

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Biologia Celular

Felipe Batista Leão

**MECANISMOS DA REGULAÇÃO TRANSCRICIONAL DE *Rasgef1b*, UM
GENE INDUZÍVEL POR RECEPTORES DO TIPO TOLL**

Belo Horizonte
2019

Felipe Batista Leão

**MECANISMOS DA REGULAÇÃO TRANSCRICIONAL DE *Rasgef1b*, UM
GENE INDUZÍVEL POR RECEPTORES DO TIPO TOLL**

Versão Final

Tese apresentada ao Programa de Pós-Graduação em Biologia Celular do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Ciências.

Área de concentração: Biologia Celular
Orientador: Dr. Aristóbolo Mendes da Silva

Belo Horizonte
2019

043 Leão, Felipe Batista.
Mecanismos da regulação transcricional de RasGEF1b, um gene induzível por receptores do tipo Toll [manuscrito] / Felipe Batista Leão. – 2019.
112 f. : il. ; 29,5 cm.

Orientador: Dr. Aristóbolo Mendes da Silva.
Tese (doutorado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Biologia Celular.

1. Biologia Celular. 2. Expressão gênica. 3. Fatores ras de Troca de Nucleotídeo Guanina. 4. NF-kappa B. I. Silva, Aristóbolo Mendes da. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.

CDU: 576



ATA DA DEFESA DE TESE DE DOUTORADO DE

FELIPE BATISTA LEÃO

213/2019
entrada
2º/2014
2016667529

Às nove horas do dia 29 de julho de 2019, reuniu-se, no Instituto de Ciências Biológicas da UFMG, a Comissão Examinadora da Tese, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho final intitulado: "MECANISMOS DA REGULAÇÃO TRANSCRICIONAL DE RASGEF1B, UM GENE INDUZÍVEL POR RECEPTORES DO TIPO-TOLL", requisito final para obtenção do grau de Doutor em Biologia Celular. Abrindo a sessão, o Presidente da Comissão, **Dr. Aristóbolo Mendes da Silva**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	Instituição	Indicação
Dr. Aristóbolo Mendes da Silva	UFMG	APROVADO
Dr. João Paulo de Biaso Viola	INCA	APROVADO
Dr. Daniel Santos Mansur	UFSC	APROVADO
Dr. Breno de Mello	UFOP	APROVADO
Dr. Enrrico Bloise	UFMG	APROVADO

Pelas indicações, o candidato foi considerado: APROVADO
O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora. **Belo Horizonte, 29 de julho de 2019.**

Dr. Aristóbolo Mendes da Silva (Orientador) Aristóbolo Mendes da Silva
Dr. João Paulo de Biaso Viola João Paulo de Biaso Viola
Dr. Daniel Santos Mansur Daniel Santos Mansur
Dr. Breno de Mello Breno de Mello Silva
Dr. Enrrico Bloise Enrrico Bloise

Obs: Este documento não terá validade sem a assinatura e carimbo do Coordenador

Prof. Erika Cristina Jorge
Coordenadora do Programa de Pós Graduação:
em Biologia Celular ICB/UFMG

Erika Cristina Jorge

Não precisamos saber pra onde iremos, se soubermos de onde viemos.

AGRADECIMENTOS

Agradeço ao professor Aristóbolo pela orientação ao longo de todos esses anos, essencial para a conclusão desse trabalho. Pela confiança, paciência, pelo conhecimento transmitido com boa vontade e entusiasmo, e de um modo geral por guiar minha trajetória pelo caminho da ciência, fundamental em minha formação profissional como pesquisador.

Ao Programa de Pós-graduação em Biologia Celular e seu corpo docente e discente, funcionários, e coordenadores que fizeram parte dessa trajetória e contribuíram de alguma forma para a realização do presente trabalho.

À sociedade Brasileira em nome das agências financiadoras CNPq, Capes, FAPEMIG e INCTV que financiaram e possibilitaram a realização desse trabalho.

À Universidade de Columbia (NY – USA), ao Dr. Sankar Ghosh e a toda sua equipe, Crystal, Carla, Thomas, Lauren, Chung, Alyssa, Maryam, Tan, Bella e Chris, pela generosa receptividade e colaboração.

Aos membros da banca, Dr. Daniel Mansur, Dr. João Viola, Dr. Breno de Mello Silva, Dr. Enrico Bloise, e aos suplentes Dr. Marco Campos e Dr. Eduardo Tarazona, por gentilmente cederem seu tempo e contribuírem com o trabalho.

Aos colegas de trabalho que passaram pelo Laboratório de Genes Inflamatórios durante a realização desse trabalho, Heliana, Brener, Isadora, Sérgio, Izabella, Alice, Giovanni, Jôsy, Fernanda, Igor, Kathleen, Rhaylla, Grazielle, Mariana, Kátia, Thalita, Paula, Solange, e Carolina pelo companheirismo, boa vontade, e por proporcionarem um ambiente de trabalho amigável, colaborativo e respeitoso. Em especial ao Brener que contribuiu ativamente para a realização desse trabalho.

Agradeço sobretudo à minha família pelo amor incondicional, pela confiança, pelo carinho, e por todo o apoio que sempre me foi dado em todas as minhas escolhas. Principalmente ao meu pai e à minha mãe por serem minha base e meu exemplo de caráter, honestidade, perseverança, generosidade, enfim, de vida. Todas as minhas conquistas eu devo a vocês e sou eternamente grato pela sorte de tê-los como pais. Às minhas queridas avós Ana Dulce e Ana Emirena, e aos meus saudosos avôs Leão e Ênio.

À Mariane pelo amor, carinho, dedicação, amizade, companheirismo e cumplicidade. Pela paciência e pelo apoio fundamental nas horas difíceis de necessidade, mas também por compartilhar e me proporcionar os melhores momentos. Obrigado pela coragem em abraçar os desafios junto comigo e sempre me ajudar a superá-los, essencial não só para a conclusão desse trabalho, como para tantos outros.

Aos amigos de Belo Horizonte, aos velhos melhores amigos de Formiga, e aos amigos companheiros de república, por compartilharem alguns dos melhores momentos da vida. São tantas pessoas que é impossível citar o nome de cada um. Quem é, sabe. Obrigado por tornarem essa viagem mais leve e prazerosa.

Aos grandes amigos e companheiros de banda com os quais tive o prazer e o privilégio de compartilhar dessa paixão que é fazer música. Obrigado pela paciência e compreensão nos momentos em que não pude estar presente e por compartilharem de cada momento de êxtase que a música nos proporciona.

A todas às mentes brilhantes que pavimentaram o caminho da ciência até aqui.

À oportunidade da vida e à minha própria força de vontade, que me faz perseverar e sempre tentar entregar meu melhor.

Enfim, agradeço a todos que fizeram e que fazem a diferença em meu desenvolvimento pessoal e profissional, sem a contribuição de cada um, certamente esse caminho seria infinitamente mais árduo.

SUMÁRIO

ABREVIATURAS	10
ENDEREÇOS ELETRÔNICOS	13
LISTA DE TABELAS	14
LISTA DE FIGURAS	15
RESUMO	17
ABSTRACT	19
1. INTRODUÇÃO.....	21
1.1 O sistema imune inato e os receptores de reconhecimento de padrão (PRRs)	21
1.2 Receptores do tipo Toll (TLRs)	22
1.3 Vias de sinalização de TLRs.....	25
1.4 Família de fatores de transcrição NF- κ B	29
1.5 A superfamília das Ras-GTPases.....	32
1.6 Os fatores de troca de nucleotídeos guanina das proteínas Ras (RasGEFs) e o membro RasGEF1b.....	36
2. JUSTIFICATIVA E HIPÓTESE.....	42
3. OBJETIVOS.....	43
3.1 Objetivo Geral.....	43
3.2 Objetivos específicos	43
4. METODOLOGIA.....	44
4.1 Animais.....	44
4.2 Obtenção e cultivo de macrófagos derivados de medula óssea (BMDMs)	44
4.3 Cultura de células imortalizadas	45
4.4 Detecção de contaminação por <i>Mycoplasma ssp</i>	46
4.5 Extração do RNA total.....	47
4.6 Obtenção da primeira fita de cDNA	48
4.7 RT-PCR quantitativa em tempo real (RT-qPCR).....	48
4.8 <i>Nested</i> RT-PCR.....	49
4.9 Oligonucleotídeos e condições termocíclicas	49
4.10 Análises <i>in silico</i>	51
4.11 Amplificação e clonagem do promotor putativo de <i>Rasgef1b</i>	51
4.12 Geração de mutantes por deleção e síntese de DNA	53
4.13 Transformação bacteriana.....	54

4.14	Obtenção de DNA plasmidiano em pequena e média escala (midiprep).....	54
4.15	Digestão com enzimas de restrição.....	54
4.16	Sequenciamento de DNA plasmidial.....	55
4.18	Transfecção transiente.....	55
4.19	Obtenção do extrato celular e ensaio de gene repórter.....	56
4.20	Plasmídeos.....	57
4.21	Ensaio para cromatina acessível por transposase seguido de sequenciamento (ATAC-seq, do inglês <i>Transposase-Accessible Chromatin with high-throughput sequencing</i>).....	58
4.22	Ensaio de desvio de mobilidade eletroforética (EMSA, do inglês, <i>Electrophoretic mobility shift assay</i>).....	59
4.23	Imunoprecipitação da cromatina seguido de sequenciamento (ChIP-seq, do inglês <i>Chromatin Immunoprecipitation Sequencing</i>).....	61
4.24	Análise estatística.....	62
5.	RESULTADOS.....	63
5.1	<i>Rasgef1b</i> é preferencialmente expresso em macrófagos.....	63
5.2	Análises <i>in silico</i> da sequência nucleotídica de <i>RasGEF1b</i> e identificação de potenciais elementos regulatórios.....	66
5.3	Identificação de regiões de cromatina acessível, ilhas CpG, e conservação evolutiva ao longo da sequência de <i>Rasgef1b</i>	70
5.4	Caracterização da região regulatória de <i>Rasgef1b</i> e identificação do promotor funcional.....	76
5.5	Elementos κ B são necessários para a ativação máxima do promotor de <i>Rasgef1b</i>	81
5.6	Em macrófagos ativados, NF- κ B/RelA é recrutado para o promotor <i>Rasgef1b</i> e contribui para a expressão induzida por TLRs.....	85
6.	DISCUSSÃO.....	89
7.	CONCLUSÕES.....	97
8.	TRABALHOS PUBLICADOS OU EM SUBMISSÃO.....	98
9.	REFERÊNCIAS.....	99

ABREVIATURAS

AP-1 – Proteína ativadora 1 (*Activator protein 1*)

ATAC-seq - Ensaio para cromatina acessível por transposase seguido de sequenciamento (*Transposase-Accessible Chromatin with high-throughput sequencing*)

BMDMs – Macrófagos derivados de medula óssea (*Bone marrow derived macrophages*)

BSA - Albumina sérica bovina (*Bovine Serum Albumin*)

C/EBP – Proteínas de ligação ao *enhancer* CCAAT (*CCAAT/enhancer binding protein*)

cDNA – DNA complementar (*complementary DNA*)

ChIP-seq - Imunoprecipitação da cromatina seguido de sequenciamento (*Chromatin Immunoprecipitation Sequencing*)

CpG ODN – CpG Oligodeoxinucleotídeo (*CpG Oligodeoxynucleotide*)

DMEM - Meio Eagle modificado de Dulbecco (*Dulbecco Modified Eagle's Media*)

DMSO - Dimetilsulfóxido (*Dimethylsulphoxide*)

dNTPs - Desoxirribonucleotídeos trifosfato (*Deoxynucleotides Triphosphate*)

EDTA - Ácido etilendiamino tetra-acético (*Ethylene-Diamine-Tetraacetic Acid*)

ERK – Cinase regulado por sinal extracelular (*Extracellular-signal Regulated Kinase*)

GAPs - Proteínas ativadoras de GTPases (*GTPase-activating proteins*)

GDP- Guanosina difosfato (*Guanosine Diphosphate*)

GEFs – Fatores de troca de nucleotídeos guanina (*Guanine-nucleotide exchange factors*)

GTP – Guanosina trifosfato (*Guanosine triphosphate*)

H3K27ac – Histona H3 acetilada na lisina 27 (*histone H3 lysine 27 acetylation*)

H3K4me3 – Histona H3 trimetilada na lisina 4 (*histone H3 lysine 4 trimethylation*)

H3K9ac - Histona H3 acetilada na lisina 9 (*histone H3 lysine 9 acetylation*)

HEK - Célula embrionária de rim humano (*Human Embryonic Kidney 293 cell*)

IL-1R - Receptor de Interleucina-1 (*Interleukin-1 receptor*)

IL-6 - Interleucina-6 (*Interleukin-6 cytokine*)

I κ B – Inibidor de kappa B (*Inhibitor of κ B*)

IKK - Cinase de I κ B (*I κ B kinase*)

IRAK - Cinase associada ao IL-1R (*IL-1R-associated kinase*)

IRF - Fator regulador de Interferon (*Interferon regulatory factor*)

LPS – Lipopolissacarídeo (*Lipopolysaccharide*)

MAPKs - Proteínas cinases ativadas por mitógenos (*Mitogen-activated protein kinases*)

Mal/TIRAP - Adaptador tipo MyD88/Proteína adaptadora contendo o domínio TIR (*TIR domain-containing adapter protein/MyD88-adaptor-like*)

mRNA - RNA mensageiro (*messenger RNA*)

MyD88 – Proteína de resposta primária de diferenciação mieloide (do inglês *myeloid differentiation factor 88*)

NF- κ B – Fator nuclear KappaB (*nuclear factor kappaB*)

PAMPs - Padrões moleculares associados a patógenos (*Pathogen-associated molecular patterns*)

PBS – Salina tamponada em fosfato (*phosphate-buffered saline*)

PCR – Reação em cadeia da polimerase (*polymerase chain reaction*)

PRRs – Receptores de reconhecimento padrão (*pattern recognition receptors*)

Poly (I:C) - *Polyinosinic:polycytidylic acid*

Ras – Sarcoma de rato (*rat sarcoma*)

RasGEF - Fatores de troca de nucleotídeos guanina das proteínas Ras (*ras associated GEFs*)

RPMI - *Roswell Park Memorial Institute culture medium*

Sp1 – Proteína de especificidade 1 (*specificity protein 1*)

TAK – Cinase ativada por TGF- β (*TGF- β activated kinase*)

TIR – domínio receptor Toll/IL-1 (*intracelular Toll- Interleukin1 (IL-1) receptor*)

TLRs – Receptores do tipo *Toll* (*toll-like receptors*)

TNF - Fator de necrose tumoral α (*Tumor necrosis factor*)

TRAM - Molécula adaptadora relacionada à TRIF (*TRIF-related adapter molecule*)

TRIF – Proteína adaptadora contendo o domínio TIR indutora de interferon- β (*TIR domain containing adaptor protein inducing interferon*)

TSS – Sítio de início da transcrição (*transcription start site*)

ENDEREÇOS ELETRÔNICOS

UniProtKB

<http://www.uniprot.org>

Clustal Omega versão 1.2.1

www.ebi.ac.uk/Tools/msa/clustalo

TreeFam

<http://www.treefam.org/>

TFBIND

<http://tfbind.hgc.jp/>

TFsearch

www.cbrc.jp/research/db/TFSEARCH.html

mVISTA

<http://genome.lbl.gov/vista/mvista/submit.shtml>

MethPrimer

<http://www.urogene.org/cgi-bin/methprimer/methprimer.cgi>

EMBOSS-Cpgplot

http://www.ebi.ac.uk/Tools/seqstats/emboss_cpgplot/

EMBOSS-Newcpgreport

http://www.ebi.ac.uk/Tools/seqstats/emboss_newcpgreport/

BLASTn

<http://www.blast.ncbi.nlm.nih.gov/Blast.cgi>

NF- κ B PBM Dataset

<http://thebrain.bwh.harvard.edu/nfkb/>

UCSC Genome Browser

<https://genome.ucsc.edu/>

Primer3 v.0.4

<http://bioinfo.ut.ee/primer3-0.4.0/>

LISTA DE TABELAS

Tabela 1: Localização celular de TLRs e seus principais ligantes	24
Tabela 2: Listagem dos genes alvos, dos oligonucleotídeos utilizados para amplificação, e das condições termocíclicas utilizadas em cada reação.....	50
Tabela 3: Listagem dos oligonucleotídeos utilizados para o sequenciamento da região flanqueadora de <i>Rasgef1b</i> clonada no vetor pGL3-Basic.	55
Tabela 4: Listagem e descrição dos plasmídeos utilizados para transfecção.	57
Tabela 5: Sondas contendo os elementos κ B do promotor de <i>Rasgef1b</i> utilizados em EMSA	61

LISTA DE FIGURAS

Figura 1: Representação esquemática dos receptores IL-1R e TLRs.....	23
Figura 2: Representação esquemática das vias de sinalizações de TLRs.....	28
Figura 3: Representação esquemática de dímero NF- κ B e da estrutura dos membros das famílias NF- κ B/Rel e I κ B.	31
Figura 4: Representação esquemática da regulação da atividade de Ras-GTPases.	34
Figura 5: Principais efetores de Ras.	35
Figura 6: Representação esquemática da estrutura de diferentes RasGEFs.....	37
Figura 7: Análise da expressão de <i>Rasgef1b</i> induzida por LPS em macrófagos.....	63
Figura 8: A homologia entre os membros da família RasGEF1.....	65
Figura 9: Análise da expressão basal e induzida por LPS de <i>Rasgef1a</i> , <i>Rasgef1b</i> e <i>Rasgef1c</i> em macrófagos.....	66
Figura 10: Validação do proposto TSS de <i>Rasgef1b</i> em macrófagos.	68
Figura 11: Elementos regulatórios associados à resposta imune identificados na sequência de <i>Rasgef1b</i> através de análise in silico.....	69
Figura 12: Identificação de regiões de cromatina acessível ao longo da sequência de <i>Rasgef1b</i> através de ATAC-seq e H3K27ac ChIP-seq.....	71
Figura 13: Marcadores epigenéticos de regiões transcricionalmente ativas presentes ao longo da sequência de <i>Rasgef1b</i>	72
Figura 14: Alinhamento da sequência de RasGEF1b murino e humano através de BLASTn e identificação de ilha CpG.....	74
Figura 15: Análise de regiões evolutivamente conservadas entre a sequência de <i>Rasgef1b</i> murino e seus ortólogos em mamíferos.....	75
Figura 16: Representação esquemática do repórter pGL3-2747/+119 e dos mutantes gerados por deleção.	78
Figura 17: Confirmação da geração dos plasmídeos repórter do promotor de <i>Rasgef1b</i> através de digestão e eletroforese.	79
Figura 18: Caracterização funcional da região promotora de <i>Rasgef1b</i>	80
Figura 19: Análise do efeito de NF- κ B sobre a ativação do promotor de <i>Rasgef1b</i>	82
Figura 20: Análise da ligação de NF- κ B aos elementos κ B do promotor de <i>Rasgef1b</i>	84
Figura 21: Identificação do recrutamento de NF- κ B/RelA ao longo da sequência de <i>Rasgef1b</i> através de ChIP-seq.	86

Figura 22: NF- κ B contribui para a expressão de *Rasgef1b* induzida por TLRs em macrófagos..... 88

RESUMO

O fator de troca de nucleotídeos guanina RasGEF1b é codificado por um gene de expressão induzida em células do sistema imune inato em resposta à ativação de receptores do tipo Toll (TLRs). O acúmulo de seu mRNA ocorre em macrófagos humanos e murinos estimulados com diferentes agonistas de TLRs, assim como em camundongos infectados com protozoários parasitas, sugerindo que a proteína codificada desempenha uma função na resposta imune. No entanto, seu papel nesse contexto ainda não é totalmente compreendido. Além disso, os mecanismos transcricionais que regulam sua expressão são também desconhecidos, dada a ausência de estudos sobre o promotor e os potenciais elementos regulatórios envolvidos nesse processo. Uma vez que as diversas funções biológicas mediadas pelo módulo RasGEF–RasGTPase são, em parte, dependentes dos seus padrões de expressão em tipos celulares ou tecidos específicos, a compreensão dos mecanismos transcricionais que determinam sua expressão dirigida pode fornecer informações importantes a respeito de seu papel na célula. Portanto, o objetivo do presente trabalho foi identificar e caracterizar a região promotora que regula a expressão do gene *Rasgef1b* em resposta à ativação de TLRs. Para isso, clonamos uma sequência de 2.8 kb (-2747/+119) a montante da região codificadora de *Rasgef1b* no plasmídeo repórter da luciferase pGL3-basic. Mutantes por deleções progressivas foram gerados e a atividade de cada construção foi avaliada através de estudos de transfecção transiente e ensaio de gene repórter. Nossos resultados mostram que uma curta sequência proximal contendo 302 pb (-183/+119) é suficiente e necessária para promover eficientemente a expressão da luciferase em células HEK293 e macrófagos Raw264.7, exibindo uma atividade basal robusta, mas também induzida por TLRs. Dessa forma, fornecemos evidências diretas que essa região é responsável pela atividade do promotor de *Rasgef1b*. Interessantemente, identificamos que a sequência é evolutivamente conservada, está contida em uma região de cromatina acessível, contém ilha CpG, parece recrutar constitutivamente a RNA Polimerase II, e abriga essencialmente elementos de ligação a Sp1 e ao principal fator de transcrição ativado por TLRs, o NF- κ B. Mostramos que a superexpressão das subunidades de NF- κ B RelA ou cRel, mas não os complexos de RelB, induzem a um aumento significativo na atividade do promotor, e que mutações nos sítios κ B localizados a montante ou a jusante do TSS resultam em significativa redução de sua atividade induzida por LPS em macrófagos murinos. Além disso, demonstramos através de ensaios de mobilidade eletroforética (EMSA) que parte dos

elementos κ B são capazes de recrutar dímeros NF- κ B, e através de CHIP-seq, observamos que RelA é recrutado não somente à região do promotor, mas também a regiões distais. Finalmente, encontramos que a expressão de *Rasgef1b* induzida por TLRs é reduzida em macrófagos desprovidos de RelA ou tratados com o inibidor farmacológico Bay11-7082 indicando que sua expressão é parcialmente dependente de NF- κ B. Em conjunto, reportamos pela primeira vez a identificação e caracterização funcional do promotor *Rasgef1b*, e demonstramos a contribuição do NF- κ B em sua expressão induzida por TLRs.

Palavras-chave: RasGEF1b. NF- κ B. Promotor. Expressão gênica.

ABSTRACT

The guanine-nucleotide exchange factor RasGEF1b is encoded by a gene whose expression is induced in innate immune cells in response to the activation of Toll-like receptors (TLRs). Its mRNA accumulates in human and mouse macrophages upon stimulation with various TLRs agonists as well as in mice infected with protozoan parasites, suggesting that *Rasgef1b* plays a role in innate immunity. However, its role in this context is not yet fully understood. In addition, the transcriptional mechanisms that regulate its expression are still unknown due to the lack of studies about the promoter and the potential regulatory elements involved in this process. As the diverse biological functions mediated by the RasGEF/RasGTPases modules are in part dependent on their expression profile in specific cell types or tissues, understanding the transcriptional mechanisms that drive their directed expression can provide important information about their role in the cell. Therefore, our objective here was to characterize the promoter region that regulates the TLR-inducible *Rasgef1b* expression. For this, we cloned a 2.8 kb sequence (-2747/+119) upstream from *Rasgef1b* coding region in the luciferase reporter plasmid pGL3-basic. Progressive 5' and 3' deletions were performed and the activity of each construct was assessed by transient transfection and reporter gene assay. Our results show that a short proximal sequence containing 302 bp (-183/+119) is sufficient and necessary to efficiently promote luciferase expression in HEK293 cells and Raw264.7 macrophages, exhibiting robust basal and TLR-induced activity. Therefore, we provide direct evidence that this region harbors *Rasgef1b* active promoter. Interestingly, we have identified that the sequence is evolutionarily conserved, contained in an accessible chromatin region with a CpG island, appears to constitutively recruit the RNA Polymerase II, and harbors essentially Sp1 and TLR-inducible NF- κ B binding sites. We showed that the overexpression of NF- κ B RelA or cRel subunits, but not of RelB complexes, induced a significant increase in promoter activity, and that mutation of the κ B sites located upstream or downstream from *Rasgef1b* TSS resulted in a significant reduction in the LPS-inducible promoter activation in mouse macrophages. In addition, we showed through electrophoretic mobility shift assays (EMSA) that some of the κ B elements can recruit NF- κ B dimers, and through CHIP-seq, we showed that RelA is recruited not only to *Rasgef1b* promoter, but also to additional distal regions. Finally, we found that the TLR-inducible *Rasgef1b* expression is reduced in RelA deficient or Bay11-7082-treated macrophages, indicating that its expression is partially dependent on NF-

κ B. Altogether, we report for the first time the identification and functional characterization of *Rasgef1b* promoter, demonstrating the contribution of NF- κ B in its TLR-inducible expression.

Keywords: RasGEF1b. NF- κ B. Promoter. Gene expression.

1. INTRODUÇÃO

1.1 O sistema imune inato e os receptores de reconhecimento de padrão (PRRs)

O papel do sistema imune inato é reconhecidamente fundamental para a manutenção da homeostase e sobrevivência do organismo, uma vez que atua como a primeira linha de defesa contra a infecção, sendo essencial para a proteção contra uma ampla gama de patógenos e potenciais ameaças (Kawai and Akira, 2006). É um sistema ancestral de reconhecimento microbiano, evolutivamente conservado na maioria dos organismos multicelulares, presente tanto em animais invertebrados e vertebrados, assim como em plantas (Albiger et al., 2007; Dempsey et al., 2003).

Classicamente, a imunidade inata era considerada uma resposta inespecífica mediada por células fagocíticas, incluindo macrófagos e neutrófilos, que englobam e lisam patógenos microbianos (Akira et al., 2006). No entanto, diversos estudos mostraram que o sistema imune inato apresenta certa especificidade decorrente de um repertório limitado de receptores de reconhecimento de padrão (PRRs, do inglês *Pattern recognition receptors*), que possibilita a ativação celular em resposta a uma ampla gama de patógenos. Esses receptores reconhecem componentes moleculares microbianos conservados, comuns a vários grupos de microorganismos, conhecidos como padrões moleculares associados a patógenos (PAMPs, do inglês *pathogen-associated molecular pattern*), assim como estruturas moleculares endógenas derivadas de células danificadas, conhecidas como padrões moleculares associadas a dano celular (DAMPs – do inglês, *damage associated molecular patterns*) (Janeway and Medzhitov, 2002). Os PRRs são comumente expressos em células do sistema imune inato, tais como células dendríticas, macrófagos, neutrófilos, e o reconhecimento de seus ligantes leva ao rápido disparo de vias de sinalização intracelular que culmina na indução da resposta imune através da produção de citocinas, quimiocinas, moléculas co-estimulatórias, peptídeos antimicrobianos, proteínas de sinalização intracelular, entre outros mediadores. Esse processo desencadeia não só a ativação de respostas imediatas do hospedeiro, como a inflamação, mas também auxilia na montagem e orquestração da resposta imune adaptativa (Kawasaki and Kawai, 2014). Entre as classes caracterizadas de PRRs estão os receptores do tipo Toll (TLRs - do inglês *Toll like receptors*), receptores do tipo RIG-I (RLRs - do inglês *RIG-I-like receptors*), receptores do tipo NOD (NLRs - do inglês *NOD-like receptors*), receptores do tipo AIM2 (ALRs – do inglês *AIM2-like receptors*),

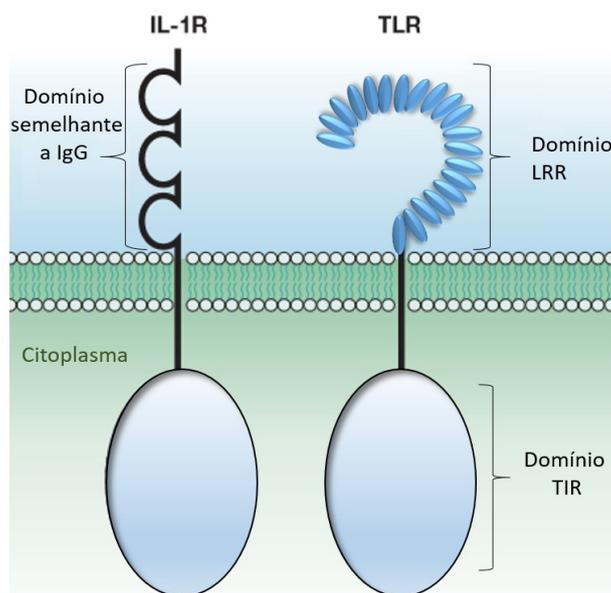
receptores do tipo lectina C (CLRs – do inglês C-type lectin receptors) e sensores citosólicos de DNA (Kawasaki and Kawai, 2014). A combinação específica dos receptores envolvidos no processo pode transmitir informações básicas sobre o tipo de patógeno invasor, permitindo a montagem de uma resposta celular adequada (Kanzler et al., 2007).

1.2 Receptores do tipo Toll (TLRs)

Os receptores do tipo Toll (TLRs) são PRRs evolutivamente conservados, presentes desde animais invertebrados até mamíferos (Medzhitov, 2007). A proteína transmembrana Toll foi inicialmente identificada em estudos com drosófilas que demonstraram sua função essencial na regulação da polaridade dorsoventral do embrião (Anderson et al., 1985). Posteriormente, Lemaitre e colaboradores forneceram a primeira evidência de que Toll estaria envolvida no processo da resposta imune, demonstrando que drosófilas mutantes com perda da função no receptor Toll apresentavam expressão defeituosa de um peptídeo antifúngico e eram altamente susceptíveis à infecção por fungo (Lemaitre et al., 1996). Esses resultados forneceram as primeiras evidências de que essa proteína deveria ser responsável pela detecção e mediação de respostas à infecção causada por fungo em drosófila. Em seguida, Medzhitov e colaboradores identificaram um homólogo humano para a proteína Toll de drosófila - o que cunhou o termo “receptor do tipo Toll” (TLRs). Eles mostraram que um mutante constitutivo ativo do homólogo humano de Toll era capaz de induzir a ativação de NF- κ B em células humanas (do inglês *nuclear factor kappa-light-chain-enhancer of activated B cells*) e conseqüentemente a expressão de genes regulados por esse fator de transcrição, como citocinas inflamatórias e moléculas co-estimulatórias (Medzhitov et al., 1997). A partir daí o campo da imunobiologia dos TLRs se expandiu rapidamente, e atualmente essa é uma das classes mais estudadas entre os PRRs.

É notório atualmente que os TLRs são cruciais para a ativação do sistema imune no combate a microorganismos patogênicos. Esses receptores estão presentes em células residentes e imunes das superfícies corporais de interface hospedeiro-patógeno, como na pele e em mucosas dos tratos respiratório, gastrointestinal e urogenital (Anwar et al., 2013; Mifsud et al., 2014). Atualmente, são descritos 10 membros da família dos TLRs em humanos (TLR1-TLR10) e 12 em camundongos (TLR1-TLR9, TLR11-TLR13) (Kawai and Akira, 2010). Os TLRs são glicoproteínas transmembrana do tipo I caracterizados por um domínio extracelular (ou luminal quando presentes em

endossomos) em “forma de ferradura” contendo repetições ricas em leucina (LRR, do inglês *Leucine rich repeats*), responsável pelo reconhecimento de diversos PAMPs provenientes de bactérias, protozoários, fungos e vírus, assim como de alguns DAMPs, uma única hélice transmembrana, e um domínio intracelular ou citoplasmático homólogo ao receptor de interleucina-1 (IL-1R, do inglês *Interleukin-1 receptor*) nomeado domínio receptor Toll/IL-1 (TIR – do inglês *Intracelular Toll- Interleukin1 (IL-1) receptor*), responsável pela ativação de vias de sinalização intracelular (Albiger et al., 2007) (**Fig. 1**). Apesar de compartilharem estrutura semelhante, o conjunto dos TLRs pode reconhecer uma grande variedade de patógenos devido a variada composição de aminoácidos entre seus domínios extracelulares. (Akira et al., 2006; Albiger et al., 2007; Kawai and Akira, 2010; Takagi, 2011).



Baseado em Akira & Takaeda (2004) e Albiger et al (2007)

Figura 1: Representação esquemática dos receptores IL-1R e TLRs. Os receptores de interleucina 1 (IL-1R) e os receptores do tipo Toll (TLRs) apresentam um domínio intracelular ou citoplasmático semelhante denominado TIR. No entanto, o domínio extracelular desses dois receptores é diferente. Os receptores de IL-1R possuem três domínios semelhantes à imunoglobulina e os TLRs possuem repetições ricas em leucina (LRR) em uma estrutura característica que se assemelha a forma de uma ferradura.

Os TLRs são expressos tanto na superfície celular quanto associados a vesículas intracelulares (Kawai and Akira, 2010). Segundo Kawasaki (2014), TLR1, TLR2, TLR4, TLR5, TLR6 e TLR10 estão localizados na superfície celular e reconhecem uma ampla gama de componentes de membrana derivados de agentes microbianos como lipídeos, lipoproteínas e proteínas, enquanto que TLR3, TLR7, TLR8, TLR9, TLR11, TLR12 e

TLR13 estão localizados em endossomos e reconhecem principalmente moléculas de ácidos nucleicos (Kawasaki and Kawai, 2014; Raetz et al., 2015). Ademais, alguns TLRs como TLR2 e TLR4, podem ser removidos da superfície celular após o reconhecimento de moléculas ligantes e recrutados para o fagossomo (Anwar et al., 2013). A compartimentalização desses receptores permite à célula a capacidade de reconhecer tanto patógenos extracelulares e intracelulares, atribuir informações espaciais à essa detecção, e confere um nível extra de regulação específica da resposta imune (Brubaker et al., 2015). Ademais, diferentes subgrupos celulares podem exibir uma expressão diferencial de determinados TLRs, o que faz com que as células sejam capazes de desenvolver padrões específicos de resposta imune a diferentes patógenos. Por exemplo, em humanos, células dendríticas plasmocitóides expressam preferencialmente TLR7 e TLR9 e são altamente responsivas a infecções virais, ao passo que células dendríticas derivadas de monócitos expressam os demais TLRs, incluindo o TLR4, permitindo que essas células sejam ativadas por diferentes tipos de agonistas, incluindo LPS (do inglês, *Lipopolysaccharide*) (Akira et al., 2001; Schreiber et al., 2010). A **Tabela 1** sumariza alguns dos mais importantes ligantes de cada TLR e sua localização celular.

Tabela 1: Localização celular de TLRs e seus principais ligantes

Receptor do tipo Toll	Localização celular	Ligantes
TLR1/2, TLR2/6	Superfície Celular	<ul style="list-style-type: none"> - Lipoproteínas - Ácido Lipoteóico (LTA) - Peptídioglicano derivado de bactérias Gram-positivas - Lipoarabidomanana - Glicosilfosfatidilinositol mucina - Zymozan - Mannan
TLR3	Endossomo	<ul style="list-style-type: none"> - dsRNA viral - siRNA - <i>Self</i> RNA derivado de células danificadas
TLR4	Superfície Celular	<ul style="list-style-type: none"> - Lipopolissacarídeo (LPS) - Proteínas estruturais de vírus - Manose

TLR5	Superfície Celular	- Flagelina
TLR7	Endolisossomo	- ssRNA viral - RNA de streptococcus B
TLR8	Endolisossomo	- ssRNA viral e bacteriano
TLR9	Endolisossomo	- CpG DNA viral e bacteriano - Hemozoína
TLR10/12	Superfície Celular	- Listeria
TLR11	Endolisossomo	- Componentes protéicos de bactérias uropatogênicas - Moléculas do tipo profilina de <i>Toxoplasma gondii</i>
TLR12	Endolisossomo	- Moléculas do tipo profilina de <i>Toxoplasma gondii</i>
TLR13	Endossomo	- rRNA bacteriano 23S - Componentes virais de vírus da estomatite vesicular

Baseado em (Kumar et al., 2011) e (Kawasaki and Kawai, 2014).

1.3 Vias de sinalização de TLRs

A sinalização celular disparada pela ativação dos TLRs se inicia quando um PAMP se liga à porção extracelular (ou luminal) do receptor. Ocorre então sua dimerização e o recrutamento de moléculas adaptadoras através do domínio citossólico TIR. Esse processo vai disparar vias de sinalização intracelular que culminarão na ativação de fatores de transcrição centrais na resposta imune, como NF- κ B, AP-1 (do inglês *activator protein 1*) através da ativação de MAP cinases, e IRFs (do inglês *Interferon regulatory factor*). Uma vez ativados, esses fatores são translocados para o núcleo onde se ligam à região promotora e regulam a transcrição de genes inflamatórios que codificam citocinas pró-inflamatórias, quimiocinas, interferons do tipo I, entre outras

várias moléculas, incluindo proteínas reguladoras intracelulares (Kahana and Cleveland, 1999; Kumar et al., 2011; Takagi, 2011).

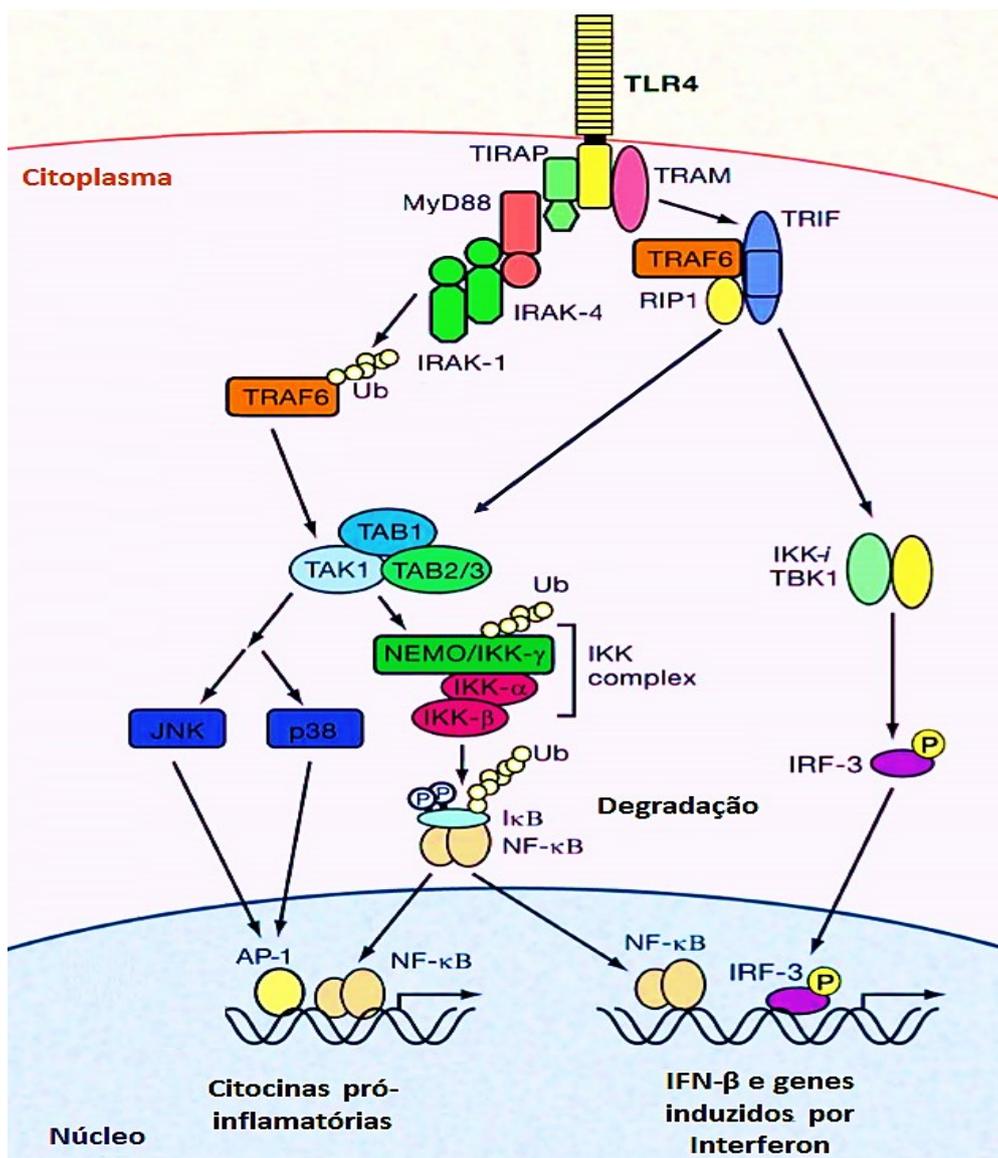
Como mencionado, após o reconhecimento dos ligantes, a sinalização disparada por TLRs é primariamente mediada via recrutamento de moléculas adaptadoras pelo domínio TIR. Dependendo do receptor ativado, diferentes moléculas adaptadoras serão recrutadas, e devido à complexidade dessa via, a sinalização dos TLRs é categorizada em vias MyD88-dependente e TRIF-dependente (Foster and Medzhitov, 2009; Wang et al., 2010) (**Fig. 2**). A via MyD88-dependente é utilizada por todos os TLRs, exceto por TLR3. A via TRIF-dependente se dá apenas por TLR3 e TLR4.

Na via MyD88-dependente, a primeira proteína adaptadora a ser recrutada é a proteína de resposta primária de diferenciação mielóide 88 (MyD88, do inglês *Myeloid differentiation primary response protein 88*), a qual contém um domínio TIR que se liga ao domínio TIR do TLR. Após sua ligação, MyD88 recruta as cinases IRAK4 (do inglês *IL-1R-associated kinase 4*) e IRAK1 (do inglês *IL-1R-associated kinase 1*). Com a formação desse complexo, IRAK4 fosforila inicialmente IRAK1, o que induz à sua atividade cinase e intensa autofosforilação em múltiplos resíduos de serina e treonina, resultando em sua dissociação de MyD88. Em seguida, IRAK-1 liga-se a várias outras proteínas, incluindo o fator 6 associado ao receptor do fator de necrose tumoral (TRAF6, do inglês *TNF receptor-associated factor 6*). A atividade ubiquitina ligase de TRAF6 é ativada, resultando em sua auto-ubiquitinação e na ativação da cinase TAK1 (do inglês *transforming growth factor-beta activated protein Kinase 1*) que faz parte de um complexo formado pelas proteínas ligadas à TAK1, TAB1, TAB2, TAB3 (do inglês *TAK1-binding protein*). Esse complexo então fosforila IKK β , componente do complexo IKK (do inglês, *I κ B kinase*) formado também por IKK α e NEMO (IKK γ). O complexo IKK promove então a fosforilação do inibidor citosólico de NF- κ B, o I κ B (do inglês *inhibitor of κ B*) que será então marcado por ubiquitinação e degradado em proteossomo, liberando por fim o fator de transcrição NF- κ B, que será translocado para o núcleo para promover a transcrição de citocinas pró-inflamatórias e diversos genes essenciais na resposta imune. Paralelamente à ativação de NF- κ B, o complexo TAK1 ativa ainda membros da família das MAP cinase (MAPK, do inglês *Mitogen-activated protein kinases*) como ERK1/2 (do inglês *extracellular-signal-regulated kinases*) p38 e JNK (do inglês *c-Jun N-terminal kinases*) que medeiam a ativação e translocação do fator de transcrição AP-1 para o núcleo (Kawasaki and Kawai, 2014; Lu et al., 2008; O'Neill and Bowie, 2007). Nos casos de TLR1, TLR2, TLR4 e TLR6, a sinalização MyD88-

dependente requer um adaptador adicional, a proteína adaptadora contendo o domínio TIR (TIRAP, do inglês *TIR domain containing adaptor protein*) que é importante para o recrutamento de MyD88 ao domínio TIR do receptor **(Figura 2)**

A ativação dos receptores TLR3 ou TLR4 ativa a via TRIF-dependente (do inglês *TIR domain containing adaptor protein inducing interferon IFN- β*) sendo que TLR4 recruta TRIF através da molécula adaptadora relacionada a TRIF (TRAM, do inglês *TRIF-related adaptor molecule*) (Kanzler et al., 2007; Kumar et al., 2011). TRIF interage tanto com TRAF6 quanto TRAF3. TRAF6 recruta a cinase RIP-1 (do inglês *Receptor-interacting protein 1*) que vai levar a ativação do complexo TAK1 promovendo a ativação de NF- κ B e MAPKs, similarmente à via MyD88-dependente (Kawasaki and Kawai, 2014). Já a interação de TRIF com TRAF3 culmina no recrutamento de TBK1 (do inglês *TANK binding kinase 1*) e IKK-epsilon que juntas promoverão a fosforilação e subsequentemente a dimerização e translocação nuclear de IRF3, que irá regular a transcrição de IFN- β e de genes responsivos a interferon (Hayden and Ghosh, 2012; Kawasaki and Kawai, 2014; Kumar et al., 2011) **(Figura 2)**.

Para que seja efetiva, a reposta de TLRs contra uma variedade grande de ameaças recai sobre a indução coordenada de centenas de produtos gênicos com inúmeras funções regulatórias. Todo esse processo é essencial para a indução da resposta imune inata do hospedeiro, assim como para auxiliar na montagem e orquestração da reposta imune adaptativa (Kawasaki and Kawai, 2014). Nesse sentido, os receptores do tipo Toll desempenham um papel essencial na manutenção da homeostase do organismo, e, portanto, é fundamental o estudo das respostas celulares mediadas por esses receptores.



Modificado de Akira et al. (2006)

Figura 2: Representação esquemática das vias de sinalizações de TLRs. A sinalização de TLR4 é representativa da sinalização dos TLRs, uma vez que dispara tanto a via MyD88-dependente quanto TRIF-dependente e utiliza todos os adaptadores principais. A ativação da via MyD88-dependente leva ao recrutamento de TIRAP e MyD88, que vai recrutar IRAK-4 e IRAK-1. IRAK-1 sofre fosforilação e recruta TRAF6, o que, por sua vez, vai ativar o complexo TAK1, TAB1, TAB2/3. Esse complexo então fosforila a IKK β promovendo a ativação do complexo IKK. O complexo IKK ativo irá fosforilar I κ B, que é então ubiquitinado e degradado, liberando NF- κ B que se transloca para o núcleo. Simultaneamente o complexo TAK1 ativa membros da família das MAP cinases, ERK1/2, p38 e JNK que vão mediar a ativação e translocação do fator de transcrição AP-1 para o núcleo. Na via TRIF-dependente, TRIF interage com TRAF6 que recruta a cinase RIP-1 levando à ativação do complexo TAK1, o que promove a ativação NF- κ B e MAPKs. A interação de TRIF com TRAF3, culmina no recrutamento de TBK1 e IKK-epsilon que juntas promoverão a fosforilação, e subsequentemente a dimerização e translocação nuclear de IRF3, que vai regular a transcrição de IFN- β e de genes responsivos a interferon.

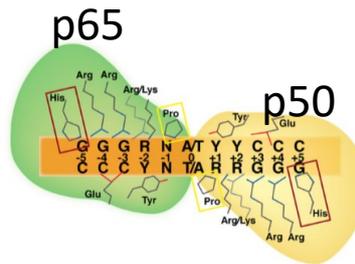
1.4 Família de fatores de transcrição NF- κ B

O NF- κ B foi descrito em 1986 por Rajan Sen e David Baltimore como um fator de transcrição nuclear que se ligava a uma região reguladora do gene da cadeia leve kappa (sítio κ B) em células B (Sen and Baltimore, 1986). Desde então, se tornou um dos mais estudados fatores de transcrição devido ao seu papel central e complexo na sinalização intracelular (Ghosh and Karin, 2002), na hematopoiese (Grossmann et al., 1999), na regulação na imunidade inata e adaptativa (Vallabhapurapu and Karin, 2009) e em outros processos fundamentais como proliferação celular, diferenciação e sobrevivência (Hayden and Ghosh, 2012). Em relação à imunidade inata e adaptativa, seu papel inclui a regulação de centenas de genes que codificam para citocinas, quimiocinas, moléculas de adesão celular, peptídeos antimicrobianos, receptores, entre outros (Baker et al., 2011; Doyle and O'Neill, 2006; Hayden and Ghosh, 2012). No contexto da resposta mediada por TLRs, NF- κ B é particularmente importante pois todos os TLRs de mamíferos conhecidos induzem sua ativação (Brubaker et al., 2015; Kawai and Akira, 2007; Kawasaki and Kawai, 2014).

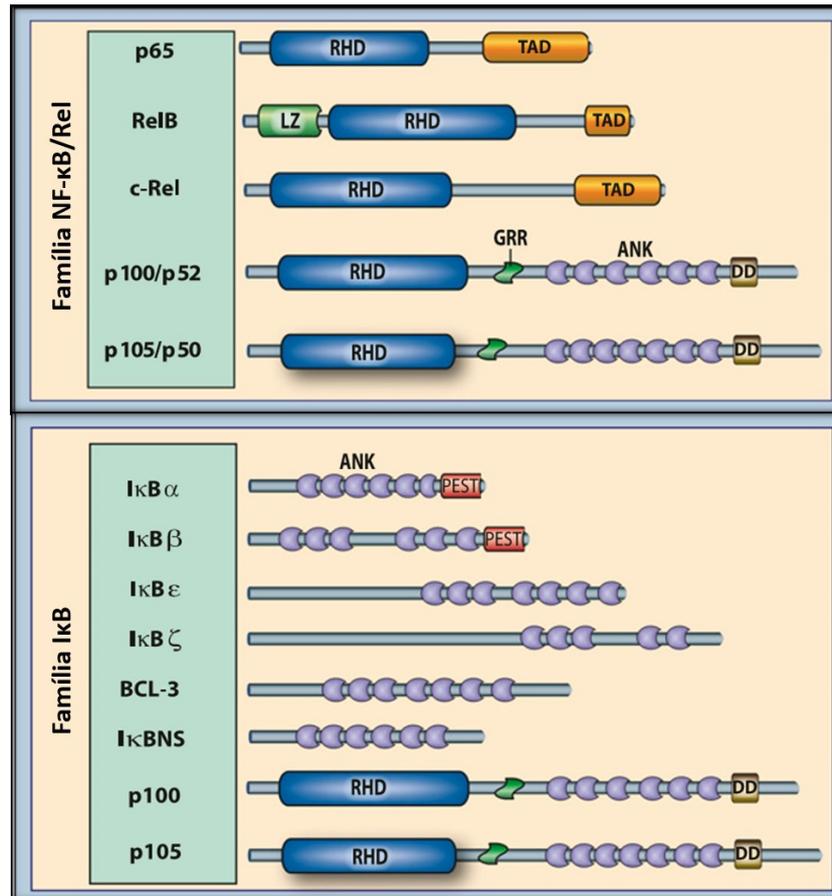
A família Rel/NF- κ B é composta por cinco membros, sendo eles, RelA (p65), RelB, c-Rel, as proteínas precursoras NF- κ B1 (p105) e NF κ B2 (p100), que são processadas em p50 e p52, respectivamente (Gerondakis et al., 1999). Os membros se combinam de forma homo ou heterodimérica e podem se ligar a um conjunto de elementos específicos no DNA conhecidos como sítios κ B (consenso: 5' GGGRNWYYCC 3', onde N é qualquer nucleotídeo, R é uma purina, W é adenina ou timina, e Y é pirimidina), presentes em *enhancers* e/ou promotores de uma série de genes alvos, regulando sua transcrição (**Fig. 3**) (Wang et al., 2012). A habilidade de ligação ao DNA deve-se ao domínio N-terminal altamente conservado de aproximadamente 300 aminoácidos chamado de domínio de homologia a Rel (RHD, do inglês *Rel-homology domain*) (Gerondakis et al., 1999; Wang et al., 2012). Esse domínio, compartilhado entre todas as subunidades, é o responsável pela dimerização, ligação ao DNA e interação com a proteína inibitória de NF- κ B, o I κ B (do inglês *inhibitor of kappa B*) (Hoffmann et al., 2003). No entanto, apenas as subunidades RelA, c-Rel e RelB possuem os domínios C-terminais de transativação (TADs, do inglês *transactivation domains*), que conferem a habilidade de induzir a transcrição (**Fig. 3**). Embora as subunidades p50 e p52 não possuam os domínios TAD, elas podem regular a transcrição quando se combinam na forma de heterodímeros com outras subunidades de NF- κ B ou com proteínas co-

ativadoras não relacionadas que possuam a capacidade de transativação. Alternativamente, homodímeros de p50 e p52 regulam negativamente a transcrição ao competir pelos sítios κ B do DNA com os dímeros que contém TADs (Hayden and Ghosh, 2012). Nesse sentido, o NF- κ B pode tanto induzir quanto reprimir a transcrição, e a composição do dímero confere, em partes, as propriedades do potencial de transativação e da especificidade de ligação ao DNA. Entre os possíveis heterodímeros, o mais freqüente é formado pelos monômeros RelA (p65) e p50, que possui a capacidade de regulação transcricional de uma ampla gama de genes, muitos deles envolvidos na resposta imune inata e adaptativa, como mencionado previamente (Hayden and Ghosh, 2004; Vallabhapurapu and Karin, 2009). No entanto, as cinco subunidades de NF- κ B podem se associar em pelo menos 12 dímeros diferentes. Cada dímero pode potencialmente se ligar preferencialmente a motivos característicos no DNA e, embora muitas vezes desempenhem função redundante e compensatória, é também possível que regulem especificamente alguns genes alvos (Saccani et al., 2003; Smale, 2012; Wang et al., 2012).

Em células não estimuladas, os dímeros de NF- κ B são encontrados associados a proteínas inibitórias da família I κ B, sendo I κ B α , I κ B β e I κ B ϵ as principais a desempenhar esse papel (Hayden and Ghosh, 2004). Além dos I κ Bs clássicos, há um grupo de proteínas atípicas que também fazem parte da família dos inibidores de NF- κ B. São elas, BCL3 (do inglês *B cell CLL/lymphoma 3*), I κ B ζ , I κ BNS e ainda os monômeros não-processados de NF- κ B, p105 e p100 (Hayden and Ghosh, 2012; Vallabhapurapu and Karin, 2009) (**Fig. 3**). Os I κ Bs são caracterizados pela presença de cinco a sete motivos de anquirina que medeiam a ligação ao domínio de dimerização de NF- κ B (Hatada et al., 1992). Eles interagem com o domínio RHD de NF- κ B mascarando parcialmente a sua sequência de localização nuclear e impedindo o seu transporte ao núcleo e conseqüente ligação ao DNA (Bonizzi and Karin, 2004). Em particular, estruturas cristalográficas de I κ B α e I κ B β ligado aos dímeros p65/p50, revelaram que o I κ B mascara somente a sequência de localização nuclear de p65, enquanto p50 permanece acessível (Huxford et al., 1998; Jacobs and Harrison, 1998; Malek et al., 2003).



Wang et al. *Cell Reports*
2012



Modificado de Hayden & Gosh (2012)

Figura 3: Representação esquemática de dímero NF- κ B e da estrutura dos membros das famílias NF- κ B/Rel e I κ B. Painel superior. Representação do heterodímero mais comum de NF- κ B ligado à sequência consenso de um sítio κ B. **Painel inferior:** A família dos fatores de transcrição NF- κ B/Rel consiste em cinco membros: RelA (p65), RelB, c-Rel (Rel), e nas proteínas precursoras p100 (NF- κ B2) and p105 (NF- κ B1), que dão origem a p52 e p50, respectivamente. Todos os membros possuem o domínio RHD, o que caracteriza a família. A família I κ B consiste de oito membros, I κ B α , I κ B β , I κ B ϵ , I κ B ζ , BLC3, I κ BNS, p100 e p105. Essa família é caracterizada pela presença de múltiplos domínios de repetições de anquirina. Os domínios que caracterizam cada família, assim como domínios específicos de alguns membros estão indicados. (ANK) Domínios de Anquirina; (DD) Domínio de morte (*death domain*); (RHD) Domínio de homologia a Rel; (TAD) Domínio de transativação; (LZ) Domínio de zíper de leucina; (GRR) Região rica em glicina.

Para que NF- κ B possa regular a transcrição gênica é necessário que haja a degradação do I κ B. A ativação de NF- κ B pode se dar através de diversos estímulos que vão levar a ativação de vias de sinalização distintas. Essas vias podem ser classificadas em canônica e não-canônica. A via canônica, ou clássica, pode se iniciar por meio de estímulos reconhecidos por diferentes receptores, como os receptores de TNF (TNFR, TNF receptors), de IL-1 (IL-1R, IL-1 receptor), receptores de antígenos, ou PRRs como os já mencionados TLRs. Nessa via, o complexo IKK é necessário e suficiente para a fosforilação de I κ B (Hayden and Ghosh, 2004). Esse processo vai levar à degradação do I κ B possibilitando a translocação dos dímeros de NF- κ B normalmente formados pelas subunidades RelA, c-Rel e p50 (Rothwarf and Karin, 1999). A via não-canônica, ou alternativa, é induzida por membros específicos da família de TNF, como CD40L, LT $\alpha\beta$, RANKL, BAFF, e tem um mecanismo de regulação diferente. Em contraste com a via clássica, a via alternativa é independente de NEMO (IKK γ) e medeia a ativação persistente de dímeros de NF- κ B formados sobretudo por RelB/p52 (Hayden and Ghosh, 2012). Tal via está envolvida com a formação dos órgãos linfóides e com o desenvolvimento e sobrevivência das células B (Karin and Ben-Neriah, 2000; Sun, 2011).

Devido ao seu papel central em diversos processos, vários estudos mostraram a associação de NF- κ B a doenças autoimunes, inflamatórias, metabólicas e com o câncer (Baker et al., 2011; Courtois and Gilmore, 2006; Karin, 2006; Zhang et al., 2017). Portanto, o estudo do conjunto de genes controlados por esse fator de transcrição é de suma importância para o melhor entendimento das consequências de sua desregulação na homeostase do organismo.

1.5 A superfamília das Ras-GTPases

A superfamília das Ras-GTPases (do inglês *rat sarcoma small guanosine triphosphatases*) compreende um grupo de pequenas proteínas G com mais 150 membros identificados em humanos, e com base nas funções e similaridade de sequência, são divididas em cinco subfamílias principais: Ras, Rho, Rab, Arf e Ran. São proteínas monoméricas, de baixo peso molecular (20 a 25 kDa), evolutivamente conservadas e presentes em virtualmente qualquer eucarioto, que regulam diversos processos celulares fundamentais (Bos et al., 2007; Vigil et al., 2010). Por exemplo, as proteínas da subfamília Rho desempenham um papel na organização do citoesqueleto, atuando na formação dos filamentos de actina e na reorganização destes filamentos em resposta a estímulos por fatores de crescimento (Ridley and Hall, 1992). Os membros da subfamília

Rab e Arf regulam o tráfego entre membranas citoplasmáticas e o tráfego de vesículas, respectivamente (Novick et al., 1980). As proteínas da subfamília Ran regulam o tráfico nucleoplasmático e a organização de microtúbulos (Kahana and Cleveland, 1999). (Cherfils and Zeghouf, 2013). A subfamília Ras compreende 36 membros em mamíferos e é subdividida classicamente em três grupos principais: Ras, Ral, Rap (Cherfils and Zeghouf, 2013). Essas proteínas estão envolvidas no controle da expressão gênica, proliferação, diferenciação, morfologia e morte celular, sendo H-Ras, K-Ras e N-Ras os membros melhor caracterizados em humanos (Bar-Sagi and Feramisco, 1985; Coleman et al., 2004; Kauffmann-Zeh et al., 1997; Whitman and Melton, 1992). Por regularem essas funções críticas, os genes *Ras* são alvo de intensa pesquisa, e quando mutados, frequentemente levam à oncogênese e por isso são considerados proto-oncogenes (Repasky et al., 2004). Foi reportado que em pelo menos 15 a 20% dos tumores humanos são portadores de mutações oncogênicas em *Ras* (Goitre et al., 2014).

As GTPases de baixo peso molecular funcionam como “interruptores moleculares”, ciclando entre um estado inativo, quando ligadas à guanosina difosfato (GDP), e ativo, ligadas à guanosina trifosfato (GTP). A conversão entre as formas ativa e inativa é estritamente regulada por dois tipos de proteínas: fatores de troca de nucleotídeos guanina (GEFs, do inglês *Guanine nucleotide exchange factors*) que possibilitam a ativação das GTPases por promover a transição do estado inativo ligado a GDP ao estado ativo ligado a GTP; e as proteínas ativadoras de GTPases (GAPs, do inglês *GTPase-Activating Proteins*), que estimulam a hidrólise de GTP, fazendo com que retornem ao seu estado inativo ligado a GDP (Downward et al., 1990; Wolfman and Macara, 1990). Existe ainda um terceiro componente na regulação das GTPases, os inibidores da dissociação de guanina (GDI, do inglês *guanine nucleotide dissociation inhibitor*), que podem inibir ou bloquear as ações das GAPs e dos GEFs (**Fig. 4**) (Campbell et al., 1998). Embora todas as proteínas da superfamília Ras apresentem o mesmo mecanismo de regulação, cada subfamília está associada ao seu próprio grupo de GEFs e GAPs (Bos et al., 2007; Cherfils and Zeghouf, 2013; Takai et al., 2001; van Dam et al.; Wolfman and Macara, 1990).

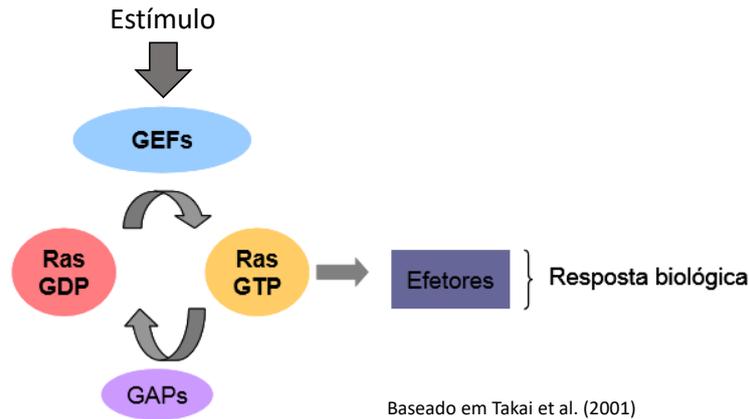


Figura 4: Representação esquemática da regulação da atividade de Ras-GTPases. Diferentes estímulos modulam a atividade dos RasGEFs e RasGAPs que, por sua vez, promovem a alteração entre o estado ativo e inativo das Ras-GTPases através da troca de GDP/GTP. As proteínas Ras ativadas são capazes de interagir com vários efetores, desencadeando diversas respostas biológicas como, por exemplo, proliferação celular e rearranjo de citoesqueleto

As proteínas da superfamília Ras são inicialmente sintetizadas no citossol mas são logo endereçadas a membranas celulares ao sofrerem importantes processamentos pós-traducionais como farnesilação/geranilgeranilação, proteólise, carboxi-metilação e palmitoilação nos motivos CAAX (C=cisteína, A=aminoácido alifático; X=qualquer aminoácido) contidos em seu domínio carboxi-terminal (Ahearn et al., 2011; Choy et al., 1999). A localização subcelular específica de cada proteína Ras é absolutamente necessária para sua atividade biológica, pois é determinante para seu recrutamento e ativação pelos GEFs, e em seguida para o recrutamento e ativação de seus efetores, capacitando a transmissão do sinal proveniente de receptores para a cascata de sinalização a jusante (Willumsen et al., 1984). A ativação de proteínas Ras se inicia através de algum estímulo, como fatores de crescimento ou hormônios por exemplo, seguido pelo recrutamento de GEFs junto ao compartimento de membrana onde Ras está localizada, permitindo então a indução transitória do complexo Ras-GTP (Boguski and McCormick, 1993). A forma ativa Ras-GTP é então capaz de interagir com seus efetores propagando o sinal. Um exemplo de via de sinalização bem caracterizado é a via Ras-Raf-MEK-ERK1/2, que se dá pela ativação de proteínas da subfamília Ras. A via é disparada a partir do reconhecimento de estímulos, entre eles o fator de crescimento epidermal (EGF), por seus cognatos receptores tirosina-cinases (i.e EGFR). Ocorre a dimerização dos receptores, o que permite a autofosforilação dos resíduos de tirosina contidos em seus domínios citossólicos através de suas cinases intrínsecas. Essa alteração permite o recrutamento da proteína adaptadora Grb2, que por sua vez recruta o RasGEF SOS. A

aproximação de SOS à membrana leva à ativação de Ras, que por sua vez recruta e ativa o efetor Raf-1, uma serina-treonina cinase. Raf-1 fosforila e ativa MEK1/2 (MAPK/ERK-cinase) que, por sua vez, fosforila e ativa a MAPK ERK1/2. Finalmente ERK1/2 fosforila e ativa fatores de transcrição que vão regular a expressão de genes associados a progressão do ciclo celular (Goitre et al., 2014; Lemmon and Schlessinger, 2010). Outras importantes vias efetoras de proteínas da subfamília Ras são Ras-PI3K, Ras-RalGDF e Ras-PLC ϵ (fosfolipase C épsilon), que em conjunto - e adicionalmente a outros efetores – promovem a regulação dos cruciais processos celulares mediados por essa subfamília (Fig.5) (Downward, 2003).

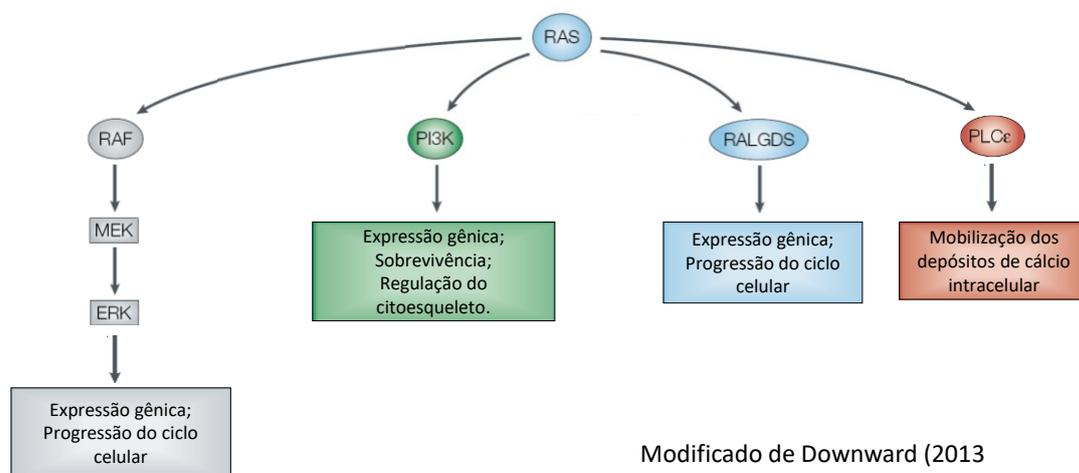


Figura 5: Principais efetores de Ras. Uma vez ligada a GTP, Ras interage com diversos efetores estimulando sua atividade. Para efeito de simplificação, a figura mostra apenas alguns dos principais efetores diretos e seus principais efeitos. A via mais bem caracterizada é de Ras-Raf-MEK-Erk. A cinase Erk possui diversos alvos citoplasmáticos e nucleares, incluindo fatores de transcrição que vão regular a progressão do ciclo celular entre outros processos. A ativação de PI3K leva à geração de segundos mensageiros lipídicos que possuem como alvo diversas proteínas intracelulares, incluindo a cinase anti-apoptótica AKT. PI3K leva ainda à ativação de RAC, envolvida com a regulação do citoesqueleto de actina. RalGDS é um GEF que leva à ativação de RalGTPases. A ativação dessa via contribui para regulação da expressão de genes que controlam o ciclo celular. A fosfolipase C ϵ (PLC ϵ) catalisa a hidrólise do fosfatidilinositol-4,5-bisfosfato em diacilglicerol e inositol trisfosfato, resultando na ativação da proteína cinase C (PKC) e na mobilização das reservas de cálcio intracelulares. ERK, *extracellular regulated kinase*; MEK, *mitogen-activated kinase/ERK kinase*; PI3K, *phosphoinositide 3-kinase*;

1.6 Os fatores de troca de nucleotídeos guanina das proteínas Ras (RasGEFs) e o membro RasGEF1b

A relevância dos estudos relacionados aos fatores de troca de nucleotídeo guanina (GEFs) se deve ao fato de essas proteínas serem capazes de ativar proteínas G, por exemplo Ras, e modular, portanto, as importantes vias de sinalização disparadas por seus diversos efetores (Vigil et al.). Os GEFs são classificados em diferentes famílias de acordo com a similaridade em sua sequência de aminoácidos e de acordo com o tipo de proteína GTPase que eles ativam (Zheng and Quilliam, 2003).

Os RasGEFs (RasGEFs, do inglês *Ras associated GEFs*) têm em comum um domínio catalítico carboxi-terminal de aproximadamente 250 aminoácidos, denominado CDC25 (do inglês, *cell division cycle 25 homology domain* - também chamado domínio RasGEF), e um domínio estrutural adjacente, amino-terminal, de aproximadamente 50 aminoácidos, chamado de domínio REM (do inglês, *Ras Exchange Motif*) (Quilliam et al., 2002). Os GEFs que possuem o domínio CDC25 regulam exclusivamente a subfamília Ras (Broek et al., 1987; Robinson et al., 1987). Além do domínio catalítico CDC25, alguns RasGEFs possuem ainda outros domínios funcionais importantes. Por exemplo, RasGRF1, Sos1 e o Sos2 possuem os domínios de homologia DH (DH - do inglês *DbI homology*) e PH (PH - do inglês *pleckstrin homology*), que estão comumente arranjados em conjunto e possibilitam adicionalmente a ativação de GTPases subfamília Rho (**Fig. 6**) (Lemmon et al., 1996; Quilliam et al., 1995). Outros domínios podem ser importantes para interações proteína-proteína, proteína-lipídeos, ligação a segundos mensageiros, sítios de fosforilação, entre outros (Vigil et al., 2010)

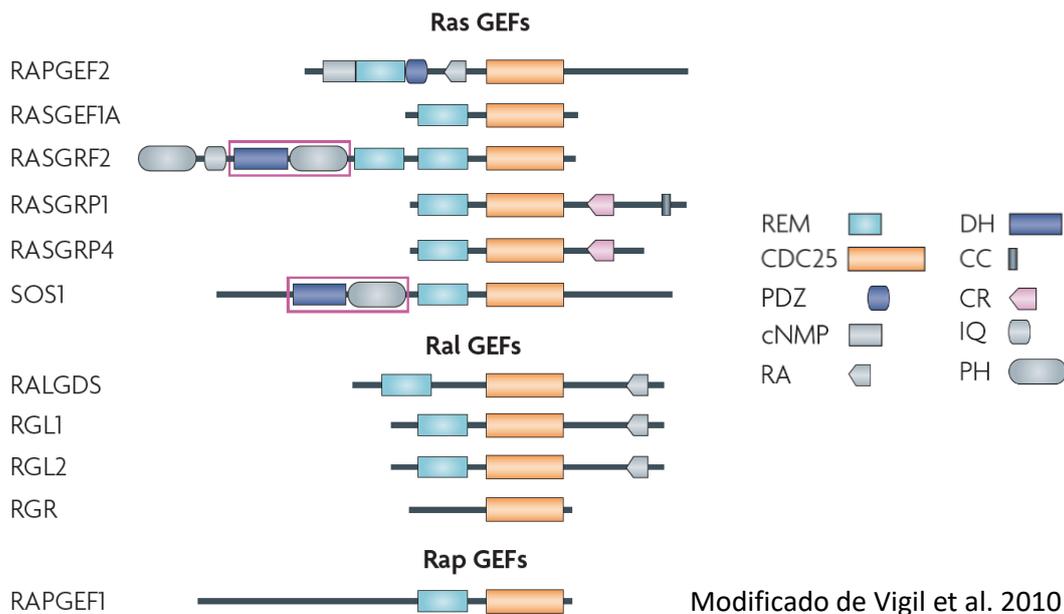


Figura 6: Representação esquemática da estrutura de diferentes RasGEFs. Os membros da família RasGEF apresentam o domínio estrutural N-terminal REM e o domínio catalítico C-terminal, Cdc25, que caracterizam a família. Diferentes RasGEFs podem apresentar ainda outros domínios específicos que desempenham variadas funções, representados na figura. (REM) *Ras Exchange Motif*; (cNMP) *Cyclic nucleotide-monophosphate binding domain*; (Cdc25) *cell division cycle 25 homology domain*; (PDZ) *Dlg homologous region*; (RA) *Ras-association domain*; (DH) *Dbl homology domain*; (CC) *Coiled-coil domain*; (CR) *Cyclic nucleotide-monophosphate binding domain*; (IQ) *Short calmodulin-binding motif containing conserved Ile and Gln residues*; (PH) *Pleckstrin homology domain*.

Como já mencionado, proteínas Ras inativas na célula encontram-se associadas a GDP. Mediante estímulos específicos, os RasGEFs promovem a dissociação do GDP da proteína Ras, tornando-a acessível ao GTP que está em excesso molar no citoplasma. Associada ao GTP, Ras adquire sua conformação estrutural ativa, e pode iniciar a cascata de sinalização através da ativação de outras proteínas, que culmina na ativação da transcrição de vários genes ligados à proliferação, diferenciação e morte celular (Boguski and McCormick, 1993). Dessa forma, esses sistemas compostos por GTPases, GAPs e GEFs permitem grande versatilidade na construção de vias de sinalização. Os sinais podem ser amplificados (um GEF pode ativar várias GTPases), integrados (várias vias ativam os mesmos GEF e GAPs) ou divididos (uma GTPase induz vários efeitos) (Wilkins and Insall, 2001). Essa versatilidade permite que, indiretamente, os GEFs mediem diferentes funções biológicas na célula.

Embora os mecanismos de ativação de Ras ocasionada pela troca GDP/GTP sejam semelhantes, a regulação dos diferentes RasGEFs pode se dar de diferentes maneiras,

incluindo interação com outras proteínas e lipídeos, sinalização por segundos mensageiros como diacilglicerol e cálcio, e modificações pós-traducionais como fosforilação. Esses eventos podem ter como consequência a translocação dos RasGEFs para um compartimento subcelular onde entrará em contato com seu alvo Ras, ou levar a alterações conformacionais no domínio catalítico ou em domínios inibitórios, permitindo sua ativação (Bos et al., 2007). No caso do RasGEF1 Sos, por exemplo, sua ativação ocorre através da interação com a proteína adaptadora Grb2 que recruta Sos à membrana plasmática permitindo o acoplamento da sinalização disparada por receptores tirosina-cinase à via Ras-Raf-Mek-Erk, como previamente descrito (Aronheim et al., 1994). Em feedback negativo, Erk fosforila Sos resultando em sua dissociação de Grb2 e interrompendo a sinalização (Waters et al., 1995). Um outro exemplo é RasGRF1, que possui um motivo de ligação à calmodulina e induz a ativação de Ras mediante influxo de cálcio (Farnsworth et al., 1995; Shou et al., 1992). Em adição aos domínios de ligação ao cálcio, membros da família RasGRP possuem domínios de ligação à diacilglicerol que são cruciais para o regulamento da atividade desses GEF através de seu endereçamento à membrana plasmática (Ebinu et al., 1998).

Outro ponto importante de regulação dos processos mediados pelos RasGEFs é o próprio padrão de expressão dessas moléculas e seus efetores Ras. Por exemplo, enquanto Sos1 é ubiquamente expressa em mamíferos, outros RasGEFs têm sua expressão predominante em tecidos ou células específicas. RasGRF1 é predominantemente expresso no sistema nervoso central, enquanto RasGRF2 tem uma expressão mais ampla (Quilliam et al., 2002). Já a família RasGRP é predominantemente expressa na linhagem de células hematopoiéticas, e embora ocorra parcial sobreposição, o perfil de expressão dos próprios membros da família é distinto (Buday and Downward, 2008; Stone, 2011). RasGRP1 e RasGRP3 são expressos predominantemente em linfócitos T e B (Stone, 2011), RasGRP2 parece ocorrer predominantemente em plaquetas (Crittenden et al., 2004), e RasGRP4 é restrito a mastócitos (Adachi et al., 2012; Yang et al., 2002). Assim como para os RasGEFs, as diversas funções biológicas desempenhadas por seus efetores Ras também são, em parte, tecido específicas e dependem da expressão diferencial dessas moléculas. Portanto, para além dos eventos de sinalização, da necessidade de colocalização subcelular, e da especificidade de interação entre determinado RasGEF e seu(s) alvo(s) Ras específico(s), a atividade dos módulos RasGEF-Ras vai depender do *pool* celular dessas moléculas em determinado tecido/célula (Gloerich and Bos, 2011).

Vários GEFs associados aos membros da subfamília Ras foram isolados e caracterizados em células de eucarióticas, como Sos, RasGRFs, RasGRPs, RalGDS, C3G, entre outros. Dada a sua variedade e importância, naturalmente, os mais diversos processos biológicos vão depender de seu funcionamento. Por exemplo, a ativação de Ras através RasGEF-CDC25 em *S. cerevisiae* é essencial para a proliferação e germinação de esporos (Broek et al., 1987). Em *Dictyostelium*, o desenvolvimento embrionário é regulado por RasGEFM, um GEF caracterizado nessa espécie (Arigoni et al., 2005). O GEF Sos, ao ativar Ras em *Drosófila*, estabelece a diferenciação de fotoreceptores (Simon et al., 1991). Nesse sentido, estudos demonstraram a importância de RasGEFs também no processo da resposta imune em mamíferos. Até o momento, o papel da família RasGRP tem sido predominante nesse contexto. Foi mostrado que RasGRP1 é importante na transdução de sinal disparada pela ativação de TCR, desempenhando um papel determinante na maturação e seleção positiva de células T (Stone, 2011). Camundongos deficientes de RasGRP1 são incapazes de montar uma resposta robusta mediada por células T contra infecções virais e bacterianas, e desenvolvem síndrome linfoproliferativa autoimune em idade avançada (Priatel et al., 2007). Neutrófilos deficientes de RasGRP2 apresentam baixa ativação de Rap1 e de integrinas $\beta 1$ e $\beta 2$, reduzida adesão em vênulas, reduzido recrutamento para cavidade peritoneal induzido por tioglicolato, e deficiência de quimiotaxia provavelmente relacionada à problemas de organização do citoesqueleto de actina (Bergmeier et al., 2007; Carbo et al., 2010). Camundongos deficientes de RasGRP3 apresentam níveis séricos de imunoglobulinas reduzido, e seus linfócitos B apresentam redução da ativação da via Ras-Erk *in vitro* e proliferação reduzida em resposta a anti-IgM e CD40 (Coughlin et al., 2005). Particularmente na resposta mediada por TLRs, de modo interessante foi demonstrado que a ativação de RasGRP3 em resposta a baixos níveis de agonistas TLR prejudica a produção de citocinas pró-inflamatórias, especialmente IL-6, em macrófagos ao ativar Rap1 e suprimir ERK1/2. No entanto, o mesmo não ocorre quando essas células são estimuladas com níveis mais elevados de agonistas. Dessa maneira, supostamente RasGRP3 pode limitar a resposta inflamatória mediada por TLRs, evitando a ativação desnecessária de macrófagos em condições fisiológicas onde o contato com microorganismos patogênicos é reduzido. Esse RasGEF seria então importante para estabelecer um limite para ativação de macrófagos, reduzindo a probabilidade do desenvolvimento de doenças inflamatórias. Nesse mesmo trabalho foi demonstrado ainda que a superexpressão de RasGRP3 é capaz de aliviar a colite induzida por dextran sulfato e a artrite induzida por colágeno *in vivo*. RasGRP4 tem expressão

restrita a mastócitos e seus progenitores. Foi demonstrado que a expressão do mRNA de TNF e IL-1 β é reduzida em mastócitos de camundongos deficientes de RasGRP4 tratados com PMA, e ao contrário de RasGRP3, a colite e artrite induzida era aliviada nesses animais, demonstrando seu papel pró-inflamatório nesse contexto (Adachi et al., 2012).

A família RasGEF1 compreende os membros RasGEF1a, RasGEF1b e RasGEF1c. Assim como os outros RasGEFs, essas proteínas apresentam o domínio N-terminal com função estrutural (REM), o domínio carboxi-terminal homólogo ao domínio catalítico CDC25 específico para a subfamília Ras, e compartilham relativa similaridade na sequência de aminoácidos. Estudos funcionais, sobretudo em relação ao RasGEF1a e RasGEF1c, são ainda bastante escassos. No entanto, alguns trabalhos tratando de RasGEF1b são encontrados na literatura. O gene RasGEF1b é potencialmente expresso em diversos organismos, desde invertebrados a vertebrados superiores, em diferentes tecidos, e em condições fisiológicas normais e patológicas. Por exemplo, em Zebrafish, foi demonstrado que a expressão de RasGEF1b ocorre durante o desenvolvimento embrionário sobretudo no cérebro e parece ser regulada pelas vias de sinalização Nodal, de FGF e TGF- β (Epting et al., 2007). Em consistência, a expressão constitutiva de *Rasgef1b* no cérebro de camundongos adultos é significativa (Ferreira et al., 2002). Interessantemente, uma série de microdeleções de diferentes tamanhos no cromossomo 4q21 foram identificadas em humanos que apresentam uma nova síndrome caracterizada por retardamento mental, problemas de fala, má formação cerebral, dimorfismo facial e restrição de crescimento. Esses estudos convergiram para uma região mínima de 761 kb que inclui o locus de RasGEF1B e PRKG2 (do inglês, *Protein Kinase cGMP-Dependent 2*) (Bonnet et al., 2010; Hu et al., 2017; Komlosi et al., 2015; Zarrei et al., 2017). Portanto, esses genes são considerados os principais candidatos responsáveis pelas alterações neurológicas observadas nesses pacientes. Atualmente, estudos estão sendo conduzidos em nosso laboratório utilizando-se modelos de camundongos deficientes de *Rasgef1b* para avaliar seu papel nesse contexto.

Não obstante, desde a descoberta desse gene nossos esforços foram concentrados sobretudo em sua caracterização sob a ótica da resposta imune. Inicialmente denominado GPIy4, RasGEF1b foi identificado e descrito por nosso grupo como um GEF de expressão fortemente induzida durante a resposta imune inata mediada por receptores do tipo Toll (Ferreira et al., 2002). Foi mostrado que, *in vitro*, o mRNA de RasGEF1b se acumula em macrófagos murinos e humanos estimulados com agonistas TLR2 e TLR4, assim como *in vivo* em órgãos linfóides e não-linfóides como timo, baço, linfonodos, rim, fígado e

coração de camundongos infectados com *Trypanosoma cruzi*. Expressão constitutiva significativa de *Rasgef1b* foi observada apenas no cérebro, intestino e testículos (Ferreira et al., 2002). Em um estudo posterior, Andrade e colaboradores demonstraram que a expressão de *Rasgef1b* também é induzida pela ativação de TLR3 através do agonista Poly:IC, e que *in vivo*, a expressão de *Rasgef1b* induzida durante a infecção por *T. cruzi* ou *Plasmodium chabaudi* em camundongos é dependente dos adaptadores de TLRs MyD88 e TRIF (Andrade et al., 2010). Essas evidências sugerem a implicação de RasGEF1b na resposta imune mediada por TLRs, e por isso, estudos estão em andamento para se esclarecer essa questão atualmente.

O gene *Rasgef1b* murino contém 14 éxons e está localizado no cromossomo 5E3 de camundongo. Até o momento, duas isoformas validadas são descritas pelo banco de dados do NCBI. A isoforma 1 (ID: NM_145839.2) codifica a proteína mais longa, com 473 aminoácidos, e a isoforma 2 (ID: NM_181318.4) utiliza um sítio de *splicing* alternativo *in-frame* na entre os exons 4 e 5, resultando na eliminação de 1 único códon e consequentemente em uma proteína mais curta (472 resíduos). Recentemente, foi relatado que, além do mRNA linear, um RNA circular (circRNA) de *Rasgef1b* gerado através de "back-splicing" dos exons 2 e 4 é induzido após ativação de TLRs em macrófagos murinos. Nesse trabalho, foi demonstrado que a expressão de *Icam-1* induzida por LPS em macrófagos é prejudicada após o *knockdown* do circRNA de *Rasgef1b*. A estabilidade do mRNA de *Icam-1* é reduzida nessas células, o que sugere um papel do circRNA de *Rasgef1b* na regulação desse processo (Ng et al., 2016). Análises transcriptômicas seguintes mostraram que o *knockdown* do circRNA de *Rasgef1b* levou à desregulação de centenas de genes adicionais induzidos por LPS, incluindo genes envolvidos nas próprias vias de sinalização de NF- κ B e LPS e na ativação de macrófagos (Ng and Marinov, 2017). Ao nível proteico, foi demonstrado que RasGEF1b se localiza em endossomos primários (Andrade et al., 2010) e funciona como um GEF para ativação de Rap2a-GTPase *in vitro* (Yaman et al., 2009). Em nosso trabalho mais recente, confirmamos que RasGEF1b é suficiente para a ativação de Rap2a em células humanas HEK293 e, embora não seja totalmente necessário, contribui para a ativação de Rap2a induzida por TLRs em macrófagos murinos. Demonstramos também que o *knockdown* de Rap2a em macrófagos murinos Raw264.7 prejudica a expressão e produção das citocinas pró-inflamatórias IL6 e KC/Cxcl1 induzida por LPS, e que tanto o silenciamento quanto a superexpressão de Rap2a prejudicam a ativação de NF- κ B, sugerindo a implicação de RasGEF1b na resposta mediada por TLRs como efetor de Rap2a em macrófagos (Carvalho et al., 2019).

Em conjunto essas evidências sugerem que RasGEF1b deve desempenhar um papel durante a resposta imune inata mediada por TLRs. Portanto, é importante que se investigue diferentes aspectos da regulação e funcionamento desse gene de modo a melhor entender o seu papel na célula.

2. JUSTIFICATIVA E HIPÓTESE

Entender a estrutura e o comportamento da expressão de um gene é um problema fundamental na biologia. Os trabalhos voltados para a análise da expressão e função de *Rasgef1b* acumulam evidências sugerindo que esse GEF está intimamente associado à resposta imune inata mediada por receptores do tipo Toll (Andrade et al., 2010; Carvalho et al., 2019; Ferreira et al., 2002; Ng and Marinov, 2017; Ng et al., 2016). No entanto, seu papel nesse processo ainda não é totalmente compreendido. Além disso, mecanisticamente, a regulação transcricional de *Rasgef1b* permanece por ser fornecida, principalmente devido à falta de informação sobre o promotor e dos elementos regulatórios responsáveis por sua expressão induzida por TLRs. Uma vez que a resposta celular apropriada mediada por RasGEFs-RasGTPase depende, em alguns casos, da expressão dirigida desses genes a tipos celulares e tecidos específicos (Buday and Downward, 2008; Gloerich and Bos, 2011; Hennig et al., 2015), a compreensão dos mecanismos que promovem a regulação transcricional pode fornecer informações importantes a respeito da função não caracterizada de *Rasgef1b* em um determinado tipo de célula e contexto. Portanto, é de grande importância a investigação da região genômica regulatória envolvida nesse processo, assim como dos potenciais fatores de transcrição que se ligam a ela, de forma a melhor compreender a regulação da expressão de *Rasgef1b* na resposta imune mediada por TLRs em macrófagos. Esses estudos podem fornecer a base para investigações adicionais sobre a regulação do gene RasGEF1b em outros tipos celulares e tecidos, além de contribuir para o melhor entendimento dos mecanismos básicos da regulação gênica.

Análises preliminares conduzidas *in silico* indicam a presença de elementos regulatórios para a ligação de fatores de transcrição ativados pela cascata de sinalização disparada por TLRs, como NF- κ B e AP-1, na sequência não-codificadora de *Rasgef1b*. Adicionalmente, trabalhos anteriores sugeriram um potencial papel de NF- κ B na regulação da expressão de *Rasgef1b* em macrófagos e fibroblastos murinos (Andrade et al., 2010; Mukherjee et al., 2013; Ng et al., 2016), e resultados preliminares obtidos

durante a fase inicial desse projeto corroboram essa suposição. Portanto, hipotetizamos que a sequência de *Rasgef1b* abriga uma sequência regulatória responsiva à ativação de TLRs e que fatores de transcrição NF- κ B se ligam à essa região e regulam sua expressão.

3. OBJETIVOS

3.1 Objetivo Geral

Identificar e caracterizar funcionalmente a região promotora que regula a expressão do gene *Rasgef1b* em resposta à ativação de TLRs e investigar o papel do fator de transcrição NF- κ B sobre sua ativação transcricional.

3.2 Objetivos específicos

1- Avaliar os níveis basais e induzidos por LPS do mRNA de RasGEF1a, RasGEF1b e RasGEF1c em macrófagos murinos.

2- Conduzir análises *in silico* da região genômica a montante da sequência codificadora de *Rasgef1b* para a identificação de elementos regulatórios de ligação a fatores gerais de transcrição e fatores de transcrição específicos associados à resposta imune inata.

3- Analisar a estrutura da cromatina ao longo da sequência do gene *Rasgef1b* para identificação de marcadores de regiões transcricionalmente ativas através de ATAC-seq e CHIP-seq de histonas acetiladas.

4- Avaliar, através de ensaios de gene repórter, o efeito da ativação de TLRs sobre a atividade de segmentos da potencial região regulatória de *Rasgef1b*.

5- Analisar o papel de NF- κ B sobre a ativação transcricional de *Rasgef1b* induzida por LPS avaliando: a) O efeito das principais subunidades de NF- κ B sobre a atividade da potencial região regulatória de *Rasgef1b*; b) O efeito de mutações sítio-dirigidas a elementos κ B sobre a atividade da potencial região regulatória de *Rasgef1b*; c) O recrutamento e a formação de complexos transcricionais aos elementos kappaB presentes na potencial região regulatória de *Rasgef1b*.

6- Analisar o papel de NF- κ B sobre a expressão de *Rasgef1b* induzida por LPS avaliando comparativamente os níveis de seu mRNA em macrófagos selvagens e deficientes de NF- κ B/RelA, ou submetidos ao tratamento com inibidor farmacológico de NF- κ B.

4. METODOLOGIA

4.1 Animais

Camundongos C57BL/6 machos com idade entre 8 e 12 semanas foram obtidos do Centro de Bioterismo da Universidade Federal de Minas Gerais (CEBIO, UFMG) ou do Biotério da Universidade de Columbia (William Black Medical Research Building – New York, USA), onde eram mantidos em gaiolas, em ambiente com luminosidade e temperatura controlada (22 a 25°C), recebendo água e ração *ad libitum* e constantemente monitorados para identificação de possíveis infecções e zoonoses. O uso desses animais esteve de acordo com resolução do Comitê de Ética no Uso de Animais (CEUA – UFMG) – Protocolo N° 69/2014 e com as Políticas Institucionais de Cuidados e Uso Animal da Universidade de Columbia.

4.2 Obtenção e cultivo de macrófagos derivados de medula óssea (BMDMs)

Camundongos C57BL/6 eram previamente anestesiados e sedados com uma solução de quetamina (80 mg/Kg) e xilazina (15 mg/Kg) administrada via intraperitoneal e em seguida eutanasiados via deslocamento cervical. Os fêmures e tíbias eram cuidadosamente extraídos e adicionados a um tubo de polipropileno contendo PBS 1x estéril (do inglês *Phosphate buffered saline* - NaCl 140mM; Na₂HPO₄ 10mM; KH₂PO₄ 1mM; KCl 2mM) e, posteriormente, deixados em etanol 70% por 1 minuto. Em fluxo laminar, utilizando-se meio RPMI 1640 (LGC Biotecnologia) e uma seringa com agulha 0,70 x 25 mm, era realizado o *flush* das células da medula em um tubo de polipropileno estéril. Para a diferenciação dos macrófagos, as células eram centrifugadas a 200 x g por 10 minutos a 4°C e ressuspensas em meio de diferenciação (R20/30) contendo 10 mL de meio RPMI 1640 suplementado com 20% soro fetal bovino (SFB; Gibco - USA), 30% de sobrenadante de cultura de células L929 (fonte de M-CSF), 2 mM de glutamina (Gibco), 100 U/mL de penicilina, 100 µg/mL de estreptomicina (Gibco - USA) e 50µM de 2-mercaptoetanol (Vetec, Brasil). As células eram contadas pelo método de exclusão com Azul de Trypan (Gibco®), e 10 mL de uma suspensão de 7 x 10⁶ células eram

plaqueados em placas de Petri de poliestireno (JProlab®) e incubadas a 37°C e 5% CO₂. O meio de diferenciação era trocado a cada 2-3 dias, e ao sétimo dia em cultura, o meio era desprezado e as células eram lavadas com 5 mL de PBS 1x estéril para eliminação dos *debris* e das células não aderentes. As células aderidas na placa, já diferenciadas em BMDMs (do inglês *Bone marrow derived macrophages*) eram coletadas adicionando-se 3 mL de PBS-EDTA 10mM gelado à cultura, seguido de incubação no gelo por 5 minutos e ressuspensão. Os BMDMs eram centrifugados a 200 x g por 5 minutos a 4 °C, ressuspensos em meio de manutenção (R10/5) contendo 5mL de RPMI 1640, 10% de SBF, 5% de sobrenadante de L929, 2 mM de glutamina, 100 U/mL de penicilina, 100 µg/mL de estreptomicina e 50µM de 2-mercaptoetanol. Finalmente as células eram contadas e plaqueadas de acordo com a necessidade para a realização dos experimentos.

4.3 Cultura de células imortalizadas

Foram utilizadas nesse trabalho as linhagens de células de embrião de rim humano HEK293 (cedido por Dr. Oscar B. Romero, ICB/UFMG), fibroblastos murinos L929 (cedidos pelo Laboratório de Vírus, ICB/UFMG), macrófagos murinos Raw264.7 (cedidos por Dr. Bryan Williams, The Cleveland Clinic Foundation, Cleveland, USA) e macrófagos murinos derivados de medula óssea selvagens (WT iBMDMs) ou deficientes de RelA (RelA -/- iBMDMs).

Para o cultivo de HEK293 e L929, as células eram retiradas do estoque armazenado em nitrogênio líquido, descongeladas imediatamente em banho de água a 37°C e transferidas para garrafas de superfície tratada para cultura de células (TPP, 25 cm²) contendo 5 mL de meio DMEM (Meio Eagle Modificado Dulbecco's, LGC), suplementado com 10% de soro fetal bovino (Gibco - USA) e antibiótico (100ug/mL estreptomicina / 100 U/mL penicilina – Gibco - USA). Após cerca de 4 horas em estufa a 37°C e 5% de CO₂ o meio era trocado para eliminação de *debris* celulares e DMSO proveniente do meio de congelamento. Ao atingirem a subconfluência (entre 80-90%) era realizada a passagem para uma garrafa de maior área (TPP, 75 cm²) para a manutenção da cultura. Para a realização das passagens, o meio era desprezado e a monocamada de células era lavada com PBS 1X. Em seguida, 1,5mL de Tripsina (0,5%) – EDTA (Gibco) era adicionado, e a garrafa era incubada a 37°C por 1 minuto. Posteriormente, 4,5 mL de meio DMEM completo fresco era adicionado e as células eram ressuspensas em um volume total de 6 mL. Parte da suspensão era desprezada e o restante era ressuspenso em um volume final de 18 mL. Para a manutenção das linhagens o estoque era

constantemente renovado através do congelamento das culturas. Ao atingirem entre 80 a 90% de confluência, as células eram lavadas com PBS e tripsinizadas como já descrito, e então ressuspensas em meio de congelamento contendo 90% de soro fetal bovino e 10% de DMSO (Hybri-Max, Sigma). Um mililitro dessa suspensão era adicionado em criotubo (TPP) e transferido para um freezer a -80°C . Após 24 horas, os tubos eram transferidos para botijão de nitrogênio líquido a -196°C .

As células L929 eram cultivadas para obtenção do sobrenadante rico em M-CSF (do inglês, *Macrophage colony-stimulating factor*), essencial na diferenciação de macrófagos, utilizado na diferenciação de BMDMs (**tópico 4.2**). Brevemente, ao atingirem a subconfluência, essas células eram mantidas por 10 dias em cultura com 30 mL de meio DMEM completo. Ao final desse período o sobrenadante era coletado, centrifugado a $400 \times g$, 4°C por 10 minutos, filtrado em filtro estéril de $0.22 \mu\text{m}$, e utilizado para suplementação dos meios de diferenciação e manutenção de BMDMs ou congelado a -80°C .

As células RelA^{-/-} foram gentilmente cedidas pelo Dr. Sankar Ghosh (Columbia University, New York – USA). Brevemente, essas células foram geradas através da imortalização de BMDMs provenientes de camundongos C57BL/6 WT ou *Rela*^{Flox/Flox}/Cre-ERT2⁺ pela infecção com retrovírus recombinante J2, como previamente descrito (Blasi et al., 1989). Para a deleção do alelo *Rela*, os iBMDMs *Rela*^{Flox/Flox}/Cre-ERT2⁺ foram tratados com 100 nM de 4-OH-tamoxifeno (Sigma H6278) por 2 dias para ativação da Cre, gerando a linhagem de RelA^{-/-} iBMDMs. Seu cultivo, assim como dos macrófagos Raw264.7 seguiu o mesmo procedimento descrito para HEK293 e L929, porém as células eram cultivadas sem antibióticos e em garrafas NUNC®, Delta Surface®; ou Sarsted® sem tratamento.

4.4 Detecção de contaminação por *Mycoplasma ssp.*

Cada lote de células utilizado era previamente submetido a testes para a detecção de contaminação por *Mycoplasma ssp.* através de PCR para amplificação do rRNA 16s conservado entre as espécies de *Mycoplasma* (van Kuppeveld et al., 1994). Para isso, as células eram cultivadas, como descrito no item 4.3, em meio livre de antibióticos por no mínimo 48 horas. Após esse período o meio era desprezado, a cultura era lavada com PBS e tripsinizada como descrito, e uma alíquota de 500 μL da suspensão era coletada. Essa alíquota era incubada em banho seco a 100°C por 10 minutos para o rompimento celular e extravasamento de DNA. As amostras eram brevemente centrifugadas e as reações eram

preparadas com uma mistura de 2µL da amostra, 0,125 µL da enzima Taq DNA polimerase (5U/µL, Phoneutria), 2,5 µL de Tampão IO 10X (Phoneutria), 0,75 µL de MgCl₂ 50mM (Phoneutria), 0,5 µL de uma mistura de dNTPs (dATP, dCTP, dGTP, dTTP) a 10mM (Invitrogen), 16,625 µL de H₂O Milli-Q autoclavada, e 1,25 µL a 10µM dos oligonucleotídeos senso GPO-3 e antisenso MGSO (**Tabela 2**). As condições termocíclicas são descritas na **Tabela 2**. Em seguida, após a adição de tampão Loading Buffer 1X (Tris-HCl 10mM pH 7.5, EDTA 50mM pH 8.0, glicerol 10%, Bromophenol blue 0,25% e Xylenecyanol 0,25%) as amostras foram fracionadas a 100 V em gel de agarose 1,5 % preparado em solução tampão TAE 1X (0,04M Tris-acetato, 0,001M EDTA) contendo brometo de etídeo (0,5µg/mL), e as imagens foram capturadas através de câmera acoplada a sistema de fotodocumentação UV (MultiDoc-It™, UVP).

4.5 Extração do RNA total

BMDMs e células Raw264.7 eram plaqueados a 2×10^5 células/poço em placas de superfície tratada de 24 wells (JET BIOFIL). Após 24 horas de plaqueamento era feita a estimulação com 100 ng/mL de LPS (O55:B5 – Sigma-Aldrich, EUA), 10µg/mL de Poly;IC (Invivogen), ou 1µM de CpG ODN 1826 (Invivogen) durante os tempos indicados nas figuras. O meio de cultura era então desprezado e imediatamente era adicionado 500µL do reagente TRIzol (Invitrogen), seguido de homogeneização por 15 minutos em shaker orbital. O conteúdo era coletado e as amostras eram processadas para extração do RNA total pelo método do TRIzol seguindo-se o procedimento descrito pelo fabricante. Brevemente, eram adicionados 0,1 mL de clorofórmio (Vetec), seguida de agitação vigorosa por inversão por 15 segundos e incubação à temperatura ambiente durante 3 minutos. Em seguida, a amostra era centrifugada a 12.000 x g por 15 minutos, obtendo-se, ao final, uma fase aquosa (superior) e uma fase orgânica (inferior). A fase aquosa era coletada e então adicionados 0,25 mL de álcool isopropílico (Sigma) para a precipitação do RNA. A amostra era incubada à temperatura ambiente durante 10 minutos e centrifugada a 12.000 x g por 10 minutos. O sobrenadante era descartado e o RNA precipitado era lavado com 0,5 mL de etanol 75% por agitação suave, seguida de centrifugação a 7.500 x g por 5 minutos. O sobrenadante era descartado e o RNA totalmente solubilizado em água DEPC (do inglês Diethylpyrocarbonate - Invitrogen) com o auxílio de uma micropipeta. Em seguida as amostras eram incubadas a 55°C durante 10 minutos e posteriormente estocadas a -80°C. Para a quantificação do RNA

total, 2 μL de cada amostra era lida em espectrofotômetro NanoDrop Lite (Thermo Scientific - EUA).

4.6 Obtenção da primeira fita de cDNA

Um micrograma do RNA total era adicionado a uma solução contendo 0,5 μg de oligo-dT₁₈ (Prodimol), 1 μL de uma mistura de dNTPs a 10mM (dATP; dCTP; dGTP; dTTP) (Invitrogen) e água Milli-Q em um volume final 13 μL . Após incubação a 65°C durante 5 minutos e 2 minutos em gelo, eram adicionados 7 μL de um mix contendo 4 μL de tampão *M-MLV First strand buffer* (Invitrogen), 2 μL de DTT 0.1 M (Invitrogen), e 1 μL (200U) da enzima transcriptase reversa MMLV-RT (Invitrogen). A mistura era incubada a 37°C por 50 minutos para obtenção do cDNA e posteriormente a 70°C durante 15 minutos para inativação da MMLV-RT. Alternativamente, devido à sua maior eficiência e fidelidade, 1 μL (200U) da enzima Superscript II (Invitrogen) foi adicionado e a reação incubada a 42°C por 50 minutos para a geração do cDNA que foi utilizado como molde nas reações de *nested* RT-PCR.

4.7 RT-PCR quantitativa em tempo real (RT-qPCR)

Para análise dos níveis de mRNA, 2 μL de uma diluição 1:10 de cDNA era utilizado como DNA molde para a reação de RT-qPCR juntamente a uma mistura contendo 5 μL de iQTM SYBR® Green Supermix (Bio-Rad), 2 μL de água milli-Q e 0,5 μL de cada iniciador a 10 mM (**Tabela 2**). Os procedimentos eram realizados em placas de 96 poços (Axygen), que eram seladas, vortexadas e centrifugadas a 1000 x g por 1 minutos antes de serem transferidas para termociclador de tempo real (CFX96 Touch, Bio-Rad). Os dados eram analisados no CFX Manager Software (Bio-Rad) e salvos em formato Microsoft Excel para realização dos cálculos. Condições termocíclicas estão descritas na **Tabela 2**. A especificidade das reações era confirmada através da curva de *melting*. Os dados de expressão foram normalizados pelo gene de referência *Rpl32* (proteína ribossomal L32) e os níveis relativos de expressão foram analisados de acordo com a equação $2^{-\Delta\text{Ct}}$, onde $\Delta\text{Ct} = \text{Ct}(\text{gene alvo}) - \text{Ct}(\text{gene normalizador})$ de acordo com o método de Livak e Schmittgen, 2001 (Livak and Schmittgen, 2001).

4.8 *Nested* RT-PCR

Uma primeira RT-PCR foi realizada utilizando-se uma mistura de 2 μ L de cDNA de células Raw264.7 tratadas ou não com 100 ng/mL de LPS, 0,125 μ L da enzima Taq DNA polimerase (5U/ μ L, Phoneutria), 2,5 μ L de Tampão IO 10X (Phoneutria), 0,75 μ L de MgCl₂ 50mM (Phoneutria), 0,5 μ L de uma mistura de dNTPs (dATP, dCTP, dGTP, dTTP) a 10mM (Invitrogen), 16,625 μ L de H₂O Milli-Q, e 1,25 μ L a 10 μ M dos iniciadores senso “Fwd+1” e antisenso “Rev 1” (**Tabela 2**). A reação foi fracionada em gel de agarose 1,5 %, a região do gel na altura da banda esperada foi cisada, o DNA foi purificado utilizando-se o “QIAquick Gel Extraction Kit” (Qiagen) e uma segunda reação de RT-PCR (*nested* RT-PCR) foi realizada a partir do produto purificado utilizando-se os iniciadores internos senso “Nested +22” e antisenso “Nested Rev2” (**Tabela 2**). As condições termocíclicas são descritas na **Tabela 2**. Em seguida, após a adição de tampão Loading Buffer 1X (Tris-HCl 10mM pH 7.5, EDTA 50mM pH 8.0, glicerol 10%, Bromophenol blue 0,25% e Xylenecyanol 0,25%) as amostras foram fracionadas a 100 V em gel de agarose 1,5 % preparado em solução tampão TAE 1X (0,04M Tris-acetato, 0,001M EDTA) contendo brometo de etídeo (0,5 μ g/mL), e as imagens foram capturadas através de câmera acoplada a sistema de fotodocumentação UV (MultiDoc-It™, UVP).

4.9 Oligonucleotídeos e condições termocíclicas

Os oligonucleotídeos utilizados nesse trabalho assim como as condições termocíclicas utilizadas em suas respectivas reações estão listadas a seguir na **Tabela 2**:

Tabela 2: Listagem dos genes alvos, dos oligonucleotídeos utilizados para amplificação, e das condições termocíclicas utilizadas em cada reação.

Alvo e Reação	Iniciadores	Sequência 5' → 3'	Condições termocíclicas
rRNA 16 (<i>Mycoplasma ssp</i>) PCR	GPO-3 (fwd)	GGGAGCAAACAGGATTAGATACCCT	94°C/5'
	MGSO (rev)	TGCACCATCTGTCACTCTGTTAACCTC	94°C/30'' 55°C/30'' (35 ciclos) 72°C/30'' 72°C/5'
Rasgef1a RT-qPCR	qRT-mRGF1a (fwd)	GTCACCCGTTGTGGACAAAG	
	qRT-mRGF1a (rev)	TTGACCCGTTCCAGTTCAATG	
Rasgef1b RT-qPCR	qRT-mRGF1b (fwd)	GCTACAGCGAACGGAAGAA	
	qRT-mRGF1b (rev)	CTTCACAGGCATGCAGATTTTC	
Rasgef1c RT-qPCR	qRT-mRGF1c (fwd)	GCTGGAGCGACTAAGGCAC	
	qRT-mRGF1c (rev)	CTGTTGAACCACTTCACGTAGG	
Rpl32 RT-qPCR	qRT-RPL32 (fwd)	GCTGCCATCTGTTTTACGG	95°C/10'
	qRT-RPL32 (rev)	TGACTGGTGCCTGATGAACT	95°C/15'' (40 ciclos) 60°C/60''
Rela-p65 RT-qPCR	qRT-Rela (fwd)	CTTCACAGGCATGCAGATTTTC	
	qRT-Rela (rev)	CATAGGTCCTTTTGCGCTTC	
Il6 RT-qPCR	qRT-Il6 (fwd)	TAGTCCTTCTACCCCAATTTCC	
	qRT-Il6 (rev)	TTGGTCCTTAGCCACTCCTTC	
Cxcl1 RT-qPCR	qRT-Cxcl1 (fwd)	ACTGCACCCAAACCGAAGTC	
	qRT- Cxcl1 (rev)	TGGGGACACCTTTTAGCATCTT	
Tnf RT-qPCR	qRT-Tnf (fwd)	ACAGAAAGCATGATCCGCG	
	qRT-Tnf (rev)	GCCCCCATCTTTTGGG	
Rasgef1b RT-PCR	Fwd +1 (fwd)	GGCGGGCGGGGGCGTGGCCTG	94°C/4'
	Rev 1 (rev)	GGAATTCAGAGGGCGGCGAGTTTTTC	94°C/30'' 60°C/30'' (30 ciclos) 72°C/1'45'' 72°C/5'
Rasgef1b RT-PCR (<i>nested</i>)	Nested Fwd +22 (fwd)	CGGCAGGAGGCGGGCATTTC	94°C/4'
	Nested Rev 2 (rev)	CCAAGTGTGGATAAGGGCTTC	94°C/30'' 55°C/30'' (30 ciclos) 72°C/1'45'' 72°C/5'
Rasgef1b PCR	<u>Xho</u> -P-mFWD1	CCGCTCGAGTTTAAACATCCTTCGCAAAC	94°C/3'
	<u>Xho</u> -PD1	CCGCTCGAGCGGGAAAACCCGCCGATC	94°C/15'' 52°C/30'' (35 ciclos) 68°C/4' 68°C/10'

4.10 Análises *in silico*

As seqüências de aminoácidos de RasGEF1a, RasGEF1b e RasGEF1c foram recuperadas a partir do banco de dados UniProtKB (números de acesso E9Q1N6, Q8JZL7 e Q9D300, respectivamente) (2019), seu alinhamento se deu do programa Clustal Omega versão 1.2.1 (Sievers et al., 2011), e o dendograma foi gerado através do programa TreeFam (Li et al., 2006). A seqüência genômica a montante da seqüência codificadora de *Rasgef1b* foi recuperada a partir do banco de dados de seqüências nucleotídicas do National Center for Biotechnology Information (NCBI) (Número de acesso: NC_000071.6). Para análise dos potenciais elementos regulatórios presentes na seqüência foram utilizados os programas TFBind (parâmetros padrão) (Tsunoda and Takagi, 1999) e TFSearch 1.3 (*cutoff* ≥ 85 – atualmente fora de serviço) (Akiyama, 1995) baseados na 3.3 do banco de dados TRANSFAC, e Genomatrix MatInspector 8.21 baseado na Matrix Family Library versão 9.4 (*cutoff* ≥ 0.85 otimizado) (Quandt et al., 1995). Os elementos não-relacionados a resposta imune ou que não atenderam os requisitos mínimos de similaridade às matrizes foram ignorados. Para análises da presença de ilhas CpG foram utilizados os programas MethPrimer, EMBOSS-Cpgplot, e EMBOSS-Newcpgreport (Li and Dahiya, 2002; Rice et al., 2000). As análises comparativas de regiões não-codificadoras conservadas entre o gene RasGEF1b murino e seus ortólogos foi realizada através dos programas BLASTn-Megablast descontínuo (Altschul et al., 1990), mVISTA (Frazer et al., 2004), e das ferramentas multiz/phyloP do navegador de genomas UCSC (Blanchette et al., 2004; Chiaromonte et al., 2002; Schwartz et al., 2003; Siepel et al., 2005). Os dados de *protein binding microarray* para a ligação de dímeros NF- κ B às sondas κ B foram recuperados do banco de dados público “NF- κ B PBM Dataset” (Siggers et al., 2011). Os dados de ChIP-seq, assim como as faixas de hipersensibilidade por DNaseI-seq foram recuperadas do banco de dados ENCODE /LICR ou ENCODE/University of Washington, respectivamente, utilizando-se o navegador do genoma UCSC (Consortium, 2012; UCSC, 2012)

4.11 Amplificação e clonagem do promotor putativo de *Rasgef1b*

Para a construção dos plasmídeos repórter contendo parte da seqüência genômica a montante da seqüência codificadora de *Rasgef1b*, um segmento de 3.000 pares de bases (-2881 a +119, onde a posição +1 corresponde ao sítio de início da transcrição - TSS) foi amplificado a partir de DNA genômico de macrófagos peritoniais de camundongos

C57BL/6. A reação foi preparada em um volume final de 25 μL de solução contendo 2,5 μL do tampão *Pfx* Amplification Buffer 10X (Invitrogen), 0,75 μL de mistura de dNTPs 10 mM (Invitrogen), 0,5 μL de MgSO_4 50 mM (Phoneutria), 2,5 μL de PCRx Enhancer Solution 10X (Invitrogen), 0,5 μL da Platinum Pfx DNA polimerase (Invitrogen), 1,25 μL dos oligonucleotídeos senso XHO-P-mFWD1 e antisenso XHO-PD1 a 10 μM , e 2 μL de DNA genômico murino. As condições termocíclicas estão descritas na **Tabela 2**. Ao final da reação, foram adicionados 0,5 μL de dATP 10mM e 0,33 μL Taq DNA polimerase 5 U/ μL (Phoneutria), seguidos de incubação a 72°C por 15 minutos.

O amplicon foi clonado inicialmente no plasmídeo pGEM-T-Easy® (Promega). Para isso foram misturados 5 μL do tampão Rapid ligation 2X T4 DNA ligase (Promega), 1 μL plasmídeo pGEM-T-Easy, 3 μL do amplicon e 1 μL de T4 DNA ligase (Promega), totalizando 10 μL , e a reação foi incubada por 16 horas a 4°C. O produto da reação foi então transformado em bactérias quimiocompetentes XL1-Blue, para replicação do plasmídeo e obtenção do inserto em maior quantidade, como descrito nos **tópicos 4.13 e 4.14**. A confirmação da clonagem se deu através de digestão com as enzimas *XhoI* (Promega) e *XmnI* (Promega) (**tópico 4.16**) e posterior fracionamento em gel de agarose 1,5 % preparado em solução tampão TAE 1X (0,04M Tris-acetato, 0,001M EDTA) contendo brometo de etídeo (0,5 $\mu\text{g}/\text{mL}$), e as imagens foram capturadas através de câmera acoplada a sistema de fotodocumentação UV MultiDoc-It™ UVP mostrado anteriormente em (Leão, 2013).

A fim de viabilizar a clonagem unidirecional do inserto no repórter de expressão da luciferase pGL3-Basic (Promega), procedeu-se uma reação de digestão com as enzimas *KpnI* e *XhoI*. Através dessa reação, 134 nucleotídeos foram removidos da extremidade 5' do inserto, gerando um segmento de 2.866 pb (-2747/+119). O vetor pGL3-basic foi previamente preparado para a clonagem através da digestão com as mesmas enzimas e posteriormente incubado com a enzima CIAP 1/ μL (do inglês *Calf Intestinal Alkaline Phosphatase* – Promega) seguindo-se o protocolo do fabricante. Para a ligação do inserto foram preparadas reações contendo 4 μL (150 ng) de pGL3-Basic, 4 μL de Ligation Buffer 5x (Invitrogen) 1 μL de T4 DNA Ligase (1 U/ μL – Invitrogen), e inserto nas proporções molares 3:1, 2:1 e 1:1 (inserto:vetor). As reações foram completadas com água Milli-Q para um volume final de 20 μL e incubadas por 60 minutos a 25°C. O inserto foi finalmente clonado no sítio múltiplo de clonagem do plasmídeo pGL3-Basic entre os sítios de *KpnI* e *XhoI*, procedeu-se então a transformação em

bactérias XL1-Blue e finalmente a obtenção do plasmídeo pGL3-2774/+119 (2.8 kb) em média escala como descrito nos **tópicos 4.13 e 4.14**.

4.12 Geração de mutantes por deleção e síntese de DNA

Uma série de deleções a 5' e 3' foram realizadas a partir do plasmídeo pGL3-2747/+119 (2.8 kb) através das seguintes reações de digestão com enzimas de restrição: para gerar a construção pGL3-Δ487 (-2.747/-368) utilizamos a enzima *BglIII* (Promega). A amostra foi fracionada em gel de agarose 1% e a banda de interesse foi recuperada e purificada utilizando-se o kit “QIAquick Gel Extraction Kit” da QIAGEN. O plasmídeo foi religado através de incubação com 1 μL de T4 DNA Ligase (1 U/μL – Invitrogen) por 60 minutos a 25°C. Para gerar a construção pGL3-1621/+119 (1.7 kb) foi realizada uma digestão dupla com as enzimas *EcoRI* e *KpnI*. Para a construção do pGL3-706/+119 (0.8 kb) foi realizada uma digestão dupla com as enzimas *MluI* e *KpnI*. As amostras foram fracionadas em gel de agarose 1% e os fragmentos de interesse foram recuperados e purificados utilizando-se o kit “QIAquick Gel Extraction Kit” da QIAGEN. Uma vez que essas digestões geraram extremidades coesivas, mas não complementares, as amostras foram submetidas ao tratamento com T4 DNA Polimerase. Para isso, 10 μL de cada amostra foi misturado a 5 unidades de T4 DNA polimerase (Invitrogen), 0,6 μL de dNTPs 10mM (Invitrogen), 12 μL de 5x T4 DNA Polymerase Buffer (Invitrogen) e 6μL de BSA 10x (Invitrogen) em um volume final de 60 μL completado com água Milli-Q. As reações foram incubadas por 15 minutos a 11°C, 5 minutos a 37°C e finalmente por 10 minutos a 75°C para inativação da enzima. Os plasmídeos foram religados através de incubação com 1 μL de T4 DNA Ligase (1 U/μL – Invitrogen) por 60 minutos a 25°C. A construção pGL3-183/+119 (0.302 kb) assim como as construções contendo mutação nos sítios κB foram geradas através de síntese de DNA pela empresa Genescript (USA). As seguintes mutações κB foram realizadas (bases mutadas em negrito e minúsculo): κB#1- **GGGTTTTCCC** / **tctTTTTCaag**; κB#2- **TGGAGTTTCA** / **TGGAGTaaCt**; κB#3- **GGGCATTTCCC** / **GttCATTTCaag**; κB#4- **GGCAAGTCCC** / **ttCAAGTaaC**; κB#5 **CGGGAGTCCC**/ **CGttAGTaCC**. As construções geradas foram transformadas em bactérias quimiocompetentes XL1-Blue e obtidas em pequena e média escala como descrito nos **tópicos 4.13 e 4.14**. Para confirmação as construções foram sequenciadas e também realizadas digestões seguidas de fracionamento em gel de agarose 1,5 % preparado em solução tampão TAE 1X (0,04M Tris-acetato, 0,001M EDTA) contendo

brometo de etídeo ($0,5\mu\text{g}/\text{mL}$), e as imagens foram capturadas através de câmera acoplada a sistema de fotodocumentação UV MultiDoc-It™ UVP.

4.13 Transformação bacteriana

Os plasmídeos utilizados nesse trabalho foram transformados em bactérias quimiocompetentes XL1-Blue para replicação e posterior obtenção dos mesmos em pequena e média escala. Para isso, as bactérias XL1-Blue quimiocompetentes eram retiradas do botijão de nitrogênio líquido e deixadas em gelo para descongelamento. Dez nanogramas de plasmídeo eram misturados a $100\mu\text{L}$ da suspensão de bactérias. A amostra era incubada durante 30 minutos em gelo, seguido de choque térmico em banho de água a 37°C por 45 segundos e 2 minutos em gelo. Em seguida, $900\mu\text{L}$ de meio Luria Bertani (LB - Triptona $10\text{g}/\text{L}$; NaCl $5\text{g}/\text{L}$; extrato de levedura $5\text{g}/\text{L}$) sem antibiótico era adicionado e a amostra era incubada por 1 hora, 37°C , e 220rpm em Shaker Orbital (MA410 – Marconi). A amostra era então centrifugada a $8.000 \times g$ por 1 minuto, o sobrenadante era descartado, o sedimento era ressuscitado em $100\mu\text{L}$ de LB líquido e plaqueado em placas de petri contendo ágar a 1,5% preparado em meio LB com $100\mu\text{g}/\mu\text{L}$ de Ampicilina (Sigma-Aldrich). As placas eram incubadas em estufa a 37°C por 16 horas.

4.14 Obtenção de DNA plasmidiano em pequena e média escala (midiprep)

Colônias eram selecionadas das placas de transformação e individualmente inoculadas em tubos de polipropileno (15mL) contendo meio LB com Ampicilina a $100\mu\text{g}/\text{mL}$. Os tubos eram incubados sob agitação a 37°C por 16 horas. Para a extração dos plasmídeos em pequena escala o “Kit QIA prep Miniprep” (QIAGEN) era utilizado de acordo com o protocolo do fabricante. Para a obtenção dos plasmídeos em média escala eram utilizados os kits “GenElute™ HP Plasmid Midiprep” (Sigma-Aldrich) e “GenElute Endotoxin-free Plasmid Midiprep” (Sigma-Aldrich) – para obtenção de DNA plasmidiano livre de endotoxinas para os experimentos realizados em células Raw264.7 – de acordo com o protocolo do fabricante.

4.15 Digestão com enzimas de restrição

Alíquotas de $10\mu\text{L}$ de plasmídeo – concentração entre 0,2 e $0,5\mu\text{g}/\mu\text{L}$ – eram adicionadas às soluções contendo $2\mu\text{L}$ do tampão padrão de cada enzima (Promega), $1\mu\text{L}$ de enzima de restrição ($10\text{U}/\mu\text{L}$), $2\mu\text{L}$ de BSA ($10\text{mg}/\text{mL}$) e água Milli-Q para um

volume final de 20 μL . As reações eram incubadas a 37°C por 90 minutos em banho de água. Para digestões duplas em que as enzimas não compartilhavam do mesmo tampão padrão, a primeira digestão era realizada com a enzima cujo tampão padrão contém a menor quantidade de sal. Em seguida era realizada a inativação da enzima a 65° C por 15 minutos e, então, era adicionado 1 μL da enzima de restrição adicional, 3 μL de seu tampão padrão, 1 μL de BSA (10 mg/mL), e o volume era completado com água Milli-Q para 30 μL final seguido de nova incubação a 37°C por 90 minutos em banho de água.

4.16 Sequenciamento de DNA plasmidial

Os segmentos clonados no vetor pGL3-Basic (Promega, USA) foram sequenciados utilizando-se os oligonucleotídeos RVprimer3 e GLprimer2 (Promega), e 8 oligonucleotídeos adicionais desenhados através do programa Primer3 versão 0.4 (Untergasser et al., 2012), descritos na **Tabela 3**, por eletroforese capilar em aparelho ABI3130, utilizando-se polímero POP7 e BigDye v3.1 pela empresa Myleus Biotechnology. As sequências nucleotídicas foram comparadas à sequência referência do banco de dados do NCBI e confirmadas.

Tabela 3: Listagem dos oligonucleotídeos utilizados para o sequenciamento da região flaqueadora de *Rasgef1b* clonada no vetor pGL3-Basic.

Iniciadores	Sequência 5' → 3'
RVprimer3 fwd (Promega)	CTAGCAAATAGGCTGTCCC
SEQ- 2.8kb-FWD2	TGTATGTTGGGTCACCATGA
SEQ- 2.8kb-FWD3	CACTTTTGCACGCCTGAGTA
SEQ- 2.8kb-FWD4	TGGAGCTAGAAGGACCCTGA
SEQ- 2.8kb-REV1	CCTGATGTGGCTGATGGAAT
SEQ- 2.8kb-REV2	TGACGGGTCCCATGAGTTAT
SEQ- 2.8kb-REV3	CCTGTGAAACCCTCCTCAAG
SEQ- 2.8kb-REV4	CCTAAGCCCCGCAGTTTAT
SEQ- 2.8kb-REV5	TTTCGCAAAGGAAAGGAAGA
GLprimer2 rev (Promega)	ACTGACTGACTGACTGACTG

4.18 Transfecção transiente

Células HEK293 ou Raw264.7 eram plaqueadas a 1×10^5 ou 2×10^5 células/poço, respectivamente, em placas de 24 poços e as transfecções eram realizadas após 24 horas

de plaqueamento, quando as células apresentavam confluência 80-90%. Quatrocentos nanogramas do plasmídeo vazio pGL3-Basic, do promotor do gene da E-Selectina (pGL3-Esel), ou dos repórteres contendo segmentos da sequência flanqueadora de *Rasgef1b* eram transfectados juntamente com 100ng do plasmídeo pRL-TK que promove a expressão constitutiva da luciferase de *Renilla*, usado como normalizador nos ensaios de gene repórter. Onde indicado, as células foram transfectadas com 500 ng de plasmídeos para expressão de Mal/TIRAP, MyD88, TRIF, RelA, RelB ou cRel para avaliar os efeitos da ativação de TLRs e de NF- κ B na atividade dos segmentos. A quantidade de DNA transfectado era mantida constante pela transfecção do plasmídeo pFLAG-CMV4 (vetor vazio) nas células que não recebiam os plasmídeos de expressão citados. As transfecções foram realizadas através do método polímero catiônico por meio do uso do reagente MAX-PEI (Polietilenoimina, PolySciences, EUA). Brevemente, os plasmídeos a serem transfectados em cada poço eram misturados a 100 μ L de meio DMEM sem soro juntamente com o reagente MAX-PEI na proporção de 5:1 (μ L de MAX-PEI : μ g de DNA/poço). A mistura era vigorosamente vortexada por 15 segundos e incubada por 10 minutos à temperatura ambiente. Eram adicionados 200 μ L de DMEM suplementado com 10% de SFB, e o volume total de 300 μ L da mistura era adicionado a cada poço. As placas eram incubadas em estufa, e a transfecção ocorria por 24 horas até a obtenção do extrato para ensaio de gene repórter. No caso das células Raw264.7, após 24 horas de transfecção e previamente à obtenção do extrato, o meio era trocado por DMEM fresco, e as células eram estimuladas ou não com 100 ng/mL de LPS (O55:B5 – Sigma-Aldrich, EUA) por 24 horas.

4.19 Obtenção do extrato celular e ensaio de gene repórter

Para a obtenção do extrato celular, o meio de cultura era aspirado por sucção a vácuo e eram adicionados por poço 200 μ L de tampão de lise passiva (PLB 1X, do inglês *Passive Lysis Buffer*) fornecido pelo kit “Dual-Luciferase Reporter System” (Promega). A placa era então incubada à temperatura ambiente sob baixa agitação em shaker orbital por 15 minutos e então armazenada em freezer a -20°C até que a leitura da luminescência fosse realizada. Para isso, após descongelamento, as amostras eram transferidas para tubos de microcentrífuga de 1,5mL (Axygen Scientific) e centrifugadas a 8.000 x g por 5 minutos. Cinco microlitros dos extratos de HEK293 ou 20 μ L dos extratos de Raw264.7 eram adicionados por poço em uma placa de 96 wells opaca, de cor branca (FluoroNunc, Nunc, EUA). Em cada poço eram adicionados 40 μ L do reagente “LARII” (Promega),

seguido da leitura das amostras em luminômetro (LumiCount™, Packard Instruments). Imediatamente após a leitura da luciferase de vagalume (*Photinus pyralis*) a placa era retirada do aparelho e eram adicionados 40µL do reagente “Stop & Glo” (Promega) a cada poço para a leitura da luciferase de *Renilla*. Os valores obtidos das leituras eram registrados e salvos automaticamente pelo programa “Reader” (Packard Instruments, USA). A atividade relativa da luciferase de cada repórter transfectado era determinada e normalizada pela razão do valor da leitura da luciferase do Vagalume sobre o valor da leitura da luciferase da *Renilla* (dada pelo plasmídeo pRL-TK).

4.20 Plasmídeos

Os plasmídeos utilizados nesse trabalho estão listados a seguir:

Tabela 4: Listagem e descrição dos plasmídeos utilizados para transfecção.

Plasmídeos	Descrição
pFLAG-CMV4	Plasmídeo de expressão eucariótica contendo epítipo FLAG (Sigma, USA).
pFLAG-Mal/TIRAP	Contém sequência codificadora de Mal/TIRAP com epítipo FLAG (Fitzgerald et al, 2001).
pCMV-HA-MyD88	Contém sequência codificadora de MyD88 com epítipo HA sob o controle do promotor CMV (Addgene # 12287).
pEF-Bos TRIF Flag	Contém sequência codificadora de TRIF com epítipo Flag (Addgene # 41554).
pFLAG-RelA	Contém sequência codificadora do monômero p65 (RelA) de NF-κB com epítipo FLAG, gentilmente cedido por Daniel Mansur (UFSC, Brasil).
pFLAG-c-Rel	Contém sequência codificadora do monômero c-Rel de NF-κB com epítipo FLAG (Addgene #20013).
RelB cFlag pcDNA3	Contém sequência codificadora do monômero RelB de NF-κB com epítipo FLAG (Addgene #20017).
pcDNA3-p50-flag	Contém sequência codificadora do monômero p50 de NF-κB com epítipo FLAG, (Addgene #20018)
p52 cFlag pcDNA3	Contém sequência codificadora do monômero p52 de NF-κB com epítipo FLAG (Addgene #20019).
pGL3-2747/+119	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 2.866 nucleotídeos (-2747/+119) da sequência não-codificadora de <i>Rasgef1b</i> murino.
pGL3- Δ487	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 2.374 nucleotídeos (-2747/-368) da sequência não-codificadora de <i>Rasgef1b</i> murino.

pGL3-1621/+119	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 1740 nucleotídeos (-1621/+119) da sequência não-codificadora de <i>Rasgef1b</i> murino.
pGL3-706/+119	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 825 nucleotídeos (-706/+119) da sequência não-codificadora de <i>Rasgef1b</i> murino.
pGL3-183/+119	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 302 nucleotídeos (-183/+119) da sequência não-codificadora de <i>Rasgef1b</i> murino.
pGL3-183/+119 (κb #4 e #5 mut)	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 302 nucleotídeos (-183/+119) da sequência não-codificadora de <i>Rasgef1b</i> murino contendo os elementos κb a montante do TSS mutados (κb #4 e #5).
pGL3-183/+119 (κb #1, #2 e #3 mut)	Contém o gene da luciferase de vagalume (<i>Pluc</i>) sob o controle transcricional de um segmento de DNA de 302 nucleotídeos (-183/+119) da sequência não-codificadora de <i>Rasgef1b</i> murino contendo os elementos κb a jusante do TSS mutados (κb #1, #2 e #3).
pGL3-Basic	Plasmídeo comercial (Promega) contendo o gene da luciferase de vagalume <i>Photinus pyralis</i> , ausente de promotor.
pRL-TK	Contém o gene da luciferase de celenterado <i>Renilla</i> sob o controle transcricional do promotor do gene da timidina cinase do vírus HSV-1 (Promega).

4.21 Ensaio para cromatina acessível por transposase seguido de sequenciamento (ATAC-seq, do inglês *Transposase-Accessible Chromatin with high-throughput sequencing*)

Para análise de regiões de cromatina acessível ao longo do genoma de camundongo, adaptamos os métodos de ATAC-seq descritos previamente (Bao et al., 2015; Buenrostro et al., 2013). BMDMs murinos primários foram plaqueados a 5×10^4 células/poço em placa de 96 poços por 24 horas e então tratados ou não com 100 ng/mL de LPS (O55:B5 - Sigma) por 2 horas. O meio foi desprezado, a monocamada foi lavada duas vezes com PBS à temperatura ambiente, e as células foram lisadas com 50 µL de tampão de lise (10 mM Tris-HCl pH 7,4, 10 mM NaCl, 3 mM MgCl₂, 0,05% NP-40) em incubação por 10 minutos à temperatura ambiente. A placa foi centrifugada, o sobrenadante foi descartado, 50 µL do Master Mix de transposição foi adicionado por

poço (Illumina Nextera DNA Library Prep Kit - 25uL de tampão 2X Tagment DNA, 5uL da enzima transposase TDE1 Tagment DNA Enzyme, 20uL de água Mili-Q), a placa foi gentilmente homogeneizada e incubada por 30 minutos a 37 °C. O DNA tagmentado foi purificado utilizando-se o kit “MinElute PCR Purification” (QIAgen – USA), então amplificado através de PCR utilizando-se o Master Mix de PCR High-Fidelity 2X e 1.25 µM dos iniciadores com *barcode* “Nextera Customized PCR primers 1 and 2” descritos por Buenrostro et al. seguindo seu protocolo para amplificação (Buenrostro et al., 2013; Buenrostro et al., 2015). O DNA amplificado foi purificado com o kit “QIAquick PCR Purification” (QIAgen) e em sequência utilizando-se o kit “Agencourt AMPure XP PCR Purification” (Beckman Coulter - USA) baseado em esferas magnéticas, seguido de análise em chip “Bioanalyzer HS” (Agilent – USA) para garantir a seleção de tamanho dos fragmentos. As amostras foram sequenciadas pelo New York Genome Center com leituras pareadas (*paired-end*) de 50 pb. As sequências foram alinhadas ao à versão mm10 do genoma referência de camundongo e o programa IGV versão 2.4.14 (Robinson et al., 2011) foi utilizado para identificação dos picos e plotagem dos gráficos.

4.22 Ensaio de desvio de mobilidade eletroforética (EMSA, do inglês, *Electrophoretic mobility shift assay*)

Macrófagos Raw264.7 foram plaqueados em placas de petri de 15cm² (20 x 10⁶ células) por 24 horas, e então estimulados ou não com 100 ng/mL de LPS (O55:B5 – Sigma) por 30 minutos, 1 hora ou 3 horas. Em seguida procedeu-se a obtenção do extrato nuclear. Para isso, o meio foi removido, as células foram lavadas 2 vezes com PBS 1x, coletadas gentilmente com o auxílio de um removedor de células (*cell scraper*), e centrifugadas a 1.500 x g por 5 minutos. O sobrenadante foi descartado e as células foram ressuspensas em 500 µL de tampão hipotônico gelado (10 mM HEPES, 10 mM KCl, 0.1 mM EDTA, 0.1 mM EGTA, 1 mM DTT, pH 7.9) e imediatamente adicionado 2 mg/mL de leupeptina, 0.03 TiU/mL de aprotinina, 1 mg/mL de pepstatina, e 1 mM de PMSF. As células foram incubadas em gelo por 20 minutos, 50µL de solução Nonideto NP-40 (10%) foi adicionado e vortexado vigorosamente por 10 segundos. A solução foi centrifugada a 18.000 x g por 30 minutos a 4°C, o sobrenadante foi descartado e o precipitado ressuspensado em 200µL de tampão hipertônico gelado (20 mM HEPES, 0.4 M NaCl, 1 mM EDTA, 1 mM EGTA, 1 mM DTT, pH 7.9) com adição imediata de 2 µg/mL de leupeptina, 0.03 TiU/mL de aprotinina, 1 mg/mL de pepstatina, e 3 mM de PMSF. A solução foi homogeneizada em shaker orbital por 15 minutos a 4°C,

centrifugada a 18.000 x g por 5 minutos a 4°C e o sobrenadante coletado. A concentração de proteína foi determinada através de ensaio de BCA com o “Micro BCA™ Protein Assay Kit” (Thermo – USA).

Para a realização do EMSA, sondas contendo os elementos κB presentes no promotor de *Rasgef1b* (Tabela 5) ou no promotor do gene *H2k* do complexo de histocompatibilidade principal (Meininger et al., 2016) foram sintetizadas pela empresa Fisher Scientific (USA) com purificação por HPLC. Procedeu-se o anelamento dos oligonucleotídeos senso e anti-senso para geração da sonda de dupla fita incubando-se 5 µg de cada oligo em 50 µL de tampão de anelamento (50 mM Tris-HCl, pH 8, 70 mM NaCl) a 90 °C por 10 minutos, e posterior incubação à temperatura ambiente por 16 horas. O anelamento foi confirmado através de fracionamento das sondas em gel de poliacrilamida 12% em TBE 1x. (10,8 g Tris-base, 5,5 g ácido bórico, 7,5 g EDTA, 1 L água mili-Q, pH 8.3). Os procedimentos a seguir foram realizados utilizando-se os equipamentos proteção pessoal necessários, em ambiente adequado e de acordo com as normas do *Research Radiation Safety Program* da Universidade de Columbia (NY). Para marcação radioativa através de preenchimento de *overhang* 5' dos oligos, 1 µL da sonda (100 ng/µL) foi incubado com 2.5 µL do tampão NEB Buffer 2, 1.8 µL de dCTP, dGTP, dTTP (0.5 mM), 3 µL de dATP [α -32P] (10 µCi/µL), 1 µL da enzima DNA Pol I (Klenow) (NEB – USA), e água mili-Q para um volume final de 25 µL a 37°C por 30 minutos. As sondas marcadas foram purificadas utilizando-se o “QIAquick® Nucleotide Removal Kit” (QIAGEN) e a taxa de decaimento de cada sonda foi determinada utilizando-se o Tri-Carb® 2810TR LSA (Perkin-Elmer). As sondas foram incubadas (30.000 cpm por amostra) com 10 µL de tampão *2x Shift* (40 mM HEPES pH 7.9, 120 mM KCl, 8% Ficoll), 0,2 µL de BSA (10 mg/mL), 0,4 µL de DTT (100mM), 1 µL de Poly (dI:dC) (2 mg/mL), e 10 µg de extrato e água mili-Q para um volume final de 20 µL à temperatura ambiente por 30 minutos. Para análise de especificidade de ligação, uma sonda não-marcada contendo a sequência κB consenso (Santa Cruz Biotechnology – USA) foi adicionada à reação em excesso em relação à sonda marcada (50 ou 100 vezes mais). Para *supershift*, 1µg de anticorpos anti-p50 (sc-114x), anti-RelA (sc-8008), ou anti-cRel (sc-71x) (Santa Cruz Biotechnology – USA) foram pré-incubadas com o extrato nuclear por 30 minutos em gelo. As amostras foram então fracionadas a 200 V em gel de poliacrilamida não-desnaturante 5% em TBE1x (44 mL de água mili-Q, 6 mL TBE 10x, 10 mL acrilamida 30%, 450 µL APS, 45 µL TEMED). Ao final da corrida, o gel foi seco a 80 °C por 2 horas em “Gel Dryer Model 583” (BioRad), exposto por 16 horas em

“KODAK Phospor Screen”, e a imagem foi analisada no Typhoon Trio Variable Mode Imager (Amersham Biosciences).

Tabela 5: Sondas contendo os elementos κ B do promotor de Rasgef1b utilizados em EMSA

Sonda	Sequência 5' → 3'
κ B #1	(Fwd) gatcCTGGATCGGCGGGTTTTCCCTTCTCTGG (Rev) gatcCCAGAAGAAAGGGAAAACCCGCCGATCCAG
κ B #2	(Fwd) gatcCGGCTCGGGCTGGAGTTTCAGTATAAACAA (Rev) gatcTTGTTTATACTGAAACTCCAGCCCGAGCCG
κ B #3	(Fwd) gatcGGCAGGAGGCGGGCATTTCGCCAGCAAGGG (Rev) gatcCCCTTGCTGGGGAAATGCCCGCCTCCTGCC
κ B #4	(Fwd) gatcCGCGGAACCCTGGCAAGTCCCCGAGCGCCCG (Rev) gatcGCGGGCGCTCGGGGACTTGCCAGGGTTCCGCG
κ B #5	(Fwd) gatcCGCAGCGTCACGGGAGTCCCTACCCGCGGA (Rev) gatcTCCGCGGGTAGGGACTCCCGTGACGCTGCG
H2k	(Fwd) gatcCAGGGCTGGGGATTCCCATCTCCACAGG (Rev) gatcCCTGTGGAGATGGGGAATCCCCAGCCCTG
κ B consenso	(Fwd) gatcAGTTGAGGGGACTTTCCAGG (Rev) gatcCCTGGGAAAGTCCCTCAACT

4.23 Imunoprecipitação da cromatina seguido de sequenciamento (ChIP-seq, do inglês *Chromatin Immunoprecipitation Sequencing*)

BMDMs primários foram plaqueados em placas de petri de 15cm² (20 x 10⁶ células) por 24 horas, e então estimulados ou não com 100 ng/mL de LPS (O55:B5 – Sigma) por 2 horas. Procedia-se a fixação (*cross-link*) com 1% de formaldeído (Sigma) por 10 minutos à temperatura ambiente, seguido de interrupção da reação com glicina a 0.125M por 10 minutos e lavagem com PBS. As células fixadas eram lisadas em tampão de lise (25mM HEPES pH 7.8, 1.5mM MgCl₂, 10mM KCl, 0.1% NP-40) contendo 1mM de DTT e de inibidores de protease em gelo por 10 minutos. Os lisados eram centrifugados a 5.400 rpm por 5 minutos para obtenção do núcleo. O núcleo isolado era então lisado em tampão de lise nuclear (50mM HEPES pH 7.9, 1mM EDTA, 1% TX-100, 0.1% Na-DOC, 140mM NaCl, 1% SDS, e 5mM EGTA adicionado na hora) contendo inibidores de protease e então o lisado era sonificado utilizando-se Bioruptor

(Diagenode) para obtenção de fragmentos entre 150 a 300 pb. Um micrograma de Anticorpos anti-RelA (C-20, Santa Cruz Biotechnology) e anti-H3K27Ac (abcam ab4729) foram incubados com *Dynabeads* de proteína A (Invitrogen). Os lisados sonicados eram misturados com as *beads* ligadas aos anticorpos e imunoprecipitado a 4 °C por 16 horas. As *beads* foram lavadas e incubadas com RNase por 30 minutos a 37 °C, incubadas com proteinase K por 2 horas, seguido de incubação a 65 °C por 16 horas para reversão do *cross-link*. O DNA imunoprecipitado então purificado com “QIAquick PCR Purification kit” (QIAGEN). Pequenas alíquotas de DNA eram usadas para confirmar a eficiência da reação de ChIP. A concentração do DNA imunoprecipitado e do DNA *input* era determinada por Bioanalyzer (Agilent). A preparação da biblioteca e o sequenciamento foram realizados pelo Centro de Genoma JP Sulzberger da Universidade de Columbia e pela empresa Admera Health. As amostras foram sequenciadas em HiSeq (Illumina) com leituras pareadas (*paired-end*) de 150 pb. As sequências foram alinhadas ao à versão mm10 do genoma referência de camundongo e o programa IGV versão 2.4.14 (Robinson et al., 2011) foi utilizado para identificação dos picos e plotagem dos gráficos.

4.24 Análise estatística

As análises estatísticas e plotagem gráfica foram realizadas utilizando o software Prism 5.0 (GraphPad). Os resultados são dados como média \pm DP. Comparações entre dois grupos foram feitas usando o teste T de Student. Múltiplas comparações foram feitas com ANOVA ou Two-Way ANOVA seguido de pós-testes de Bonferroni. Significância estatística foi determinada como * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Os experimentos de RT-qPCR foram realizados em triplicata técnica, os ensaios de gene repórter foram realizados em duplicata técnica. As réplicas biológicas e repetições de experimentos independentes variavam e estão indicados em cada figura.

5. RESULTADOS

5.1 *Rasgef1b* é preferencialmente expresso em macrófagos

A expressão de *Rasgef1b* é induzida em macrófagos peritoneais após ativação de TLRs e *in vivo* em camundongos infectados com protozoários parasitas (Andrade et al., 2010; Ferreira et al., 2002). No entanto, sua expressão em macrófagos murinos derivados de medula óssea (BMDMs), assim como na linhagem de macrófagos Raw264.7, é pouco conhecida. Uma vez que esses tipos celulares foram elencados como modelo em nossos estudos, avaliamos os níveis do mRNA de RasGEF1b nessas células estimuladas ou não com LPS através de RT-PCR quantitativa em tempo real (RT-qPCR). Como mostrado na **figura 7**, a estimulação com LPS rapidamente induz a um aumento de até 7 vezes nos níveis do mRNA de RasGEF1b em BMDMs e até 9 vezes em macrófagos Raw264.7 (**Fig. 7**).

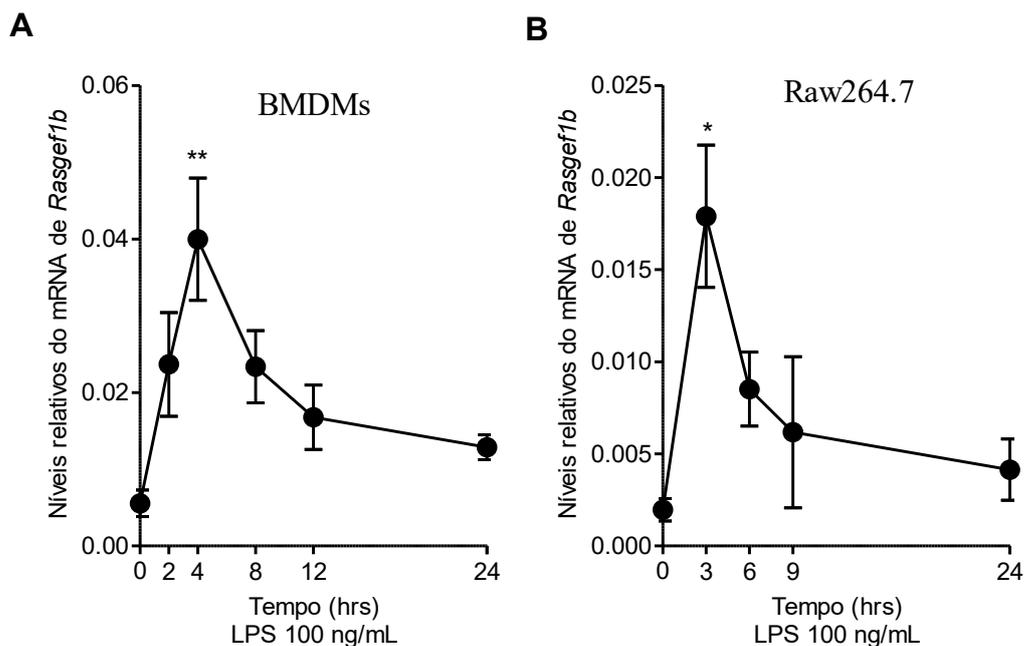


Figura 7: Análise da expressão de *Rasgef1b* induzida por LPS em macrófagos. Macrófagos murinos derivados da medula óssea (BMDMs) ou macrófagos Raw264.7 foram deixados sem tratamento (ctrl) ou estimulados com 100 ng/mL de LPS pelos tempos indicados. O RNA total foi extraído e os níveis relativos do mRNA de RasGEF1b foi avaliado por RT-qPCR, normalizados em relação ao gene referência Rpl32. Os dados representam a média \pm D.P. de pelo menos três experimentos independentes. * $p \leq 0,05$; ** $p \leq 0,01$; em relação ao não tratado (0hr), determinado através de One-way ANOVA seguido por pós-teste de Bonferroni.

O gene *Rasgef1b* é conservado evolutivamente, possuindo cerca de 240 potenciais ortólogos identificados em vertebrados (NCBI *orthologs*) e em alguns invertebrados como *C. elegans* e *D. melanogaster*. O produto gênico em humanos, por exemplo, compartilha cerca de 97% de identidade em relação ao RasGEF1b murino em sua sequência de aminoácidos. Além disso, RasGEF1b possui dois parálogos, RasGEF1a e RasGEF1c, intimamente relacionados por compartilharem relativa similaridade de sequência assim como os dois domínios característicos da família RasGEF (REM e CDC25) conservados (Andrade et al., 2010; Yaman et al., 2009). A sequência de aminoácidos de RasGEF1b murino apresenta aproximadamente 60% e 66% de identidade entre RasGEF1a e RasGEF1c, respectivamente (**Fig. 8**). Entretanto, sua expressão em macrófagos ativados é pouco conhecida, sobretudo em relação aos níveis bem estabelecidos de RasGEF1b. Elucidando essa questão, mostramos que os níveis basais do mRNA de RasGEF1b em BMDMs primários é 81 vezes mais elevada que RasGEF1a, e 1.953 vezes mais elevada que RasGEF1c (**Fig. 9A**). Assim como ocorre com RasGEF1b, RasGEF1c parece ser induzido após estímulo com LPS, enquanto RasGEF1a parece ser inibido (**Fig. 9B**). Esses resultados indicam que macrófagos expressam preferencialmente RasGEF1b em relação aos parálogos RasGEF1a e RasGEF1c, reforçando a hipótese de que esse GEF desempenha um papel na resposta imune inata mediada TLR em macrófagos

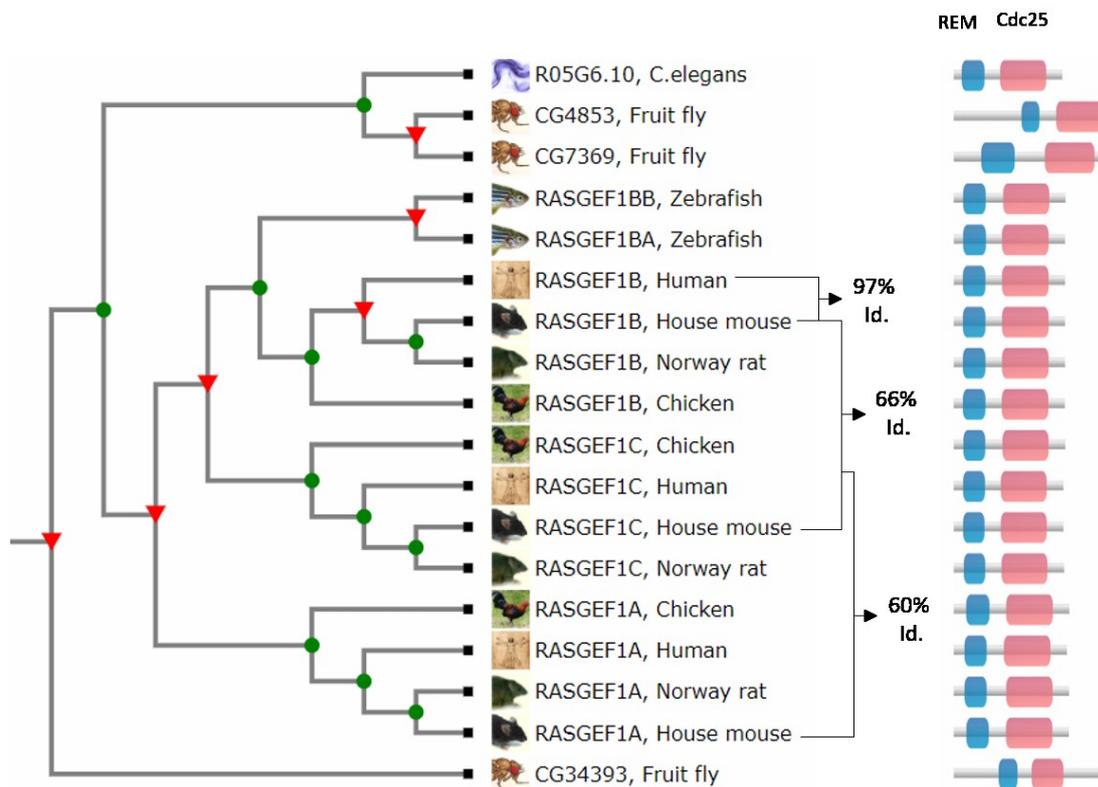


Figura 8: A homologia entre os membros da família RasGEF1. A sequência de aminoácidos de RasGEF1a, RasGEF1b e RasGEF1c de diferentes espécies foram alinhadas através do programa Clustal Omega 1.2.1 e o dendograma foi gerado através do programa TreeFam. A similaridade entre os ortólogos murino e humano de RasGEF1b e entre os parálogos RasGEF1a e RasGEF1c de camundongo está representada (Id = identidade). À direita, esquema representativo dos domínios REM e CDC25 e sua localização ao longo da sequência peptídica de cada membro RasGEF1 em diferentes espécies.

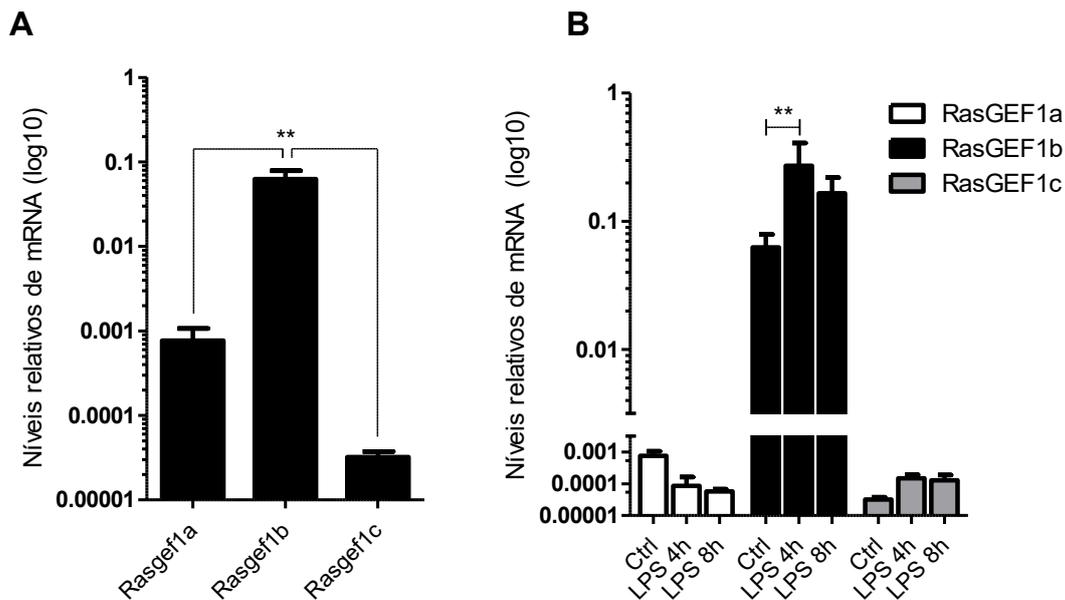


Figura 9: Análise da expressão basal e induzida por LPS de *Rasgef1a*, *Rasgef1b* e *Rasgef1c* em macrófagos. Macrófagos murinos derivados da medula óssea (BMDMs) foram deixados sem tratamento (A) (ctrl) ou estimulados com 100 ng/mL de LPS pelos tempos indicados (B). O RNA total foi extraído e os níveis relativos do mRNA de RasGEF1a, RasGEF1b e RasGEF1c foram avaliados por RT-qPCR, normalizados em relação ao gene referência Rpl32. Os dados representam a média \pm D.P. de pelo menos três experimentos independentes. ** $p \leq 0,01$; determinado através de One-way ANOVA (A) ou Two-way ANOVA (B) seguido por pós-teste de Bonferroni.

5.2 Análises *in silico* da sequência nucleotídica de *RasGEF1b* e identificação de potenciais elementos regulatórios

Análises *in silico* utilizando-se ferramentas de bioinformática podem fornecer informações importantes sobre sequências nucleotídicas e peptídicas ainda não caracterizadas, permitindo guiar o trabalho experimental e fornecendo subsídios para perguntas biológicas pertinentes (Jareborg et al., 1999). Nesse sentido, a sequência nucleotídica referência de *Rasgef1b* foi recuperada a partir do Genbank (NCBI Refseq: NC_000071.6) para a condução de análises sobre sua provável região regulatória transcricional. Em consistência com outros bancos de dados como Ensembl (número de acesso: ENSMUST00000031276.14) e o UCSC genome browser (número de acesso: uc008ygl.1) essa anotação curada sugere um nucleotídeo guanina localizado 9.687 pb a montante do sítio de início da tradução (ATG, localizado no exon 2) como o primeiro nucleotídeo transcrito de *Rasgef1b* (Fig. 10A e Fig. 11). Esse nucleotídeo considerado como o sítio de início da transcrição (TSS, posição relativa +1) foi utilizado no presente

trabalho como referência em relação ao início do mRNA de *Rasgef1b* para a numeração e mapeamento dos demais nucleotídeos a montante (região 5' flanqueadora) e a jusante.

Para fornecer evidências da posição +1 predita na anotação, realizamos reações de RT-PCR a partir do cDNA de Raw264.7 estimuladas ou não com LPS. Uma reação inicial foi realizada utilizando-se um iniciador senso (Fwd + 1) cuja extremidade 5' se alinha ao nucleotídeo sugerido como TSS (posição +1 até +21) e um iniciador reverso (Rev 1) que se alinha à região codificadora de *Rasgef1b* (posição +720 até +705) (**Fig. 10A**). Embora uma banda de 720 pb fosse esperada, observamos um rastro indicando o acúmulo de produtos inespecíficos amplificados, sobretudo na amostra proveniente de células estimuladas (**Fig. 10B**). O ocorrido se deve provavelmente em razão da sequência pouco específica do oligonucleotídeo senso, rica em repetições de Guanina e Citosina. Para fornecer um diagnóstico mais esclarecedor, foi realizada uma segunda PCR (*nested* PCR) a partir do produto purificado da primeira reação usando iniciadores senso e anti-senso internos (posição +22/+41 e +371/350, respectivamente). O resultado (**Fig. 10B**) revela que há a amplificação de um produto de tamanho esperado de 350 pb na amostra tratada com LPS, sugerindo que parte dos transcritos de RasGEF1b induzidos por LPS em macrófagos abriga a região proximal ao TSS anotado.

Em seguida, utilizamos as ferramentas TFsearch, TFBIND e MatInspector para investigar a presença de potenciais elementos regulatórios em uma ampla sequência de 2.8 kb que inclui 2.747 nucleotídeos a montante e 119 nucleotídeos a jusante do sítio de início da transcrição de *Rasgef1b* (-2747 a +119). Em concordância com o perfil de expressão de *Rasgef1b* em macrófagos, identificamos vários elementos relacionados à resposta imune como C/EBP (do inglês *CCAAT-enhancer-binding proteins*), STAT1 (do inglês, *Signal transducer and activator of transcription 1*), Sp1 (do inglês, *Specificity protein 1*), assim como sítios de ligação para NF- κ B e AP-1 (**Fig. 11**). Esses dados sugerem, portanto, um potencial papel dessa região na regulação da expressão de *Rasgef1b* durante a ativação de TLRs.

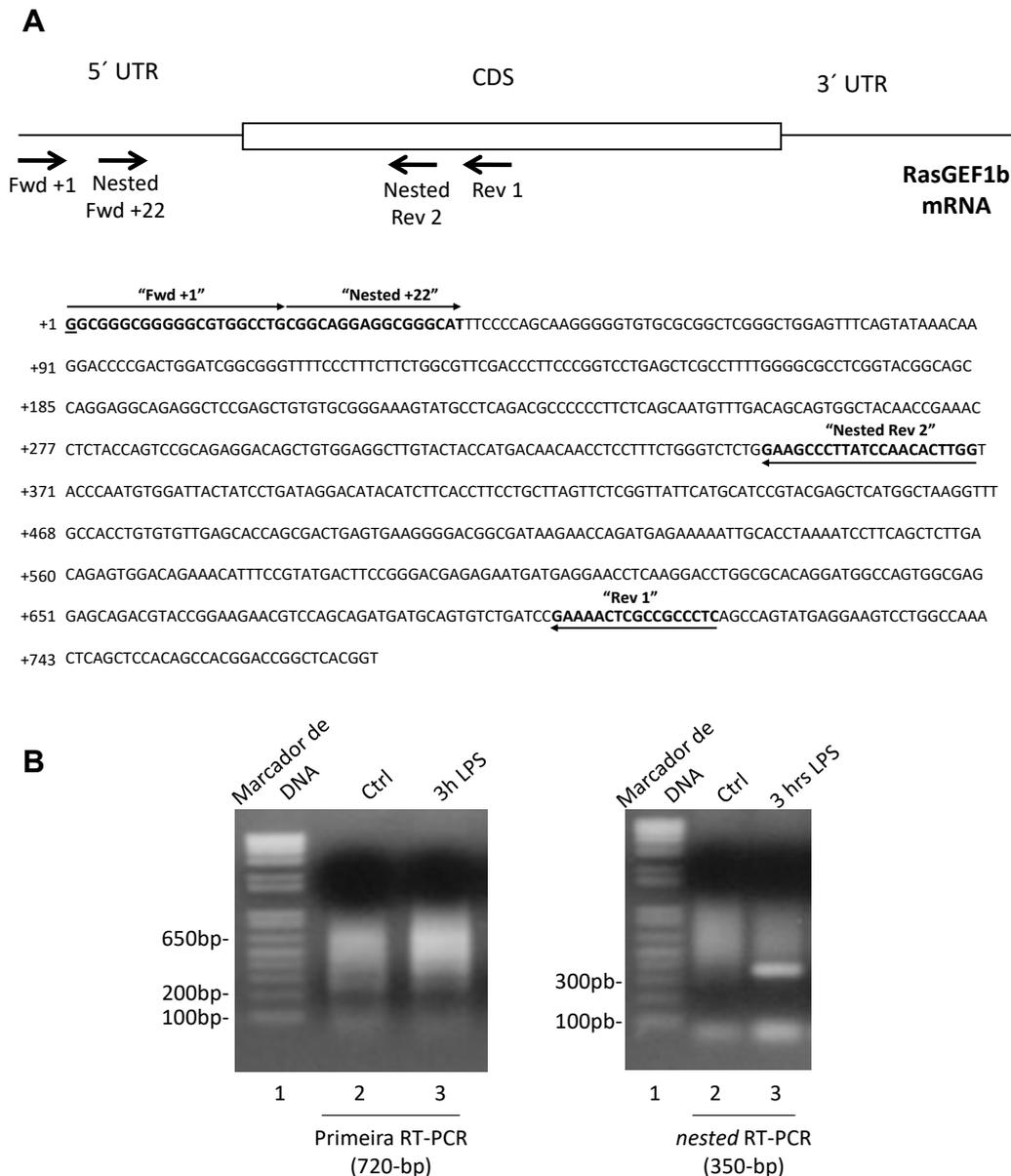


Figura 10: Validação do proposto TSS de *Rasgef1b* em macrófagos. Reações de RT-PCR foram conduzidas a partir do RNA total de macrófagos Raw264.7 estimulados ou não (ctrl) com 100ng/mL de LPS a fim de se avaliar a amplificação da extremidade 5' do mRNA de RasGEF1b. **(A)** Uma primeira reação de RT-PCR foi realizada utilizando-se o primer forward +1, cuja extremidade 5' se alinha ao sítio de início da transcrição (TSS) de *Rasgef1b* segundo a sequência referência mais atual (NCBI - Refseq: NC_000071.6), e o primer reverso "Rev 1" que se alinha na sequência codificadora de *Rasgef1b* (exon 5). Os primers internos "Nested fwd +22" e "Nested Rev 2" foram utilizados em uma segunda reação (*nested* PCR) a partir do produto purificado da primeira reação. **(B)** Fracionamento gel de agarose 1,5% para análise das reações. Uma única banda de 720 pb era esperada a partir da primeira reação, mas os resultados foram inconclusivos devido à amplificação inespecífica. Na segunda reação uma única banda induzível de 350 pb foi detectada nas amostras provenientes de células tratadas com LPS como esperado (canaleta 3), confirmando a amplificação da primeira reação. Marcador de DNA: Invitrogen 1 Kb Plus DNA Ladder

5.3 Identificação de regiões de cromatina acessível, ilhas CpG, e conservação evolutiva ao longo da sequência de *Rasgef1b*

Regiões de cromatina acessível e modificações pós-traducionais de nucleossomos como metilação e acetilação de histonas são amplamente utilizadas para a identificação de regiões regulatórias transcricionais, como os promotores e *enhancers* (Kim and Shiekhhattar, 2015; Thurman et al., 2012). Decidimos então avaliar essas características ao longo da sequência de *Rasgef1b* utilizando as técnicas de ATAC-seq e ChIP-seq anti-H3K27ac a partir do DNA genômico de BMDMs primários estimulados ou não com LPS. As análises de ATAC-seq revelaram uma região de cromatina acessível (baixa ocupação de nucleossomos) ao longo de parte do exon 1 e da sequência imediatamente à montante, aproximadamente entre +151 a -263 (**Fig. 12**). Adicionalmente, identificamos uma extensa região rica em nucleossomos contendo acetilação da lisina 27 na histona H3 (H3K27ac), aproximadamente entre +4219 a -1831. Curiosamente, identificamos um segmento adicional de cromatina acessível circundada por nucleossomos H3K27ac em uma região distal 3', a jusante do último exon de *Rasgef1b*. Observa-se ainda que o tratamento com LPS causou apenas um modesto aumento nos sinais de ATAC e H3K27ac ChIP, mas a estrutura da cromatina ao longo da sequência permaneceu praticamente inalterada após estímulo (**Fig. 12**). Em consistência com esses resultados, através de busca no banco de dados ENCODE identificamos marcadores epigênicos de regiões transcricionalmente ativas adicionais como H3K4me3 e H3K9ac, assim como dados de DNase-seq que confirmam o estado de cromatina acessível na região. Esses dados foram obtidos não só em BMDMs, mas também em uma grande variedade de células e tecidos como MEFs, linfócitos B e T, timo, fígado, coração, cérebro, rins, entre outros, e surpreendentemente, mostram ainda a ocupação constitutiva da RNA Polimerase II (RNAP II) na região (**Fig 13**).

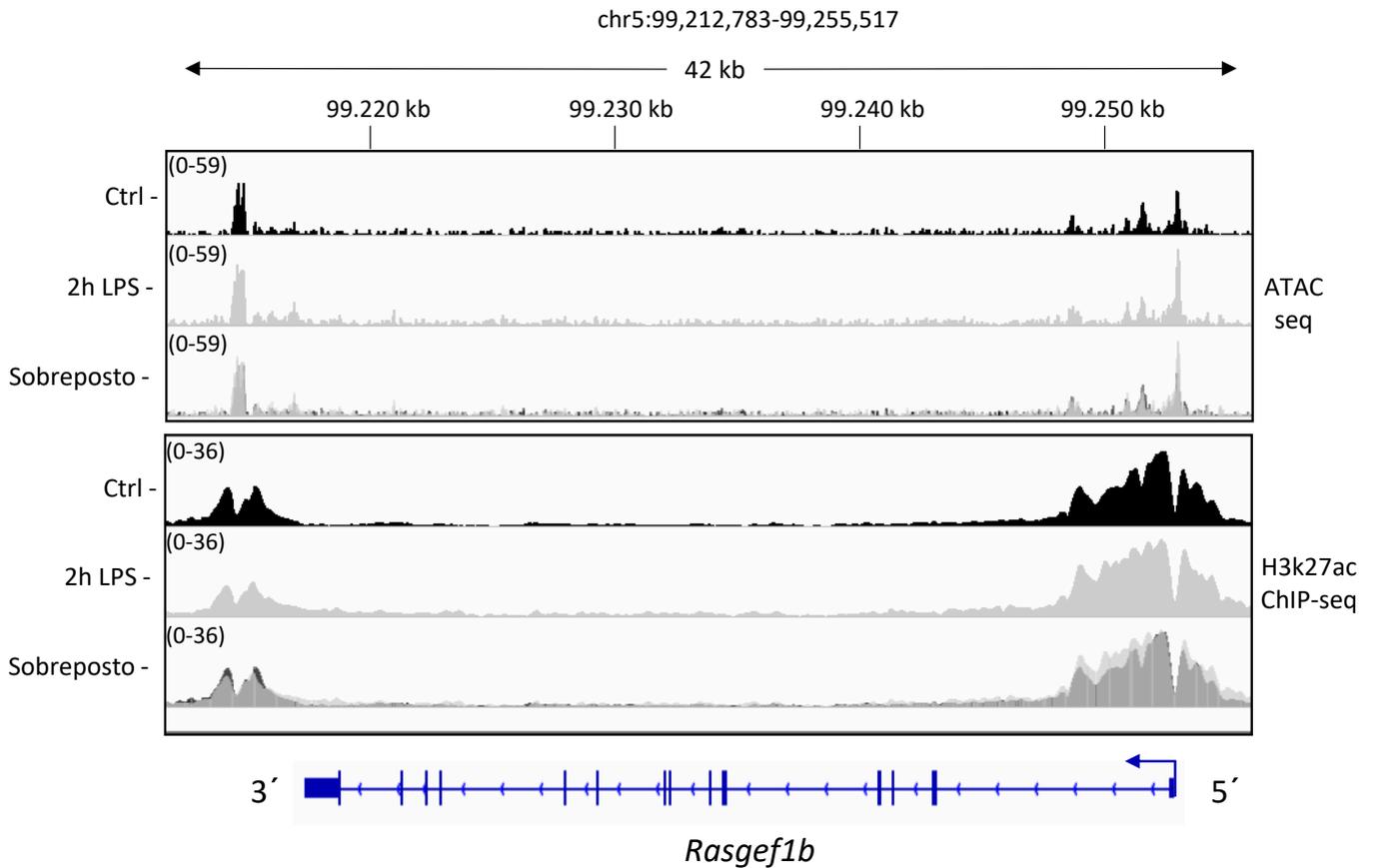


Figura 12: Identificação de regiões de cromatina acessível ao longo da sequência de *Rasgef1b* através de ATAC-seq e H3K27ac ChIP-seq. Macrófagos murinos primários derivados de medula óssea (BMDMs) foram tratados ou não (ctrl) com LPS a 100 ng/mL por 2 horas. **Painel superior:** O DNA genômico foi obtido, tagmentado através da enzima TDE1 (Illumina), purificado, e então amplificado por PCR utilizando-se *primers* customizados com *barcode* (Nextera Customized PCR primers 1 and 2). Processos de purificação adicionais foram realizados para a seleção de tamanho dos fragmentos, e as amostras foram então sequenciadas pelo New York Genome Center com leituras pareadas de 50 pb. **Painel inferior:** Após tratamentos, as células foram fixadas com formaldeído 1% (*cross-link*). O DNA genômico foi obtido e sonicado para obtenção de fragmentos de 150 a 300 pb. Os fragmentos foram imunoprecipitados utilizando-se o anticorpo anti-H3k27ac ligado a *beads* de Proteína A. O *cross-link* foi desfeito, o DNA imunoprecipitado foi purificado e sequenciado com leituras pareadas de 150 pb. As sequências foram alinhadas ao à versão mm10 do genoma referência de camundongo e o programa IGV versão 2.4.14 foi utilizado para identificação dos picos e plotagem dos gráficos. O resultado é representativo de dois experimentos independentes. **Nota:** *Rasgef1b* é transcrito a partir da fita negativa (-), portanto seu sentido 5'-3' é sempre representado da direita para a esquerda nos navegadores de genoma. Mantivemos essa representação convencional ao mostrar figuras geradas nesses programas.

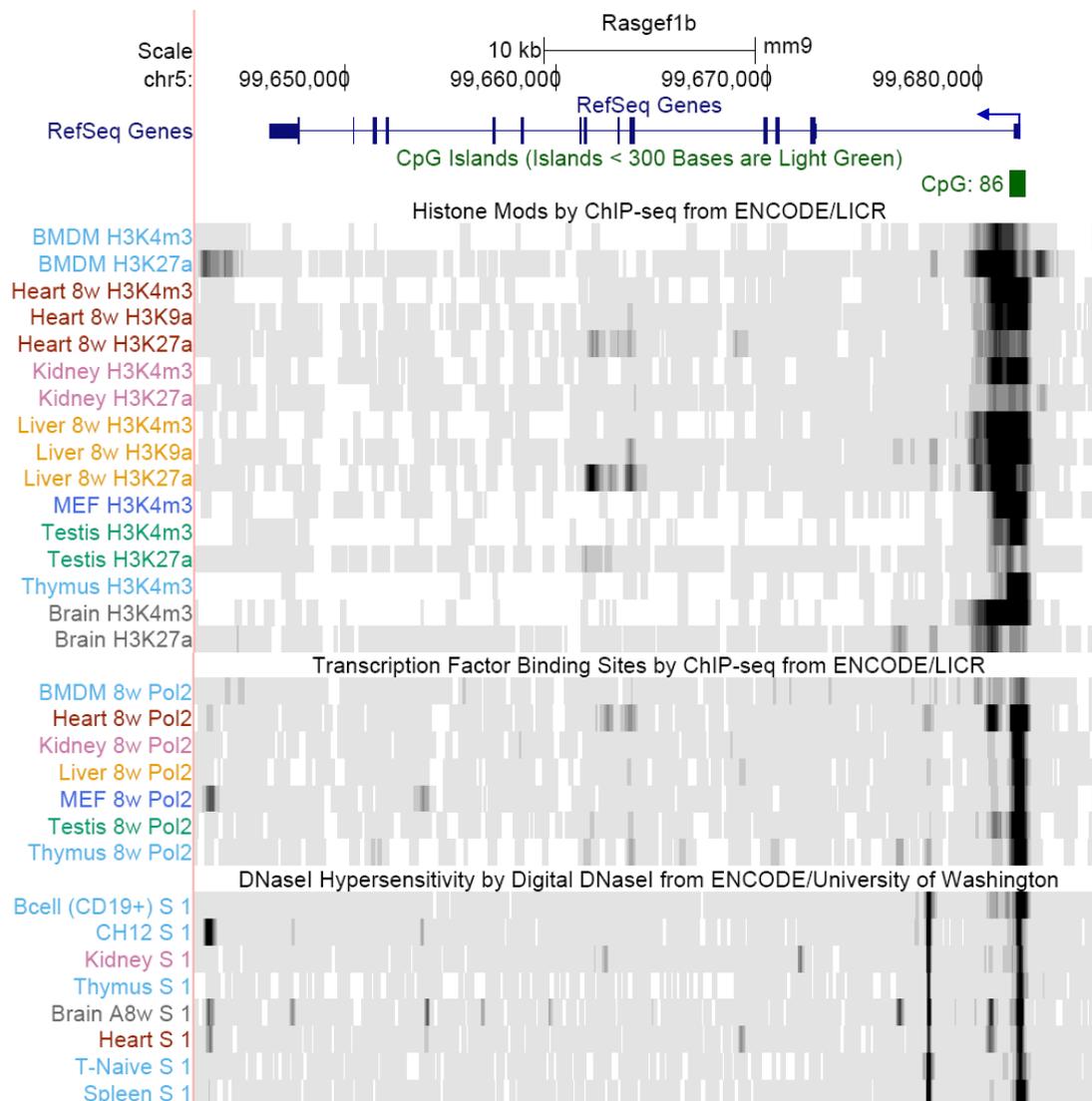


Figura 13: Marcadores epigenéticos de regiões transcricionalmente ativas presentes ao longo da sequência de *Rasgef1b*. Faixas de ChIP-seq anti-H3K4m1, H3K4m3, H3K9ac, H3K27a, e RNA Polimerase II publicamente disponíveis provenientes do banco de dados ENCODE/LICR e faixas de Hipersensibilidade a DNase I provenientes do ENCODE/ University of Washington são mostradas para vários tipos celulares e tecidos de camundongo ao longo do locus do gene *Rasgef1b* utilizando-se o navegador de genomas da UCSC. Os trechos escuros ao longo das faixas mostram a região enriquecida para cada um dos marcadores investigados.

Análises manuais e *in silico* não revelaram elementos regulatórios comuns de ligação a fatores gerais de transcrição como o TATA box, Inr (do inglês, *initiator*), BRE (do inglês, *TFIIB recognition element*), e DPE (do inglês, *downstream core promoter element*) em suas posições canônicas na sequência de *Rasgef1b*. No entanto, identificamos a presença de uma região rica em dinucleotídeos Citosina-Guanina (ilha CpG) que se estende da posição -281 ao +158, uma característica típica de promotores que não contém TATA box e usualmente associada à regulação da expressão gênica (**Fig. 14**) (Butler and Kadonaga, 2002; Deaton and Bird, 2011; Saxonov et al., 2006). Além disso, através de alinhamento por BLASTn identificamos que a região de -401 a +204 de *Rasgef1b* murino apresenta 70% de identidade com a sequência humana, ao passo que nenhuma outra região adicional à montante (até a posição -3000 analisada) apresenta identidade significativa. Consistentemente, utilizando ferramentas adicionais de bioinformática que empregam distintos algoritmos de alinhamento para identificar regiões conservadas ao longo de todo genoma, encontramos um resultado semelhante mostrando elevada conservação do exon 1 de *Rasgef1b* e da região imediatamente à montante entre a sequências de camundongo e de vários ortólogos em mamíferos (**Fig. 15**). Esses dados mostram um grau de conservação relativamente elevado para sequências não-codificadoras, sugerindo uma importância funcional.

Em conjunto, as características descritas aqui são marcadores bem estabelecidos de regiões transcricionalmente ativas. Notavelmente, todos os dados convergem para uma região que inclui parte do exon 1 e algumas centenas de nucleotídeos da sequência 5' flanqueadora proximal, sugerindo uma potencial função regulatória dessa região.

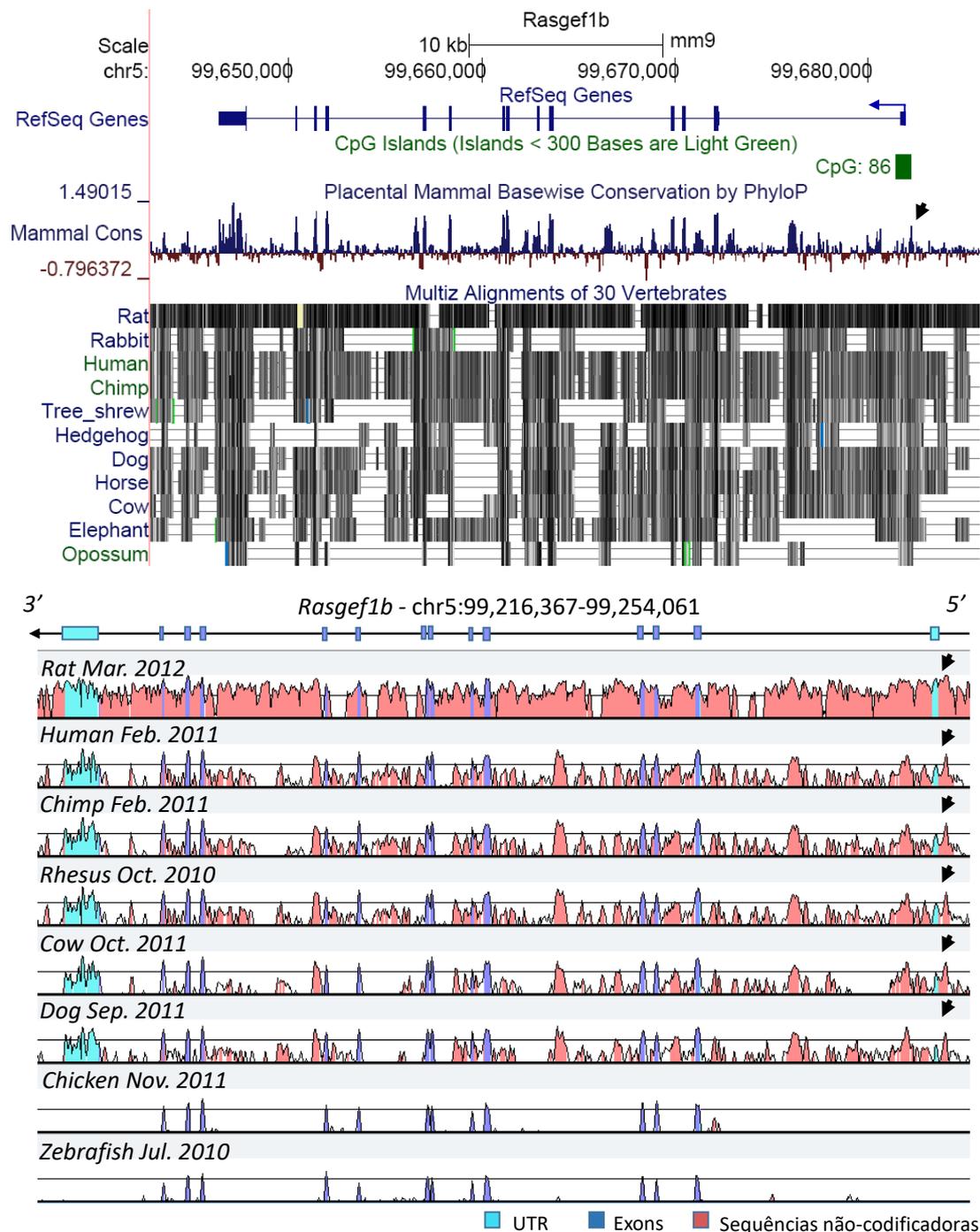


Figura 15: Análise de regiões evolutivamente conservadas entre a sequência de *Rasgef1b* murino e seus ortólogos em mamíferos. (Painel superior): Alinhamento múltiplo entre a sequência do gene *Rasgef1b* murino e seus ortólogos em mamíferos e medidas de conservação evolutiva utilizando as ferramentas multiz e phyloP do navegador do genoma da UCSC. Na trilha de conservação (phyloP), os picos com maior pontuação (Azul) indicam regiões altamente conservadas, enquanto os picos com menor pontuação (Marrom) indicam regiões não conservadas. Nas faixas de alinhamento (multiz), os trechos mais escuros representam regiões com maior identidade em relação à sequência de camundongo. **(Painel inferior):** Alinhamento e análise de conservação de *Rasgef1b* através do programa mVISTA. Regiões de no mínimo 50 nucleotídeos que exibem pelo menos 70% de identidade entre as sequências comparadas são destacadas como regiões evolutivamente conservadas. O eixo horizontal representa a posição relativa das regiões genômicas e o eixo vertical representa a percentagem de identidade entre os ortólogos comparados com a sequência de camundongo. Setas pretas indicam as regiões não-codificadoras conservadas imediatamente à montante do exon 1 de *Rasgef1b*.

5.4 Caracterização da região regulatória de *Rasgef1b* e identificação do promotor funcional

Para entender os mecanismos transcricionais responsáveis pela expressão induzível de *Rasgef1b* por TLRs, focamos em identificar as principais regiões regulatórias em sua sequência e os potenciais componentes moleculares envolvidos neste processo. Para isso, nós amplificamos uma sequência de DNA de 2.866 pb contendo parte da região 5' flanqueadora e parte do exon 1 de *Rasgef1b* (de -2747 a +119), seguido da clonagem no vetor de expressão de luciferase, pGL3-Basic, dando origem assim à construção pGL3-2747/+119 (**Fig. 16**). Deleções progressivas à 5' e 3' foram realizadas e confirmadas através de digestão com enzimas de restrição e sequenciamento (**Figs. 16 e 17**). A atividade das sequências foi então analisada através de transfecção transiente e ensaio de gene repórter. Em células HEK293 observamos que a construção pGL3-2747/+119 apresentou atividade basal elevada quando comparada ao vetor vazio pGL3-basic (**Fig. 18A**). Nenhuma alteração de atividade foi observada quando removidos 1.126 bp da extremidade 5' (construção pGL3-1621/+119). Porém, deleção adicional até a posição -706 (pGL3-706/+119) resultou em um aumento de atividade da construção, sugerindo que a região deletada de -1621 a -706 pode potencialmente abrigar elementos de regulação negativa. Deleção adicional da extremidade 5' até a posição -183 (pGL3-183/+119) resultou em uma curta construção contendo 302 pb que reteve a atividade máxima. De modo importante, quando essa região foi removida (deleção de +119 a -368, construção pGL3 Δ 487) a atividade foi completamente abolida, fornecendo evidências que esse segmento é crítico para recrutar a maquinaria de fatores gerais de transcrição. Avaliamos ainda a atividade das construções em resposta a ativação das vias de sinalização de TLRs através da superexpressão da molécula adaptadora Mal/TIRAP. O resultado (**Fig. 18A**) revela que todas as construções funcionais foram responsivas à ativação da sinalização de TLRs, exibindo um aumento de atividade de até 2 vezes.

Uma vez que a construção pGL3-183/+119 é suficiente e necessária para a expressão constitutiva e induzida da luciferase por TLRs, esse resultado indica que o promotor funcional de *Rasgef1b* está contido nessa região. Notavelmente, essa sequência está contida na região de cromatina acessível, ilha CpG, e conservação previamente descrita, suportando essa conclusão. Curiosamente, entre todos os potenciais elementos regulatórios identificados na sequência de *Rasgef1b*, essa sequência abriga principalmente um cluster de elementos Sp1 e NF- κ B (**Fig. 16**). Em linha com estudos

anteriores que mostram que a expressão de *Rasgef1b* é dependente dos adaptadores TRIF e MyD88 (Andrade et al., 2010), a ativação do promotor foi induzida através da superexpressão dessas moléculas, com MyD88 resultando em um aumento de até 3.8 vezes na atividade (**Figs. 18B e C**). Mais importante, LPS induziu um aumento de 3.5 vezes na atividade do promotor em macrófagos Raw264.7, indicando que essa sequência é suficiente para a ativação da expressão de *Rasgef1b* induzida por LPS (**Fig. 18D**).

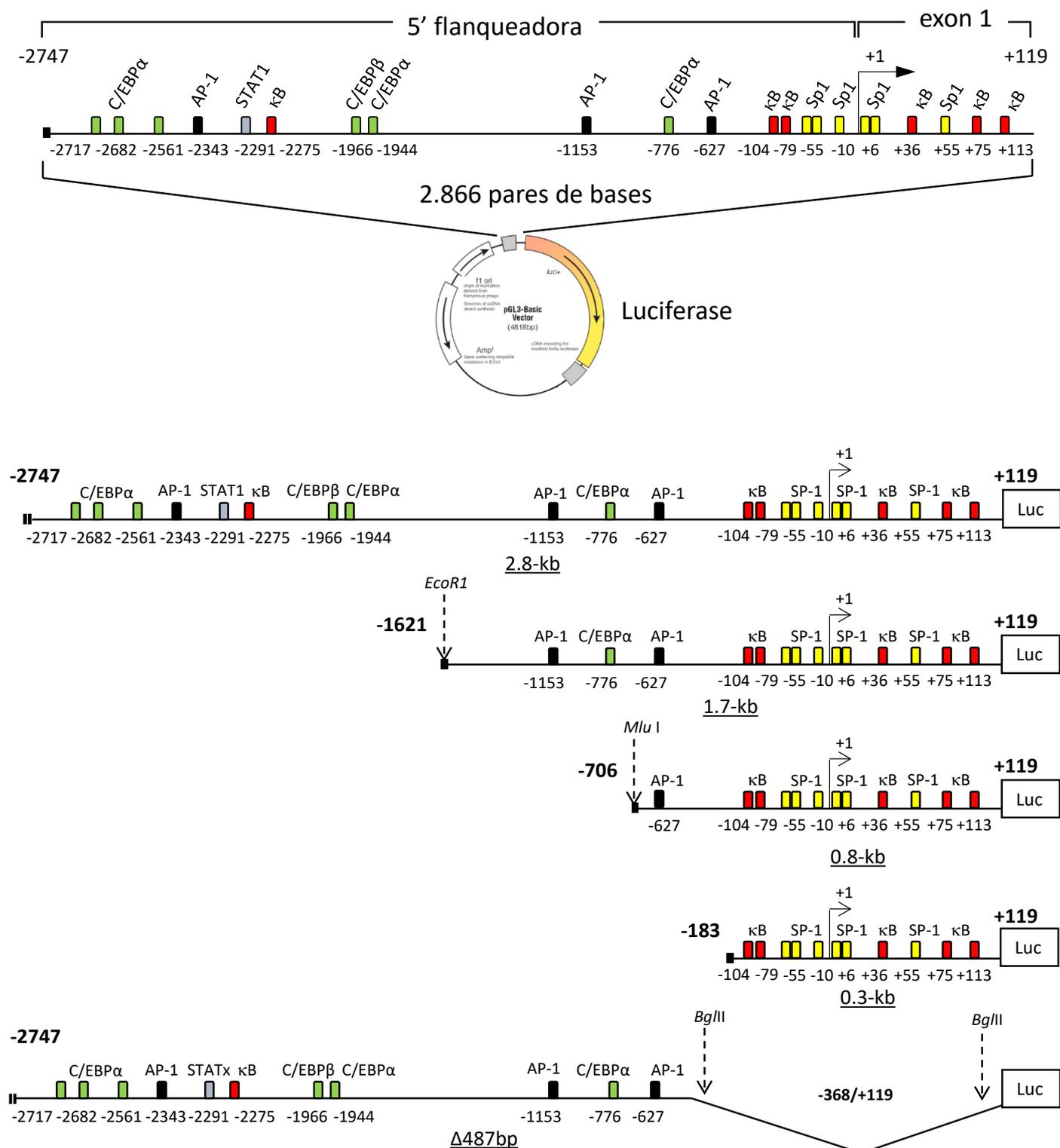


Figura 16: Representação esquemática do repórter pGL3-2747/+119 e dos mutantes gerados por deleção. O segmento de 2.866 pb (-2747/+119) da região não-codificadora de *Rasgef1b* foi clonado entre os sítios de KpnI e XhoI do plasmídeo pGL3-Basic. No esquema está indicada a posição dos principais sítios de ligação para fatores de transcrição associados à resposta imune identificados in silico em relação ao TSS (seta; +1). Mutantes por deleções progressivas à 5' e 3' foram gerados a partir dessa construção através da digestão com as enzimas de restrição indicadas ou através de síntese de DNA (pGL3-183/+119).

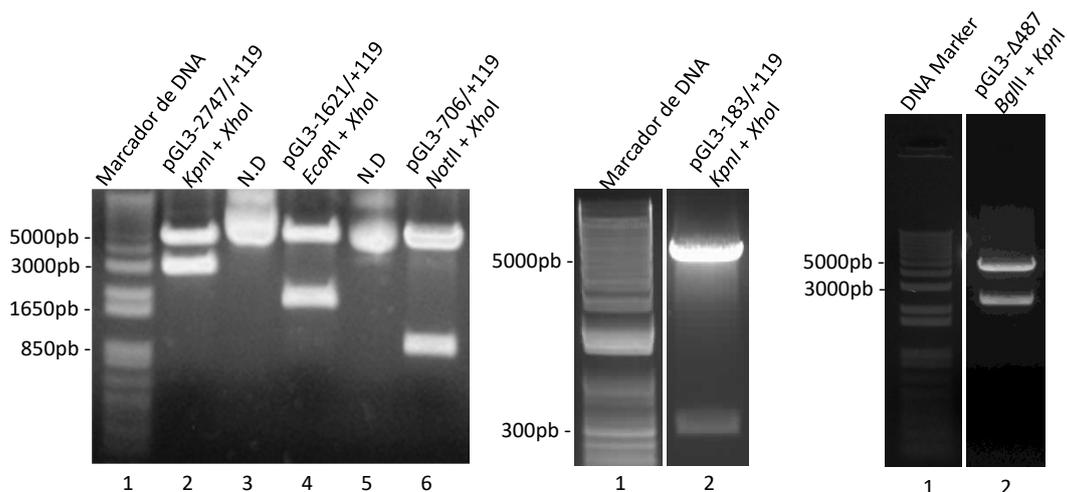


Figura 17: Confirmação da geração dos plasmídeos repórter do promotor de *Rasgef1b* através de digestão e eletroforese. Alíquotas provenientes de miniprep de cada plasmídeo foram digeridas para liberação do inserto com as enzimas utilizadas para clonagem. No caso onde esse sítio foi perdido durante o processo (pGL3-706/+119), foi utilizada uma enzima cujo sítio de restrição é adjacente ao sítio de clonagem (*NotI* ao invés de *MluI*). As amostras foram fracionadas em gel de agarose 1% contendo brometo de étido para diagnóstico. Marcador de DNA: 1kb DNA ladder plus (Invitrogen). N.D = Não digerido

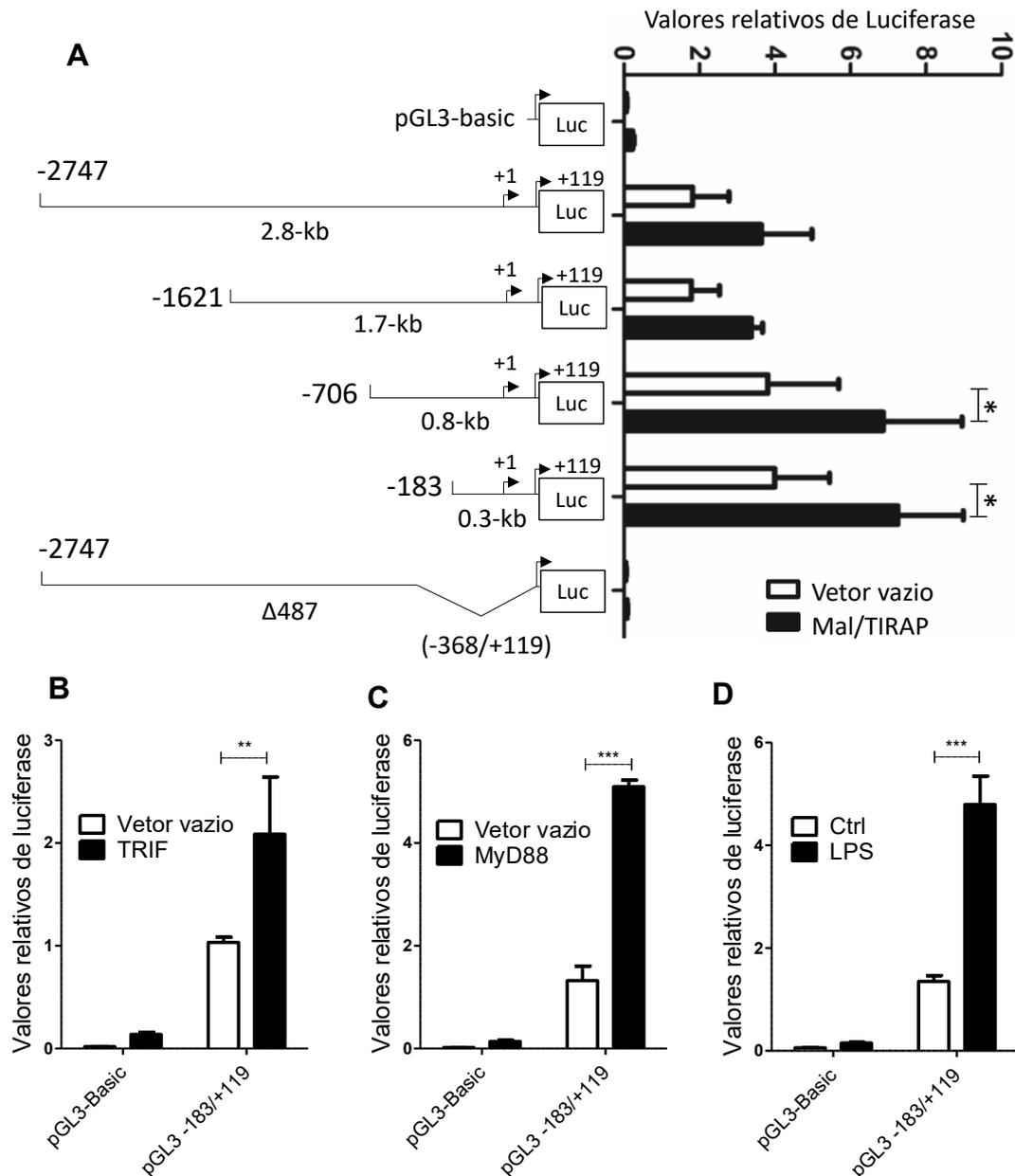


Figura 18: Caracterização funcional da região promotora de *Rasgef1b*. (A) Painel da esquerda: Representação esquemática das construções repórter do promotor de *Rasgef1b* ligado ao gene da luciferase no vetor pGL3-Basic. Os nucleótidos da extremidade de cada construção estão numerados em relação ao TSS (+1). (A-C) Células HEK293 foram co-transfectadas com as construções representadas (ou com vetor vazio pGL3-Basic) juntamente com pRL-TK durante 24 horas. Onde indicado, as células foram co-transfectadas com plasmídeos para expressão de Mal/TIRAP, TRIF ou MyD88, ou com o vetor vazio pcDNA 3.1 para equiparação da quantidade de DNA transfectado. (D) Células Raw264.7 foram co-transfectadas com o repórter do promotor de *Rasgef1b* (pGL3-183/+ 119) e pRL-TK durante 24 horas, e então tratadas ou não (ctrl) com 100 ng/mL de LPS por 24 horas. Os lisados foram coletados e a atividade da luciferase medida em luminômetro. Os resultados são apresentados como valores relativos da atividade de luciferase do Vagalume normalizado pela atividade da luciferase de Renilla (pRL-TK). Os dados representam a média \pm D.P. de três experimentos independentes (A) ou dois experimentos independentes (B-D) com duas réplicas biológicas cada ($n = 4$). * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$ determinado por Two-Way ANOVA seguido por pós-teste de Bonferroni (A) ou teste de T de Student não-pareado (B-D).

5.5 Elementos κ B são necessários para a ativação máxima do promotor de *Rasgeflb*

Como a sequência do promotor de *Rasgeflb* abriga cinco potenciais sítios de ligação para NF- κ B, nós questionamos se esse fator de transcrição central na resposta de TLRs desempenha um papel na regulação transcricional de *Rasgeflb*. Inicialmente, avaliamos o efeito das subunidades RelA (p65), c-Rel, ou RelB de NF- κ B em células HEK293 sobre a atividade do promotor de *Rasgeflb*. Mostramos que tanto a expressão forçada de RelA quanto de c-Rel foi suficiente para induzir a ativação do promotor, causando um aumento de 4.6 vezes e 2.5 vezes, respectivamente, na atividade da luciferase (**Fig. 19A e B**). Como esperado, a superexpressão de RelB por si só não foi capaz de induzir a ativação do promotor de *Rasgeflb* ou de E-Selectina (controle positivo) (**Fig. 19C e anexo**) (Huang et al., 2005; Ruben et al., 1992). No entanto, ao ser combinado com as subunidades p50 ou p52, RelB é capaz de induzir a ativação do promotor de E-Selectina, mas não do promotor de *Rasgeflb* de modo significativo (**Fig. 19C e anexo**). Em seguida, avaliamos através de mutação sítio-dirigida, o papel dos elementos κ B a montante e a jusante da posição +1 na ativação do promotor induzida por LPS em macrófagos Raw264.7. Como observa-se na **figura 19D**, a mutação dos elementos a montante (nomeados κ B # 5 e # 4) ou a jusante (κ B # 3, # 2 e # 1) resultou em 46,5% ou 51,3% de redução na atividade do promotor induzida por LPS, respectivamente.

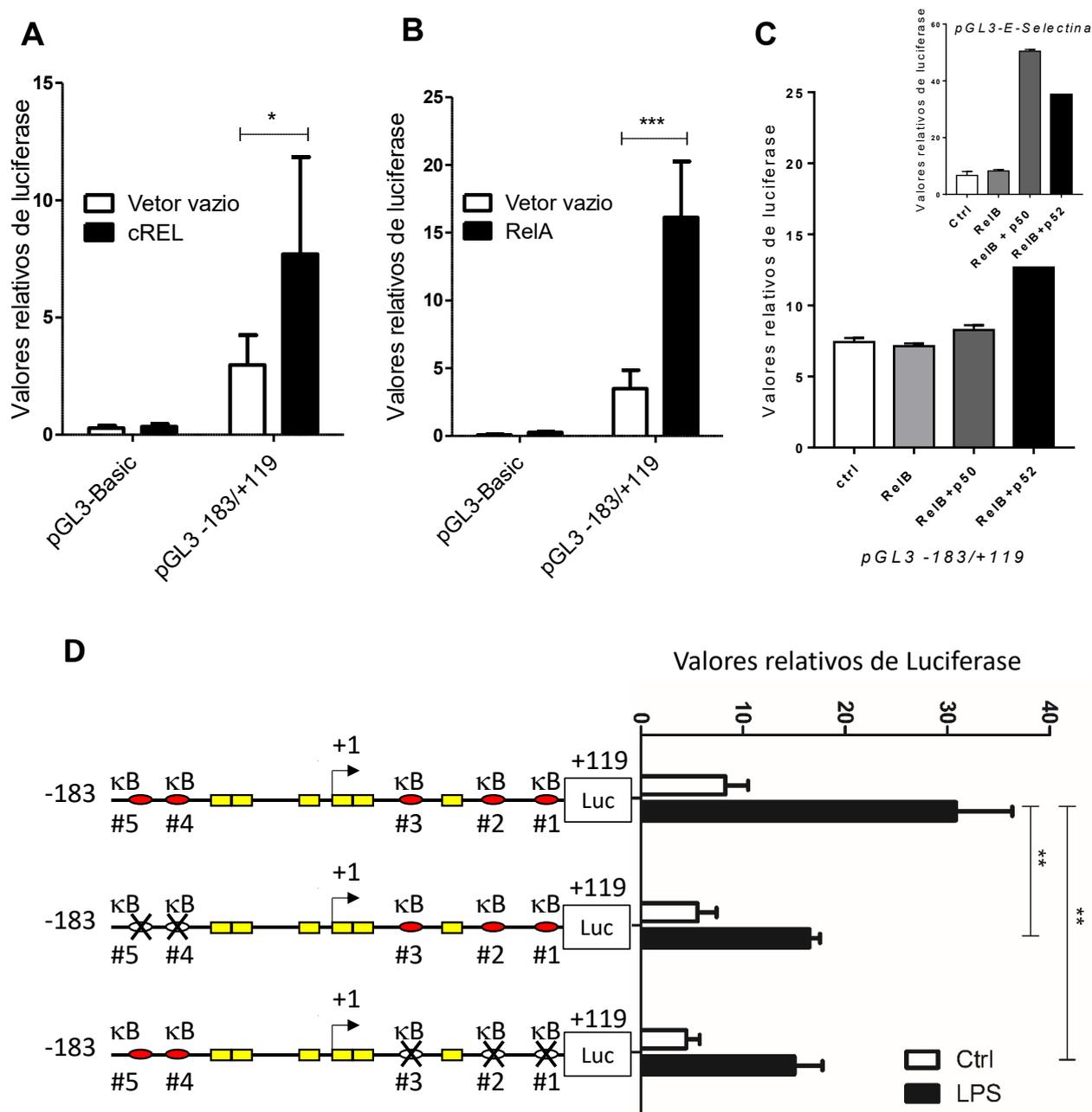


Figura 19: Análise do efeito de NF- κ B sobre a ativação do promotor de *Rasgef1b*. (A, B e C) Células HEK293 foram co-transfectadas com o repórter do promotor de *Rasgef1b* (-183/+119) ou vetor vazio pGL3-Basic juntamente com pRL-TK durante 24 horas. Onde indicado, as células foram co-transfectadas com plasmídeos para expressão de RelA (p65), cRel, RelB, p50, p52, ou com o vetor vazio pcDNA 3.1 para equiparação da quantidade de DNA transfectado. (D) Células Raw264.7 foram co-transfectadas com a construção original do promotor de *Rasgef1b* (-183/+119) ou com construções contendo os elementos κ B mutados (painel esquerdo) juntamente com pRL-TK por 24 horas, e posteriormente tratadas ou não com 100 ng/mL de LPS. Os lisados foram coletados após 24 horas e a atividade da luciferase medida em luminômetro. Os resultados são apresentados como valores relativos da atividade de luciferase do Vagalume normalizado pela atividade da luciferase de Renilla (pRL-TK). Os dados representam a média \pm D.P. de cinco (A e B), dois (C) ou quatro (D) experimentos independentes. * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; determinado pelo teste T de Student não-pareado (A e B) ou Two-way ANOVA seguido por pós-teste de Bonferroni (D).

Para determinar se os elementos κ B apresentam preferencia de ligação a NF- κ B, analisamos um conjunto de dados públicos gerados através de experimentos de PBM (do inglês, *protein-binding microarray*), onde foi examinada a afinidade de ligação de vários dímeros de NF- κ B a milhares de sondas contendo potenciais sítios κ B (Siggers et al., 2011). Como mostrado na **figura 20A**, os sítios κ B #1, #3, #4, #5 do promotor de *Rasgef1b* foram incluídos nesses experimentos, porém não foi investigada sonda contendo o elemento #2. De acordo com essas análises, há uma preferência de ligação dos dímeros analisados aos elementos κ B #1 e #3, embora ainda ocorra ligação aos elementos κ B #4 e #5 em menor intensidade. Homodímeros de cRel e RelA, assim como os heterodímeros cRel/p50 e RelB/p50 apresentam elevada afinidade pelo κ B #1, enquanto os demais dímeros apresentam uma afinidade moderada. Em relação ao κ B #3, a afinidade do homodímero cRel é elevada, RelA/RelA e Relb/p50 moderada, e dos demais relativamente baixa. Quanto aos κ B #4 e κ B #5, dímeros cRel/cRel e Relb/p50 apresentam moderada afinidade, enquanto os demais apresentam baixa afinidade (**Fig. 20A**). Para fornecer evidências experimentais dessas observações, nós então realizamos estudos de EMSA a partir do extrato nuclear de células Raw264.7 estimuladas ou não com LPS. Nas amostras provenientes de células tratadas com LPS por 1 hora, similarmente ao controle positivo (H2k), detectamos a ligação de complexos às sondas dos elementos κ B #1, #3, e #5. No entanto, a confirmação da especificidade de NF- κ B ocorreu apenas para os elementos κ B #1 e #3, uma vez que a ligação foi abolida quando foi adicionada a sonda não-marcada do consenso de NF- κ B em excesso à reação (competidor específico), o que não ocorreu para o elemento κ B #5 (**Fig. 20B**). Através de ensaio de *supershift*, confirmamos a capacidade de ligação de homo ou heterodímeros contendo RelA, cRel ou p50 aos elementos κ B #1 e κ B #3 (**Fig. 20C**). Ademais, encontramos nenhuma ligação ou ligação de baixa intensidade aos elementos κ B #2 e #4 (**Figs. 20B e D**).

Em conjunto, estes resultados sugerem que, atuando através dos sítios κ B específicos na sequência proximal de *Rasgef1b*, NF- κ B é suficiente para induzir a ativação do promotor e necessário para sua atividade máxima induzida por LPS.

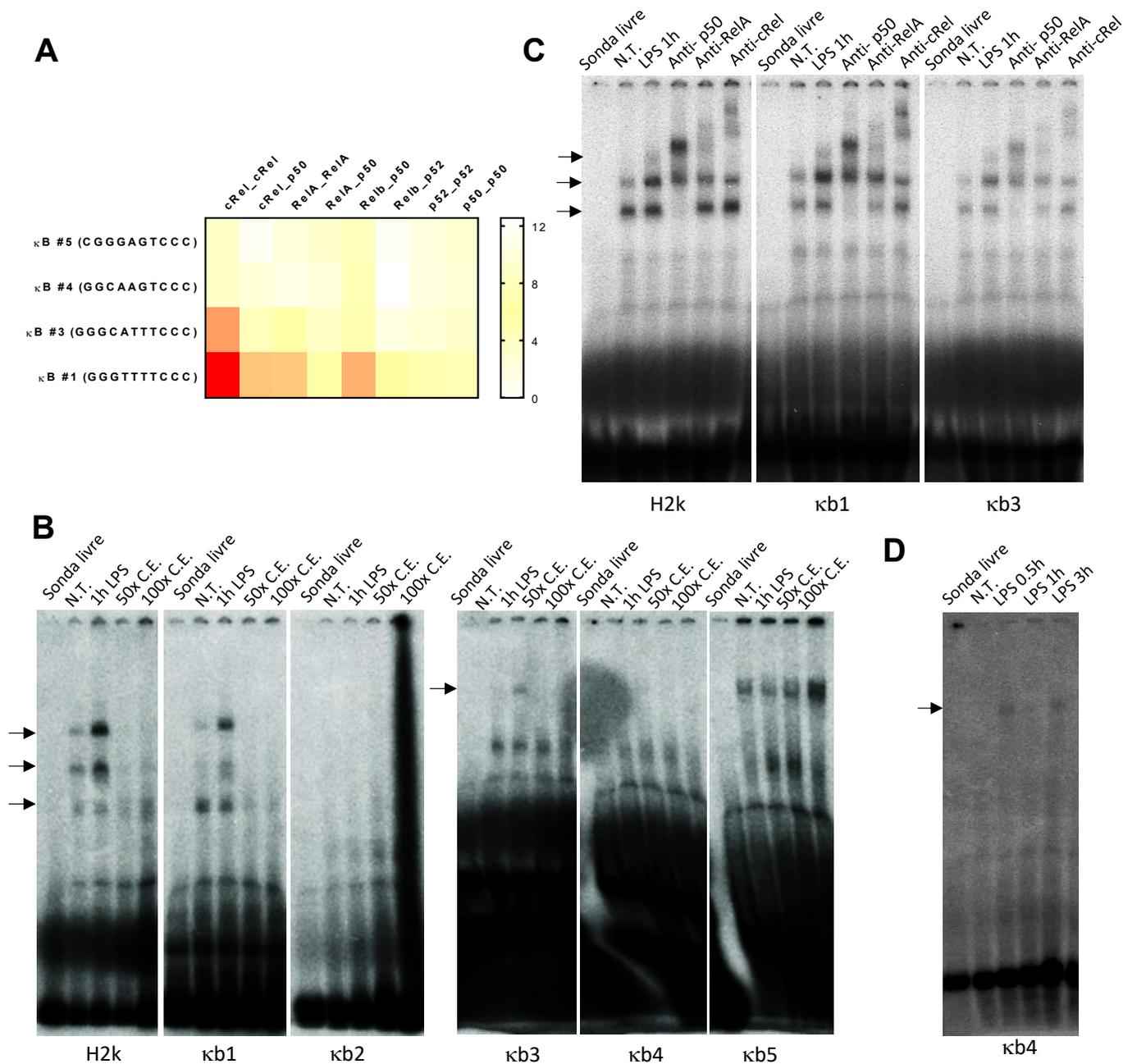


Figura 20: Análise da ligação de NF- κ B aos elementos κ B do promotor de *Rasgef1b*. (A) Foi avaliada a afinidade de ligação de vários dímeros NF- κ B purificados a 3.285 sondas contendo variações da sequência κ B consenso através de *Protein Binding Microarray*. Recuperamos os resultados disponíveis online para avaliar a afinidade de ligação das sondas contendo os elementos κ B do promotor de *Rasgef1b*. Os valores dados são calculados com base na fluorescência média de cada sonda, normalizada pelos valores obtidos em sondas inespecíficas. (B-D) Raw264.7 foram tratadas ou não (ctrl) com LPS a 100 ng/mL pelos tempos indicados. O extrato nuclear foi isolado e incubado com sondas dupla-fita radioativamente marcadas contendo a sequência dos elementos κ B do promotor de *Rasgef1b*, ou a sequência κ B do gene *H2k* (controle positivo). Onde indicado, foi adicionada uma sonda não marcada contendo a sequência κ B consenso em excesso (50x ou 100x em relação à sonda marcada). Para *supershift* (C), foram adicionados 1 μ g dos anticorpos anti-p50, anti-Rela, ou anti-cRel (Santa Cruz). As amostras foram fracionadas em gel não-desnaturante de poliacrilamida a 5%. Em B e D, foram utilizados extratos nucleares de um mesmo experimento, e em C de um segundo experimento independente.

5.6 Em macrófagos ativados, NF- κ B/RelA é recrutado para o promotor *Rasgef1b* e contribui para a expressão induzida por TLRs

Para fornecer evidências adicionais e diretas do recrutamento de NF- κ B ao promotor de *Rasgef1b*, nós conduzimos análises de um ensaio de ChIP-seq anti-RelA em que foi utilizado o DNA genômico de BMDMs murinos primários estimulados ou não com LPS. Consistente com os nossos resultados funcionais, o recrutamento de RelA para a região promotora foi observado após 2 horas de estímulo com LPS, levando a um aumento de 5 vezes no sinal de ChIP quando comparado aos macrófagos não estimulados (**Fig. 21**). Interessante notar que essa região se sobrepõe à região de cromatina acessível que mostramos previamente. Além disso, três picos adicionais de RelA-ChIP foram identificados ao longo do locus de *Rasgef1b*. Especificamente, LPS induziu um aumento de 22 vezes na ligação de RelA a uma região flanqueadora distal a montante (de -2345 a -1724), 40 vezes a uma região codificadora no exon 14, e 20 vezes a uma região flanqueadora distal, a jusante do último exon de *Rasgef1b*. Esse último pico também sobrepõe-se aos picos adicionais de ATAC-seq e H3K27ac distais a jusante descritos anteriormente, sugerindo um potencial *enhancer* responsivo ao LPS (**Fig. 21**).

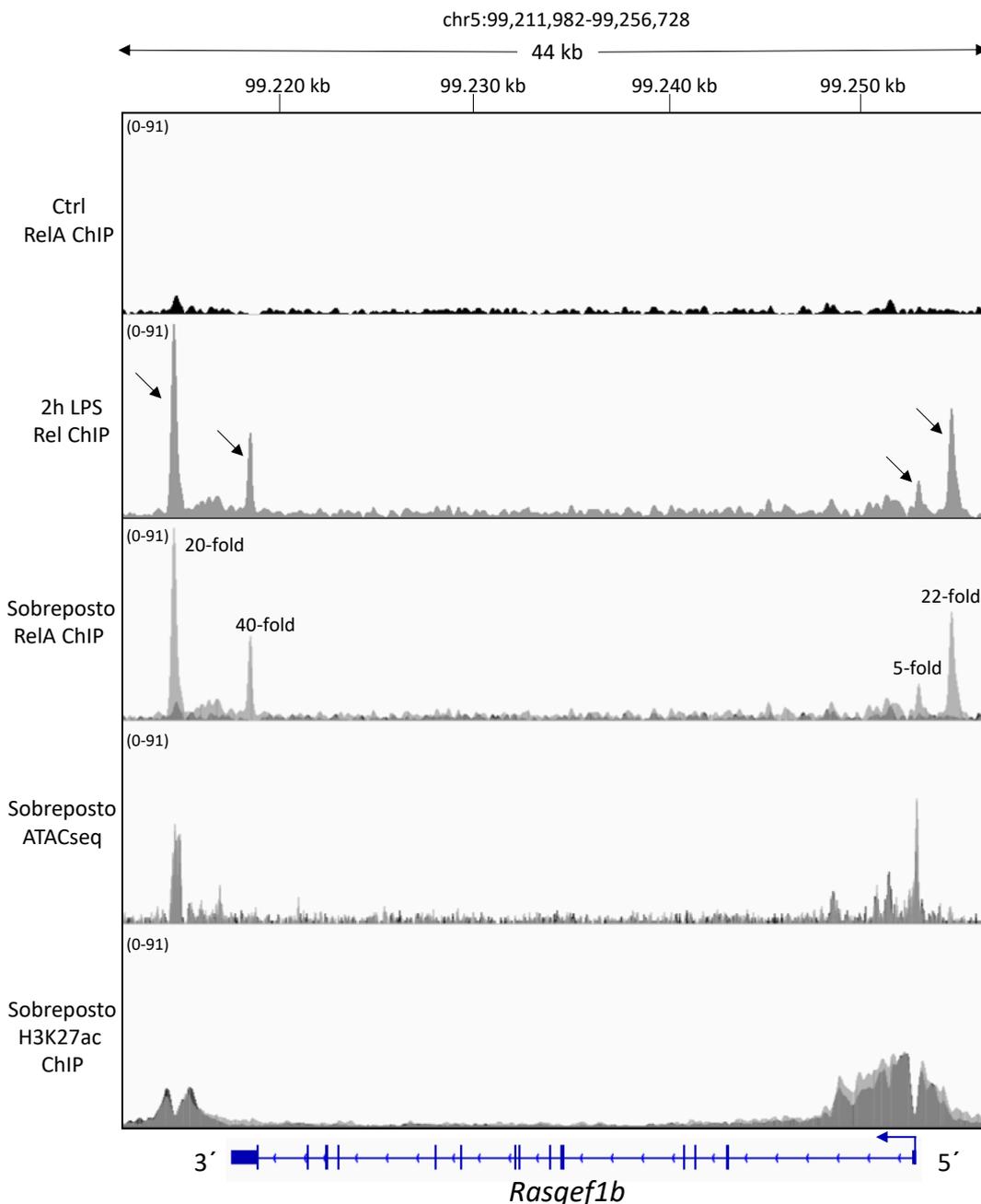


Figura 21: Identificação do recrutamento de NF- κ B/RelA ao longo da sequência de *Rasgef1b* através de ChIP-seq. Macrófagos murinos primários derivados de medula óssea (BMDMs) foram tratados ou não (ctrl) com LPS a 100 ng/mL por 2 horas. Após tratamentos, as células foram fixadas com formaldeído 1% (*cross-link*). O DNA genômico foi obtido e sonicado para obtenção de fragmentos de 150 a 300 pb. Os fragmentos foram imunoprecipitados utilizando-se o anticorpo anti-RelA ligado a *beads* de Proteína A. O *cross-link* foi desfeito, o DNA imunoprecipitado foi purificado e sequenciado com leituras pareadas de 150 pb. As sequências foram alinhadas ao à versão mm10 do genoma referência de camundongo e o programa IGV versão 2.4.14 foi utilizado para identificação dos picos e plotagem dos gráficos. As setas mostram os quatro principais picos indicando sítios de ligação à RelA após tratamento com LPS. O enriquecimento de cada pico após tratamento está indicado. Também estão representados os picos de ATAC-se e H3K27ac previamente demonstrados. O resultado é representativo de dois experimentos independentes.

Finalmente, nós avaliamos se NF- κ B desempenha um papel na expressão induzível de *Rasgef1b* ao nível transcricional. Para isto, avaliamos os níveis do mRNA de RasGEF1b em BMDMs imortalizados do tipo selvagem (WT) ou deficientes de RelA (RelA-KO) após estimulação com agonistas de TLR. Inicialmente, a deficiência de RelA foi confirmada por RT-qPCR, e como esperado, a expressão induzida por LPS de genes sabidamente dependentes de NF- κ B, *Cxcl1* e *Il6*, é praticamente abolida nessas células (**Fig. 22A-C**). Em relação a *Rasgef1b*, embora não abolida, sua expressão induzida por TLRs é significativamente prejudicada nessas células em comparação com os iBMDMs WT, exibindo uma redução de 50% nos níveis de mRNA após 1 hora de estimulação com LPS ou poly (I:C) e 46% após estimulação com CpG ODN (**Fig. 22D-F**). Experimentos adicionais utilizando o inibidor farmacológico de NF- κ B, Bay11-7082, mostraram que enquanto os níveis do mRNA de *Tnf* são reduzidos de forma dose-dependente – de 33% com 5 μ M até 67% com 20 μ M – os níveis do mRNA de RasGEF1b foram reduzidos em média 35%, independente da concentração do inibidor após 3 horas de estímulo com LPS em células Raw264.7 (**Fig. 22G-H**). Em conjunto, esses resultados mostram que NF- κ B/RelA é recrutado ao promotor de *Rasgef1b* e a regiões distais logo após estimulação com LPS e contribui para a sua expressão induzível por TLRs.

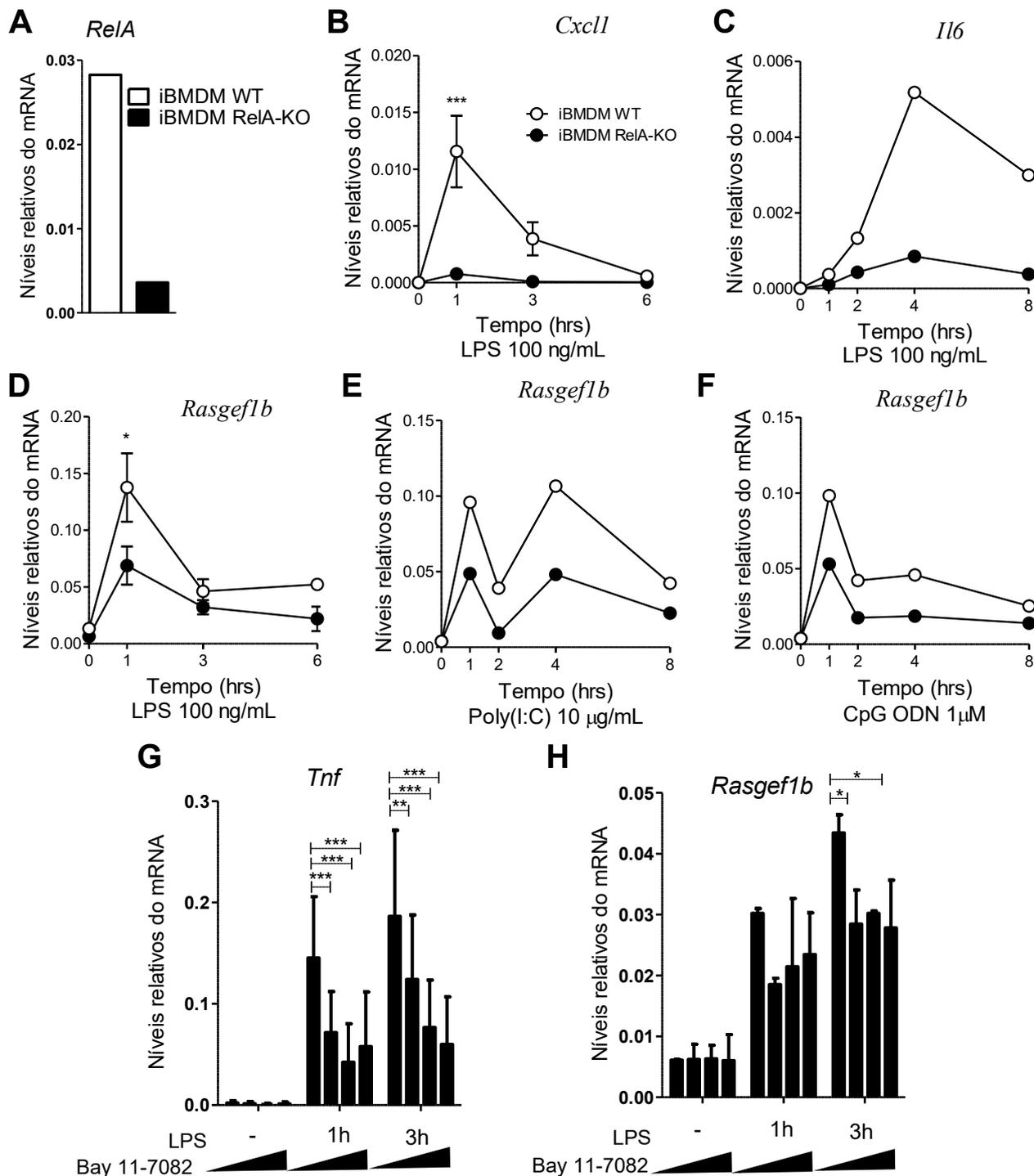


Figura 22: NF- κ B contribui para a expressão de *Rasgef1b* induzida por TLRs em macrófagos. (A-F) BMDMs imortalizados selvagens (WT) ou deficientes de RelA (RelA-KO) foram deixados sem tratamento (0 hr) ou tratados com 100 ng/mL de LPS, 10µg/mL de Poly (I: C) ou 1µM de CpG ODN pelos tempos indicados. O RNA total foi extraído e os níveis relativos do mRNA de *Rasgef1b*, *Il6*, *Cxcl1* e *Rela* (*p65*) foram avaliados por RT-qPCR. **(G e H)** Macrófagos Raw264.7 foram deixados sem tratamento ou pré-tratados por 1 hora com concentração crescente (5, 10 ou 20 µM) do inibidor de NF- κ B Bay11-7082 e então tratados ou não com 100 ng/mL de LPS pelos tempos indicados. Os dados representam a média \pm D.P. de duas **(G e H)** ou quatro **(B e D)** réplicas biológicas em dois experimentos independentes (n=4). **A, C, E e F** foram experimentos únicos. * $p \leq 0,05$; *** $p \leq 0,001$; determinado por Two-way ANOVA seguido de pós-teste de Bonferroni.

6. DISCUSSÃO

Por ser um GEF, RasGEF1b deve estar envolvido em importantes e diversos processos celulares mediados por seus potenciais efetores Ras-GTPases. Sob a ótica imunológica, tem sido reportado que os níveis do mRNA de RasGEF1b aumentam rápida e significativamente em macrófagos humanos e murinos após ativação de TLRs, e em órgãos linfóides e não-linfóides de camundongos infectados com protozoários parasitas, enquanto que apenas o cérebro, intestinos e testículos expressam níveis basais significativos de seu mRNA (Andrade et al., 2010; Ferreira et al., 2002). Esses dados mostram, portanto, que o perfil de expressão de *Rasgef1b* é tecido-específico e regulado ao nível da transcrição após ativação de TLRs em células imunes e em alguns órgãos. No entanto, até o momento, a região promotora e os mecanismos regulatórios responsáveis pela sua expressão induzível por TLRs permaneciam desconhecidos.

No presente trabalho, demonstramos que segmentos de uma sequência não-codificadora flanqueadora de *Rasgef1b* podem promover eficientemente a expressão da luciferase em células HEK293 e macrófagos Raw264.7, exibindo uma robusta atividade basal e induzida por TLRs, em consistência com o perfil de expressão de *Rasgef1b*. Os principais efetores responsáveis pela regulação do complexo programa de expressão gênica resultante da ativação de TLRs são as proteínas cinase ativadas por mitógenos (MAPK), como p38, JNK, e ERK1/2, que levam à ativação de AP-1, e os fatores de transcrição IRFs e NF- κ B (Kawai and Akira, 2010; Kumar et al., 2011). Mostramos que vários elementos regulatórios relacionados à resposta imune estão presentes na sequência flanqueadora de *Rasgef1b*, incluindo sítios para ligação de fatores diretamente ativados por TLRs como AP-1 e NF- κ B, mas não para IRFs. Através de deleções progressivas, demonstramos que a maioria dos elementos identificados, incluindo AP-1, não são necessários para a ativação do promotor mediada por TLRs, sugerindo que a via MAPK-AP1 não deve estar envolvida na regulação transcricional de *Rasgef1b*. Além disso, é importante notar que a deleção da região -1621 a -706 levou a um aumento na atividade relativa da sequência. Portanto, é possível que essa região abrigue sítios putativos para a ligação de reguladores negativos da transcrição constitutivamente expressos. Um dos elementos presente nessa região, mas ausente na sequência proximal, é um sítio para ligação de C/EBP. A família C/EBP contém 6 membros (C/EBP α , β , γ , δ , ϵ , ζ), que podem ser expressos constitutivamente em diferentes tipos celulares, e que em conjunto desempenham funções variadas no organismo, atuando desde a regulação de genes

associados a processos básicos como o controle do ciclo celular, até na imunidade (Lekstrom-Himes, 1998; Mikita et al., 1998; Udofa et al., 2013). Esses fatores podem atuar tanto como ativadores ou repressores da transcrição, dependendo do promotor em que ligam. Por exemplo, foi mostrado que C/EBP β pode atuar como um regulador negativo da expressão do miRNA-145 em células humanas (Sachdeva et al., 2012). Além disso, C/EBP β é capaz de se ligar à região regulatória e reprimir a replicação do vírus HPV11 em queratinócitos primários de prepúcio humano (Ralph et al., 2006; Wang et al., 1996). Em vista desses dados, é razoável supor que fatores de transcrição constitutivamente expressos, como C/EBP, podem potencialmente atuar na regulação negativa de *Rasgef1b* mediante elementos diversos presentes ao longo de sua sequência.

Importantemente, mostramos que uma curta sequência proximal contendo 302 pb (-183/+119) é suficiente e necessária para promover a expressão constitutiva e induzível da luciferase, revelando, portanto, a presença do promotor ativo de *Rasgef1b* nessa região. Suportando essa análise funcional, demonstramos que tal sequência é evolutivamente conservada, está incluída numa região de cromatina constitutivamente acessível, circundada por nucleossomos H3k27ac, e capaz de recrutar constitutivamente a RNA Pol II. Dentre todos os elementos regulatórios identificados, destaca-se a prevalência de sítios κ B e Sp1 nessa sequência. Curiosamente, Mukherjee e colaboradores apresentaram dados de RNA-seq que sugerem que a expressão de *Rasgef1b* pode ser regulada pela subunidade RelA de NF- κ B em interação com a proteína co-ativadora CBP/p300 após a estimulação de fibroblastos murinos com TNF (Mukherjee et al., 2013). Mais recentemente, Ng e colaboradores mostraram que os níveis do mRNA de RasGEF1b em células Raw264.7 estimuladas com LPS são reduzidos em resposta ao pré-tratamento com o inibidor IKK VII, sugerindo uma dependência parcial de NF- κ B na expressão de *Rasgef1b* (Ng et al., 2016). Em consistência com esses relatos, fornecemos no presente trabalho evidências diretas de que o NF- κ B é importante para a expressão de *Rasgef1b* induzida por TLRs, atuando através de sítios κ B na região promotora e possivelmente através de regiões distais adicionais.

Nossos dados mostram que as subunidades centrais de NF- κ B, RelA ou cRel, são suficientes para induzir a ativação do promotor de *Rasgef1b*, ao passo que dímeros contendo de RelB são incapazes de promover tal ativação. Essa diferença é de certa forma esperada, visto que dímeros contendo RelB são ativados predominantemente pela via não-canônica de NF- κ B em contraste com a ativação da via canônica mediada por TLRs e a

consequente ativação de dímeros contendo RelA ou cRel (Hayden and Ghosh, 2012). Adicionalmente, mostramos que a mutação dos sítios κ B a montante ou a jusante do TSS resultou em uma redução significativa da atividade do promotor induzida por LPS, revelando a sua importância funcional. Vale notar que, como esses elementos regulatórios podem ter funções sinérgicas e/ou redundantes (Plevy et al., 1997), é possível que a mutação simultânea de todos os sítios κ B resultaria em uma diminuição ainda maior da atividade do promotor. Estudos futuros poderão esclarecer essa questão. Consistentemente, mostramos o recrutamento de NF- κ B/RelA em resposta a LPS, não apenas para o promotor de *Rasgef1b*, mas também a regiões adicionais. A região distal de ligação a RelA a montante (de -2345 a -1724), embora altamente enriquecida em ensaio de CHIP-seq, não foi necessária para a ativação do promotor por TLRs em nosso modelo, como mostrado através de deleções e análise de gene repórter. Por outro lado, de forma semelhante ao promotor, o sítio mais distal a jusante (3') se localiza com uma região de cromatina acessível flanqueada por marcadores H3K27ac, o que poderia indicar um papel funcional, potencialmente como um enhancer (Kim and Shiekhattar, 2015). Embora passível de especulação, a função de cada um desses elementos permanece obscura. É possível que eles possam desempenhar um papel na expressão de *Rasgef1b in vivo*, em tipos celulares distintos, em diferentes contextos, ou até mesmo funcionem como elementos regulatórios distantes para genes adicionais. Para esclarecer essa questão, estudos adicionais serão necessários.

Importantemente, mostramos que a expressão de *Rasgef1b* induzida por TLRs é prejudicada em macrófagos deficientes em RelA. Vale notar, no entanto, que a expressão de *Rasgef1b* não é completamente abolida nessas células, ocorrendo parcial indução mediante ativação de TLRs. Embora RelA:p50 seja o heterodímero predominante de NF- κ B em situações inflamatórias, os dímeros de NF- κ B podem ser agrupados em várias formas distintas (Hoffmann et al., 2003). Composições distintas desses dímeros não apenas exibem papéis redundantes e compensatórios na regulação gênica, mas também possibilitam a regulação específica alguns genes alvos (Hayden and Ghosh, 2012; Sanjabi et al., 2000; Wang et al., 2012). Ao contrário de RelA que é ubiquamente expresso, cRel é constitutivamente expresso principalmente em monócitos maduros e em linhagens linfocitárias, enquanto que, em outros tipos celulares como fibroblastos, sua expressão requer estímulos variados (Grumont and Gerondakis, 1990; Liou et al., 1994). Dímeros que contêm a subunidade c-Rel parecem desempenhar um papel mais restrito, sobretudo

na resposta imune adaptativa. No entanto, esses fatores também são importantes e requeridos para a regulação de alguns promotores e para o funcionamento normal do organismo. Por exemplo, Sanjabi e colaboradores mostraram que c-Rel é necessário para a indução de IL-12-p40 em macrófagos (Sanjabi et al., 2000). Kontgen e colaboradores mostraram que camundongos deficientes de c-Rel, embora viáveis, apresentam problemas na proliferação de linfócitos e na imunidade humoral (Kontgen et al., 1995). Portanto, poderíamos especular que dímeros adicionais, independentes de RelA, tais como os dímeros contendo cRel, poderiam desempenhar um papel compensatório na expressão de *Rasgef1b* em macrófagos deficientes em RelA. Embora seja possível que isso ocorra em tipos celulares ou contextos distintos, parece improvável em macrófagos, pois mostramos que o uso de um inibidor abrangente de NF- κ B, Bay11-7082, não é suficiente para abolir a expressão de *Rasgef1b* induzida por LPS em Raw264.7. Portanto, a expressão de *Rasgef1b* induzível por TLRs em macrófagos é apenas parcialmente dependente de NF- κ B e provavelmente envolve fatores de transcrição heterólogos adicionais que atuem sinergicamente.

Nossas análises *in silico* revelaram notável consistência com nossos dados funcionais de caracterização do promotor de *Rasgef1b*. Análises filogenéticas mostraram que a sequência genômica ao redor do TSS de *Rasgef1b* é conservada entre o genoma de diversos mamíferos, incluindo primatas, o que é uma característica esperada em sequências que detêm importância funcional devido à forte seleção. Corroborando as observações preliminares (Ferreira et al., 2002), mostramos que a sequência promotora abriga uma ilha CpG (CGI – do inglês, *CpG island*). Ilhas CpG são sequências de DNA dispersas geralmente com várias centenas de pares de bases, predominantemente ricas em repetições Citosina-Guanina não metiladas e altamente associadas a regiões transcricionalmente ativas em vertebrados (Butler and Kadonaga, 2002; Deaton and Bird, 2011; Saxonov et al., 2006). Geralmente, CGIs são caracterizadas por um estado de cromatina transcricionalmente permissivo e colocalizam-se com a maioria dos promotores em vertebrados, incluindo humanos e camundongos. É relatado que cerca de 70% dos promotores de genes humanos contêm CGIs, vários deles sendo genes *housekeeping*, mas muitos são genes de expressão tecido-específica ou regulados transcricionalmente (Illingworth and Bird, 2009; Ramirez-Carrozzi et al., 2009; Smale, 2001). Ao contrário dos promotores contendo TATA ou InR que tendem a exibir um único TSS, os promotores contendo CGIs apresentam usualmente um padrão de iniciação de transcrição dispersa, isto é, o início da transcrição pode ocorrer em múltiplos locais ao

longo de uma curta região (cerca de 50 a 100 nucleotídeos) com diferentes taxas de iniciação para os diferentes TSSs (Carninci et al., 2006; Deaton and Bird, 2011; Juven-Gershon et al., 2008). De fato, embora aqui consideremos o TSS de *Rasgef1b* como o primeiro nucleotídeo da sequência referência atual (NCBI ID: NC_000071.6), e tenhamos demonstrado a presença de tal região entre os transcritos de Raw264.7, utilizando-se a técnica de CAGE (do inglês, *Cap analysis of gene expression*) foi demonstrado a presença de múltiplos supostos TSSs de *Rasgef1b* em diferentes células e tecidos de camundongo, como pode ser observado no banco de dados dbTSS (Suzuki et al., 2015). Isso indica que o início da transcrição de *Rasgef1b* ocorre provavelmente em múltiplos sítios distintos, o que pode depender do tipo de célula ou tecido em que o gene é expresso e de determinado estímulo ou contexto. Curiosamente, embora não tenhamos encontrado sítios para fatores gerais de transcrição em posições condizentes com o TSS tomado como referência no presente trabalho, um elemento TATA-box encontra-se 80 nucleotídeos a jusante (+80 a +85). Vale notar que os dados públicos de CAGE mencionados mostram uma elevada concentração de possíveis transcritos que se iniciam entre 23 a 36 nucleotídeos a jusante desse elemento TATA-box em embriões de camundongo, sugerindo um papel funcional. De fato, aparentemente a determinação do início da transcrição de *Rasgef1b* é complexa e pode depender de diversos elementos distintos que potencialmente promovem a produção de transcritos alternativos a depender do tipo celular e contexto.

É importante ressaltar que os fatores de transcrição da família Sp1 estão geralmente associados a ilhas CpG. Esses fatores de expressão ubíqua se ligam a sequências ricas em repetições de dinucleotídeos C-G e tem sido relatado que eles contribuem para a manutenção do estado hipometilado das ilhas CpG e para o recrutamento de proteína de ligação a TATA (TBP), funcionando, portanto, em sincronia com a maquinaria de transcrição basal para induzir o início da transcrição em promotores sem TATA-box (Brandeis et al., 1994; Korner et al., 1997; Pugh and Tjian, 1991; Wierstra, 2008). Como o promotor de *Rasgef1b* abriga seis elementos regulatórios para Sp1, é possível que esse fator de transcrição se ligue a essa região e auxilie no recrutamento da maquinaria de transcrição basal para iniciar a transcrição de *Rasgef1b*. Ademais, embora comumente associado à ativação transcricional constitutiva, evidências crescentes mostram que modificações pós-traducionais dependentes de estímulos, incluindo a fosforilação por MAPKs, podem impactar positiva e negativamente na ligação de Sp1 ao DNA e em sua capacidade de transativação (Benasciutti et al., 2004; Tan and Khachigian, 2009; Ueno-Shuto et al., 2014; Yurochko et al., 1997). Curiosamente, tem

sido relatado que os sítios de ligação de Sp1 são frequentemente localizados próximos a elementos κ B e que NF- κ B e Sp1 podem agir em cooperação para promover a indução de genes alvo, como *VCAM-1*, *IL6*, *Tlr2* murino, e possivelmente genes do vírus HIV (Neish et al., 1995; Oeckinghaus et al., 2011; Perkins et al., 1993; Sanceau et al., 1995; Wang et al., 2001). Como a estrutura do promotor de *Rasgef1b* é semelhante à dos genes mencionados nesse aspecto, é possível que Sp1 possa agir em sinergia com o NF- κ B para induzir sua expressão.

Trabalhos notáveis publicados pelos grupos de S.Smale e R.Medzhitov forneceram informações significativas sobre os mecanismos de regulação transcricional e da cinética de expressão de genes induzíveis por TLRs em macrófagos. Resumidamente, esses trabalhos mostraram que um grande grupo de genes de resposta precoce (PRGs, do inglês, *primary response genes*), incluindo *Tnf*, *Cxcl2*, e *Ptgs2* contém promotores com ilhas CpG, o que faz com que a ocupação de nucleossomos na região seja constitutivamente baixa (cromatina acessível) e, portanto, esses genes são independentes do remodelador de cromatina SWI/SNF (do inglês, *Switch/Sucrose Non-Fermentable*) para a sua rápida expressão induzida por LPS. Esses promotores também contêm, constitutivamente, elevados níveis de marcadores ativos H3K9/K14Ac e H3K4me3 e RNA-PolIII recrutada. Em contraste, um segundo grupo de PRGs incluindo *Ccl6*, *Saa3*, e *Ifnb1*, são expressos um pouco mais tardiamente, não possuem promotores com ilhas CpG, e são dependentes de SWI/SNF. Um terceiro grupo é formado por genes de expressão ainda mais tardia, que não incluem quaisquer das características do primeiro grupo descrito, e ainda requerem a síntese de proteínas adicionais para sua expressão. São categorizados como genes de resposta secundária e incluem *Il12b*, *Nos2*, e *Il6* (Hargreaves et al., 2009; Ramirez-Carrozzi et al., 2009; Ramirez-Carrozzi et al., 2006). Importaneamente, Hargreaves e colaboradores mostraram que Sp1 é responsável pelo recrutamento constitutivo da RNA PolIII fosforilada na Serina-5 (S5-P) ao promotor dos genes do primeiro grupo. Surpreendentemente, foi demonstrado que a Pol II-S5-P é capaz de gerar, embora em baixos níveis, transcritos imaturos (não processados) que são rapidamente degradados e provavelmente incapazes de serem traduzidos eficientemente. Somente após estímulo com LPS, NF- κ B é recrutado possibilitando a acetilação das histonas H4K5/K8/K12 ao auxiliar no recrutamento da acetiltransferase GCN5 para a região, o que é requerido para a ligação de BRD4 (do inglês, *Bromodomain-containing protein 4*) que por sua vez vai recrutar P-TEFb (do inglês, *Positive Transcription*

Elongation Factor). P-TEFb finalmente promove a fosforilação da Serina-2 da RNA Pol II, culminando em uma eficiente elongação transcricional, processamento do mRNA e produção de altos níveis de transcritos maduros para codificação de proteínas (Hargreaves et al., 2009). Portanto, nesse caso, esses trabalhos mostraram que NF- κ B não é necessário para o recrutamento da Pol II para promover o início da transcrição, mas controla eventos críticos pós-iniciação e é requerido nos promotores dos PRGs para a elongação, processamento e produção do mRNA maduro. Como mostramos aqui, o promotor de *Rasgef1b* está localizado em uma região de cromatina permissiva com baixa ocupação de nucleossomos, flanqueada por nucleossomos positivamente marcados, abriga uma ilha de CpG e uma RNA Pol II constitutivamente recrutada. As semelhanças com o modelo descrito são notáveis. Portanto, especulamos que este seria um potencial mecanismo pelo qual se dá a transcrição de *Rasgef1b* induzida por TLRs e mediada por NF- κ B.

Embora pouco se saiba sobre o perfil de expressão de *Rasgef1a* e *Rasgef1c* sob um contexto imunológico, trabalhos anteriores do nosso grupo abordaram essa questão em macrófagos peritoneais (Andrade et al., 2010). No presente trabalho, ao empregar um método altamente sensível e quantitativo, demonstramos que a expressão de *Rasgef1a* e *Rasgef1c* é extremamente baixa, próxima de níveis indetectáveis, em macrófagos derivados de medula óssea não estimulados. Após a estimulação com LPS, a expressão de *Rasgef1a* é regulada negativamente, enquanto que *Rasgef1c* é induzido. Ainda assim, quando comparados com os níveis do mRNA de *Rasgef1b* e com outros genes conhecidos induzidos por TLRs, tais como citocinas inflamatórias, seus níveis são insignificantes. Essas observações corroboram a ideia de que esses GEFs podem ter papéis muito distintos, e no caso de RasGEF1a e RasGEF1c, provavelmente não relacionados à resposta imune, ao menos em macrófagos. A função de RasGEF1b também permanece em grande parte ainda desconhecida, no entanto, sua expressão dominante em relação aos seus parálogos indica que esse GEF deve desempenhar um papel dominante nessas células.

A ativação de NF- κ B mediada por TLRs é responsável pela regulação de centenas de genes-alvo (Zhang et al., 2017). O resultado é um enorme impacto, não somente na resposta imune, mas também em outros processos importantes, como proliferação, diferenciação e sobrevivência celular (Gerondakis et al., 1999; Hayden and Ghosh, 2012). Conseqüentemente, o ajuste fino da resposta de NF- κ B é essencial para a homeostase do organismo, e naturalmente, a desregulação desse processo está associada a várias doenças

metabólicas, inflamatórias, autoimunes e câncer (Baker et al., 2011; Courtois and Gilmore, 2006; Karin, 2006). Expandir nosso conhecimento sobre o conjunto de genes alvo de NF- κ B é de importância crucial para entender melhor o amplo impacto fisiológico de tal resposta, o que também poderia, potencialmente, levar à identificação de novos alvos terapêuticos para os distúrbios associados. No presente trabalho mostramos pela primeira vez a identificação e caracterização funcional do promotor *Rasgef1b*, demonstrando a contribuição do NF- κ B em sua expressão induzida por TLRs. Além de contribuir para entender melhor o conjunto de genes regulados por NF- κ B durante a resposta imune inata mediada por TLRs, este trabalho pode fornecer a base para estudos adicionais sobre a regulação do gene RasGEF1b em outros tipos celulares e tecidos, assim como pode auxiliar no entendimento dos potenciais mecanismos regulatórios de genes estruturalmente semelhantes.

7. CONCLUSÕES

Com base no conjunto de dados apresentados nesse trabalho, concluímos que:

- 1- RasGEF1b é predominantemente expresso em BMDMs murinos em relação aos parálogos RasGEF1a e RasGEF1c.
- 2- A sequência proximal que inclui o TSS reportado na sequência referência está presente nos transcritos de *Rasgef1b* em macrófagos.
- 3- O promotor ativo de *Rasgef1b* está contido em uma sequência proximal entre -183 e +119 que é conservada evolutivamente, abriga elementos regulatórios responsivos à ativação de TLRs, está contida em uma região de cromatina acessível e ilha CpG, e potencialmente recruta constitutivamente a RNA PolII.
- 4- O fator de transcrição NF- κ B se liga a elementos κ B específicos no promotor de *Rasgef1b* assim como a regiões distais e contribui para sua expressão induzida por TLRs.

O presente trabalho fornece um conjunto de dados inéditos que contribuem para a compreensão da regulação transcricional de *Rasgef1b*, da região genômica envolvida nessa função, e de parte dos componentes moleculares que atuam sobre ela, fornecendo subsídios para estudos futuros sobre o papel desse gene na resposta imune inata mediada por TLRs em macrófagos.

8. TRABALHOS PUBLICADOS OU EM SUBMISSÃO

1. CARVALHO, BRENER C. ; OLIVEIRA, LEONARDO C. ; ROCHA, CAROLINA D. ; FERNANDES, HELIANA B. ; OLIVEIRA, ISADORA M. ; **LEÃO, FELIPE B.** ; VALVERDE, THALITA M. ; REGO, IGOR M.G. ; GHOSH, SANKAR ; SILVA, ARISTÓBOLO M. . **Both knock-down and overexpression of Rap2a small GTPase in macrophages result in impairment of NF- κ B activity and inflammatory gene expression.** MOLECULAR IMMUNOLOGY, v. 109, p. 27-37, 2019.

2. CARVALHO, BRENER C. ; OLIVEIRA, LEONARDO C. ; ROCHA, CAROLINA D. ; FERNANDES, HELIANA B. ; OLIVEIRA, ISADORA M. ; **LEÃO, FELIPE B.** ; VALVERDE, THALITA M. ; REGO, IGOR M.G. ; GHOSH, SANKAR ; SILVA, ARISTÓBOLO M. . **Data in support of Rap2a GTPase expression, activation and effects in LPS-mediated innate immune response and NF- κ B activation.** DATA IN BRIEF, v. 24, p. 103965, 2019.

3. ZAMUDIO, ROXANA ; PEREIRA, LATIFE ; ROCHA, CAROLINA D. ; BERG, DOUGLAS E. ; MUNIZ-QUEIROZ, THAÍS ; SANT ANNA, HANAISA P. ; CABRERA, LILIA ; COMBE, JUAN M. ; HERRERA, PHABIOLA ; JAHUIRA, MARTHA H. ; **LEÃO, FELIPE B.** ; LYON, FERNANDA ; PRADO, WILLIAM A. ; RODRIGUES, MAÍRA R. ; RODRIGUES-SOARES, FERNANDA ; SANTOLALLA, MEDDLY L. ; ZOLINI, CAMILA ; SILVA, ARISTÓBOLO M. ; GILMAN, ROBERT H. ; TARAZONA-SANTOS, EDUARDO ; KEHDY, FERNANDA S. G. . **Population, Epidemiological, and Functional Genetics of Gastric Cancer Candidate Genes in Peruvians with Predominant Amerindian Ancestry.** DIGESTIVE DISEASES AND SCIENCES, v. 61, p. 107-116, 2016.

4. **LEÃO, FELIPE B.** ; CARVALHO, BRENER C.; OLIVEIRA, LEONARDO C.; VAUGHAN, LAUREN S. ; BHATT, DEV M. ; LIAO, WILL ; MALONEY, DILLON. **Transcriptional mechanisms of TLR-inducible Rasgef1b: Promoter characterization and the role of NF- κ B.** (Em preparo para submissão).

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

9. REFERÊNCIAS

(2019). UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 47, D506-d515.

Adachi, R., Krilis, S.A., Nigrovic, P.A., Hamilton, M.J., Chung, K., Thakurdas, S.M., Boyce, J.A., Anderson, P., and Stevens, R.L. (2012). Ras guanine nucleotide-releasing protein-4 (RasGRP4) involvement in experimental arthritis and colitis. *The Journal of biological chemistry* 287, 20047-20055.

Ahearn, I.M., Haigis, K., Bar-Sagi, D., and Philips, M.R. (2011). Regulating the regulator: post-translational modification of RAS. *Nat Rev Mol Cell Biol* 13, 39-51.

Akira, S., Takeda, K., and Kaisho, T. (2001). Toll-like receptors: critical proteins linking innate and acquired immunity. *Nature immunology* 2, 675-680.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. *Cell* 124, 783-801.

Akiyama, Y. (1995). TFSEARCH: Searching Transcription Factor Binding Sites. <http://www.rwcp.or.jp/lab/pdappl/papiahtml>.

Albiger, B., Dahlberg, S., Henriques-Normark, B., and Normark, S. (2007). Role of the innate immune system in host defence against bacterial infections: focus on the Toll-like receptors. *Journal of internal medicine* 261, 511-528.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *Journal of molecular biology* 215, 403-410.

Anderson, K.V., Bokla, L., and Nusslein-Volhard, C. (1985). Establishment of dorsal-ventral polarity in the *Drosophila* embryo: the induction of polarity by the Toll gene product. *Cell* 42, 791-798.

Andrade, W.A., Silva, A.M., Alves, V.S., Salgado, A.P., Melo, M.B., Andrade, H.M., Dall'Orto, F.V., Garcia, S.A., Silveira, T.N., and Gazzinelli, R.T. (2010). Early endosome localization and activity of RasGEF1b, a toll-like receptor-inducible Ras guanine-nucleotide exchange factor. *Genes and immunity* 11, 447-457.

Anwar, M.A., Basith, S., and Choi, S. (2013). Negative regulatory approaches to the attenuation of Toll-like receptor signaling. *Exp Mol Med* 45, e11.

Arigoni, M., Bracco, E., Lusche, D.F., Kae, H., Weeks, G., and Bozzaro, S. (2005). A novel *Dictyostelium* RasGEF required for chemotaxis and development. *BMC Cell Biol* 6, 43.

Aronheim, A., Engelberg, D., Li, N., al-Alawi, N., Schlessinger, J., and Karin, M. (1994). Membrane targeting of the nucleotide exchange factor Sos is sufficient for activating the Ras signaling pathway. *Cell* 78, 949-961.

- Baker, R.G., Hayden, M.S., and Ghosh, S. (2011). NF-kappaB, inflammation, and metabolic disease. *Cell Metab* *13*, 11-22.
- Bao, X., Rubin, A.J., Qu, K., Zhang, J., Giresi, P.G., Chang, H.Y., and Khavari, P.A. (2015). A novel ATAC-seq approach reveals lineage-specific reinforcement of the open chromatin landscape via cooperation between BAF and p63. *Genome biology* *16*, 284.
- Bar-Sagi, D., and Feramisco, J.R. (1985). Microinjection of the ras oncogene protein into PC12 cells induces morphological differentiation. *Cell* *42*, 841-848.
- Benasciutti, E., Pages, G., Kenzior, O., Folk, W., Blasi, F., and Crippa, M.P. (2004). MAPK and JNK transduction pathways can phosphorylate Sp1 to activate the uPA minimal promoter element and endogenous gene transcription. *Blood* *104*, 256-262.
- Bergmeier, W., Goerge, T., Wang, H.W., Crittenden, J.R., Baldwin, A.C., Cifuni, S.M., Housman, D.E., Graybiel, A.M., and Wagner, D.D. (2007). Mice lacking the signaling molecule CalDAG-GEFI represent a model for leukocyte adhesion deficiency type III. *The Journal of clinical investigation* *117*, 1699-1707.
- Blanchette, M., Kent, W.J., Riemer, C., Elnitski, L., Smit, A.F., Roskin, K.M., Baertsch, R., Rosenbloom, K., Clawson, H., Green, E.D., *et al.* (2004). Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res* *14*, 708-715.
- Blasi, E., Radzioch, D., Merletti, L., and Varesio, L. (1989). Generation of macrophage cell line from fresh bone marrow cells with a myc/raf recombinant retrovirus. *Cancer biochemistry biophysics* *10*, 303-317.
- Boguski, M.S., and McCormick, F. (1993). Proteins regulating Ras and its relatives. *Nature* *366*, 643-654.
- Bonizzi, G., and Karin, M. (2004). The two NF-kappaB activation pathways and their role in innate and adaptive immunity. *Trends in immunology* *25*, 280-288.
- Bonnet, C., Andrieux, J., Beri-Dexheimer, M., Leheup, B., Boute, O., Manouvrier, S., Delobel, B., Copin, H., Receveur, A., Mathieu, M., *et al.* (2010). Microdeletion at chromosome 4q21 defines a new emerging syndrome with marked growth restriction, mental retardation and absent or severely delayed speech. *Journal of medical genetics* *47*, 377-384.
- Bos, J.L., Rehmann, H., and Wittinghofer, A. (2007). GEFs and GAPs: critical elements in the control of small G proteins. *Cell* *129*, 865-877.
- Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Nemes, A., Temper, V., Razin, A., and Cedar, H. (1994). Sp1 elements protect a CpG island from de novo methylation. *Nature* *371*, 435-438.
- Broek, D., Toda, T., Michaeli, T., Levin, L., Birchmeier, C., Zoller, M., Powers, S., and Wigler, M. (1987). The *S. cerevisiae* CDC25 gene product regulates the RAS/adenylate cyclase pathway. *Cell* *48*, 789-799.

Brubaker, S.W., Bonham, K.S., Zanoni, I., and Kagan, J.C. (2015). Innate immune pattern recognition: a cell biological perspective. *Annual review of immunology* 33, 257-290.

Buday, L., and Downward, J. (2008). Many faces of Ras activation. *Biochimica et biophysica acta* 1786, 178-187.

Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature methods* 10, 1213-1218.

Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. *Current protocols in molecular biology* 109, 21.29.21-29.

Butler, J.E., and Kadonaga, J.T. (2002). The RNA polymerase II core promoter: a key component in the regulation of gene expression. *Genes & development* 16, 2583-2592.

Campbell, S.L., Khosravi-Far, R., Rossman, K.L., Clark, G.J., and Der, C.J. (1998). Increasing complexity of Ras signaling. *Oncogene* 17, 1395-1413.

Carbo, C., Duerschmied, D., Goerge, T., Hattori, H., Sakai, J., Cifuni, S.M., White, G.C., 2nd, Chrzanowska-Wodnicka, M., Luo, H.R., and Wagner, D.D. (2010). Integrin-independent role of CalDAG-GEFI in neutrophil chemotaxis. *Journal of leukocyte biology* 88, 313-319.

Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C.A., Taylor, M.S., Engstrom, P.G., Frith, M.C., *et al.* (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nature genetics* 38, 626-635.

Carvalho, B.C., Oliveira, L.C., Rocha, C.D., Fernandes, H.B., Oliveira, I.M., Leao, F.B., Valverde, T.M., Rego, I.M.G., Ghosh, S., and Silva, A.M. (2019). Both knock-down and overexpression of Rap2a small GTPase in macrophages result in impairment of NF-kappaB activity and inflammatory gene expression. *Molecular immunology* 109, 27-37.

Cherfils, J., and Zeghouf, M. (2013). Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* 93, 269-309.

Chiaromonte, F., Yap, V.B., and Miller, W. (2002). Scoring pairwise genomic sequence alignments. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing*, 115-126.

Choy, E., Chiu, V.K., Silletti, J., Feoktistov, M., Morimoto, T., Michaelson, D., Ivanov, I.E., and Philips, M.R. (1999). Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi. *Cell* 98, 69-80.

Coleman, M.L., Marshall, C.J., and Olson, M.F. (2004). RAS and RHO GTPases in G1-phase cell-cycle regulation. *Nat Rev Mol Cell Biol* 5, 355-366.

Consortium, T.E.P.e.a. (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74.

Coughlin, J.J., Stang, S.L., Dower, N.A., and Stone, J.C. (2005). RasGRP1 and RasGRP3 regulate B cell proliferation by facilitating B cell receptor-Ras signaling. *J Immunol* 175, 7179-7184.

Courtois, G., and Gilmore, T.D. (2006). Mutations in the NF-kappaB signaling pathway: implications for human disease. *Oncogene* 25, 6831-6843.

Crittenden, J.R., Bergmeier, W., Zhang, Y., Piffath, C.L., Liang, Y., Wagner, D.D., Housman, D.E., and Graybiel, A.M. (2004). CalDAG-GEFI integrates signaling for platelet aggregation and thrombus formation. *Nat Med* 10, 982-986.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. *Genes & development* 25, 1010-1022.

Dempsey, P.W., Vaidya, S.A., and Cheng, G. (2003). The art of war: Innate and adaptive immune responses. *Cellular and molecular life sciences : CMLS* 60, 2604-2621.

Downward, J. (2003). Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer* 3, 11-22.

Downward, J., Riehl, R., Wu, L., and Weinberg, R.A. (1990). Identification of a nucleotide exchange-promoting activity for p21ras. *Proc Natl Acad Sci U S A* 87, 5998-6002.

Doyle, S.L., and O'Neill, L.A. (2006). Toll-like receptors: from the discovery of NFkappaB to new insights into transcriptional regulations in innate immunity. *Biochem Pharmacol* 72, 1102-1113.

Ebinu, J.O., Bottorff, D.A., Chan, E.Y., Stang, S.L., Dunn, R.J., and Stone, J.C. (1998). RasGRP, a Ras guanyl nucleotide- releasing protein with calcium- and diacylglycerol-binding motifs. *Science* 280, 1082-1086.

Epting, D., Vorwerk, S., Hageman, A., and Meyer, D. (2007). Expression of rasgef1b in zebrafish. *Gene expression patterns : GEP* 7, 389-395.

Farnsworth, C.L., Freshney, N.W., Rosen, L.B., Ghosh, A., Greenberg, M.E., and Feig, L.A. (1995). Calcium activation of Ras mediated by neuronal exchange factor Ras-GRF. *Nature* 376, 524-527.

Ferreira, L.R., Abrantes, E.F., Rodrigues, C.V., Caetano, B., Cerqueira, G.C., Salim, A.C., Reis, L.F., and Gazzinelli, R.T. (2002). Identification and characterization of a novel mouse gene encoding a Ras-associated guanine nucleotide exchange factor:

- expression in macrophages and myocarditis elicited by *Trypanosoma cruzi* parasites. *Journal of leukocyte biology* 72, 1215-1227.
- Foster, S.L., and Medzhitov, R. (2009). Gene-specific control of the TLR-induced inflammatory response. *Clinical immunology* 130, 7-15.
- Frazer, K.A., Pachter, L., Poliakov, A., Rubin, E.M., and Dubchak, I. (2004). VISTA: computational tools for comparative genomics. *Nucleic Acids Res* 32, W273-279.
- Gerondakis, S., Grossmann, M., Nakamura, Y., Pohl, T., and Grumont, R. (1999). Genetic approaches in mice to understand Rel/NF-kappaB and IkappaB function: transgenics and knockouts. *Oncogene* 18, 6888-6895.
- Ghosh, S., and Karin, M. (2002). Missing pieces in the NF-kappaB puzzle. *Cell* 109 *Suppl*, S81-96.
- Gloerich, M., and Bos, J.L. (2011). Regulating Rap small G-proteins in time and space. *Trends Cell Biol* 21, 615-623.
- Goitre, L., Trapani, E., Trabalzini, L., and Retta, S.F. (2014). The Ras superfamily of small GTPases: the unlocked secrets. *Methods in molecular biology (Clifton, NJ)* 1120, 1-18.
- Grossmann, M., Metcalf, D., Merryfull, J., Beg, A., Baltimore, D., and Gerondakis, S. (1999). The combined absence of the transcription factors Rel and RelA leads to multiple hemopoietic cell defects. *Proc Natl Acad Sci U S A* 96, 11848-11853.
- Grumont, R.J., and Gerondakis, S. (1990). Murine c-rel transcription is rapidly induced in T-cells and fibroblasts by mitogenic agents and the phorbol ester 12-O-tetradecanoylphorbol-13-acetate. *Cell Growth Differ* 1, 345-350.
- Hargreaves, D.C., Horng, T., and Medzhitov, R. (2009). Control of inducible gene expression by signal-dependent transcriptional elongation. *Cell* 138, 129-145.
- Hatada, E.N., Nieters, A., Wulczyn, F.G., Naumann, M., Meyer, R., Nucifora, G., McKeithan, T.W., and Scheidereit, C. (1992). The ankyrin repeat domains of the NF-kappa B precursor p105 and the protooncogene bcl-3 act as specific inhibitors of NF-kappa B DNA binding. *Proc Natl Acad Sci U S A* 89, 2489-2493.
- Hayden, M.S., and Ghosh, S. (2004). Signaling to NF-kappaB. *Genes & development* 18, 2195-2224.
- Hayden, M.S., and Ghosh, S. (2012). NF-kappaB, the first quarter-century: remarkable progress and outstanding questions. *Genes & development* 26, 203-234.
- Hennig, A., Markwart, R., Esparza-Franco, M.A., Ladds, G., and Rubio, I. (2015). Ras activation revisited: role of GEF and GAP systems. *Biological chemistry* 396, 831-848.

Hoffmann, A., Leung, T.H., and Baltimore, D. (2003). Genetic analysis of NF-kappaB/Rel transcription factors defines functional specificities. *The EMBO journal* 22, 5530-5539.

Hu, X., Chen, X., Wu, B., Soler, I.M., Chen, S., and Shen, Y. (2017). Further defining the critical genes for the 4q21 microdeletion disorder. *American journal of medical genetics Part A* 173, 120-125.

Huang, D.B., Vu, D., and Ghosh, G. (2005). NF-kappaB RelB forms an intertwined homodimer. *Structure (London, England : 1993)* 13, 1365-1373.

Huxford, T., Huang, D.B., Malek, S., and Ghosh, G. (1998). The crystal structure of the IkappaBalpha/NF-kappaB complex reveals mechanisms of NF-kappaB inactivation. *Cell* 95, 759-770.

Illingworth, R.S., and Bird, A.P. (2009). CpG islands--'a rough guide'. *FEBS Lett* 583, 1713-1720.

Jacobs, M.D., and Harrison, S.C. (1998). Structure of an IkappaBalpha/NF-kappaB complex. *Cell* 95, 749-758.

Janeway, C.A., Jr., and Medzhitov, R. (2002). Innate immune recognition. *Annual review of immunology* 20, 197-216.

Jareborg, N., Birney, E., and Durbin, R. (1999). Comparative analysis of noncoding regions of 77 orthologous mouse and human gene pairs. *Genome Res* 9, 815-824.

Juven-Gershon, T., Hsu, J.Y., Theisen, J.W., and Kadonaga, J.T. (2008). The RNA polymerase II core promoter - the gateway to transcription. *Current opinion in cell biology* 20, 253-259.

Kahana, J.A., and Cleveland, D.W. (1999). Beyond nuclear transport. Ran-GTP as a determinant of spindle assembly. *J Cell Biol* 146, 1205-1210.

Kanzler, H., Barrat, F.J., Hessel, E.M., and Coffman, R.L. (2007). Therapeutic targeting of innate immunity with Toll-like receptor agonists and antagonists. *Nat Med* 13, 552-559.

Karin, M. (2006). Nuclear factor-kappaB in cancer development and progression. *Nature* 441, 431-436.

Karin, M., and Ben-Neriah, Y. (2000). Phosphorylation meets ubiquitination: the control of NF-[kappa]B activity. *Annual review of immunology* 18, 621-663.

Kauffmann-Zeh, A., Rodriguez-Viciana, P., Ulrich, E., Gilbert, C., Coffey, P., Downward, J., and Evan, G. (1997). Suppression of c-Myc-induced apoptosis by Ras signalling through PI(3)K and PKB. *Nature* 385, 544-548.

Kawai, T., and Akira, S. (2006). TLR signaling. *Cell death and differentiation* 13, 816-825.

Kawai, T., and Akira, S. (2007). Signaling to NF-kappaB by Toll-like receptors. *Trends in molecular medicine* *13*, 460-469.

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nature immunology* *11*, 373-384.

Kawasaki, T., and Kawai, T. (2014). Toll-like receptor signaling pathways. *Front Immunol* *5*, 461.

Kim, T.K., and Shiekhatar, R. (2015). Architectural and Functional Commonalities between Enhancers and Promoters. *Cell* *162*, 948-959.

Komlosi, K., Duga, B., Hadzsiev, K., Czako, M., Kosztolanyi, G., Fogarasi, A., and Melegh, B. (2015). Phenotypic variability in a Hungarian patient with the 4q21 microdeletion syndrome. *Molecular cytogenetics* *8*, 16.

Kontgen, F., Grumont, R.J., Strasser, A., Metcalf, D., Li, R., Tarlinton, D., and Gerondakis, S. (1995). Mice lacking the c-rel proto-oncogene exhibit defects in lymphocyte proliferation, humoral immunity, and interleukin-2 expression. *Genes & development* *9*, 1965-1977.

Korner, K., Wolfrain, L.A., Lucibello, F.C., and Muller, R. (1997). Characterization of the TATA-less core promoter of the cell cycle-regulated *cdc25C* gene. *Nucleic Acids Res* *25*, 4933-4939.

Kumar, H., Kawai, T., and Akira, S. (2011). Pathogen recognition by the innate immune system. *Int Rev Immunol* *30*, 16-34.

Leão, F. (2013). Clonagem e caracterização funcional da região promotora e de deleções progressivas da 5' flanqueadora de *Rasgef1b*, um gene induzido por receptores do tipo Toll (TLRs) (Belo Horizonte: Universidade Federal de Minas Gerais).

Lekstrom-Himes, J. (1998). Biological Role of the CCAAT/Enhancer-binding Protein Family of Transcription Factors. *Journal of Biological Chemistry* *273*, 28545-28548.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., and Hoffmann, J.A. (1996). The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell* *86*, 973-983.

Lemmon, M.A., Ferguson, K.M., and Schlessinger, J. (1996). PH domains: diverse sequences with a common fold recruit signaling molecules to the cell surface. *Cell* *85*, 621-624.

Lemmon, M.A., and Schlessinger, J. (2010). Cell signaling by receptor tyrosine kinases. *Cell* *141*, 1117-1134.

Li, H., Coghlan, A., Ruan, J., Coin, L.J., Heriche, J.K., Osmotherly, L., Li, R., Liu, T., Zhang, Z., Bolund, L., *et al.* (2006). TreeFam: a curated database of phylogenetic trees of animal gene families. *Nucleic Acids Res* *34*, D572-580.

Li, L.C., and Dahiya, R. (2002). MethPrimer: designing primers for methylation PCRs. *Bioinformatics (Oxford, England)* 18, 1427-1431.

Liou, H.C., Sha, W.C., Scott, M.L., and Baltimore, D. (1994). Sequential induction of NF-kappa B/Rel family proteins during B-cell terminal differentiation. *Mol Cell Biol* 14, 5349-5359.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-408.

Lu, Y.C., Yeh, W.C., and Ohashi, P.S. (2008). LPS/TLR4 signal transduction pathway. *Cytokine* 42, 145-151.

Malek, S., Huang, D.B., Huxford, T., Ghosh, S., and Ghosh, G. (2003). X-ray crystal structure of an IkappaBbeta x NF-kappaB p65 homodimer complex. *The Journal of biological chemistry* 278, 23094-23100.

Medzhitov, R. (2007). Recognition of microorganisms and activation of the immune response. *Nature* 449, 819-826.

Medzhitov, R., Preston-Hurlburt, P., and Janeway, C.A., Jr. (1997). A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature* 388, 394-397.

Meininger, I., Griesbach, R.A., Hu, D., Gehring, T., Seeholzer, T., Bertossi, A., Kranich, J., Oeckinghaus, A., Eitelhuber, A.C., Greczmiel, U., *et al.* (2016). Alternative splicing of MALT1 controls signalling and activation of CD4(+) T cells. *Nature communications* 7, 11292.

Mifsud, E.J., Tan, A.C., and Jackson, D.C. (2014). TLR Agonists as Modulators of the Innate Immune Response and Their Potential as Agents Against Infectious Disease. *Front Immunol* 5, 79.

Mikita, T., Kurama, M., and Schindler, U. (1998). Synergistic activation of the germline epsilon promoter mediated by Stat6 and C/EBP beta. *J Immunol* 161, 1822-1828.

Mukherjee, S.P., Behar, M., Birnbaum, H.A., Hoffmann, A., Wright, P.E., and Ghosh, G. (2013). Analysis of the RelA:CBP/p300 interaction reveals its involvement in NF-kappaB-driven transcription. *PLoS biology* 11, e1001647.

Neish, A.S., Khachigian, L.M., Park, A., Baichwal, V.R., and Collins, T. (1995). Sp1 is a component of the cytokine-inducible enhancer in the promoter of vascular cell adhesion molecule-1. *The Journal of biological chemistry* 270, 28903-28909.

Ng, W.L., and Marinov, G.K. (2017). Transcriptomic analysis of the role of RasGEF1B circular RNA in the TLR4/LPS pathway. *7*, 12227.

- Ng, W.L., Marinov, G.K., Liao, E.S., Lam, Y.L., Lim, Y.Y., and Ea, C.K. (2016). Inducible RasGEF1B circular RNA is a positive regulator of ICAM-1 in the TLR4/LPS pathway. *RNA biology* *13*, 861-871.
- Novick, P., Field, C., and Schekman, R. (1980). Identification of 23 complementation groups required for post-translational events in the yeast secretory pathway. *Cell* *21*, 205-215.
- O'Neill, L.A., and Bowie, A.G. (2007). The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat Rev Immunol* *7*, 353-364.
- Oeckinghaus, A., Hayden, M.S., and Ghosh, S. (2011). Crosstalk in NF-kappaB signaling pathways. *Nature immunology* *12*, 695-708.
- Perkins, N.D., Edwards, N.L., Duckett, C.S., Agranoff, A.B., Schmid, R.M., and Nabel, G.J. (1993). A cooperative interaction between NF-kappa B and Sp1 is required for HIV-1 enhancer activation. *The EMBO journal* *12*, 3551-3558.
- Plevy, S.E., Gemberling, J.H., Hsu, S., Dorner, A.J., and Smale, S.T. (1997). Multiple control elements mediate activation of the murine and human interleukin 12 p40 promoters: evidence of functional synergy between C/EBP and Rel proteins. *Mol Cell Biol* *17*, 4572-4588.
- Priatel, J.J., Chen, X., Zenewicz, L.A., Shen, H., Harder, K.W., Horwitz, M.S., and Teh, H.S. (2007). Chronic immunodeficiency in mice lacking RasGRP1 results in CD4 T cell immune activation and exhaustion. *J Immunol* *179*, 2143-2152.
- Pugh, B.F., and Tjian, R. (1991). Transcription from a TATA-less promoter requires a multisubunit TFIID complex. *Genes & development* *5*, 1935-1945.
- Quandt, K., Frech, K., Karas, H., Wingender, E., and Werner, T. (1995). MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. *Nucleic Acids Res* *23*, 4878-4884.
- Quilliam, L.A., Khosravi-Far, R., Huff, S.Y., and Der, C.J. (1995). Guanine nucleotide exchange factors: activators of the Ras superfamily of proteins. *Bioessays* *17*, 395-404.
- Quilliam, L.A., Rebhun, J.F., and Castro, A.F. (2002). A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases. *Progress in nucleic acid research and molecular biology* *71*, 391-444.
- Raetz, M., Kibardin, A., Sturge, C.R., Pifer, R., Li, H., Burstein, E., Ozato, K., Larin, S., and Yarovinsky, F. (2015). Cooperation of TLR12 and TLR11 in the IRF8-dependent IL-12 response to *Toxoplasma gondii* profilin. *J Immunol* *191*, 4818-4827.
- Ralph, W.M., Jr., Liu, K., and Auborn, K.J. (2006). CCAAT/enhancer-binding protein beta represses human papillomavirus 11 upstream regulatory region expression through a promoter-proximal YY1-binding site. *J Gen Virol* *87*, 51-59.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* 138, 114-128.

Ramirez-Carrozzi, V.R., Nazarian, A.A., Li, C.C., Gore, S.L., Sridharan, R., Imbalzano, A.N., and Smale, S.T. (2006). Selective and antagonistic functions of SWI/SNF and Mi-2beta nucleosome remodeling complexes during an inflammatory response. *Genes & development* 20, 282-296.

Repasky, G.A., Chenette, E.J., and Der, C.J. (2004). Renewing the conspiracy theory debate: does Raf function alone to mediate Ras oncogenesis? *Trends Cell Biol* 14, 639-647.

Rice, P., Longden, I., and Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16, 276-277.

Ridley, A.J., and Hall, A. (1992). The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors. *Cell* 70, 389-399.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nature biotechnology* 29, 24-26.

Robinson, L.C., Gibbs, J.B., Marshall, M.S., Sigal, I.S., and Tatchell, K. (1987). CDC25: a component of the RAS-adenylate cyclase pathway in *Saccharomyces cerevisiae*. *Science* 235, 1218-1221.

Rothwarf, D.M., and Karin, M. (1999). The NF-kappa B activation pathway: a paradigm in information transfer from membrane to nucleus. *Sci STKE* 1999, RE1.

Ruben, S.M., Klement, J.F., Coleman, T.A., Maher, M., Chen, C.H., and Rosen, C.A. (1992). I-Rel: a novel rel-related protein that inhibits NF-kappa B transcriptional activity. *Genes & development* 6, 745-760.

Saccani, S., Pantano, S., and Natoli, G. (2003). Modulation of NF-kappaB activity by exchange of dimers. *Mol Cell* 11, 1563-1574.

Sachdeva, M., Liu, Q., Cao, J., Lu, Z., and Mo, Y.Y. (2012). Negative regulation of miR-145 by C/EBP-beta through the Akt pathway in cancer cells. *Nucleic Acids Res* 40, 6683-6692.

Sanceau, J., Kaisho, T., Hirano, T., and Wietzerbin, J. (1995). Triggering of the human interleukin-6 gene by interferon-gamma and tumor necrosis factor-alpha in monocytic cells involves cooperation between interferon regulatory factor-1, NF kappa B, and Sp1 transcription factors. *The Journal of biological chemistry* 270, 27920-27931.

Sanjabi, S., Hoffmann, A., Liou, H.C., Baltimore, D., and Smale, S.T. (2000). Selective requirement for c-Rel during IL-12 P40 gene induction in macrophages. *Proc Natl Acad Sci U S A* 97, 12705-12710.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A* *103*, 1412-1417.

Schreibelt, G., Tel, J., Sliepen, K.H., Benitez-Ribas, D., Figdor, C.G., Adema, G.J., and de Vries, I.J. (2010). Toll-like receptor expression and function in human dendritic cell subsets: implications for dendritic cell-based anti-cancer immunotherapy. *Cancer immunology, immunotherapy : CII* *59*, 1573-1582.

Schwartz, S., Kent, W.J., Smit, A., Zhang, Z., Baertsch, R., Hardison, R.C., Haussler, D., and Miller, W. (2003). Human-mouse alignments with BLASTZ. *Genome Res* *13*, 103-107.

Sen, R., and Baltimore, D. (1986). Multiple nuclear factors interact with the immunoglobulin enhancer sequences. *Cell* *46*, 705-716.

Shou, C., Farnsworth, C.L., Neel, B.G., and Feig, L.A. (1992). Molecular cloning of cDNAs encoding a guanine-nucleotide-releasing factor for Ras p21. *Nature* *358*, 351-354.

Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., *et al.* (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res* *15*, 1034-1050.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., *et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* *7*, 539.

Siggers, T., Chang, A.B., Teixeira, A., Wong, D., Williams, K.J., Ahmed, B., Ragoussis, J., Udalova, I.A., Smale, S.T., and Bulyk, M.L. (2011). Principles of dimer-specific gene regulation revealed by a comprehensive characterization of NF-kappaB family DNA binding. *Nature immunology* *13*, 95-102.

Simon, M.A., Bowtell, D.D., Dodson, G.S., Laverty, T.R., and Rubin, G.M. (1991). Ras1 and a putative guanine nucleotide exchange factor perform crucial steps in signaling by the sevenless protein tyrosine kinase. *Cell* *67*, 701-716.

Smale, S.T. (2001). Core promoters: active contributors to combinatorial gene regulation. *Genes & development* *15*, 2503-2508.

Smale, S.T. (2012). Dimer-specific regulatory mechanisms within the NF-kappaB family of transcription factors. *Immunological reviews* *246*, 193-204.

Stone, J.C. (2011). Regulation and Function of the RasGRP Family of Ras Activators in Blood Cells. *Genes & cancer* *2*, 320-334.

Sun, S.C. (2011). Non-canonical NF-kappaB signaling pathway. *Cell Res* *21*, 71-85.

Suzuki, A., Wakaguri, H., Yamashita, R., Kawano, S., Tsuchihara, K., Sugano, S., Suzuki, Y., and Nakai, K. (2015). DBTSS as an integrative platform for transcriptome, epigenome and genome sequence variation data. *Nucleic Acids Res* 43, D87-91.

Takagi, M. (2011). Toll-like receptor--a potent driving force behind rheumatoid arthritis. *J Clin Exp Hematop* 51, 77-92.

Takai, Y., Sasaki, T., and Matozaki, T. (2001). Small GTP-binding proteins. *Physiol Rev* 81, 153-208.

Tan, N.Y., and Khachigian, L.M. (2009). Sp1 phosphorylation and its regulation of gene transcription. *Mol Cell Biol* 29, 2483-2488.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., *et al.* (2012). The accessible chromatin landscape of the human genome. *Nature* 489, 75-82.

Tsunoda, T., and Takagi, T. (1999). Estimating transcription factor bindability on DNA. *Bioinformatics (Oxford, England)* 15, 622-630.

UCSC (2012). An integrated encyclopedia of DNA elements in the human genome. *Nature* 489, 57-74.

Udofa, E.A., Stringer, B.W., Gade, P., Mahony, D., Buzza, M.S., Kalvakolanu, D.V., and Antalis, T.M. (2013). The Transcription Factor C/EBP- β Mediates Constitutive and LPS-Inducible Transcription of Murine SerpinB2. *PloS one* 8.

Ueno-Shuto, K., Kato, K., Tasaki, Y., Sato, M., Sato, K., Uchida, Y., Sakai, H., Ono, T., Suico, M.A., Mitsutake, K., *et al.* (2014). Lipopolysaccharide decreases single immunoglobulin interleukin-1 receptor-related molecule (SIGIRR) expression by suppressing specificity protein 1 (Sp1) via the Toll-like receptor 4 (TLR4)-p38 pathway in monocytes and neutrophils. *The Journal of biological chemistry* 289, 18097-18109.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res* 40, e115.

Vallabhapurapu, S., and Karin, M. (2009). Regulation and function of NF-kappaB transcription factors in the immune system. *Annual review of immunology* 27, 693-733.

van Dam, T.J., Bos, J.L., and Snel, B. Evolution of the Ras-like small GTPases and their regulators. *Small GTPases* 2, 4-16.

van Kuppeveld, F.J., Johansson, K.E., Galama, J.M., Kissing, J., Bolske, G., van der Logt, J.T., and Melchers, W.J. (1994). Detection of mycoplasma contamination in cell cultures by a mycoplasma group-specific PCR. *Appl Environ Microbiol* 60, 149-152.

Vigil, D., Cherfils, J., Rossman, K.L., and Der, C.J. Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat Rev Cancer* 10, 842-857.

- Vigil, D., Cherfilis, J., Rossman, K.L., and Der, C.J. (2010). Ras superfamily GEFs and GAPs: validated and tractable targets for cancer therapy? *Nat Rev Cancer* *10*, 842-857.
- Wang, H., Liu, K., Yuan, F., Berdichevsky, L., Taichman, L.B., and Auborn, K. (1996). C/EBPbeta is a negative regulator of human papillomavirus type 11 in keratinocytes. *J Virol* *70*, 4839-4844.
- Wang, T., Lafuse, W.P., and Zwilling, B.S. (2001). NFkappaB and Sp1 elements are necessary for maximal transcription of toll-like receptor 2 induced by *Mycobacterium avium*. *J Immunol* *167*, 6924-6932.
- Wang, V.Y., Huang, W., Asagiri, M., Spann, N., Hoffmann, A., Glass, C., and Ghosh, G. (2012). The transcriptional specificity of NF-kappaB dimers is coded within the kappaB DNA response elements. *Cell reports* *2*, 824-839.
- Wang, Y., Liu, L., Davies, D.R., and Segal, D.M. (2010). Dimerization of Toll-like receptor 3 (TLR3) is required for ligand binding. *J Biol Chem* *285*, 36836-36841.
- Waters, S.B., Yamauchi, K., and Pessin, J.E. (1995). Insulin-stimulated disassociation of the SOS-Grb2 complex. *Mol Cell Biol* *15*, 2791-2799.
- Whitman, M., and Melton, D.A. (1992). Involvement of p21ras in *Xenopus* mesoderm induction. *Nature* *357*, 252-254.
- Wierstra, I. (2008). Sp1: emerging roles--beyond constitutive activation of TATA-less housekeeping genes. *Biochemical and biophysical research communications* *372*, 1-13.
- Wilkins, A., and Insall, R.H. (2001). Small GTPases in *Dictyostelium*: lessons from a social amoeba. *Trends Genet* *17*, 41-48.
- Willumsen, B.M., Christensen, A., Hubbert, N.L., Papageorge, A.G., and Lowy, D.R. (1984). The p21 ras C-terminus is required for transformation and membrane association. *Nature* *310*, 583-586.
- Wolfman, A., and Macara, I.G. (1990). A cytosolic protein catalyzes the release of GDP from p21ras. *Science* *248*, 67-69.
- Yaman, E., Gasper, R., Koerner, C., Wittinghofer, A., and Tazebay, U.H. (2009). RasGEF1A and RasGEF1B are guanine nucleotide exchange factors that discriminate between Rap GTP-binding proteins and mediate Rap2-specific nucleotide exchange. *The FEBS journal* *276*, 4607-4616.
- Yang, Y., Li, L., Wong, G.W., Krilis, S.A., Madhusudhan, M.S., Sali, A., and Stevens, R.L. (2002). RasGRP4, a new mast cell-restricted Ras guanine nucleotide-releasing protein with calcium- and diacylglycerol-binding motifs. Identification of defective variants of this signaling protein in asthma, mastocytosis, and mast cell leukemia patients and demonstration of the importance of RasGRP4 in mast cell development and function. *The Journal of biological chemistry* *277*, 25756-25774.

Yurochko, A.D., Hwang, E.S., Rasmussen, L., Keay, S., Pereira, L., and Huang, E.S. (1997). The human cytomegalovirus UL55 (gB) and UL75 (gH) glycoprotein ligands initiate the rapid activation of Sp1 and NF-kappaB during infection. *J Virol* *71*, 5051-5059.

Zarrei, M., Merico, D., Kellam, B., Engchuan, W., Scriver, T., Jokhan, R., Wilson, M.D., Parr, J., Lemire, E.G., Stavropoulos, D.J., *et al.* (2017). A de novo deletion in a boy with cerebral palsy suggests a refined critical region for the 4q21.22 microdeletion syndrome. *173*, 1287-1293.

Zhang, Q., Lenardo, M.J., and Baltimore, D. (2017). 30 Years of NF-kappaB: A Blossoming of Relevance to Human Pathobiology. *Cell* *168*, 37-57.

Zheng, Y., and Quilliam, L.A. (2003). Activation of the Ras superfamily of small GTPases. Workshop on exchange factors. *EMBO Rep* *4*, 463-468.