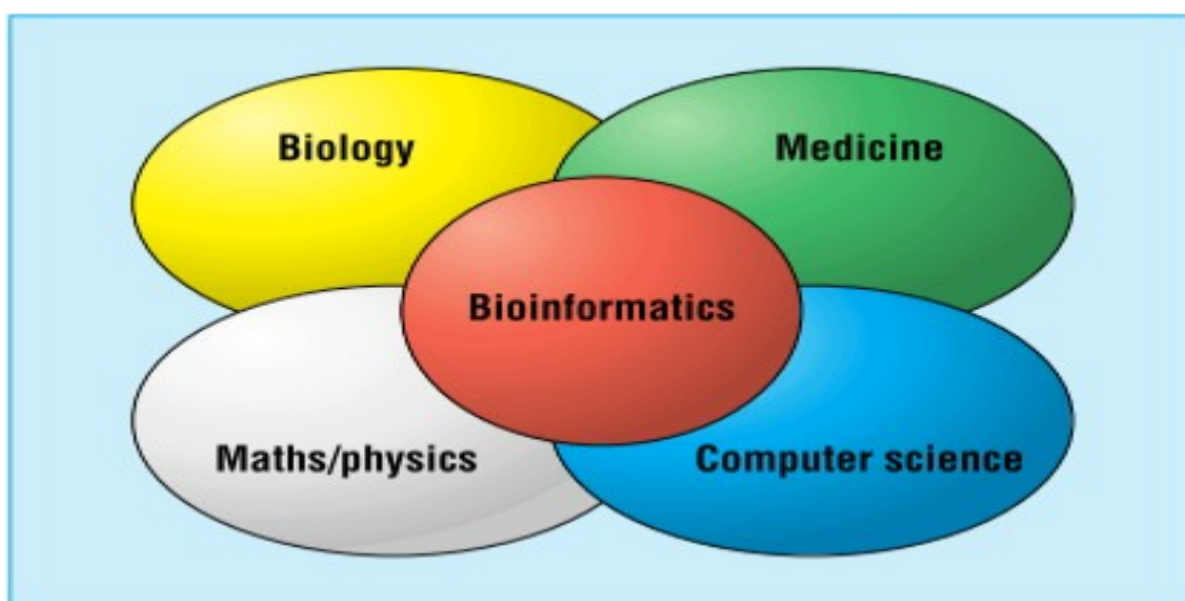


Figura 7 – Interação de disciplinas que tem contribuído para o desenvolvimento da bioinformática (BAYAT, 2002).

Através da bioinformática é possível mapear sequências para bancos de dados, criar modelos de interação entre moléculas, avaliar compatibilidade estrutural, encontrar diferenças entre DNA hospedeiro e de patógenos e identificar regiões de conservação em estruturas proteicas. Além disso, a bioinformática representa uma nova forma de expandir o conhecimento acerca de doenças, ligação entre genótipos e fenótipos, resolver a causa de doenças poligenéticas, e identificar

novos alvos terapêuticos. Ela é dividida em dois grandes ramos: uso da informação armazenada em bancos de dados e análise da informação através de ferramentas mineradoras de dados (DEBES; URRUTIA, 2004).

No geral a bioinformática possui três objetivos: organizar os dados obtidos de forma que permitam aos pesquisadores acesso e possam submeter novos que venham a produzir; desenvolver ferramentas e recursos que auxiliem na análise de dados; usar as ferramentas para analisar os dados e interpretar os resultados de uma forma biologicamente significativa (LUSCOMBE; GREENBAUM; GERSTEIN, 2001).

Diversos algoritmos computacionais usando métodos bioquímicos e genéticos tem sido desenvolvidos para tentar prever quais modificações que ocorrem na função da proteína, que podem ou não causar uma determinada doença. Estes, dependem largamente de informação acerca da estrutura tridimensional e de regiões conservadas da proteína (VALDMANIS; VERLAAN; ROULEAU, 2009). É importante considerar a conservação das proteínas entre as espécies quando novos polimorfismos missenses são identificados como variantes que podem afetar posições que são conservadas através da evolução, são susceptíveis a serem patogênicas. A posição da variação na proteína pode também ser importante na determinação da sua patogenicidade, como por exemplo, se está localizada em um domínio funcional ou se afeta um aminoácido que já foi identificado como uma mutação patogênica (FLANAGAN; PATCH; ELLARD, 2010). Métodos automáticos que podem fazer a predição dos efeitos das mutações irá permitir a redução da caracterização experimental destas economizando tempo e dinheiro (WORTH; PREISSNER; BLUNDELL, 2011).

Com o surgimento da grande quantidade de variações, o desenvolvimento de métodos computacionais sofisticados que predizem a funcionalidade das alterações provocadas por nsSNPs estão crescendo ao passar dos anos (CLINE; KARCHIN, 2011). Os servidores para a predição de dano na proteína são construídos através de conhecimentos da ciência da computação, matemática aplicada, genética de populações, propabilidade, modelagem, estatística, engenharia de software e filogenia (KARCHIN, 2009). Os servidores web tais como o Sorting intolerant from tolerant (SIFT), polymorphism phenotyping (PolyPhen), Site Directed Mutator (SDM) e MutationAssessor, são servidores que utilizam informação ou da conservação ou

estrutura para realizar análises de predição do impacto de nsSNPs no desenvolvimento de doenças (ADZHUBEI et al., 2010).

O programa Polyphen é uma ferramenta computacional que é baseada em uma combinação de informações filogenéticas, estruturais e anotações de sequência para fazer a predição do possível impacto da substituição de um aminoácido na estrutura e função da proteína. Ele utiliza oito alinhamentos de sequências homólogas e critérios baseados na estrutura da proteína que são selecionados automaticamente por um algoritmo iterativo com uma acurácia de 80%. É mais limitado quando dados estruturais não são disponíveis. Se a estrutura da proteína que está sendo testada é desconhecida, então o programa usa uma proteína homóloga com a estrutura já elucidada. Para cada troca de aminoácido, o Polyphen realiza vários passos: 1- extração de características baseadas na sequência do local de substituição do banco de dado UniProt; 2 – calcula os valores da troca das duas variações de aminoácido; 3 – calcula os parâmetros estruturais. As saídas do programa são classificadas como “benigno”, “possivelmente danoso”, “provavelmente danoso” e o cálculo da contagem de cada variante e da diferença das duas variantes for igual ou maior a 1.5 é considerado danoso (DOSS; SETHUMADHAVAN, 2009).

SIFT é um servidor que faz a distinção entre mutação funcional e não funcional em regiões codificadoras, se a substituição do aminoácido terá um efeito fenotípico. É baseado na premissa que a evolução da proteína é correlacionada com a função desta. Os polimorfismos que ocorrem em posições que são conservadas no alinhamento, espera-se que sejam menos toleradas que as que ocorrem em posições diversas. Para a realização da predição do impacto do SNP na proteína o programa realiza vários passos. Inicialmente procura por sequências similares; procura sequências que dividem função similar a sequência que está servindo de comparação; alinha as sequências escolhidas; e calcula a probabilidades para todas as substituições possíveis de um alinhamento. Se o valor do cálculo da probabilidade for menor que 0.05 são considerados como prejudiciais para a proteína, enquanto que maior do que ou igual a 0.05 são considerados tolerados. SIFT tem sido usado para analisar nsSNPs em genes associados com doenças humanas, com uma acurácia entre 65 e 92% (CHAN P. A. et al, 2007).

O servidor SDM é uma função de energia potencial estatística que usa frequências de substituição de aminoácidos com famílias de proteínas homólogas

para calcular o valor da estabilidade, que é análoga a diferença de energia livre entre a proteína selvagem e a mutante (WORTH; BURKE; BLUNDELL, 2007). Os resultados obtidos por esse programa são estruturas tridimensionais contendo os resíduos de aminoácido selvagens e mutantes destacados, o valor da predição da estabilidade com uma acurácia de 61%, assim como a predição da associação com doença (WORTH; PREISSNER; BLUNDELL, 2011).

O software Mutation Assessor tem uma aproximação da conservação mais elaborado. Ele distingue entre padrões de conservação em alinhamento de famílias (valor de conservação) e sub-famílias (valor de especificidade) de homólogos e também tenta explicar mudanças funcionais entre subfamílias de proteínas. A especificidade é definida pelos resíduos baseados no agrupamento de identificação das subfamílias de sequências homólogas para determinar a especificidade funcional no contexto de toda conservação. Ao analisar as sequências homólogas esse servidor possui uma acurácia de 81% (GNAD et al., 2013).

Pesquisas tais como a de Chan P. A. et al 2007, usaram servidores computacionais para avaliar mutações. Eles compararam os métodos de SIFT, Polyphen e A-GVGD na predição das consequências de 254 variações missenses em cinco diferentes genes (CDKN2A, MLH1, MSH2, MECP2 e TYR) e mostraram que a predição é útil para interpretar os genes que estão envolvidos em doenças humanas. Masoodi et al 2012, selecionou um conjunto de polimorfismos missenses do gene ePHA2 que pode estar envolvido na susceptibilidade de formação da catarata. Para saber a funcionalidade destes foram usados os servidores SIFT e Polyphen. Eles encontraram que o SNP rs2291806 pode ter um impacto funcional no gene EPHA2.

O uso de servidores computacionais que utilizam a predição do impacto de SNPs missenses pode ser útil na seleção de polimorfismos que podem estar associados com alguma doença e podem ser usados em investigações mais detalhadas.

3. OBJETIVOS

3.1. Geral

Analisar, através de ferramentas de bioinformática, os polimorfismos de base única (SNP) nos genes Toll Like Receptor (TLR) e as consequências estruturais das proteínas, buscando por polimorfismos que possam estar relacionados com a predisposição ao câncer.

3.2. Específicos

I. Realizar “data mining” de polimorfismos no banco de dados de SNP do NCBI (<http://www.ncbi.nlm.nih.gov/SNP/>);

II. Realizar literature mining no banco de dados de literatura do NCBI, PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>);

III. Caracterizar os polimorfismos encontrados quanto a troca de aminoácidos, troca de frame e alterações que levem a uma mudança estrutural na proteína;

IV. Analisar os SNPs missenses relacionados com a predisposição ao câncer na literatura;

V. Analisar os SNPs missenses coletados do banco de dados do NCBI.

4. REFERÊNCIAS

- ABDELSADIK, A.; TRAD, A. Toll-like receptors on the fork roads between innate and adaptive immunity. **Human immunology**, v. 72, n. 12, p. 1188–93, dez. 2011.
- ADZHUBEI, I. A et al. A method and server for predicting damaging missense mutations. **Nature methods**, v. 7, n. 4, p. 248–9, abr. 2010.
- AKIRA, S.; UEMATSU, S.; TAKEUCHI, O. Pathogen recognition and innate immunity. **Cell**, v. 124, n. 4, p. 783–801, 24 fev. 2006.
- ANTON G KUTIKHIN ARSENIY E YUZHALLIN. Are Toll-like receptor gene polymorphisms associated with prostate cancer? p. 23–29, 2012.
- BASSETT, E. H. Toll Receptors and the Renaissance of Innate Immunity. p. 1–17, 2004.
- BAYAT, A. Clinical review. v. 324, n. April, p. 1018–1022, 2002.
- BELL, J. K. et al. The molecular structure of the Toll-like receptor 3 ligand-binding domain. **Proceedings of the National Academy of Sciences of the United States of America**, v. 102, n. 31, p. 10976–80, 2 ago. 2005.
- BHATTACHARYA, D.; YUSUF, N. Expression of toll-like receptors on breast tumors: taking a toll on tumor microenvironment. **International journal of breast cancer**, v. 2012, p. 716564, jan. 2012.
- BUTTE, A. J. Translational bioinformatics applications in genome medicine. **Genome medicine**, v. 1, n. 6, p. 64, jan. 2009.
- CHAN P. A. et al. Interpreting Missense Variants: Comparing Computational Methods in Human Disease Genes CDKN2A, MLH1, MSH2, MECP2 and Tyrosinase (TYR). v. 28, n. 7, p. 683–693, 2007.
- CHANG, P. L. Clinical bioinformatics. **Chang Gung medical journal**, v. 28, n. 4, p. 201–11, abr. 2005.
- CHANG, Z. L. Important aspects of Toll-like receptors, ligands and their signaling pathways. **Inflammation research: official journal of the European Histamine Research Society ... [et al.]**, v. 59, n. 10, p. 791–808, out. 2010.
- CHMIELECKI, J.; MEYERSON, M. DNA sequencing of cancer: what have we learned? **Annual review of medicine**, v. 65, p. 63–79, 14 jan. 2014.
- CLINE, M. S.; KARCHIN, R. Using bioinformatics to predict the functional impact of SNVs. **Bioinformatics (Oxford, England)**, v. 27, n. 4, p. 441–8, 15 fev. 2011.
- DEBES, J. D.; URRUTIA, R. Bioinformatics tools to understand human diseases. **Surgery**, v. 135, n. 6, p. 579–85, jun. 2004.
- DOSS, C. G. P.; SETHUMADHAVAN, R. Investigation on the role of nsSNPs in HNPCC genes--a bioinformatics approach. **Journal of biomedical science**, v. 16, p. 42, jan. 2009.
- FLANAGAN, S. E.; PATCH, A.-M.; ELLARD, S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. **Genetic testing and molecular biomarkers**, v. 14, n. 4, p. 533–7, ago. 2010.

GEORGE PRIYA DOSS, C. et al. Applications of computational algorithm tools to identify functional SNPs. **Functional & integrative genomics**, v. 8, n. 4, p. 309–16, nov. 2008.

GNAD, F. et al. Assessment of computational methods for predicting the effects of missense mutations in human cancers. **BMC genomics**, v. 14 Suppl 3, n. Suppl 3, p. S7, jan. 2013.

HAJISHENGALLIS, G.; LAMBRIS, J. D. Crosstalk pathways between Toll-like receptors and the complement system. **Trends in immunology**, v. 31, n. 4, p. 154–63, abr. 2010.

IWASAKI, A.; MEDZHITOV, R. Toll-like receptor control of the adaptive immune responses. **Nature immunology**, v. 5, n. 10, p. 987–95, out. 2004.

JOHNSON, A. D. Single-nucleotide polymorphism bioinformatics: a comprehensive review of resources. **Circulation. Cardiovascular genetics**, v. 2, n. 5, p. 530–6, out. 2009.

JOHNSON, M. M.; HOUCK, J.; CHEN, C. Screening for deleterious nonsynonymous single-nucleotide polymorphisms in genes involved in steroid hormone metabolism and response. **Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology**, v. 14, n. 5, p. 1326–9, maio 2005.

KARCHIN, R. Next generation tools for the annotation of human SNPs. **Briefings in bioinformatics**, v. 10, n. 1, p. 35–52, jan. 2009.

KAWAI, T.; AKIRA, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. **Nature immunology**, v. 11, n. 5, p. 373–84, maio 2010.

KAWAI, T.; AKIRA, S. Toll-like receptors and their crosstalk with other innate receptors in infection and immunity. **Immunity**, v. 34, n. 5, p. 637–50, 27 maio 2011.

KIM, S. K. et al. A missense polymorphism (rs11466653, Met326Thr) of toll-like receptor 10 (TLR10) is associated with tumor size of papillary thyroid carcinoma in the Korean population. **Endocrine**, v. 43, n. 1, p. 161–9, fev. 2013.

KIRK, P.; BAZAN, J. F. Pathogen recognition: TLRs throw us a curve. **Immunity**, v. 23, n. 4, p. 347–50, out. 2005.

KNAUP, P. et al. Towards clinical bioinformatics: advancing genomic medicine with informatics methods and tools. **Methods of information in medicine**, v. 43, n. 3, p. 302–7, jan. 2004.

KUMAR, H.; KAWAI, T.; AKIRA, S. Toll-like receptors and innate immunity. **Biochemical and biophysical research communications**, v. 388, n. 4, p. 621–5, 30 out. 2009.

KUTIKHIN, A. G. Association of polymorphisms in TLR genes and in genes of the Toll-like receptor signaling pathway with cancer risk. **Human immunology**, v. 72, n. 11, p. 1095–116, nov. 2011.

LAZARUS, R. et al. Single nucleotide polymorphisms in innate immunity genes: abundant variation and potential role in complex human disease. **Immunological reviews**, v. 190, p. 9–25, dez. 2002.

LAZARUS, R. et al. Single-nucleotide polymorphisms in the Toll-like receptor 9 gene (TLR9): frequencies, pairwise linkage disequilibrium, and haplotypes in three U.S. ethnic groups and exploratory case–control disease association studies☆☆This work was supported by Programs. **Genomics**, v. 81, n. 1, p. 85–91, jan. 2003.

LI, X.; JIANG, S.; TAPPING, R. I. Toll-like receptor signaling in cell proliferation and survival. **Cytokine**, v. 49, n. 1, p. 1–9, jan. 2010.

LIU, S. et al. Generation of genome-scale gene-associated SNPs in catfish for the construction of a high-density SNP array. **BMC genomics**, v. 12, n. 1, p. 53, jan. 2011.

LUSCOMBE, N. M.; GREENBAUM, D.; GERSTEIN, M. What is bioinformatics? A proposed definition and overview of the field. **Methods of information in medicine**, v. 40, n. 4, p. 346–58, jan. 2001.

MANAVALAN, B.; BASITH, S.; CHOI, S. Similar Structures but Different Roles - An Updated Perspective on TLR Structures. **Frontiers in physiology**, v. 2, n. July, p. 41, jan. 2011.

MANICASSAMY, S.; PULENDRAN, B. Modulation of adaptive immunity with Toll-like receptors. **Seminars in immunology**, v. 21, n. 4, p. 185–93, ago. 2009.

MASOODI, T. A. et al. Screening and structural evaluation of deleterious Non-Synonymous SNPs of ePHA2 gene involved in susceptibility to cataract formation. **Bioinformation**, v. 8, n. 12, p. 562–7, jan. 2012.

MEDVEDEV, A. E. Toll-Like Receptor Polymorphisms, Inflammatory and Infectious Diseases, Allergies, and Cancer. **Journal of interferon & cytokine research: the official journal of the International Society for Interferon and Cytokine Research**, v. 00, n. 00, 15 maio 2013.

MONTERO VEGA, M. T.; DE ANDRÉS MARTÍN, A. Toll-like receptors: a family of innate sensors of danger that alert and drive immunity. **Allergologia et Immunopathologia**, v. 36, n. 6, p. 347–357, dez. 2008.

MUCCIOLI, M. et al. Toll-like receptors as novel therapeutic targets for ovarian cancer. **ISRN oncology**, v. 2012, p. 642141, jan. 2012.

NIETERS, A et al. Gene polymorphisms in Toll-like receptors, interleukin-10, and interleukin-10 receptor alpha and lymphoma risk. **Genes and immunity**, v. 7, n. 8, p. 615–24, dez. 2006.

O'NEILL, L. A J.; BOWIE, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. **Nature reviews. Immunology**, v. 7, n. 5, p. 353–64, maio 2007.

PROSDOCIMI, F.; COUTINHO, G. Bioinformática: manual do usuário. **Biotecnologia ...**, 2002.

RAMENSKY, V.; BORK, P.; SUNYAEV, S. Human non-synonymous SNPs: server and survey. **Nucleic acids research**, v. 30, n. 17, p. 3894–900, 1 set. 2002.

REVA, B.; ANTIPIN, Y.; SANDER, C. Predicting the functional impact of protein mutations: application to cancer genomics. p. 1–14, 2011.

SATO, Y. et al. Cancer Cells Expressing Toll-like Receptors and the Tumor Microenvironment. **Cancer microenvironment: official journal of the**

International Cancer Microenvironment Society, v. 2 Suppl 1, p. 205–14, set. 2009.

SINGH, V. et al. Single-nucleotide polymorphisms in genes encoding toll-like receptor -2, -3, -4, and -9 in a case-control study with bladder cancer susceptibility in a North Indian population. **Archives of medical research**, v. 44, n. 1, p. 54–61, jan. 2013.

SUNYAEV, S. et al. Prediction of deleterious human alleles. **Human molecular genetics**, v. 10, n. 6, p. 591–7, 15 mar. 2001.

SUO, S.-B. et al. Proteome-wide analysis of amino acid variations that influence protein lysine acetylation. **Journal of proteome research**, v. 12, n. 2, p. 949–58, 1 fev. 2013.

TAKEDA, K.; AKIRA, S. TLR signaling pathways. **Seminars in Immunology**, v. 16, n. 1, p. 3–9, fev. 2004.

TAKEDA, K.; KAISHO, T.; AKIRA, S. Toll-like receptors. **Annual review of immunology**, 2003.

TAKEUCHI, O.; AKIRA, S. Genetic approaches to the study of Toll-like receptor function. **Microbes and Infection**, v. 4, n. 9, p. 887–895, jul. 2002.

TAKEUCHI, O.; AKIRA, S. Pattern recognition receptors and inflammation. **Cell**, v. 140, n. 6, p. 805–20, 19 mar. 2010.

TENG, S. et al. Modeling effects of human single nucleotide polymorphisms on protein-protein interactions. **Biophysical journal**, v. 96, n. 6, p. 2178–88, 18 mar. 2009.

TENG, S.; MICHONOVA-ALEXOVA, E.; ALEXOV, E. Approaches and resources for prediction of the effects of non-synonymous single nucleotide polymorphism on protein function and interactions. **Current pharmaceutical biotechnology**, v. 9, n. 2, p. 123–33, abr. 2008.

TENG, S.; SRIVASTAVA, A. K.; WANG, L. Sequence feature-based prediction of protein stability changes upon amino acid substitutions. **BMC genomics**, v. 11 Suppl 2, n. Suppl 2, p. S5, jan. 2010.

TORKAMANI, A.; SCHORK, N. J. Predicting functional regulatory polymorphisms. **Bioinformatics (Oxford, England)**, v. 24, n. 16, p. 1787–92, 15 ago. 2008.

VALDMANIS, P. N.; VERLAAN, D. J.; ROULEAU, G. A. The proportion of mutations predicted to have a deleterious effect differs between gain and loss of function genes in neurodegenerative disease. **Human mutation**, v. 30, n. 3, p. E481–9, mar. 2009.

WEISCHENFELDT, J. et al. Phenotypic impact of genomic structural variation: insights from and for human disease. **Nature reviews. Genetics**, v. 14, n. 2, p. 125–38, fev. 2013.

WORTH, C. L.; BURKE, D.; BLUNDELL, T. L. Beilstein-Institut 11. p. 11–26, 2007.

WORTH, C. L.; PREISSNER, R.; BLUNDELL, T. L. SDM--a server for predicting effects of mutations on protein stability and malfunction. **Nucleic acids research**, v. 39, n. Web Server issue, p. W215–22, jul. 2011.

YANG, H. et al. Reduced expression of Toll-like receptor 4 inhibits human breast cancer cells proliferation and inflammatory cytokines secretion. **Journal of experimental & clinical cancer research : CR**, v. 29, n. 1, p. 92, jan. 2010.

YU, L.; WANG, L.; CHEN, S. Dual character of Toll-like receptor signaling: pro-tumorigenic effects and anti-tumor functions. **Biochimica et biophysica acta**, v. 1835, n. 2, p. 144–54, abr. 2013.

ZHANG, R. et al. SNP Cutter: a comprehensive tool for SNP PCR-RFLP assay design. **Nucleic acids research**, v. 33, n. Web Server issue, p. W489–92, 1 jul. 2005.

ZHANG, Z. D. et al. Identification of genomic indels and structural variations using split reads. **BMC genomics**, v. 12, n. 1, p. 375, jan. 2011.

ZWART, H. Genomics and identity: the bioinformatisation of human life. **Medicine, health care, and philosophy**, v. 12, n. 2, p. 125–36, jun. 2009.

5. ARTIGO

Título: Functional *in silico* analysis of Single Nucleotide Polymorphisms in the genes of the Toll Like Receptors (TLR) family and the susceptibility to development of cancer

Revista: Human Mutation

Fator de Impacto: 5.213

Title: Functional *in silico* analysis of Single Nucleotide Polymorphisms in the genes of the Toll Like Receptors (TLR) family and the susceptibility of the development of cancer

Carolina da Rocha Simões^{a,b}; Danyelly Bruneska Gondim Martins^{a,b}; José Luiz de Lima Filho^{a,b}; Carlos Henrique Madeiros Castelletti^{a,c}

^a Molecular Prospection and Bioinformatics Group (ProspecMol), Laboratory of Immunopathology Keizo Asami (LIKA), Federal University of Pernambuco (UFPE), Recife, PE, 50670-901, Brazil;

^b Department of Biochemistry, Federal University of Pernambuco (UFPE), Recife, PE, 50670-901, Brazil

^c Agronomic Institute of Pernambuco (IPA), Recife, PE, 50761-000, Brazil;

Corresponding author:

Carlos Henrique Madeiros Castelletti

Grupo de Prospecção Molecular e Bioinformática (ProspecMol)

Laboratório de Imunopatologia Keizo Asami (LIKA)

Universidade Federal de Pernambuco (UFPE)

Av. Prof. Moraes Rego, S/N

Cidade Universitária, Recife, Pernambuco, Brasil

CEP: 50670-901

Phone: + 55 81 2126 8484

Fax: + 55 81 2126 8485

e-mail: hcastelletti@prospecmol.org

ABSTRACT

The increase of information from genome sequencing permitted the discovery of millions of variations, like Single Nucleotide Variations and the challenge is to identify the relationship between genotype and phenotype. Computational tools can help to identify these relationships and to support the prediction of functionality alterations. The Toll Like Receptors (TLR) are proteins from the immune system and are involved in the regulation of inflammation and the development of cancer. In this work we collected 5,839 SNPs in 10 TLRs genes and performed structural analysis in 1,017 missense polymorphisms to evaluate the consequences of amino acid changes by four computational methods (SIFT, Polyphen, MutationAssessor and SDM). The analysis revealed that in 223 polymorphisms the structural changes suffered by TLR proteins were classified as damaged and consequently can conduct to a malfunction, and associated to a disease. Results showed the importance of the computational prediction of missense SNPs to help the interpretation of variations on genes associated with cancer and in a personalized diagnose.

KEYWORDS: structural variation, SNP, conservation

Nowadays with the advent of high-throughput technologies (NGS) is possible sequencing the human genome at a significantly lower cost, enabling novel medical diagnostics and treatment (Dalca and Brudno 2010). The NGS platform provided a high amount of structural variations and knowledge of genetic variants to the susceptibility of diseases (Koboldt et al. 2010). The challenge is to identify the relationship between genotype and phenotype (Li and Durbin 2010, (Zhang et al. 2011), providing information about clinical-pathological indexes and molecular profiling to create diagnostic, prognostic, and therapeutic strategies tailored to each patient, the precision medicine (Mirnezami Reza 2012).

Nonsynonymous SNPs (nsSNP) are the most investigated polymorphism, implicated in 50% of known mutations in inherited diseases which may lead to drastic phenotypic consequences (Suo et al. 2013), varying the functional impact, depending on their positions, function and nature of the replacement amino acid (Gnad et al. 2013). Polymorphisms in genes like Toll Like Receptor (TLR), play an important role in the innate immunity and are involved in the regulation of inflammatory responses (Minmin et al. 2011), implicating in the development and initiation of several kinds of cancers by chronic and recurrent inflammations (Sun et al. 2005b). They are a transmembrane protein type I and have three domains; the ectodomain that contains multiple leucine-rich repeats that recognize the pathogens; the single transmembrane domain and the intraceclular domain homologous to human interleukin 1 that interacts with signaling machinery of imune response (Cheng et al. 2007). SNPs situated in LRRs can affect the ability of receptors to bind pathogens and in the transmembrane domain may conduct to defects in intracellular receptor transport. In

the cytoplasmic domain, polymorphisms can outcome in altered interaction with adapter proteins or in disrupted dimerization (Anton G Kutikhin Arseniy & Yuzhalin 2012). In this work we conducted *in silico* analyses of nonsynonymous SNPs in the genes of TLR family, investigating the structural variations in the proteins, and correlating with the literature mining about association with cancer.

We collected and manual curated the SNP information about the 10 TLRs members available in dbSNP/NCBI database (build 138), using the parameters “clinical source” and “in gene region”. The missenses SNPs are annotated for the prediction of functional impact by different methodologies of computational analyses. We used four *in silico* web servers to perform the functional analysis: The *SIFT* prediction (<http://sift.jcvi.org>) is based on the degree of conservation of amino acid residues in sequence alignments derived from closely related sequences, collected through PSI-BLAST (Kumar et al. 2009); The *PolyPhen-2*, available in <http://genetics.bwh.harvard.edu/pph2/>, use eight sequence-based and three structure-based predictive features, including structural, phylogenetic and sequence annotation information to characterize the nsSNP (Adzhubei et al. 2010). The *Mutation Assessor* (<http://mutationassessor.org/>) can distinguishes between conservation patterns within aligned families and subfamilies of homologs and so attempts to account for functional shifts between subfamilies of proteins to predict the functional impact of amino acid substitutions in proteins (Reva et al. 2011). The *Site Directed Mutator* (SDM) available in <http://mordred.bioc.cam.ac.uk/~sdm/sdm.php>, use amino acid substitution frequencies within homologous protein families to calculate stability score between a wild type and mutant protein (Worth et al. 2011).

We found a total of 5,839 SNPs in all TLRs genes distributed into 3' near gene, 3' untranslated region, 5' near gene, 5' untranslated region, frameshift, intron, nonsense, synonymous, and missense (Table 1).

The assays realized with the web servers for TLR1 showed that 36 SNPs were predicted to cause a malfunction on the protein. A study developed by Stevens et al. (2008) tested a gene cluster (TLR1-TLR6-TLR10) and found that four polymorphism in TLR1 (rs4833095, rs5743596, rs5743595 and rs5743551) could be associated with the reduced risk of prostate cancer. Breyer et al. 2009 did not find association with this cancer. Another study developed by Sun et al. 2005 suggested that these gene cluster is associated with the increased risk of prostate cancer. These differences in the researches most exist because of unknown differences in the study population or with the distinction of stages of the development of cancer in the two works. Prediction analysis of rs4833095 showed that, in neither web server, it was associated with the development of disease, corroborating with the findings of Stevens et al (2008).

For the analysis realized to TLR2 gene the results revealed that 21 were considered to injure the protein. A work released by Slattery et al. 2012 tested SNPs in the gene of TLR2 and shows a few statistically significant associations with colon and rectal cancer. The TLR2 rs5743708 polymorphism did not have an association with colorectal cancer (Figure 1)(Boraska Jelavić et al. 2006, Davoodi and Seow 2011). Even so, Ashton et al. 2010 did not found an association with endometrial cancer. In an assay of the gene of TLR2 conducted by Kim et al. 2012 was found an association with two nonsynonymous SNPs (rs3804099 and rs3804100) and Papillary Thyroid Cancer. As well the work of Junjie et al. 2012, found an association with these SNPs and hepatocellular carcinoma susceptibility. A study conducted by

Purdue et al. 2009 showed that the rare variant rs3804100 was significantly associated with marginal zone lymphoma. Immunohistochemical expression study shows that TLR2 is present in carcinoma and adenomas of follicular thyroid neoplasms and gastric dysplasia (Hagström et al. 2012, Pimentel-Nunes et al. 2011). The web servers analysis of rs5743708 showed that it was considered damaging by three of the four methods utilized (Polyphen, SIFT and MutationAssessor). Thus, new researches will be necessary to clarify if this polymorphism can conduct to a malfunction of the protein leading to the development of cancer in others populations.

Computational tests accomplished for the TLR3 showed that 11 SNPs could be dangerous for the protein. Expression studies with the TLR3 gene have indicated that it show high expression by breast and prostate cancer cells (González-Reyes et al. 2010, González-Reyes et al. 2011). The association with TLR3 and bladder cancer susceptibility and the nonsynonymous rs3775290 did not exhibit correlation also with the risk of developing cervical cancer in Indians populations (Figure 2) (Pandey et al. 2011, Singh et al. 2013). An investigation performed in a relatively large and homogeneous German population demonstrated that a missense polymorphism (rs3775291) is an independent prognostic marker for disease survival in patients with stage II colorectal cancer (Castro et al. 2011, Slattery et al. 2012). In the computational analysis, rs3775291 was evaluated as damaging by the Polyphen and SIFT servers. These results can corroborate with the findings of Castro et al and Slattery et al 2012, that this polymorphism could be a marker for the development of colorectal cancer.

The analyses of TLR4 demonstrated that 38 SNPs were predicted to be harmful for the protein. The role of the TLR4 gene polymorphisms in genetic susceptibility to gastric cancer has been intensively investigated. The work

elaborated by Hold et al. 2007 found that rs4986790 could be considered a risk factor for noncardia gastric carcinoma and its precursors (Figure 3). On the other hand, Achyut et al. 2007 and Santini et al. 2008 did not detect association between rs4986790 and the development of gastritis and precancerous lesions but rs4986791 is linked with an increased susceptibility to gastric cancer development. A study in a Brazilian population did not find association with rs4986790 and gastric cancer (de Oliveira and Silva 2012). Trejo-de la O et al. 2008 and Rigoli et al. 2010 found that rs4986790 and rs4986791 appear to play a role in gastric carcinogenesis. However, Garza-Gonzalez et al. 2007 shows that both rs4986790 and rs4986791 were accepted as a risk factor. As well as, the TLR4 rs11536889 is a factor of risk of gastric atrophy and gastric cancer (Hishida et al. 2009). Huang et al. 2010 genotyped two SNPs in TLR4, rs10759932 may play a protective role in gastric carcinogenesis and rs1927914 that did not have any association. Ohara et al. 2006 suggests that T135A is related with the development of gastric adenocarcinoma. Further works tried to demonstrate the association of TLR4 with other cancers. Kim et al. 2012a and Kelly et al. 2006 observed that TLR4 might be a risk factor for developing ovarian epithelial cancer. In a study released in Indian women revealed that rs4986791 was found to be significantly associated with early stage (Stage II) of cervical cancer (Figure 4) (Pandey et al. 2009). The results of Theodoropoulos et al. 2012 indicate that rs4986790 may confer an increased susceptibility to breast cancer development. Six heterozygous genotypes (rs10759930, rs2737190, rs10116253, rs1927914, rs12377632 and rs1927911) had significantly decreased risk of Hepatocellular carcinoma (Minmin et al. 2011). A variant allele substitution (11350C) may act as a genetic susceptibility factor for nasopharyngeal carcinoma risk (Song et al. 2006). The association with bladder cancer susceptibility was revealed in an Indian

population (Singh et al. 2013). Logistic regression analysis showed an association of rs4986791 with gallbladder cancer (Srivastava et al. 2010). A study conducted by Purdue et al. 2009 did not find association with non-Hodgkin lymphoma (NHL). Bergmann et al. 2011 found that rs4986790 and rs4986791 are involved in the tumor development of Head and Neck cancer. Boraska Jelavić et al. 2006 and Pimentel-Nunes et al. 2013 reveal that rs4986790 may alter the risk to colorectal cancer but Landi et al. 2006 and Guo et al. 2006 display a lack of association. The TLR4 rs11536898 was significant associated with colon cancer (Slattery et al. 2012). Türe-Ozdemir et al. 2008 did not found association with rs4986790 and gastric mucosa associated Lymphoid Tissue Lymphoma. Some polymorphisms rs11536889, rs10759932, rs1927911 and rs11536858 showed significant association with risk of prostate cancer (Zheng et al. 2004, Kim et al. 2012a, Cheng et al. 2007, Chen et al. 2005, Song et al. 2009). Lindström et al. 2010 and Breyer et al. 2009 did not observe risk of prostate cancer with TLR4. In computational analysis, rs4986790 was classified as damaging by two different servers (SDM and SIFT). Whereas rs4986791 was considered as benign for all of the methods tested. There are some discrepancies between various studies that can be possible because there will be exist certain differences in ascertainment, study design, and populations, as well as limited sample size. Replication of findings in independent populations is a widely accepted method to exclude false-positive associations in genetic association studies. Unfortunately, it is complicated to have a conclusion about those SNPs and the development of these cancers. There are increasing evidences suggesting that TLR4 is also expressed in many tumors. They are expressed in neuroblastoma cells, follicular thyroid neoplasms, lung cancer cells, meningiomas, melanomas, pancreatic adenocarcinoma, follicular thyroid neoplasms, breast cancer cells, prostate cancer,

Head and Neck cell lines, gastric dysplasia and carcinoma, intestinal metaplasia and dysplasia, colorectal neoplasia and cervical neoplasia (Hassan et al. 2006, Hagström et al. 2012, He et al. 2007, Zhang et al. 2009, Tichomirowa et al. 2008, Molteni et al. 2006, Zhang 2010, Yang et al. 2010, González-Reyes et al. 2011b, González-Reyes et al. 2010b, Szczepanski et al. 2009, Pimentel-Nunes et al. 2011, Schmausser et al. 2005, Fukata et al. 2007, Yu et al. 2010).

The exams conducted for TLR5 showed that the quantity of polymorphism damaging for this protein were 24. In a case-control study, Zeng et al. 2011 evaluated 9 different polymorphisms of TLR5 rs759303, rs1773766, rs5744135, rs851181, rs45528236, rs2072493, rs5744174, rs2302597 and the allele –889T>C. They found an association only with rs5744174 with risk of gastric cancer. We analysed the missenses polymorphisms by computational methods, and none of those were considered damaging by all servers used. An independent case–control study including a more appropriate control group must be realized to other racial and ethnic groups. Immunohistochemical staining shows that TLR5 expression cervical neoplasia and gastric dysplasia (Kim et al. 2008, Pimentel-Nunes et al. 2011).

Investigations about the prediction of the variations in TLR6 exhibited that 25 nsSNPs could be deleterious. An investigation in the TLR6 gene did not found association with prostate cancer (Breyer et al. 2009). An analysis using a panel with 9412 SNPs from 1253 genes in a study of case-control with patients with non-Hodgkin lymphoma show that the TLR6 rs5743815 may be important in this cancer (Cerhan et al. 2007). When verified the prediction of this polymorphism by computational tools, the results evidence that it is benign to the protein. Further analysis of this variation is required.

Evaluations of TLR7 showed that 15 variations can possibly prejudice the protein. The analyses of functional prediction for TLR8 found only 8 SNPs predicted as damaging. An investigation made by Monroy et al. 2011 evaluated a missense SNP rs179008 and the risk of Hodgkin Disease. This SNP can change from a Glutamine to a Leucine or Proline. When analyzed by this work, these amino acid alterations did not show to be deleterious to the protein. These results strengthens the findings of Monroy et al that this polymorphism is benign and do not modify the activity of the protein. Moreover, TLR7 and TLR8 expression were associated with tumor progression in colorectal cancer (Grimm et al. 2010).

In TLR9 the nsSNPs considered as potential dangerous for the protein for all computational methods used were computed as 23. The investigation about prostate cancer mortality conducted in a Sweden population show that the TLR9 rs187084 was statistically significant for this malignancy (Stark et al. 2009). Analysis of the TLR9 showed that rs5743836 was significantly associated as a risk factor for the development of premalignant gastric changes (Ng et al. 2010). In contrast, (Hold 2010) an association with this polymorphism and increased risk of gastric cancer. TLR9 polymorphisms, rs5743836 and rs187084 was associated with colorectal cancer. The rs352140 did not exhibit correlation with bladder cancer susceptibility (Singh et al. 2013). The parsing of TLR9 demonstrated a lack of association between rs352140 and the risk of developing cervical cancer in a population from North India (Pandey et al. 2011). However, a logistic regression analysis showed that two polymorphisms, rs187084 and rs352140 may be a genetic risk factor for cervical cancer in a Polish population (Roszak et al. 2012). A research accomplished in a Cantonese population ascertains that the allele C for the SNP 829A/C increased the overall risk of nasopharyngeal carcinoma. Expression studies indicated that TLR9

show expression by breast, prostate cancer cells, cervical, gastric carcinoma and lung cancer cells (González-Reyes et al. 2010a, González-Reyes et al. 2011B, Lee et al. 2007, Schmausser et al. 2005, He et al. 2007, Zhang et al. 2009, Droemann et al. 2005).

A total of 22 of polymorphism were considered injurious in TLR10. Two SNPs could be associated with the reduced risk of prostate cancer, rs11096955 and rs11096957 (Stevens et al. 2008) but not in a caucasian population (Breyer et al. 2009). The rs11096955 was considered benign by all bioinformatic tools and the rs11096957, was evaluated as benign just by Polyphen. According the findings of Stevens et al these missense variations did not cause modification in the functionality of the protein and may, therefore, reduce the risk of prostate cancer. The missense polymorphism rs11466653 was associated with papillary thyroid carcinoma in the gene of TLR10 and may be a risk factor for the development of tumors in the Korean population (Kim et al. 2013). Computational analyses also were made by this work and they saw that this polymorphism were measured as damaging by SIFT, Polyphen and SNPs3D as the same as ours results.

The mainly interest in the field of medical research is to distinguish mutations that are functionally neutral from those that contribute to disease. The evaluation of the polymorphisms of TLRs and functional assessments is needed to address the question of cancer predisposition. Conducting large-scale studies is very difficult because of the vast amount of biases like the scarcity of data, the incorrect interpretation of the results, the technical level applied, among others. The use of computational tools to discover deleterious nsSNPs out of a pool of all the SNPs will be very useful for epidemiological population based studies even as for the analyze of single nucleotide variants and their prioritization for experimental characterization.

The methods used to predict the consequences of the variation on a protein uses different principles, thus, the combination of them can be useful to give a single prediction with improved accuracy. Here, for the most accurate of the polymorphisms that could be involved in the development of a disease were obtained when considering the mutations for which all four or three (in the case of lack of the SDM) programs predicted the same result. We show an amount of polymorphisms through prediction tools that can be harmful for the protein (see table 1).

Other researches such as developed by Chan P. A. et al 2007 used computational methods to evaluate mutations. They compared the methods of SIFT, Polyphen and A-GVGD in predicting the consequences of 254 missense variants in five different genes (CDKN2A, MLH1, MSH2, MECP2 and TYR). The results show that the use of these methods that use evolutionary sequence conservation, with or without considering protein structural change, to predict the clinical consequences of missense variants is useful for to interpret genes that are involved in human diseases. Even as, Masoodi et al. 2012, screened a pool of deleterious Non-Synonymous SNPs of the ePHA2 gene that was expect to be involved in the susceptibility of cataract formation. Using I-Mutant to calculate the protein stability change, SIFT and Polyphen to predict the potentially functional nsSNPs and their effect on protein that was found, where SNP rs2291806 could have functional impact on the EPHA2 gene. Valdmanis P N 2009 and coworkers using SIFT, Polyphen and Panther, saw the profile of benign and damaging changes in genes of neurodegenerative disease. An assay conducted for 5,500 genes carrying nearly 24,000 nsSNPs they show that the prediction tools are able to discern from mutations able to cause a disease and from those which could have a neutral effect. They saw

that rare mutations are consistently predicted to be deleterious as often as commonly occurring nsSNPs (Burke et al. 2007).

Using four high quality bioinformatics tools, we could select from a large pool of variations the candidates that are interesting for detailed investigation for the gene of the TLR family, similarly the evaluation of SNPs that were considered associated with the risk or protection of cancer. This research has a valuable contribution because the amount of variation data available for this protein evaluated here may be hereafter investigated experimentally as well as help to disclose somewhat the conflicts that exist in the literature about these genes and the predisposition of development of cancer and provided a firm basis for the clinical use of computational methods as one tool to help interpret the huge quantity of polymorphisms.

ACKNOWLEDGMENTS

This study was supported by grants for Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and developed on Laboratório de Imunopatologia Keizo Asami (LIKA).

REFERENCES

- Achyut BR, Ghoshal UC, Moorchung N, Mittal B. 2007. Association of Toll-like receptor-4 (Asp299Gly and Thr399Ileu) gene polymorphisms with gastritis and precancerous lesions. *Hum. Immunol.* 68: 901–7.
- Adzhubei I a, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. *Nat. Methods* 7: 248–9.
- Anton G Kutikhin Arseniy E Yuzhalin. 2012. Are Toll-like receptor gene polymorphisms associated with prostate cancer? 23–29.
- Ashton K a, Proietto A, Otton G, Symonds I, McEvoy M, Attia J, Scott RJ. 2010. Toll-like receptor (TLR) and nucleosome-binding oligomerization domain (NOD) gene polymorphisms and endometrial cancer risk. *BMC Cancer* 10: 382.

- Bergmann C, Bachmann HS, Bankfalvi A, Lotfi R, Pütter C, Wild CA, Schuler PJ, Greve J, Hoffmann TK, Lang S, Scherag A, Lehnerdt GF. 2011. Toll-like receptor 4 single-nucleotide polymorphisms Asp299Gly and Thr399Ile in head and neck squamous cell carcinomas. *J. Transl. Med.* 9: 139.
- Boraska Jelavić T, Barisić M, Drmic Hofman I, Boraska V, Vrdoljak E, Peruzović M, Hozo I, Puljiz Z, Terzić J. 2006. Microsatellite GT polymorphism in the toll-like receptor 2 is associated with colorectal cancer. *Clin. Genet.* 70: 156–60.
- Breyer JP, McReynolds KM, Yaspan BL, Bradley KM, Dupont WD, Smith JR. 2009. Genetic variants and prostate cancer risk: candidate replication and exploration of viral restriction genes. *Cancer Epidemiol. Biomarkers Prev.* 18: 2137–44.
- Burke DF, Worth CL, Priego E-M, Cheng T, Smink LJ, Todd J a, Blundell TL. 2007. Genome bioinformatic analysis of nonsynonymous SNPs. *BMC Bioinformatics* 8: 301.
- Castro F a, Försti a, Buch S, Kalthoff H, Krauss C, Bauer M, Egberts J, Schniewind B, Broering DC, Schreiber S, Schmitt M, Hampe J, et al. 2011. TLR-3 polymorphism is an independent prognostic marker for stage II colorectal cancer. *Eur. J. Cancer* 47: 1203–10.
- Cerhan JR, Ansell SM, Fredericksen ZS, Kay NE, Liebow M, Call TG, Dogan A, Cunningham JM, Wang AH, Liu W, Macon WR, Jelinek D, et al. 2007. Genetic variation in 1253 immune and inflammation genes and risk of nonHodgkin lymphoma. *110: 1–7.*
- Chan P. A. et al. 2007. Interpreting Missense Variants: Comparing Computational Methods in Human Disease Genes CDKN2A, MLH1, MSH2, MECP2 and Tyrosinase (TYR). *28: 683–693.*
- Chen Y-C, Giovannucci E, Lazarus R, Kraft P, Ketkar S, Hunter DJ. 2005. Sequence variants of Toll-like receptor 4 and susceptibility to prostate cancer. *Cancer Res.* 65: 11771–8.
- Cheng P-L, Eng H-L, Chou M-H, You H-L, Lin T-M. 2007. Genetic polymorphisms of viral infection-associated Toll-like receptors in Chinese population. *Transl. Res.* 150: 311–8.
- Cheng I, Plummer SJ, Casey G, Witte JS. 2007. Toll-like receptor 4 genetic variation and advanced prostate cancer risk. *Cancer Epidemiol. Biomarkers Prev.* 16: 352–5.
- Dalca A V, Brudno M. 2010. Genome variation discovery with high-throughput sequencing data. *Brief. Bioinform.* 11: 3–14.
- Davoodi H, Seow HF. 2011. Variant Toll-like receptor4 (Asp299Gly and Thr399Ile alleles) and Toll-like receptor2 (Arg753Gln and Arg677Trp alleles) in colorectal cancer. *Iran. J. Allergy. Asthma. Immunol.* 10: 91–9.
- Droemann D, Albrecht D, Gerdes J, Ulmer AJ, Branscheid D, Vollmer E, Dalhoff K, Zabel P, Goldmann T. 2005. Human lung cancer cells express functionally active Toll-like receptor 9. *Respir. Res.* 6: 1.
- Fukata M, Chen A, Vamadevan AS, Cohen J, Breglio K, Krishnareddy S, Hsu D, Xu R, Harpaz N, Dannenberg AJ, Subbaramaiah K, Cooper HS, et al. 2007. Toll-like receptor-4 promotes the development of colitis-associated colorectal tumors. *Gastroenterology* 133: 1869–81.

- Garza-Gonzalez E, Bosques-Padilla FJ, Mendoza-Ibarra SI, Flores-Gutierrez JP, Maldonado-Garza HJ, Perez-Perez GI. 2007. Assessment of the toll-like receptor 4 Asp299Gly, Thr399Ile and interleukin-8 -251 polymorphisms in the risk for the development of distal gastric cancer. *BMC Cancer* 7: 70.
- Gnad F, Baucom A, Mukhyala K, Manning G, Zhang Z. 2013. Assessment of computational methods for predicting the effects of missense mutations in human cancers. *BMC Genomics* 14 Suppl 3: S7.
- González-Reyes S, Fernández JM, González LO, Aguirre A, Suárez A, González JM, Escaff S, Vizoso FJ. 2011a. Study of TLR3, TLR4, and TLR9 in prostate carcinomas and their association with biochemical recurrence. *Cancer Immunol. Immunother.* 60: 217–26.
- González-Reyes S, Fernández JM, González LO, Aguirre A, Suárez A, González JM, Escaff S, Vizoso FJ. 2011b. Study of TLR3, TLR4, and TLR9 in prostate carcinomas and their association with biochemical recurrence. *Cancer Immunol. Immunother.* 60: 217–26.
- González-Reyes S, Marín L, González L, González LO, Casar JM del, Lamelas ML, González-Quintana JM, Vizoso FJ. 2010a. Study of TLR3, TLR4 and TLR9 in breast carcinomas and their association with metastasis. *BMC Cancer* 10: 665.
- González-Reyes S, Marín L, González L, González LO, Casar JM del, Lamelas ML, González-Quintana JM, Vizoso FJ. 2010b. Study of TLR3, TLR4 and TLR9 in breast carcinomas and their association with metastasis. *BMC Cancer* 10: 665.
- Grimm M, Kim M, Rosenwald A, Heemann U, Germer C-T, Waaga-Gasser AM, Gasser M. 2010. Toll-like receptor (TLR) 7 and TLR8 expression on CD133+ cells in colorectal cancer points to a specific role for inflammation-induced TLRs in tumourigenesis and tumour progression. *Eur. J. Cancer* 46: 2849–57.
- Guo Q, Zhu J, Xia B. 2006. Polymorphism of CD14 gene but not the mutation of TLR4 gene is associated with colorectal cancer in Chinese patients. *J. Gastroenterol. Hepatol.* 21: 92–7.
- Hagström J, Heikkilä A, Siironen P, Louhimo J, Heiskanen I, Mäenpää H, Arola J, Haglund C. 2012. TLR-4 expression and decrease in chronic inflammation: indicators of aggressive follicular thyroid carcinoma. *J. Clin. Pathol.* 65: 333–8.
- Hassan F, Islam S, Tumurkhuu G, Naiki Y, Koide N, Mori I, Yoshida T, Yokochi T. 2006. Intracellular expression of toll-like receptor 4 in neuroblastoma cells and their unresponsiveness to lipopolysaccharide. *BMC Cancer* 6: 281.
- He W, Liu Q, Wang L, Chen W, Li N, Cao X. 2007. TLR4 signaling promotes immune escape of human lung cancer cells by inducing immunosuppressive cytokines and apoptosis resistance. *Mol. Immunol.* 44: 2850–9.
- Hishida A, Matsuo K, Goto Y, Mitsuda Y, Hiraki A, Naito M, Wakai K, Tajima K, Hamajima N. 2009. Toll-like receptor 4 +3725 G/C polymorphism, *Helicobacter pylori* seropositivity, and the risk of gastric atrophy and gastric cancer in Japanese. *Helicobacter* 14: 47–53.
- Hold GL. 2010. CD14-159C/T and TLR9-1237T/C polymorphisms are not associated with gastric cancer risk in Caucasian populations. 18: 117–119.

- Hold GL, Rabkin CS, Chow W-H, Smith MG, Gammon MD, Risch HA, Vaughan TL, McColl KEL, Lissowska J, Zatonski W, Schoenberg JB, Blot WJ, et al. 2007. A functional polymorphism of toll-like receptor 4 gene increases risk of gastric carcinoma and its precursors. *Gastroenterology* 132: 905–12.
- Huang H, Wu J, Jin G, Zhang H, Ding Y, Hua Z, Zhou Y, Xue Y, Lu Y, Hu Z, Xu Y, Shen H. 2010. A 5'-flanking region polymorphism in toll-like receptor 4 is associated with gastric cancer in a Chinese population. *J. Biomed. Res.* 24: 100–6.
- Junjie X, Songyao J, Minmin S, Yanyan S, Baiyong S, Xiaxing D, Jiabin J, Xi Z, Hao C. 2012. The association between Toll-like receptor 2 single-nucleotide polymorphisms and hepatocellular carcinoma susceptibility. *BMC Cancer* 12: 57.
- Kelly MG, Alvero AB, Chen R, Silasi D-A, Abrahams VM, Chan S, Visintin I, Rutherford T, Mor G. 2006. TLR-4 signaling promotes tumor growth and paclitaxel chemoresistance in ovarian cancer. *Cancer Res.* 66: 3859–68.
- Kim HJ, Bae JS, Chang IH, Kim K Do, Lee J, Shin HD, Lee JY, Kim W-J, Kim W, Myung SC. 2012a. Sequence variants of Toll-like receptor 4 (TLR4) and the risk of prostate cancer in Korean men. *World J. Urol.* 30: 225–32.
- Kim KH, Jo MS, Suh DS, Yoon MS, Shin DH, Lee JH, Choi KU. 2012b. Expression and significance of the TLR4/MyD88 signaling pathway in ovarian epithelial cancers. *World J. Surg. Oncol.* 10: 193.
- Kim MK, Park SW, Kim SK, Park HJ, Eun YG, Kwon KH, Kim J. 2012c. Association of Toll-like receptor 2 polymorphisms with papillary thyroid cancer and clinicopathologic features in a Korean population. *J. Korean Med. Sci.* 27: 1333–8.
- Kim SK, Park HJ, Hong IK, Chung J-H, Eun YG. 2013. A missense polymorphism (rs11466653, Met326Thr) of toll-like receptor 10 (TLR10) is associated with tumor size of papillary thyroid carcinoma in the Korean population. *Endocrine* 43: 161–9.
- Kim YS, Hwang YJ, Kim SY, Yang SM, Lee KY, Park IB. 2008. Rarity of TLR4 Asp299Gly and Thr399Ile polymorphisms in the Korean population. *Yonsei Med. J.* 49: 58–62.
- Koboldt DC, Ding L, Mardis ER, Wilson RK. 2010. Challenges of sequencing human genomes. *Brief. Bioinform.* 11: 484–98.
- Kumar H, Kawai T, Akira S. 2009. Toll-like receptors and innate immunity. *Biochem. Biophys. Res. Commun.* 388: 621–5.
- Landi S, Gemignani F, Bottari F, Gioia-Patricola L, Guino E, Cambray M, Biondo S, Capella G, Boldrini L, Canzian F, Moreno V. 2006. Polymorphisms within inflammatory genes and colorectal cancer. *J. Negat. Results Biomed.* 5: 15.
- Lee J, Choi J, Seo ES, Kim MJ, Kim WY, Choi CH, Kim T, Kim B, Song SY, Bae D. 2007. Increased Toll-Like Receptor 9 Expression in Cervical Neoplasia. 947: 941–947.
- Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26: 589–95.

- Lindström S, Hunter DJ, Grönberg H, Stattin P, Wiklund F, Xu J, Chanock SJ, Hayes R, Kraft P. 2010. Sequence variants in the TLR4 and TLR6-1-10 genes and prostate cancer risk. Results based on pooled analysis from three independent studies. *Cancer Epidemiol. Biomarkers Prev.* 19: 873–6.
- Masoodi TA, Shammari S a Al, Al-Muammar MN, Almubrad TM, Alhamdan A a. 2012. Screening and structural evaluation of deleterious Non-Synonymous SNPs of ePHA2 gene involved in susceptibility to cataract formation. *Bioinformation* 8: 562–7.
- Minmin S, Xiaoqian X, Hao C, Baiyong S, Xiaxing D, Junjie X, Xi Z, Jianquan Z, Songyao J. 2011. Single nucleotide polymorphisms of Toll-like receptor 4 decrease the risk of development of hepatocellular carcinoma. *PLoS One* 6: e19466.
- Mirnezami Reza NJ and DA. 2012. Preparing for Precision Medicine. 2012–2014.
- Molteni M, Marabella D, Orlandi C, Rossetti C. 2006. Melanoma cell lines are responsive in vitro to lipopolysaccharide and express TLR-4. *Cancer Lett.* 235: 75–83.
- Monroy CM, Cortes AC, Lopez MS, D'Amelio AM, Etzel CJ, Younes A, Strom SS, El-Zein RA. 2011. Hodgkin disease risk: role of genetic polymorphisms and gene-gene interactions in inflammation pathway genes. *Mol. Carcinog.* 50: 36–46.
- Ng MTH, Van't Hof R, Crockett JC, Hope ME, Berry S, Thomson J, McLean MH, McColl KEL, El-Omar EM, Hold GL. 2010. Increase in NF-kappaB binding affinity of the variant C allele of the toll-like receptor 9 -1237T/C polymorphism is associated with *Helicobacter pylori*-induced gastric disease. *Infect. Immun.* 78: 1345–52.
- Ohara T, Morishita T, Suzuki H, Hibi T. 2006. Heterozygous Thr 135 Ala polymorphism at leucine-rich repeat (LRR) in genomic DNA of toll-like receptor 4 in patients with poorly-differentiated gastric adenocarcinomas. *Int. J. Mol. Med.* 18: 59–63.
- Oliveira JG de, Silva AE. 2012. Polymorphisms of the TLR2 and TLR4 genes are associated with risk of gastric cancer in a Brazilian population. *World J. Gastroenterol.* 18: 1235–42.
- Pandey S, Mittal B, Srivastava M, Singh S, Srivastava K, Lal P, Mittal RD. 2011. Evaluation of Toll-like receptors 3 (c.1377C/T) and 9 (G2848A) gene polymorphisms in cervical cancer susceptibility. *Mol. Biol. Rep.* 38: 4715–21.
- Pandey S, Mittal RD, Srivastava M, Srivastava K, Singh S, Srivastava S, Mittal B. 2009. Impact of Toll-like receptors [TLR] 2 (-196 to -174 del) and TLR 4 (Asp299Gly, Thr399Ile) in cervical cancer susceptibility in North Indian women. *Gynecol. Oncol.* 114: 501–5.
- Pimentel-Nunes P, Afonso L, Lopes P, Roncon-Albuquerque R, Gonçalves N, Henrique R, Moreira-Dias L, Leite-Moreira AF, Dinis-Ribeiro M. 2011. Increased expression of toll-like receptors (TLR) 2, 4 and 5 in gastric dysplasia. *Pathol. Oncol. Res.* 17: 677–83.
- Pimentel-Nunes P, Teixeira AL, Pereira C, Gomes M, Brandão C, Rodrigues C, Gonçalves N, Boal-Carvalho I, Roncon-Albuquerque R, Moreira-Dias L, Leite-Moreira AF, Medeiros R, et al. 2013. Functional polymorphisms of Toll-like receptors 2 and 4 alter the risk for colorectal carcinoma in Europeans. *Dig. Liver Dis.* 45: 63–9.

- Purdue MP, Lan Q, Wang SS, Kricker A, Menashe I, Zheng T-Z, Hartge P, Grulich AE, Zhang Y, Morton LM, Vajdic CM, Holford TR, et al. 2009. A pooled investigation of Toll-like receptor gene variants and risk of non-Hodgkin lymphoma. *Carcinogenesis* 30: 275–81.
- Reva B, Antipin Y, Sander C. 2011. Predicting the functional impact of protein mutations: application to cancer genomics. 1–14.
- Rigoli L, Bella CDI, Fedele F, Procopio V, Amorini M, Giudice GLO, Romeo P, Pugliatti F, Finocchiaro G, Caruso RA. 2010. TLR4 and NOD2/CARD15 Genetic Polymorphisms and their Possible Role in Gastric Carcinogenesis. 518: 513–517.
- Roszak A, Lianeri M, Sowińska A, Jagodziński PP. 2012. Involvement of Toll-like Receptor 9 polymorphism in cervical cancer development. *Mol. Biol. Rep.* 39: 8425–30.
- Santini D, Angeletti S, Ruzzo a, Dicuonzo G, Galluzzo S, Vincenzi B, Calvieri a, Pizzagalli F, Graziano N, Ferraro E, Lorino G, Altomare a, et al. 2008. Toll-like receptor 4 Asp299Gly and Thr399Ile polymorphisms in gastric cancer of intestinal and diffuse histotypes. *Clin. Exp. Immunol.* 154: 360–4.
- Schmausser B, Andrulis M, Endrich S, Müller-Hermelink H-K, Eck M. 2005. Toll-like receptors TLR4, TLR5 and TLR9 on gastric carcinoma cells: an implication for interaction with *Helicobacter pylori*. *Int. J. Med. Microbiol.* 295: 179–85.
- Singh V, Srivastava N, Kapoor R, Mittal RD. 2013. Single-nucleotide polymorphisms in genes encoding toll-like receptor -2, -3, -4, and -9 in a case-control study with bladder cancer susceptibility in a North Indian population. *Arch. Med. Res.* 44: 54–61.
- Slattery ML, Herrick JS, Bondurant KL, Wolff RK. 2012. Toll-like receptor genes and their association with colon and rectal cancer development and prognosis. *Int. J. Cancer* 130: 2974–80.
- Song C, Chen L-Z, Zhang R-H, Yu X-J, Zeng Y-X. 2006. Functional variant in the 3'-untranslated region of Toll-like receptor 4 is associated with nasopharyngeal carcinoma risk. *Cancer Biol. Ther.* 5: 1285–91.
- Song J, Kim DY, Kim CS, Kim HJ, Lee DH, Lee HM, Ko W, Lee G. 2009. The association between Toll-like receptor 4 (TLR4) polymorphisms and the risk of prostate cancer in Korean men. *Cancer Genet. Cytogenet.* 190: 88–92.
- Srivastava K, Srivastava A, Kumar A, Mittal B. 2010. Significant association between toll-like receptor gene polymorphisms and gallbladder cancer. *Liver Int.* 30: 1067–72.
- Stark JR, Wiklund F, Grönberg H, Schumacher F, Sinnott J a, Stampfer MJ, Mucci L a, Kraft P. 2009. Toll-like receptor signaling pathway variants and prostate cancer mortality. *Cancer Epidemiol. Biomarkers Prev.* 18: 1859–63.
- Stevens VL, Hsing AW, Talbot JT, Zheng SL, Sun J, Chen J, Thun MJ, Xu J, Calle EE, Rodriguez C. 2008. Genetic variation in the toll-like receptor gene cluster (TLR10-TLR1-TLR6) and prostate cancer risk. *Int. J. Cancer* 123: 2644–50.
- Sun J, Wiklund F, Zheng SL, Chang B, Bälter K, Li L, Johansson J-E, Li G, Adami H-O, Liu W, Tolin A, Turner AR, et al. 2005a. Sequence variants in Toll-like receptor gene cluster (TLR6-TLR1-TLR10) and prostate cancer risk. *J. Natl. Cancer Inst.* 97: 525–32.

- Sun J, Wiklund F, Zheng SL, Chang B, Li L, Johansson J, Li G, Adami H, Liu W, Tolin A, Turner AR, Meyers DA, et al. 2005b. Sequence Variants in Toll-Like Receptor Gene Cluster (TLR6-TLR1-TLR10) and Prostate Cancer Risk. 97: 525–532.
- Suo S-B, Qiu J-D, Shi S-P, Chen X, Huang S-Y, Liang R-P. 2013. Proteome-wide analysis of amino acid variations that influence protein lysine acetylation. *J. Proteome Res.* 12: 949–58.
- Szczepanski MJ, Czystowska M, Szajnik M, Harasymczuk M, Boyiadzis M, Kruk-Zagajewska A, Szyfter W, Zeromski J, Whiteside TL. 2009. Triggering of Toll-like receptor 4 expressed on human head and neck squamous cell carcinoma promotes tumor development and protects the tumor from immune attack. *Cancer Res.* 69: 3105–13.
- Theodoropoulos GE, Saridakis V, Karantanos T, Michalopoulos N V, Zagouri F, Kontogianni P, Lymperi M, Gazouli M, Zografos GC. 2012. Toll-like receptors gene polymorphisms may confer increased susceptibility to breast cancer development. *Breast* 21: 534–8.
- Tichomirowa M a, Theodoropoulou M, Daly AF, Yassouridis A, Hansen S, Lu J, Lange M, Goldbrunner RH, Stalla GK, Renner U. 2008. Toll-like receptor-4 is expressed in meningiomas and mediates the antiproliferative action of paclitaxel. *Int. J. Cancer* 123: 1956–63.
- Trejo-de la O A, Torres J, Pérez-Rodríguez M, Camorlinga-Ponce M, Luna LF, Abdo-Francis JM, Lazcano E, Maldonado-Bernal C. 2008. TLR4 single-nucleotide polymorphisms alter mucosal cytokine and chemokine patterns in Mexican patients with *Helicobacter pylori*-associated gastroduodenal diseases. *Clin. Immunol.* 129: 333–40.
- Türe-Ozdemir F, Gazouli M, Tzivras M, Panagos C, Bovaretos N, Petraki K, Giannakopoulos A, Korkolopoulou P, Mantzaris GJ. 2008. Association of polymorphisms of NOD2, TLR4 and CD14 genes with susceptibility to gastric mucosa-associated lymphoid tissue lymphoma. *Anticancer Res.* 28: 3697–700.
- Worth CL, Preissner R, Blundell TL. 2011. SDM--a server for predicting effects of mutations on protein stability and malfunction. *Nucleic Acids Res.* 39: W215–22.
- Yang H, Zhou H, Feng P, Zhou X, Wen H, Xie X, Shen H, Zhu X. 2010. Reduced expression of Toll-like receptor 4 inhibits human breast cancer cells proliferation and inflammatory cytokines secretion. *J. Exp. Clin. Cancer Res.* 29: 92.
- Yu L, Wang L, Li M, Zhong J, Wang Z, Chen S. 2010. Expression of toll-like receptor 4 is down-regulated during progression of cervical neoplasia. *Cancer Immunol. Immunother.* 59: 1021–8.
- Zeng H-M, Pan K-F, Zhang Y, Zhang L, Ma J-L, Zhou T, Su H-J, Li W-Q, Li J-Y, Gerhard M, Classen M, You W-C. 2011. Genetic variants of toll-like receptor 2 and 5, *helicobacter pylori* infection, and risk of gastric cancer and its precursors in a chinese population. *Cancer Epidemiol. Biomarkers Prev.* 20: 2594–602.
- Zhang J-J. 2010. Expression and significance of TLR4 and HIF-1 α in pancreatic ductal adenocarcinoma. *World J. Gastroenterol.* 16: 2881.

Zhang Y-B, He F-L, Fang M, Hua T-F, Hu B-D, Zhang Z-H, Cao Q, Liu R-Y. 2009. Increased expression of Toll-like receptors 4 and 9 in human lung cancer. *Mol. Biol. Rep.* 36: 1475–81.

Zhang ZD, Du J, Lam H, Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. Identification of genomic indels and structural variations using split reads. *BMC Genomics* 12: 375.

Zheng SL, Augustsson-Bälter K, Chang B, Hedelin M, Li L, Adami H-O, Bensen J, Li G, Johnsson J-E, Turner AR, Adams TS, Meyers D a, et al. 2004. Sequence variants of toll-like receptor 4 are associated with prostate cancer risk: results from the CAncer Prostate in Sweden Study. *Cancer Res.* 64: 2918–22.

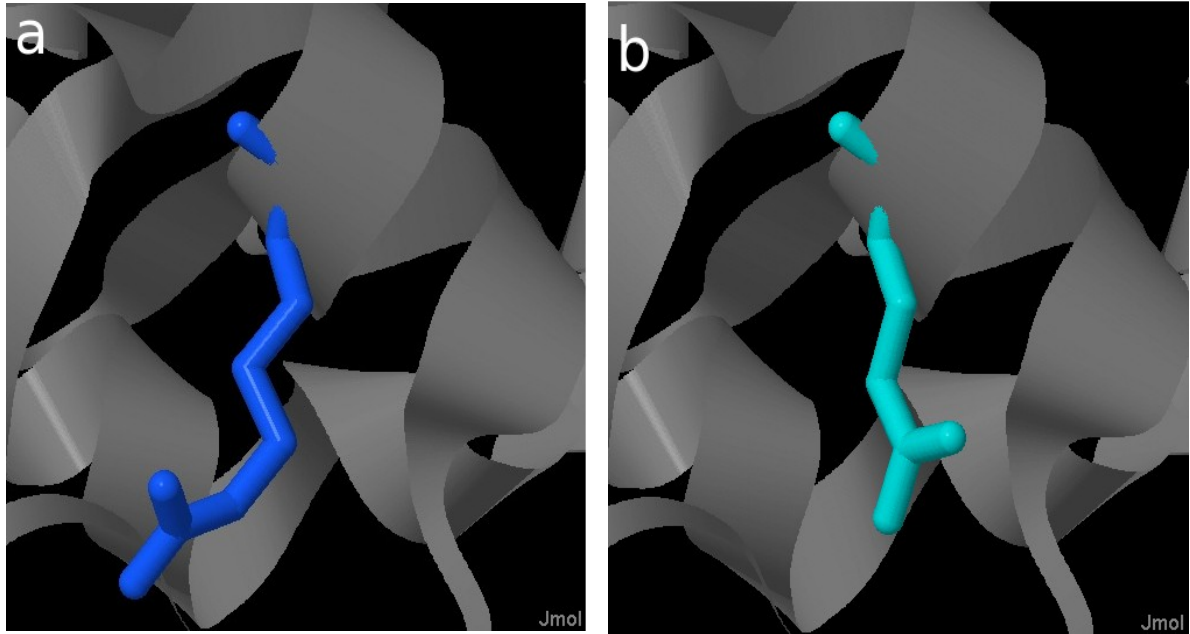


Figure 1 - a – residue of amino acid of wild protein. b- residue of amino acid of mutated protein.

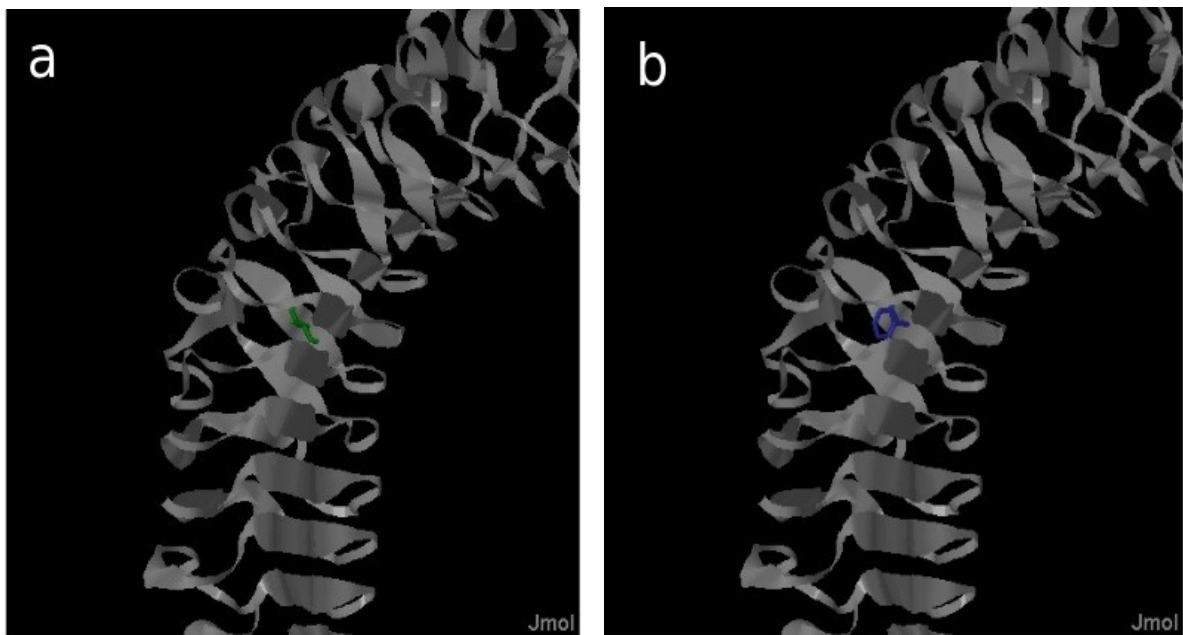


Figure 2 – a- residue of amino acid of wild protein. b- residue of amino acid of mutated protein.

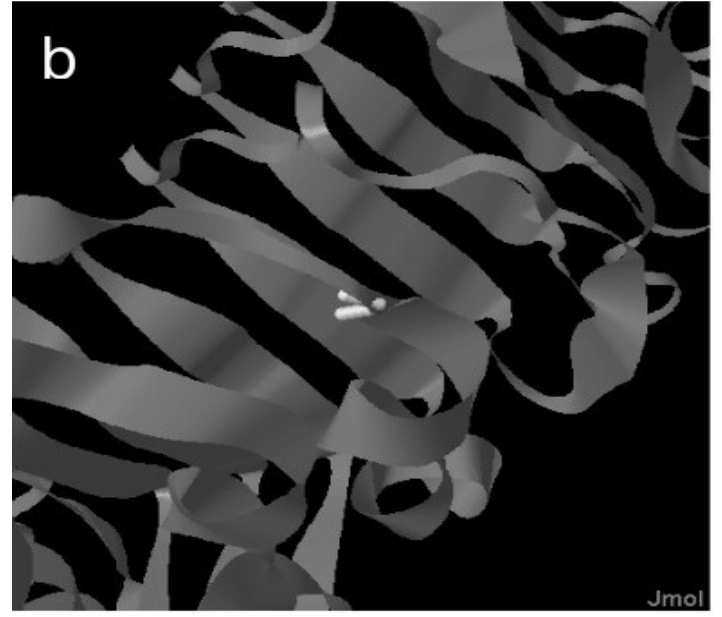


Figure 3 – a- residue of amino acid of wild protein. b- residue of amino acid of mutated protein.

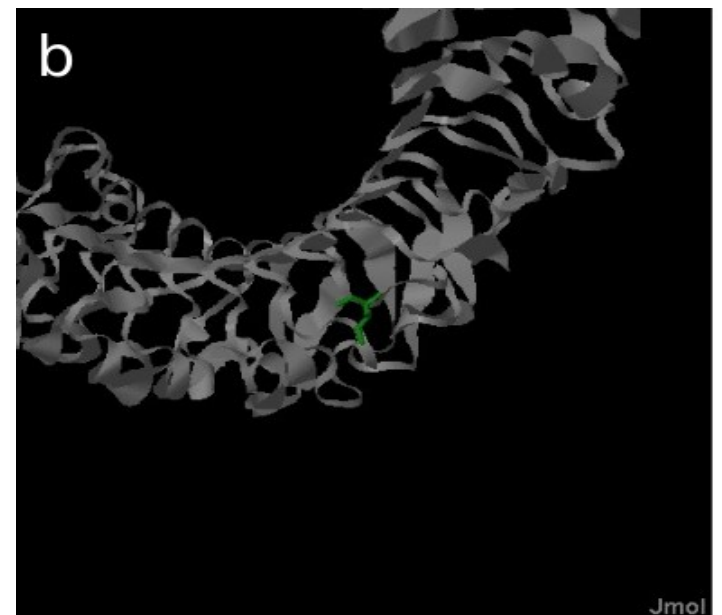


Figure 4 – a - residue of amino acid of wild protein. b- residue of amino acid of mutated protein.

Table 1. – Data mining summary of TLRs.

Gene	SNP	mRNA	Protein	Position	TOTAL	3' Near	3' UTR	5' Near	5' UTR	Frame shift	Intron	Non sense	Synony mous	Mis sense
TLR1	7096	NM_003263	NP_003254	4p14	399	10	7	70	8	1	141	7	44	111
TLR2	7097	NM_003264	NP_003255	4q32	607	8	10	42	10	2	402	7	40	86
TLR3	7098	NM_003265	NP_003256	4q35	462	11	4	44	5	1	276	1	52	68
TLR4	7099	NM_138554	NP_612564	9q33.1	672	20	216	34	24	1	139	8	73	157
TLR5	7100	NM_003268	NP_003259	1q41q-42	898	31	40	45	24	3	628	8	30	89
TLR6	10333	NM_006068	NP_006059	4p14	991	12	88	75	6	1	675	2	41	91
TLR7	51284	NM_016562	NP_057646	Xp22.3	604	11	114	33	6	1	296	2	63	78
TLR8	51311	NM_138636	NP_619542	Xp22	330	4	24	40	5	0	178	0	38	41
TLR9	54106	NM_017442	NP_059138	3p21.3	441	0	15	28	33	5	47	6	138	169
TLR10	81793	NM_030956	NP_112218	4p14	435	15	28	61	16	1	170	7	36	101
TOTAL					5839	122	546	472	137	16	2952	48	555	991

6. PATENTE

Título: Painel genético na detecção de genes inibidores da via de sinalização das proteínas Toll Like Recptor (TLR) para o tratamento do câncer

Número do depósito: BR 10 2013 027576 0

Data do depósito: 25/10/2013

Inventores: José Luiz de Lima Filho

Ana Carolina da Rocha Simões

Carlos Henrique Madeiros Castelletti

Danyelly Bruneska Gondim Martins



< Uso exclusivo do INPI >

Espaço reservado para o protocolo

Espaço reservado para a etiqueta

Espaço reservado para o código QR



INPI INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
Sistema de Gestão da Qualidade
Diretoria de Patentes

DIRPA

Tipo de Documento:

Formulário

DIRPA

Página:

1/3

Título do Documento:

Depósito de Pedido de Patente

Código:

FQ001

Versão:

2

Procedimento:

DIRPA-PQ006

Ao Instituto Nacional da Propriedade Industrial:

O requerente solicita a concessão de um privilégio na natureza e nas condições abaixo indicadas:

1. Depositante (71):

1.1 Nome: **José Luiz de Lima Filho**

1.2 Qualificação: **Médico**

1.3 CNPJ/CPF: **21638241449**

1.4 Endereço Completo: **Av Prof. Moraes Rego, s/n**

1.5 CEP: **50670-901**

1.6 Telefone: **81 2126 8484**

1.7 Fax: **81 2126 8485**

1.8 E-mail: **joseluiz60@gmail.com**

☐ continua em folha anexa

2. Natureza:

☒ Invenção

☐ Modelo de Utilidade

☐ Certificado de Adição

3. Título da Invenção ou Modelo de Utilidade (54):

PAINEL GENÉTICO NA DETECÇÃO DE GENES INIBIDORES DA VIA DE SINALIZAÇÃO DAS PROTEÍNAS TOLL LIKE RECEPTOR (TLR) PARA O TRATAMENTO DO CÂNCER

☐ continua em folha anexa

4. Pedido de Divisão: do pedido N°

Data de Depósito:

5. Prioridade:

☐ Interna (66)

☐ Unionista (30)

O depositante reivindica a(s) seguinte(s):

Pais ou Organização do depósito	Número do depósito (se disponível)	Data de depósito

☐ continua em folha anexa



INPI INSTITUTO
NACIONAL
DA PROPRIEDADE
INDUSTRIAL

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
Sistema de Gestão da Qualidade
Diretoria de Patentes

 Título do Documento:	Tipo de Documento:	DIRPA	Página:
	Formulário		3/3
Depósito de Pedido de Patente		Código:	Versão:
		FQ001	2
		Procedimento:	DIRPA-PQ006

11. Documentos Anexados:

(Assinale e indique também o número de folhas):

(Deverá ser indicado o número total de somente uma das vias de cada documento).

	Documentos Anexados	folhas
☒	11.1 Guia de Recolhimento da União (GRU).	1
☐	11.2 Procuração.	
☐	11.3 Documentos de Prioridade.	
☐	11.4 Documento de contrato de trabalho.	
☒	11.5 Relatório descritivo.	8
☒	11.6 Reivindicações.	1
☐	11.7 Desenho(s) (se houver). Sugestão de figura a ser publicada com o resumo: nº, _____ por melhor representar a invenção (sujeito à avaliação do INPI).	
☒	11.8 Resumo.	1
☐	11.9 Listagem de sequências em arquivo eletrônico: _____ nº de CDs ou DVDs (original e cópia).	
☐	11.10 Código de controle alfanumérico no formato de código de barras referente às listagem de sequências.	
☐	11.11 Listagem de sequências em formato impresso.	
☐	11.12 Declaração relativa à Listagem de sequências.	
☒	11.13 Outros (especificar) Lista de inventores	1

12. Total de folhas anexadas: 12 fls.

13. Declaro, sob as penas da Lei que todas as informações acima prestadas são completas e verdadeiras.

Local e Data

Assinatura e Carimbo

< Uso exclusivo do INPI >

Espaço reservado para o protocolo

Espaço reservado para a etiqueta

Espaço reservado para o código QR



INPI
INSTITUTO
NACIONAL
DA PROPRIEDADE
INDUSTRIAL

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
Sistema de Gestão da Qualidade
Diretoria de Patentes

DIRPA

Tipo de Documento:

Fomulário

DIRPA

Página:

1/3

Título do Documento:

Depósito de Pedido de Patente

Código:

FQ001

Versão:

2

Procedimento:

DIRPA-PQ006

Ao Instituto Nacional da Propriedade Industrial:

O requerente solicita a concessão de um privilégio na natureza e nas condições abaixo indicadas:

1. Depositante (71):

1.1 Nome: **José Luiz de Lima Filho**

1.2 Qualificação: **Médico**

1.3 CNPJ/CPF: **21638241449**

1.4 Endereço Completo: **Av Prof. Moraes Rego, s/n**

1.5 CEP: **50670-901**

1.6 Telefone: **81 2126 8484**

1.7 Fax: **81 2126 8485**

1.8 E-mail: **joseluiz60@gmail.com**

☐ continua em folha anexa

2. Natureza:

☒ Invenção

☐ Modelo de Utilidade

☐ Certificado de Adição

3. Título da Invenção ou Modelo de Utilidade (54):

**PAINEL GENÉTICO NA DETECÇÃO DE GENES INIBIDORES DA VIA DE
SINALIZAÇÃO DAS PROTEÍNAS TOLL LIKE RECEPTOR (TLR) PARA O
TRATAMENTO DO CÂNCER**

☐ continua em folha anexa

4. Pedido de Divisão: do pedido Nº

Data de Depósito:

5. Prioridade:

☐ Interna (66)

☐ Unionista (30)

O depositante reivindica a(s) seguinte(s):

País ou Organização do depósito	Número do depósito (se disponível)	Data de depósito

☐ continua em folha anexa



INPI INSTITUTO
NACIONAL
DA PROPRIEDADE
INDUSTRIAL

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
Sistema de Gestão da Qualidade
Diretoria de Patentes

DIRPA Título do Documento:	Tipo de Documento:	DIRPA	Página:
	Formulário		2/3
	Depósito de Pedido de Patente	Código: FQ001	Versão: 2
		Procedimento: DIRPA-PQ006	

6. **Inventor (72):**

☐ Assinale aqui se o(s) mesmo(s) requer(em) a não divulgação de seus nome(s), neste caso não preencher os campos abaixo.

6.1 Nome: **José Luiz de Lima Filho**

6.2 Qualificação: **Médico**

6.3 CPF: **21638241449**

6.4 Endereço Completo: **Av. Prof. Moraes Rego s/n**

6.5 CEP: **50670-901**

6.6 Telefone: **81 2126-8484**

6.7 FAX: **81 2126-8485**

6.8 E-mail: **joseluiz60@gmail.com**

☒ continua em folha anexa

7. **Declaração de divulgação anterior não prejudicial.**

Artigo 12 da LPI – período de graça.

Informe no item 11.13 os documentos anexados, se houver.

☐

8. **Declaração na forma do item 3.2 da Instrução Normativa PR nº 17/2013:**

☐ Declaro que os dados fornecidos no presente formulário são idênticos ao da certidão de depósito ou documento equivalente do pedido cuja prioridade está sendo reivindicada.

9. **Procurador (74):**

9.1 Nome:

9.2 CNPJ/CPF:

9.3 API/OAB:

9.4 Endereço Completo:

9.5 CEP:

9.6 Telefone:

9.7 FAX:

9.8 E-mail:

☐ continua em folha anexa

10. **Listagem de sequências biológicas.**

Informe nos itens 11.9 ao 11.12 os documentos anexados, se houver.

☐



INPI
INSTITUTO
NACIONAL
DA PROPRIEDADE
INDUSTRIAL

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
Sistema de Gestão da Qualidade
Diretoria de Patentes

DIRPA	Tipo de Documento: Formulário	DIRPA	Página: 3/3
Título do Documento: Depósito de Pedido de Patente		Código: FQ001	Versão: 2
		Procedimento: DIRPA-PQ006	

11. Documentos Anexados:

(Assinale e indique também o número de folhas):

(Deverá ser indicado o número total de somente uma das vias de cada documento).

	Documentos Anexados		folhas
☒	11.1	Guia de Recolhimento da União (GRU).	1
☐	11.2	Procuração.	
☐	11.3	Documentos de Prioridade.	
☐	11.4	Documento de contrato de trabalho.	
☒	11.5	Relatório descritivo.	8
☒	11.6	Reivindicações.	1
☐	11.7	Desenho(s) (se houver). Sugestão de figura a ser publicada com o resumo: nº, _____ por melhor representar a invenção (sujeito à avaliação do INPI).	
☒	11.8	Resumo.	1
☐	11.9	Listagem de sequências em arquivo eletrônico: _____ nº de CDs ou DVDs (original e cópia).	
☐	11.10	Código de controle alfanumérico no formato de código de barras referente às listagem de sequências.	
☐	11.11	Listagem de sequências em formato impresso.	
☐	11.12	Declaração relativa à Listagem de sequências.	
☒	11.13	Outros (especificar) Lista de inventores	1

12. Total de folhas anexadas: 12 fls.

13. Declaro, sob as penas da Lei que todas as informações acima prestadas são completas e verdadeiras.

Local e Data

Assinatura e Carimbo

6. Inventores:

6.1 Nome: Carolina da Rocha Simões
6.2 Qualificação: Bióloga
6.3 CPF: 05808640490
6.4 Endereço Completo: Av Morais Rego s/n
6.5 CEP: 50670-901
6.6 Telefone: 81 2126 8484
6.7 Fax: 81 2126 8485
6.8 Email: crocha@prospecmol.org

6.1 Nome: Carlos Henrique Madeiros Castelletti
6.2 Qualificação: Biólogo
6.3 CPF: 02107117446
6.4 Endereço Completo: Av Morais Rego s/n
6.5 CEP: 50670-901
6.6 Telefone: 81 2126 8484
6.7 Fax: 81 2126 8485
6.8 Email: hcastelletti@prospecmol.org

6.1 Nome: Danyelly Bruneska Gondim Martins
6.2 Qualificação: Biólogo
6.3 CPF: 89562453472
6.4 Endereço Completo: Av Morais Rego s/n
6.5 CEP: 50670-901
6.6 Telefone: 81 2126 8484
6.7 Fax: 81 2126 8485
6.8 Email: bruneska@prospecmol.org

PAINEL GENÉTICO NA DETECÇÃO DE GENES INIBIDORES DA VIA DE SINALIZAÇÃO DAS PROTEÍNAS TOLL LIKE RECEPTOR (TLR) PARA O TRATAMENTO DO CÂNCER

RELATÓRIO DESCRITIVO

Campo da invenção

A presente invenção refere-se ao campo métodos e dispositivos baseados em um painel genético para a identificação de marcadores moleculares relacionados à inibição e/ou regulação negativa das proteínas Toll Like Receptors (TLRs), não excluindo outras formas de regulação.

Antecedentes da invenção

As proteínas Toll Like Receptor (TLR) são uma classe de receptores de reconhecimento de patógenos denominadas de receptores de reconhecimento de padrão (PRRs). São glicoproteínas transmembranares tipo 1 que possuem três domínios: o extracelular que é formado por múltiplas repetições de leucina; o transmembranar; e o domínio intracelular que é homólogo a interleucina 1 humana. As TLR1, TLR2, TLR4, TLR6 e TLR10 estão usualmente localizadas na superfície celular, enquanto que as TLR3, TLR7, TLR8 e TLR9 estão na membrana do retículo endoplasmático ou na membrana do lisossomo ou endossomo.

São consideradas um link entre a imunidade inata e adaptativa e contribuem com o sistema imunológico a combater eficientemente os patógenos que invadem o organismo, desempenhando um papel importante na resposta inflamatória, surgindo como marcadores em diversos tipos de câncer e doenças, pois reconhecem diferentes tipos de ligantes. As TLR7, TLR8 e TLR9 reconhecem ácidos nucleicos derivados de vírus e bactérias; a TLR1 reconhece lipoproteína triacetilada; a TLR2 lipoproteínas bacterianas, ácido lipotecóico e zymosan de fungos; TLR3 reconhece dsRNA de vírus; já a TLR4 lipopolissacarídeo de bactérias Gram negativas; a TLR5 reconhece a flagelina. Até o momento nenhum ligante específico foi identificado para a TLR 10.

Como moléculas sensores, as TLRs detectam padrões moleculares associados a patógenos e se acopla a diferentes proteínas adaptadoras que desencadeiam específicas respostas nas vias de sinalização. Quando ativadas desencadeia-se uma cascata de sinalização que levará à ativação do NfκB, que regula a expressão de citocinas, quimioquinas, moléculas de adesão e moléculas co-estimulatórias que coordenam a resposta celular e humoral. Todas as TLRs, exceto a TLR3, fazem o recrutamento da proteína adaptadora MyD88 e as quinases IRAK1 e IRAK4. Para ocorrer a via de sinalização das TLR2 e 4, é necessária a proteína adaptadora TIRAP/MAL para ligar estas a MyD88. Existe uma outra via de sinalização que dar-se através das TLR3 e 4 que é independente do adaptador MyD88. Na TLR3 é realizada pelo adaptador TRIF e na TLR4, TRAM.

Vários tipos de cânceres surgem em locais que possuem inflamação crônica, onde esta resposta inflamatória é responsável pelo crescimento do tumor, gerando processos neoplásicos como proliferação, sobrevivência e migração. Várias evidências mostram a associação entre a superexpressão das TLRs e a progressão do câncer, assim como as moléculas da via de sinalização das TLRs estão envolvidas na progressão de tumores. Desta forma, as TLRs são consideradas como alvos terapêuticos em vários tipos de câncer e doenças.

A patente **WO2013087937 A1** intitulada “Treatment of cancer by inhibition of the myd88/erk map kinase interaction” de 20 de Junho de 2013, que consiste em um método para seleção *in vitro* de compostos que são capazes de potenciar o efeito do dano no DNA causado por agentes quimioterápicos para o tratamento do câncer compreendendo compostos que inibem a interação entre a MyD88 e ERK MAP KINASE. Utiliza métodos tais como: dois sistemas híbridos, cromatografia de afinidade, co-imunoprecipitação, fração e isolamento de grandes complexos moleculares, imunoblotting, imunohistoquímica, ensaio de imunoprecipitação, um ensaio biacore ou um ensaio GST pulldown para a identificação dos compostos. No entanto, esse método é indicado apenas para a inibição da via de sinalização dependente da MyD88 e erk map kinase e não engloba outras proteínas da resposta imunológica independentes da via da MyD88.

A patente **WO2013042137 A1** intitulada “Bicyclic heterocycles as irak4 inhibitors” de 28 de Março de 2013, versa sobre o fornecimento de um inibidor enzimático quinase heterocíclio bicíclico que poderá ser terapêuticamente útil como

um inibidor de quinase, particularmente inibidor da IRAK4. Entretanto, esta patente não fornece a inibição de outras moléculas que são importantes na cascata de sinalização em resposta a ativação da TLR.

A patente **WO2011043740 A1** intitulada “Peptides for inhibition of myd88 dependent signaling pathways” de 14 de Abril de 2011, caracteriza um inibidor peptídico que contém um segmento, que pode inibir a ativação da MyD88 utilizado para o tratamento de doenças associadas com a excessiva ativação do sistema imune. Não obstante, fornece peptídeos que inibam apenas a via de sinalização da MyD88 e não compreende proteínas que pertencem a uma via de sinalização independente desta proteína.

A patente **WO2009143987 A1** intitulada “Viral encoded pellino protein and its uses” de 03 de Dezembro de 2009, relata uma nova composição farmacêutica, de uma forma codificada viralmente da proteína Pelino e seu uso na regulação nas vias de sinalização que são associadas com desordens imunológicas. Porém, este fármaco interage apenas com a proteína IRAK-1 que pertence a uma das vias de ativação do NfκB.

A patente **WO2006070860 A1** intitulada “Novel screening method for inflammatory cytokine inhibitor” de 06 de Julho de 2006, destina-se a mostrar um novo método de triagem para um inibidor de citocinas inflamatórias. No entanto, limitou-se a descrever a inibição da ativação de citocinas focando apenas nas proteínas IRF-4 e IRF-5.

A patente **US20130203649 A1** intitulada “Inhibitors of TLR signaling by targeting TIR domain interfaces” de 08 de Agosto de 2013, descreve peptídeos do domínio TIR, como também, métodos de uso dos peptídeos na regulação da ativação e sinalização da TLR. Esta patente, destina-se a desenvolver peptídeos que inibam a ativação da TLR apenas através de seu domínio de reconhecimento de patógenos e não a inibição de proteínas da via de sinalização.

A patente **EP2477641 A2** intitulada “Inhibition of endosomal toll like receptor activation” de 25 de Julho de 2012, relata receptores de reconhecimento de padrões (PRRs), incluindo as TLRs, e, em particular, um método de inibição da ativação induzida por ácidos nucleicos como, por exemplo, de TLR endossomais que usam um agente que se liga ao ácido nucleico, preferencialmente de uma maneira que é

independente da sequência nucleotídica, a química e ou a estrutura do ácido nucléico responsável por induzir a ativação da TLR. Também relata métodos de identificação de agentes de ligação de ácidos nucleicos adequados para o uso em tais métodos. Porém, enfoca na inibição apenas de proteínas que apenas são ativadas por ácidos nucléicos.

A patente **WO2013116590 A1** de 08 de Agosto de 2013 intitulada “Inhibition of pattern recognition receptors in pancreatic cancer treatment using tlr inhibitors”, descreve uma nova descoberta de que a ativação dos receptores de reconhecimento padrão é central na progressão do câncer pancreático, e fornece antagonistas dos receptores de reconhecimento padrão (PRRs), incluindo TLRs 4, 7, e 9, e CLR dectin1, para o tratamento e prevenção do câncer e inflamação pancreático. Essa invenção descobriu que o desenvolvimento e progressão do câncer podem ser prevenidos pela inibição de receptores de reconhecimento de padrão. Esta patente, no entanto, descreve a inibição da ativação de três TLRs através do domínio de ativação destas, como também de apenas uma molécula da via de sinalização.

A patente **WO2008131457 A1** intitulada “Functional toll like receptors (TLR) on melanocytes and melanoma cells and uses there of” de 30 de outubro de 2008, relata um método de detecção da expressão gênica ou da atividade proteica da TLR em células de melanoma ou melanócito. Como também revelou métodos de modulação da expressão gênica ou atividade proteica em uma célula de melanócito ou melanoma pelo contato da célula com um agente modulador da TLR e um método de inibição da migração da célula de melanoma pelo contato desta com um inibidor de TLR. Entretanto, descreveu métodos apenas para o câncer de pele.

Nenhuma das patentes citadas acima possui semelhança com o método proposto de um painel genético na detecção de genes inibidores da via de sinalização das proteínas Toll Like Receptor (TLR) para o tratamento do câncer.

Por fim, o objetivo do invento proposto é disponibilizar a identificação e predição precisa de alteração na expressão dos genes relacionados à via de sinalização das proteínas Toll Like Receptors (TLRs), sendo uma forma de diagnóstico precoce ao risco de desenvolver o câncer e outras doenças.

Descrição da Invenção

A presente invenção descreve um método baseado em painel genético para a identificação de marcadores moleculares relacionados à inibição e/ou regulação negativa das proteínas Toll Like Receptors (TLRs), que podem ser utilizados no tratamento de diversos tipos de câncer e outras doenças.

O método visa a detecção gênica da produção de moléculas da resposta imunológica tais como citocinas pró-inflamatórias que são responsáveis pela inflamação provocada pelo excesso de ativação das TLRs. Além disso o método visa a detecção gênica das proteínas envolvidas na cascata de sinalização que regulam vários processos celulares tais como proliferação, diferenciação e progressão do ciclo celular em resposta a uma variedade de sinais extracelulares.

A presente invenção baseia-se em análises de dados, genômicos, metabolômicos, interactômicos e da literatura para a análise das interações gênicas da via de sinalização das TLRs. O método de busca para a elaboração do painel compreendeu a seleção de potenciais genes relacionados à via de sinalização referentes à inibição e/ou regulação negativa das TLRs.

O painel consistirá em um grupo de 37 genes: FLIP, TRAIL-R, LGR4, GPR48, CREB, ERK1/2, OPTN, A20, SIKE, RNF125, TGBK1, FIP2, NLRX1, IRAK-1/4, SARM, IRAK-M, SHIP-1, A20 mRNA, IRAK-4, TOLLIP, p105, RIG-I, DNTBK1, DNIKK- ϵ , MAL/TIRAP, DUSP, YOPJ, MEK1, KFKB1, TPL-2, p50, PI3K, DNTIRAP, DNRAC1N17, TRIM30a, COT/TLP2, NLRC3, TBK1/KK- ϵ , que serão caracterizados como marcadores moleculares e analisados quanto a inibição e/ou regulação negativa da cascata de sinalização das TLRs.

A análise compreenderá da detecção de diferentes níveis de expressão gênica, de mRNA ou proteico baseando-se em diferentes métodos como por exemplo RTPCR, qRT, RNA-Seq, e microarranjos, não excluindo outros métodos de detecção de expressão gênica, dos genes citados acima presentes no painel genético proposto nesta invenção. A tabela 1 descreve o painel genético de forma completa da presente invenção com informações pertinentes aos genes.

Tabela 1

Símbolo	Gene	Genes inibidos	Genes regulados negativamente
FLIP	CASP8 and FADD-like apoptosis regulator	FADD-caspase-8 interactions and DR apoptosis	
TRAIL-R	tumor necrosis factor receptor superfamily, member 10b		TLR2, TLR3, and TLR4 signaling.
LGR4/GPR48	leucine-rich repeat containing G protein-coupled receptor 4		TLR2/4 signaling
CREB	cAMP response element binding protein		CD14
ERK1/2	mitogen-activated protein kinase 3/mitogen-activated protein kinase 1	p38	
OPTN	optineurin	IFNb/ helicase (RIG-I/Mda-5)	
A20	tumor necrosis factor, alpha-induced protein 3	IFNb	TLR4
SIKE	suppressor of IKBKE 1	IFNb	
RNF125	ring finger protein 125, E3 ubiquitin protein ligase	IFNb	IFNb
TBK1	TANK-binding kinase 1	TNFa	
FIP2	Optineurin	TNFa	
NLRX1	NLR family member X1		NF-kB
IRAK-1/4	interleukin-1 receptor-associated kinase 1/interleukin-1 receptor-associated kinase 4	benzimidazole	
SARM	sterile alpha and TIR motif containing 1		TRIF
A20 mRNA	tumor necrosis factor, alpha-induced protein 3		TLR4
IRAK-4	interleukin-1 receptor-associated kinase 4		TLR signaling
TOLLIP	toll interacting protein	Nfkb,TNF and IL-6 production,IRAK	

p105	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	NfκB	
RIG-I	DEAD (Asp-Glu-Ala-Asp) box polypeptide 58	Il12b transcription	
DNTBK1	Dominant negative TANK-binding kinase 1	IRF5	
DNIKK-ε		IRF5	
MAL/TIRAP	toll-interleukin 1 receptor (TIR) domain containing adaptor protein	IRF7	
DUSP	dual specificity phosphatase	MAPK/p38	
YopJ		MAPK and NF-κB signalling	
MEK1	mitogen-activated protein kinase kinase 1		TLR2 signaling.
Nfκb1	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	IFN-β	
TPL-2	mitogen-activated protein kinase kinase kinase 8		IFN-β
p50/p105	nuclear factor of kappa light polypeptide gene enhancer in B-cells 1		IFN signaling
PI3K	phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit alpha	IL-12	
DNTIRAP	Dominant negative toll-interleukin 1 receptor (TIR) domain containing adaptor protein	Rac1V12	
DNRac1N17	Dominant negative ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1)	NF-κB transactivation	
TRIM30a		NF-κB activation	
Cot/tpl2	carnitine O-octanoyltransferase/mitogen-activated protein kinase kinase kinase 8	IKK, different MAP kinase pathways and PI3K-Akt-mTOR-p70 S6k pathway	
NLRC3	NLR family, CARD domain containing 3	TRAF6	

TBK1/IKKε	TANK-binding kinase 1	IKKs	
-----------	-----------------------	------	--

REIVINDICAÇÕES

1. Processo que prevê a inibição ou a regulação negativa da cascata de sinalização da TLR através da detecção da expressão, nível de mRNA ou de proteína em um ou mais biomarcadores selecionados em um grupo consistindo de FLIP, TRAIL-R, LGR4, GPR48, CREB, ERK1/2, OPTN, A20, SIKE, RNF125, TGBK1, FIP2, NLRX1, IRAK-1/4, SARM, IRAK-M, SHIP-1, A20 mRNA, IRAK-4, TOLLIP, p105, RIG-I, DNTBK1, DNIKK-ε, MAL/TIRAP, DUSP, YOPJ, MEK1, KFKB1, TPL-2, p50, PI3K, DNTIRAP, DNRAC1N17, TRIM30a, COT/TLP2, NLRC3, TBK1/IKK-ε, em que o aumento ou diminuição de um ou mais dos biomarcadores, em comparação com um padrão, indicará uma probabilidade de inibir e/ou regular negativamente a via de sinalização da TLR.
2. A análise de amostras biológica a partir do método da reivindicação 1.
3. Processo na forma de detecção dos padrões dos níveis de expressão, de mRNA ou proteico dos potenciais biomarcadores descritos no método de reivindicação 1.

PAINEL GENÉTICO NA DETECÇÃO DE GENES INIBIDORES DA VIA DE SINALIZAÇÃO DAS PROTEÍNAS TOLL LIKE RECEPTOR (TLR) PARA O TRATAMENTO DO CÂNCER

RESUMO

A invenção refere-se a um painel de marcadores genéticos (biomarcadores) que possuem potencial de diagnóstico e predição das alterações na expressão dos genes relacionados à via de sinalização das proteínas Toll Like Receptors (TLRs), fornecendo uma forma de diagnóstico precoce ao risco de desenvolver o câncer e outras doenças. A invenção consiste na identificação da variação da expressão de genes que estão ligados à inibição e/ou regulação negativa das proteínas TLRs e que de desta forma podem ser utilizados no tratamento de diversos tipos de câncer e outras doenças.

7. CONCLUSÕES

Podemos por conseguinte, concluir que:

- A integração de dados disponíveis na busca por proteínas que possam inibir a via de sinalização das TLRs, podem auxiliar na melhor estratégia terapêutica, visto que cada proteína irá desativar parcialmente as proteínas da via responsáveis pela resposta inflamatória e, consequentemente, a diminuição do câncer;
- Das 37 proteínas encontradas, 7 podem diretamente regular negativamente ou inibir as TLR2, 3 ou 4, sendo estas alvos relevantes na construção de inibidores moleculares. As 30 restantes estão relacionadas com a via de sinalização tais como do NfκB, citocinas inflamatórias, proteínas fosforiladoras e apoptose;
- A predição computacional dos SNPs missense, associando-os a alterações danosas podem auxiliar na interpretação das variações que ocorreram nas TLRs visto que dos 5.839 SNPs encontrados nas TLRs, 991 estavam na região do exon e alteravam os aminoácidos, e somente 223 foram classificados com alterações estruturais danosas;
- A análise dos polimorfismos missense selecionou SNPs que podem ser candidatos a investigações detalhada sobre as consequências das mudanças estruturais e sua associação com doenças como o câncer;
- Estudos *in silico* prévios que possam inferir quanto as alterações estruturais nas proteínas e as consequências destas mudanças são uma ferramenta importante no uso pré-clínico podendo ajudar na interpretação da grande quantidade de polimorfismos, contribuindo na busca de uma medicina de precisão, onde pequenas alterações podem ter grandes efeitos;
- Em relação aos dados coletados da literatura, vários estudos mencionaram a ligação das variações genéticas nos genes das TLRs com a susceptibilidade ao câncer, mas essas interações são bastante complexas e dependem de diversos fatores ambientais, étnicos e de gênero;
- Em alguns casos os resultados encontrados através da predição estrutural é discordante com o encontrado nos dados da literatura, que associa o SNP ao câncer, enquanto que a predição não indica uma mudança estrutural importante. Estes conflitos podem estar relacionados, entre outros fatores, à estreita associação entre as variações e estudos

de grupos específicos e/ou a mudanças estruturais não perceptíveis pelos softwares.

ANEXO

- Apresentação do trabalho: Literature and data mining of SNP in the gene of human TLR4 related with cancer no congresso regional de bioquímica,
- Apresentação do trabalho: Analysis through literature and data mining of SNPs related with cancer in the genes of human TLR family no congresso nacional de bioquímica;
- Apresentação do trabalho: Analysis of SNPs missenses in the stability of the human TLR4 protein no simpósio internacional em diagnóstico e terapêutica;
- Participação no resumo: expression profile of a moxotomale prostate gland na conferência internacional de cabrinos;
- Participação no workshop on drug desing and neglected tropical diseases;
- Cursou a disciplina introdução à computação no Centro de informática da UFPE.