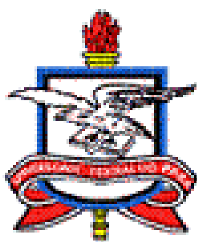
Universidade Federal do Pará
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética e Biologia Molecular

BIOLOGIA MOLECULAR APLICADA À HANSENÍASE: ESTUDO DE PARÂMETROS GENÉTICOS E EPIGENÉTICOS EM UMA AMOSTRA DO ESTADO DO PARÁ.

Pablo Diego do Carmo Pinto

Belém/PA
2016



Universidade Federal do Pará
Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética e Biologia Molecular

BIOLOGIA MOLECULAR APLICADA À HANSENÍASE: ESTUDO DE PARÂMETROS GENÉTICOS E EPIGENÉTICOS EM UMA AMOSTRA DO ESTADO DO PARÁ.

Pablo Diego do Carmo Pinto

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFPA como requisito para obtenção do grau de Doutor em Genética e Biologia Molecular.

Orientadora: Prof^a. Dr^a. Ândrea Kely Campos Ribeiro dos Santos

Belém/PA
2016

DEDICATÓRIA

Aproveito este espaço para fazer uma dedicatória a todas as pessoas que se fizeram importantes nestes momentos e que contribuíram para esta realização.

Dedico esta Disertação a minha mãe **Edna Martins do Carmo**, ao meu pai **Alvaro Pinto Filho,** ao meu irmão **Danilo Lucas Martins do Carmo** e a minha namorada **Carolina Veloso Barroso** por todos esses anos de convivência, de aprendizagem, de lutas e derrotas, mas que serviram de ponte para que fosse construído uma passagem para o sucesso que traz consigo as vitórias, os encantos, as alegrias e acima de tudo o poder de vive-las da forma mais intensa possível, ao lado de quem se ama.

Tenho um grande prazer ao dedicar esta tese também à minha professora e amiga **Ândrea Kely Campos Ribeiro dos Santos**, por ter me aceito como aluno e por ter me aconselhado e indicado os caminhos da ciência e da pesquisa com toda audácia, vigor, ética e coerência que lhe são inerentes, sendo assim agradeço do fundo do meu coração por ter aceito ser minha mãe cientifica.

Agradeço enormemente, ainda, a **Drª Ândrea Kely Campos Ribeiro dos Santos** como orientadora, por ter aceitado a tarefa, não só de me orientar, mas também de me ajudar a trilhar nos caminhos da pesquisa científica com conselhos, questões e dicas valorosas. O seu aceite à minha orientação muito me honra, e sua protagonização em minha vida científica será lembrada com louvor e muita felicidade.

AGRADECIMENTOS

Agradeço a todos os colegas que sempre estiveram a disposição para ajudar e contribuir na minha formação acadêmica para que este resultado fosse alcançado, aos eternos amigos do LGHM (Laboratório de Genética Humana e Médica) – Marcos Amador, Milene Raiol, Dayse Alencar, Antonette El Husny, Rafael Resque, Igor Hamoy, Isabella Guerreiro, Leandro Magalhães, Amanda Vidal, Roberta Borges, André Santos, Deborá Ricardo, Jaime Vieira, Fabiano Cordeiro, Ana Paula Schann, Glória Tatiana, Camile Lopes e tantos outros não citados que sempre me auxiliaram nos processos mais difíceis do trabalho, com boas conversas, piadas, auxílios e conselhos. Além disso, deixo um agradecimento especial ao colega Vinícius Albuquerque Sortica pelo auxilio laboratorial e analise de dados, essencial para a conclusão desta Tese.

Agradeço também ao laboratório de Dermatoimunologia, e as grandes contribuições do Prof. Dr. Claudio Guedes Salgado, a quem devoto grande admiração, pelas lições, dicas e auxilio que muito me serviram e servirão durante toda uma jornada que pretendo seguir.

Por fim, agradeço as agencias de fomento CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) e CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) por financiar o projeto e permitir que o mesmo fosse realizado, assim como pela UFPA (Universidade Federal do Pará) onde pude cursar meus estudos de graduação e Pós-graduação, que me traz tanto orgulho e vontade de continuar.

EPÍGRAFE

Se a humanidade é feita de circunstâncias,

então que façamos as circunstâncias humanamente.

(KARL MARX E FRIEDRICH ENGELS)

RESUMO

A hanseníase é causada pelo *Mycobacterium leprae* e os indivíduos acometidos pela hanseníase podem ser classificados, em Paucibacilares e Multibacilares. Alternativamente, segundo Ridley-Jopling (1966), com base em critérios clínicos e imuno-hitológicos em outros dois pólos: (i) o pólo Tuberculóide (TT); e (ii) pólo Lepromatoso (LL), e seus intermediários. Independente de sua classificação, este espectro parecer ser inflluenciado por moléculas moduladoras da resposta imune, como os genes que codificam estes mediadores, e por um grupo de pequenos RNAs (microRNAs) que são responsáveis pela regulação destes genes, portanto essas investigações podem adensar o conhecimento sobre o mecanismo de resposta ao processo infecsioso, assim como possibilitar a identificação de novos biomarcadores no auxilio ao diagnóstico da hanseníase. O objetivo foi investigar oito polimorfismos do tipo INDEL nos genes *CYP19A1*, *NFKβ1*, *IL1α*, *CASP8*, *UGT1A1*, *PAR1*, *CYP2E1*, e *IL4*, para identificar possiveis marcadores de susceptibilidade e a influência da ancestria genética neste risco, além disso foi realizado o primeiro miRnoma da hanseníase por sequenciamento massivo em plataforma de alto desempenho, afim de elucidar o perfil epigenético presente na hanseníase. Nosso estudo revelou que os genes *NFKβ1*, *CASP8*, *PAR1* e *IL4*, são potenciais marcadores de susceptibilidade para a hanseníase, enquanto que *NFKβ1*, *CASP8*, *PAR1* e *CYP19A1* são potenciais marcadores da forma clínica multibacilar. Adicionalmente, a análise da ancestralidade genômica mostrou que a contribuição Européia elevou o risco ao desenvolvimento da doença, enquanto a contribuição Africana aumentou proteção. No que diz respeito a análise diferencial do perfil de expressão dos microRNAs de pacientes com hanseníase, por meio da análise de biopsias de pele, revelaram-se 67 miRNAs diferencialmente expressos, dos quais 43 apresentavam um padrão de expressão *downregulated* e 24 *upregulated.* Quando analisamos amostras de sangue desses mesmos pacientes, observaram-se 10 miRNAs

diferencialmente expressos, dos quais 9 com padrão de expressão *downregulated* e 1 *upregulated*. Os alvos pesquisados, em análise *in silico,* a partir desses resultados sugeriu os genes *(IL1β, IL6, IL8, IL12, TLR2, TLR4, IL17RB, IFNGR1, TGFBR1, NFκβ, família SMAD, STAT3, CASP8, CYP19A1, BCL-2, entre outros)* como envolvidos na patologia da hanseníase. Por fim, monstrou-se pela primeira vez o perfil de microRNAs em *genome wide* da Hanseníase.

ABSTRACT

Leprosy is caused by *Mycobacterium leprae* and patients can be grouped in Paucibacillary and Multibacillary. Alternatively, according by Ridley-Jopling (1966), using immune-hystogical criteria, grouped in two distinct pole: (i) Tuberculóide (TT); and (ii) lepromatous (LL), and your intermediaries. Independently these classification, the disease can be affected by molecules that modulates immune response, like genes that encode these molecules, and by small RNA (micro-RNA), wich regulated these genes, thus these study can improve the knowledge about the mechanism of response to infectious process, as well as enable the identification of new possibles biomarkers to assist diagnosis in leprosy. The objective of this study was to investigate eight INDEL polymorphisms on genes *CYP19A1*, *NFKβ1*, *IL1α*, *CASP8*, *UGT1A1*, *PAR1*, *CYP2E1*, and *IL4*, to identify possible susceptibility markers of leprosy and evaluate the influence of genetic ancestry on disease risk. Besides was performed the first genome wide miRNA profiling of Leprosy by next generation sequencing (NGS), assessing and describing the expression standard in leprosy. Our study shows that the *NFKβ1, CASP8, PAR1, IL4* and CYP19A1 genes are possible markers for the susceptibility to development of leprosy and the severe clinical form MB. Moreover, after correcting for population structure within an admixture population, the results show that different levels of ethnic group composition can generate different OR rates for leprosy susceptibility. The differential expression profile from tissue samples reveal 67 miRNAs differentially expression, with 43 *down* and 24 *upregulated* and from blood sample were found a total of 10 miRNAs differentially expression with 9 *down* and one *upregulated*. Moreover was performed in silico target analysis and detect the genes (*IL1β, IL6, IL8, IL12, TLR2, TLR4, IL17RB, IFNGR1, TGFBR1, NFκβ, família SMAD, STAT3, CASP8, CYP19A1, BCL-2, in others*) involved on pathological of leprosy. Lastly, was showed for the first time the *genome wide* microRNA of leprosy.

SUMÁRIO

RESUMO

ABSTRACT

1 - INTRODUÇÃO 1

1.1 Epidemiologia da Hanseníase 4

1.2 Biologia do Agente 10

1.3 Resposta Imunológica 11

1.4 Aspectos Genéticos 13

1.4.1 Biomarcadores gênicos relacionados à Hanseníase 14

1.5 RNA Reguladores 18

1.5.1 Micro-RNA (*mi*RNA) 18

1.5.2 Biogênese dos *mi*RNA's 19

1.5.3 Mecanismo de regulação da expressão por *mi*RNA 23

1.6 Biomarcadore epigenéticos associados à Hanseníase 24

2 - OBJETIVOS 28

2.1 Objetivo Geral 28

2.2 Objetivos Específicos 28

3 – CAPÍTULO 1 - INFLUENCE OF GENETIC ANCESTRY ON INDEL MARKERS OF NFKB1, CASP8, PAR1, IL4 AND CYP19A1 GENES IN LEPROSY PATIENTS 29

4 – CAPÍTULO 2 - GENOME WIDE MICRORNAS IN LEPROSY PATIENTS BY NEXT GENERATION SEQUENCING 45

5- DISCUSSÃO GERAL 77

6- Referência 84

7- ANEXO – OUTRA PUBLICAÇÃO NO PERIODO 96

1 - INTRODUÇÃO

Cerca de trinta e cinco mil pessoas morrem diariamente acometidas por doenças infectocontagiosas, como a hanseníase, que afetam desproporcionalmente as populações pobres da América Latina, África e Ásia. Essas doenças são comumente referidas como "doenças negligenciadas", termo cunhado pela Organização Mundial da Saúde, porque de uma forma geral foram esquecidas pela comunidade científica global, governos e indústria farmacêutica (Torreele, 2006; Suzuki *et al.*, 2012).

A hanseníase é uma doença infectocontagiosa, de progressão lenta, que se manifesta principalmente por meio de lesões cutâneas ocasianada pelo *Mycobacterium leprae,* uma bactéria intracelular obrigatória (Dickson *et al.,* 2011). Esta patologia se apresenta como uma infecção crônica, com tropismo por macrófagos e células de Schwann de nervos periféricos (onde a bactéria pode causar alterações dermatológicas), com evolução a um quadro de deficiência sensitiva e progressivo comprometimento motor (Alcais *et al*., 2005) que estigmatiza socialmente os grupos de indivíduos que manifestam esta doença.

A manifestação clínica da hanseníase origina uma patologia com características imunológicas variáveis, das quais emergem o aparecimento de dois espectros distintos da mesma doença, o pólo Tuberculóide e o Lepromatoso. O pólo Tuberculóide (TT) envolve uma intensa resposta imune celular realizada por células com atividade fagocitárias, como macrófagos e células NK (Natural Killer), além da presença de citocinas inflamatórias secretadas por linfócitos $TCD4^{+}$. Adicionalmente, constitue um tipo de resposta imune Th1, eficiente no combate ao crescimento e proliferação do *M. leprae,* que auxilia a manutenção de um baixo índice bacilar (IB)[*1], e melhora o

prognóstico dos pacientes, diminuindo as chances de lesões neurológicas (Elizabeth *et al.*, 2010; Araújo, 2003).

O pólo Lepromatoso (LL) ocasiona mudança do padrão de resposta inflamatória imune celular para um padrão de resposta humoral (Th1 $\rightarrow$ Th2). Este padrão é formado por linfócitos B, os quais passam a produzir anticorpos específicos para o *M. leprae*, como o anti-PGL1 (Lipídio Glico-fenólico 1), além de ser mediado por citocinas anti-inflamatórias secretadas por linfócitos $TCD4^+$. Constitue o tipo de resposta imune ineficiente no combate ao crescimento e proliferação do *M. leprae*, uma vez que o bacilo se encontra na região intracelular. Este padrão imunológico gera Índice Bacilar (IB) elevado que dificulta o tratamento e aumenta a probabilidade de ocorrência de deficiências neurológicas (Figura 1) (Alcais *et al.*, 2005].

Por outro lado, muitos pacientes apresentam características intermediárias entre estes dois pólos. O pólo indeterminado (I), como definidos pela classificação de Madrid, que poderá progredir para os pólos TT ou LL da doença, ou as categorias intermediárias, de acordo com classificação de Ridley-Jopling: Boderline-Tuberculóide (TB), Boderline-Boderline (BB) e Boderline-Lepromatoso (BL) (Elizabeth *et al.*, 2010; Araújo, 2003; Ridley and Jopling, 1966].

Recentes estudos de células T em hanseníase, revelaram que as células Th17 e Foxp3, como envolvidas na patogênese da doença. O subconjunto Th17 produz IL-17, descritas em reações de ENH (Eritema nodoso hansênico) (Martiniuk, 2012). Segundo Saini *et al* (2013), as células Th17 podem constituir um terceiro tipo de perfil Th na hanseníase, e pode ser uma via alternativa em pacientes incapazes de montar uma resposta Th definida (Saini *et al.,* 2015). As células Foxp3 ou Th3 ($CD4^+$, $CD25^+$, $Foxp3^+$), são induzidas pelo aumento de TGFβ e regulam negativamente as células Th17, apresentando função supressora (Kumar, 2013). Acredita-se que estas células T

reguladoras podem ser ativadas na presença de uma grande quantidade de IL-4, como na forma LL, induzindo a produção de TGF-β1 e IL-10, contribuindo com a persistência do parasito e a indução da cronicidade da doença (Goulart *et al.*, 2002; Venturini, 2011).

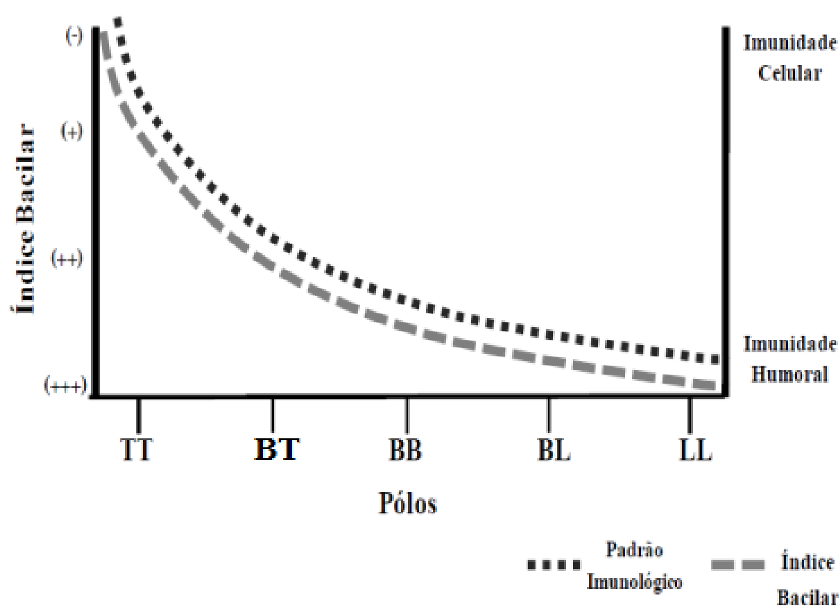


Figura 1: Espectro clínico da hanseníase: O padrão Imunológico, medido pelo teste de Mitsuda, e o Ìndice Bacilar, medido pelo índice baciloscópico (IB), entre os distintos pólos da hanseníase, segundo a classificação de Ridley & Jopling.

O padrão epidemiológico das doenças infecciosas é multifatorial e é modulado por componentes intrínsecos e extrínsecos ao indivíduo, envolvendo relações entre os dois genomas, o do hospedeiro e do *M. leprae* . Os fatores intrínsecos são os biológicos por natureza, tais como a composição genética e a resposta imune (Suzuki *et al.*, 2012). Esses fatores geralmente associados aos extrínsecos, tais como a falta de saneamento básico, falta de acesso aos serviços de saúde (presentes principalmente em grupos sociais mais empobrecidos), diferentes atividades humanas, e o auto-tratamento inadequado, podem determinar o padrão de prevalência das doenças negligenciadas. Estes fatores mostram que as políticas públicas para a eliminação dessas doenças,

devem necessariamente apresentar uma abragência multiprofissional para que o processo seja bem sucedido (Ehrenberg and Ault, 2005).

1.1 Epidemiologia da Hanseníase

Durante meados da década de 1990, a assembléia mundial da saúde aprovou uma ousada e ambiciosa campanha, que propunha a eliminação da hanseníase com o slogan "*eliminação da hanseníase como um problema de saúde pública até o ano 2000*". A eliminação foi definida como uma redução na prevalência de pacientes com hanseníase, através do tratamento com a poliquimioterapia (PQT) antimicrobiana, para menos de 1 por 10.000 habitantes (World Health Assembly, 1991). A justificativa para esta definição reside no fato de que o tratamento com PQT seria altamente eficaz, e portanto, a incidência de hanseníase seria reduzida, o que geraria o desaparecimento gradual da doença (Suzuki *et al.*, 2012).

Na contra-mão ao combate a hanseníase, apesar de todos os esforços da comunidade internacional, 15 países endêmicos ainda apresentam uma prevalência de mais de um indivíduo por 10.000 habitantes, principalmente nos continentes da Ásia, África e América do Sul. Por outro lado, 116 dos 122 países endêmicos para a hanseníase, identificados em 1985, atingiram o controle da doença (Dickson *et al.,* 2011). Há ainda uma elevada taxa endêmica em algumas áreas da Índia, Brasil, Indonésia, Bangladesh, República Democrática do Congo, Etiópia, Nepal, Nigéria, Myanmar, República da Tanzânia, Sudan, Sri Lanka, Filipinas, China, Madagascar e Mozambique (Figura 2), com uma contribuição de 93% do total de casos de hanseníase registrados no mundo em 2009, com as taxas de detecção por 10.000 habitantes mais elevadas: India (133.717 casos), Brasil (37.610 casos), Indonésia (17.260 casos), Bangladesh (5.239 casos), República Democrática do Congo (5.062 casos), Etiópia

(4.417 casos), Nepal (4.394 casos), Nigéria (4.219 casos), Myanmar (3.147 casos), Republica da Tanzania (2.654 casos), Sudan (2.100 casos), Sri Lanka (1.875 casos Filipinas (1.795 casos), China (1.597 casos), Madagascar (1.572 casos) e Mozambique (1.191 casos) (Dickson *et al.,* 2011 and Suzuki *et al.*, 2012).

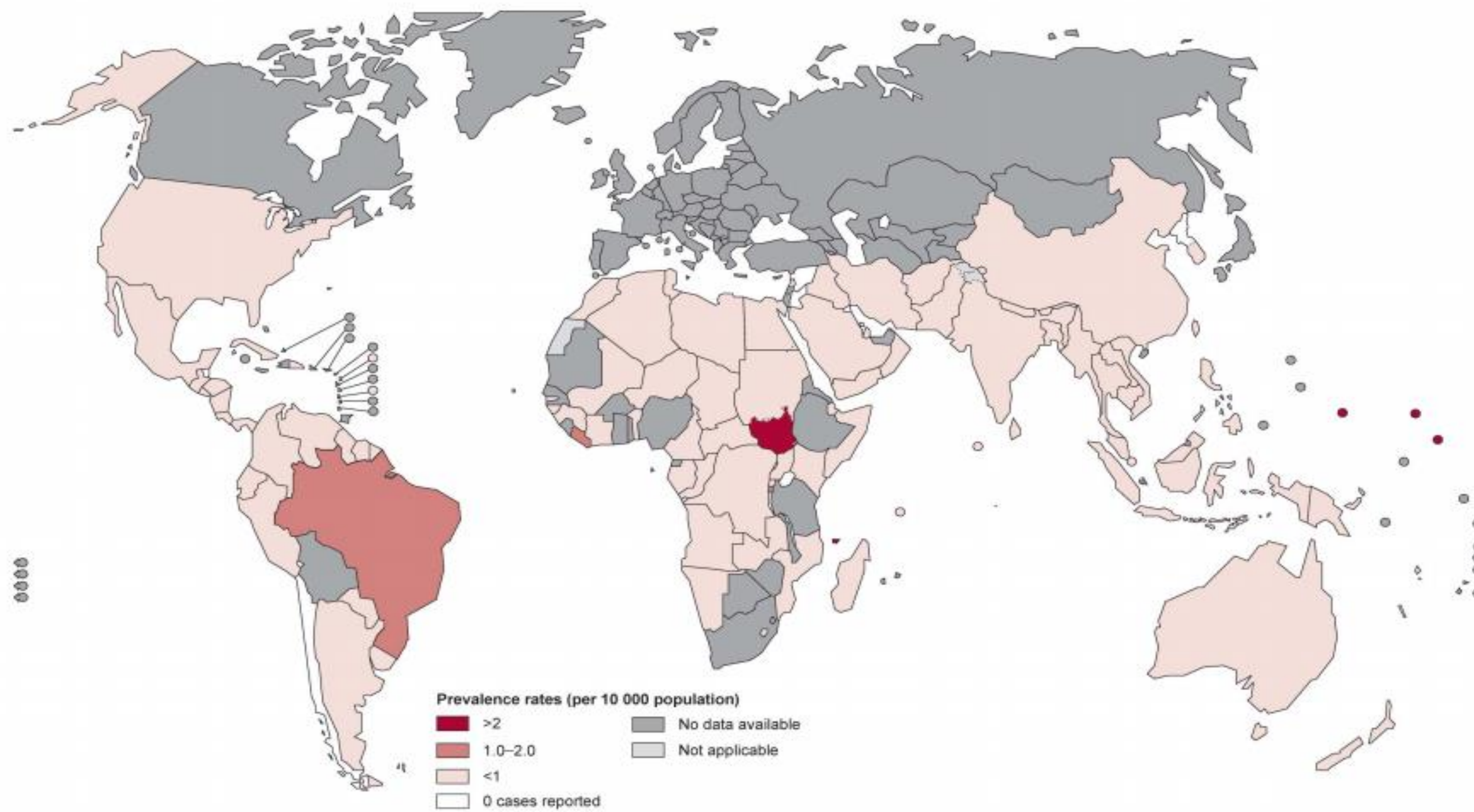
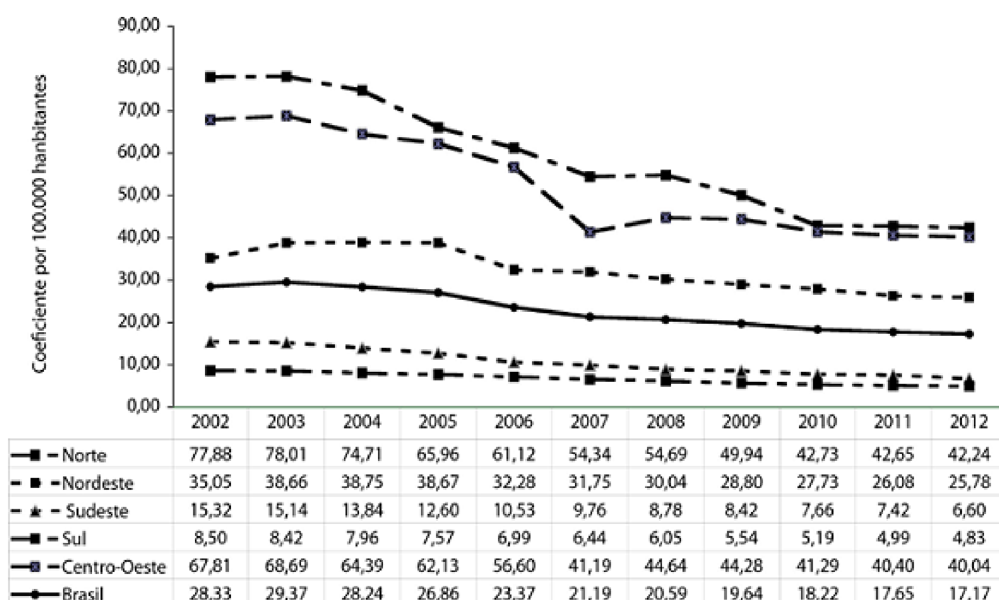


Figura 2: Taxas de prevalência da Hanseníase no mundo no ano de 2011; registro de caso/10.000 Habitantes. Fonte: WHO, 2011. Disponível: Janeiro/2012.

A Figura 3 apresenta um gráfico com a evolução do coeficiente de detecção geral de casos novos no Brasil e regiões entre os anos de 2003 a 2012. Observa-se, no período, uma diminuição do coeficiente de detecção de casos em todas as regiões geográficas, inclusive na região Sul, que historicamente apresenta os menores coeficientes. Os maiores coeficientes de detecção por região geográfica foram verificados nas regiões Norte e Centro-Oeste, com respectivamente 42,3 e 40,1 casos novos por 100 mil habitantes em 2012, o que caracteriza hiperendemicidade (Ministério da Saúde, 2012).



	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Norte	77,88	78,01	74,71	65,96	61,12	54,34	54,69	49,94	42,73	42,65	42,24
Nordeste	35,05	38,66	38,75	38,67	32,28	31,75	30,04	28,80	27,73	26,08	25,78
Sudeste	15,32	15,14	13,84	12,60	10,53	9,76	8,78	8,42	7,66	7,42	6,60
Sul	8,50	8,42	7,96	7,57	6,99	6,44	6,05	5,54	5,19	4,99	4,83
Centro-Oeste	67,81	68,69	64,39	62,13	56,60	41,19	44,64	44,28	41,29	40,40	40,04
Brasil	28,33	29,37	28,24	26,86	23,37	21,19	20,59	19,64	18,22	17,65	17,17

Figura 3: Coeficientes de detecção de casos novos de hanseníase por 100.000 habitantes, regiões e Brasil,2002/2012.
Fonte: SINAN/SVS-MS. Disponível: Abril/2013.

A Figura 4, mostra um gráfico com os coeficientes de detecção de casos novos registrados nos estados em 2012. Evidencia-se o comprometimento da região da Amazônia Legal em relação à hanseníase, que comporta uma população correspondente, em 2012, a 17.5% da população do Brasil, porém concentra um coeficiente de detecção de casos novos de 42,24/100.000 habitantes, ou seja, a região com menor contribuição populacional detém a maior taxa de detecção do país. A Amazônia Legal apresenta barreiras físicas e sociais que dificultam o acesso aos serviços de saúde, e tem aspectos demográficos e ambientais que a tornam historicamente vinculada à evolução da endemia no Brasil (Ministério da Saúde - coordenação geral de hanseníase e doenças em eliminação, 2012).

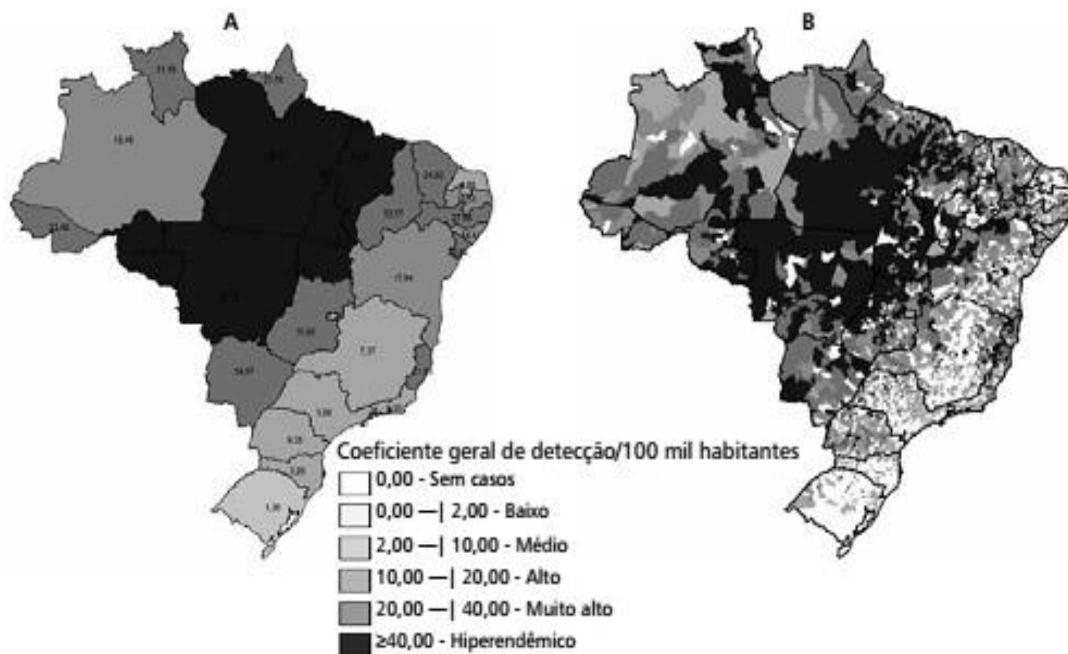


Figura 4: Coeficiente de detecção de casos novos de hanseníase por 100.000 habitantes,(A) estados da Federação, e (B) municipios Brasil, 2012.
Fonte: SINAN/SVS-MS. Disponível: Abril/2013.

Em particular no Estado do Pará, tem-se registrado avanços importantes no combate à hanseníase, devido a busca ativa e tratamento dos pacientes em regiões endêmicas e em regiões ao redor destas. Entretanto, em 2010 um total de 34.894 novos casos foram registrados no Brasil correspondendo a uma taxa de incidência de 18,22 casos por 100.000 habitantes. Neste cenário o Pará foi responsável por 10,2% dos casos (3.562 casos), uma taxa de incidência de 46,93 por 100.000 habitantes. Considerando apenas crianças menores de 15 anos, o Estado registrou 389 novos casos de hanseníase em 2010, representando 10,9% de todos os casos, uma taxa de incidência 16,52 por 100.000 habitantes, indicando que o *Mycobacterium leprae* está circulando entre as crianças, consideradas como focos ativos da doença (Salgado *et al*., 2012).

Estes dados são agravados pelo pouco acesso ao sistema público de saúde, uma vez que no Pará somente 42% da população total é atingida por estes serviços, o que dificulta a periodicidade do monitoramento clínico e sugere a existência de muitos pacientes com hanseníase não diagnosticados no Pará. Além disso, um estudo de 1,592 estudantes entre 09 a 15 anos no estado do Pará diagnosticou 63 novos casos de hanseníase (4%). Considerando que o Pará tinha no período do estudo, dois milhões de estudantes matriculados nas escolas públicas, e se extrapolarmos os dados, estamos falando de aproximadamente 80.000 estudantes não diagnosticados no Estado, o que pode aumentar a incidência de hanseníase nas próximas décadas (Salgado *et al*., 2012; Barreto *et al*., 2012).

Em muitas dessas áreas, 18% dos casos novos são diagnosticados com elevado comprometimento do nervo motor, o que gera incapacidades físicas, e demonstra o diagnóstico tardio. Este fato acarreta um custo social muito elevado para o Estado e a população atingida. A hanseníase constitui, para o Pará, um grave problema de saúde

pública, mesmo com todo o trabalho que vem sendo realizado em prol do controle dessa doença (Ministério da Saúde - coordenação geral de hanseníase e doenças em eliminação, 2012).

1.2 Biologia do Agente

O Mycobacterium leprae é um bacilo de BAAR (bacilo álcool-acido resistente) e parasita intracelular obrigatório (Suzuki *et al.,* 2012). O genoma do *M. leprae* inclui 1.605 genes que codificam proteínas e 50 genes para moléculas estáveis de RNA. Mais da metade dos genes funcionais do genoma do *Mycobacterium tuberculosis* estão ausentes, ou são pseudogenes no *M. leprae*, o bacilo parece ter excluído genes normalmente exigidos para a replicação fora de organismos vivos, assumindo assim um nicho ecológico com hospedeiros e a necessidade de crescimento dentro de suas células (Cole *et al., 2001*).

A perda de material genético pode ter afetado vários genes, particularmente aqueles envolvidos no catabolismo, porém os genes essenciais para a formação da parede celular de micobactérias permaneceram inalterados. Este fato pode colaborar para que o *M. leprae,* tenha tornado-se dependente de produtos metabólicos das células que parasita, isto poderia explicar o longo período de duração da infecção e a incapacidade de crescer em meio de cultura. Contudo, a perda de variabilidade genética do *M. leprae* em comparação ao *M. tuberculosis* não parece influenciar em sua virulência (Brennan *et al.,* 2001; Laura *et al.,* 2011).

A parede celular de micobactérias contém alvos importantes da resposta imune do hospedeiro, que incluem o glicolipido fenólico I (PGL-I), que estimula uma potente resposta de anticorpos IgM, que é proporcional à carga bacilar dos pacientes infectados,

e tende a decrescer com a terapia (Thole *et al.,* 1999). Alguns genes codificadores de vários antígenos foram identificados, incluindo antígenos compartilhados com *Mycobacterium tuberculosis* e com outras micobacterias ambientais (Roche *et al.,* 1992; Suzuki *et al.,* 2012).

A hanseníase é uma doença infecciosa de curso crônico e de baixa patogenicidade, apresenta um quadro clínico variável, que depende basicamente do padrão de resposta imunológica individual. A transmissão da doença ainda não é clara, mas se acredita que ocorra pelo contato pessoal, preferencialmente prolongado, com maior risco para os contatos intradomiciliares de pacientes hansênicos (Britton *et al.,* 2004).

1.3 Resposta Imunológica

Segundo Siddiqui *et al.* (*2001*), fatores genéticos do hospedeiro assumem efeitos parciais no desenvolvimento da hanseníase. O sequenciamento do genoma humano identificou locos relacionados com a susceptibilidade `a hanseníase, no cromossomo 6 na região do MHC (Major histocompatibility complex), assim como vários outros genes envolvidos na regulação da resposta imune, tais como os codificantes de interleucinas, e alguns receptores essenciais à função imunológica como o transportador associado ao processamento de antigeno (TAP) que é responsável pelo transporte de moléculas antigênicas para o MHC de classe I no citoplasma das células imunológicas (Shaw *et al.,* 2001).

Os linfócitos T com fenótipo $CD4^{+}$, possuem atividades imuno-reguladoras específicas mediadas por diferentes interleucinas, as quais podem gerar diferenciação em subpopulações como Th1, Th2, Th17, Th3, Th9, correlacionadas com a

susceptibilidade a diferentes tipos de doenças. A subpopulação Th1 produz um padrão de resposta imune mediado principalmente por IL-2 (Interleucina 2), IFN-γ (Interferon gamma) TNF-α (Fator de necrose Tumoral), IL-12 (Interleucina 12) e IL1β (Interleucina 1 beta), as quais ativam a resposta imune celular. A IL-2 atua em receptores dos linfócitos CD4+, e mantém estável ritmo de produção destas citocinas (IL-2, IFN-γ e TNF-α), além de estimular células NK a produzirem IFN-γ. O interferon atua diretamente sobre macrófagos e estimula mecanismos que promovem a fagocitose, além da produção de TNF-α e IL12 (Figura 5). Esse mecanismo integrado, determina uma resposta celular eficiente no combate e eliminação do bacilo, e ocasiona o padrão de resposta imune observado na maioria dos pacientes do pólo Tuberculóide (TT) (Foss *et al.,* 1997, Britton *et al.,* 2004, Scollard *et al.,* 2006, Elizabeth *et al.,* 2010].

De maneira distinta, a subpopulação Th2 produz um padrão de resposta imune mediado principalmente por IL-4 (Interleucina 4), IL-5 (Interleucina 5), IL-6 (Interleucina 6), IL-8 (Interleucina 8), IL-10 (Interleucina 10), e VDR (receptor de vitamina D). Estes mediadores suprimem a atividade do macrófago, e células NK com consequente desvio do padrão de resposta imunológica de celular para humoral (Figura 5). Adicionalmente, a IL-4 estimula linfócitos B, a produzirem imunoglobulinas e a se diferenciarem em mastócitos para produzir mais IL-4, esse mecanismo imunorregulador irá aumentar a resposta supressora macrofágica, que consequentemente ativa a resposta humoral ineficaz para a eliminação dos bacilos, como observado na maioria dos pacientes do pólo lepromatoso (LL] (Foss *et al.,* 1997; Salgame *et al.,* 1992; Yamamura *et al*., 1992; Elizabeth *et al*., 2010).

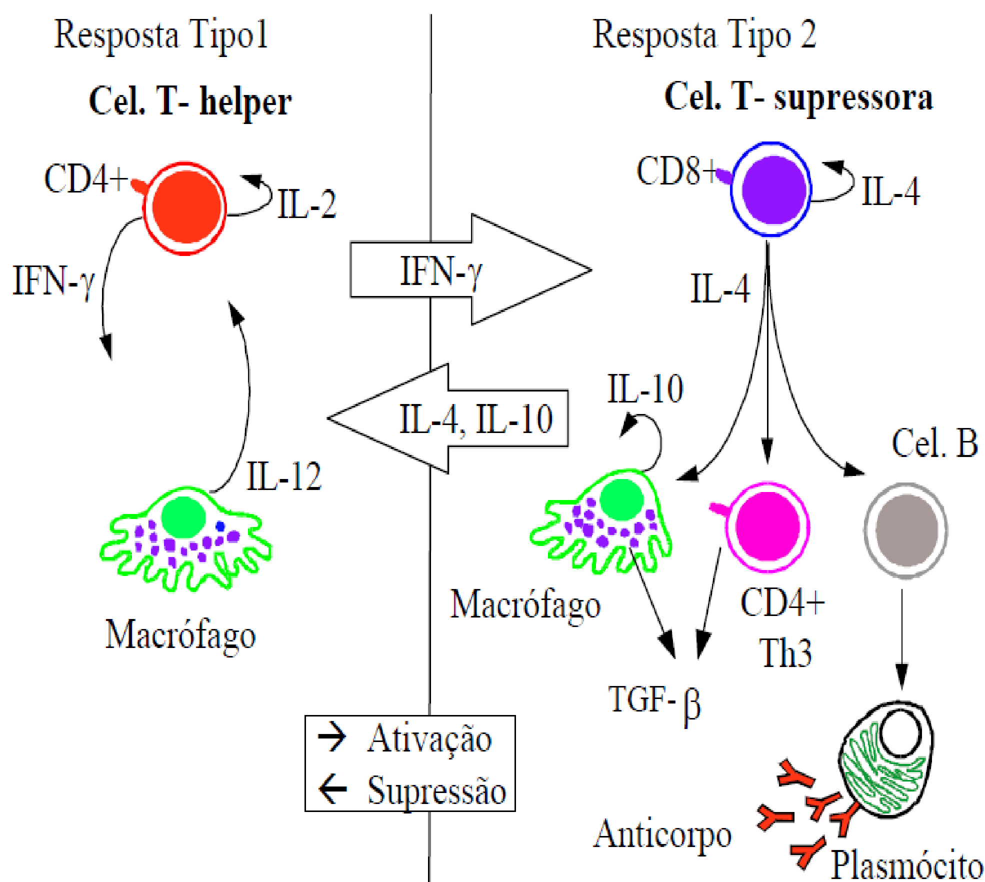


Figura 5: Mecanismo de regulação imunológica, mediada por interleucinas, as quais influenciam o balanço entre os dois principais tipos resposta imune, envolvido na susceptibilidade á hanseníase, o tipo 1 (Th1) e o tipo 2 (Th2). Fonte: Goulart et al., 2002.

1.4 Aspectos Genéticos

Os mecanismos imunológicos mediados por citocinas envolvem o sensível balanço entre a resposta immune inflamatória (Th1), mediada principalmente por IL-12, IFN-γ, IL-1β, e TNF-α, e anti-inflamatória (Th2), mediada principalmente por IL-6, IL-10, IL-4 e IL-8 (Stretch *et al.*, 1999). O entendimento deste sensível balanço e de seus

fatores regulatórios imunológicos podem elucidar o processo de manifestação da doença, susceptibilidade e estabelecimento do polo BB ou LL. Adicionalmente, estes mecanismos regulatórios estão relacionados com os processos pelos quais o hospedeiro pode oferecer resistência ao crescimento e proliferação do *M. leprae* (Elizabeth *et al.,* 2010).

1.4.1 Biomarcadores gênicos relacionados à Hanseníase

Um dos principais fatores de sinalização ativado por danos teciduais e patógenos microbianos é o fator nuclear kB (NFkB), o qual realiza um papel central na indução transcricional de genes da resposta imune inata, controlando a expressão de uma rede de indutores e efetores que irão definir as respostas aos patógenos via sinalização intracelular. Neste sentido, a capacidade de regular a expressão gênica como fator de proteção, possibilita a adaptabilidade e sobrevivência dos organismos multicelulares frente as infecções. O NFkB é, principalmente, ativado após infecção por microorganismos via sinalização de vários receptores como TLRs (receptores Toll-like), TNFR (receptor do fator de necrose tumoral), IL1R (receptor de interleucina 1) entre outros, os quais são expressos por macrófagos, células dendriticas, e células epiteliais de mucosa (Mohamed and McFadden, 2009)

O NFkB é uma família de fatores de transcrição diméricos, os quais compreendem cinco proteínas RelA(p65), RelB, c-Rel, NFkB1(p50) e NFkB2(p52), que se associam aos pares pelo domínio homologo Rel n-terminal (RHDs), para formar os fatores de transcrição que irão ativar regiões diferentes do genoma (Mohamed and McFadden, 2009). Em estágios celulares não sensibilizados, a proteína IkBα (inibidor da atividade de NFkB) sequestra NFkB no citoplasma e evita que haja a ativação

transcricional, porém em resposta a estímulos específicos (via receptores TLR, TNFR etc), IkBα é ubquitinado e degradado o que libera o NFkB para migrar para o núcleo e estimular a transcrição de genes pró-inflamatórios (Ali *et al.,* 2013; Andersen *et al.,* 2011). Neste contexto, estudos genéticos de associação apontam que o alelo deleção (DEL, rs28362491) do gene *NFkB* dimune a atividade transcricional da enzima, e por conseguinte a transcrição de uma variedade de genes da imunidade inata inflamatória (Karban *et al.,* 2004; Salim *et al.,* 2013; Cen *et al.,* 2013).

Receptores ativados por proteases (PAR) podem ser ativados por proteases plasmáticas circulantes comuns em processos infecciosos, e contém um mecanismo molecular de ativação que inclui clivagem proteolítica irreversível de sua porção N-terminal, o qual o torna um ligante ativo. Esses receptores compreendem uma família com quatro classes (PAR1, PAR2, PAR3, PAR4) transcritos pelos genes *PAR1*, *PAR2*, *PAR3* e *PAR4*, e a excessção de *PAR2*, todos os outros são ativados por trombina, uma protease presente no plasma em estágios de infecção (Aerts *et al.,* 2013).

Os receptores PAR de ação ubiquitinante estão presentes na superfície de plaquetas, células endoteliais e epiteliais, neurônios, fibroblastos, músculo esquelético, leucócitos e células tumorais (Adams *et al.,* 2011). Estes receptores estão envolvidos em alguns processos fisiológicos como função cardiovascular, respiratória, sistema nervoso central, câncer e inflamação (Aerts *et al.,* 2013). PAR1 suprime a atividade de células T auxiliares do tipo 1 (Th1) e do tipo 17 (Th17), e por conseguinte a secreção de IL-12 e IL-13 o que resulta em uma inibição da atividade celular inflamatória, importante para o combate ao crescimento do bacilo (Chionh *et al.,* 2014). O alelo de inserção (INS) do INDEL (rs11267092) do gene *PAR1* foi relatado como responsável por aumentar o nível de transcrição (Arnald *et al*., 2000), e assim pode representar um fator de risco ao desenvlvimento da hanseníase.

A família das caspases (*cysteine-aspartic-acid-proteases*) são fisiologicamente envolvidas na regulação da apoptose celular, e tem duas funções básicas: caspases-1,-4,-5 e -11, como caspases iniciadoras, estão envolvidas no processessamento e ativação de citocinas pro-inflamatórias. Por outro lado, caspases -2, -3, -6, -7, -8 e -9, como caspases executoras, desenvolvem um papel direto na execução da apoptose. Neste último caso, há dois caminhos metabólicos distintos, porém complementares: i) o caminho extrínseco ou mediado por receptor, e ii) o caminho intrínseco ou mitocondrial (Chen *et al.,* 2012).

Macrófagos com alta carga bacilar de *M. leprae* são submetidos à apoptose, e este mecanismos está sob controle de citocinas inflamatórias (Klingler *et al*., 1997). Em pacientes com hanseníase, a contínua ativação de células T por antígenos de *M. leprae*, induz a apoptose e redução de linfócitos em tecidos periféricos e outras células, nestes casos a regulação da apoptose envolve estimulação e ativação de caspase-8, codificado pelo gene *CASP8* (Chattree *et al*., 2008). O alelo DEL (rs3834129) do gene *CASP8* causa diminuição do nível de transcrição e está associado a redução da apoptose induzida por caspase-8 (Sun *et al.,* 2007).

A interleucina 4 (IL-4) estimula o padrão de resposta Th2 ao promover a diferenciação e proliferação de linfócitos B e elevar o nível de produção de IgE, através da ligação desta citocina ao seu receptor (IL-4R), que é uma proteína transmembranar codificada pelo gene *IL4R*, que contém 12 éxons (Paul *et al.*, 1991). Este receptor é formado por duas subunidades, uma alfa (IL4Rα), que forma um sitio de ligação específico para IL-4, localizado na zona intracelular da proteína, com a função de conduzir a informação para o interior da célula, a partir da ativação de proteínas kinases (JAKs) que fosforilam uma série de substratos, dentre estes o fator de transcrição STAT6, que regula a transcrição de vários genes de resposta Th2. E outra subunidade

denominada γ (gama), que é comum a vários outros receptores e exerce função importante na condução do sinal (Holgate, 1999; Izuhara *et al*., 2004; Robaee *et al*., 2012).

Neste contexto, por ativar o padrão de resposta imune humoral (Th2), através da ativação e diferenciada de células B, caracteriza esta interleucina como um marcardor de ineficiência da resposta imunológica contra o bacilo e desenvolvimento da hanseníase (Teles *et al.,* 2010). O polimorfismo do tipo VNTR (rs79071878) do gene *IL-4* tem dois alelos mais frequentes, com duas repetições (A_1) e com três repetições (A_2). Adicionalmente, o alelo A_2 é responsável por elevar o nível de produção de IL-4 (Nakashima *et al.,* 2002), e pode estar relacionado como um fator de risco à doença.

A conversão de andrógenos em estrogénos, é catalisada pela aromatase codificada pelo gene *CYP19A1*, esta é a via primária de produção de estrogéno em seres humanos (Beitelshees *et al.,* 2010). Os níveis desses hormônios são importante em pacientes com hanseníase e tem sido demonstrado que os níveis de andrógenos são significativamente menores nos pacientes com hanseníase em comparação com indivíduos sem hanseníase (Foss *et al.,* 2012). Além disso, existe uma correlação inversa entre os níveis de androgénos plasmáticos e a secreção de citoquinas inflamatórias, sugerindo que os níveis de androgénos elevados no plasma pode ser menos eficaz na inibição do crescimento do Bacilo (Leal *et al.,* 2003). O alelo DEL (rs11575899), do gene *CYP19A1*, foi previamente relatada como tendo um efeito negativo sobre a actividade da aromatase (Limer *et al*., 2009), e pode assim ser um fator de ricso ao desenvolvimento da doença.

1.5 RNA Reguladores

Os RNAs reguladores englobavam tanto entidades expressas constitutivamente (RNA ribossômico, transportador, telomérico e pequenos RNAs nucleares que possuíam funções), quanto uma nova classe de RNA não codificantes, cuja expressão ocorre em fase específica do desenvolvimento ou no processo de diferenciação de um organismo. Estes atuam na coordenação de funções celulares eucarióticas importantes em nível transcricional e/ou nível pós-transcricional (Houwing *et al.,* 2007).

A descoberta de um controle regulatório associado ao papel catalítico do RNA descrito há 20 anos por Szymanski *et al.,* reconfigurou um novo e convincente conceito às moléculas de RNA como participantes ativos na regulação, catálise e controle de muitas reações que definem processos fundamentais celulares, papéis que, anteriormente, eram reservados às proteínas (Taft *et al.,* 2010), inserindo-se assim, a necessidade de desbravar um novo caminho na biologia molecular.

1.5.1 MicroRNA (miRNA)

No inicio de 1990, um grupo de biólogos que trabalhava com a manipulação de genes da plantas petúnias, buscavam acentuar a pigmentação roxa de suas pétalas por meio da síntese de antocianina via superexpressão do transgene da enzima chalcona sintase. Ao contrário da manifestação da coloração esperada, percebeu-se uma interação entre o bloqueio na síntese de pigmentos e a redução em 50 vezes dos níveis de mRNA transgênico, denominando-o efeito de co-supressão (Jorgensen, 1990; Napoli *et al.,* 1990).

Ambros *et al.,* 2003 ao buscarem defeitos genéticos no controle do calendário de desenvolvimento pós-embrionário em *Caenorhabditis elegans*, identificaram *lin-4*, um

gene incomum que originou um transcrito bifilamentar de pequeno RNA, responsável por controlar o processo de desenvolvimento no nematóide (Lee *et al.,* 1993). Em experimentos consecutivos, notou-se a complementaridade antisentido de *lin-4* a vários sítios de 3'*UTR* do mRNA de lin-14 (Lee *et al.,* 1993) sugerindo que este havia sofrido controle traducional por *lin-4*.

Estudos posteriores revelaram a regulação traducional de *lin-14* e *lin-28*, envolvidos nos estágios de desenvolvimento inicial e tardio de nematóide, respectivamente, os quais eram realizados por um mecanismo de interação entre RNA e RNA (Sevignani *et al.,* 2006), recebendo a denominação de "interferência de RNA" (iRNA) em nematóides (Hüttenhofer *et al.,* 2004).

Em 2000, foi descrita a molécula *let-7* que tinha como alvo o mRNA do gene *lin-41*, responsável por conduzir a transição do nematóide à fase adulta (Reinhart *et al.,* 2000). A mutação induzida no sítio de ligação de *lin-41* promovia a estagnação do desenvolvimento, ilustrando que a arquitetura global, bem como o sítio de ligação eram essenciais para a função eficaz do miRNA (Vella *et al.,* 2004). Neste mesmo ano, *let-7* foi identificado em humanos (Pasquinelli *et al.,* 2000).

1.5.2 Biogênese dos miRNA's

A biogênese de miRNAs caracteriza-se por um complexo processo de múltiplos passos ainda não totalmente esclarecidos, compostos pela associação de diferentes grupos de enzimas e proteínas, ocorridos no núcleo e no citoplasma celular. As fases desse processamento compreendem cinco passos sintetizados na:

a) Processo transcricional efetuado pela POLIMERASE II:

Estímulos temporais alternativos ocorrem para que haja a transcrição de gene de miRNA pela RNA POLIMERASE II. O produto gerado é um longo transcrito de miRNA primário (pri-miRNA) com tamanho aproximado de 700 a 1000 nucleotídeos, dispondo de 5'cap e cauda poli(A) semelhante aos demais RNAs mensageiros (Lee *et al.,* 2004; Sevignani, 2006). Devido à presença de regiões palindrômicas, esta longa molécula sobrepõem sobre si (pri-miRNA), como objetivo de aumentar sua estabilidade e proteger-se da ação de endonucleases (Kim, 2005), de onde futuramente, será originada uma estrutura madura chamada de *hairpin.*

b) Processamento nuclear pela DROSHA

DROSHA é uma proteína que possui dois domínios para RNase III duplicados (RIIDa e RIIIDb), um domínio de ligação C-terminal ao RNA duplex (dsRNA), região rica em PROLINA (PRR) e um domínio rico em ARGININA/SERINA (RS) na região N-terminal (Wu *et al.,* 2000; Lee *et al.,* 2003). A RNase III é capaz de reconhecer inúmeras moléculas de pri-miRNA, por apresentarem estruturas terciária que lhes designam como substrato, valorizando tanto o sítio de clivagem próximo à base, quanto o largo terminal do *stem-loop* (Lee *et al.,* 2003; Kim, 2005). A estrutura em grampo (pri-miRNA) após a ação catalítica da DROSHA gera uma molécula com aproximadamente 70 nt, denominada de pré-miRNA, (Lee *et al.,* 2003).

c) Transporte núcleo-citoplasma pela EXPORTINA-5 (Figura)

Ao término do processamento nuclear pelo microprocessador, o pré-miRNA é gerado e transportado ao citoplasma pela EXPORTINA-5 (EXP-5), uma proteína de exportação nuclear que utiliza Ran-GTP como cofator necessário para a ligação específica de pré-miRNAs à EXPORTINA-5. Esta interação requer que o terminal 3'*UTR* seja específico e o tronco do RNA duplex seja maior que 16 pb de tamanho (Cullen *et al.,* 2004).

d) Ação catalítica da DICER no citoplasma

O pré-miRNA é novamente processado por outra enzima da família RNase III, chamada DICER, uma proteína altamente conservada com variadas isoformas, que apresentam papéis distintos em organismos eucariotos. A DICER contém dois sítios catalíticos RIIIDa e RIIIDb na região C-terminal para RNA duplex, um domínio na região N-terminal para ATPase, um domínio helicase contendo caixa de DECH, e um domínio DUF 283 com função desconhecida, um domínio AGO, PIWI e PAZ. (Bernstein *et al.,* 2001; Provost *et al.,* 2002; Zhang *et al.,* 2002; Zhang *et al.,* 2004; Kim, 2005).

Zhang *et al.,* (2004) propuseram um modelo no qual a DICER era detentora de um único centro catalítico, executando sua função por meio da dimerização intramolecular de seus dois elementos catalíticos RIIID, assistidos pelo acompanhamento dos domínios PAZ e de ligação ao RNA duplex (dsRNA). O domínio PAZ participava do reconhecimento do terminal 3' do pré-miRNA, em que cada elemento catalítico RIIID clivava uma fita do duplex, gerando um saliência de 2 nucleotídeos (nt), devido ao alinhamento do dímero com um novo terminal 3'

hidroxilados e outro terminal 5’ fosforilado. A nova molécula gerada pela ação catalítica da DICER apresentava tamanho aproximado de 21 pb determinada pela distância entre o domínio PAZ e o sítio ativo de clivagem.

e) Formação do complexo efetor RISC

A molécula gerada a partir desse processamento é oriundo de um pequeno e imperfeito duplex (miRNA: miRNA*) que contém tanto a fita do miRNA maduro quanto a sua fita antisentido em vertentes separadas (Lee *et al.,* 2003). A relação de estabilidade termodinâmica é a responsável por determinar qual vertente será ligada ao complexo RISC e qual será degradada por enzimas. A preferência é concedida a vertente relativamente instável no terminal 5’*UTR*, determinador da montagem assimétrica do complexo RISC e, portanto a especificidade de destino pós-transcricional. No entanto, apesar de raro, pode ocorrer similaridade termodinâmica de ambos, igualando a frequência de veiculação ao RISC (Kim, 2005).

Após a escolha do unifilamento, este será incorporado a ribonucleoproteínas (RNPs) para formar RISC (RNA-induced silenciando o complexo) ou miRNP (Schwarz *et al.,* 2003). As ribonucleoproteínas são denominadas ARGONAUTAS 2 (AGO2) que pesam cerca de 100 kDa e dispõem de dois domínios: PAZ e PIWI. O domínio PAZ, localizado no centro da proteína, hibridiza-se à saliência 3' *UTR* do RNA unifilamentar processado anteriormente pela DICER (Song *et al.,* 2003;. Yan *et al.,* 2003), enquanto que o domínio PIWI apresenta uma homologia estrutural RNase H, agindo diretamente sobre a região de dominio do mesmo RNA unifilamentar, inibindo-o (Song *et al.,* 2004).

1.5.3 Mecanismo de regulação da expressão por miRNA

Há dois distintos mecanismos regulatórios pós-transcricionais apontado por ensaios sobre o mRNA alvo, clivando-o ou reprimindo-o traducionalmente. A escolha dependerá da interação molecular entre a extremidade 5'*UTR* do regulador e extremidade 3'*UTR* do regulado. A interação é determinada quando a sequência chamada *seed* (2-8 nucleotídeos iniciais da extremidade 5'UTR do miRNA) reconhece e pareia-se com mRNA alvo. Assim, observa-se que o pareamento perfeito entre as bases desencadeia a clivagem/degradação do mensageiro, enquanto que o pareamento imperfeito causa repressão traducional, observado respectivamente em siRNA e miRNA (Bartel, 2004; Nilsen *et al.,* 2007). Um resumo esquemático do processo é apresentado na Figura 6.

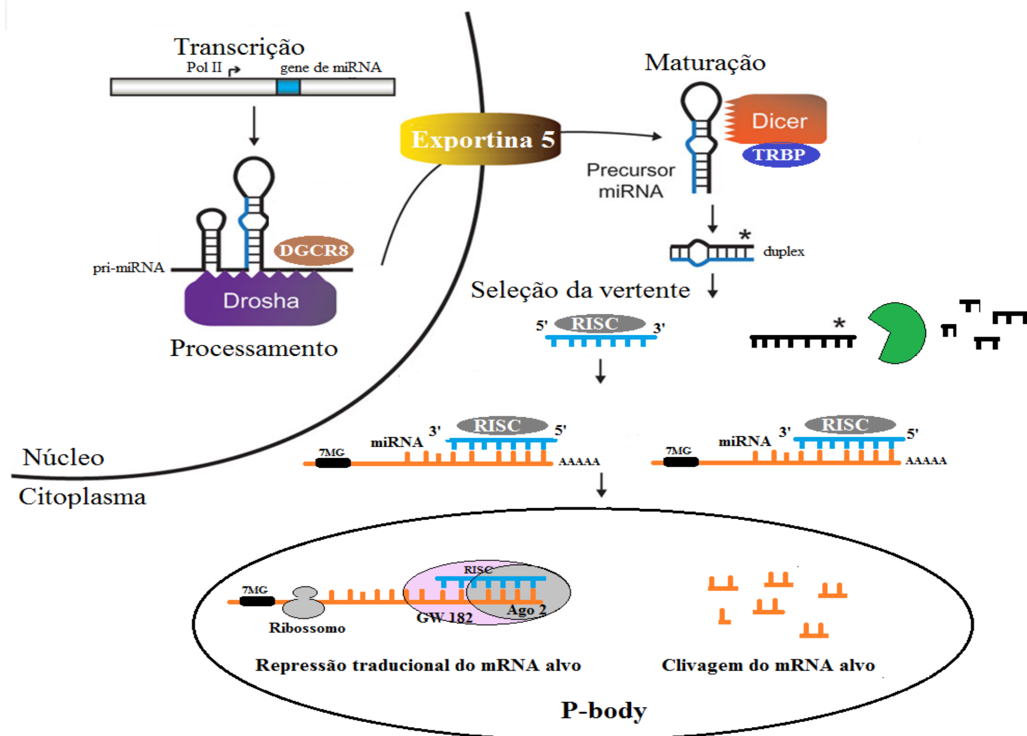


Figura 6. Resumo esquemático da biogênese do miRNA Fonte: Kai and Pasquinell, 2010.

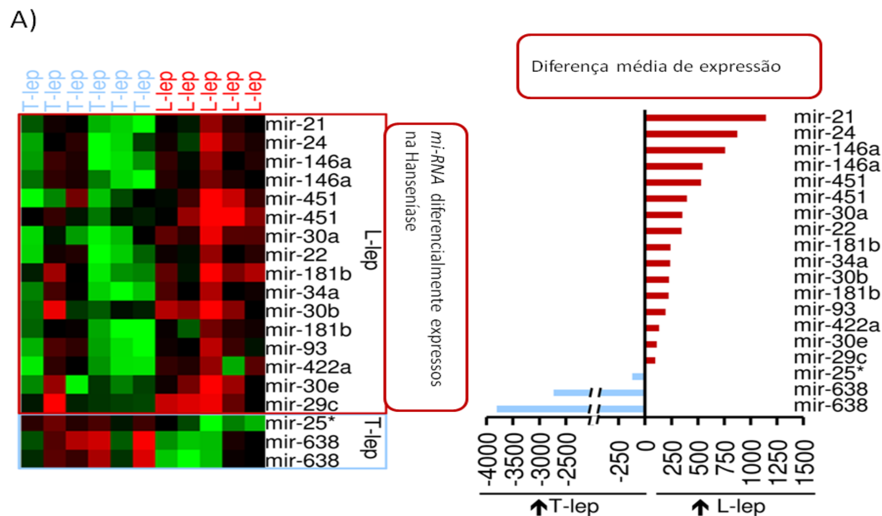
1.6 Biomarcadore epigenéticos associados à Hanseníase

Variados estudos já provaram que a função básica dos *miRNAs* é regular a expressão de um grande número de genes, por isso o perfil de expressão dessas moléculas é investigado em várias doenças infecciosas, como patologias bacterianas, virais e parasitárias, logo os estudos que buscam associar um padrão de expressão de *mi*RNA a determinadas doenças infectocontagiosas, estão revelando mecanismos de defesa até o momento desconhecidos, tal fato esta atraindo o interesse em se pesquisar a relação destas moléculas como reguladoras da resposta imune do hospeiro, frente a uma infecção bacteriana como a causada pelo *M. leprae* (Ding *et al.,* 2007; Cullen *et al.,* 2011; Hakimi *et al.,* 2011).

Embora os estudos atuais estabelecerem, com clareza, que determinadas interações entre humanos e o ambiente podem determinar um grau mais elevado de susceptibilidade à hanseníase, o entendimento de quais mecanismos moleculares estão envolvidos neste processo é incerto (Sing *et al*., 2013). Pesquisas recentes tem demonstrado o papel central que os *mi*RNAs exercem no desenvolvimento de doenças micobacterianas, assim o conhecimento destes mecanismos regulatórios podem auxiliar a elucidar a biologia da doença, e determinar como a resposta do hospedeiro à infecção por *M. leprae* pode modular os níveis de expressão de *mi*RNA e ocasionar um maior risco de desenvolvimento desta patologia (Sing *et al*., 2013; Hakimi *et al.,* 2011)

A avaliação de expressão de miRNA em distintos padrões de lesões hansênicas, encontrou níveis diferenciados de expressão em 16 miRNAs (Figura 7a) quando comparados indivíduos do pólo tuberculoíde (T-lep) *versus* individuos do pólo lepromatoso (L-lep) da doença (Liu *et al.,* 2012). Dentre os miRNAs estudados, o *hsa-miR-21* foi o mais diferencialmente expresso, com um padrão de hiperexpressão em

lesões L-lep, assim como quando avaliado monócitos infectado e mantidos em cultura para análise (Liu *et al.,* 2012). Experimentos adicionais demonstraram que o *hsa-miR-21* inibe a atividade de dois genes que traduzem peptídeos antimicrobianos dependentes de vitamina-D, *CAMP* e *DEFB4A*. A regulação ocorre por inibição direta da expressão dos genes *CYP27B1* e *IL1β* (Figura 7b). O *hsa-miR-21* não apresenta sítios alvo na região 3'UTR do gene *IL10*, porém regula um mecanismo que parece ser mediado pela interação entre *hsa-miR-21* e o receptor heterodimérico Toll-like 2/1 (TLR2/1), o qual é ativado durante o processo de infecção pelo *M. leprae*. Nestes mecanismos a conseqüência é a elevação do nível de mRNA do gene *IL10,* que diminuem a capacidade do hospedeiro de responder a infecção e desenvolver o pólo mais severo da doença. Quando o *hsa-miR-21* foi inibido em cultura de monócitos infectados, a expressão de *CAMP, DEFB4A, TLR2/1, CYP27B1* e *IL1β* voltaram para os níveis normais e as células passaram a exercer ações para combater o progresso da infecção (Liu *et al.,* 2012).



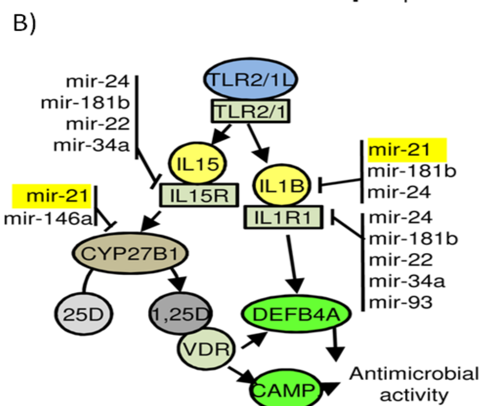


Figura 7.Perfil de expressão de *mi*RNA. A) entre os pólos distintos da hanseníase e B) predição de genes modulados pelos target de *mi*RNA. Fonte: Liu *et al.,* 2012. (com modificações).

Adicionalmente, o *hsa-miR-21* também pode suprimir a expressão de IL12, uma importante interleucina que é induzida em processos inflamatórios e estimula a resposta imune do tipo Th1 do hospedeiro, a qual determina um padrão de controle e combate da proliferação do *M. leprae*. Essa ação supressora também induz a apoptose de células denriticas, por inibição de *BCL2* (B-cell lymphoma 2), o que pode estar associado as sequelas sensitivas e motoras ocasionadas pela hanseníase (Kumar *et al*., 2011). Apesar do papel crítico do *hsa-miR-21* na hanseníase, o *hsa-miR-181a*, o qual esta hipoexpresso

em L-lep, desempenha um papel importante na progressão da doença. Este permite a hiperexpressão do gene *SHP2* (um tipo de receptor tirosina quinase) que ocasiona uma hiporesponsividade das células T em resosta a infecção micobacteriana (Kumar *et al*., 2011). O conjunto de *hsa-miR-21* e *hsa-miR-181a* parece modular a resposta imune do hospeiro (Tabela 1), dimunuindo os mecanismos de defesa e facilitanto a instalação e proliferação do *M. leprae*, assim como o desenvolvimento da doença.

Tabela1. Genes, alvos dos *hsa-mir-21* e *hsa-mir-181a*, moduladores da resposta imune do hospedeiro durante a infecção pelo *M. leprae*.

<table>
<tr><th>Gene
miRNA</th><th>Contig na
região 3’UTR</th><th>Região consenso</th></tr>
<tr><td>IL1β</td><td>517-523</td><td>5' ...UAAGACUGAAAAUAUAUAAGCUC...</td></tr>
<tr><td></td><td></td><td>| | | | | | |</td></tr>
<tr><td>hsa-mir-21</td><td></td><td>3' AGUUGUAGUCAGACUAUUCGAU</td></tr>
<tr><td>CYP27B1</td><td>369-375</td><td>5' ...AGCAUUUAUCAAAGC- AUAAGCUC...</td></tr>
<tr><td></td><td></td><td>| | | | | | | | | | |</td></tr>
<tr><td>hsa-mir-21</td><td></td><td>3' AGUUGUAGUCAGACUAUUCGAU</td></tr>
<tr><td>IL12</td><td>256-263</td><td>5' ...GAAGGGCAAAUAUUUAUAAGCUA...</td></tr>
<tr><td></td><td></td><td>| | | | | | | | |</td></tr>
<tr><td>hsa-mir-21</td><td></td><td>3' AGUUGUAGUCAGACUAUUCGAU</td></tr>
<tr><td>SHP2</td><td>64-71</td><td>5' ...AGAAGUGUGAACUUGUGAAUGUA...</td></tr>
<tr><td></td><td></td><td>| | | | | | |</td></tr>
<tr><td>hsa-mir-181a</td><td></td><td>3' UGAGUGGCUGUCGCAACUUACAA</td></tr>
</table>

2 - OBJETIVOS

2.1 Objetivo Geral

Investigar sete polimorfismos de sete genes moduladores do sistema imune (*CYP19A1* (rs11575899), *NFKβ1* (rs28362491), *CASP8* (rs3834129), *PAR1* (rs11267092), *CYP2E1* (INDEL 96pb), *IL4* (rs79071878), assim como o perfil de expressão diferencial dos microRNAs, identificados no miRnoma de amostras de pacientes com hanseníase na população do Estado do Pará.

2.2 Objetivos Específicos

- Investigar a associação de polimorfismos presentes em sete genes envolvidos na modulação da resposta imune, como fatores de risco à hanseníase, assim como à progressão clínica da doença;
- Realizar o primeiro sequenciamento de alto desempenho em grupos de pacientes acometidos pela hanseníase;
- Investigar e descrever o perfil de expressão dos microRNAs em diferentes tipos de lesões hansênicas (TT, e LL);
- Comparar o perfil de expressão diferencial dos microRNAs dos distintos grupos de pacientes com hanseníase;
- Realizar a predição de alvos *in silico*, na busca de possíveis biomarcadores de risco ao desenvolvimento da doença;
- Melhorar o conhecimento acerca dos mecanismos biológicos do hospedeiro no combate a doença.

3 – CAPÍTULO 1 - INFLUENCE OF GENETIC ANCESTRY ON INDEL MARKERS OF *NFKB1, CASP8, PAR1, IL4* AND *CYP19A1* GENES IN LEPROSY PATIENTS.

Artigo publicado na revista Plos Neglected Tropical Disease

doi: 10.1371/journal.pntd.0004050. eCollection 2015 (Fator de Impacto 4.71 – Qualis A1).

PLOS | NEGLECTED TROPICAL DISEASES

RESEARCH ARTICLE

Influence of Genetic Ancestry on INDEL Markers of *NFKβ1*, *CASP8*, *PAR1*, *IL4* and *CYP19A1* Genes in Leprosy Patients

Pablo Pinto[1,2], Claudio Salgado[3], Ney Pereira Carneiro Santos[2], Sidney Santos[1,2], Ândrea Ribeiro-dos-Santos[1,2]*

1 Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Pará, Brasil, **2** Núcleo de Pesquisas em Oncologia - NPO, Universidade Federal do Pará, Belém, Pará, Brasil, **3** Laboratório de Dermatoimunologia, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, Pará, Brasil

* akely@ufpa.br, akelyufpa@gmail.com

OPEN ACCESS

Citation: Pinto P, Salgado C, Santos NPC, Santos S, Ribeiro-dos-Santos Â (2015) Influence of Genetic Ancestry on INDEL Markers of *NFKβ1*, *CASP8*, *PAR1*, *IL4* and *CYP19A1* Genes in Leprosy Patients. PLoS Negl Trop Dis 9(9): e0004050. doi:10.1371/journal.pntd.0004050

Editor: Christian Johnson, Fondation Raoul Follereau, FRANCE

Received: March 25, 2015

Accepted: August 12, 2015

Published: September 14, 2015

Copyright: © 2015 Pinto et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available within the manuscript and Supporting Information files.

Funding: This study was supported by FAPESPA (Fundação Amazônia Paraense de Amparo à Pesquisa) ICAAF 155/2014, CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), CNPQ 481652/2012-4, CAPES (Coordenação de Aperfeiçoamento Pessoal de Nível Superior), Rede de Pesquisa em Genômica Populacional Humana (RPGPH)-CAPES, CAPES-Proamazonia 3288/2013 and FADESP/PROPESP/

Abstract

Background

Leprosy is an insidious infectious disease caused by the obligate intracellular bacteria *Mycobacterium leprae*, and host genetic factors can modulate the immune response and generate distinct categories of leprosy susceptibility that are also influenced by genetic ancestry.

Methodology/Principal Findings

We investigated the possible effects of *CYP19A1* [rs11575899], *NFKβ1* [rs28362491], *IL1α* [rs3783553], *CASP8* [rs3834129], *UGT1A1* [rs8175347], *PAR1* [rs11267092], *CYP2E1* [INDEL 96pb] and *IL4* [rs79071878] genes in a group of 141 leprosy patients and 180 healthy individuals. The INDELs were typed by PCR Multiplex in ABI PRISM 3130 and analyzed with GeneMapper ID v3.2. The *NFKβ1*, *CASP8*, *PAR1* and *IL4* INDELs were associated with leprosy susceptibility, while *NFKβ1*, *CASP8*, *PAR1* and *CYP19A1* were associated with the MB (Multibacilary) clinical form of leprosy.

Conclusions/Significance

NFKβ1 [rs28362491], *CASP8* [rs3834129], *PAR1* [rs11267092] and *IL4* [rs79071878] genes are potential markers for susceptibility to leprosy development, while the INDELs in *NFKβ1*, *CASP8*, *PAR1* and *CYP19A1* (rs11575899) are potential markers for the severe clinical form MB. Moreover, all of these markers are influenced by genetic ancestry, and European contribution increases the risk to leprosy development, in other hand an increase in African contribution generates protection against leprosy.

UFPA (Universidade Federal do Pará). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

Author Summary

Leprosy is an infectious disease caused by *Mycobacterium leprae*, which can carry to skin lesions and affect peripheral nerves, which cause physical and motor injuries on the patients. Moreover, leprosy, may be classified in two major groups, based on clinical manifestations in Paucibacillary (PB) or Multibacillary (MB), and these phenotype may be influenced by host immune response; that can be controlled by genetics factors that can be useful like future panel of biomarkers to leprosy, and it's related with the different genetic background of population studied. Therefore, we conducted a study to evaluate seven INDEL polymorphisms in seven genes involved in modulation of the host immune response, and consequently can modulated o phenotype showed through the disease, to identify possible susceptibility markers of leprosy. However this analysis can be spurious on presence of population structure, common in admixture population like the Brazilian, thus we evaluate like the influence of genetic ancestry can modulated the disease risk.

Introduction

Leprosy is an insidious infectious disease caused by the obligate intracellular bacteria *Mycobacterium leprae* that affects the skin and peripheral nerves, causing a chronic granulomatous infection [1]. Leprosy patients may be classified in two major groups, based on clinical manifestations using a simple system introduced by the WHO (World Health Organization) in 1982. Paucibacillary (PB) is the primary characteristic of Tuberculoid (TT) leprosy and is characterized by a few lesions and scarce bacilli, and Multibacillary (MB) is the primary characteristic of anergic Lepromatous (LL) leprosy. From an epidemiological perspective, the situation in Brazil is critical because, along with India and Indonesia, it has the highest rate of new cases detected worldwide [2, 3, 4].

In addition to the system introduced by WHO in 1982, the use of histological and immunological criteria as described by Ridley-Jopling further improves definition of Borderline cases. According to this classification, TT (tuberculoid-tuberculoid) patients, who have the PB type, exhibit a strong cellular immune response (CIR) mediated by Th1, and a negative skin smear test. In contrast, LL (lepromatous-lepromatous) patients have a weak or absent CIR and a highly positive skin smear associated to an humoral immune response. In the middle of this spectrum are a large number of borderline patients, which together with LL comprise the MB pole, with symptoms varying from weak to strong CIR and negative to positive skin smears [5, 6].

The regulation of the host immune response and manifestation of disease clinical between types PB (better) and MB (severe) involves cytokine and others mediators produced by various subtypes of T cells. In PB, an inflammatory immune response is mediated by Th1 cells that express pro-inflammatory interleukins that stimulate macrophages and phagocytosis mechanisms to inhibit bacillary growth and kill mycobacteria [2,7–9]. On the other hand, MB patients have an intense Th2 immune response with production of anti-inflammatory cytokines in addition to the specific anti-PGL-1 (phenolic glycolipid 1) antibody. This mechanism does not block bacillary growth and contributes to the host's inability to resist the development of severe disease [2,8,9–11].

Recent studies have investigated genetic markers, usually innate immune response genes, as possible susceptibility factors for leprosy because the SNPs in these genes can modulated the host immune response and consequently lower host resistance to bacillus growth [6,12,13]. However, few studies have investigated INDEL polymorphisms (insertion-deletion) in immune response genes in leprosy. Moreover, such polymorphisms present interesting features as

genetic markers because i) INDELs are spread throughout the human genome, ii) INDELs derive from a single event (they do not present homoplasy), iii) small INDELs can be analyzed using short amplicons, which improves amplification of degraded DNA and facilitates multiplexing reaction, iv) INDELs can create abrupt changes in the normal function of the gene and v) INDELs can be easily genotyped using a simple dye-labeling electrophoretic approach [14]. The current study select eight INDEL in seven genes (*CYP19A1*, *NFKβ1*, *IL1α*, *CASP8*, *UGT1A1*, *PAR1*, *CYP2E1*, and *IL4*), which have relation with the immune response modulation in leprosy patients, beside literature that demonstrate these molecular markers like functional polymorphisms that alter transcriptional activity of the gene, and consequently the immunological phenotype against the bacilli. Additionally these INDELs can be able to contribution to construction a possible panel of susceptibility markers.

However, from the genetic point of view, Brazil is recognized as having one of the most heterogeneous populations in the world, with important genetic information being contributed by three main continental groups, Europeans, Africans and Amerindian, resulting in a genetically very diverse modern Brazilian population [15]. Therefore, analysis of genetic markers in complex diseases may result in spurious results due to population substructure [16], and it is important to perform the genomic ancestry control, especially in populations with a high degree of interethnic admixture [14].

The objective of this study was to investigate eight INDEL polymorphisms in seven genes involved in modulation of the host immune response, including *CYP19A1* [rs11575899], *NFKβ1* [rs28362491], *IL1α* [rs3783553], *CASP8* [rs3834129], *UGT1A1* [rs8175347], *PAR1* [rs11267092], and *CYP2E1* [INDEL 96pb], besides one VNTR (variable number tandem repeat) of 70 bp on intron 3 of *IL4* [rs79071878] in a group consisting of 141 leprosy patients and 180 healthy individuals, to identify possible susceptibility markers of leprosy and evaluate the influence of genetic ancestry on disease risk.

Materials and Methods

Ethics statement

The project was approved by the Pará Federal University ethics committee (N° 197/07).

Samples

We investigated 141 leprosy patients who attended the Dr Marcello Candia Reference Unit in Sanitary Dermatology of the State of Pará (UREMC), in Marituba, Pará, Brazil between January 2008 and December 2009. All patients were informed about the study before they signed informed consent forms. Since 2002, UREMC registered between 308 and 472 leprosy patients (mean: 408 cases per year). Of the 765 leprosy cases registered in 2008 and 2009 alone, 141 (18.43%) were randomly selected for this study. These patients were divided according to Ridley-Jopling classification [5] into Paucibacillary (TT: PB 31) and Multibacillary (BT, BB, BL and LL: MB 110) groups. A total of 180 healthy individuals who were unrelated, without leprosy or other chronic diseases and from the same geographic area as each other were chosen for the control group. Leprosy patient's descriptions were made previously [6]. These subjects were asked to participate in the study after being informed about the study objectives and signing informed consent forms.

DNA extraction

DNA extraction was performed as previously described by phenol-chloroform method [6, 17]. The DNA concentration was determined by spectrophotometry (Themo Scientific NanoDrop 1000, NanoDrop Technologies, Wilmington, US).

Multiplex Typing. DNA samples were typed for the 7 biallelic INDELs and 1 VNTR of 70 bp. Each multiplex PCR was performed in a final volume of 10 μL containing 5 μL of QIAGEN Multiplex PCR Kit, 1 μL of Q-solution, 1 μL of primer mix (Forward + Reverse primers at a concentration of 2 μM each), 1 μL of DNA (10 ng) and 2 μL of water. The fluorescent molecules 6FAM and HEX were inserted at the 5' position of each primer (forward or reverse). All PCR reactions were performed on the Thermocycle Veriti-96 Well Thermal Cycler (Life Technologies, CA, USA). Before capillary electrophoresis, 1.0 μL PCR product was added to 8.5 μL deionized formamide HI-DI (Life Technologies, CA, USA) and 0.5 μL GeneScan 500 LIZ size standard (Life Technologies, CA, USA). DNA fragments were separated using an ABI PRISM 3130 Genetic Analyzer (Life Technologies, CA, USA) and analyzed with GeneMapper ID v3.2 software (Life Technologies, CA, USA). This method is similar like described in a paper of INDEL markers of our group [14].

Ancestry Informative Marker (AIM)

Individual interethnic admixture was estimated using a panel of 48 ancestry informative markers (AIMs) as previously described [6, 14].

Statistical analysis

The allelic frequencies between healthy individuals and leprosy patients and between PB and MB patients were estimated by gene counting. Deviation from the Hardy-Weinberg equilibrium was assessed using chi-squared tests, using the Arlequin v3.5 software [18], and p-value of HWE was corrected by Bonferroni methods.

Differences between leprosy patients and healthy individuals and between PB and MB patients with respect to age, gender and genetic ancestry were estimated using Student's t-Test, Fisher's exact test and Mann-Whitney tests, respectively. The association of markers between groups was analyzed by logistic regression tests, all the test were corrected by FDR (False Discovery Rate) method, and all tests were performed using the statistical package under R calculation. A two-tailed p-value < 0.05 was considered statistically significant.

The individual contributions of European, African and Amerindian genetic ancestry were estimated using the STRUCTURE 2.3.3 program assuming three parental populations (European, African and Amerindian), a burn-in period of 200,000, and 200,000 Markov Chain Monte Carlo repetitions after burn-in [16]. The differences in allelic frequencies between leprosy cases and the healthy individuals for markers analyzed following an adjustment for population stratification was performed using the STRAT software program with 10,000 simulations [16].

Results

The data of clinical and demographic distribution of leprosy patients and healthy individuals is shown in Table 1. The mean age was higher in healthy individuals (55.7±12 versus 43.3±21, $p<0.001$), and male patients were more frequent among leprosy patients (97 [68.8%] versus 65 [36.1%], $p<0.001$). Analysis of ethnicity showed that the mean frequency of Africans was higher among leprosy patients (0.284 versus 0.236, $p<0.001$) and Europeans were more frequent in healthy individuals (0.461 versus 0.427, $p = 0.004$).

The frequencies of INDELs for the eight (8) genes analyzed in leprosy patients and healthy individuals are show in Table 2. For the polymorphism in *IL4* (VNTR of 70 bp), only two alleles were identified in the sample. One allele had two repeats of 70 bp (allele A1) and the other had three repeats of 70 bp (allele A2), suggesting theses alleles are biallelic markers. All the polymorphisms analyzed were according to the Hardy Weinberg equilibrium, therefore the

Table 1. Demographic and clinical characteristic of the sample of leprosy patients and healthy individuals.

Variables	LEPROSY PATIENTS (n = 141)	HEALTHY INDIVIDUALS (n = 180)	*p value* (IC-95%)
Age[a]	43.3±21	55.7±12	<0.001
Gender[b] (M/F)	97(68.8%)/44(31.2%)	65(36.1%)/115(63.9%)	<0.001
Genetic Ancestry[c]			
African	0.284±0.11	0.236±0.04	<0.001
European	0.427±0.13	0.461±0.06	0.004
Ameridian	0.289±0.11	0.303±0.09	0.094

[a] *t*-Test of Student;

[b] Fisher's Exact Test;

[c] Mann-whitney test; The data are show like mean ± standard deviation.

doi:10.1371/journal.pntd.0004050.t001

association analysis were performed with regression logistic test and differences in allelic frequencies were corrected by frequencies of ancestry markers informative.

When the INDELs were analyzed by logistic regression, the genes *NFKβ1* and *PAR1* showed statistically significant differences associated with the presence of the DEL allele (p = 0.016 and p = 0.022, respectively) and both were associated like protection factors to not developing the disease (OR[IC95%] = 0.50[0.27–0.88] and OR[IC95%] = 0.35[0.14–0.86], respectively), for these genes was found a dominance effect DEL allele, that increase your protection capacity in general population. The *CASP8* showed significant differences associated with the presence of the DEL/DEL homozygous genotype and was associated with a risk factor for leprosy development (p = 0.017; OR[IC95%] = 2.33[1.16–4.69]) (Table 3). The analysis of allele frequency differences was then corrected for the influence of genetic ancestry on population structure, and the results showed that the DEL allele of *PAR1* gene and the allele A_1 of *IL4* is more frequent in healthy individuals (p = 0.018 and p = 0.019, respectively) (Table 3), these results shown the importance of statistical correction in admixture population, in order to exhibit differences covert by structure population.

Table 4 summarizes the clinical and demographic characteristics of leprosy patients grouped according to clinical manifestation in PB (Paucibacillary) and MB (Multibacillary) groups, and the only significant difference was observed for age (p = 0.003), with a higher mean age in MB patients (45.7±22 versus 34.9±15). When the INDELs were analyzed by logistic regression, *NFKβ1* showed significant differences like risk factor associated with the presence of the allele DEL in MB patients (p = 0.024; OR[IC95%] = 2.64[1.13–6.19]), of contradictory way the dominance effect of DEL allele seem protect against the development of leprosy, but when the disease is established your effect seem inefficient to combat to bacilli. *PAR1* showed significant differences associated with the presence of homozygous DEL/DEL genotype in PB patients (p = 0.031; OR[IC95%] = 0.41[0.17–0.96]) (Table 5). The analysis of allele frequency differences were corrected for population structure and showed that the DEL allele of *CASP8* is more frequent in PB patients (p = 0.003), while the DEL allele of *CYP19A1* is more frequent in MB patients (p = 0.007) (Table 5).

Fig 1 shows the OR (odds ratio) values obtained from leprosy patients and healthy individuals within groups having distinct level of ancestry composition. The figure shows that greater frequency of European ethnic between the groups (leprosy patients and healthy individuals), higher is the risk for developing leprosy, while the smaller the frequency of the African ethnic, lower is the risk for developing leprosy. No statistically significant values were obtained for the analysis of the Amerindian group.

Table 2. Allele frequencies of INDELs for the eight investigated genes.

GENE	LEPROSY PATIENTS (n = 141)	HEALTHY INDIVIDUALS (n = 180)
CYP19A1 (rs11575899)		
INS	0.574	0.558
DEL	0.425	0.442
HWE	0.998	0.998
NFKβ1 (rs28362491)		
INS	0.546	0.467
DEL	0.454	0.533
HWE	0.281	0.266
IL1α (rs3783553)		
INS	0.617	0.544
DEL	0.383	0.456
HWE	0.898	0.779
CASP8 (rs3834129)		
INS	0.557	0.586
DEL	0.443	0.414
HWE	0.215	0.996
UGT1A1 (rs8175347)		
INS	0.411	0.325
DEL	0.589	0.675
HWE	0.280	0.487
IL4 (rs79071878)		
A_2[a]	0.660	0.581
A_1[b]	0.340	0.419
HWE	0.521	0.876
PAR1 (rs11267092)		
INS	0.320	0.217
DEL	0.680	0.783
HWE	0.06	0.196
CYP2E1		
INS	0.088	0.083
DEL	0.911	0.916
HWE	0.600	0.994

HWE = p-value for Hardy Weinberg equilibrium after Bonferroni correction;
[a] Allele with three repeats of 70 pb;
[b] Allele with two repeats of 70 pb

doi:10.1371/journal.pntd.0004050.t002

These results are better understood on frequencies distribution, according with range of ancestry contribution (S1 Table). For African ancestry 99.4% of health individual is closed between 0% and 50% of African contribution (range that have $p<0.05$ on Fig 1), moreover the contribution range of 10% to 30% is closed 81.7% of health individual, in this range the Fig 1 have more decline of OR value, that showed the higher protection effect of African ancestry. To European ancestry, 61.7% of leprosy patients is closed between 40% to 80% of European contribution, while 87.2% of health individuals is closed between 0% to 50% of European contribution. Additionally, for the contribution range between 60% to 80% we observed 17% of all patients, while no healthy individual was observed this range, these data show that leprosy patients have higher European contribution compared with healthy individuals. Take together

Table 3. Allelic and genotypic distribution between leprosy patients and healthy individuals to markers associated whit susceptibility to leprosy.

GENE	LEPROSY PATIENTS n(%)	HEALTHY INDIVIDUALS n(%)	p^a	OR (IC95%)[b]	p_{STRAT} [c]
NFKβ1 (rs28362491)					
INS/INS	45(31.9%)	35(19.4%)		1	
INS/DEL	64(45.4%)	98(54.4%)			
DEL/DEL	32(22.7%)	47(26.1%)	0.559	0.83(0.64–2.34)	
[DEL]carriers	96(68.1%)	145(80.5%)	**0.016**	**0.50(0.27–0.88)**	
INS	0.546	0.467			
DEL	0.454	0.533			0.243
PAR1 (rs11267092)					
INS/INS	19(13.5%)	11(6.1%)		1	
INS/DEL	52(36.9%)	56(31.1%)			
DEL/DEL	70(49.6%)	113(62.8%)	0.094	0.64(0.39–0.86)	
[DEL]carriers	122(86.5%)	169(93.9%)	**0.022**	**0.35(0.14–0.86)**	
INS	0.320	0.217			
DEL	0.680	0.783			**0.018**
CASP8 (rs3834129)					
INS/INS	50(35.5%)	62(34.4%)		1	
INS/DEL	57(40.4%)	87(48.3%)			
DEL/DEL	34(24.1%)	31(17.2%)	**0.017**	**2.33(1.16–4.69)**	
[DEL]carriers	91(64.5%)	118(65.5%)	0.413	0.80(0.47–1.36)	
INS	0.557	0.586			
DEL	0.443	0.414			0.123
IL4(rs79071878)					
A_2/A_2	63(44.7%)	60(33.3%)		1	
A_2/A_1	60(42.6%)	89(49.4%)			
A_1/A_1	18(12.8%)	31(17.2%)	0.132	0.56(0.26–1.18)	
[A_1]carriers	78(55.4%)	120(66.6%)	0.088	0.63(0.37–0.84)	
A_2[d]	0.660	0.581			
A_1[d]	0.340	0.419			**0.019**

[a] *p-value* obtained for logistic regression adjusted by age, gender and genetic ancestry;

[b] Adjusted Odds Ratio (OR);

[c] p-value after correction for population structure;

[d] A_1—allele with two tandem repeats A_2—allele with three tandem repeats.

doi:10.1371/journal.pntd.0004050.t003

the Fig 1 and S1 Table shown that to leprosy patients of an admixture population, like Brazil, African ethnic generates protection against the development of disease, and the opposite is also truth for European ethnic.

Discussion

NF-κB belongs to family of protein transcription factors that modulate many inflammatory processes. In the resting state, IκBα (inhibitor of NF-kβ activity) sequesters NF-κB in the cytoplasm and prevents its activity, but in response to specific stimuli, IkBα is ubiquitinated and degraded allowing NF-kB to migrate to the nucleus and stimulate the transcription of proinflammatory genes [19,20]. The allele DEL (rs28362491) has been shown to be associated with a decrease of transcriptional activity of variety genes of immune response [21] and with auto immune disease such as Systemic Sclerosis [22] and lupus erythematosus [23].

Table 4. Demographic and clinical characteristics of the sample according with clinical form of leprosy.

Variables	LEPROSY PATIENTS		*p value* (IC-95%)
	PB (n = 31)	MB(n = 110)	
Age[a]	34.9±15	45.7±22	**0.003**
Gender[b] (M/F)	19(63.3%)/11(36.7%)	78(70.9%)/32(29.1%)	0.504
Genetic Ancestry[c]			
Afric	0.290±0.10	0.282±0.12	0.534
European	0.419±0.09	0.429±0.13	0.964
Ameridian	0.289±0.10	0.288±0.11	0.648

[a]*t*-Test of Student;

[b]Fisher's Exact Test;

[c] Mann-whitney test; The data are show like mean ± standard deviation.

doi:10.1371/journal.pntd.0004050.t004

The role of NF-kβ in leprosy is not clear, and studies linked to expression of NF-kβ have suggested that lower expression is common in leprosy patients [24,25]. Our results suggest that the DEL carries genotype induces protection against leprosy (Table 3), although a comparison of PB and MB patients also suggests that DEL behaves like a risk factor for the development of the severe clinical form of MB (Table 5). Because the transcription of NF-kβ is mediated by specific stimuli, such as the presence of *M. leprae* [24], it is conceivable that the presence of DEL confers risk to MB leprosy.

PAR1 is a receptor of the PAR family of proteins that belong to a unique group of G protein —coupled receptors. In particular, *PAR1* protein is present in a variety of cells like platelets, endothelia, epithelial, neurons, fibroblasts, smooth muscle, leukocytes and tumor lines [26]. This receptor has been shown to be involved in many natural physiological processes, that involve inflammation like the systems cardiovascular, respiratory and central nervous and in embryogenesis, cancer and inflammation [27]. *PAR1* suppresses T helper type 1 (Th1) and T helper type 17 (Th17) cells and the secretion of IL-12 and IL-23, thereby resulting in the inhibition of pro-inflammatory responses [28]. The allele of insertion (INS) of INDEL studied (rs11267092) has been shown to increase gene transcription [29] and therefore, it is a risk allele for leprosy. Our results suggest that the presence of DEL induces protection against leprosy (Table 3), and the DEL/DEL genotype confers protection against the development of clinical forms of MB (Table 5), thus this genotype of *PAR1* gene can suppresses cellular infiltration and increase both Th1 and Th17 responses to infection. Moreover, analyses of macrophages revealed that secretion of IL-12 and IL-13, two cytokine that play role key on cellular immunity Th1 and Th17, can be suppressed by *PAR1* activation. Furthermore, *PAR1* can suppress interferon regulatory factor 5 (IRF5), that play role key like transcription factor for IL-12 and IL-23, which modulates the sub sets of cellular immunity. Thereby the suppression of IRF5 and IL-12/23 secretion by *PAR1* gene, can provides a novel mechanism by which the host suppresses the Th1 and Th17response to infection, and dysregulation of this process can likely an important factor in the susceptibility of some individuals to leprosy [28].

Macrophages with a high load of *M. leprae* have been shown to undergo apoptosis, and this mechanism is under the control of cytokines [30]. In leprosy patients, the immune system is overburdened with bacilli, and most likely the continuous activation of T cells by circulating *M. leprae* antigens leads to apoptosis and to a reduction of peripheral lymphocytes and other immune effector cells in these patients with the regulation of apoptosis involved in the stimulation and activation of caspase-8 [31]. The allele DEL (rs3834129) cause a decrease in *CASP8*

Table 5. Allelic and genotypic distribution between leprosy patients grouped according clinical form PB or MB.

GENE	PATIENTS PB n(%)	PATIENTS MB n(%)	p^a	OR (IC95%)[b]	p_{STRAT} [c]
NFKβ1 (rs28362491)					
INS/INS	15(48.4%)	30(27.3%)		1	
INS/DEL	13(41.9%)	51(46.4%)			
DEL/DEL	3(9.7%)	30(27.3%)	0.119	2.78(0.76–10.07)	
[DEL]carriers	16(51.6%)	81(73.7%)	**0.024**	**2.64(1.13–6.19)**	
INS	0.694	0.495			
DEL	0.306	0.504			0.410
PAR1 (rs11267092)					
INS/INS	5(16.1%)	14(12.7%)		1	
INS/DEL	5(16.1%)	47(42.7%)			
DEL/DEL	21(67.7%)	49(44.5%)	**0.031**	**0.41(0.17–0.96)**	
[DEL]carriers	26(83.9%)	96(87.2%)	0.259	1.98(0.60–6.55)	
INS	0.241	0.340			
DEL	0.759	0.660			0.223
CASP8 (rs3834129)					
INS/INS	7(22.6%)	43(39.1%)		1	
INS/DEL	16(51.6%)	41(37.3%)			
DEL/DEL	8(25.8%)	26(23.6%)	0.579	0.76(0.29–1.97)	
[DEL]carriers	24(77.4%)	67(60.9%)	0.114	0.46(0.18–0.90)	
INS	0.484	0.577			
DEL	0.516	0.423			**0.003**
CYP19A1(rs11575899)					
INS/INS	14(45.2%)	32(29.1%)		1	
INS/DEL	17(54.8%)	53(48.2%)			
DEL/DEL	-	25(22.7%)	0.998	-	
[DEL]carriers	17(54.8%)	78(70.9%)	0.082	2.11(1.00–4.93)	
INS	0.726	0.532			
DEL	0.274	0.468			**0.007**

[a] *p-value* obtained for logistic regression adjusted by age, gender and genetic ancestry;

[b] Adjusted Odds Ratio (OR);

[c] p-value after correction for population structure.

doi:10.1371/journal.pntd.0004050.t005

transcription and a reduction in apoptosis [32], thereby improving the bacillary load. Our results suggested that the DEL/DEL genotype (Table 3) and the high frequency of DEL allele (Table 5) can raise the bacillary load and thus confers a risk to leprosy development.

Interleukin-4 (IL-4) is a key cytokine secreted by Th2 lymphocytes, eosinophils and mast cells that induces the activation and differentiation of B cells and the development of the Th2 subset of lymphocytes, which is ineffective in combating leprosy [33]. Our analysis of the VNTR on intron 3 of the *IL4* gene (rs79071878) revealed two common alleles with two (A_1) and three (A_2) tandem repeats. Of these, A_2 allele is known to be a high producer of IL-4 [34]. Our results indicate that allele A_2 is more frequent in leprosy patients compared to healthy individuals, consistent with the fact that higher levels of *IL4* would be ineffective in controlling the growth of bacilli (Table 3).

The conversion of androgens to estrogens, catalyzed by aromatase encoded by the *CYP19A1* gene, is the primary pathway of estrogen production in humans [35]. The levels of these hormones are important in leprosy patients and it has been demonstrated that androgen levels are

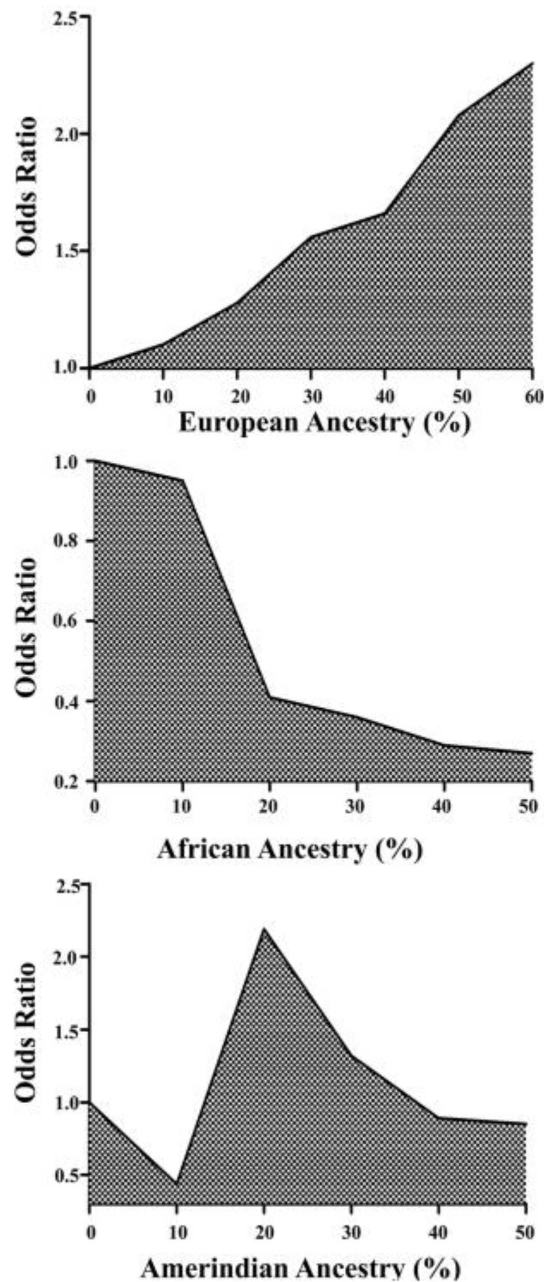


Fig 1. In a comparison of 141 leprosy patients and 180 healthy individuals, events with statistically significant ($p<0.05$) differences can be categorized into six categories of individuals with African and European genetic ancestry (10%>20%> 30%>40%>50%>60%), the p-values were adjusted by age and gender. The analysis in individuals with Amerindian ancestry was not statistically significant.

doi:10.1371/journal.pntd.0004050.g001

significantly lower in leprosy patients compared to healthy control subjects [36]. Moreover, there is an inverse correlation between plasma androgen levels and secretion of inflammatory cytokines, suggesting that high plasma androgen levels can be less effective in inhibiting bacillus growth [37]. The DEL allele (rs11575899) has previously been reported to have a negative effect on aromatase activity [38], and our results show that the DEL allele is more frequent in MB patients (Table 5). We hypothesize that the DEL allele can decrease aromatase activity and increase androgen levels, resulting in an overall reduction in effective combat of bacillary growth and development of the severe clinical form MB.

It is unclear whether leprosy originated in Asia or Africa. However, leprosy is believed to have been introduced into Europe from India, and the incidence was high in Europe during the Middle Ages until approximately 1870 when the number of cases dramatically reduced because of socioeconomic development [39,40,41]. It is believed that leprosy was introduced in Brazil primarily by the Spanish and Portuguese [41]. Estimates indicate that before the arrival of colonizers, approximately 2.5 million natives lived in Brazil, and during the European immigration in the first three centuries, approximately 500,000 individuals came from Portugal and approximately 3.5 million Africans were brought into Brazil through slave trade [14]. Therefore, there is evidence of a so-called directed admixture process involving predominantly European, Native American and African people [42–45].

Our data indicates that the contribution of different ethnic groups to the composition of the current Brazilian population can generate different rates of risk for leprosy development according to the level of inter-ethnic composition of the individuals involved. Our analysis suggests that an increase in European contribution increases the risk of leprosy development, while an increase in African contribution decreases the risk for leprosy development and the Amerindian contribution does not result in any statistically significant differences (Fig 1).

The introduction of leprosy in Brazil primarily can it be accredited to the slave trade, but no only for this reason. Slaves were firstly there from Africa, and in succeeding years the number these slaves were increased, but was not common between they the clinical manifestation of leprosy, because these slaves were from region of the Africa where leprosy was comparatively rare. Moreover isn't doubt that the Portuguese and, to a less degree, Dutch, French and Spaniards were responsible by introduction of leprosy in Brazil, on period of country colonization. Additionally, data showed that as early, as 1419, the disease was common in Portuguese and epidemiologically in this time the leprosy was very prevalent in Europe, and particularly in Portugal [41]. Therefore, our data of risk of leprosy according the different ethnic groups compositions is consistent with the higher numbers of settlers Portuguese that came to Brazilian that probably increases the frequencies of alleles of susceptibility on Brazilian population [14, 42–45]. In other hand, the African contribution may have increase the frequencies of allele that confer protection against to leprosy.

Comparative analyses of the four *M. leprae* genomes (India, Thailand, Brazil and US) have revealed little clonal differences. Thus, the patterns of global human migration routes, during the past 100,000 years, corroborate and suggest that leprosy probably originated in Africa [46]. African-descendants in admixture populations can be less susceptible to the leprosy bacilli, probably because of genetic polymorphisms accumulated during these times, in gene that can modulate the immune response on infection combat. Furthermore African humans are the more genetically diverse population in the world consequently, by selection bias, genetic polymorphisms accumulated that confer protection against disease, can be present in this population and your descendants.

Of point view epidemiological, the situation of African and Americas region is critical, and is associated the socioeconomic challenges related to the disease, but genetics components also are important to disease knowledge [4]. Thus understanding of like genetic ancestry, in

admixture population, can to influence genetic susceptibility is essential to avoid spurious results. In conclusion, our study shows that the *NFKβ1* [rs28362491], *CASP8* [rs3834129], *PAR1* [rs11267092] and *IL4* [rs79071878] genes are possible markers for the susceptibility to development of leprosy and the severe clinical form MB. Moreover, after correcting for population structure within an admixture population, the results show that different levels of ethnic group composition can generate different OR rates for leprosy susceptibility.

Supporting Information

S1 Table. Distribution of percent range of African and European genetic ancestry between leprosy patients and healthy individuals.
(DOCX)

Acknowledgments

We extend special thanks to the sample donors (leprosy patients from northern Brazil) who enabled this study to be carried out.

Author Contributions

Conceived and designed the experiments: PP NPCS SS ÂRdS. Performed the experiments: PP. Analyzed the data: PP NPCS SS. Contributed reagents/materials/analysis tools: CS SS ÂRdS. Wrote the paper: PP CS ÂRdS.

References

1. Alcais A, Mira M, Casanova JI, Schurr E, Abel I. Genetic dissection of immunity in leprosy. Current Opinion in Immunology. 2005; 17: 44–48. PMID: 15653309
2. Elizabeth AM, William RB, James CV, Hawn TR. Leprosy and the Human Genome. Microbiol. Mol. Biol. 2010; 74(4):589–620.
3. Salgado CG, Ferreira DVG, Frade MAC, Guimarães LS, Silva MB, Barreto JG. High Anti—Phenolic Glycolipid-I IgM Titers and Hidden Leprosy Cases, Amazon Region. Emerging Infectious Diseases. 2012; 19: 46–49.
4. World Health Organization (WHO). Global leprosy: update on the 2012 situation. WHO N°35. 2013; 88: 365–380. Available: http://www.who.int/wer/2013/wer8835.pdf?ua=1
5. Ridley DS, Jopling W H. Classification of leprosy according to immunity: a five-group system. Int. J. Leprosy. 1966; 34:255–73.
6. Pinto P, Salgado CG, Santos N, Alencar DO, Santos S, et al. Polymorphisms in the CYP2E1 and GSTM1 Genes as Possible Protection Factors for Leprosy Patients. PLoS ONE. 2012; 7(10): e47498. doi: 10.1371/journal.pone.0047498 PMID: 23077626
7. Foss NT.Immunological aspects of Leprosy. Medicina Ribeirão Preto. 2007; 30: 335–339.
8. Britton WJ, Lockwood DN. Leprosy Lancet. 2004; 363:1209–1219. PMID: 15081655
9. Scollard DM, Adams LB, Gillis TP, Krahenbuhl JL, Truman RW, Williams DL.The continuing challenges of leprosy. Clin. Microbiol. Ver. 2006; 19:338–381.
10. Salgame PM, Bloom YBR, Modlin RL. Evidence for functional subsets of CD4_ and CD8_ T cells in human disease: lymphokine patterns in leprosy. Chem. Immunol. 1992; 54:44–59. PMID: 1358110
11. Yamamura M, Uyemura K, Deans RJ, Weinberg K, Rea TH, Bloom BR,et al. Defining protective responses to pathogens: cytokine profiles in leprosy lesions. Science. 1992; 254:277–279.
12. Garcia P, Alencar D, Pinto P, Santos N, Salgado C, Sortica VA, et al. Haplotypes of the IL10 gene as potential protection factors in leprosy patients. Clin Vaccine Immunol. 2013; Oct; 20(10):1599–603. doi: 10.1128/CVI.00334-13 PMID: 23966553
13. Bobosha K, Wilson L, van Meijgaarden KE, Bekele Y, Zewdie M, et al. T-Cell regulation in Lepromatous Leprosy. PLoS Negl Trop Dis. 2014; 8(4): e2773. doi: 10.1371/journal.pntd.0002773 PMID: 24722473

14. Santos NPC, Ribeiro-Rodrigues EM, Ribeiro-dos-Santos AKC. Assessing individual interethnic admixture and population substructure using a 48 insertion-deletion ancestry informative markers panel. Hum Mutat. 2010; 31(2): 184–90. doi: 10.1002/humu.21159 PMID: 19953531

15. Palha T, Gusmão L, Ribeiro-Rodrigues E, Guerreiro JF, Ribeiro-dos-Santos A, et al. Disclosing the Genetic Structure of Brazil through Analysis of Male Lineages with Highly Discriminating Haplotypes. PLoS ONE. 2012; 7(7): e40007. doi: 10.1371/journal.pone.0040007 PMID: 22808085

16. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data.Genetics. 2000; 155:945–959. PMID: 10835412

17. Sambrook J, Fritsch F, Maniatis T. Molecular Cloning: A Laboratory Manual. Cold SprHarb Laboratory. 1989; NY. 2nd edition.

18. Excoffier L. Laval G., and Schneider S.. Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online. 2005; 1:47–50.

19. Ali S, Hirschfeld AF, Mayer ML, Fortuno ES, Corbett A, Kaplan M, et al. Functional Genetic Variation in NFKBIA and Susceptibility to Childhood Asthma, Bronchiolitis, and Bronchopulmonary Dysplasia. J Immunol. 2013; 190: 3949–3958. doi: 10.4049/jimmunol.1201015 PMID: 23487427

20. Andersen V, Christensen J, Ernst A, Jacobsen BA, Tjønneland TA, Krarup HB, et al. Polymorphisms in NF-κB, PXR, LXR, PPARγ and risk of inflammatory bowel disease. World J Gastroenterol. 2011; 17(2): 197–206. doi: 10.3748/wjg.v17.i2.197 PMID: 21245992

21. Karban AS, Okazaki T, Panhuysen CI, Gallegos T, Potter JJ, Bailey-Wilson JE, et al. Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis. Hum Mol Genet. 2004; 13: 35–45. PMID: 14613970

22. Salim PH, Jobim M, Bredemeier M, Chies JAB, Brenol JCT, Jobim LF, ET al. Interleukin-10 Gene Promoter and NFKB1 Promoter Insertion/Deletion Polymorphisms in Systemic Sclerosis. Scandinavian J. Immunology. 2013; 77:162–168.

23. Cen H, Zhoua M, Lenga RX, Wanga W, Fenga C, Li B et al. Genetic interaction between genes involved in NF-KB signalingpathway in systemic lupus erythematosus. Molecular Immunology. 2013; 56:643–648. doi: 10.1016/j.molimm.2013.07.006 PMID: 23911423

24. Zea AH, Ochoa MT, Ghosh P, Longo DL, Alvord WG, Valderrama L et al. Changes in Expression of Signal Transduction Proteins in T Lymphocytes of Patients with Leprosy. Infection and immunity. 1998; 66(2):499–504. PMID: 9453602

25. Pereira RMS, Calegari-Silva TC, Hernandez MO, Saliba AM, Redner P, Pessolani MCV et al. Mycobacterium leprae induces NF-jB-dependent transcription repression in human Schwann cells. Biochemical and Biophysical Research Communications. 1995; 335: 20–26.

26. Adams MN, Ramachandran R, Yau MK, Suen JY, Fairlie DP et al. Structure, function and pathophysiology of protease activated receptors. Pharmacol Ther. 2011; 130: 248–282. doi: 10.1016/j.pharmthera.2011.01.003 PMID: 21277892.

27. Aerts L, Hamelin M-È, Rhéaume C, Lavigne S, Couture C, et al. Modulation of Protease Activated Receptor 1 Influences Human Metapneumovirus Disease Severity in a Mouse Model. PLoS ONE. 2013; 8(8): e72529. doi: 10.1371/journal.pone.0072529 PMID: 24015257

28. Chionh YT, Ng GC, Ong L, Arulmuruganar A, Stent A, Saeed MA, Wee JLK, Sutton P. Protease-activated receptor 1 suppresses Helicobacter pylori gastritis via the inhibition of macrophage cytokine secretion and interferon regulatory factor 5. Mucosal Immunology. 2014 doi: 10.1038/mi.2014.43

29. Arnaud E, Nicaud V, Poirier O. Protective effect of a thrombin receptor (protease-activated receptor 1) gene polymorphism toward venous thromboembolism. Arterioscler Thromb Vasc Biol. 2000; 20: 585–592 PMID: 10669659

30. Klingler K, Tchou-Wong KM, Brandli O. Effects of mycobacteria on regulation of apoptosis in mononuclear phagocytes. Infect Immun. 1997; 5:5272–5278.

31. Chattree V, Khanna N, Bisht V, Rao DN. Inhibition of apoptosis, activation of NKT cell and upregulation of CD40 and CD40L mediated by M. leprae antigen(s) combined with Murabutide and Trat peptide in leprosy patients. Mol Cell Biochem. 2008; 309:87–97. doi: 10.1007/s11010-007-9646-8 PMID: 18008143

32. Sun T, Gao Y, Tan W, Ma S, Shi Y, et al. A six-nucleotide insertiondeletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers. Nat Genet. 2007; 39: 605–613. PMID: 17450141

33. Teles RMB, Krutzik SR, Ochoa MT, Oliveira RB, Sarno EN, Modlin RL. Interleukin-4 Regulates the Expression of CD209 and Subsequent Uptake of Mycobacterium leprae by Schwann Cells in Human Leprosy. Infection and immunity. 2010; 78(11):4634–4643. doi: 10.1128/IAI.00454-10 PMID: 20713631

34. Nakashima H, Miyake K, Inoue Y. Association between IL4 genotype and IL-4 production in the Japanese population. Genes Immun. 2002; 3:107–9. PMID: 11960309

35. Beitelshees AL, Johnson JA, Hames ML, Gong Y, Cooper-DeHoff RM, et al. Aromatase Gene Polymorphisms Are Associated with Survival among Patients with Cardiovascular Disease in a Sex-Specific Manner. PLoS ONE. 2010 5(12): e15180. doi: 10.1371/journal.pone.0015180 PMID: 21170323

36. Foss NT, Motta ACF. Leprosy, a neglected disease that causes a wide variety of clinical conditions in tropical countries. Mem Inst Oswaldo Cruz. 2012; 107: 28–33. PMID: 23283450

37. Leal AMO, Magalhaes PKR, Souza CS, Foss NT. Adrenocortical hormones and interleukin patterns in leprosy. Parasite Immunol. 2003; 25: 457–461. PMID: 14651593

38. Limer KL, Pye SR, Thomson W, Boonen S, Borghs H, Vanderschueren D et al. Genetic Variation in Sex Hormone Genes Influences Heel Ultrasound Parameters in Middle-Aged and Elderly Men: Results From the European Male Aging Study (EMAS). J. Bone Miner Res. 2009; 24(2): 314–323. doi: 10.1359/jbmr.080912 PMID: 18767927

39. Eidt LM. Breve história da hanseníase: sua expansão do mundo para as Américas, o Brasil e o Rio Grande do Sul e sua trajetória na saúde pública brasileira. Saúde Soc. 2004; 13:76–88.

40. Lastória JC, Abreu MAMM. Leprosy: review of the epidemiological, clinical, and etiopathogenic aspects —Part 1. An Bras Dermatol. 2014; 89(2):205–18. doi: 10.1590/abd1806-4841.20142450 PMID: 24770495

41. Scott HH. The influence of the slave-trade in the spread of tropical disease. Transactions of the society of tropical and hygiene medicine. Trans R Soc Trop Med Hyg. 1943; 37(3):169–188.

42. Ribeiro-dos-Santos AK, Carvalho BM, Feio-dos-Santos AC, Santos SE. Nucleotide variability of HV-I in Afro-descendants populations of the Brazilian Amazon Region. Forensic Sci Int. 2007; 167:77–80. PMID: 16448796

43. Ribeiro-Rodrigues EM, dos Santos NP, dos Santos AK, Pereira R, Amorim A, Gusmão L, Zago MA, Santos SE. Assessing interethnic admixture using na X-linked insertion-deletion multiplex. Am J Hum Biol. 2009; 21:707–709. doi: 10.1002/ajhb.20950 PMID: 19533621

44. Santos SEB, Rodrigues JD, Ribeiro-dos-Santos AK, Zago MA. Differential contribution of indigenous men and women to the formation of an urban population in the Amazon region as revealed by mtDNA and Y-DNA. Am J Phys Anthropol. 1999; 109:175–180. PMID: 10378456

45. Ribeiro-dos-Santos AKC, Carvalho BM, Santos ACF, Santos SEB. Nucleotide variability of HV-I in Afro-descendents populations of the Brazilian Amazon Region. Forensic Science International. 2007; 167: 77–80. PMID: 16448796

46. Han XY, Silva FJ. On the Age of Leprosy. PLoS Negl Trop Dis. 2014; 8(2): e2544. doi: 10.1371/journal.pntd.0002544 PMID: 24551248

S1 Table. Ranges of African and European genetic ancestry contributions in leprosy patients and healthy individuals.

African	**Leprosy patients (N=141)**		**Healthy Individuals (N=180)**			
Range (%) of contribution	Percent	Cumulative Percent	Percent	Cumulative Percent	*p value*[a]	OR (95% CI)[b]
0 ├─ 10	4.3	4.3	1.1	1.1		1
10 ├─ 20	24.1	28.4	41.7	42.8	0.035	0.93 (1.00-0.76)
20 ├─ 30	37.6	66	40	82.8	0.010	0.41 (0.85-0.15)
30 ├─ 40	17	83	12.2	95	0.001	0.38 (0.75-0.10)
40 ├─ 50	9.9	92.9	4.4	99.4	0.005	0.17 (0.82-0.12)
50 ├─ 60	5.7	98.6	0.6	100	-	-
60 ├─ 70	1.4	100	-	-	-	-

European	**Leprosy patients (N=141)**		**Healthy Individuals (N=180)**			
Range (%) of contribution	Percent	Cumulative Percent	Percent	Cumulative Percent	*p value*[a]	OR (95% CI)[b]
0 ├─ 10	1.4	1.4	1.1	1.1		1
10 ├─ 20	2.8	4.3	5	6.1	0.001	1.25 (1.01-1.86)
20 ├─ 30	12.1	16.3	16.1	22.2	0.009	1.63 (1.05-1.98)
30 ├─ 40	22	38.3	33.3	55.6	0.035	1.80 (1.25-2.10)
40 ├─ 50	27.7	66	31.7	87.2	0.026	2.22 (1.85-3.5)
50 ├─ 60	17	83	12.8	100	0.045	2.45 (1.23-12.5)
60 ├─ 70	12.1	95.1	-	-	-	-
70 ├─ 80	5	100	-	-	-	-

[a]*p value* obtained for logistic regression; [b]odds ratio (OR).

3 – CAPÍTULO 2 - GENOME WIDE MICRORNAS IN LEPROSY PATIENTS BY NEXT GENERATION SEQUENCING.

Artigo em preparação para submissão em revista científica

GENOME WIDE MICRORNAS IN LEPROSY PATIENTS BY NEXT GENERATION SEQUENCING.

Pablo Pinto[1,2], Tatiana Vinasco Sandoval[1], André Mauricio Ribeiro dos Santos[1], Fabiano Cordeiro Moreira[2], Luiz Ricardo Goulart[3], Sidney Santos[1,2], Claudio Guedes Salgado[4], Ândrea Ribeiro-dos-Santos[1,2]

[*1]Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, PA, Brasil, 66.075-970.

[*2]Núcleo de Pesquisas em Oncologia - NPO, Universidade Federal do Pará, Belém, PA,Brasil, 66.075-970.

[*3]Laboratório de Nanobiotecnologia, Instituto de Genética e Bioquímica, Universidade Federal de Uberlândia.

[*4]Laboratório de Dermatoimunologia, Instituto de Ciências Biológicas, Universidade Federal do Pará, Belém, PA, Brasil, 66.075-970.

*Corresponding author: Ândrea Ribeiro-dos-Santos

Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federal do Pará

Cidade Universitária Prof. José Silveira Netto – Av. Augusto Corrêa, 01

CEP: 66.0750970, Guamá - Belém, PA, Brasil.

E-mail: akelyufpa@gmail.com or akely@ufpa.br

Abstract

Leprosy is a contagious infectious disease that is caused by *Mycobacterium leprae*, which primarily affects macrophages and Schwann cells in peripheral nerves. Leprosy is characterized by two major types of clinical manifestations. tuberculoid (T) leprosy, and lepromatous (L) leprosy. In Amazon region an epidemiological situation is critical, where approximately 80,000 new cases were reported the last 20 years, and in the state of Pará with annual case detection rate of 50/100,000 peoples, measured in 2012. Evidences show that miRNAs are able to modulated the host antibacterial pathways during the infection process and can influence the outcome of disease. The analysis of miRNA differentially expressed in distinct pole of disease can provide a known about target key to immune response efficient on prevent infection. In the present study was conduct with 17 skin tissue biopsies samples (6 L-pole and 5 T-pole) and 11 (5 L-pole and 4 T-pole) peripheral blood samples of leprosy patients, which were extract RNA, sequencing in next generation sequencing on Illumina platform. The results obtained were normalized and analyzed using the package Bioconductor-*DESeq2* with statistic software R. The differential expression profile from tissue samples reveal 67 miRNAs differentially expression, with 43 downregulated and 24 upregulated and from blood sample were found a total of 10 miRNAs differentially expression with 9 downregulated and one upregulated. Moreover was performed *in silico* target analysis for detect genes involved on pathological of leprosy. In the present, we demonstrated a first genome wide miRNA profile of leprosy in two sample type skin lesions biopsies and peripheral blood of leprosy patients and non-leprosy individuals, and highlight for new miRNA no yet describe, and corroborated same find of literature.

Introduction

Leprosy is a contagious infectious disease that is caused by *Mycobacterium leprae*, which primarily affects macrophages and Schwann cells in peripheral nerves [1]. Leprosy is characterized by two major types of clinical manifestations. Paucibacillary (PB) is primarily characterized by tuberculoid (T) leprosy, with a few lesions and scarce bacilli, and multibacillary (MB) is primarily characterized by anergic lepromatous (L) leprosy, with multiple lesions and many bacilli inside macrophages. In addition, the three following intermediate forms exist: borderline-tuberculoid (BT), borderline-borderline (BB) and borderline-lepromatous (BL) [2]. According the World Health Organization (WHO), MB leprosy includes the BB, BL and LL forms, while PB leprosy encompasses the TT and BT forms [3,4].

In epidemiological world context more than 200,000 cases are registered every year and a great idea of its elimination is questioned [5-7]. the Brazil has the highest burden of disease on American continent and the second most affected country in the world, with 31,064 new cases reported in 2014 and case detection rate of 15.4/100.000 people [8]. In Amazon region, north of the country, an epidemiological situation is critical, where approximately 80,000 new cases were reported the last 20 years, and in the state of Pará this is an old problem, with annual case detection rate of 50/100,000 peoples, measured in 2012, three times the national average [9].

Exist consensus to affirm that detection and treatment would reduce transmission of leprosy, and an idea of prophylaxis like way. Thus new tools are needed for detection of new cases, include subclinical infections [10]. There is no laboratory test that detects all forms of leprosy, but some biomarkers have been developed since the isolation and characterization in the 1980s of phenolic Glycolipid-I (PGL-1) [11,12]. Serology could potentially be used to detect antibodies against PGL-I to assist the classification of patients for treatment and monitoring, identify the risk of relapse and the healthy household contacts of leprosy patients who are most at risk of contracting the disease, besides also a marker of subclinical infection [9,10]. However, evidences do not straightforward, with variation in the validity this test as a predictor of who will contracting leprosy [10].

Genetics studies of portions of genome that not encode protein, revealed one class de small non-coding RNAs (named microRNAs or miRNAs) involved in control post-transcriptional of gene expression [13]. The knowledge about the interaction between miRNA and leprosy disease is limited, recent study demonstrated that miRNA can influence the mechanism whereby the cell host can prevent the bacillary growth and generate natural barriers against infection by *M. leprae* [14]. Evidences show that miRNAs are able to modulated the host antibacterial pathways during the infection process and can influence the outcome of disease. The analysis of miRNA differentially expressed in distinct pole of disease can provide a known about target key to immune response efficient on prevent infection, besides can generate novels possible biomarkers to leprosy, as well as to subclinical infection and one possible predictor of who will contracting leprosy [15, 16].

In view of this new field of study on Leprosy, with possibilities of increase the knowledge about the mechanism that involve the progression and installation of disease, as to point to possible new markers to diagnose, our study aim performs the first genome wide miRNA profiling of Leprosy by next generation sequencing (NGS), assessing and describing the expression standard in two distinct sample, skin tissue and peripheral blood of leprosy patients.

Methods

Samples

A total of 28 biological samples from leprosy patients and individuals without leprosy, who attended the Dr Marcello Candia Reference Unit in Sanitary Dermatology of the State of Pará (UREMC), in Marituba, Pará, Brazil, treatment virgin, were included in the present study: a) 17 tissue biopsies samples (11 from from leprosy patients [6 L-pole and 5 T-pole] and 6 ephitelial tissue from individuals without leprosy for controls); b) 11 peripheral blood samples (9 from from leprosy patients [5 L-pole and 4 T-pole] and 2 from individuals without leprosy for controls). This study adhered to the Declaration of Helsinki and was approved by the Institute of Health Sciences Research Ethics Committee at Universidade federal do Pará. A written informed consent to publish was obtained from every individual who accepted to participate in this study. (No. 197/07 UFPA)

Total RNA storage, extraction and quantification

The peripheral blood sample were collected in Tempus Blood RNA Tube (Thermo Fisher Scientific, US) and storage at -20ºC until extraction. The skin tissue biopsies samples were collected in propylene tube of 2 mL with RNAlater (Thermo Fisher Scientific, US) and storage in liquid nitrogen until extraction.

Extraction of tissue sample total RNA was performed using Trizol reagent (Ivitrogen, US), and samples were eluate in DEPC water and storage in liquid nitrogen. For extraction of blood sample total RNA was performed using MagMAX RNA Isolation Kit (Thermo Fisher Scientific, US)

Total RNA quantification and quality were measured with Nanodrop 1000 (Thermo Scientific, US) and Agilent 2200 TapeStation (Agilent Technologies, US).

Library preparation and Next Generation Sequencing

The library was prepared using TruSeq Small RNA Library Preparation Kit (Illumina, Inc, US), according to the manufacturer's instructions, and all samples using for library, obtained initial concentration of 1 µg/5µL of total RNA

The library was validated and quantifications with Agilent 2200 TapeStation (Agilent Technologies, US) platform, and by real time PCR with KAPA Library Quantification Kit (KAPABIOSYSTEM, USA). After, the libraries were diluted for 4 nM concentration and sequencing using MiSeq Reagent Kit v3 150 cycle (Illumina, Inc, US) on MiSeq System (Illumina Inc, US). The tissue and blood sample, were sequencing separately.

Sequencing data processing and Analysis-Small RNA-Seq Pipeline

The data of sequencing, were processed on Illumina MiSeq reporter and extract in FASTQ format. A pipeline of pre-processing using Fastx_toolkit was applied for filter low quality, Trimer the extremes 3' reads obtained and removal contaminants. The pipeline was performed according the chronogram: a) Quality score average Phred (Q) greater than 30, b) Reads with more of 17 nucleotides of longitude, c) A base calling error probabilities (P) greater than 80.

Next, was performed the reads alignment with Human genome (GRCh37) in combination with microRNA data base (MirBase v.19) using STAR (Spliced Transcripts Alignment to a Reference). The score of microRNA was performed with htseq-count toll, results obtained were normalized and analyzed using the package Bioconductor-*DESeq2* (Love MI. Et al, 2014) with statistic software R. Thus were conducted following comparisons: a) Leprosy patient's *vs* Individuals without leprosy; b) T-pole Leprosy *vs vs* Individuals without leprosy; c) L-pole Leprosy *vs vs* Individuals without

leprosy; d) T-pole *vs* L-pole Leprosy. Values of *p adjusted* $\leq$ 0,05 e log2*fold change* > 2 were considered statistically significant.

In silico Target genes identification

The identification of targets regulated by miRNA differentially expressed detected at the analysis, were performed using four toll: i) TargetCompare (lghm.ufpa.br/targetcompare); ii) miRTarBase (http://mirtarbase.mbc.nctu.edu.tw); iii) DIANA miRPath v.3 (Vlachos et al., 2012) and iv) TargetScan (Agarwal et al., 2015). The Target selected share more than one miRNA and they are related with a pathological of leprosy.

Results

This study evaluated different types of leprosy samples, skin biopsies and peripheral blood, in distinct next generation sequencing. For two sample types were analyzed the microRNA differential expression profile, aiming identify possible leprosy biomarkers, that assist the understand of pathology, progression and risk factor of disease.

miRNA sequencing and differential expression profile from Tissue samples.

This sequencing a yield of 4 million of reads were obtained. After the process pipeline, more than 96% of reads were alignment with Human genome and the count of miRNA was performed using *htseq-count* (miRNA count $\geq$ 10), with average of 36.745 reads per sample and 656 miRNAs expressed at least one sample.

The analysis of differential expression was conducted comparisons between a) Leprosy patient's *vs* Individuals without leprosy; b) T-pole Leprosy *vs* Individuals

without leprosy; c) L-pole Leprosy *vs* Individuals without leprosy. Ergo were found a total of 67 miRNAs differentially expression, with 43 downregulated and 24 upregulated (Table 1, and S1).

For all miRNA differentially expressed, a heatmap was constructed using the expression in RPKM (Reads Per Kilobase per Million mapped reads) and was visualized a clustering and standards of expression able to differential leprosy patients of Individuals without leprosy, and T-pole of L-pole (Fig 1).

The Figure 2 show the 39 miRNAs, according by fold change, presents on three comparisons conducted like demonstrate in Table 1, with 24 downregulated e 15 upregulated (Fig 2.)

The analysis performed between the extreme pole of leprosy, T and L, five miRNAs differentially expressed were obtained, of which three were downregulated and two upregulated. The *hsa-miR-362-5p* is only present in this comparisons (Fig. 3)

For in silico target genes identification, among the miRNAs differentially expressed obtained on sequencing, were selected those with p adjusted $< 1x10^{-5}$ for a prediction of target gene, and were selected targets regulated for more than one miRNA or validated for western blot, reporter assay or qPCR. Thus were found that the 12 miRNAs more differentially expressed may be regulating 53 target genes (Fig 4), in common these genes are involved in modulation inflammatory.

miRNA sequencing and differential expression profile from peripheral blood samples.

This sequencing a yield of 6 million of reads were obtained. After the process pipeline, more than 95% of reads were alignment with Human genome and the count of miRNA was performed using *htseq-count* (miRNA count ≥ 10), with average of 371.325 reads per sample and 527 miRNAs expressed at least one sample.

The analysis of differential expression was conducted comparisons between a) Leprosy patient's *vs* Individuals without leprosy; b) T-pole Leprosy *vs vs* Individuals without leprosy; c) L-pole Leprosy *vs vs* Individuals without leprosy. Ergo were found a total of 10 miRNAs differentially expression (Fig 5), with nine downregulated and one upregulated (Table 2, and S2).

With the miRNA differentially expressed identified, a heatmap was constructed using the expression in RPKM and was visualized a clustering and standards of expression able to differential leprosy patients of Individuals without leprosy (Fig 6), but when the comparisons were performed between the T-pole *vs* L-pole, not miRNA differentially expressed were found.

For *in silico* target genes identification, the miRNAs differentially expressed obtained on sequencing were selected for a prediction of target gene, and were selected targets regulated for more than one miRNA or validated for western blot, reporter assay or qPCR. Thus were found that the 10 miRNAs more differentially expressed may be regulating 48 target genes (Fig 7), in common these genes too are involved in modulation inflammatory.

Discussion

Our study found miRNAs differentially expression profile in two types of leprosy samples, and showed like it's have ability to differentiate leprosy patients and individuals without leprosy, as well as got discern specific miRNAs clustering of T-pole or L-pole leprosy. Besides showed the target genes regulated by these miRNAs, which be involved in various inflammatory process, modulatory mechanism of host immunity, include immune response against leprosy, and reveal common targets in skin tissue and peripheral blood miRNA profile, regulated by different miRNAs. Thus reveal new miRNAs not describes, as a profile of blood samples that can show possible new markers of disease.

The evaluated of expression of microRNA in skin biopsies, by microarray method, in distinct leprosy lesion (T-pole *vs* L-pole) showed differentially expression in 16 miRNAs (*hsa-miR-21, hsa-miR-146a, hsa-miR-30a, hsa-miR-22, hsa-miR-30b, hsa-miR-451, hsa-miR-34, hsa-miR-181b, hsa-miR-181a, hsa-miR-155, hsa-miR-29c, hsa-miR-24* e *hsa-miR-638*). The *hsa-miR-21* was the more differentially expressed like upregulated in L-pole, and was demonstrated inhibit antimicrobial mechanism D-vitamin dependent and stop production of *CAMP* e *DEFB4A*, two peptides with anti-leprosy activity. Besides *hsa-miR-21* can modulated the *TLR2/1* (receptor of toll-like family) and generate an increase of expression of IL10 that decrease the host ability to combat the *M. leprae* infection [14].

Additionally, *hsa-miR-21* can decrease IL12 expression, an important cytokine that is stimulated on inflammation and improve Th1 host immune response, and generate a standard of control to infection by *M. leprae*. The *hsa-miR-21* also inhibit *BCL2* (B-cell lymphoma) and can be associated with sensitive motor sequel in leprosy [17, 18]. Further the *hsa-miR-181a* when downregulated permit an expression of *SHP2* (Tyrosine-kinase receptor family) that decrease a T-cell activity in response to *M. leprae* and reduction the host defense mechanism against infection [17].

Tuberculosis study model in vitro, demonstrated that high expression of *hsa-miR-125b* and low expression of *hsa-miR-155* can low TNF (Tumoral Necrosis Factor) production, with *hsa-miR-125b* act directly on mRNA of TNF, and *hsa-miR-155* acting in Bach1 (transcriptional repressor of heme oxygenase-1) and SHIP1 (inhibitor of activation of the serine/threonine kinase AKT), take together this events are critical on modulation of host innate immunity and protection against infection [17, 16]. Besides *hsa-miR-99b* can also modulate TNF expression and affect the cell power in combat the bacterial growth [19], and overexpression of *has-miR-144-5p* can modulated the T cells proliferation and inhibit TNF and IFN-γ, that modulated the anti-tuberculosis immunity and support the disease development [20]. Futher a study of rs2910164 in *hsa-miR-146a*, in nerve biopsies sample of leprosy patients, reveal that high expression this miRNA is correlated with lowest level of TNF and minor capacity of combat to *M. leprae* [21]

A recent study demonstrated that the microRNA *let-7f* modulates the immune response against *Mycobaterium tuberculosis* (Mtb), and has as target the protein A20, which inhibits the NFκβ pathway and contributes to survival of the infected cell. During infection process, microRNA *let-7f* is downregulated, which results in the decrease of expression of key genes to adequate immune response, like *TNF, IL1β* and *IL6*, ergo Mtb has a larger survival. Further, let-7f overexpression diminishes Mtb survival and augments the production of cytokines including TNF and IL-1β [22].

Our analysis showed an overview of profile of all miRNA differentially express in skin tissue and blood peripheral samples of virgin treatment patients, and realized comparisons include individuals without leprosy, therefore the first case-control revealing miRNA genome wide of leprosy. In tissue sample a high diversities of miRNA were found and able to differentiate leprosy of non-leprosy, as the distinct pole T or L of non-leprosy and the pole in itself. Thus can classify the miRNA: i) involved on progression and manifestation of leprosy (for those that be present on Leprosy patient's *vs* non-leprosy, T-pole Leprosy *vs* non-leprosy and L-pole Leprosy *vs* non-

leprosy comparisons); ii) markers of clinical form Tuberculoide (for those that be present on Leprosy patient's *vs* non-leprosy and T-pole Leprosy *vs* non-leprosy comparisons); iii) markers of clinical form Lepromatous (for those that be present on Leprosy patient's *vs* non-leprosy and L-pole Leprosy *vs* non-leprosy comparisons), iv) markers specific of Lepromatous (for those that be present on L-pole Leprosy *vs* non-leprosy comparisons only) and v) markers of clinical form (for those that be present on T-pole Leprosy *vs* L-pole Leprosy comparisons).

The findings of miRNAs differentially express in our study corroborate with data of literature like *has-miR-146a* [14,21], *has-miR-34* [14], *has-miR-155* [14,16,17], *has-miR-29c* [14], *has-miR-125b* [17,16], *has-miR-99* [19], that act regulated key genes to establishment of adequate immune response, generate conditions to host cell combat the infection and bacillary growth, and doing a micro-environment that can lead to remiss and elimination of *M. leprae* or to leprosy. Besides the *in silico* target analysis proposed by us (Fig 4), showed same important validated markers that modulated immune response against leprosy like *IL1β, IL6, IL8, TLR2, TLR4, IL17RB, IFNGR1* and *TGFBR1* that regulated a sensitive balance between pro-inflammatory or humoral immunity and contribute for establishment of leprosy and your clinical form T-pole or L-pole, additionally same important mediators like *NFκβ, SMAD* family, *STAT3*, that regulated the cellular signaling and expression in inflammatory process like showed above. Thus this overview of differentially expression profile can highlight new perspectives e knowledge about the leprosy and host mechanism of defense.

In peripheral blood sample a low diversities of miRNA were found, but able to differentiate leprosy of non-leprosy, as the distinct pole T or L of non-leprosy, and showed an potential field for search of new biomarkers. In this context an classify the miRNA: i) involved on progression and manifestation of leprosy (for those that be present on Leprosy patient's *vs* non-leprosy, T-pole Leprosy *vs* non-leprosy and L-pole Leprosy *vs* non-leprosy comparisons); ii) markers of clinical form Tuberculoide (for those that be present on Leprosy patient's *vs* non-leprosy and T-pole Leprosy *vs* non-leprosy comparisons); iii) markers of clinical form Lepromatous (for those that be

present on Leprosy patient's *vs* non-leprosy and L-pole Leprosy *vs* non-leprosy comparisons), iv) markers specific of Tuberculoide (for those that be present on T-pole Leprosy *vs* non-leprosy comparisons only) and v) markers of leprosy (for those that be present on Leprosy *vs* non-leprosy comparisons only).

The findings of miRNAs in blood (Fig 7 and S2) in our study corroborate with two data of literature the *has-miR-144-5p* [20] and *has-let-7f-5p* [22], that act regulated T cells differentiation and NFκβ pathway respectively, that demonstrate importance to survival of viable mycobacterial in host cell, and to mechanism like the innate immune response will perform your defense against infection. The *in silico* target analysis reveal targets validated markers like *NFκβ, CASP8* and *CYP19A1*, which were describe in a genetic study of case-control, like associated to development of disease [23], this corroborate the idea that these miRNA regulated genes already validated in population study and high the your evidence like potential biomarkers for leprosy, besides also regulated *IL12* and *BCL-2* that were showed how very important to improve Th1 host immune response [17,18].

Differentially expression profile in skin tissue and blood have different standard of miRNA and exhibit different markers on comparisons conducted, but can converge in some important issues, both seem regulated a marker in common, the NFκβ, studies how show above demonstrated an importance this mediator for establishment of an effective host immunity against infection by *M. leprae*. So NFκβ necessity be more study in leprosy to know like it can assist to discovery of new markers for both the academic search like clinical practices. Further the *has-miR-1291* was the only differentially express in skin tissue and peripheral blood, so more study about this miRNA are necessary to validated us like potential biomarker for leprosy.

In summarize, we demonstrated a first genome wide miRNA profile of leprosy in two sample type skin lesions biopsies and peripheral blood of leprosy patients and non-leprosy individuals, and highlight for new miRNAs no yet describe, and corroborated same find of literature. We understand that this study showed a new way for study of new biomarkers of leprosy, that assist to development of diagnostics

measures with objective of identify precociously who will sicken, like healthy household contacts of patients, individual in subclinical infection, to improve the diagnostic and combat to leprosy.

Reference

1- A, Mira M, Casanova JI, Schurr E, Abel I (2005) Genetic dissection of immunity in leprosy. Current Opinion in Immunology 17: 44–48.

2- Ridley DS, Jopling WH (1966) Classification of leprosy according to immunity: a five-group system. Inernational Journal of Leprosy Other Mycobact. Disease 34:255–273.

3- WHO. 2009. Global leprosy situation. Wkly. Epidemiol. Rec. 84:333–340.

4- Foss NT (2007).Immunological aspects of Leprosy.MedicinaRibeirãoPreto; 30: 335-339.

5- CUNHA Sergio Souza da, BIERRENBACH Ana Luiza, BARRETO Vitor Hugo Lima. Chemoprophylaxis to control leprosy and the perspective of its implementation in brazil: a primer for non-epidemiologists. Rev. Inst. Med. trop. S. Paulo; 57(6): 481-487.

6- Khan Burki T. Leprosy and the rhetoric of elimination. BMJ. 2013;347:f6142

7- Lockwood DN, Shetty V, Penna GO. Hazards of setting targets to eliminate disease: lessons from the leprosy elimination campaign. BMJ. 2014;348:g1136.

8- Araujo S, Rezende, MMF, S DCR, Rosa MR, Santos DC, Goulart LR and Goulart IMB. Risk-benefit assessment of Bacillus Calmette-Guérin vaccination, anti-phenolic glycolipid I serology, and Mitsuda test response: 10-year follow-up of household contacts of leprosy patients. Rev. Soc. Bras. Med. Trop., Uberaba , v. 48, n. 6, p. 739-745, Dec. 2015.

9- Barreto JG, Bisanzio D, Frade MAC, et al. Spatial epidemiology and serologic cohorts increase the early detection of leprosy. BMC Infectious Diseases. 2015;15:527. doi:10.1186/s12879-015-1254-8.

10- Penna MLF, Penna GO, Iglesias PC, Natal S, Rodrigues LC. Anti-PGL-1 Positivity as a Risk Marker for the Development of Leprosy among Contacts of Leprosy Cases: Systematic Review and Meta-analysis. Small PLC, ed. PLoS Neglected Tropical Diseases. 2016;10(5):e0004703.

11- Hunter SW, Brennan PJ. A novel phenolic glycolipid from Mycobacterium leprae possibly involved in immunogenicity and pathogenicity. J Bacteriol. 1981;147(3):728–35.

12- Spencer JS, Brennan PJ. The role of Mycobacterium leprae phenolic glycolipid I (PGL-I) in serodiagnosis and in the pathogenesis of leprosy. Lepr Rev. 2011;82(4):344–57.

13- Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function.Cell 2004, 116(2):281–297.

14- Liu PT, Wheelwright M, Teles R, Komisopoulou E, Edfeldt K, Ferguson B, Mehta MD, Vazirnia A, Rea TH, Sarno EN, Graeber TG, Modlin RL: MicroRNA-21 targets the vitamin D-dependent antimicrobial pathway in leprosy. Nat Med 2012, 18(2):267–273

15- Mehta MD, Liu PT. microRNAs in mycobacterial disease: friend or foe?*Frontiers in Genetics*. 2014;5:231. doi:10.3389/fgene.2014.00231.

16- Singh PK, Singh AV, Chauhan DS. Current understanding on micro RNAs and its regulation in response to Mycobacterial infections. *Journal of Biomedical Science*. 2013;20(1):14. doi:10.1186/1423-0127-20-14.

17- Kumar R, Halder P, Sahu SK, Kumar M, Kumari M, Jana K, Ghosh Z, Sharma P, Kundu M, Basu J: Identification of a novel role of ESAT-6-dependent miR-155 induction during infection of macrophages with Mycobacterium tuberculosis. Cell Microbiol 2012, 14(10):1620–1631.

18- Wu Z, Lu H, Sheng J, Li L: Inductive microRNA-21 impairs antimycobacterial responses by targeting IL-12 and Bcl-2. FEBS Lett 2012, 586(16):2459–2467.

19- Singh Y, Kaul V, Mehra A, Chatterjee S, Tousif S, Dwivedi VP, Suar M, Kaer LV, Bishai WR, Das G: Mycobacterium tuberculosis controls miR-99b expression in infected murine dendritic cells to modulate host immunity. J Biol Chem 2012. Epub ahead of print.

20- Liu Y, Wang X, Jiang J, Cao Z, Yang B, Cheng X: Modulation of T cell cytokine production bymiR-144* with elevated expression in patients with pulmonary tuberculosis. Mol Immunol 2011, 48(9–10):1084–90.

21- Cezar-de-Mello PFT, Toledo-Pinto TG, Marques CS, Arnez LEA, Cardoso CC, et al. (2014) Pre-miR-146a (rs2910164 G.C) Single Nucleotide Polymorphism Is Genetically

and Functionally Associated with Leprosy. PLoS Negl Trop Dis 8(9): e3099. doi:10.1371/journal.pntd.0003099

22- Kumar M, Sahu SK, Kumar R, Subuddhi A, Maji RK, Jana K, Gupta P, Raffetseder J, Lerm M, Ghosh Z, van Loo G, Beyaert R, Gupta UD, Kundu M, Basu J. MicroRNA let-7 modulates the immune response to Mycobacterium tuberculosis infection via control of A20, an inhibitor of the NF-κB pathway. Cell Host Microbe. 2015 Mar 11;17(3):345-56. doi: 10.1016/j.chom.2015.01.007.

23- Pinto P, Salgado C, Santos NPC, Santos S, Ribeiro-dos-Santos Â. Influence of Genetic Ancestry on INDEL Markers of NFKβ1, CASP8, PAR1, IL4 and CYP19A1 Genes in Leprosy Patients. Johnson C, ed. PLoS Neglected Tropical Diseases. 2015;9(9):e0004050. doi:10.1371/journal.pntd.0004050.

Table 1. Number of differential expression profile between the all comparisons.

Analysis	miRNAs	miRNAs downregulated	miRNAs upregulated
Leprosy patient's *vs* Individuals without leprosy	43	26	17
T-pole Leprosy *vs* Individuals without leprosy	14	7	7
L-pole Leprosy *vs* Individuals without leprosy	60	41	19

Table 2. Number of miRNAs differential expression profile between the all comparisons.

Analysis	miRNAs	miRNAs downregulated	miRNAs upregulated
Leprosy patient's *vs* Individuals without leprosy	7	7	0
T-pole Leprosy *vs* Individuals without leprosy	5	5	0
L-pole Leprosy *vs* Individuals without leprosy	4	3	1

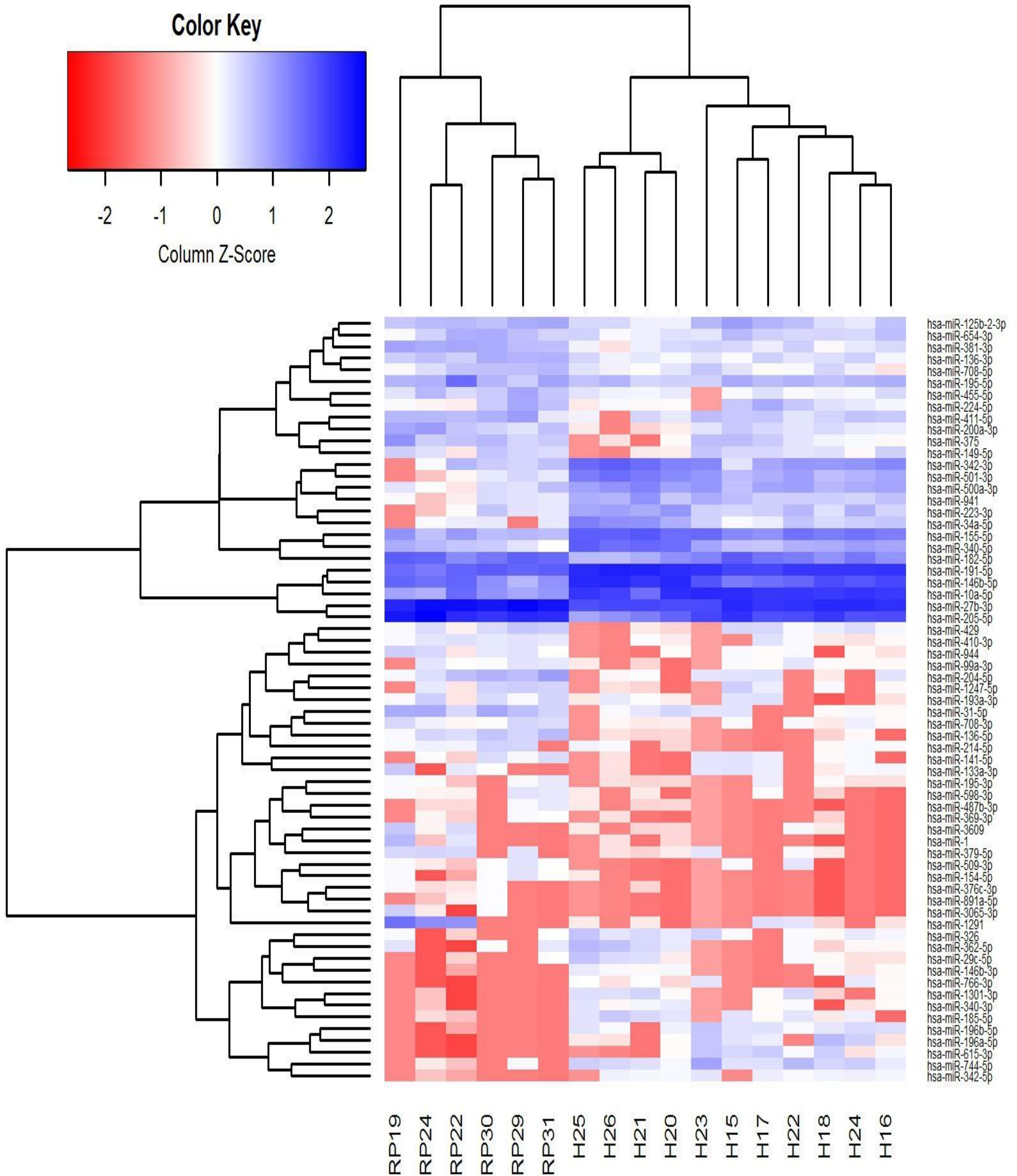


Fig. 1: Heatmap in log2 RPKM of 67 miRNAs deferentially expressed in Leprosy (RP: Individuals without leprosy, H: Leprosy patients with H25, H26, H21, H20, H23, H15 belonging to L-pole).

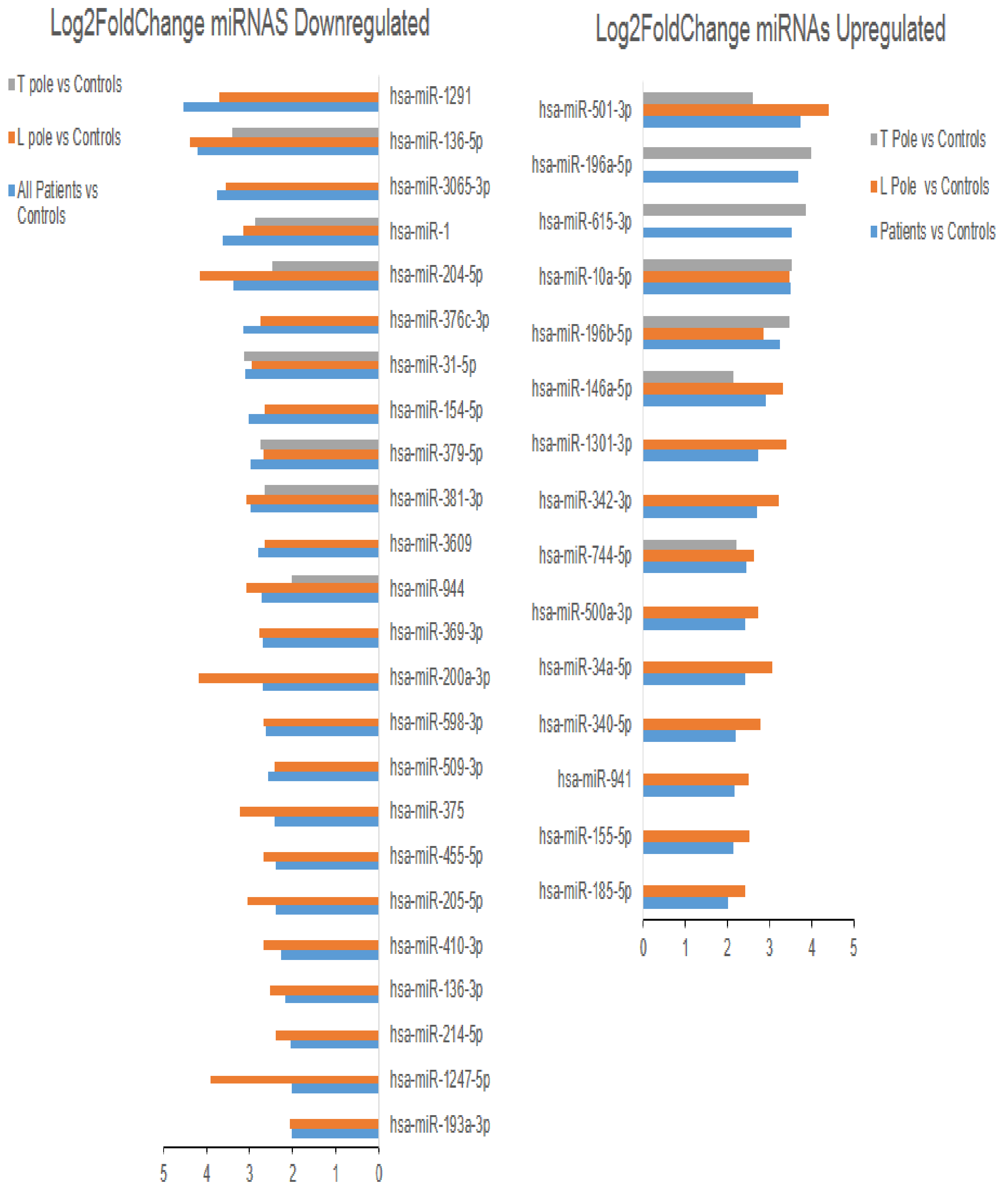


Fig 2: miRNAs differentially expressed according your profile in down or up regulated in three distinct comparisons (data in |log2FoldChange| >2 e p *adjusted* ≤ 0,05

Fig. 3: Volcano plot of miRNAs differentially expressed between T-pole *vs* L-pole.

Fig. 4: Targets found for 12 miRNAs in tissue sample (red box: validated miRNAs like regulators of targets; gray box: predicted miRNAs like regulators of targets).

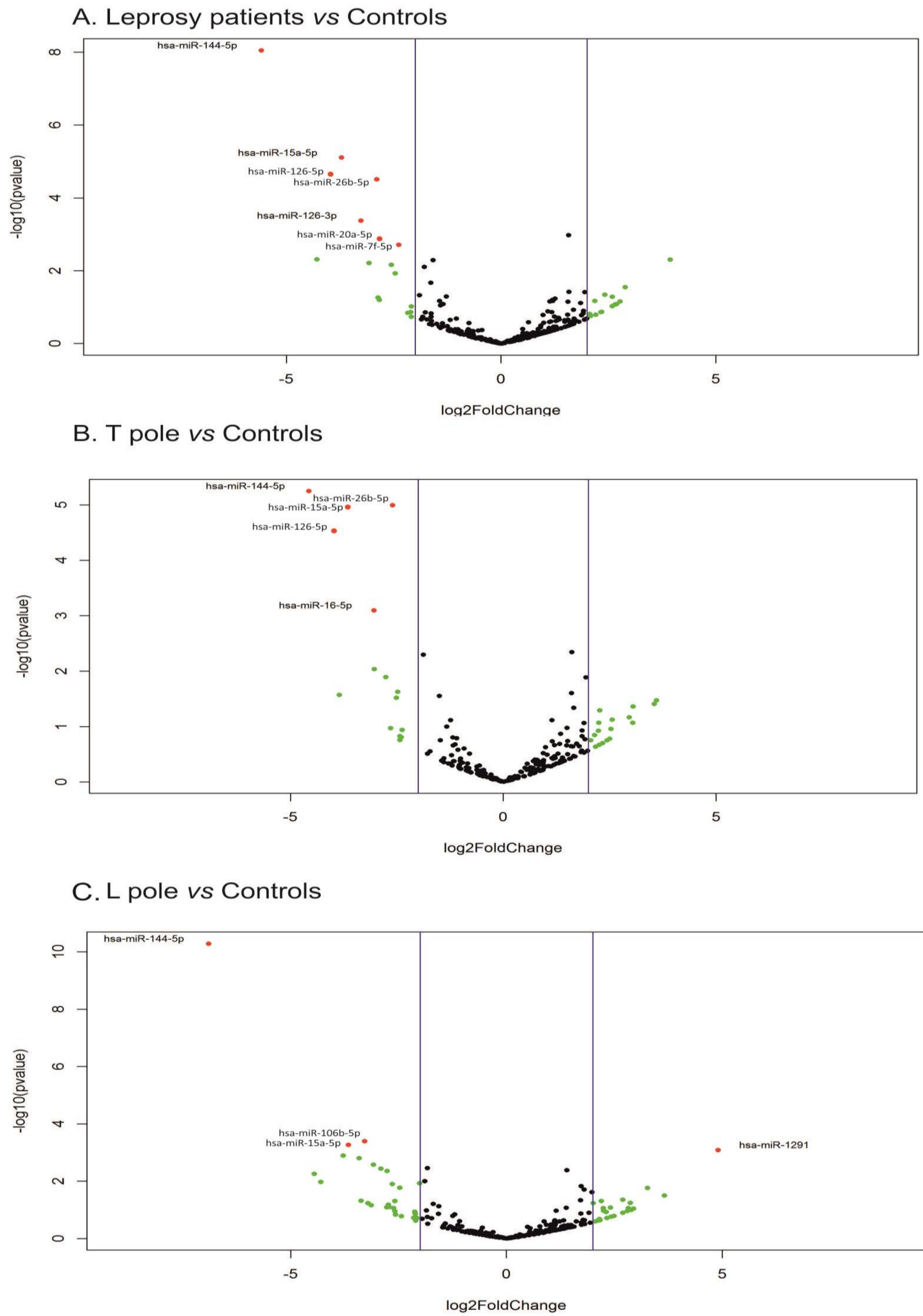


Fig. 5. Volcano plot for miRNAs differentially expressed in all comparisons from blood samples.

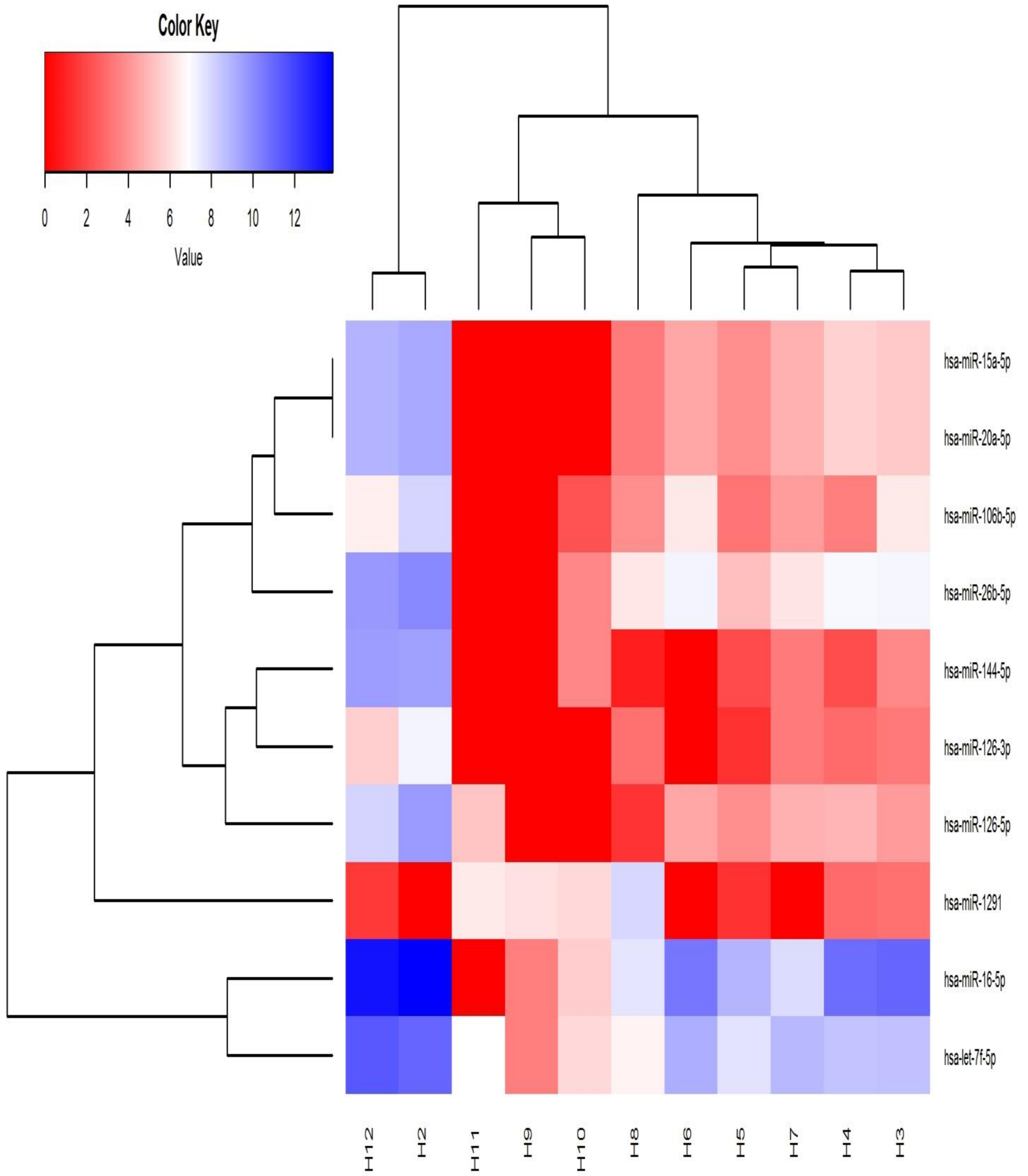


Fig. 6: Heatmap in log2 RPKM of 10 miRNAs deferentially expressed in Leprosy (H12 and H2: Individuals without leprosy).

Fig. 7: Targets found for 10 miRNAs in tissue sample (red box: validated miRNAs like regulators of targets; gray box: predicted miRNAs like regulators of targets).

S1. Profile of differential expression for four comparisons conduct in Tissue sample.

	Leprosy vs Individual without Leprosy		T-pole vs Individual without Leprosy		L-pole vs Individual without Leprosy		L-pole vs T-pole Leprosy	
microRNAs	Fold Change*1	*P*adj*2	Fold Change	*p*adj	Fold Change	*p*adj	Fold Change	*p*adj
hsa-miR-1	-3,64	3,63E-03	-2,87	2,90E-02	-3,16	1,61E-02		
hsa-miR-10a-5p	3,51	4,48E-13	3,53	1,23E-12	3,47	1,53E-08		
hsa-miR-1247-5p	-2,01	2,32E-02			-3,90	1,55E-04		
hsa-miR-1291	-4,53	9,08E-05			-3,70	1,16E-02		
hsa-miR-1301-3p	2,72	2,58E-02			3,39	3,87E-03		
hsa-miR-136-3p	-2,17	3,53E-08			-2,53	3,52E-08		
hsa-miR-136-5p	-4,22	1,04E-09	-3,41	8,51E-05	-4,39	2,11E-07		
hsa-miR-146a-5p	2,92	8,95E-07	2,15	2,07E-03	3,32	1,88E-08		
hsa-miR-154-5p	-3,02	1,63E-02			-2,64	4,88E-02		
hsa-miR-155-5p	2,13	1,27E-06			2,52	3,54E-10		

hsa-miR-185-5p	2,02	3,89E-02			2,43	9,35E-03		
hsa-miR-193a-3p	-2,01	1,06E-02			-2,07	1,28E-02		
hsa-miR-196a-5p	3,69	6,06E-03	3,99	5,62E-04				
hsa-miR-196b-5p	3,25	5,02E-03	3,47	9,34E-04	2,85	2,83E-02		
hsa-miR-200a-3p	-2,70	2,12E-04			-4,19	1,97E-07	2,09	5,81E-03
hsa-miR-204-5p	-3,38	1,59E-06	-2,48	1,88E-03	-4,17	2,78E-07		
hsa-miR-205-5p	-2,40	3,54E-04			-3,04	7,41E-05		
hsa-miR-214-5p	-2,05	3,65E-03			-2,39	4,69E-03		
hsa-miR-3065-3p	-3,75	6,17E-03			-3,56	1,55E-02		
hsa-miR-31-5p	-3,11	9,29E-08	-3,12	4,28E-06	-2,96	6,00E-05		
hsa-miR-340-5p	2,18	6,88E-03			2,79	1,38E-04	-2,18	2,32E-03
hsa-miR-342-3p	2,70	1,79E-04			3,23	3,55E-07		
hsa-miR-342-5p	2,16	4,48E-02						

hsa-miR-34a-5p	2,43	6,17E-03			3,06	7,76E-05	-2,16	3,56E-03
hsa-miR-3609	-2,80	6,88E-03			-2,65	1,67E-02		
hsa-miR-369-3p	-2,71	1,95E-02			-2,78	3,26E-02		
hsa-miR-375	-2,43	2,38E-03			-3,24	2,42E-04		
hsa-miR-376c-3p	-3,16	9,99E-03			-2,76	3,49E-02		
hsa-miR-379-5p	-2,99	7,59E-04	-2,76	1,60E-02	-2,67	4,58E-03		
hsa-miR-381-3p	-2,97	4,48E-13	-2,66	1,12E-07	-3,09	3,98E-09		
hsa-miR-410-3p	-2,27	7,59E-04			-2,67	7,40E-04		
hsa-miR-455-5p	-2,40	1,27E-06			-2,68	1,20E-05		
hsa-miR-487b-3p	-2,41	2,69E-02						
hsa-miR-500a-3p	2,43	1,88E-06			2,74	1,26E-07		
hsa-miR-501-3p	3,72	5,11E-08	2,60	1,03E-05	4,41	1,04E-13		
hsa-miR-509-3p	-2,58	1,28E-02			-2,43	3,29E-02		

hsa-miR-598-3p	-2,63	3,34E-03			-2,66	8,84E-03	
hsa-miR-615-3p	3,52	8,95E-03	3,86	8,38E-04			
hsa-miR-744-5p	2,44	2,35E-03	2,23	1,21E-02	2,63	2,72E-03	
hsa-miR-766-3p	2,78	4,47E-02					
hsa-miR-891a-5p	-3,42	1,22E-02					
hsa-miR-941	2,17	6,71E-04			2,49	1,63E-04	
hsa-miR-944	-2,72	6,94E-04	-2,01	3,94E-02	-3,07	1,52E-03	
hsa-miR-125b-2-3p					-2,08	8,08E-06	
hsa-miR-133a-3p					-2,62	4,67E-02	
hsa-miR-141-5p					-3,01	5,78E-03	
hsa-miR-146b-5p					2,42	3,14E-04	
hsa-miR-149-5p					-2,27	5,18E-03	
hsa-miR-182-5p					-2,32	1,54E-04	

hsa-miR-191-5p					2,23	2,37E-18		
hsa-miR-195-3p					-2,24	3,87E-02		
hsa-miR-195-5p					-2,22	1,90E-03		
hsa-miR-199a-3p					-2,36	5,60E-09		
hsa-miR-223-3p					2,34	7,40E-04		
hsa-miR-224-5p					-2,03	2,02E-02		
hsa-miR-27b-3p					-2,51	1,88E-12		
hsa-miR-29c-5p					2,20	3,53E-02		
hsa-miR-326					2,18	9,94E-03		
hsa-miR-340-3p					2,74	2,79E-02		
hsa-miR-411-5p					-2,55	1,22E-04		
hsa-miR-429					-3,50	7,09E-05	2,43	3,56E-03
hsa-miR-654-3p					-2,14	4,98E-04		

hsa-miR-708-3p					-2,40	1,59E-03		
hsa-miR-708-5p					-2,12	1,04E-04		
hsa-miR-889-3p					-2,47	2,58E-02		
hsa-miR-99a-3p					-2,73	3,74E-03		
hsa-miR-362-5p							-2,12	1,90E-02

*1. Log2 *fold change*; *2 *p* adjusted by Bomferroni.

S2. Profile of differential expression for three comparisons conduct in blood sample.

	Leprosy vs Individual without Leprosy		**T-pole vs Individual without Leprosy**		**L-pole vs Individual without Leprosy**	
miRNAs	Fold Change*1	Padj*2	Fold Change	padj	Fold Change	padj
hsa-let-7f-5p	-2,38	4,15E-02				
hsa-miR-126-3p	-3,27	1,44E-02				
hsa-miR-126-5p	-3,97	1,29E-03	-3,94	9,64E-04		
hsa-miR-144-5p	-5,58	1,52E-06	-4,57	5,05E-04	-6,91	1,03E-08
hsa-miR-15a-5p	-3,72	6,70E-04	-3,61	5,05E-04	-3,67	3,58E-02
hsa-miR-20a-5p	-2,83	3,27E-02				
hsa-miR-26b-5p	-2,90	1,32E-03	-2,60	5,05E-04		
hsa-miR-106b-5p					-3,29	3,58E-02
hsa-miR-1291					4,90	4,07E-02
hsa-miR-16-5p			-3,04	2,18E-02		

*1. Log2 *fold change*; *2 *p* adjusted by Bomferroni.

4 - DISCUSSÃO GERAL

No Brasil, os primeiros casos de hanseníase foram notificados no ano de 1600, na cidade do Rio de Janeiro. Posteriormente, outros focos da doença também foram identificados, na Bahia e no Pará (Eidt, 2004). Após a introdução da moléstia por diversos pontos da costa brasileira a infecção se disseminou a parir de Pernambuco à Paraíba e a Alagoas, provavelmente em função do desenvolvimento agrícola dessas regiões. E ao Ceará, Maranhão, Pará e Amazonas pela ocupação desses Estados. A hanseníase no estado do Pará, do século XIX, tinha grande prevalência e as relações entre Belém, Santarém e Manaus, as quais eram intensas nessa época, auxiliaram na disseminação da doença na região amazônica (Eidt, 2004).

A situação epidemiológica da hanseníase no contexto amazônico é antiga, tanto quanto problemática, e os dados atuais, revelam-na ainda latente como um problema de saúde pública com baixa resolubilidade e alta incidência. No estado do Pará, em dias atuais, a taxa de detecção de novos casos é de 50.7 por 100 000 indivíduos, quase o triplo da taxa nacional (17/100 000) (Barreto *et al.*, 2015). Além disso, a taxa de detecção da doença entre escolares no estado do Pará, é em média de 4%, e entre contatos intradomiciliares dos pacientes de 8% (Barreto *et al.*, 2012). Todos estes fatores desenham uma doença em plena atividade no estado e em processo de expansão, de tal modo que ao se extrapolarem os dados, considerando o tempo de incubação da hanseníase, a realidade socioeconômica do estado e dos pacientes acometidos, assim como as atuais políticas em saúde de combate à doença, pode-se apontar potenciais novos casos no estado do Pará por mais algumas décadas (Salgado *et al.,* 2016).

A distribuição espacial da hanseníase no Brasil é geograficamente desigual, atingindo principalmente as regiões norte, nordeste e centro-oeste do país, em que a doença é hiper-endêmica e subdiagnosticada, porém convergente com a pobreza e a falta de assistência e promoção à saúde, haja vista que aproximadamente metade dos casos detectados no país estão em municípios empobrecidos que compreendem apenas 17% do total populacional do Brasil (Barreto *et al.*, 2015, Barreto *et al.*, 2012, Penna *et al.*, 2009). Este histórico problema de saúde pública na região amazônica, que detectou aproximadamente 80.000 novos casos, nos últimos 20 anos, mostra sinais nítidos de continuidade com falha na quebra da cadeia de transmissão e alta endemicidade em crianças menores de 15 anos , indicando foco ativo da infecção na comunidade (Barreto

et al., 2015, Barreto *et al*., 2012). Adicionalmente, cerca de 50% da população não é coberta pela estratégia de saúde da família, a qual deveria – no âmbito da assistência primaria à saúde - ser responsável pela detecção e tratamento dos pacientes acometidos pela hanseníase, assim pode ser explicado a falha no diagnóstico e alta incidência da doença em nosso estado (Salgado *et al.,* 2016).

A compreensão de que a falta do diagnóstico da hanseníase não represente sua não existência como agravo de saúde na população, cria o imaginário de que a erradicação da doença é objetivo palpável e próximo da nossa realidade (Barreto *et al*., 2015, Penna *et al*., 2009, Salgado *et al.,* 2016). A negligência das políticas públicas em saúde para hanseníase alicerçado com a falha diagnóstica, configura para a nossa região uma endemia oculta e ocultada, em que o enfrentamento do problema se torna ponto importante de pesquisas, como a realizada aqui, que visem desenvolver técnicas e metodologias que aprimorem o controle e o auxílio diagnóstico da doença.

O presente trabalho investigou dois diferentes parâmetros biológicos que podem influenciar o mecanismo de defesa do hospedeiro frente ao proesso infeccioso promovido pelo *M. lepra*: (i) os genéticos intrínsecos ao hospedeiro, que está associado à suscetibilidade à doença; e (ii) os epigenéticos que congregam mecanismos presentes no hospedeiro e no agente infeccioso.

Os fatores genéticos, per si, estão associados a susceptibilidade à doença e a maneira como a interação parasita-hospedeiro pode gerar distintas formas clínicas da patologia, neste sentido encontramos associação dos marcadores do tipo INDEL dos genes *NFκβ1, CASP8, PAR1, IL4* e *CYP19A1* como fatores de susceptibilidade à hanseníase e como potenciais marcadores ao desenvolvimento da forma clínica MB. Adicionalmente, observa-se que o padrão de contribuição Inter-étnica individual, pode afetar de maneira distinta o risco de desenvolvimento da doença.

A análise genética de nossos resultados, evidenciou que o alelo DEL (rs28362491) do gene *NFκβ1* é responsável pela diminuição de sua atividade transcricional e, por conseguinte de uma conjunto de genes pró-inflamatórios que ativa (Karban *et al.,*2004; Salim *et al.,* 2013, Cen *et al.,* 2013). Esta observação sugere que a presença deste polimorfismo induz uma proteção contra a o desenvolvimento da hanseníase, embora tenha sido associado como fator de risco ao desenvolvimento do polo MB. Baseado no fato de que a transcrição de *NFκβ1* não é constitutiva, mas mediado por estímulos antigênicos específicos como antígenos de *M. leprae* (Cen *et al.,*

2013), a ideia de que o alelo DEL confere risco para o desenvolvimento da doença em seu pólo mais grave é viável.

O gene *PAR1* está envolvido em vários momentos do processo inflamatório, e parece modulá-lo em muitos estágios da relação parasita-hospedeiro (Adams *et al.,* 2011). *PAR1 é* um receptor da família PAR e estabelece a supressão da atividade de linfócitos do tipo Th1, Th17, e secreção de IL12 e IL23 (Aerts *et al.,* 2013; Chionh *et al.,* 2014), o que resulta em diminuição da imunidade mediada por células, logo uma diminuição da capacidade de combater a infeção causada pelo *M. leprae*. O alelo INS (rs112667092) deste gene eleva sua capacidade de transcrição. Nossos dados apontam que o genótipo DEL\DEL confere proteção ao desenvolvimento da doença e ao estabelecimento da forma clinica MB, por melhorar o padrão de resposta imune celular e as classes Th1 e Th17. Portanto, a ação do gene *PAR1* pode revelar um novo mecanismo pelo qual o hospedeiro modula sua resposta imune frente ao parasito, e levar ao desenvolvimento da doença.

A elevada carga bacilar em células apresentadoras de antígeno gera um estímulo à apoptose, e este mecanismo é controlado por citocinas e mediadores apoptóticos como BCL-2 e CASP8 (Klinger *et al.,* 1997; Chattree *et al.,* 2008). Como pacientes acometidos pela hanseníase têm sobrecarga bacilar e\ou continua ativação de células T por antígenos circulantes no plasma, estas células são estimuladas à apoptose e a diminuição da taxa de infiltração tissular (Sun *et al.,* 2007). Assim, o alelo DEL (rs3834129) diminuí a taxa transcricional da caspase 8, e consequente diminuição da apoptose de células imunes infectadas com o bacilo. Portanto, nossos dados sugerem que o genótipo DEL\DEL pode elevar a sobrevida de células infectadas e favorecer a o estabelecimento da infecção, sendo assim um fator de risco à hanseníase.

A IL4 é um mediador largamente estudado e analisado em várias patologias, é secretado por linfócitos, eosinófilos e mastócitos, os quais induzem a diferenciação e maturação de linfócitos B e o subconjunto de células Th2, que são ineficazes ao combate da proliferação bacilar e controle da doença (Teles *et al.,* 2010; Nakashima *et al.,* 2002). Os resltados observados no presente trabalho mostram que o VNTR no intron 3 do gene da *IL4*, revela que o alelo A_2, é significativamente frequente em pacientes acometidos pela doença, além disso o alelo A_2 é responsável por uma elevada produção da proteína e consistente com o fato de ser um fator de susceptibilidade à doença.

O gene *CYP19A1,* produz uma enzima da superfamília dos citocromo P450 conhecida como aromatase, a qual é responsável pelo metabolismo primário dos precursores de hormônios esteróides e andrógenos. Em pacientes com hanseníase o nível de hormônios androgênicos é significativamente menor que em indivíduos sem a doença (Leal *et al.,* 2003; Limer *et al.,* 2009). A elevada concentração de hormônios androgênicos no plasma desses pacientes diminui a secreção de citocinas inflamatórias, e sugere que quanto maior a concentração deste hormônio e menor for a atividade metabólica da aromatase, menos efetivo é a inibição à infecção e crescimento do bacilo (Limer *et al.,* 2009). Nossos dados sugerem que o alelo DEL (rs11575899) do gene *CYP19A1* diminue a atividade da aromatase e eleva a concentração de androgênios, resultando em uma redução global da capacidade do hospedeiro de oferecer barreiras imunológicas eficazes contra o desenvolvimento da doença e da forma clinica MB.

Do ponto de vista da contribuição genômica dos diferentes grupos étnicos que constituem a atual população brasileira, podemos sugerir que estas podem gerar taxas de risco ao desenvolvimento de distintas doenças (Santos *et al.,* 1999; Ribeiro-dos-Santos *et al.,* 2007; Ribeiro-Rodrigues *et al.,* 2009), como a hanseníase. Nossos dados sugerem que quanto maior for a contribuição de genes de origem europeia, no genoma do indivíduo, maior será o risco ao desenvolvimento da hanseníase. E quanto maior for a contribuição africana, no genoma do indivíduo, menor é o risco ao desenvolvimento da doença. Do ponto de vista epidemiológico a hanseníase tem estreita ligação com as condições de vida e o desenvolvimento sócio econômico, porém fatores genéticos também são importantes para a susceptibilidade à patologia, logo o fato de indivíduos africanos serem a população mais geneticamente diversa do mundo, associado ao fato de terem entrado em contato com o bacilo a mais tempo (Eidt, 2004; Han *et al.,* 2014), sugre-se que por vias de seleção, polimorfismos genéticos que conferem proteção podem ter se acumulado nesta população e em seu descendentes. É interessante ressaltar que a principal entrada de africanos no Brasil, foi de forma compulsória em navios negreiros que realizavam tráfico de escravos. Nesta situação a comercialização de escravos que apresentassem aspectos que desfavorecessem sua comercialização como aspectos de desnutrição, ausência de dentes, manchas ou lesão de pele, eram considerados como fator de perda de valor; logo os escravos que chegaram ao Brasil, e por conseguinte espalharam seus genes na população atual, em sua maioria eram livres da hanseníase e poderiam ter uma imunidade natural contra o desenvolvimento da doença (Scott *et al.,* 1943; Han *et al.,* 2014).

O segundo parâmetro analisado foram as alterações epigenéticas, por meio de análise da expressão diferencial de miRNAs em amostras de biopsia da lesão hansênica, equanto de sangue periférico de pacientes diagnosticados com hanseníase. Neste contexto o estudo mostrou uma diversidade de 67 miRNAs diferencialmente expressos em amostras de biopsias, as quais diferenciaram o grupo de indivíduos acometidos pela doença e o grupo sem a doença (controle), assim como distinguir os pólos Tuberculóide e Lepromatoso.

As amostras de sangue periférico mostraram uma diversidade menor de microRNAs (10) diferencialmente expressos, mas que também foram hábeis em diferenciar indivíduos acometidos pela doença e o grupo controle, distinguindo os pólos Tuberculóide e Lepromatoso do controle. Adicionalmente a análise *in silico* dos alvos regulados por esses miRNAs, revelou importantes genes que podem influenciar a capacidade do hospedeiro de combater o crescimento bacilar, a progressão da infecção e a manifestação clínica da doença.

O estudo de miRNAs como fatores epigenéticos que podem influenciar o estabelecimento da doença e suas formas clinicas, realizou o primeiro miRnoma da hanseníase, incluindo amostras de pacientes virgens de tratamento - apresentando clinicamente o polo Tuberculóide ou Lepromatoso - e indivíduos não acometidos pela doença. O dado de amostra de biopsia de pele mostrou uma elevada diversidade de miRNA, permitindo ainda o agrupamento destes miRNAs em:

i) Envolvidos na manifestação e progressão da doença;
ii) Marcadores do polo Tuberculóide;
iii) Marcadores do polo Lepromatoso;
iv) Marcadores específicos do polo Lepromatoso;
v) Marcadores que diferenciam os polos da hanseníase.

Os miRNAs diferencialmente expressos nessas amostras corroboraram com alguns dados da literatura como o *has-miR-146a, has-miR-34, has-miR-155, has-miR-29c, has-miR-125b* e *has-miR-99*, que regulam genes importantes para o estabelecimento de uma adequada resposta imunológica, a qual gere condições para a célula do hospedeiro elimine e contenha o bacilo, construindo um micro-ambiente que possa levar a remissão ou eliminação da hanseníase.

Adicionalmente, a analise *in silico* dos alvos revelou importantes genes – já validados – que modulam a resposta imune frente a infecção como *IL1β, IL6, IL8, TLR2, TLR4, IL17RB, IFNGR1* e *TGFBR1*, capazes de influenciar o sensível balanço entre imunidade celular e humoral, por conseguinte a influenciam na susceptibilidade à doença e ao estabelecimento dos polos Tuberculóide ou Lepromatoso. Além disso alguns mediadores importantes como NFκβ (citado acima), proteínas da família SMAD e STAT3, que regulam a sinalização celular e a expressão de citocinas inflamatórias, também são regulados por este *hall* de miRNAs. Portanto, esta visão global do perfil de miRNAs diferencialmente expressos poderá melhorar nosso entendimento e perspectivas sobre novos estudos dos fatores genéticos envolvidos nos mecanismos de defesa do hospedeiro e de combate ao desenvolvimento da hanseníase.

Os dados das amostras de sangue periférico mostraram pouca diversidade de miRNA, porém conseguiu diferenciar pacientes de não pacientes, e os distintos polos da hanseníase de não pacientes, permitindo ainda o agrupamento destes miRNAs em:

vi) Envolvidos na manifestação e progressão da doença;

vii) Marcadores do polo Tuberculóide;

viii) Marcadores do polo Lepromatoso;

ix) Marcadores específicos do polo tuberculóide;

Os miRNAs diferencialmente expressos nessas amostras corroboraram com alguns dados da literatura como o *has-miR-144-5p* e *has-let-7f-5p*, que regulam a diferenciação de células T e a via de ação do NFκβ respectivamente, que são importantes para a sobrevida de micobactérias viáveis na célula do hospedeiro, e para a resposta imune inata estabelecer uma correta barreira imune contra o bacilo. Adicionalmente, a analise *in silico* dos alvos revelou importantes genes – já validados – como *NFκβ, CASP8* e *CYP19A1*, já discutidos acima, mostrando a relação entre os fatores genéticos e epigenéticos envolvidos na relação entre o hospedeiro-parasita que podem influenciar no estabelecimento da doença, além disso mostra que os miRNAs atuam em genes já validados em estudos populacionais e melhoram sua evidencia como possíveis biomarcadores para a hanseníase. Para além a fácil obtenção de amostras de sangue para auxilio diagnóstico, principalmente em áreas remotas e longe de equipamentos modernos de análise, poderiam ser usadas para validar biomarcadores que melhorassem o *screen* epidemiológico de populações vulneráveis à hanseníase, podendo

se estabelecer como uma possível novo marcador de desenvolvimento da patologia e de suas formas clinicas.

Em conclusão, o estudo aponta para marcadores genéticos importantes e suas variantes na população que podem influenciar no estabelecimento da doença, e podem indicar formas alternativas pelas quais o hospedeiro pode oferecer barreira imunológica à infecção pelo *M. leprae* e o desenvolvimento dos distintos pólos da hanseníase. Ademais os dados de miRnoma apontam para um novo cainho para o estudo de marcadores e biomarcadores da hanseníase, revelando que futuras pesquisas podem elucidar novas moléculas que auxiliem o diagnóstico, e até prevê quem em situação de vulnerabilidade e risco poderá desenvolver a doença, ou até encontra infecções subclínicas que necessitem de intervenção profilática. Por fim, a hanseníase ainda configura um problema de saúde pública e anseia por novas pesquisas capazes de gerar resposta para a sociedade, como foi o objetivo desta tese.

6- Referência

Adams MN, Ramachandran R, Yau MK, Suen JY, Fairlie DP et al (2011). **Structure, function and pathophysiology of protease activated receptors.** Pharmacol Ther. 130: 248-282. doi:10.1016/j.pharmthera. 2011.01.003. PubMed: 21277892.

Aerts L, Hamelin M-È, Rhéaume C, Lavigne S, Couture C, et al (2013). **Modulation of Protease Activated Receptor 1 Influences Human Metapneumovirus Disease Severity in a Mouse Model.** PLoS ONE. 8(8): e72529. doi:10.1371/journal.pone.0072529.

Alcais, A.; Mira, M.; Casanova, J. L.; Schurr, E.; Abel, L. (2005). **Genetic dissection of immunity in leprosy**. Current Opinion in Immunology 17: 44–48.

Ali S, Hirschfeld AF, Mayer ML, Fortuno ES, Corbett A, Kaplan M, et al. (2013). **Functional Genetic Variation in NFKBIA and Susceptibility to Childhood Asthma, Bronchiolitis, and Bronchopulmonary Dysplasia**. J Immunol. 190: 3949 - 3958. doi:10.4049/jimmunol.1201015.

Ambros V, Bartel B, Bartel DP *et al.,* (2003). **A uniform system for miRNA annotation. RNA**. 9 (3):277-9.

Andersen V, Christensen J, Ernst A, Jacobsen BA, Tjønneland TA, Krarup HB, et al (2011). **Polymorphisms in NF-κB, PXR, LXR, PPARγ and risk of inflammatory bowel disease**. World J Gastroenterol. 17(2): 197-206.

Araújo, MG. (2003). **Hanseníase no Brasil**. Revista da Sociedade Brasileira de Medicina Tropical.

Arnaud E, Nicaud V, Poirier (2000). **O. Protective effect of a thrombin receptor (protease-activated receptor 1) gene polymorphism toward venous thromboembolism**. Arterioscler Thromb Vasc Biol. 2000; 20: 585–592.

Bach EA, Aguet M, Schreiber RD (1997). **The IFN gamma receptor: a paradigm for cytokine receptor signaling**. Annu Rev Immunol;15:563-91.

Barreto JG, Guimarães LS, Frade MAC, Rosa PS5 Salgado CG (2012). High rates of undiagnosed leprosy and subclinical infection amongst school children in the Amazon Region. Mem Inst Oswaldo Cruz, Rio de Janeiro, Vol. 107(Suppl. I): 60-67.

Bartel, DP (2004) MiRNAs**: genomics, biogenesis, mechanism, and function.** Cell 116(2): 281-297.

Bayley J, Ottenhoff PTH, Verweij CL, (2004). **Is there a future for TNF promoter polymorphisms?**.Genes Immun. 5:315–329.

Bernstein E, Kim SY, Carmell MA *et al.,* (2003) **Dicer is essential for mouse development**. Nature Genetics 35(3):215-217.

Brennan PJ, Vissa VD. **Genomic evidence for the retention of the essential mycobacterial cell wall in the otherwise defective *Mycobacterium leprae***. *Lepr Rev* 2001; 72: 415–28.Britton WJ, Lockwood DN (2004). **Leprosy Lancet**; 363:1209–1219.

Britton WJ, Lockwood DN (2004). Leprosy Lancet; 363:1209–1219.

Salgame PM, Bloom YBR, Modlin RL (1992). **Evidence for functional subsets of CD4_ and CD8_ T cells in human disease: lymphokine patterns in leprosy**. Chem. Immunol; 54:44–59.

Carvalho VCV, Mace JL, Lima CAD, Lima MCG, Hera´clio AS, Amorim M, Maia MMD et al, (2012). **IFN-gamma and IL-12B polymorphisms in women with**

cervical intraepithellial neoplasia caused by human papillomavirus. Mol Biol Rep. 39:7627–7634.

Cen H, Zhoua M, Lenga RX, Wanga W, Fenga C, Li B *et al* (2013). **Genetic interaction between genes involved in NF-KB signalingpathway in systemic lupus erythematosus.** Molecular Immunology. 2013; 56:643– 648.

Chattree V, Khanna N, Bisht V, Rao DN (2008). **Inhibition of apoptosis, activation of NKT cell and upregulation of CD40 and CD40L mediated by M. leprae antigen(s) combined with Murabutide and Trat peptide in leprosy patients.** Mol Cell Biochem. 309:87–97. DOI 10.1007/s11010-007-9646-8

CHEN D, MA T, LIU XW, LIU Z (2012). **CASP-8 -652 6N ins/del polymorphism and cancer risk: A literature-based systematic HuGE review and meta-analysis.** EXP.THER. MED. 4: 762-770. DOI: 10.3892/etm.2012.661.

Chionh YT, Ng GC, Ong L, Arulmuruganar A, Stent A, Saeed MA, Wee JLK, Sutton P (2014). **Protease-activated receptor 1 suppresses Helicobacter pylori gastritis via the inhibition of macrophage cytokine secretion and interferon regulatory factor 5**. Mucosal Immunology. 2014 doi:10.1038/mi.2014.43.

Cole ST, Eiglmeier K, Parkhill J, et al. **Massive gene decay in the leprosy bacillus**. *Nature* 2001; 409: 1007–11.

Cullen, BR (2004) **Transcription and processing of human miRNA precursors.**Mol. Cell.16(6): 861–865.

Cullen BR (2011) **Viruses and microRNAs: RISCy interactions with serious consequences.** Genes Dev 2011, 25(18):1881–1894.

Dickson SN, Elijah AB, Jules CNA, Anna LN, Henri LFK, Anne-Cécile ZKB (2011). **Elimination of Leprosy as a public health problem by 2000 AD: an epidemiological Perspective.** Pan African Medical Journal. 9:4.

Ding SW, Voinnet O (2007) **Antiviral immunity directed by small RNAs.** Cell, 130(3):413–426.

Elizabeth AM, William RB, James CV, Hawn TR (2010). **Leprosy and the Human Genome**. *Microbiol. Mol. Biol. Ver*; 74(4):589.

Ehrenberg J.P.; Ault S.K. (2005). **Neglected diseases if neglected populations: Thinking to reshape the determinants of health in Latin America and the Caribbean**. BMC Public Health 5: 119-131

Foss NT (1997).**Immunological aspects of Leprosy**.Medicina Ribeirão Preto; 30: 335-339.

Foss NT, Motta ACF (2012). **Leprosy, a neglected disease that causes a wide variety of clinical conditions in tropical countries. Mem Inst Oswaldo Cruz.** 107: 28-33.

Goulart IMB, Figueiredo F, Coimbra T, Foss NT (1996). **Detection of transforming growth factor-beta 1 in dermal lesions of diVerent clinical forms of leprosy**. Am J Pathol; 148:911–917.

Goulart IMB, Mineo JR, Foss NT (2000). **Production of transforming growth factor-beta 1 (TGF-_1) by blood monocytes from patients with diVerent clinical forms of leprosy**. Clin Exp Immunol; 122:330–334.

Goulart LR, Goulart IMB (2009). **Leprosy pathogenetic background: a review and lessons from other mycobacterial diseases**. Arch Dermatol Res; 301:123–137.

Hakimi MA, Cannella D (2011) **Apicomplexan parasites and subversion of the host cell microRNA pathway.** Trends Parasitol, 27(11):481–486.

Hershey GK, Friedrich MF, Esswein LA, Thomas ML, Chatila TA (1997). **The association of atopy with a gain-of-function mutation in the alpha subunit of the interleukin-4 receptor**. New Eng. J. Med; 337: 1720-1725.

Holgate ST (1999). **The epidemic of allergy and asthma**. Nature; 402:B2-4.

Houwing S, Kamminga LM, Berezikov E *et al.,* (2007) **A role of piRNA in germ cell maintenance and transposon silencing in zebrafish.** Cell, 129 (1):69-82.

Hüttenhofer A. (2004) **RNomics: identification and function of small non-protein-coding RNAs in model organisms.** Cold Spring Harb Symp Quant Biol. 71:135-40.

Izuhara K, Arima K (2004). **Signal transduction of IL -13 and its role in the pathogenesis of bronchial asthma**. Drug News Perspect; 17:91-8.

Karban AS, Okazaki T, Panhuysen CI, Gallegos T, Potter JJ, Bailey-Wilson JE, et al (2004). **Functional annotation of a novel NFKB1 promoter polymorphism that increases risk for ulcerative colitis.** Hum Mol Genet. 13: 35-45.

Kim VN (2005) **miRNA biogenesis: coordinated cropping and dicing**. Nat Rev Mol Cell Biol. (5): 376-85.

Klingler K, Tchou-Wong KM, Brandli O (1997). **Effects of mycobacteria on regulation of apoptosis in mononuclear phagocytes**. Infect Immun. 5:5272–5278.

Kumar S, Naqvi RA, Khanna N, Rao DN (2011) **Disruption of HLA-DR raft, deregulations of Lck-ZAP-70-Cbl-b cross-talk and miR181a towards T cell hyporesponsiveness in leprosy**. Mol Immunol, 48(9–10):1178–1190.

Leal AMO, Magalhaes PKR, Souza CS, Foss NT (2003). **Adrenocortical hormones and interleukin patterns in leprosy**. Parasite Immunol. 25: 457-461.

Lee DJ, Li H, Ochoa MT, Tanaka M, Carbone RJ, Damoiseaux R *et al, (*2010). **Integrated Pathways for Neutrophil Recruitmentand Inflammation in Leprosy**. The Journal of Infectious Diseases; 201:558–69.

Lee Y, Ahn C, Han J *et al.,* (2003) **The nuclear RNase III Drosha initiates miRNA processing.** Nature 425(6956):415-9.

Lee Y, Kim M, Han J *et al.,* (2004) MiRNA **genes are transcribed by RNA polymerase II**. EMBO J.23(20):4051-60.

Limer KL, Pye SR, Thomson W, Boonen S, Borghs H, Vanderschueren D *et al* (2009). **Genetic Variation in Sex Hormone Genes Influences Heel Ultrasound Parameters in Middle-Aged and Elderly Men: Results From the European Male Aging Study (EMAS)**. J. bone and mineral research. 24:2.

Liu PT, Wheelwright M, Teles R, Komisopoulou E, Edfeldt K, Ferguson B, Mehta MD, Vazirnia A, Rea TH, Sarno EN, Graeber TG, Modlin RL (2012) MicroRNA-21 targets the vitamin D-dependent antimicrobial pathway in leprosy.Nat Med, 18(2):267–273.

Livak KJ, Schmittgen TD (2001) **Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT Method**. Methods. (4):402-8.

Love MI, Huber W, Anders S (2014) **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.** Genome Biology. 15:550. DOI-10.1186/s13059-014-0550-8

MINISTÉRIO DA SAÚDE. (2012). **coordenação geral de hanseníase e doenças em eliminação, 2012.** Brasília: Ministério da Saúde.

Mohamed Ragaa Mohamed & Grant McFadden (2009). **NFκB inhibitors: Strategies from poxviruses**. Cell Cycle, 8(19): 125-3132. DOI: 10.4161/cc.8.19.9683

Moro JR, Iwata M, Andriano UH (2008).**Vitamin effects on the immune system: vitamins A and D take centre stage**. Nat. Rev. Immunol; 8:685–698.

Mosmann TR, (1994). **Properties and functions of interleukin-10**. AdvImmunol;56:1–26.

Muller K, Herner EB, Stagg A, Bendtzen K, Woo P (1998). **Inflammatory cytokines and cytokine antagonists in whole blood cultures of patients with systemic juvenile chronic arthritis**. Br J Rheumatol; 37:562–9.

Nakashima H, Miyake K, Inoue Y (2002). **Association between IL4 genotype and IL-4 production in the Japanese population**. Genes Immun. 2002; 3:107–9.

Nilsen TW (2007) **Mechanisms of miRNA-mediated gene regulation in animal cells.** Trends Genet. (5):243-9.

Pasquinelli AE, Reinhart BJ, Slack F *et al.,* (2000) **Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA.** Nature. 408(6808):86-9.

Paul WE (1991). **Interleukin-4: A prototypic immune regulatory lymphokine.** Blood. 77:1859-70.

Penna G, Adorini L (2000). **1_,25-dihydroxyvitamin D3 inhibits differentiation, maturation, activation, and survival of dendritic cells leading to impaired alloreactive T cell activation**. J. Immunol; 164:2405–2411.

Penna M, Wand-Del-Rey-de-Oliveira M, Penna G. **Spatial distribution of leprosy in the Amazon region of Brazil**. Emerg Infect Dis. 2009;15(4):650–2

Provost P, Dishart D, Doucet J *et al.,* (2002) **Ribonuclease activity and RNA binding of recombinant human Dicer.** EMBO J. 21(21):5864-74.

Reinhart BJ, Slack FJ, Basson M *et al.,* (2000). **The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.** Nature. 403(6772):901-6.

Ridley DS, Jopling WH (1966) **Classification of leprosy according to immunity: a five-group system**. Inernational Journal of Leprosy Other Mycobact. Disease 34:255–273.

Robaee AAA, Zolibani AA, Shobaili HA, Settin A (2012). **Association of interleukin 4 (-590 T/C) and interleukin 4 receptor (Q551R A/G) gene polymorphisms with acne vulgaris**. Ann Saudi Med; 32(4): 349-354.

Roche PW, Theuvenet WJ, Britton WJ. **Cellular immune responses to mycobacterial heat shock proteins in Nepali leprosy patients**. *Int J Lepr Other Mycobact Dis.* 1992; 60: 36–43.

Sahoo M, Olvera IC, Barrio L, Re F (2011). **Role of the Inflammasome, IL-1β, and IL-18 in Bacterial Infections**. The Scientific World Journal; 11: 2037–2050.

Salgado CG, Ferreira DVG, Frade MAC, Guimarães LS, Silva MB, Barreto JG (2012). **High Anti–Phenolic Glycolipid-I IgM Titers and Hidden Leprosy Cases, Amazon Region**. Emerging Infectious Diseases. Vol. 18, Nº5.

Salgado CG, Barreto JG, Silva MB, Frade MAC, Spencer JS, , (2016). **What do we actually know about leprosy worldwide?**. Lancet Infect Dis. Vol. 16.

Salim PH, Jobim M, Bredemeier M, Chies JAB, Brenol JCT, Jobim LF, ET al. **Interleukin-10 Gene Promoter and NFKB1 Promoter Insertion/Deletion Polymorphisms in Systemic Sclerosis**. Scandinavian J. Immunology. 2013; 77:162–168.

Schurr E, Alcais A, Singh M, Mehra N, Abel L (2007). M**ycobacterial infections: PARK2 and PACRG associations in leprosy.**

Schwarz DS, Hutvágner G, Du T *et al.,* (2003) **Asymmetry in the assembly of the RNAi enzyme complex.** Cell 115(2):199-208.

Scollard D, Joyce M, Gillis PT (2006). **Development of leprosy and type 1 leprosy reactions after treatment with infliximab: a report of 2 cases**. Clin. Infect. Dis. 43:e19–e22.

Sevignani C, Calin GA, Siracusa LD *et al.,* (2006) **Mammalian miRNAs: a small world for fine-tuning gene expression.** Mamm Genome 17(3):189-202.

Shaw MA, Donaldson IJ, Collins A, *et al* (2001). **Association and linkage of leprosy phenotypes with HLA class II and tumour necrosis factor genes**. *Genes Immun.* 2001; 2: 196–204.

Singh PK, Singh AV, Chauhan DS (2013) **Current understanding on micro RNAs and its regulation in response to Mycobacterial infections.** Jour. Bio. Sci. 20:14

Song JJ, Liu J, Tolia NH *et al.,* (2003) **The crystal structure of the Argonaute2 PAZ domain reveals an RNA binding motif in RNAi effector complexes.** Nat Struct Biol. 10(12):1026-32.

Song JJ, Smith SK, Hannon GJ *et al.,* (2004) **Crystal structure of Argonaute and its implications for RISC slicer activity.** Science. 305(5689):1434-7.

Sun T, Gao Y, Tan W, Ma S, Shi Y, et al (2007). **A six-nucleotide insertiondeletion polymorphism in the CASP8 promoter is associated with susceptibility to multiple cancers**. Nat Genet. 39: 605–613.

Suzuki K, Akama T, Kawashima A, Yoshihara A, Yotsu RR, Ishii N (2012). **Current status of leprosy: Epidemiology, basic science and clinical perspectives.** Journal of Dermatology; 39: 121–129.

Szymański M, Erdmann VA, Barciszewski J (2003) **Noncoding regulatory RNAs database**. Nucleic Acids Res. 31(1):429-31.

Taft RJ, Pang KC, Mercer TR *et al.,* (2010) **Non-coding RNAs: regulators of disease.** J Pathol. 220(2):126-39.

Teles RMB, Krutzik SR, Ochoa MT, Oliveira RB, Sarno EN, Modlin RL (2010). **Interleukin-4 Regulates the Expression of CD209 and Subsequent Uptake of Mycobacterium leprae by Schwann Cells in Human Leprosy.** Infection and immunity. 78(11):4634–4643.

Thole JE, Janson AA, Cornelisse Y, et al**. HLA-class II-associated control of antigen recognition by T cells in leprosy: a prominent role for the 30/31-kDa antigens**. *J Immunol.* 1999; 162: 6912–18.

Torreele, E. **No drugs in age of plenty: urging governments to redress the balance**. Interdisciplinary Science Reviews 31: 3-8. 2006.

Tomomitsu DDS, Yoichi MD, Koichi MS, Kazuhiko MS, Akira MD, Mitsuteru MD, *et al* (2005). **Functional haplotypes ofIL-12Bare associatedwith childhood atopic asthma**. American Academy of Allergy, Asthma and Immunology;10:789-795.

Trinchieri. G (2003). **Interleukin-12 and the regulate on of innate resistance and adaptive immunity**. Nat Rev Immunol;3:133–46.

Vaishali BN, Usha BN, Swati M, Rao P (2011). **Evaluation of significance of skin smears in leprosy for diagnosis follow-up, assessment of treatment outcome and relapse**. Asiatic Jour. of Biotech.Resources 2(5): 547–552.

Van Roon JAG, van Roy JLAM, Gmelig-Meyling FHJ, LafeberFPJG, Bijlsma JWJ, (1996). **Prevention and reversal of cartilagedegradationin rheumatoid arthritis by interleukin-10 and interleukin-4**. ArthritisRheum; 39:829–35.

Vella MC, Choi EY, Lin SY *et al.,* (2004) **The C. elegans miRNAlet-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR.** Genes Dev. 2004 Jan 15;18(2):132-7.

Vidyagouri S, Patrick M, Deepa SR, Sharada RS, Sundaramoorthy A, Suman J, Indira N, Kumarasamy T, T.P. Velavan, Vijaya LV (2013) **Genetic evidence of TAP1 gene variant as a susceptibility factor in Indian leprosy patients.** Human Immunology, Volume 74, Issue 6, June, p.803-807

World Health Organization (2011). **Enhanced Global Strategy for Further Reducing the Disease Burden due to Leprosy**. Regional Office for South-East Asia. Plan Period 2011-2015.

World Health Assembly. **Leprosy resolution WHA 44.9, Forty-fourth World Health Assembly**, May 13, 1991.

WHO. 2009. **Global leprosy situation**. Wkly. Epidemiol. Rec. 84:333–340.

Wolf SF, Temple PA, Kobayashi M, Young D, Dicig M, Lowe L, et al (1991). **Cloning of cDNA for natural killer cell stimulatory factor, a heterodimeric cytokine with multiple biologic effects on T and natural killer cells**. J Immunol;146:3074–81.

Wu H, Xu H, Miraglia LJ *et al.,* (2000) **Human RNase III is a 160-kDa protein involved in preribosomal RNA processing.** J Biol Chem. 275(47):36957-65.

Yamamura M, K Uyemura, RJ Deans, K Weinberg, TH Rea, BR Bloom,*et al* (1992). **Defining protective responses to pathogens:cytokine profiles in leprosy lesions**. Science ;254:277–279.

Yan KS, Yan S, Farooq A *et al.,* (2003) **Structure and conserved RNA binding of the PAZ domain**. Nature.426(6965):468-74.

Zhang H, Kolb FA, Brondani V *et al*.,(2002) **Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP.** EMBO J. 21(21):5875-85.

Zhang H, Kolb FA, Jaskiewicz L *et al*.,(2004) **Single processing center models for human Dicer and bacterial RNase III.** 118(1):57-68.

7- ANEXO – OUTRA PUBLICAÇÃO NO PERÍODO

Genetics and Molecular Biology, 39, 2, 189-198 (2016)
Copyright © 2016, Sociedade Brasileira de Genética. Printed in Brazil
DOI: http://dx.doi.org/10.1590/1678-4685-GMB-2014-0342

Research Article

CDH1 mutations in gastric cancer patients from northern Brazil identified by Next- Generation Sequencing (NGS)

Antonette El-Husny[1,2], Milene Raiol-Moraes[1,2], Marcos Amador[1,2], André M. Ribeiro-dos-Santos[1,2], André Montagnini[3], Silvanira Barbosa[4], Artur Silva[2,4], Paulo Assumpção[2,5], Geraldo Ishak[2,5,6], Sidney Santos[1,2,5], Pablo Pinto[1,5], Aline Cruz[1,5] and Ândrea Ribeiro-dos-Santos[1,2,5]

[1]*Laboratório de Genética Humana e Médica (LGHM), Instituto de Ciências Biológicas, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.*
[2]*Rede de Pesquisa em Genômica Populacional Humana, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.*
[3]*Instituto Sírio-Libanês de Ensino e Pesquisa, Hospital Sírio-Libanês, São Paulo, SP, Brazil.*
[4]*Laboratório de Polimorfismo de DNA - Instituto de Ciências Biológicas, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.*
[5]*Núcleo de Pesquisas em Oncologia, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.*
[6]*Hospital Universitário João de Barros Barreto, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.*

Abstract

Gastric cancer is considered to be the fifth highest incident tumor worldwide and the third leading cause of cancer deaths. Developing regions report a higher number of sporadic cases, but there are only a few local studies related to hereditary cases of gastric cancer in Brazil to confirm this fact. *CDH1 g*ermline mutations have been described both in familial and sporadic cases, but there is only one recent molecular description of individuals from Brazil. In this study we performed Next Generation Sequencing (NGS) to assess *CDH1* germline mutations in individuals who match the clinical criteria for Hereditary Diffuse Gastric Cancer (HDGC), or who exhibit very early diagnosis of gastric cancer. Among five probands we detected *CDH1* germline mutations in two cases (40%). The mutation c.1023T > G was found in a HDGC family and the mutation c.1849G > A, which is nearly exclusive to African populations, was found in an early-onset case of gastric adenocarcinoma. The mutations described highlight the existence of gastric cancer cases caused by *CDH1* germline mutations in northern Brazil, although such information is frequently ignored due to the existence of a large number of environmental factors locally. Our report represent the first *CDH1* mutations in HDGC described from Brazil by an NGS platform.

Keywords: *CDH1,* germline mutations, HDGC, Gastric Cancer, NGS.

Received: November 30, 2014; Accepted: July 23, 2015.

Introduction

Gastric cancer was considered to be the fifth highest incident tumor in both genders worldwide in 2012 and the third leading cause of cancer deaths (8.8% of the total). Although East Asia represents the region of greatest gastric cancer mortality, high rates are also observed in both genders in Central and Eastern Europe and in Central and South America (Ferlay *et al.*, 2010). In Brazil, it was estimated that gastric cancer was the fourth most frequent in men and the fifth most common in women in 2014. In the northern region, gastric cancer is the second most common in men and the third most common in women (INCA, 2014).

Corso *et al.* (2012) stated that developing regions of the world have a greater number of sporadic cases. Although only a few local studies have addressed this question, their results suggest that this statement is also applicable to Brazil. The gene related to the gastric cancer predisposition syndrome (Hereditary Diffuse Gastric Cancer - HDGC) is *CDH1* which encodes the E-cadherin protein, responsible for cell adhesion in non-neural epithelial cells, among other functions. Germline mutations of this gene as a cause of familial cases of diffuse gastric cancer were first described in the Maori tribe in New Zealand, in which the existence of a genetic syndrome predisposing to gastric cancer was first detected (Guilford *et al.*, 1998, 1999; Blair *et al.*, 2013).

Send correspondence to Ândrea Ribeiro-dos-Santos. Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federal do Pará, Cidade Universitária Prof. José da Silveira Netto, 01, 66075-970 Belém, PA, Brazil. E-mail: akelyufpa@gmail.com or akely@ufpa.br.

To date, *CDH1* mutations have been described in various populations around the world. Noteworthy are the descriptions of mutations associated with carcinogenesis of different tumor types, such as breast cancer, prostate cancer, and gastric cancer (Richards *et al.*, 1999; Ikonen *et al.*, 2001; Masciari *et al.*, 2007; McVeigh *et al.*, 2014). Germline *CDH1* mutations have been described in approximately 30-40% of familial cases and in a smaller proportion of sporadic cases (Kaurah *et al.*, 2007; Garziera *et al.*, 2013;). Although several mutations have been detected in distinct families, no hotspot has been characterized. To date, there is only one molecular description of individuals from Brazil (Moreira-Nunes *et al.*, 2014).

Little has been discussed regarding the ancestral origin of pathogenic mutations described in cases of diffuse gastric cancer. However, a population approach is important when dealing with rare disorders like HDGC, because specific mutations observed in a population can guide the testing approach in other individuals of the same group.

Given the high incidence of gastric tumors in Brazil, specifically in the state of Pará, it is of great importance to focus the attention of clinicians and researchers on genetic factors potentially associated with gastric cancer in this population.

Subjects and Methods

Subjects

The study included six individuals of which only two were related (cases 1 and 3). Among the patients, four (cases 1, 3, 4 and 5) matched the clinical criteria for HDGC of the International Gastric Cancer Linkage Consortium (Fitzgerald *et al.*, 2010) and two (cases 2 and 6) had early onset (< 40 years) diffuse-type gastric cancer (Fitzgerald *et al.*, 2010; Kluijt *et al.*, 2012). Five patients (cases 1 to 5) were from northern Brazil and one individual was from southeastern Brazil (case 6).

This study was approved by the Research Ethics Committee of Universidade Federal do Pará - Hospital João de Barros Barreto (protocol number 359.927), obeying the principles of the Declaration of Helsinki and Nuremberg Code. All individuals signed an Informed Consent form.

Genotyping analysis

DNA extraction was performed with the PureLink Genomics™ Mini Kit(Life Technologies, Foster City, CA, USA) according to the manufacturer's protocol. Amplification of the coding regions of the *CDH1* gene was performed by PCR, with a total of 20 amplicons per patient. The amplicons of the five index individuals (Table 1) were sequenced on an Ion Torrent PGM™ platform (Life Technologies).

Next Generation Sequencing (NGS)

Each of the 20 amplicons, including the promoter region of *CDH1* and its 16 exons (Table 1), were combined into a single equimolar pool with a total of 100 ng of DNA in a final volume of 35 μL.

The fragmentation of samples for 200 bp sequencing was performed with the ION shear plus reagent kit (Life Technologies) followed by purification with Agencourt™ AMPure Reagent XP™ (Agencourt Bioscience Corporation, Beverly, MA, USA). The connection of barcode adapters and repairs were made with the following kits: Ion Plus kit fragment library (Life Technologies) and Ion Xpress Barcode Adapters 1-16 kits (Life Technologies) in each one of the individual samples.

After purification, the samples passed through size selection in an e-Gel size selection 2% agarose gel, from which a band of 200 base pairs was retrieved. The library was then amplified, purified and assembled in the same pool concentration for emulsion PCR with the Ion PGM 200 template reagent kit (Life Technologies).

The sequencing reaction was performed with the ION sequencing reagent kit (Life Technologies). The resulting data was aligned to the reference genome hg19 (available at http://genome.ucsc.edu) and the mutations were identified using the GATK v.2.6. Toolkit. The variants were filtered by low quality calling (less than 50X depth and homopolymer runs) and analyzed using the Integrative Genomics Viewer software (IGV v.2.3) (Broad Institute; https://www.broadinstitute.org/igv).

Validation

After identifying the familial mutation (index individuals), first-degree relatives were screened for the specific mutation (Figure 1) by Sanger sequencing. Although being a limited analyses for clinical purposes, every missense mutation detected had its potential pathogenicity tested using Polyphen-2 (Adzhubei *et al.*, 2010) and SIFT (Ng and Hennikoff, 2003) for predicting functional effects of human nsSNPs.

The main detected mutations were investigated by Sanger sequencing of 100 samples from the local population of Belém, PA, Brazil, for comparative purposes. Additionally, all populational data for mutations were checked using the 1,000 genomes project data (McVean *et al.*, 2012).

For each exon, the sequencing reaction was performed with 1 μL of purified PCR product of each exon, 0.5 μL of the reverse specific primer, 0.5 μL of Big Dye Terminator V3.1 Cycle Sequencing Kit (Life Technologies), and 3.0 μL of SaveMoney buffer to a final volume of 5 μL. The thermocycling reaction proceeded as follows: 95 °C for 2 min, followed by 35 cycles of 95 °C for 30 s, 60 °C for 20 s and 72 °C for 2 min.

Table 1 - Polymerase chain reaction primers used for amplification of *CDH1* promoter region and its 16 exons, their product size and flanking regions size.

Region	Forward and reverse primers	Product size (bp)	Flanking regions size (bp)
Promoter	5' GAGAACTCAGTAAAGGGGCTGA 3'	853	5'- 108
	5' ACTAAGACCTGGGATCAGAAAGG 3'		-
Exon 1	5' CCATCTCCAAAACGAACAAAC 3'	752	-
	5' GAACTTTCTTGGAAGAAGGGAAG 3'		3' - 113
Exon 2	5' CTAGGTCTTGAGGGGGTGACT 3'	486	5' – 236
	5' GTAAATTCCAAGGGGTGTCGT 3'		3' – 135
Exon 3	5' GTAAATTCCAAGGGGTGTCGT 3'	421	5' – 80
	5' CAACCCCTACAACACAAAATCA 3'		3' - 117
Exon 4	5' TCAAACTGTACACTGCCCACA 3'	347	5' – 117
	5' ATCCCAACACTGGGTCTTTTC 3'		3' - 86
Exon 5	5' TCTGTTTCTCTGGGAGGGATT 3'	383	5' – 111
	5' TCAAGTTAAGCTCCTCATGTGTTC 3'		3' – 106
Exon 6	5' GTCACCCTCACTTGGTTCTTTC 3'	280	5' – 22
	5' CCGTAGGAAGGATCAGCTTTAGT 3'		3' – 111
Exon 7	5' TTCTTTCTCCCCTAGCACTTTG 3'	436	5' – 169
	5' ACAACTGGCCTAGCAGGATTT 3'		3' – 91
Exon 8	5' CTTGGTTGTGTCGATCTCTCTG 3'	194	5' – 103
	5' GACCTTTCTTTGGAAACCCTCT 3'		3' – 40
Exon 9	5' ATGATCGCTCAAATACACTCCA 3'	429	5' – 148
	5' CTGCCAAAGCGAATCTACTTCT 3'		3' – 99
Exon 10	5' CATTGAAAGTCATGGCAGAAAC 3'	420	5' – 142
	5' GCTGCAAGTCAGTTGAAAAATC 3'		3' – 33
Exon 11	5' GCTTAAGCCGTTTTCAGCTACA 3'	303	5' – 70
	5' AACTCTTCCCTCCAAAAGAAGG 3'		3' – 87
Exon 12	5' CTAGACTTGGTCTGGTGGAAGG 3'	430	5' – 79
	5' GGAAGCAAGTATCAATGGAAGG 3'		3' - 126
Exon 13	5' AAGCAGCTCTGCTCTCTTCACT 3'	470	5' – 122
	5' CTCTTTCCCACATCAGCTAACC 3'		3' – 120
Exon 14	5' TCTGTGATAGCTGCTGCTTCTG 3'	294	5' – 75
	5' AGCTGTTTCAAATGCCTACCTCT 3'		3' – 88
Exon 15	5' AAGGCATCATCCAACCATAATC 3'	311	5' – 100
	5' TTTTTGACACAACTCCTCCTGA 3'		3' – 67
Exon 16.1	5' AAGTCTGGGTGCATTGTCGTA 3'	690	5' – 110
	5' AGCTGACTTCTCCCCTTCTTTT 3'		-
Exon 16.2	5' CAGCACCTTGCAGATTTTCTTA 3'	840	-
	5' CTAGTCAAGATGTGGCCAGACA 3'		-
Exon 16.3	5' CAGTTGCTTTGCCCAAGATAG 3'	817	-
	5' TAGCTTGAACTGCCGAAAAATC 3'		-
Exon 16.4	5' GGTAGTGAGGATCTTGATTTGGA 3'	398	-
	5' CCTCTTTCTCCACGTTTTGACT 3'		3' - 90

After thermocycling, the product was prepared for sequencing in an ABI 3130 automatic sequencer (Life Technologies). The sequence information was interpreted by ABI Analysis Software^{TM}. The electropherograms were analyzed using the ChromasPro1.49 software and compared with the reference sequence obtained from GenBank (NM_004360.3).

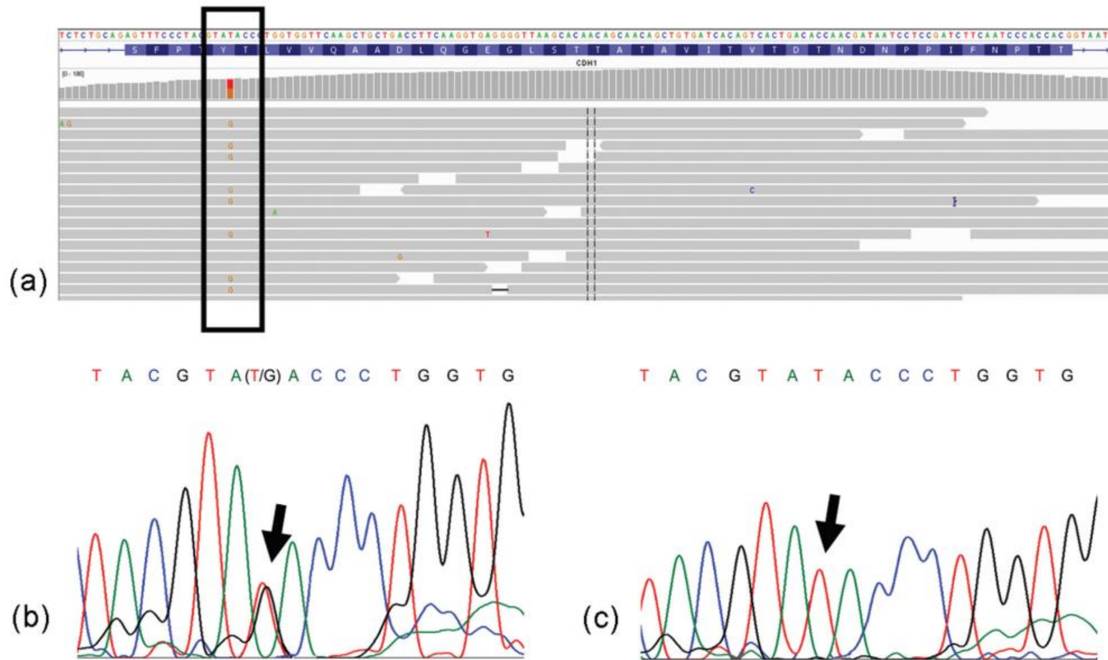


Figure 1 - Molecular analysis presenting exon 8 *CDH1* mutation, c.1023T > G. (a) Integrative Genomics Viewer - IGV™ software result; (b) Sanger sequencing analysis of a patient; (c) Sanger sequencing analysis of a control case.

Analysis of genetic ancestry

All five index cases had their genetic ancestry tests performed using 48 INDEL informative markers of ancestry (American Indians, Europeans and Africans), following the methodology previously described by Santos *et al.* (2010). Three multiplex PCR reactions were performed, each with 16 markers, followed by electrophoresis on an ABI-PRISM 3130 sequencer and analysis using GeneMapper ID v. 3.2 (Life Technologies). The individual proportions of European, African and Amerindian genetic ancestry were estimated using the STRUCTURE software v.2.3.3, assuming three parental populations (Europeans, Africans and Amerindians) and running with a 200,000 burn-in period and 200,000 Markov Chain Monte Carlo repetitions after burning.

Results

Results of the molecular analyses of the *CDH1* gene are summarized in Table 2. Among five probands, we detected *CDH1* germline mutations in two cases (40%). Case 1 exhibited a heterozygous *CDH1* exon 8 germline mutation c.1023T > G. This patient is member of a classic HDGC family. Case 3, who belongs to the same family, confirmed the presence of the c.1023T > G mutation.

Figure 1a presents the molecular analysis of the *CDH1* germline mutation c.1023T > G by Integrative Genomics Viewer (IGV v.2.5) software, which was validated by Sanger sequencing (index case 1; Figure 1b). In a sample of 100 individuals from the local population of Belém, PA (Brazil) no instances of the mutation were found (Figure 1c). This mutation was identified as familial by analyzing first-degree relatives of cases 1 and 3 by Sanger sequencing. The family pedigree with 46 individuals is presented in Figure 2. Ten of the family members were tested for the c.1023T > G mutation and all exhibited the heterozygous mutation.

The molecular analysis of index case 2 revealed the *CDH1* germline mutation exon 12 c.1849G > A in heterozygosis (Table 2), which was confirmed by Sanger sequencing. This patient exhibited early-onset gastric adenocarcinoma (by 28 years of age) without any other similar case in the family. Figure 3a presents the molecular analysis for the *CDH1* germline mutation c.1849G > A by IGV v.2.3. Similar to index case 1, this mutation was not identified in the sample from local population (Figure 3c).

Previous population studies performed by the 1000 Genomes Project (McVean *et al.*, 2012) did not detected the c.1023T > G mutation, but described the c.1849A mutation as almost exclusively African, with an allele frequency of 0.045 and overall database frequency (MAF) of 0.01.

NGS alignment presented two INDEL variations with good quality (more than 50X depth) in heterozygous state among all individuals: c.1649delG and c.2218delC located

Table 2 - Germline mutations and polymorphisms detected in *CDH1* exons in hereditary diffuse gastric cancer in Brazil.

	Exon	*CDH1* mutation	Type of mutation	Prediction	Ancestry panel			Previous References	
					EUR	AME	AFR	Tumor type	Origin
Case 1*	8	c.1023T > G / p.Tyr341*	Nonsense	Pathogenic	0,517	0,334	0,149	Gastric cancer (McVean *et al.*, 2012)	Unavailable
	13	rs1801552	Silent	-				Populational studies (Adzhubei *et al.*, 2010)	Unavailable
		c.2076T > C / p.Ala692=							
Case 2	12	rs33935154	Missense	Benign**	0,520	0,167	0,313	Endometrial cancer (Risinger *et al.*, 2010)	Unavailable
				(SIFT score 0.19)				Sporadic diffuse gastric cancer (Guilford *et al.*, 2010)	African-American
				(PSIC score 0.04)				Early onset diffuse gastric cancer (Risinger *et al.*, 1994)	African-American
		c.1849G > A / p.Ala617Thr						Populational studies (Adzhubei *et al.*, 2010)	Unavailable
	13	rs1801552	Silent	-					
		c.2076T > C / p.Ala692=							
Case 3*	8	c.1023T > G / p.Tyr341*	Nonsense	Pathogenic	0,391	0,385	0,224	Gastric cancer (McVean *et al.*, 2012)	Unavailable
	13	rs1801552	Silent	-				Populational studies (Adzhubei *et al.*, 2010)	Unavailable
		c.2076T > C / p.Ala692=							
Case 4	13	rs1801552	Silent	-	0,392	0,356	0,252	Populational studies (Adzhubei *et al.*, 2010)	Unavailable
		c.2076T > C / p.Ala692=							
Case 5	13	rs1801552	Silent	-	0,511	0,399	0,090	Populational studies (Adzhubei *et al.*, 2010)	Unavailable
		c.2076T > C / p.Ala692=							
	16	rs2229044	Silent	-				Populational studies (Adzhubei *et al.*, 2010)	
		c.2634C > T / p.Gly878=							
Case 6	13	rs1801552	Silent	-	0,656	0,186	0,158	Populational studies (Adzhubei *et al.*, 2010)	Unavailable
		c.2076T > C / p.Ala692=							

• Individuals from the same family; ** Pathogenicity prediction evaluated by SIFT (sorting intolerant from tolerant) and Polyphen2 (PSIC) softwares.

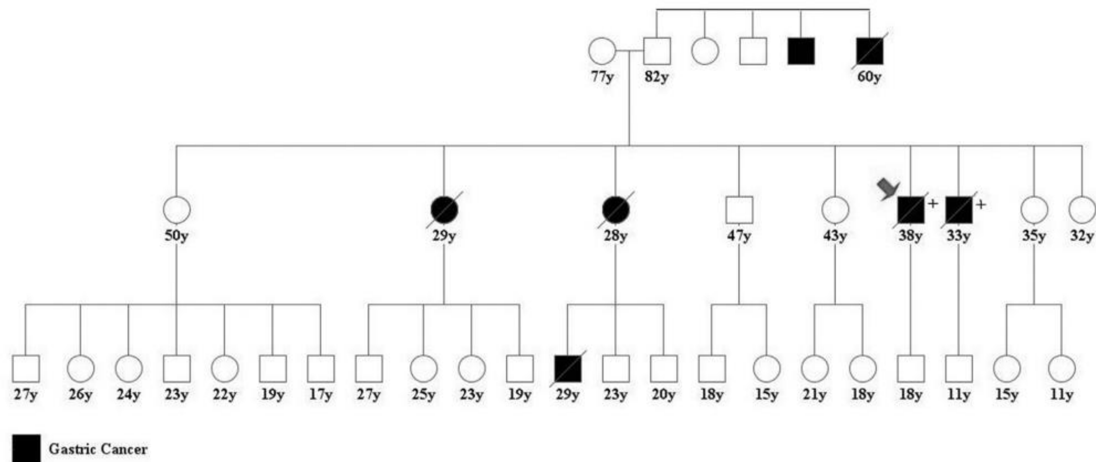


Figure 2 - Pedigree of the *HDGC* Northern Brazilian family described in this paper, red arrow showing the index case (case 1); (+) represents individuals with molecular analysis showing the *CDH1* c.1023T > G mutation.

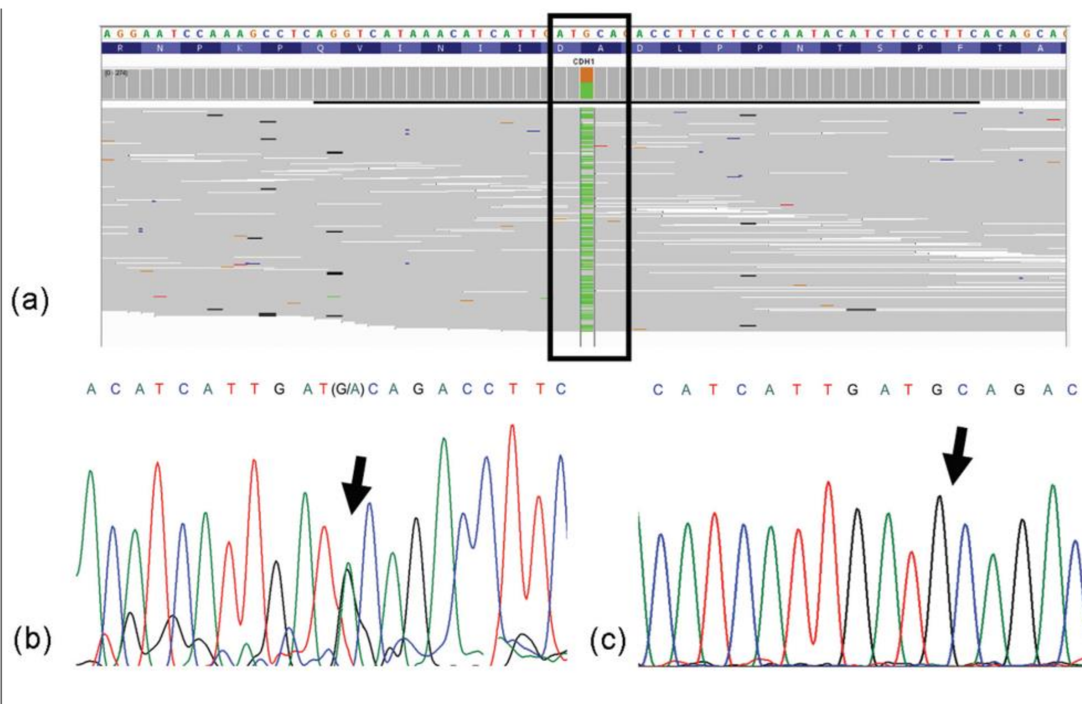


Figure 3 - Molecular analysis presenting exon 12 *CDH1* mutation, c.1849T > G. (a) Integrative Genomics Viewer - IGV^{TM} software result; (b) Sanger sequencing analysis of the patient; (c) Sanger sequencing analysis of a control case.

in exon 11 and exon 14, respectively. Such kind of frame shift mutations would cause drastic damage for the final protein, and to us they seem to be incompatible with the biological and epidemiological background of gastric cancer. These variants probably represent sequencing errors of the platform chemistry, since they were found in highly repetitive regions. Thus, we choose to validate all our findings by Sanger sequencing.

No pathogenic *CDH1* germline mutation was confirmed in the index cases 4, 5 and 6, and only polymorphisms were observed (Tables 2 and 3).

When analyzing the genetic ancestry contribution, index case 1 exhibited a 52% European, 33% Amerindian and 15% African contribution. Index case 2 exhibited a 52% European, 17% Amerindian and 31% African contribution. These results are presented in Table 2.

Besides the exonic alterations, NGS was also able to detect intronic modifications. The ones already described as polymorphisms and registered in NCBI dbSNP are shown in Table 3.

Discussion

The *CDH1* germline mutation of index case 1 (c.1023T > G) had been previously described in New Zealand patients, with only few clinical details available (Guilford *et al.*, 2010). It is responsible for the introduction of a premature stop codon at position 341 of the E-cadherin (p.Y341*) protein and is therefore pathogenic. This region encodes the second cadherin domain that is normally located in the extracellular portion of the protein and is essential for its juxtacellular adhesion function. Significant levels of mature protein cease to be translated in the presence of this mutation. Given the typ-

Table 3 - Known polymorphisms detected in *CDH1* introns and flanking regions.

CDH1 region	Cases					
	Case 1*	Case2	Case 3*	Case 4	Case 5	Case 6
5' flanking region						
rs7194355 (C > A)	*wt*	A/A	*wt*	C/A	C/A	*wt*
rs35582463 (C > T)	C/T	*wt*	C/T	*wt*	*wt*	*wt*
rs33945903 (C > T)	C/T	*wt*	C/T	*wt*	*wt*	*wt*
rs5030625 (GA > G)	G/G	G/G	G/G	G/G	G/G	GA/G
rs3395334 (C > A)	C/A	*wt*	C/A	*wt*	*wt*	*wt*
rs16260 (C > A)	*wt*	A/A	A/A	C/A	C/A	*wt*
Intron 1						
rs3743674 (C > T)	T/T	T/T	*	T/T	T/T	*wt*
rs147838237 (C > CGCCCCAGCCCCGT)	hoz	hoz	*	hoz	hoz	*wt*
rs286579983 (T > C)	T/C	T/C	*	T/C	T/C	*wt*
rs12928281 (C > T)	C/T	C/T	*	C/T	C/T	C/C
Intron 6						
rs8059669 (A > C)	*	*wt*	A/C	*wt*	*	A/C
Intron 7						
rs34374107 (T > C)	*wt*	*wt*	*wt*	*wt*	*wt*	*wt*
Intron 9						
rs35423758 (C > T)	*wt*	*wt*	*wt*	C/T	*wt*	*wt*
rs339509003 (G > C)	*wt*	*wt*	*wt*	*wt*	*wt*	*wt*
Intron 12						
rs2276330 (T > C)	T/C	*wt*	*wt*	*wt*	T/C	*wt*
3'UTR						
rs1801026 (C > T)	C/T	*wt*	*wt*	*wt*	C/T	*wt*
rs8049282 (C > T)	C/T	*wt*	C/T	C/T	*wt*	*wt*
rs33956791 (C > T)	*wt*	*wt*	*wt*	*wt*	*wt*	*wt*
rs9282653 (G > A)	*wt*	*wt*	*wt*	*wt*	*wt*	*wt*
rs13689 (T > C)	T/C	*wt*	*wt*	*wt*	*wt*	*wt*
3' flanking region						
rs8045438 (A > G)	G/G	G/G	G/G	G/G	G/G	G/G
rs181705992 (T > A)	*wt*	*wt*	*wt*	*wt*	*wt*	*wt*
rs17690554 (C > G)	C/G	*wt*	*wt*	*wt*	*wt*	*wt*

wt, wildtype; *, no call; *hoz*, mutant homozygous

ical genetic admixture in Brazil, evidenced by studies of population ancestry (*e.g.* Santos *et al.*, 2010), it was not possible to associate ethnicity with the presence of the c.1023T > G mutation. However, the absence of this mutation in 100 individuals of the local population (Belém, PA, Brazil) and in the 1,000 genome project dataset (McVean *et al.*, 2012) puts in evidence that the mutation must be a rare mutation rather than a polymorphic variant.

Case 2 exhibited the c.1849G > A mutation previously detected by Risinger *et al.* (1994) in a tissue sample of endometrial cancer, and by Ascano *et al.* (2001) in patients with diffuse gastric cancer. Subsequently, Suriano *et al.* (2003) demonstrated the functional inactivation *in vitro* of c.1849G > A in cases of early onset gastric cancer like the present case. Supported by these studies, the pathogenicity of this mutation is described in the NCBI SNP database (http://www.ncbi.nlm.nih.gov/projects/SNP) under the code rs33935154.

However, analysis performed by Polyphen-2 (Adzhubei *et al.*, 2010) and SIFT (Ng and Hennikoff, 2003) software suggest low pathogenicity for this mutation (PSIC score 0.04; SIFT score 0.19), since different species present substitutions in the same protein position (p.617). Population studies performed in the 1,000 genomes project (McVean *et al.*, 2012) described the overall frequency (MAF) of allele c.1849A as 0.01. In the same database, the presence of this mutation is almost exclusive of African populations, with an allele frequency increasing to 0.045.

This divergent information can be explained under at least four hypotheses: (i) the mutation has incomplete penetrance; (ii) the mutation is not truly pathogenic, given the high frequency in African populations, where there is no significant increase in the HDGC case number or diffuse type gastric cancer at a young age; although scientific data on this subject are minimal; (iii) the mutation is pathogenic for other population groups but not among Africans, given its local frequency of 4.5%; (iv) the mutation is only pathogenic in the presence of other genetic and/or epigenetic factors not yet studied. As there is no study testing this mutation in African patients with diffuse gastric cancer, all hypotheses above must be considered.

In case 2, it is likely that the mutation occurs as an effect of the patient's African ancestry contribution of 31%. Interestingly, Suriano *et al.* (2003) described the c.1849A mutation in two African-American-unrelated cases. The geographical origin of these two cases was not available.

Regarding the polymorphisms described in this paper, we highlight the exonic rs1801552 and the intronic: rs13689, rs16260 and rs17690554, already analyzed by Zhan *et al.* (2012) in a case-control study without any difference detected between genotypes in gastric cancer patients and control group.

Previous studies on rs16260 and also on rs1801026 have called attention to the possibility of being markers for genetic susceptibility to cancer (Wang *et al.*, 2007; Li *et al.*, 2011). But conflicting results were shown when comparing different populations (Li *et al.*, 2012). Additional population studies of Brazilian subjects from different geographic regions of the country should be performed to find out whether these polymorphisms can provide useful susceptibility information in our country.

Recent studies of *CDH1* with NGS focused mostly on hereditary breast cancer (Castera *et al.*, 2014; Yang *et al.*, 2015). Dang *et al.* (2014) studied *CDH1* mutations in gastric cancer tissues with NGS, aiming to clarify their particular pathogenesis. All of these studies were performed on different sequencing platforms than the one used here, and concluded that NGS technology was an excellent method for their investigations.

Regarding the mutations c.1649delG and c.2218delC detected by NGS in all cases, it is unlikely that they represent low mosaicism with clinical significance, as they presented large read depth and calling quality in every case analyzed. Rather, they seem to be due to a sequencing error of a highly repetitive region of the genome. Although NGS techniques generate highly reliable data, they still can produce several miss calling due to specifics of their chemistry and software limitations. Thus, a critical analysis is necessary when interpreting variant reports, and suspicious results must be double checked by techniques such as Sanger sequencing or others.

Furthermore, although three of the six index cases presented no pathogenic mutation, the proportion of diffuse gastric cancer cases with mutations detected is similar to previously reports in the literature (40%; Kaurah *et al.*, 2007; Garziera *et al.*, 2013). Quite possibly, other genes or epigenetics factors may be the cause behind these undefined cases.

Conclusion

The mutations described in this paper demonstrate the existence of gastric cancer cases caused by *CDH1* germline mutations in an endemic region of gastric cancer (northern Brazil), and such information is frequently ignored due to the significant number of environmental factors present.

The presence of the c.1849G > A mutation, a mutation almost African-exclusive, demonstrates the importance of considering ancestry and ethnicity when studying genetic disorders.

These *CDH1* germline mutations (c.1849G > A; c.1023T > G) are the first described in association with HDGC and early onset gastric cancer from Brazil revealed by a Next-Generation Sequencing platform. Larger studies that examine the frequency of gastric cancer cases associated with an abnormal *E-cadherin* gene will be of great value to determine the true importance of this genetic factor for gastric cancer in this area.

Acknowledgments

The authors thank the pacients from NPO (Núcleo de Pesquisas em Oncologia) and HUJBB/UFPA (Hospital Universitário João de Barros Barreto - Universidade Federal do Pará) for the participation of this study. Thank you Aline Cruz and Pablo Pinto for help. This work was part of the REDE DE PESQUISA EM GENÔMICA POPULACIONAL HUMANA (RPGPH - BioComputacional/CAPES). Financial support was provided by PROPESP/UFPA-FADESP, MS/Decit, CNPq, FAPESPA, SESPA and CAPES. ASEH was supported by Programa de Bolsa Especial para Doutorado em Pesquisa Médica (PBE-DPM/CAPES); SS and ARDS by CNPq/Produtividade grants.

References

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7:248-249.

Ascano JJ, Frierson H, Moskaluk CA, Harper JC, Roviello F, Jackson CE, El-Rifai W, Vindigni C, Tosi P and Powell, SM (2001) Inactivation of the E-cadherin gene in sporadic diffuse-type gastric cancer. Mod Pathol 14:942-949.

Blair V, Kahokehr A and Sammour T (2013) Cancer in Mãori: Lessons from prostate, colorectal and gastric cancer and progress in hereditary gastric cancer in New Zeland. ANZ J Surg 83:42-48.

Castéra L, Krieger S, Rousselin A, Legros A, Baumann J-J, Bruet O, Brault B, Fouillet R, Goardon N, Letac O, *et al.* (2014) Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. Eur J Hum Genet 22:1305-1313.

Corso G, Marrelli D, Pascale V, Vindigni C and Roviello, F (2012) Frequency of CDH1 germline mutations in gastric carcinoma coming from high- and low-risk areas: Metanalysis and systematic review of the literature. BMC Cancer 12:e8.

Dang Y, Wang YC and Huang QJ (2014) Microarray and next-generation sequencing to analyse gastric cancer. Asian Pac J Cancer Prev 15:8033-8039.

Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Fitzgerald RC, Hardwick R, *et al.* (2010) Hereditary diffuse gastric cancer: Updated consensus guidelines for clinical management and directions for future research. J Med Genet 47:436-444.

Garziera M, De Re V, Geremia S, Seruca R, Figueiredo J, Melo S, Simões-Correia J, Caggiari L, De Zorzi M, Canzonieri V, *et al.* (2013) A novel CDH1 germline missense mutation in a sporadic gastric cancer patient in north-east of Italy. Clin Exp Med 13:149-157.

Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A and Reeve AE (1998) E-cadherin germline mutations in familial gastric cancer. Nature 392:402-405.

Guilford P, Hopkins JBW, Grady WM, Markowitz SD, Willis J, Lynch H, Rajput A, Wiesner GL, Lindor NM, Burgart LJ, *et al.* (1999) E-cadherin germline mutations define an inherited cancer syndrome dominated by diffuse gastric cancer. Hum Mutat 14:249-255.

Guilford P, Humar B and Blair V (2010) Hereditary diffuse gastric cancer: Translation of CDH1 germline mutations into clinical practice. Gastric Cancer 13:1-10.

Ikonen T, Matikainen M, Mononen N, Hyytinen E, Helin HJ and Tommola S (2001) Association of E-cadherin germ-line alterations with prostate cancer. Clin Cancer Res 7:3465-3471.

Instituto Nacional de Câncer José Alencar Gomes da Silva (2014) Estimativa 2014: Incidência de Câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva, Coordenação de Prevenção e Vigilância. INCA, Rio de Janeiro, 126 pp.

Kaurah P, MacMillan A and Boyd N (2007) Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. JAMA 297:2360-2372.

Kluijt I, Sijmons RH, Hoogerbrugge N, Plukker JT, de Jong D, van Krieken JH, van Hillegersberg R, Ligtenberg M, Bleiker E and Cats A (2012) Familial gastric cancer: Guidelines for diagnosis, treatment and periodic surveillance. Familial Cancer 11:363-369.

Li Y, Tang Y, Zhou R, Sun D, Duan YA, Wang N, Chen ZF and Shen N (2011) Genetic polymorphisms in the 3'-untranslated region of the E-cadherin gene is associated with risk of different cancers. Mol Carcinog 50:857-862.

Li YL, Tian Z, Zhang JB and Fu BY (2012) CDH1 promoter polymorphism and gastric cancer susceptibility. Mol Biol Rep 39:1283-1286.

Masciari S, Larsson N, Senz J, Boyd N, Kaurah P, Kandel MJ, Harris LN, Pinheiro HC, Troussard A, Miron P, *et al.* (2007) Germline E-cadherin mutations in familial lobular breast cancer. J Med Genet 44:726-731.

McVean GA, Donnelly P, Lunter G, Marchini JL, Myers S, Gupta-Hinch A, Iqbal Z, Mathieson I, Rimmer A, Xifara DK, *et al.* (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491:56-65.

McVeigh T, Choi JK, Miller NM, Green AJ and Kerin MJ (2014) Lobular breast cancer in a CDH1 splice site mutation carrier: Case report and review of the literature. Clin Breast Cancer 14:47-51.

Moreira-Nunes CA, Barros MB, Borges BN, Montenegro RC, Lamarão LM, Ribeiro HF, Bona AB, Assumpção PP, Rey JA, Pinto GR, *et al.* (2014) Genetic screening analysis of patients with hereditary diffuse gastric cancer from northern and northeastern Brazil. Hered Cancer Clin Pract 12:e18.

Ng PC and Henikoff S (2003) SIFT-predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812-3814.

Richards FM, McKee SA, Rajpar MH, Cole TR, Evans DG, Jankowski JA, McKeown C, Sanders DS and Maher ER (1999) Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. Hum Mol Genet 8:607-610.

Risinger J, Berchuck A, Kohler M and Boyd J (1994) Mutations of the E-cadherin gene in human gynecologic cancers. Nat Genet 7:98-102.

Santos NPC, Ribeiro-Rodrigues E, Ribeiro-dos-Santos A, Pereira R, Gusmão L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, *et al.* (2010) Assessing individual interethnic admixture and population substructure using a 48-insertion-

deletion (INSEL) ancestry-informative marker (AIM) panel. Hum Mutat 31:184-190.

Suriano G, Oliveira C, Ferreira P, Machado J, Bordin M, De Wever O, Bruyneel EA, Moguilevsky N, Grehan N, Porter TR, *et al.* (2003) Identification of CDH1 germline missense mutations associated with functional inactivation of the E-cadherin protein in young gastric cancer probands. Hum Mol Genet 12:575-582.

Wang GY, Lu CQ, Zhang RM, Hu XH and Luo ZW (2007) The E-cadherin gene polymorphism -160C → A and cancer risk: A HuGe review and meta-analysis of 26 case-control studies. Am J Epidemiol 167:7-14.

Yang X, Wu J, Lu J, Liu G, Di G, Chen C, Hou Y, Sun M, Yang W, Xu X, *et al.* (2015) Identification of a comprehensive spectrum of genetic factors for hereditary breast cancer in a Chinese population by next-generation sequencing. PLoS One 10:e0125571.

Zhan Z, Wu J, Zhang J, Yang Y, Ton S, Zhang C, Li J, Yang X and Dong W (2012) CDH1 gene polymorphisms, plasma CDH1 levels and risk of gastric cancer in a Chinese population. Mol Biol Rep 39:8107-8113.

Internet Resources

Forman D and Bray F (2012) Globocan v1.0 Cancer Incidence and Mortality Worldwide - IARC CancerBase No 11. International Agency for Research on Cancer-France, Lyon, http://globocan.iarc.fr (accessed Febr 10, 2014).

Associate Editor: Maria Isabel Achatz

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.