

**INSTITUTO OSWALDO CRUZ**  
**Mestrado em Ciências de Saúde**

**Estudo de polimorfismos dos genes PKLR e VDR na  
população de Moçambique**

**Graça Salomé**

**Rio de Janeiro**  
**2010**



14095

Instituto Oswaldo Cruz  
Mestrado em Ciências de Saúde

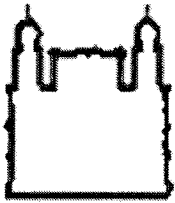
**Estudo de polimorfismos dos genes *PKLR* e *VDR* na  
população de Moçambique**

$\frac{69}{1}$

Graça Salomé

Rio de Janeiro

2010



**Ministério da Saúde**

**FIOCRUZ**

**Fundação Oswaldo Cruz**

**INSTITUTO OSWALDO CRUZ**  
**Pós-graduação em Ciências de Saúde**

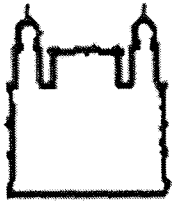
**GRAÇA SALOMÉ**

Estudo de polimorfismos dos genes *PKLR* e *VDR* na população de Moçambique

Dissertação apresentada ao Instituto Oswaldo Cruz como parte dos requisitos para obtenção do título de Mestre em Ciências, área de Biologia Molecular.

Orientadores: Prof. Milton Ozório Moraes  
Prof. Ricardo Thompson

Rio de Janeiro  
2010



Ministério da Saúde

**FIOCRUZ**

**Fundação Oswaldo Cruz**

**INSTITUTO OSWALDO CRUZ**  
**Pós-graduação em Ciências de Saúde**

**GRAÇA SALOMÉ**

Estudo de polimorfismos dos genes *PKLR* e *VDR* na população de Moçambique

Orientadores: Dr. Milton Ozório Moraes  
Dr. Ricardo Thompson

Aprovada em: \_09/\_11/\_2010\_

Banca examinadora:

Dr. Wilson Savino – Fundação Oswaldo Cruz – Presidente

Dra. Sónia Enosse – Instituto Nacional de Saúde – Titular

Dr. Bettencourt Capece – Universidade Eduardo Mondlane – Titular

Dr. Ilesh Jani – Instituto Nacional de Saúde – Suplente

Dra. Cynthia Cardoso – Fundação Oswaldo Cruz - Revisora

Rio de Janeiro, \_09/\_11/\_2010

*Dedico este trabalho:*

*A minha mãe e a memória do meu pai, pessoas que me explicaram a importância do conhecimento para a vida*

*Ao Manuel, Daniel, Valdo*

*Aos meus irmãos, sobrinhos e restante família.*

## **Agradecimentos**

Ao meu orientador, Professor Milton Moraes sempre presente e ao meu co-orientador Professor Ricardo Thompson pelo apoio incondicional e pelos ensinamentos que me emprestaram durante estes dois anos de trabalho.

A minha mãe Victória, minha razão de ser, por tudo mais.

Ao Manuel meu companheiro, Valdo e Daniel meus queridos, meus irmãos, sobrinhos e restantes familiares pelo apoio e pela compreensão pelos muitos momentos de ausência.

A Cynthia Cardoso pela valiosa contribuição em todo o trabalho, especialmente pela revisão minuciosa deste trabalho.

Ao pessoal do Laboratório de Hanseníase do IOC: Cynthia, Alejandra, Caroline Marques, Caroline Carvalho, Suelen, Luana, Paula, Lucia, Cláudia, Ana Beatriz, Ricardo, Marcelo pela amizade, convivência e apoio prestado durante o período de trabalho no laboratório e não só.

Ao Luís Filipe, a Jacinta, a Tânia, e a todo o pessoal dos bancos de sangue, da Imunologia e Parasitologia do INS e da Bioquímica e da Parasitologia da Faculdade de Medicina da UEM pelo apoio dado durante o período de colheita e processamento das amostras em Moçambique.

A Amina, Helena Chavale, Rafael, Cinthia Semá e Alice e os restantes colegas do mestrado pelo apoio material e didático e pelos bons momentos de convivência que tivemos.

Ao Dr Sérgio Chibute e restantes colegas do sector de Bioquímica da Faculdade de Medicina da UEM pelo apoio e pela força que me deram neste período.

Aos coordenadores deste mestrado pela oportunidade que me deram de frequentá-lo.

Ao pessoal técnico-administrativo ligado a Faculdade de Medicina, ao INS e a FIOCRUZ pelo apoio prestado.

A Margarida, Fátima e Dulce minhas companheiras e irmãs de todos os momentos pelo apoio e força dados incansavelmente apesar da minha ausência.

A todos os demais que não mencionei mas que de alguma forma contribuíram para que este trabalho fosse realizado com sucesso o meu muito obrigado.

A DEUS por tudo.

*“A satisfação está no esforço feito para alcançar o objectivo, e não em te-lo alcançado”.*

*Ghandi*

## LISTA DE ABREVIATURAS E SÍMBOLOS

A – Adenina

*Acil* – Enzima de restrição isolada de *Arthrobacter citreus*

AD – Depois de Cristo

*Apal* – Enzima de Restrição isolada de *Acetobacter pasteurianus*

ATP – Adenosina Trisfosfato

BSA – Albumina de soro Bovino

*BsmI* – Enzima de Restrição isolada de *Bacillus stearothermophilus*

C – Citosina

CD - Molécula diferenciadora de subgrupos celulares (“Cluster differentiator”)

cDNA – DNA complementar

CEU – Residentes do Utah na Pensilvania com ancestrais no Oeste e Norte da Europa

CHB – Chineses Ham de Beijing

dATP – desoxiadenosina trifosfato

dbSNP – Base de dados de polimorfismos de um único nucleótido

dCTP – desoxicitidina trifosfato

dGTP – desoxiguanosina trifosfato

DNSP – Direcção Nacional de Saúde Pública

DNA – Ácido desoxiribonucléico

dNTP – Desoxinucleósido trifosfato

dTTP – desoxitimidina trifosfato

EC – Comissão de Enzimas

EDTA – Ácido etilenodiamino tetra acético

FEP – Fosfoenolpiruvato

FIOCRUZ – Fundação e Instituto Oswaldo Cruz

*FokI* – Enzima de Restrição isolada de *Flavobacterium okeanoikoites*

$F_{ST}$  – Índice de fixação usado como medida de diferenciação genética entre subpopulações

G – Guanina

GATA – Caixa na região promotora de genes a que se ligam factores de transcrição

GWAS – Estudos de associação do genoma em larga escala

HCB – Hospital Central da Beira

HCl - Ácido Clorídrico

HCM – Hospital Central de Maputo



HCN – Hospital Central de Nampula  
HGJM – Hospital Geral José Macamo  
HGM – Hospital Geral de Mavalane  
HLA – Antígeno Leucocitário Humano  
HPP – Hospital Provincial de Pemba  
HRC – Hospital Rural de Chókwe  
EHW – Equilíbrio de Hardy-Weinberg  
IL – Interleucina  
INE – Instituto Nacional de Estatística  
INF – Interferon  
INS – Instituto Nacional de Saúde  
IOC – Instituto Oswaldo Cruz  
JPT – Japoneses de Tokyo  
Kb – kilobases  
mfa – Menor frequência alélica  
MCP – Proteína quimiotática para Monócitos  
MgCl<sub>2</sub> – Cloreto de Magnésio  
MISAU – Ministério da Saúde  
RNA<sub>m</sub> – Ácido ribonucléico mensageiro  
mtDNA – Ácido desoxiribonucléico mitocondrial  
NEB – New England Biolabs  
*NlaIII* – Enzima de restrição isolada de *Neisseria lactamica*  
PAMP – Padrões Moleculares Associados a Patógenos  
pb – Pares de bases  
PCR – Reacção de polimerase em cadeia  
PK – Piruvato quinase  
PKL – Piruvato quinase isoenzima hepática  
*PKLR* – Gene que codifica a piruvato quinase no fígado e reticulócito  
PKM<sub>1</sub> – Piruvato quinase isoenzima do músculo  
PKM<sub>2</sub> – Piruvato quinase isoenzima dos tecidos fetais  
PKR – Piruvato quinase isoenzima eritrocitária  
PNCM – Programa Nacional de Controlo da Malária  
PNCT – Programa Nacional de Controlo da Tuberculose

PRR – Receptor de Reconhecimento Padrão  
RFLP – Polimorfismo do comprimento do fragmento de restrição  
RNA – Ácido ribonucléico  
rpm – Rotações por minuto  
rt-PCR – PCR em tempo real  
rs – número de referência para identificação de SNPs  
SDS – Dodecil Sulfato de Sódio  
SNP – Polimorfismo de base única  
STR – sequência repetitiva curta de DNA  
T – Timina  
Taq – *Thermus aquaticus*  
TaqI – Enzima de Restrição isolada de *Thermus aquaticus*  
TATA – sequência na região promotora de genes na qual se ligam factores de transcrição  
TBE – Tris-Borato-EDTA  
TE – Tris-EDTA  
TLR – Receptor semelhante a Toll  
TNF – Factor de Necrose Tumoral  
tRNA – Ácido ribonucléico de transferência  
UTR – Região não traduzida  
VDR – Receptor da vitamina D  
VDRE – Elemento responsivo ao complexo Vitamina D e seu receptor  
YRI – Yoruba em Ibadan na Nigéria

## LISTA DE FIGURAS

<b>Figura 1.1</b> – Mapa de Moçambique ilustrando a divisão político-administrativa e as 3 regiões (1-Sul; 2-Centro e 3-Norte).....	3
<b>Figura 1.2</b> - Mapa de parte da África sub-Ssahariana ilustrando o movimento dos bantu para o Sul de África.....	5
<b>Figura 1.3</b> – Representação esquemática dos principais passos do processo da regulação da transcrição de genes alvo da vitamina D. ....	9
<b>Figura 3.1</b> – Mapa de Moçambique com indicação do locais de colheita de amostras. ....	17
<b>Figura 3.2</b> – Representação esquemática dos genes <i>PKLR</i> e <i>VDR</i> destacando os SNPs estudados.....	19
<b>Figura 4.1</b> – Imagens dos geis após a digestão.....	27
<b>Figura 4.2</b> - Representação gráfica do resultado da discriminação alélica por PCR em tempo real do SNP rs 4760658 do gene <i>VDR</i> .....	31

## LISTA DE QUADROS

<b>Quadro 3.1:</b> Sequência dos iniciadores, condições de ciclagem e tamanho dos fragmentos amplificados em pares de base (pb), utilizados para a amplificação nos sistemas do gene <i>PKLR</i> . .....	21
<b>Quadro 3.2:</b> Condições da reação de digestão utilizadas para determinação dos genótipos nos sistemas que utilizam a técnica de PCR-RFLP.....	22
<b>Quadro 3.3:</b> Sequência dos iniciadores, sondas e condições de ciclagem utilizadas para a amplificação e discriminação alélica por PCR em tempo real.....	24



## LISTA DE TABELAS

<b>Tabela 1.1</b> – Sítios polimórficos do gene <i>PKLR</i> mais frequentemente estudados na população em geral.....	13
<b>Tabela 4.1</b> - Distribuição de indivíduos recrutados para este estudo com base no local de colheita da amostra.....	25
<b>Tabela 4.2</b> – Características dos indivíduos recrutados para este estudo. ....	26
<b>Tabela 4.3</b> – Frequências genotípicas dos SNP rs8847 e rs1052176 na população de Moçambique.....	27
<b>Tabela 4.4</b> – Frequências genotípicas dos SNPs rs8847 e rs1052176 na população de Moçambique.....	28
<b>Tabela 4.5</b> – Frequências genotípicas do rs8847 de Moçambicanos, residentes do Utah na Pensilvania com ancestrais no Norte e Oeste da Europa (CEU), Chineses Han em Beijing (CHB), Afro-Americanos e Brasileiros.....	29
<b>Tabela 4.6</b> – Frequências genotípicas do rs1052176 de Moçambicanos, Chineses Han em Beijing (CHB), Afro-Americanos, Yoruba em Ibadan na Nigéria (YRI), Cabo-Verdianos, São Tomenses, Brasileiros e Portugueses. ....	30
<b>Tabela 4.7</b> – Frequências dos haplótipos formados pelos SNPs rs8847 e rs1052176, totais e por regiões.....	30
<b>Tabela 4.8</b> – Frequências genotípicas dos SNP rs4760658 e rs2228570 na população de Moçambique. ....	32
<b>Tabela 4.9</b> – Frequências alélicas e genotípicas dos SNP rs476065 e rs2228570 por regiões de Moçambique. ....	33

<b>Tabela 4.10</b> – Frequências genotípicas do rs4760658 de Moçambicanos, Afro-Americanos e residentes do Utah na Pensilvânia com ancestrais no Norte e Oeste da Europa (CEU). .....	34
<b>Tabela 4.11</b> – Frequências genotípicas do rs2228570 de Moçambicanos, Portugueses, Brasileiros, Chineses, Venda da África do Sul, Yoruba em Ibadan na Nigéria (YRI) e Afro-Americanos. ....	34
<b>Tabela 4.12</b> – Frequências dos haplótipos formados pelos SNPs rs4760658 e rs2228570, totais e por regiões. ....	35

## RESUMO

As doenças infecciosas constituem um dos principais problemas de saúde no continente Africano especialmente na África sub-Sahariana, da qual Moçambique faz parte. A participação de factores genéticos na susceptibilidade dos humanos às doenças infecciosas é um facto. A identificação de marcadores genéticos de susceptibilidade ou resistência a estas doenças constitui um ponto importante do processo de compreensão dos mecanismos de interacção entre agentes patogénicos e o hospedeiro humano o que poderá permitir a melhoria na abordagem preventiva e terapéutica destas doenças. Estudos de associação envolvendo genes ligados ao metabolismo e ao sistema imune têm revelado a relação entre determinados genótipos e a susceptibilidade ou a resistência a doenças infecciosas como a hanseníase, malária e a tuberculose. Estas duas últimas contribuem em grande medida para a morbi-mortalidade na população Africana particularmente na África sub-Sahariana. O presente trabalho teve como objectivo realizar uma análise descritiva de polimorfismos de base única (SNPs) em dois genes candidatos a associação com doenças infecciosas, nomeadamente o gene da enzima piruvato quinase (*PKLR*) e o do receptor da vitamina D (*VDR*). Foram determinadas as frequências alélicas, genotípicas e hapotípicas de quatro SNPs, dois do gene *PKLR* (rs8847 e rs1052176) e dois do gene *VDR* (rs4760658 e rs2228570) numa amostra composta por 395 indivíduos não relacionados recrutados em três regiões de Moçambique. A genotipagem foi realizada por PCR-RFLP e PCR em tempo real. As frequências dos alelos de maior prevalência foram 0,76 (rs8847G); 0,6 (rs1052176C) 0,88 (rs4760658A) e 0,85 (rs2228570C). Os SNPs do *PKLR* apresentavam frequências genotípicas desviadas do Equilíbrio de Hardy-Weinberg. A comparação entre as frequências obtidas no presente estudo com as de outras populações de diferentes regiões geográficas e grupos étnicos revelou que na maioria dos casos as frequências eram semelhantes às observadas em outras populações de origem Africana e diferentes das de populações Europeias, Americanas e Asiáticas para os quatro SNPs. Os haplótipos mais frequentes formados entre os SNPs de *PKLR* (rs8847/rs1052176) foram G/A e G/C, com 38% cada um, e entre os de *VDR* (rs4760658/rs2228570) foi o A/C, com 75,7%. Não foi observada evidência de desequilíbrio de ligação entre os SNPs de ambos genes. Os resultados deste estudo representam um primeira caracterização destes polimorfismos na população de Moçambique podendo ser úteis para o palnejamento de possíveis estudos de associação com doenças infecciosas.

## ABSTRACT

Mozambique is a country situated in the Sub-Saharan Africa where the infectious diseases are one of the major health concerns. The contribution of genetics factors in the human susceptibility to this type of diseases is now evident, and the identification of genetic markers associated to disease susceptibility or resistance is an important point to understand the mechanisms of host-pathogen interaction. Furthermore, the use of these markers in the clinical routine can help to develop new strategies for disease prevention and treatment. Association between single nucleotide polymorphisms (SNPs) in genes involved in metabolism and immune system have been reported in different studies and some genotypes were linked to susceptibility or resistance to infectious diseases such as leprosy, malaria and tuberculosis. Malaria and tuberculosis have a huge contribution in the morbidity and mortality burden in Africa, particularly in the sub-Saharan region. The aim of this study was to describe polymorphisms of the pyruvate kinase (*PKLR*) and vitamin D receptor (*VDR*) genes. These genes have been previously reported to be associated with infectious diseases. Four SNPs were studied in a sample of 395 unrelated individuals recruited in three different regions of Mozambique. Two polymorphisms were from *PKLR* gene (rs1052176 and rs8847) and two from *VDR* gene (rs2228570 and rs4760658). The samples were genotyped using PCR-RFLP or Real-time PCR. Allele, genotype and haplotype frequencies were determined, and linkage disequilibrium analyses were performed. The allele frequencies found were 0.76, 0.60, 0.88 and 0.85 for the most prevalent alleles rs8847G, rs1052176C, rs4760658A and rs2228570C, respectively. The genotype frequencies of the *PKLR* SNPs showed deviations from the expected under Hardy-Weinberg equilibrium conditions. For all loci studied, the genotype frequencies were in most cases similar to those previously reported in other African populations but significantly different from those found in American, European and Asian populations. The most frequent haplotypes for *PKLR* (rs8847/rs1052176) were G/A and G/C, with 38% each, and for *VDR* (rs4760658/rs2228570) was A/C, with 76%. The SNPs were not in linkage disequilibrium in both genes. These results represent an initial characterization of these SNPs in the population of Mozambique and can be useful for planning possible association studies with infectious diseases.



## SUMÁRIO

Lista de abreviaturas e símbolos.....	vi
Lista de Figuras .....	ix
Lista de Quadros.....	x
Lista de tabelas .....	xi
Resumo .....	xiii
Abstract.....	xiv
Sumário.....	xv
1. INTRODUÇÃO.....	1
1.1. Variantes genéticas e susceptibilidade à doenças.....	1
1.2. Formação do povo de Moçambique .....	2
1.3. Resposta imune à doenças infecciosas .....	6
1.4. Impacto das doenças infecciosas na evolução humana .....	7
1.5. Receptor da vitamina D .....	8
1.6. Piruvato quinase .....	11
1.7. Justificativa.....	13
2. OBJECTIVOS .....	15
2.1. Objectivo geral: .....	15
2.2. Objectivos específicos:.....	15
3. MATERIAL E MÉTODOS.....	16
3.1. População do estudo.....	16
3.2. Extração do DNA .....	18
3.3. Genotipagem dos indivíduos .....	19
3.3.1. Genotipagem por PCR-RFLP.....	20
3.3.1.1. Amplificação .....	20
3.3.1.1.1. SNP rs1052176.....	20
3.3.1.1.2. SNP rs 8847.....	21
3.3.1.2. Digestão .....	22
3.3.2. Genotipagem por discriminação alélica .....	23
3.4. Análise estatística .....	24
4. RESULTADOS .....	25
4.1. Polimorfismos do gene <i>PKLR</i> .....	26
4.2. Polimorfismos do gene <i>VDR</i> .....	31

5. DISCUSSÃO .....	36
5.1. Polimorfismos do <i>PKLR</i> .....	36
5.2. Polimorfismos do <i>VDR</i> .....	41
6. CONCLUSÕES.....	44
7. Recomendações .....	44
8. REFERÊNCIAS BIBLIOGRÁFICAS .....	45

# 1. INTRODUÇÃO

## 1.1. Variantes genéticas e susceptibilidade à doenças

A diversidade genética humana é um recurso emergente para o estudo de doenças complexas em que contribuem factores genéticos e ambientais. A análise da sequência do DNA do genoma humano demonstra que muitos genes são polimórficos. Um polimorfismo é definido como uma variação no genoma com frequência mínima de 1% na população. Estas variações podem ser de vários tipos, tais como a substituição de um par de nucleótidos (SNP, do Inglês “Single Nucleotide Polymorphism”), ou repetições de uma sequência repetitiva curta de DNA (STR, do Inglês “Short Tandem Repeats”), inserções e deleções, dentre outras. Podem aparecer em regiões codificantes ou não codificantes de um gene podendo afectar a taxa de transcrição, a estabilidade do ácido ribonucléico mensageiro (RNAm) ou actividade de uma dada proteína codificada. Assim a presença de alelos específicos em genes polimórficos poderá influenciar a susceptibilidade ou gravidade de um número de desordens no indivíduos que os detém.

A associação entre variantes genéticas e as doenças tem sido revelada em vários estudos. Na maior parte destes estudos os SNPs foram usados como marcadores genéticos, e são considerados valiosas ferramentas para a pesquisa. No genoma humano os SNPs encontram-se largamente distribuídos e actualmente estima-se que existam mais de 10 milhões (Pacheco & Moraes, 2009). Estas variantes são usadas como marcadores para o mapeamento genético fino em estudos de associação do genoma em larga escala. (GWAS, do Inglês “Genome-Wide Association Studies”) (Collins *et al.*, 1997; Risch & Merikangas, 1996).

Os estudos de associação na maioria das vezes do tipo caso-controle podem revelar uma relação entre um determinado SNP e uma doença, o que pode suportar a ideia de que uma dado gene candidato está de algum modo envolvido com a doença. Neste tipo de estudos de associação compara-se a prevalência de um ou vários marcadores genéticos em dois grupos, um de indivíduos afectados e outro de indivíduos não afectados. Uma grande prevalência de um marcador no grupo dos indivíduos afectados é considerado evidência de associação do marcador a um risco aumentado de desenvolvimento da doença (Collins *et al.*, 1997; Pacheco & Moraes, 2009).

Os polimorfismos genéticos com efeito funcional significativo no produto génico (sendo muitas vezes os que ocorrem nas regiões codificantes) são os mais úteis para os estudos de associação com doenças. Os polimorfismos cujo significado funcional ainda não é conhecido ou que não têm significado funcional são simples marcadores. A interpretação de resultados dos estudos de associação pode de algum modo ser dificultada quando se usam SNPs sem significado funcional conhecido. Assim é de se esperar que estes tenham uma ligação a polimorfismos verdadeiramente funcionais no mesmo ou em genes vizinhos. Deste modo ao se determinar que polimorfismos existem numa determinada área de um gene candidato, é importante compreender como é que estes se relacionam a nível genético e funcional (Valdivielso & Fernandez, 2006).

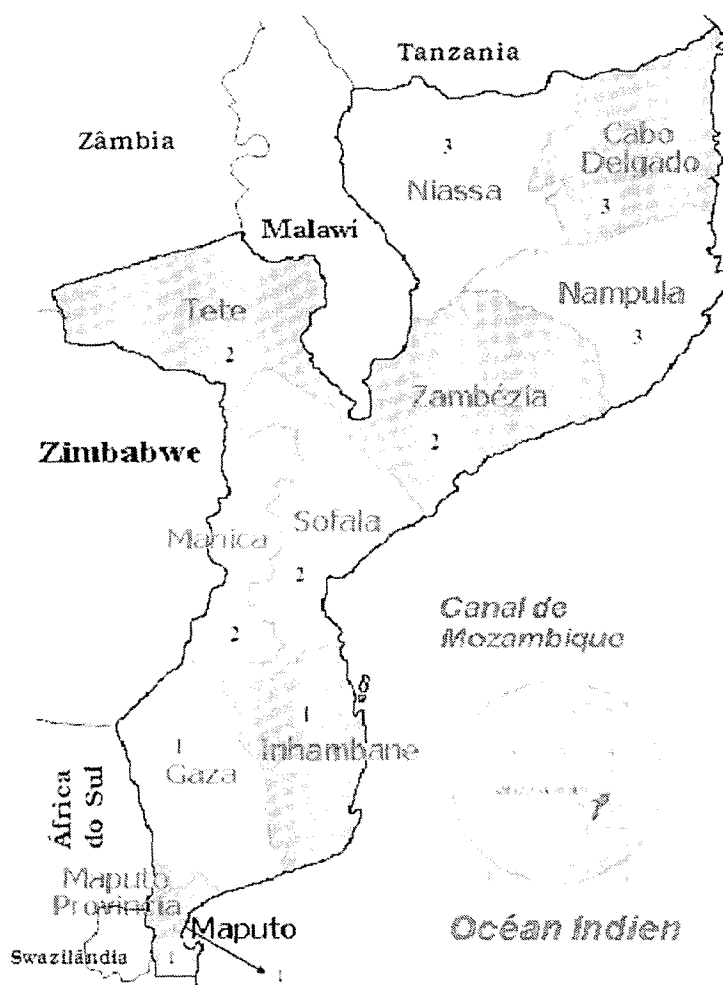
A caracterização das variações genéticas entre indivíduos e entre populações poderá permitir compreender melhor a evolução dos humanos, as diferenças na susceptibilidade às doenças e na resposta a agentes farmacológicos bem como a complexa interação entre factores genéticos e ambientais na geração de fenótipos (Campbell, 2008). Assim, a identificação de SNPs em populações humanas, étnica e geograficamente distintas torna-se importante uma vez que as frequências dos SNPs podem variar entre populações, assim como a sua utilidade como marcadores para estudos de mapeamento de genes (Tishkoff & Williams, 2002).

## **1.2. Formação do povo de Moçambique**

O continente Africano dispõe de enorme diversidade cultural, linguística e genética (Tishkoff & Williams, 2002). O padrão das variações genéticas nas populações africanas modernas é influenciado pela história demográfica (exemplo: alterações no tamanho populacional, a miscigenação e a migração), bem como por forças específicas de um dado locus tais como a selecção natural, as recombinações e as mutações. A história populacional complexa, as mudanças climáticas e dietéticas, bem como a exposição às doenças infecciosas tornam a África numa região importante para o estudo da diversidade genética humana uma vez que estes factores resultam numa grande variação genotípica e fonotípica destas populações (Campbell, 2008).



Moçambique é um país situado no sudeste da África. A Leste é banhado pelo canal de Moçambique no oceano Índico. No Norte faz fronteira com a Tanzânia; no Noroeste com o Maláwi e Zâmbia; no Oeste com o Zimbabwé e Swazilândia e no Sudoeste com a África do Sul. O país está dividido em três regiões onde se distribuem as 11 províncias que o compõem, nomeadamente Cabo Delgado, Niassa e Nampula no Norte; Zambezia, Tete, Manica e Sofala no Centro; e Inhambane, Gaza, Maputo-Província e Maputo-Cidade no Sul (INE, 2009) (Figura 1.1).

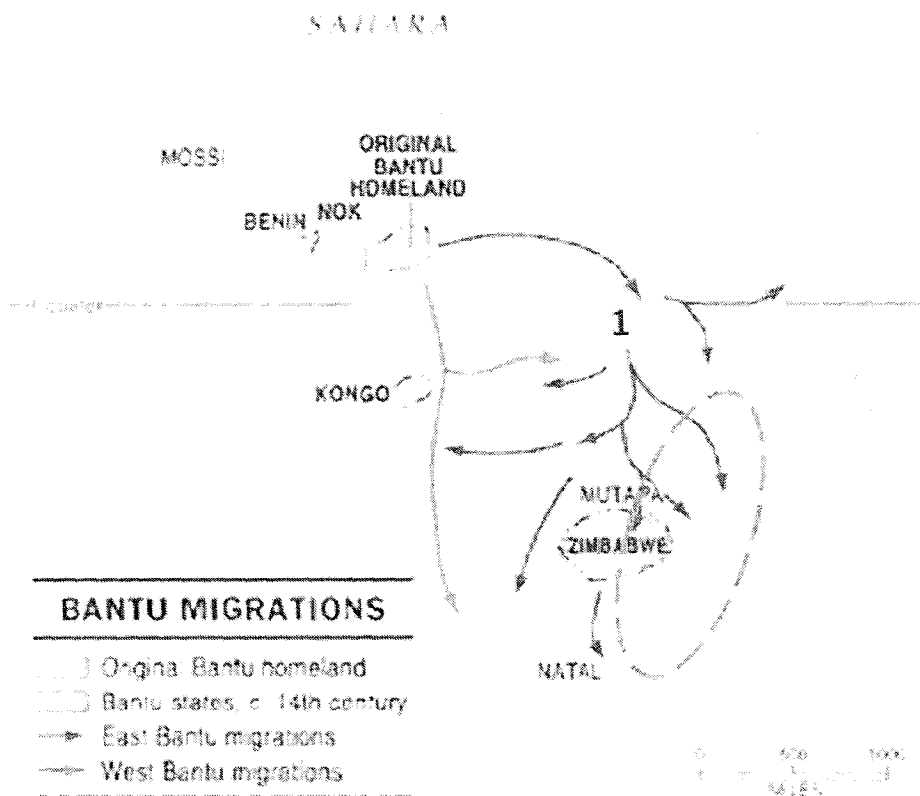


**Figura 1.1** – Mapa de Moçambique ilustrando a divisão político-administrativa e as 3 regiões (1-Sul; 2-Centro e 3-Norte). No inset se encontra o mapa do continente Africano no globo onde se destaca Moçambique no Sudoeste. Adaptado de [http://www.freewebs.com/africa\\_mozambique/pictures.htm](http://www.freewebs.com/africa_mozambique/pictures.htm) e [http://www.nationsonline.org/oneworld/map/mozambique\\_map.htm](http://www.nationsonline.org/oneworld/map/mozambique_map.htm).

Os primeiros habitantes de Moçambique eram caçadores e colectores chamados San e ancestrais dos Khoisan. Entre os séculos 1 e 5 AD, houve várias ondas de migração dos povos Bantu da região dos grandes lagos (núcleo proto-Bantu 3) para o Sudeste da África (Figura 1.2), incluindo a região que é hoje Moçambique. Durante este período houve substituição gradual das comunidades primitivas pelas comunidades bantu e nalguns casos a integração das comunidades com resultante miscigenação (Pereira *et al.*, 2001). Durante a migração bantu a região que hoje é Moçambique serviu de estrada de migração para o sul. Choques constantes entre diferentes grupos em migração nessa região fizeram com que surgissem desvios para o Oeste e até novas migrações do Sul para o Norte. Isto levou a dispersão considerável dos grupos bantu do Este resultando na existência de grandes grupos étno-linguísticos na região Oeste ao contrário dos grupos da região Este que são relativamente reduzidos (Chabert, 2003). Entre os séculos IX e XIII a expansão mercantil árabo-persa para a região da África Oriental permitiu a fixação de mercadores Árabes da península Arábica, Persas do Golfo Pérsico e Baneanes da Índia. Isto propiciou a miscigenação entre os mercadores asiáticos e a população Africana com o surgimento de bolsas de mestiçagem ao longo do litoral principalmente na região Norte de Moçambique. Mais tarde por volta do século XVI os portugueses colonizaram Moçambique, que resultou numa miscigenação entre os povos das tribos Bantu e a população de origem portuguesa (Assane *et al.*, 2010; Pereira *et al.*, 2002; Serra, 2000). O facto de Moçambique ter servido de estrada de migração na época da expansão bantu, as migrações internas e externas durante e depois do término da guerra civil, a migração sazonal de indivíduos do sexo masculino para o trabalho nas minas da África do Sul bem como a migração de populações do Zimbabwe para Moçambique especialmente na zona centro de Moçambique antes da independência do Zimbabwe, contribuíram para a distribuição étnica da população de Moçambique na actualidade.

Actualmente Moçambique tem cerca de 20 milhões de habitantes de acordo com os dados censo populacional de 2007 (INE, 2009). A língua oficial é o Português, existindo porém uma diversidade de grupos etno-linguísticos. Os principais grupos etno-linguísticos em Moçambique são: Macua-Lomue (37% da população) residentes em Nampula, Sul de Cabo Delgado, Norte da Zambézia e Sudeste do Niassa; Makonde em Cabo Delgado; Chuabo na Zambézia; Yao e Chewa/Nyanja no Niassa e parte de Tete; Sena em Sofala e parte da Zambézia; Ndau no sul de Sofala e Manica e norte de Gaza e Inhambane; Chona em Manica e Tsonga no Sul de Moçambique (Changanas em Gaza e Maputo, Rongas em Maputo, Bitongas, Xitswas e Chopis em Inhambane) (Serra, 2000). Esta distribuição dos grupos étnicos não obedece as fronteiras

político-administrativas, de modo que muitas vezes os mesmos partilham território de duas ou mais regiões geográficas distintas nacionais e/ou internacionais.



**Figura 1.2-** Mapa da ilustrando o movimento dos bantu para a região sul de África. O número 1 indica a região dos grandes lagos donde partiram os bantu para a região sudeste. A elipse circunscribe a região onde actualmente se situa Moçambique. Adaptado de <http://mathildasanthropologyblog.wordpress.com/2008/04/29/the-bantu-people/>.

### 1.3. Resposta imune à doenças infecciosas

As doenças infecciosas constituem um dos principais problemas de saúde no continente Africano especialmente na África sub-sahariana da qual Moçambique faz parte e onde doenças como a malária e a tuberculose têm altas taxas de morbidade e mortalidade (WHO, 2009b; WHO, 2009c).

O sistema imune inato fornece uma forma rápida de defesa do hospedeiro contra patógenos microbianos. A resposta imune inata aos agentes infecciosos é dada principalmente pela acção das barreiras físicas como a pele e as membranas mucosas, pelos fagócitos que reconhecem padrões moleculares característicos de microrganismos e pelos factores humorais presentes na circulação como as proteínas do complemento, os anticorpos naturais e as proteínas de fase aguda (Abbas *et al.*, 2008). O reconhecimento dos padrões moleculares associados a patógenos, (PAMP do Inglês “Pathogen-Associated Molecular Patterns”), pelos fagócitos é feito mediante receptores expressos nessas células denominados Receptores de Reconhecimento Padrão, (PRR do Inglês “Pattern Recognition Receptors”). Estes receptores podem ser agrupados em várias classes dentre as quais se destacam os receptores semelhantes a Toll (TLR do Inglês “Toll Like Receptors”), que encontram-se ligados a vias de transdução de sinal intracelulares que culminam com a activação de várias respostas celulares contra os microrganismos (Akira *et al.*, 2006; Janeway Jr., 1989). Diversos genes que codificam citocinas inflamatórias tais como o factor de necrose tumoral (TNF do Inglês “Tumor necrosis factor”), interleucinas 1 e 12 (IL do Inglês “interleukin” – IL-1, IL-12), quimiocinas (IL-8, proteína quimiotática de monócitos - MCP-1 do Inglês “Monocyte chemotatic protein”), moléculas de adesão endotelial (Selectina – E), moléculas co-estimuladoras (CD80, CD86) e citocinas anti-virais (Interferão alfa e beta – INF $\alpha/\beta$  do Inglês “Interferon”) são expressos por acção de factores de transcrição activados por via de TLR (Abbas *et al.*, 2008). Para além destes genes, os TLR activam a expressão do receptor da vitamina D (VDR, do Inglês “Vitamin D Receptor”) e da enzima Cyp27B1, que cataliza a conversão da vitamina D<sub>3</sub> inactiva (25-hidroxicolecalciferol ou 25-hidroxivitamina D<sub>3</sub>) para a sua forma activa a 1,25-dihidroxicolecalciferol (1,25-dihidroxivitamina D<sub>3</sub>). A forma activa da Vitamina D<sub>3</sub> actua por meio do VDR nos monócitos e macrófagos activando nestes a actividade microbida através da transactivação do gene que codifica um péptido antimicrobiano, a catelicidina (Liu *et al.*, 2006). A resposta inata aos agentes infecciosos pode ser influenciada por variações nos genes que

codificam proteínas com actividade específica no sistema imune tornando o indivíduo portador das mesmas susceptível ou resistente a infecção ou ao desenvolvimento de formas graves ou ligeiras de infecção. Variações em genes ligados ao metabolismo também podem estar associados a resposta do hospedeiro devido a alterações de toda a sorte. Um exemplo é a deficiência da enzima piruvato quinase (PK, do Inglês “Pyruvate Kinase”) que foi associada a susceptibilidade a infecção por *Salmonella typhimurium* (Roy *et al.*, 2007). Indivíduos com esta deficiência possuem sobrecarga de ferro devido a hemólise massiva que caracteriza esta doença. O ferro actua como factor de crescimento para *Salmonella* e quando em excesso aumenta a susceptibilidade a infecção por esta bactéria. Estudos *in vitro* mostraram que a suplementação de ferro prejudica a produção de oxido nítrico e TNF nos macrófagos resultando no aumento da proliferação bacteriana (Nairz *et al.*, 2007).

#### **1.4. Impacto das doenças infecciosas na evolução humana**

O envolvimento de factores genéticos na susceptibilidade dos humanos às doenças infecciosas é um facto. J.B.S.Haldane foi o primeiro a propor que as doenças infecciosas constituem a principal força selectiva na evolução humana (Lederberg, 1999). Para que uma determinada doença se tornasse uma força selectiva na evolução da população, o gene envolvido deveria ter um impacto significativo por longos períodos de tempo, influenciando assim a morbidade e mortalidade antes da idade reprodutiva. A tuberculose foi a maior força selectiva na evolução das populações da Europa Ocidental enquanto que a malária teve o mesmo papel na África (Cooke & Hill, 2001; Miller, 1994). Estudos de associação demonstraram a contribuição de variantes alélicas e haplotípicas de genes que codificam o complexo HLA de classe II (do Inglês “Human Leucocyte Antigen”) e o VDR na susceptibilidade a tuberculose (Goldfeld *et al.*, 1998; Lombard *et al.*, 2006). Em relação à malária estudos experimentais “*in vitro*” usando eritrócitos humanos e “*in vivo*” usando camundongos assim como estudos genéticos de associação, demonstraram que variantes polimórficas de alelos em genes que codificam proteínas eritrocitárias (Hemoglobina S, Glucose-6-fosfato desidrogenase, Piruvato quinase etc.) e os antígenos do HLA comuns na África ocidental conferem protecção contra formas graves de malária (Ayi *et al.*, 2008; Durand & Coetzer, 2008a; Hill *et al.*, 1991; Min-Oo *et al.*, 2003).

A malária é endémica em Moçambique havendo áreas mesoendémicas e outras hiperendémicas. A transmissão é contínua sendo mais intensa durante a época chuvosa. Este padrão de transmissão pode variar em dependência da quantidade de precipitação, e da temperatura ambiente. Contribuem também para esta variação as condições específicas locais como por exemplo a altitude. O *Plasmodium falciparum* é o agente responsável por 90% dos casos de malária e pelos casos mais graves de doença (MISAU-PNCM, 2007). Em 2009 ocorreram cerca de 4.300.000 casos de malária no país. Com excepção das províncias de Maputo e Maputo-Cidade as demais províncias registaram mais de 200.000 casos, estando as províncias de Zambézia e Nampula no topo com mais de 600.000 e 1.000.000 de casos respectivamente (MISAU-DNSP, 2009).

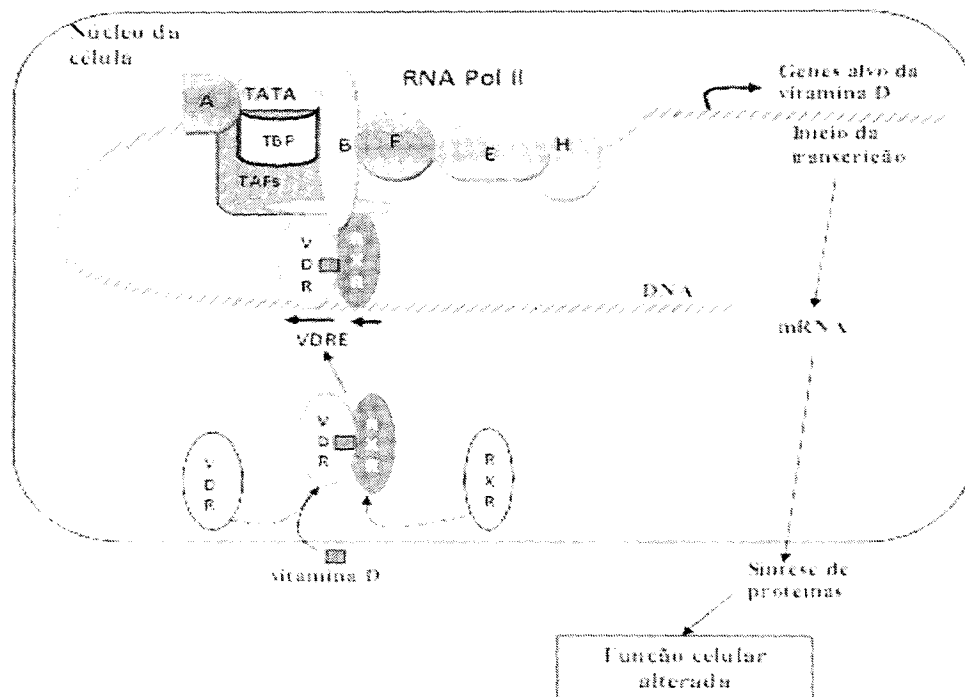
Em relação a tuberculose, pode-se referir que constitui um dos principais problemas de saúde pública em Moçambique, sendo uma das principais causas de morbi-mortalidade. Dentre os 22 países que contribuem com o maior peso da doença no mundo, Moçambique situa-se na 18ª posição (MISAU-PNCT, 2007). Em 2009 registaram-se 45.493 casos de tuberculose em Moçambique (MISAU-DNSP, 2009).

Há bastante tempo que se sabe que os indivíduos respondem de maneira diferente a determinadas infecções específicas devido a diferenças na susceptibilidade às mesmas (James *et al.*, 1932). Estas diferenças também podem ser observadas entre populações. Portanto a compreensão das bases moleculares destas diferenças na susceptibilidade às doenças infecciosas poderá permitir o aprofundamento dos conhecimentos sobre a patogénese das mesmas e trazer alguma contribuição para o controle das mesmas a nível mundial (Cooke & Hill, 2001).

### **1.5. Receptor da vitamina D**

O receptor da vitamina D, é uma fosfoproteína, responsável pelas acções biológicas da 1,25-dihidroxicolecalciferol ou calcitriol (vitamina D<sub>3</sub>). Estas acções são responsáveis pela regulação de eventos transcricionais dentro do núcleo das células alvo. O VDR é um factor de transcrição dependente do ligando a semelhança do receptor das hormonas esteróides, da vitamina A e das hormonas tiroideas, todos pertencentes a superfamília dos receptores nucleares reguladores da transcrição (Colston & Feldman, 1980; Haussler & Norman, 1969). Os diversos efeitos biológicos imputados a Vitamina D são atribuídos a ubíqua distribuição tecidual do VDR bem

como a regulação altamente selectiva da expressão génica por esta hormona em tipos celulares específicos, visto que o VDR é expresso principalmente nas células que participam no metabolismo ósseo, na modulação da resposta imune e na regulação da proliferação e diferenciação celular (Verstuyf *et al.*, 2010). Dentre os principais passos envolvidos no controle da transcrição dos genes através do VDR listam-se a união ao ligando, a heterodimerização com o Receptor retinóide X, a associação do heterodímero com os elementos responsivos a vitamina D (VDRE, do Inglês “Vitamin D Responsive Elements”) situados na região promotora dos genes alvo, e o recrutamento de outras proteínas nucleares para o complexo transcripcional pré-iniciação (Kerner *et al.*, 1989; Lian *et al.*, 1989; Morrison *et al.*, 1989) (Figura 1.3).



**Figura 1.3** – Representação esquemática dos principais passos do processo da regulação da transcrição de genes alvo da vitamina D. Adaptado de Malloy *et al.*, 1999.

Alterações ao nível do gene *VDR* podem resultar em defeitos na activação de genes, afectando o metabolismo mineral, a proliferação celular, a resposta imune dentre outras funções. Variações neste sistema foram ligadas a várias doenças incluindo a osteoartrite, diabetes, cancer, urolitíase e, doenças infecciosas tais como a tuberculose e a hanseníase (Banoei *et al.*, 2010; Goulart *et al.*, 2006; Lewis *et al.*, 2005; Wilkinson *et al.*, 2000). Sabe-se que mutações deletérias

no gene *VDR* causam raquitismo resistente a vitamina D activada, uma doença monogenética rara (Malloy *et al.*, 1999; Whitfield *et al.*, 1996 ). Existem variações mais subtis do gene *VDR*, os polimorfismos, sendo os mais frequentes os SNPs, mas o significado delas ainda não foi sistematicamente analisado e os efeitos nos níveis e função do VDR são desconhecidos na maioria dos casos. Nalguns casos como o da tuberculose sabe-se que a influência dos SNPs do gene *VDR* na actividade do VDR é colectiva e não individual como foi comprovado pelos resultados de estudos de associação que mostraram associação com haplótipos do VDR e não com os SNPs de forma individual (Lombard *et al.*, 2006; Uitterlinden *et al.*, 2004).

Os polimorfismos mais estudados do gene *VDR* são os SNPs rs2228570 (*FokI*), rs1544410 (*BsmI*), rs7975232 (*ApaI*), e o rs731236 (*TaqI*). O rs2228570 que foi descrito pela primeira vez em 1991(Saijo *et al.*, 1991), resulta da substituição do alelo T pelo C e localiza-se na região codificadora do exon 2, eliminando o primeiro dos dois potenciais sítios de iniciação da transcrição com conseqüente aparecimento de duas proteínas distintas. Na presença do alelo T (ATG), há produção da variante mais longa com 427 aminoácidos e na presença do alelo C (ACG) há produção da variante com menos 3 aminoácidos, sendo esta a mais activa (Arai *et al.*, 1997). O rs731236 resulta da substituição do alelo T pelo C, levando a uma mutação sinónima no códon 352 no exon 9. Os restantes polimorfismos acima mencionados localizam-se em regiões não codificadoras.

O gene do *VDR* localiza-se no braço longo do cromossoma 12 (12q13), tem cerca de 100Kb e 14 exons. Seis exons constituem a região 5'- não traduzida e os restantes oito exons codificam os componentes estruturais do VDR. Múltiplos promotores controlam a transcrição do gene VDR e com especificidade tecidual. Estes caracterizam-se pela ausência da caixa TATA iniciadora, pela riqueza em GC e pela presença de sítios de ligação para vários factores de transcrição (Crofts *et al.*, 1998).

As frequências alélicas e genóticas dos SNPs deste gene variam entre populações de regiões e grupos étnicos diferentes (Lombard *et al.*, 2006; Mory *et al.*, 2009; Pan *et al.*, 2009; Zmuda *et al.*, 2003). Contudo existem poucos dados sobre a distribuição destas frequências em populações em África, não existindo informação sobre Moçambique em particular.



## 1.6. Piruvato quinase

A piruvato quinase (Adenosina trifosfato: piruvato 2-o-fosfotransferase) é a enzima que cataliza a última reacção da via glicolítica em que o fosfoenolpiruvato (FEP) é convertido em piruvato com concomitante formação de uma molécula de adenosina trifosfato (ATP) (Boiteux & Hess, 1981). A PK desempenha um papel importante e central no metabolismo celular uma vez que o piruvato, produto da reacção por ela catalizada, está envolvido numa série de vias metabólicas. O funcionamento integral desta enzima é crucial no metabolismo dos eritrócitos que dependem única e exclusivamente da glicólise para a obtenção de energia. O ATP formado na reacção catalizada pela PK representa 50% da energia produzida na glicólise e este é necessário para a manutenção da integridade e função celular eritrocitária.

Nos humanos existem quatro isoenzimas da PK (PKM<sub>1</sub>, PKM<sub>2</sub>, PKL e PKR) distribuídas por tecidos específicos (Fothergill-Gilmore & Michels, 1993; Hall & Cottam, 1978) e exibindo propriedades cinéticas distintas. As isoenzimas PKM<sub>1</sub> e PKM<sub>2</sub> são produtos de dois RNAs resultantes de processamento alternativo do transcrito primário do gene *PKM* (Noguchi *et al.*, 1986) que localiza-se no braço longo do cromossoma 15 (15q22) (Tani *et al.*, 1988) e contém 12 exons. A PKM<sub>1</sub> é expressa no tecido muscular esquelético e cardíaco, no cérebro, enquanto que a PKM<sub>2</sub> é expressa primariamente nos tecidos fetais em proliferação e é progressivamente substituída por outras isoformas específicas de cada tecido, apesar de permanecer a forma principal nos rins, leucócitos, plaquetas, pulmões, baço e tecido adiposo (Takegawa *et al.*, 1983). As isoenzimas PKL e PKR são produtos de RNAs diferentes apenas na sequência da extremidade 5'. Isto resulta do uso de promotores alternativos específicos para cada tecido durante a transcrição do gene *PKLR* que codifica ambas (Noguchi *et al.*, 1987). A PKL é principalmente expressa no fígado, sendo expressa em menor quantidade no córtex renal, intestino delgado e nas células beta das Ilhotas de Langerhans no pâncreas, enquanto que a PKR é expressa exclusivamente nos eritrócitos. O gene *PKLR* localiza-se no braço longo do cromossoma 1 (1q21) (Satoh *et al.*, 1988) e possui mais de 9,5 Kb. A região codificante contém 12 exons, sendo que o primeiro e o segundo são especificamente transcritos para PKL e PKR respectivamente enquanto os últimos 10 exons são partilhados pelos dois RNAs (Kanno *et al.*, 1992a). O cDNA da PKR possui 2060pb que codificam 574 aminoácidos; o promotor situado na região a montante da extremidade 5' do primeiro exon possui duas caixas CAC e quatro motivos GATA. O motivo GATA mais proximal liga-se ao factor de transcrição eritróide GATA-1 que é

essencial para a actividade promotora do *PKR* (Manco *et al.*, 2000). A actividade promotora basal localiza-se na região proximal -120pb e o estimulador situa-se na região entre -120 e -270pb, portanto na região distal 150pb (Kanno *et al.*, 1991).

Mutações no gene *PKLR* podem levar a deficiência de piruvato quinase, que é transmitida como um traço autossómico recessivo (Kanno *et al.*, 1991). Os indivíduos portadores destas mutações podem ser homozigóticos, heterozigóticos compostos para dois alelos mutantes, ou heterozigóticos. As alterações no gene *PKLR* podem resultar em modificações nas enzimas do fígado e do eritrócito porém as manifestações clínicas das alterações são mais pronunciadas nos eritrócitos, uma vez que a deficiência nos hepatócitos é compensada pela contínua produção da isoenzima hepática ou pela expressão compensatória da PKM<sub>2</sub> (Lenzner *et al.*, 1997a; Zanella *et al.*, 2005). O quadro clínico ocorre apenas nos homozigóticos ou heterozigóticos compostos, isto é, nos indivíduos que são heterozigotos para as duas mutações. Os indivíduos heterozigóticos, salvo raras excepções, não apresentam anormalidades hematológicas (Fermo *et al.*, 2005).

Cento e oitenta mutações do gene *PKLR* associados a anemia hemolítica não esferocítica hereditária já foram descritos (Pissard *et al.*, 2006). A maior parte das mutações afecta regiões codificantes; apenas duas variantes de mutações foram descritas em regiões promotoras e funcionalmente caracterizadas (Manco *et al.*, 2000; van Wijk *et al.*, 2003). Vários sítios polimórficos foram identificados no gene *PKLR*. Informações sobre esses SNPs já identificados em várias populações podem ser acessadas em bancos de dados de SNPs como o “dbSNP”, do Inglês “Database of Single Nucleotide Polymorphisms”. Os polimorfismos mais frequentemente estudados estão apresentados na Tabela 1.1, sendo o mais estudado o rs1052176 (1705A/C). A frequência da deficiência da PK e a natureza das mutações é variável em diferentes grupos étnicos e áreas geográficas. A deficiência de PK tem distribuição geográfica mundial com predomínio na população Caucasiana do Norte da Europa. Apesar da influência do gene *PKLR* na susceptibilidade à malária, muito poucos estudos sobre as mutações e polimorfismos deste gene foram feitos em populações Africanas (Alves *et al.*, 2010; Durand & Coetzer, 2008b; Manco *et al.*, 2001a). Em Moçambique foi feito um estudo (Machado *et al.*, 2010) em que foram determinadas as frequências alélicas de 13 SNPs apenas na população da Província e cidade de Maputo.

**Tabela 1.1** – Sítios polimórficos do gene *PKLR* mais frequentemente estudados na população em geral

<b>Sítio polimórfico (cDNA)</b>	<b>Localização</b>	<b>Referência</b>
T <sub>10/19</sub>	Intron 10	Lenzner et al., 1997b
Microsatélite ATT	Intron 11	Lenzner et al., 1994b
1705 A/C	Exon 12	Kanno et al., 1992b
1738 C/T	Exon 12	Zanella et al., 1997

### 1.7. Justificativa

A descoberta de variantes genéticas ligadas a susceptibilidade a doenças pode ser a chave para avanços na medicina preventiva (Valdivielso & Fernandez, 2006). Actualmente, um dos principais usos da epidemiologia molecular talvez seja o estudo de doenças infecciosas, visando, por um lado, a caracterização de patógenos, e por outro, a identificação de indivíduos geneticamente susceptíveis. Dependendo da frequência de determinado genótipo de susceptibilidade numa população exposta ao agente etiológico, essa informação pode ser considerada útil para justificar a implantação de medidas de prevenção da doença no local (Preston, 2003). Estes estudos podem fornecer também informações que possam de algum modo permitir compreender melhor a interação entre o agente patogénico e hospedeiro. Apesar de ser crítica para compreensão da origem do homem moderno e dos factores de risco genéticos para as doenças, poucos estudos sobre a genética humana foram feitos com populações em África. E muito do que se sabe hoje sobre a diversidade genética nestas populações é resultado de estudos numa minoria de grupos populacionais africanos, sendo a maioria destes estudos de DNA mitocondrial e do DNA do cromossoma Y (Campbell, 2008).

Embora a associação entre a susceptibilidade à malária e tuberculose e os polimorfismos do genes *PKLR* e *VDR* tenha sido determinada e as epidemias de tuberculose e malária na África mostrarem uma tendência crescente, poucos estudos de associação e para a descrição das frequências alélicas e genotípicas foram feitos em populações Africanas, particularmente na África sub-Sahariana. Segundo a OMS, as taxas de morbidade e mortalidade por malária e

tuberculose na África são particularmente altas. Dos cerca de 9,3 milhões de casos de tuberculose que ocorreram em 2007, cerca de 30% foram em África (WHO, 2009a). Em 2006 a OMS estimou que 60% dos casos de malária ocorreram na África sub-Sahariana, região que lidera no número de mortes por esta doença especialmente em crianças menores de cinco anos. A existência de informação relativa a frequência destes polimorfismos na população africana poderá permitir o desenho de estudos de associação, cujos resultados poderão ajudar no desenho de novas estratégias de prevenção e novas abordagens terapêuticas para o combate e mitigação dos efeitos das mesmas nestas populações.

## **2. OBJECTIVOS**

### **2.1. Objectivo geral:**

**2.1.1.** Caracterizar as frequências de polimorfismos nos genes *PKLR* e *VDR* na população de Moçambique

### **2.2. Objectivos específicos:**

**2.2.1.** Determinar o genótipo e estimar as frequências alélicas e genotípicas dos SNPs rs8847, rs1052176 do gene *PKLR*, e rs4760658 e rs2228570 do gene *VDR* numa população de Moçambique

**2.2.2.** Comparar as frequências alélicas e genotípicas destes SNPs por regiões de Moçambique e com as de outras populações de diferentes regiões do mundo.

**2.2.3.** Determinar o grau de desequilíbrio de ligação entre os polimorfismos em cada um dos genes.

**2.2.4.** Estimar a frequência dos haplótipos formados em cada gene.

### **3. MATERIAL E MÉTODOS**

#### **3.1. População do estudo**

As 395 amostras de sangue utilizadas neste trabalho foram obtidas entre os meses de Abril e Agosto de 2009 nos bancos de sangue de sete Hospitais das três regiões de Moçambique (Figura 3.1), nomeadamente Hospitais Centrais de Maputo (HCM), Beira (HCB) e Nampula (HCN); Hospital Provincial de Pemba (HPP); Hospital Rural do Chókwe (HRC) e Hospitais Gerais de Mavalane (HGM) e José Macamo em Maputo (HGJM). Na região sul do país encontram-se o HCM, HGJM e o HGM localizados na Província de Maputo-Cidade, e o HRC na Província de Gaza. Na região Centro encontra-se o HCB localizado na Província de Sofala. O HPP e o HCN localizam-se na região Norte do país nas províncias de Cabo Delgado e Nampula respectivamente.

A escolha de dadores de sangue foi devido ao facto de ser um grupo de indivíduos aparentemente saudáveis e de acesso relativamente mais fácil podendo representar indivíduos da comunidade. A abordagem ao dador foi feita de forma sequencial conforme a ordem de chegada ao banco de sangue logo após o interrogatório e exame clínico. Só foram recrutados os dadores aptos para doação no dia em que vinham doar. Foi-lhes dada a folha de informação ao participante e foram convidados a participar do estudo após a explicação dos objectivos do mesmo, dos procedimentos para a colheita e do esclarecimento das dúvidas. Aos que aceitaram participar foi-lhes pedido a assinatura na folha de Termo de consentimento livre e esclarecido (TCLE) em duplicado ficando uma cópia com o dador e outra com o investigador. Durante a doação os voluntários responderam as questões do formulário de recolha de dados (anexo I). Só foram recrutados os dadores com 18 anos ou mais e de origem Moçambicana. O estudo teve a aprovação do Comité Nacional de Bioética para a Saúde de Moçambique (Ref.355/CNBS de 02 de Dezembro de 2008) (anexo II) e do Comité de Ética da FIOCRUZ (Protocolo 531/09).



### 3.2. Extração do DNA

Após a doação das amostras para análises de rotina efectuadas nos bancos de sangue, foram colhidos 5ml de sangue total para um tubo vacutainer K3EDTA. As amostras colhidas fora da Cidade de Maputo foram conservadas à -20°C até a data do transporte para o Departamento de Imunologia do Instituto Nacional de Saúde (INS) no HCM onde foram conservadas a mesma temperatura até a data do processamento. As amostras colhidas na Cidade de Maputo foram transportadas a 4°C até ao Departamento de Imunologia do INS no HCM e processadas no dia da colheita.

O DNA das 395 amostras de sangue total foi extraído usando-se um método padrão de precipitação por sal, com algumas modificações (Miller *et al.*, 1988). Resumidamente, 2ml de sangue total foram misturados com 10 ml de tampão Tris-EDTA (TE) (Tris-HCl 10 mM; EDTA 1 mM) e centrifugados numa centrífuga modelo Labofuge 400, a 2500 rotações por minuto (rpm) durante 20 minutos. O sobrenadante foi descartado e o precipitado foi ressuscitado com 10 ml de TE. Este processo de lavagem foi repetido duas vezes. Após a última lavagem, ao precipitado foram adicionados 600 µl de tampão para lise de leucócitos (Tris-HCl 10 mM; EDTA 2 mM), 80µl de SDS 10% e 9 µl de proteinase K (solução a 25 mg/ml). Deixou-se a mistura incubar na estufa durante a noite a 37°C. Após a incubação adicionou-se 200 µl da solução de acetato de sódio saturada e centrifugou-se a 2500 rpm durante 20 minutos. Transferiu-se o sobrenadante para outro tubo cónico e adicionou-se etanol absoluto (duas vezes o volume obtido). Misturou-se por inversão até que o DNA precipitasse. Recolheu-se o DNA precipitado para um tubo de 1,5 ml ao qual adicionou-se 200µl de etanol a 70% e centrifugou-se na microcentrífuga modelo Allegra™ X-22R, a 12000 rpm durante 5 minutos. Deixou-se o precipitado secar por inversão dos tubos durante 30 minutos a temperatura ambiente. O DNA foi ressuscitado com 300 µl de tampão TE (Tris-HCl 5 mM; EDTA 0,1 mM) e armazenado a -70°C após incubação em banho-maria modelo YCW-01, a 56°C durante 50 minutos. A extracção do DNA foi feita nos laboratórios de Biologia Molecular dos Departamentos de Imunologia e Parasitologia de sangue do INS. O DNA extraído foi conservado a -70°C até a altura que foi transportado para o Brasil. As amostras de DNA foram mantidas a temperatura ambiente durante quatro dias, período que durou o processo de transporte de Maputo para o Rio de Janeiro. No Laboratório de Hanseníase as amostras foram conservadas a -20°C até a altura do processamento. A quantidade e a pureza do

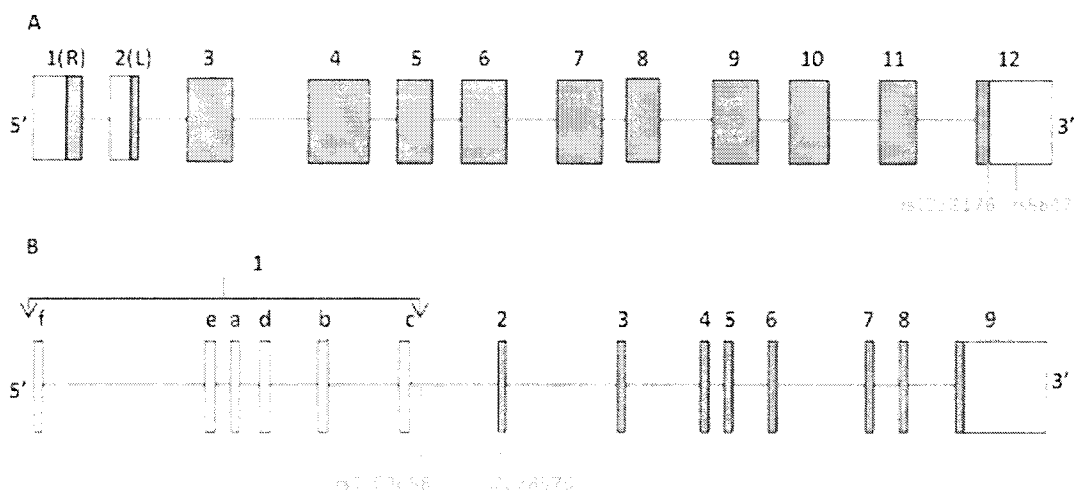


DNA foram determinadas pela medição da absorvância a 260/280 nm usando o espectrofotômetro Nanodrop ND-1000.

### 3.3. Genotipagem dos indivíduos

A genotipagem foi feita no Laboratório de Hanseníase – Grupo de pesquisa em genômica funcional e epidemiologia genética da FIOCRUZ.

A escolha dos polimorfismos de *PKLR* foi feita tendo em conta o facto do rs1052176 representar um tag para um haplótipo composto por 4 SNPs do gene *PKLR*, e ser um dos mais frequentemente usados em vários estudos, e o rs8847 por ser um polimorfismo na região UTR-3', o que poderia gerar uma implicação funcional. Em relação ao polimorfismos de *VDR* o rs2228570 foi escolhido por ser o único que resulta numa proteína polimórfica descrito até ao momento e o rs4760658 por representar um tag para um haplótipo composto por 4 SNPs deste gene. A Figura 3.2 mostra a representação esquemática dos dois genes e a localização dos respectivos SNPs analisados neste estudo.



**Figura 3.2** – Representação esquemática dos genes *PKLR* e *VDR* destacando os SNPs estudados. As caixas em branco representam regiões não traduzidas e as sombreadas as traduzidas. **A** – *PKLR*: os dois SNPs nomeadamente, rs1052176 ou c1705C>A e rs8847 ou c1040G>A localizam-se no exon 12.; **B** – *VDR* (Os primeiros 6 exons deste gene são designados de 1 seguido cada um de uma letra entre a e f). O SNP rs2228570 ou c2T>C ou FokI localiza-se no exon 2 e o rs4760658 ou -84+2252T>C localiza-se no intron 1. Adaptado de Zanella *et al.*, 1997 e Uitterlinden *et al.*, 2004.

A genotipagem dos polimorfismos rs8847 e rs1052176 do gene *PKLR* foi feita através do sistema PCR-RFLP (do inglês “Polymerase Chain Reaction-Restriction Fragment Length Polymorphism”) que consiste na amplificação por PCR (reacção em cadeia da polimerase) da região do gene a analisar, seguida de digestão dos produtos de amplificação e separação electroforética dos produtos digeridos. No caso do SNP rs8847 o oligonucleótido antisense foi desenhado de modo a induzir a substituição de um nucleótido (“mismatch”) numa posição próxima a do alelo polimórfico, resultando na criação de um sítio de restrição para a enzima *NlaIII* na presença do alelo A. Para o caso do SNP rs1052176 a presença do alelo C naturalmente criava o sítio de restrição para a enzima *AclI*.

Os polimorfismos rs4760658 e o rs2228570 do gene *VDR* foram genotipados através de um sistema de discriminação alélica por PCR em tempo real que consiste na utilização de um único par de iniciadores e duas sondas alelo-específicas marcadas com fluoróforos diferentes.

### **3.3.1. Genotipagem por PCR-RFLP**

As reacções de amplificação foram realizadas num termociclador MJ Research modelo PTC-100. O volume final de cada reacção foi de 25µl. Os oligonucleótidos utilizados, as condições de ciclagem e o tamanho dos fragmentos amplificados encontram-se discriminados no quadro 3.1.

#### **3.3.1.1. Amplificação**

##### **3.3.1.1.1. SNP rs1052176**

Para esta amplificação foram utilizados 12,5 µl do HotStarTaq Master Mix (QIAGEN), aos quais foram adicionados os oligonucleótidos (0,3 µM de cada) e aproximadamente 50 ng de DNA genómico. Esta quantidade de master mix forneceu cerca de 1,5 mM de MgCl<sub>2</sub>, 0,2 mM de cada dNTP e 2,5 U de HotStarTaq DNA polymerase com o seu respectivo tampão na concentração de 1X.

### 3.3.1.1.2. SNP rs 8847

Cada 25 µl de reacção continham dNTP a 0,2 mM (Invitrogen, contendo dATP, dCTP, dTTP e dGTP), oligonucleótidos 0,4 µM de cada, 2 mM de MgCl<sub>2</sub>, 50 ng de DNA genómico, 1 U de Platinum<sup>®</sup> Taq DNA polymerase (Invitrogen) e o respectivo tampão na concentração de 1X.

**Quadro 3.1:** Sequência dos iniciadores, condições de ciclagem e tamanho dos fragmentos amplificados em pares de base (pb), utilizados para a amplificação nos sistemas do gene *PKLR*.

.SNP	Sequência dos oligonucleótidos	Ciclagem	Tamanho do fragmento
rs8847	Senso: 5'-AGAGTAGGAGAGGGGACAGGCA	95°C→15min 95°C→30seg 65°C→30seg	201 pb
	Antisenso: 5'-CATGAGAGAGGGAGAAGGACCA-3'	72°C→60seg 72°C→10min	
rs1052176	Senso: 5'-CCTCCCCCAGGAAAGCTC-3'	95°C→15min 95°C→30seg 65°C→30seg	125 pb
	Antisenso: 5'-AGGGGCGTCTCAGGATATG-3'	72°C→40seg 72°C→5min	

Após a amplificação as amostras foram submetidas a electroforese em gel de agarose a 1,7% em tampão TBE 0,5X (Tris-Borato 0,045M, EDTA 1mM) corado com o kit Syber Safe DNA gel stain (Invitrogen) na concentração 1X. Ao gel foram aplicados 10 µl do produto amplificado após ter sido homogeneizado em 3 µl de tampão de carregamento (glicerol 30%, Azul de bromofenol a 0,25% e Xileno cianol a 0,25%). A electroforese ocorreu a 100V durante aproximadamente 30 minutos. O marcador de peso molecular (Invitrogen) utilizado foi o de 123pb. A visualização das bandas foi feita num transiluminador L-Pix, e as fotografias captadas por um sistema de fotodocumentação acoplado ao transiluminador.

### 3.3.1.2. Digestão

A digestão dos fragmentos amplificados por enzimas específicas ocorreu numa reacção com volume final de 20 µl. Os reagentes e as concentrações de cada reacção encontram-se especificados no Quadro 3.2. A incubação ocorreu no banho-maria a 37°C durante 16 horas.

**Quadro 3.2:** Condições da reacção de digestão utilizadas para determinação dos genótipos nos sistemas que utilizam a técnica de PCR-RFLP.

SNP	Reacção da digestão
rs8847	<i>Nla</i> III* (5'-CATG*-3') - 3 U Tampão #4* - 1X Produto amplificado - 15 µl Água suficiente para - 20 µl BSA - 100 µg/ml
rs1052176	<i>Acl</i> I* (5'-C*CGC-3' 3'-GGC*G-5') - 5U Tampão #3* - 1X Produto amplificado - 15 µl Água suficiente para - 20 µl

\* Sítio de clivagem pela enzima de restrição; \* New England Biolabs

Após este período de incubação os produtos da digestão foram submetidos a electroforese em gel agarose 3,5% nas mesmas condições descritas anteriormente.

A electroforese ocorreu a 90 V durante duas horas. O marcador de peso molecular (Invitrogen) utilizado foi o de 50 pb. As bandas foram visualizadas num transiluminador e as fotos captadas como anteriormente descrito.

### 3.3.2. Genotipagem por discriminação alélica

A genotipagem dos SNPs rs2228570 e rs4760658 do gene *VDR* foi realizada através de um sistema TaqMan de discriminação alélica do tipo “assay by design” (Applied Biosystems). Este tipo de discriminação utiliza um par de oligonucleótidos para a amplificação de um fragmento de 60pb flanqueando a posição em que se encontra o alelo polimórfico. A discriminação foi feita através do uso de sondas específicas marcadas com os fluoróforos VIC e FAM. Vide no quadro 3.3 as sequências dos oligonucleótidos e das sondas usadas para cada SNP.

As reacções ocorreram no termociclador modelo ABI Prism 7000 (Applied Biosystems) num volume final de 10 µl contendo 5 ng de DNA genómico, mix de iniciadores e sondas na concentração final de 1X, 5 µl de TaqMan PCR Universal master Mix e água. Todos os reagentes foram utilizados seguindo as recomendações do fabricante (Applied Biosystems).

Inicialmente as reacções foram submetidas a uma pré-leitura (60°C durante 1 minuto) a fim de detectar o “background” da fluorescência. Em seguida as amostras foram submetidas a amplificação (50°C por 2 minutos; 95°C por 10 minutos seguidos de 40 ciclos de 95°C por 15 segundos e 60°C por 60 segundos). Após a amplificação as amostras foram submetidas a uma pós-leitura nas mesmas condições da pré-leitura. A discriminação alélica foi realizada através da interpretação do gráfico resultante da subtração do “background” da fluorescência obtido na pré-leitura a partir do valor final da fluorescência obtido na pós-leitura.

**Quadro 3.3:** Sequência dos iniciadores, sondas e condições de ciclagem utilizadas para a amplificação e discriminação alélica por PCR em tempo real.

SNP	Sequência dos oligonucleótidos e sondas
rs2228570	Senso: 5'-TGGCCTGCTTGCTGTTCTTA-3' Antisenso: 5'-GGGTCAGGCAGGGAAGTG-3'
	VIC: 5'-ATTGCCTCC(G)*TCCCTG-3' FAM: 5'-TTGCCTCC(A)*TCCCTG-3'
rs4760658	Senso: 5'-GGCCCCTAGGACCTGACT-3' Antisenso: 5'-GCATCTGGAACCTCTGCTAGAAA-3'
	VIC: 5'-CTCGCTGT(T)*GCCTGTG-3' FAM: 5'-TCGCTGT(C)*GCCTGTG-3'

\* Entre parenteses encontram-se os alelos polimórficos.

### 3.4. Análise estatística

As frequências genóticas e alélicas de cada polimorfismo foram determinadas manualmente. Os desvios do equilíbrio de Hardy-Weinberg para cada SNP foram verificados a partir da comparação entre as frequências genóticas observadas e as esperadas pela lei de Hardy-Weinberg, utilizando o teste qui-quadrado de Pearson com 1 grau de liberdade ( $\chi^2_{GL=1}$ ). As frequências dos haplótipos foram estimadas por máxima verossimilhança. As comparações entre as frequências das diferentes regiões de Moçambique e entre as diferentes populações foram realizados, utilizando o teste exacto de Fisher. Um p-valor < 0,05 foi considerado significativo. Nos casos de comparações múltiplas foi feita a correcção de Bonferoni multiplicando o p-valor obtido pelo número de comparações realizadas. As análises de desequilíbrio de ligação foram feitas através do cálculo do coeficiente de determinação  $r^2$ , sendo considerados em forte desequilíbrio de ligação os SNPs com  $r^2$  maiores ou igual a 0,8.

As análises foram realizadas com o auxílio do pacote estatístico R na versão 2.10.1 (Gentleman & Ihaka, 1997) e os pacotes “genetics” e “haplostats”.

#### 4. RESULTADOS

Um total de 395 amostras de sangue total foram colhidas em igual número de dadores de sete Bancos de sangue de Moçambique. O número de amostras colhidas em cada Hospital foi calculado com base na densidade populacional da respectiva área. A distribuição das amostras com base no local de colheita encontra-se na Tabela 4.1.

**Tabela 4.1** - Distribuição de indivíduos recrutados para este estudo com base no local de colheita da amostra

<b>Banco de sangue</b>	<b>Número de indivíduos</b>
Hospital Provincial de Pemba	46
Hospital Central de Nampula	92
Hospital Central da Beira	164
Hospital Rural de Chokwé	50
Hospital Geral de Mavalane	15
Hospital Geral José Macamo	15
Hospital Central de Maputo	18

A amostra de dadores era composta por 23,8% de mulheres, e as idades variavam entre 18 e 62 anos. A distribuição por regiões e província de origem encontra-se apresentada na Tabela 4.2. Apesar da colheita não ter sido feita em todas as províncias do país, a amostra inclui indivíduos de todas as províncias.

**Tabela 4.2** – Características dos indivíduos recrutados para este estudo

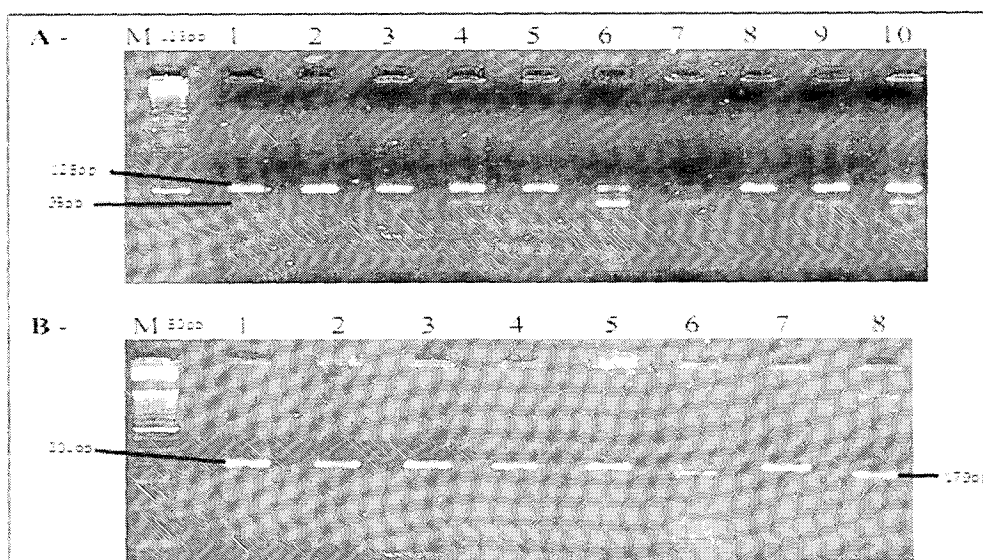
Característica	n (%)
<b>Idade</b>	<b>n = 395</b>
Média (amplitude)	27,97 (18-62)
<b>Sexo</b>	<b>n = 395</b>
Masculino	301 (76,2)
Feminino	94 (23,8)
<b>Província de origem por região</b>	<b>n = 393*</b>
<u>Norte</u>	121 (30,8)
Cabo Delgado	33 (8,4)
Niassa	7 (1,8)
Nampula	81 (20,6)
<u>Centro</u>	150 (38,2)
Zambézia	31 (7,9)
Tete	5 (1,3)
Sofala	108 (27,5)
Manica	6 (1,5)
<u>Sul</u>	122 (31,0)
Inhambane	13 (3,3)
Gaza	48 (12,2)
Maputo Província	51 (13,0)
Maputo Cidade	10 (2,5)

\* dois indivíduos recrutados não nasceram em Moçambique mas foram considerados na análise por terem origem Moçambicana

#### 4.1. Polimorfismos do gene *PKLR*

Das 395 amostras colhidas, 340 foram tipados para o SNP rs8847 e 364 para o SNP rs1052176. Os genótipos foram atribuídos aos indivíduos tendo em conta o padrão das bandas no gel depois da digestão (Figura 4.1). Para o rs1052176 a presença de uma banda (125 pb) determina o genótipo AA, a presença de três bandas (125 pb, 95 pb e 30 pb) determina o genótipo AC e a presença de duas bandas (95 pb e 30 pb) determina o genótipo CC. Para o SNP rs8847 a presença de apenas uma banda de 201 pb determina o genótipo GG, a presença de duas bandas (179 pb e 22 pb) determina o genótipo AA e a presença de 3 bandas (201 pb, 179 pb e 22 pb) determina o genótipo GA. Não foi possível amplificar o DNA de 55 amostras para a genotipagem de rs8847, 31 amostras para genotipagem de rs1052176. Na maior parte dos casos de homozigotos CC de rs1052176 e AA de rs8847 e ambos heterozigotos as bandas de menor tamanho não eram visíveis no gel.





**Figura 4.1** – Imagens dos geis após a digestão. **A** – SNP rs1052176 (digestão parcial): 2 e 5 são AA, 1,3-4,7-9 e 10 são AC, 6 é CC. **B** – SNP rs8847: 1-5 e 7 são GG, 6 é GA e 8 é AA. M representa o marcador de pares de bases.

As frequências dos alelos G e A do rs8847 foram 0,76 e 0,24 respectivamente, e dos alelos A e C do rs1052176 foram 0,4 e 0,6 respectivamente. As frequências genóticas e os números absolutos encontram-se discriminados na Tabela 4.3. A população do estudo foi considerada fora do equilíbrio para ambos os SNPs uma vez que ambas apresentam p-valores superiores a 0,05 (Tabela 4.3).

**Tabela 4.3** – Frequências genóticas dos SNPs rs8847 e rs1052176 na população de Moçambique

SNP	Genótipo	N (%)	EHW* $\chi^2_{GL=1}$ (p-valor)
<i>PKLR</i> rs8847	GG	237 (0,70)	136,57 (0,0001)
	GA	45 (0,13)	
	AA	58 (0,17)	
<i>PKLR</i> rs1052176	CC	120 (0,33)	6,61 (0,0117)
	CA	198 (0,54)	
	AA	46 (0,13)	

\*EHW – Equilíbrio de Hardy-Weinberg.

Para averiguar se há diferenças na distribuição das frequências alélicas e genotípicas por região, estas foram determinadas e comparadas. Conforme observado na Tabela 4.4. Não foram observadas diferenças significativas na distribuição das frequências dos SNPs entre as três regiões de Moçambique. Não foi possível realizar a análise estratificada por grupos etno-linguísticos em virtude do reduzido número amostral em determinados grupos.

**Tabela 4.4** – Comparação entre as frequências alélicas e genotípicas dos SNPs rs8847 e rs1052176 de acordo com as regiões de Moçambique

SNP	Genótipo	Norte n (%)	Centro n (%)	Sul n (%)	$\chi^2_{GL=4}$	p-valor	
<i>PKLR</i> rs8847	GG	70 (0,74)	89 (0,69)	76 (0,66)	4,2846	0,3689	
	GA	14 (0,15)	15 (0,12)	16 (0,14)			
	AA	10 (0,11)	25 (0,19)	23 (0,20)			
		<b>Alelo</b>					
		G	154 (0,82)	193 (0,75)	168 (0,73)	4,9232	0,0853
		A	34 (0,18)	65 (0,25)	62 (0,27)		
<i>PKLR</i> rs1052176	CC	35 (0,32)	49 (0,37)	36 (0,31)	1,6251	0,8043	
	CA	61 (0,55)	67 (0,50)	68 (0,58)			
	AA	14 (0,13)	18 (0,13)	14 (0,12)			
		<b>Alelo</b>					
		C	131 (0,60)	165 (0,62)	140 (0,59)	0,3243	0,8503
	A	89 (0,40)	103 (0,38)	96 (0,41)			

Depois da comparação das frequências dos SNPs entre as regiões do país, e uma vez que existem dados publicados sobre a frequência dos SNPs, foi feita a comparação entre a população deste estudo e populações de diversos grupos étnicos e regiões geográficas. Os resultados encontram-se discriminados nas Tabelas 4.5 e 4.6.

Para o SNP rs8847 as frequências usadas foram as encontradas no dbSNP (<http://www.ncbi.nlm.nih.gov/SNP/>, ) para as populações Han da China em Beijing (CHB), residentes do Utah na Pensilvânia com ancestrais do Norte e Oeste da Europa (CEU) e Afro-Americana. Para a população Brasileira foram usadas frequências determinadas no trabalho de Alvarado Arnez et al (dados não publicados). Depois de feitas as análises foram encontradas diferenças significativas entre a população deste estudo e as populações Afro-Americana ( $p < 0,00001$ ), CHB ( $p < 0,00001$ ), CEU ( $p < 0,003$ ) e Brasileira ( $p < 0,00001$ ), mesmo após a correção de Bonferroni.

**Tabela 4.5** – Frequências genóticas do rs8847 de Moçambicanos, residentes do Utah na Pensilvânia com ancestrais no Norte e Oeste da Europa (CEU), Chineses Han em Beijing (CHB), Afro-Americanos e Brasileiros

Grupo étnico – região	N	Alelo A* (%)	N (%)			p-valor	p <sub>c</sub> **	Referência
			AA	AG	GG			
Afro-Americanos	21	62	7 (33)	12 (57)	2 (10)	$< 10^{-5}$	$< 10^{-5}$	dbSNP
Brasileiros	386	9	27 (7)	13 (3)	34 (90)	$< 10^{-5}$	$< 10^{-5}$	Alvarado Arnez et al <sup>#</sup>
CEU	24	29	2(8)	10 (42)	12 (50)	0,003	0,009	dbSNP
CHB	24	77	14(58)	9 (38)	1 (4)	$< 10^{-5}$	$< 10^{-5}$	dbSNP
Africanos – Moçambique	340	24	58(17)	45 (13)	237(70)	-----		Presente estudo

\*Alelo com menor frequência alélica (mfa) na população de Moçambique. \*\* p<sub>c</sub> – p-valor corrigido. <sup>#</sup> Dados não publicados.

Em relação ao SNP rs1052176 foram usadas frequências publicadas no dbSNP para as populações Afro-Americana, Youruba em Ibadan na Nigéria (YRI) e CHB. Para as populações Caucásiana de Portugal (Manco *et al.*, 2001a) e Africanas de Cabo Verde (Alves *et al.*, 2010) e São Tomé e Príncipe (Manco *et al.*, 2001a) foram usadas frequências publicadas em dois estudos. Para a população Brasileira foram usadas as frequências determinadas no trabalho de Alvarado Arnez et al (dados não publicados). Feitas as análises notou-se que entre a população do estudo e a Brasileira não há diferenças significativas ( $p = 0,058$ ) bem como com a maioria das populações de origem Africana (YRI ( $p = 0,751$ ), Cabo Verdianos ( $p = 0,0734$ ) e os Afro-Americanos ( $p = 0,127$ )), havendo porém diferença significativa em relação à distribuição da população São Tomense ( $p < 0,00001$ ). Observou-se ainda diferenças significativas ao compararmos a amostra de Moçambique com as populações Caucásiana ( $p < 0,00001$ ) e CHB ( $p < 0,00001$ ).

**Tabela 4.6** – Frequências genotípicas do rs1052176 de Moçambicanos, Chineses Han em Beijing (CHB), Afro-Americanos, Yoruba em Ibadan na Nigéria (YRI), Cabo-Verdianos, São Tomenses, Brasileiros e Portugueses

Grupo étnico – região	N	Alelo A* (%)	N(%)			p-valor	p <sub>c</sub> **	Referência
			AA	AC	CC			
Afro-Americanos	62	48	14 (23)	31 (50)	17 (27)	0,127		dbSNP
YRI	60	42	7 (12)	36 (60)	17 (28)	0,751		dbSNP
Africanos – Cabo Verde	200	44	40 (20)	94 (47)	66 (33)	0,074		Alves et al, 2010
Africanos – São Tomé	95	61	36 (38)	44 (46)	15 (16)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	Manco et al, 2001
Brasileiros	333	35	37 (11)	157 (47)	139 (42)	< 10 <sup>-5</sup>		Alvarado Arnez et al #
Caucasianos – Portugal	189	23	8 (4)	70 (37)	111 (59)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	Manco et al, 2001
CHB	44	67	19 (43)	21 (48)	4 (9)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	dbSNP
Africanos – Moçambique	364	40	46 (13)	198(54)	120 (33)	-----		Presente estudo

\*Alelo com menor frequência alélica (mfa) na população de Moçambique. \*\* p<sub>c</sub> – p-valor corrigido. # Dados não publicados.

Após a análise isolada de cada polimorfismo foi feita a determinação das frequências haplotípicas totais e por regiões, e a análise de desequilíbrio ligação entre os dois SNPs. Os haplótipos mais frequentes foram o rs8847G/rs1052176C e o rs8847G/rs1052176A com 38% cada um (Tabela 4.7). Quando analisado por regiões de Moçambique o haplótipo rs8847G/rs1052176C foi o mais frequente no norte (42%) e centro (39%). O mais frequente da região sul foi o rs8847G/rs1052176A (39%). As análises de desequilíbrio de ligação mostram que os dois SNPs não se encontram em desequilíbrio de ligação ( $r^2 = 0,132$ ).

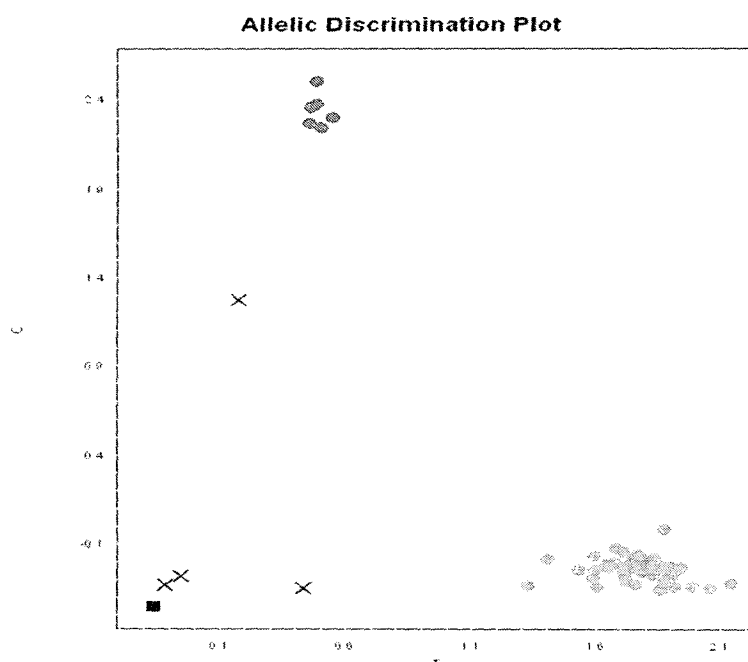
**Tabela 4.7** – Frequências dos haplótipos formados pelos SNPs rs8847 e rs1052176, na população total e nas diferentes regiões de Moçambique

Haplótipo rs8847 – rs1052176	Frequência			
	Total	Norte	Centro	Sul
A – A	0,01906	0,01656	0,02593	0,01428
A – C	0,22077	0,17577	0,22704	0,25066
G – A	0,37977	0,39006	0,35703	0,39358
G – C	0,38040	0,41760	0,39000	0,34147

## 4.2. Polimorfismos do gene *VDR*

Das 395 amostras colhidas 357 tiveram o seu genótipo determinado para o SNP rs2228570 e 375 para o SNP rs4760658. Não foi possível amplificar o DNA de 31 amostras para a genotipagem de rs2228570, e 20 amostras para genotipagem de rs4760658.

Depois da genotipagem dos SNPs por discriminação alélica por PCR em tempo real, os genótipos foram atribuídos com base na intensidade do fluoróforo nos seus respectivos comprimentos de onda. A representação gráfica utilizada para a determinação dos genótipos em experimentos de discriminação alélica por PCR em tempo real pode ser observada na Figura 4.2.



**Figura 4.2** - Representação gráfica do resultado da discriminação alélica por PCR em tempo real do SNP rs 4760658 do gene *VDR*. Os homozigotos AA (TT) e GG (CC) estão representados pelos círculos vermelhos e azul respectivamente e os heterozigotos AG com círculos verdes. Os indeterminados estão representados por X. A letra T (VIC) representa o alelo A e a letra C (FAM) representa o alelo G.

As frequências dos alelos A e G do rs4760658 e C e T do rs2228570 foram 0,88, 0,12, 0,85 e 0,15 respectivamente. Na Tabela 4.8 encontram-se discriminadas as frequências genotípicas e os números absolutos assim como os valores de  $\chi^2_{GL=1}$ . Os testes para desvios do EHW revelaram que a população do estudo encontra-se em equilíbrio para ambos os SNPs do gene *VDR*.

**Tabela 4.8** – Frequências genotípicas dos SNP rs4760658 e rs2228570 na população de Moçambique

SNP	Genótipo	N (%)	EHW $\chi^2_{GL=1}$ (p-valor)
<i>VDR</i> rs4760658	AA	294 (0,79)	1,80 (0,1377)
	AG	73 (0,19)	
	GG	8 (0,02)	
<i>VDR</i> rs2228570	CC	260 (0,73)	0,013 (1)
	CT	89 (0,25)	
	TT	8 (0,02)	

As frequências alélicas e genotípicas por região encontram-se discriminadas na Tabela 4.9. Comparando estas frequências, não se observou diferenças significativas na distribuição dos SNPs por regiões. Do mesmo modo que para os SNPs de *PKLR*, as estratificações por etnia e por província de origem geraram um número amostral muito baixo e a análise não mostrou diferenças entre os grupos (Dados não mostrados).

**Tabela 4.9** – Frequências alélicas e genótípicas dos SNP rs4760658 e rs2228570 por regiões de Moçambique

SNP	Genótipo	Norte n (%)	Centro n (%)	Sul n (%)	$\chi^2_{GL=4}$	p-valor	
VDR rs4760658	AA	83 (0,76)	112 (0,76)	98 (0,82)	1,2242	0,8741	
	AG	23 (0,21)	29 (0,20)	20 (0,17)			
	GG	3 (0,03)	3 (0,02)	2 (0,02)			
	<b>Alelo</b>						
	A	189 (0,87)	253 (0,88)	216 (0,90)	1,2549	0,5339	
G	29 (0,13)	35 (0,12)	24 (0,10)				
VDR rs2228570	CC	75 (0,72)	101 (0,75)	83 (0,71)	1,5222	0,8227	
	CT	27 (0,26)	31 (0,23)	30 (0,26)			
	TT	2 (0,02)	2 (0,01)	4 (0,03)			
	<b>Alelo</b>						
	C	177 (0,85)	233 (0,87)	196 (0,84)	1,0256	0,5988	
T	31 (0,15)	35 (0,13)	38 (0,16)				

A semelhança do que foi feito para os SNPs do *PKLR*, as frequências encontradas neste estudo foram comparadas com as de outras populações de diferentes áreas geográficas e etnias. No caso do rs4760658 as comparações foram feitas com dados publicados num estudo com a população Brasileira (Marques, 2010) e com as populações Afro-Americana e CEU cujos dados foram obtidos no dbSNP (Tabela 4.10). Para o SNP rs2228570, as comparações foram feitas com dados publicados de diversos estudos com populações Venda da África do Sul (Lombard *et al.*, 2006), Brasileira (Marques, 2010), Chinesa (Pan *et al.*, 2009), Caucasiana de Portugal (Lemos *et al.*, 2008), e com dados das populações Afro-Americana e YRI publicados no dbSNP.

Para o SNP rs4760658 não houve diferenças significativas entre a população deste estudo e a população Afro-Americana, havendo porém entre aquela e as populações Portuguesa e Brasileira.

**Tabela 4.10** – Frequências genotípicas do rs4760658 de Moçambicanos, Afro-Americanos e residentes do Utah na Pensilvânia com ancestrais no Norte e Oeste da Europa (CEU)

Grupo étnico – região	N	Alelo G* (%)	N(%)			p-valor	p <sub>c</sub> **	Referência
			AA	AG	GG			
CEU	34	41	9 (26)	22 (65)	3 (9)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	dbSNP
Afro-Americanos	61	18	40 (65)	20 (33)	1 (2)	0,065		dbSNP
Brasileiros	599	31	287(48)	257(43)	53(9)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	Marques, 2010
Africanos – Moçambique	375	12	294(78)	73 (20)	8 (2)			Presente estudo

\*Alelo com menor frequência alélica (mfa) na população de Moçambique. \*\* p<sub>c</sub> – p-valor corrigido.

Na Tabela 4.11 pode-se observar que para o SNP rs2228570 houve diferenças significativas com as populações Portuguesa, Brasileira e Chinesa. De notar que quando a população de Moçambique foi comparada a outras populações de origem Africana não houve diferença significativa na distribuição das frequências.

**Tabela 4.11** – Frequências genotípicas do rs2228570 de Moçambicanos, Portugueses, Brasileiros, Chineses, Venda da África do Sul, Yoruba em Ibadan na Nigéria (YRI) e Afro-Americanos

Grupo étnico – região	N	Alelo T* (%)	N(%)			p-valor	p <sub>c</sub> **	Referência
			CC	CT	TT			
Caucasianos - Portugal	249	38	97(39)	114(46)	38(15)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	Lemos et al, 2008
Brasileiros	596	32	277(46)	263(45)	56(9)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	Marques, 2010
Chineses	212	45	68(32)	97(46)	47(22)	< 10 <sup>-5</sup>	< 10 <sup>-5</sup>	Pan et al, 2009
Afro-Americanos	62	19	41(66)	18(29)	3 (5)	0,361		dbSNP
YRI	60	17	41(68)	18(30)	1 (2)	0,784		dbSNP
Venda – África do Sul	86	13	66(77)	18(21)	2 (2)	0,725		Lombard et al, 2006
Africanos – Moçambique	357	15	260(73)	89(25)	8 (2)	-----		Presente estudo

\*Alelo com menor frequência alélica (mfa) na população de Moçambique. \*\* p<sub>c</sub> – p-valor corrigido.

O haplótipo mais frequente total (76%) e por regiões (75% no Norte, 77% no Centro e 75% no Sul) foi o rs4760658A/rs2228570T (Tabela 4.12). As análises de desequilíbrio de ligação mostraram que os SNP rs4760658 e rs2228570 não estão associados entre si ( $r^2 = 0,002$ ).



**Tabela 4.12** – Frequências dos haplótipos formados pelos SNPs rs4760658 e rs2228570, totais e por regiões

<b>Haplótipo</b> rs4760658 – rs2228570	<b>Frequência</b>			
	<b>Total</b>	<b>Norte</b>	<b>Centro</b>	<b>Sul</b>
A – C	0,75709	0,74937	0,77140	0,752263
A – T	0,12426	0,11786	0,10691	0,1477
G – C	0,09579	0,10103	0,09818	0,08537
G – T	0,02286	0,03175	0,02351	0,01465

## 5. DISCUSSÃO

No presente estudo foram caracterizados 4 SNPs, dois do gene *PKLR* (rs1052176 e rs8847) usando a técnica de PCR-RFLP e dois do gene *VDR* (rs2228570 e rs4760658) usando a técnica de discriminação alélica por PCR em Tempo real. É sabido que a frequência dos SNPs varia nos diferentes grupos étnicos e regiões geográficas.

### 5.1. Polimorfismos do *PKLR*

As frequências do SNP rs1052176 encontram-se desviadas do equilíbrio de Hardy-Weinberg. Um facto curioso surge quando se testa o equilíbrio por regiões. A população da região sul é a única que apresenta desvio (dados não mostrados). Isto pode ser reflexo de algum tipo de estratificação críptica possivelmente presente na população estudada. A separação dos indivíduos por região foi feita tendo em conta a localização da província de origem com base na divisão político-administrativa do país e é sabido que a separação por fronteiras físicas geográficas não reflete a separação dos diferentes grupos populacionais tendo como base as diferentes origens étnicas/etno-linguísticas. Uma estratificação tendo como base cada um dos grupos etno-linguísticos de Moçambique, ou a província de origem do indivíduo resultou num número amostral muito pequeno ou nulo em alguns grupos fazendo com que os resultados das análises comparativas não fossem confiáveis. Salientar que para este SNP a população deste estudo apresenta maior percentagem de indivíduos heterozigotos quando comparado a ambos homozigotos. Este desvio do Equilíbrio de Hardy-Weinberg pode refletir algum artefacto da técnica de genotipagem ou a presença de alguma força selectiva sobre um dos alelos de forma balanceada (Miller, 1994). A chamada Lei de Hardy-Weinberg enuncia que para qualquer locus génico as frequências relativas dos genótipos permanecem constantes de geração em geração desde que a população em causa atenda às seguintes premissas: estar livre da acção de factores evolutivos (selecção, mutação, migração e deriva genética); ser suficientemente grande para que os erros de amostragem e o efeito da deriva genética se mostrem insignificantes, e ser panmítica, isto é, os cruzamentos ocorram ao acaso (Brooker, 2009). Contudo sabe-se que a população humana não obedece a todas estas premissas. Assim ao determinarmos se a população do estudo está em Equilíbrio de Hardy-Weinberg, utilizamos esta população hipotética como parâmetro de

comparação. O desvio do equilíbrio de Hardy-Weinberg pode ser devido a cruzamentos consanguíneos, seleção natural e a estratificação populacional, mas também podem ser reflexo de uma associação com doença (Balding, 2006). Comparando a frequência de heterozigotos nas populações estudadas nota-se com excepção da população caucasiana, que a maior parte apresenta frequências elevadas de heterozigotos em relação a ambos homozigotos, sendo superior ou igual a 50% na população do presente estudo (54%), nos YRI (60%) e nos afro-Americanos (50%). Comparando as frequências genóticas no geral, nota-se um fenómeno curioso. As frequências não diferem significativamente quando comparadas com as da maioria de outras populações de origem Africana nomeadamente os Afro-Americanos, YRI, e Cabo Verdianos, sendo porém diferentes das da população São Tomense. São Tomé e Príncipe é um arquipélago situado no Oeste de África mais concretamente no Golfo da Guiné. Até os finais do século XV não possuía habitantes permanentes, altura em que os comerciantes Europeus se estabeleceram. Uma grande contribuição na demografia deste arquipélago foi dada pelos povos de várias regiões da África sub-Sahariana uma vez que este arquipélago foi um importante ponto de comércio e tráfico de escravos (Caldeira, 2008; Mateu *et al.*, 1997; Pereira *et al.*, 1999; Trovoada *et al.*, 2007). Após a abolição da escravatura o movimento de força laboral trazida de Angola, Cabo Verde e Moçambique no início do século XIX também contribuiu para a demografia de São Tomé e Príncipe (Trovoada *et al.*, 2007). Isto propiciou a miscigenação entre indivíduos oriundos de várias partes do continente Africano, bem como a miscigenação com populações Europeias. Estudos das variações de mtDNA e DNA do cromossoma Y em São Tomé e Príncipe demonstraram a presença de uma sub-estratificação desta população (Trovoada *et al.*, 2001; Trovoada *et al.*, 2004). Estes factores em conjunto podem ter influenciado na distribuição dos alelos assim como das frequências genóticas podendo justificar essas diferenças. Cabo Verde também é um arquipélago, foi um ponto de comércio e tráfico de escravos (Silva, 2009) e situa-se no Oeste de África há cerca de 300 Km da costa do Senegal e da Guiné Bissau. Apesar disso as frequências genóticas não apresentam diferenças significativas. O padrão de miscigenação do povo Cabo Verdiano é diferente do de São Tomé. A população de Cabo Verde é resultante da miscigenação entre populações de várias regiões de África Ocidental trazidas a estas ilhas durante o período de tráfico de escravos com Caucasianos Europeus sendo os haplogrupos característicos destas populações mais frequentes em Cabo Verde do que em São Tomé e Príncipe (Gonçalves *et al.*, 2003; Trovoada *et al.*, 2007). O facto da população Africana possuir uma grande diversidade

genética intra e interpopulacional (Mateu *et al.*, 1997) mesmo entre indivíduos muito próximos (Schuster *et al.*, 2010) pode explicar estas diferenças observadas no padrão de distribuição genotípica de alelos entre estas populações com origem Africana. Outro facto curioso é em relação aos resultados da comparação com a população Brasileira. As análises mostram que não há diferença significativa entre estas duas populações apesar da miscigenação que caracteriza a população Brasileira conforme discutido mais adiante. Em relação as populações Caucasiana de Portugal e CHB, as frequências mostram claramente diferenças significativas.

Foi recentemente publicado um estudo de Machado *et al.* (2010) em que foi determinada a frequência alélica do rs1052176 em 150 indivíduos Moçambicanos com malária. A frequência alélica encontrada foi de 0,399 para o alelo C e 0,601 para o alelo A. Estas frequências são opostas as frequências encontradas neste estudo (A 0,40 e C 0,60). Olhando para os grupos populacionais em ambos os estudos nota-se que no estudo de Machado *et al.* a população é composta de indivíduos com malária e no presente estudo a população é composta de indivíduos aparentemente sem malária. Foi sugerido após ter sido relatado o efeito protector em modelos murínicos (Min-Oo *et al.*, 2003) e a replicação diminuída de parasitas quando cultivados *in vitro* em sangue de pessoas com deficiências de PK (Ayi *et al.*, 2008; Durand & Coetzer, 2008a), que mutações no gene *PKLR* que levam a deficiência de PK podem conferir protecção contra a malária especialmente nos indivíduos heterozigotos. É sabido que alguns defeitos tais como substituição, inserção ou deleção de nucleótidos em genes que codificam proteínas eritrocitárias (hemoglobina S, talassemia, Glucose-6-fosfato desidrogenase) tornaram-se comuns devido a força selectiva exercida pela malária uma vez que indivíduos portadores desses defeitos são resistentes a infecção (indivíduos homozigotos) pelo parasita que causa a malária ou ao desenvolvimento de formas graves da malária (heterozigotos) (Aidoo *et al.*, 2002; Ayi *et al.*, 2004; Guindo *et al.*, 2007; Mockenhaupt *et al.*, 2004; Ruwende *et al.*, 1995). Fazendo uma comparação hipotética entre as frequências alélicas do presente estudo e as frequências de pacientes Moçambicanos com malária (Machado *et al.*, 2010) pode-se sugerir que existe um efeito protector do alelo C já que a sua frequências está aumentada no grupo de indivíduos sadios. O SNP rs1052176 é resultante de uma mutação sinónima. A possível acção da força selectiva poderia ser na verdade exercida num polimorfismo verdadeiramente funcional em forte desequilíbrio de ligação com este. Ainda no estudo de Machado *et al.* (2010) foram determinadas as frequências alélicas na população

portuguesa em indivíduos saudáveis e em indivíduos com deficiência de PK. Em ambos os grupos pertencentes a uma população residente em área livre da malária o alelo predominante é o C, sendo maior nos indivíduos deficientes. Mas conforme os resultados de Machado et al (2010) e de Alves et al (2010), nenhuma associação foi encontrada através do uso de SNPs ou STRs isoladamente. Referir que a amostragem para ambos estudos foi insuficiente o que pode ter influenciado nos resultados. Porém no mesmo estudo a análise da distância genética ( $F_{ST}$ ) entre Africanos e Portugueses usando STRs e SNPs do gene *PKLR* e regiões adjacentes revelou que os valores de  $F_{ST}$  eram altos quando comparados com os valores obtidos usando STRs e SNPs de genes neutros. Este achado sugere que esta região genômica está sob efeito de um processo selectivo. As análises de desequilíbrio de ligação de STRs na região genômica estudada mostraram a presença de um forte desequilíbrio de ligação entre pares de marcadores numa região maior no grupo de indivíduos com malária não complicada em comparação com o grupo de indivíduos com malária complicada. Foi encontrada uma associação borderline entre um haplótipo de STR nessa região e a malária não complicada. Estes dois achados em conjunto sugerem que um bloco desta região genômica mais conservada possa estar associada a proteção contra formas mais graves de malária. Portanto a realização de um estudo de associação com tamanho amostral e um grupo controle adequados poderia dissipar estas dúvidas.

Em relação ao SNP rs8847 é de referir que também se encontra bastante desviada do equilíbrio de Hardy-Weinberg ( $p < 0,00001$ ). Nota-se uma deficiência de indivíduos heterozigotos em comparação aos indivíduos homozigotos para ambos alelos. Mesmo quando se testa os indivíduos agrupados por regiões de Moçambique o desvio se mantém pronunciado especialmente na região centro (dados não mostrados). Fenómenos que alteram as frequências alélicas e genotípicas provocando desvios do Equilíbrio de Hardy-Weinberg poderiam explicar esta distribuição. Este SNP localiza-se na região 3'UTR do gene *PKLR* podendo regular a taxa de transcrição deste gene. Sendo um SNP funcional ou podendo estar em desequilíbrio de ligação forte a um SNP funcional, pode-se pensar na possibilidade de estar a sofrer a acção de uma força selectiva como discutido para o rs1052176. As frequências genotípicas do SNP rs8847 na população Brasileira encontram-se desviadas do equilíbrio de Hardy-Weinberg (Arnez et al, dados não publicados). A genotipagem das amostras desta população foi feita usando a mesma técnica e no mesmo laboratório em que foram genotipadas as amostras da população de

Moçambique. Isto pode levantar a possibilidade de estes achados refletirem algum defeito na mutagenese durante a amplificação do fragmento contendo o SNP uma vez que para a genotipagem deste SNP foi desenhado um par de iniciadores que induzisse uma mutagenese na base que flanqueia o alelo polimórfico de modo que na presença do alelo polimórfico A, fosse criado o sítio de clivagem pela enzima de restrição *NlaIII*. Mas curiosamente a frequência de indivíduos homocigotos para o alelo A em ambas populações é superior a dos indivíduos heterocigotos o que pode de alguma forma refutar esta hipótese. Outra razão possível razão que pode justificar estas frequências é o efeito Wahlund em que se verifica uma diminuição na frequência de heterocigotos devido a uma subestratificação populacional (Beiguelman, 2008). Esta subestratificação pode ser resultante do aparecimento de grandes isolados populacionais causado por barreiras que impedem a troca de genes com outros grupos populacionais por meio de casamento entre os seus elementos. Estas barreiras podem ser geográficas, políticas, socio-económicas, religiosas ou culturais. Uma repetição da genotipagem usando outras técnicas e usando uma amostra populacional estratificada de diversas formas e com tamanho adequado por cada grupo poderiam clarificar estas zonas de penumbra.

O genoma humano contém estruturas semelhantes a blocos com pouca ou nenhuma recombinação dentro dos blocos de tal modo que os SNPs não se encontram distribuídos ao acaso (Rezende *et al.*, 2007). Os SNPs localizados em tais blocos haplotípicos encontram-se portanto em desequilíbrio de ligação. SNPs localizados nestes blocos são considerados mais informativos visto que a tipagem de apenas um SNP é suficiente para capturar a informação de todos os demais marcadores presentes no mesmo bloco.

O  $D'$  e o  $r^2$  são as medidas do desequilíbrio de ligação mais comumente utilizadas em estudos de desequilíbrio de ligação. Ambos são dependentes das frequências alélicas sendo menos sensíveis a estes os valores de  $D'$  (Lewontin, 1988). Os valores de  $D'$  e  $r^2$  variam de 0 a 1. Quando  $D = 0$  não há desequilíbrio de ligação (é possível observar os quatro haplótipos possíveis) e quando  $D = 1$  o desequilíbrio é completo (observa-se apenas um ou dois dos possíveis haplótipos). Quando  $r^2 = 0$  não há desequilíbrio de ligação porém quando  $r^2 = 1$  o desequilíbrio é perfeito (os alelos são sempre transmitidos num mesmo bloco). Os valores de  $r^2$

tendem a ser próximos do zero quando um ou ambos polimorfismos analisados são raros ( $mfa < 0,10$ ) (Tiret *et al.*, 2002).

As análises de desequilíbrio de ligação entre os SNPs rs8847 e rs1052176 do gene *PKLR* mostram que estes não estão associados na população de Moçambique ( $r^2 = 0,132$ ).

## 5.2. Polimorfismos do *VDR*

Este é o primeiro estudo em que se determina a frequência de SNPs do gene *VDR* em Moçambique. As frequências genóticas dos SNPs deste gene tipados na população deste estudo estavam em Equilíbrio de Hardy-Weinberg ( $p > 0,05$ ). Não houve diferença na distribuição dos genótipos por regiões de Moçambique apesar da possível estratificação acima discutida. Comparando estas frequências com a de outras populações de origem Africana, notou-se que para o rs2228570 não há diferenças com as populações Venda da África do Sul, YRI e Afro-Americana, e para o rs4760658 não há diferença com os Afro-Americanos, a única população com origem Africana de que se dispõem dados publicados para este SNP. No estudo de Marques *et al* (2010), as frequências genóticas do SNP rs2228570 foram determinadas em três grupos étnicos da população Brasileira nomeadamente negros, mestiços e brancos. A comparação destas frequências com as da população do presente estudo revelou diferenças significativas comparando as frequências dos três grupos em conjunto (Tabela 4.11) ou separadamente (dados não mostrados). É curioso o facto da população de indivíduos negros Brasileiros apresentar diferença significativa com a população deste estudo. Sabe-se que a população Brasileira é bastante miscigenada em resultado do cruzamento entre Europeus, Africanos e Ameríndios (Parra *et al.*, 2003). A maior parte da contribuição africana no pool do mtDNA é da região sub-Sahariana (Alves-Silva *et al.*, 2000). A identificação dos indivíduos segundo o grupo étnico é muitas vezes feita pelo próprio indivíduo oralmente ou tem em consideração a cor da pele. Esta forma de identificação não reflete a linhagem genética do indivíduo conforme demonstrado no estudo de Alves-Silva *et al* (2000). Provavelmente houve erros na classificação étnica dos indivíduos Brasileiros, o que pode explicar a diferença observada. Em relação as populações Chinesa e Portuguesa houve claramente diferenças significativas nas frequências do rs2228570.

As frequências do SNP rs4760658 foram significativamente diferentes quando comparadas com as da população CEU.

É de digno de nota o facto de que apesar do alelo C do rs2228570 ser o predominante em todas as populações estudadas, ele é mais frequente nos caucasianos que nos africanos (Uitterlinden *et al.*, 2004). As frequências encontradas na população deste estudo foram semelhantes às das populações africanas avaliadas.

O SNP rs2228570 não está em desequilíbrio de ligação com nenhum outro polimorfismo do gene *VDR* (Nejentsev *et al.*, 2004). Os valores de  $D'$  e de  $r^2$  encontrados para a população de Moçambique indicam-nos que os SNPs rs4760658 e rs2228570 não se encontram em desequilíbrio de ligação.

O gene *VDR* é um gene candidato para o estudo de associação com a tuberculose e com a hanseníase. A vitamina D activada, ligando do VDR, é produzida por macrófagos. Ela activa monócitos, suprime a imunidade humoral regulando a produção de citocinas e estimula a imunidade celular (Adams *et al.*, 2007). É um factor importante para a resposta imune inata contra bactérias intracelulares. Essas observações levaram à hipótese de que polimorfismos no gene *VDR* pudessem estar envolvidas no controle da resposta imune do hospedeiro contra micobactérias.

Os polimorfismos rs2228570 (*FokI*), rs1544410 (*BsmI*) e rs731236 (*TaqI*) foram usados em estudos de associação com tuberculose (Lombard *et al.*, 2006; Motsinger-Reif *et al.*, 2010; Valdivielso & Fernandez, 2006; Wilkinson *et al.*, 2000; Zhang *et al.*, 2010). Os resultados destes estudos mostram inconsistências nas associações encontradas pois diferentes alelos de risco foram associados a tuberculose em diferentes populações e alguns estudos relatam associação com alelos individualmente ou formando haplótipos. Foi recentemente publicado um artigo de revisão e de meta-análise sobre tuberculose e polimorfismos do gene *VDR* (Gao *et al.*, 2010). A meta-análise sugeriu que o genótipo TT do rs2228570 pode constituir um factor de risco para a tuberculose quando comparado com o genótipo CC. Porém ao se estratificar a população analisada esta associação foi encontrada apenas na população asiática e não nas populações



Africana e Sul Americana. Em relação ao polimorfismo rs1544410, após a estratificação a meta-análise sugeriu uma redução no risco de desenvolver tuberculose na presença do genótipo GG na população Asiática, e na população Africana nenhuma associação foi encontrada. Estes achados inconsistentes nos resultados dos estudos de associação entre os polimorfismos do gene VDR e a tuberculose podem ser devido a vários factores nomeadamente: Estratificação populacional; tamanho amostral pequeno; diferentes grupos étnicos; Uso de polimorfismos não funcionais que podem ter alguma ligação a outros polimorfismos verdadeiramente funcionais no mesmo gene ou em genes vizinhos e o facto de haver diferenças nos padrões de desequilíbrio de ligação entre diferentes populações; a seleção inapropriada de controlos; interação gene-gene e gene-ambiente (Lewis *et al.*, 2005).

Em 2006 Lombard *et al.* (Lombard *et al.*, 2006) relataram a associação da tuberculose com haplótipos do gene *VDR*. Este achado mais uma vez suporta a ideia de que a influência dos polimorfismos na actividade do VDR é colectiva e não individual (Uitterlinden *et al.*, 2004).

Contribuições valiosas à compreensão de fenómenos que participam nos processos de interação entre os agentes patogénicos e o hospedeiro humano podem ser conseguidas através da identificação de polimorfismos em genes candidatos, da determinação das frequências alélicas e genótípicas dos mesmos, de estudos de desequilíbrio de ligação entre estes polimorfismos bem como a realização de estudos de associação com doenças, especialmente as infecciosas. Esta identificação poderá ainda contribuir para o desenvolvimento de melhores estratégias preventivas e terapéuticas que visam o combate e mitigação das doenças infecciosas e dos seus efeitos.

## 6. CONCLUSÕES

Estudos de marcadores genéticos associados a doenças, especialmente as infecciosas, ainda são escassos em Moçambique. Neste trabalho investigamos as frequências genotípicas, alélicas e haplotípicas nos genes *PKLR* e *VDR*. Os resultados encontrados permitem-nos concluir que:

1. Não há diferenças na distribuição das frequências genotípicas dos SNPs estudados nas diferentes regiões de Moçambique.
2. As frequências genotípicas nos SNPs de *PKLR* estão desviadas do equilíbrio de Hardy-Weinberg.
3. As frequências dos SNPs em ambos genes estudados diferem das frequências populações não Africanas avaliadas.

## 7. RECOMENDAÇÕES

A compreensão do significado dos desvios do equilíbrio de Hardy-Weinberg observados nas frequências dos polimorfismos do gene *PKLR* demanda a realização de mais estudos do género.

São necessários estudos de associação entre polimorfismos nos genes *PKLR* e *VDR* com doenças infecciosas como a malária e tuberculose respectivamente em populações de áreas onde estas doenças são endémicas ou apresentam altas taxas de morbilidade e mortalidade.

## 8. REFERÊNCIAS BIBLIOGRÁFICAS

Abbas AK, Litchman AH and Pillai S. *Imunologia Celular e Molecular*. 6 ed. Rio de Janeiro: Elsevier; 2008. 4-6

Adams JS, Liu PT, Chun R, Modlin RL and Hewison M. Vitamin D in defense of the human immune response. *Ann NY Acad Sci*. 2007; 1117: 94-105.

Aidoo M, Terlouw DJ, Kolczak MS, McElroy PD, ter Kuile FO, Kariuki S, Nahlen BL, Lal AA and Udhayakumar V. Protective effects of the sickle cell gene against malaria morbidity and mortality. *Lancet*. 2002; 359: 1311-1312.

Akira S, Uematsu S and Takeuchi O. Pathogen recognition and innate immunity. *Cell Microbiol*. 2006; 124: 783-801.

Alves-Silva J, Santos MS, Guimarães PEM, Ferreira ACS, Bandelt H-J, Pena SDJ and Prado VF. The ancestry of Brazilian mtDNA lineages. *Am J Hum Genet*. 2000; 67: 444-461.

Alves J, Machado P, Silva J, Gonçalves N, Ribeiro L, Faustino P, do Rosário VE, Manco L, Gusmão L, Amorim A and Arez A P. Analysis of malaria associated genetic traits in Cabo Verde, a melting pot of European and sub-Saharan settlers. *Blood Cells Mol Dis*. 2010; 44: 62-68.

Arai H, Miyamoto K-I, Taketani Y, Yamamoto H, Iemori Y, Morita K, Tonai T, Nishisho T, Mori S and Takeda E. A vitamin D receptor gene polymorphism in the translation initiation codon: Effect on protein activity and relation to bone mineral density in Japanese women. *J Bone Miner Res*. 1997; 12: 915-921.

Assane AAA, Fabricio-Silva GM, Cardoso-Oliveira J, Mabunda NEJ, Sousa AM, Jani IV, Ferreira OC and Porto LCMS. HLA-A, -B and DRB1 allele and haplotype frequencies in the Mozambican population: a blood donor based population study. *Hum Immunol*. 2010;

Ayi K, Min-Oo G, Serghides L, Crockett M, Kirby-Allen M, Quirt I, Gros P and Kain KC. Pyruvate Kinase Deficiency and Malaria. *N Engl J Med*. 2008; 358: 1805-1810.

Ayi K, Turrini F, Piga A and Arese P. Enhanced phagocytosis of ring-parasitized mutant erythrocytes. A common mechanism that may explain protection against falciparum-malaria in sickle-trait and beta-thalassemia-trait. *Blood*. 2004; 104: 3364-3371.

Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet.* 2006; 7: 781-791.

Banoei MM, Mirsaedi MS, Houshmanda M, Tabarsi P, Ebrahimi G, Zargari L, Kashani BH, Masjedi MR, Mansouri SD and Ramirez J. Vitamin D receptor homozygote mutant tt and bb are associated with susceptibility to pulmonary tuberculosis in the Iranian population. *Int J Inf Dis.* 2010; 14: e84-e85.

Beiguelman B. *Genética de Populações Humanas.* ed. Ribeirão Preto, SP: Sociedade Brasileira de Genética; 2008. 134-138

Boiteux A and Hess B. Design of glycolysis. *Philos Trans R Soc Lond B Biol Sci.* 1981; 293: 5-22.

Brooker RJ. *Genetics: Analysis & Principles.* 3rd ed. New York: McGraw Hill; 2009. 665-670

Caldeira AM. Tráfico de escravos e conflitualidade: o arquipélago de São Tomé e Príncipe e o reino do Congo durante o século XVI. *Ciências e letras.* 44: Julho/Dezembro. 2008: 55-76. Disponível em <http://www.l.fapa.com.br/cienciasyletras/pdf/revista44/artigo3.pdf>.

Campbell MC and Tishkoff SA. African genetic diversity: Implications for human demographic history, modern human origins, and complex disease mapping. *Annu Rev Genomics Hum Genet.* 2008; 9: 403-433.

Chabert J. Breve história de África. 2003. [Online]. [Data de acesso: 28 de Junho de 2010]; Disponível em <http://www.carlosduarte.ecn.br/brevehistoriaafrica.htm>.

Collins FS, Guyer MS and Chakravarti A. Variations on a theme: Cataloging human DNA sequence variation. *Science.* 1997; 278: 1580 - 1581.

Colston K and Feldman D. Nuclear translocation of the 1,25-dihydroxycholecalciferol receptor in the mouse kidney. *J Biol Chem.* 1980; 255: 7510-7513.

Cooke GS and Hill AVS. Genetics of susceptibility to human infectious disease. *Nat Rev Genet.* 2001; 2: 967-977.

Crofts LA, Hancock MS, Morrison NA and Eisman JA. Multiple promoters direct the tissue-specific expression of novel N-terminal variant human vitamin D receptor gene transcripts. Proc Natl Acad Sci USA. 1998; 95: 10529-10534.

Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. (dbSNP Build ID: 131). Disponível em: <http://www.ncbi.nlm.nih.gov/SNP/>.

Durand PM and Coetzer TL. Pyruvate kinase deficiency in a South African kindred caused by a 1529A mutation in the *PK-LR* gene. S Afr Med J. 2008b; 98: 455-457.

Durand PM and Coetzer TL. Pyruvate kinase deficiency protects against malaria in humans. Haematologica. 2008a; 93: 939-940.

Fermo E, Bianchi P, Chiarelli LR, Chiarelli LR, Cotton F, Vercellati C, Writzl K, Baker K, Hann I, Rodwell R, Valentini G and Zanella A. Red cell pyruvate kinase deficiency: 17 new mutations of the PK-LR gene. Br J Haematol. 2005; 129: 839-846.

Fothergill-Gilmore LA and Michels PA. Evolution of glycolysis. Prog Biophys Mol Biol. 1993; 59: 105-235.

Gao L, Tao Y, Zhang L and Jin Q. Vitamin D receptor genetic polymorphisms and tuberculosis: updated systematic review and meta-analysis. Int J Tuberc Lung Dis. 2010; 14: 15-23.

Gentleman R and Ihaka R. R: A Programming Environment for Data Analysis and Graphics. The R Foundation for Statistical Computing: 1997. 2.10.1 (2009-12-14).ISBN 3-900051-07-0 [Data de acesso: 02.02.2010]. Disponível em <http://www.r-project.org/>.

Goldfeld AE, Delgado JC, Thim S, Bozon MV, Ugliarolo AM, Turbay D, Cohen C and Yunis EJ. Association of an HLA-DQ allele with clinical tuberculosis. JAMA. 1998; 279: 226-228.

Gonçalves R, Rosa A, Freitas A, Fernandes A, Kivisild T, Villems R and Brehm A. Y-chromosome lineages in Cabo Verde Islands witness the diverse geographic origin of its first male settlers. Hum Genet. 2003; 113: 467-472.

Goulart LR, Ferreira FR and Goulart IMB. Interaction of TaqI polymorphism at exon 9 of the vitamin D receptor gene with the negative lepromin response may favor the occurrence of leprosy. FEMS Immunol Med Microbiol. 2006; 48: 91-98.

Guindo A, Fairhurst RM, Doumbo O, Wellem TE and Diallo DA. X-linked G6PD deficiency protects hemizygous males but not heterozygous females against severe malaria. 2007. PLoS Med [Online]. 4: e66. [Data de acesso 14.06.2010]. Disponível em [www.plosmedicine.org](http://www.plosmedicine.org).

Hall ER and Cottam GL. Isozymes of pyruvate kinase in vertebrates: their physical, chemical, kinetic and immunological properties. *Int J Biochem.* 1978; 9: 785-793.

Haussler MR and Norman AW. Chromosomal receptor for a vitamin D metabolite. *Proc Natl Acad Sci USA.* 1969; 62: 155-162.

Hill AV, Allsopp CE, Kwiatkowski D, Anstey NM, Twumasi P, Rowe PA, Bennett S, Brewster D, McMichael AJ and Greenwood BM. Common west african HLA antigens are associated with protection from severe malaria. *Nature.* 1991; 352: 595.

Instituto Nacional de Estatística. Estatísticas de Moçambique. 2009. INE [Online]. [Data de acesso: 22.05.2010]; Disponível em <http://www.ine.gov.mz/>.

James SP, Nicol WD and Shute PG. A Study of induced malignant tertian malaria. *Proc R Soc Med.* 1932; 25: 1153-1186.

Janeway Jr. CA. Approaching the asymptote: Evolution and revolution in immunology. *Cold Spring Harb Symp Quant Biol.* 1989; 54: 1-13.

Kanno H, Fujii H, Hiron A, Omine M and Miwa S. Structural analysis of human pyruvate kinase L-gene and identification of the promoter activity in erythroid cells. *Biochem Biophys Res Commun.* 1992a; 188: 516-523.

Kanno H, Fujii H, Hirono A and Miwa S. cDNA cloning of human R-type pyruvate kinase and identification of a single amino acid substitution (Thr384--Met) affecting enzymatic stability in a pyruvate kinase variant (PK Tokyo) associated with hereditary hemolytic anemia. *Proc Natl Acad Sci USA.* 1991; 88: 8218-8221.

Kanno H, Fujii H, Hirono A, Omine M and Miwa S. Identical point mutations of the R-type pyruvate kinase (PK) cDNA found in unrelated PK variants associated with hereditary hemolytic anemia. *Blood.* 1992b; 79: 1347-1350.

Kerner SA, Scott RA and Pike J W. Sequence elements in the human osteocalcin gene confer basal activation and inducible response to hormonal vitamin D<sub>3</sub>. *Proc Natl Acad Sci USA.* 1989; 86: 4455-4459.

Lederberg J. Anecdotal, Historical and Critical Commentaries on Genetics - J. B. S. Haldane (1949) on Infectious Disease and Evolution. *Genetics*. 1999; 153: 1-3.

Lemos MC, Fagulha A, Coutinho E, Gomes L, Bastos M, Barros L, Carrilho F, Geraldes E, Regateiro FJ and Carvalheiro M. Lack of association of vitamin D receptor gene polymorphisms with susceptibility to type 1 diabetes mellitus in the Portuguese population. *Hum Immunol*. 2008; 69: 134-138.

Lenzner C, Jacobasch G, Reis A, Thiele BJ and Nurnberg P. Trinucleotide repeat polymorphism at the PKLR locus. *Hum Mol Genet*. 1994b; 3: 523.

Lenzner C, Nurnberg P, Jacobasch G and Thiele BJ. Complete genomic sequence of the human PK-L/R-gene includes four intragenic polymorphisms defining different haplotype backgrounds of normal and mutant PK-genes. *DNA sequence: the Journal of DNA Sequencing and Mapping*. 1997b; 8: 45-53.

Lenzner C, Nurnberg P, Jacobasch G, Gerth C and Thiele BJ. Molecular analysis of 29 pyruvate kinase-deficient patients from central Europe with hereditary hemolytic anemia. *Blood*. 1997a; 89: 1793-1799.

Lewis SJ, Baker I and Smith GD. Meta-analysis of vitamin D receptor polymorphisms and pulmonary tuberculosis risk. *Int J Tuberc Lung Dis*. 2005; 9: 1174-1177.

Lewontin RC. On measures of gametic disequilibrium. *Genetics*. 1988; 120: 849-852.

Lian J, Stewart C, Puchacz E, Mackowiak S, Shalhoub V, Collart D and Zambetti G. Structure of the rat osteocalcin gene and regulation of vitamin D- dependent expression. *Proc Natl Acad Sci USA*. 1989; 86: 1143-1147.

Liu PT, Stenger S, Li H, Wenzel L, Tan BH, Krutzik SR, Ochoa MT, Schaubert J, Wu K, Meinken C, Kamen DL, Wagner M, Bals R, Steinmeyer A, Zügel U, Gallo RL, Eisenberg D, Hewison M, Hollis BW, Adams JS, Bloom BR and Modlin RL. Toll-Like Receptor triggering of a vitamin D - mediated human antimicrobial response. *Science*. 2006; 311: 1770-1773.

Lombard Z, Dalton D-L, Venter PA, Williams RC and Bornman L. Association of HLA-DR, -DQ, and vitamin D receptor alleles and haplotypes with tuberculosis in the Venda of South Africa. *Hum Immunol*. 2006; 67: 643-654.

Machado P, Pereira R, Rocha AM, Manco L, Fernandes N, Miranda J, Ribeiro L, Rosário VE, Amorim A, Gusmão L and Arez AP. Malaria: looking for selection signatures in the human *PKLR* gene region. *Br J Haematol.* 2010; 149: 775-784.

Malloy PJ, Pike WJ and Feldman D. The vitamin D receptor and the syndrome of hereditary 1,25-dihydroxyvitamin D-resistant rickets. *Endocr Rev.* 1999; 20: 156-188.

Manco L, Oliveira AL, Gomes C, Granjo A, Trovoada MJ, Ribeiro ML, Abade A and Amorim A. Population genetics of four *PKLR* intragenic polymorphisms in Portugal and São Tomé e Príncipe (Gulf of Guinea). *Hum Biol.* 2001a; 73: 467-474.

Manco L, Ribeiro ML, Maximo V, Almeida H, Costa A, Freitas O, Barbot J, Abade A and Tamagnini G. A new *PKLR* gene mutation in the R-type promoter region affects the gene transcription causing pyruvate kinase deficiency. *Br J Haematol.* 2000; 110: 993-997.

Marques CS. Estudo de associação entre o gene VDR e a hanseníase. Rio de Janeiro; 2010. Mestrado [Dissertação em Biologia Celular e Molecular] - Instituto Oswaldo Cruz.

Mateu E, Comas D, Calafell F, Perez-Lezaun A, Abade A and Bertranpetit J. A tale of two islands: population history and mitochondrial DNA sequence variation of Bioko and São Tomé, Gulf of Guinea. *Ann Hum Genet.* 1997; 61: 507-518.

Miller LH. Impact of malaria on genetic polymorphism and genetic diseases in Africans and African Americans. *Proc Natl Acad Sci USA.* 1994; 91: 2415-2419.

Miller SA, Dykes DD and Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 1988; 16: 1215.

Min-Oo G, Fortin A, Tam M, Nantel A, Stevenson MM and Gros P. Pyruvate kinase deficiency in mice protects against malaria. *Nat Genet.* 2003; 35: 357-362.

Ministério da Saúde. Direcção Nacional de Saúde Pública. Boletim Epidemiológico N° 4/2009. 2009. [Data de acesso: 10.05.2010]. Disponível em [http://www.misau.gov.mz/pt/epidemias\\_endemias/vigilancia\\_epidemiologica](http://www.misau.gov.mz/pt/epidemias_endemias/vigilancia_epidemiologica).

Ministério da Saúde. Programa Nacional de Controlo da Malária Relatório da Malária. Maputo: 2007. [Data de acesso: 15.05.2010]. Disponível em [http://www.misau.gov.mz/pt/programas/malaria/relatorio\\_de\\_malaria\\_2007](http://www.misau.gov.mz/pt/programas/malaria/relatorio_de_malaria_2007).



Ministério da Saúde. Programa Nacional de Controlo da Tuberculose. Relatório da Tuberculose. Maputo: 2007. [Data de acesso: 15.05.2010]. Disponível em [http://www.misa.gov.mz/pt/programas/tuberculose/relatorio\\_da\\_tuberculose\\_2007](http://www.misa.gov.mz/pt/programas/tuberculose/relatorio_da_tuberculose_2007).

Mockenhaupt FP, Ehrhardt S, Cramer JP, Otchwemah RN, Anemana SD, Goltz K, Mylius F, Dietz E, Eggelte TA and Bienzle U. Hemoglobin C and resistance to severe malaria in Ghanaian children. *J Infect Dis.* 2004; 190: 1006-1009.

Morrison NA, Shine J, Fragonas J-C, Verkest V, McMenemy ML and Eisman JA. 1,25-dihydroxyvitamin D-responsive element and glucocorticoid repression in the osteocalcin gene. *Science.* 1989; 246: 1158-1161.

Mory DB, Rocco ER, Miranda WL, Kasamatsu T, Crispim F and Dib SA. Prevalence of vitamin D receptor gene polymorphisms *FokI* and *BsmI* in Brazilian individuals with type 1 diabetes and their relation to  $\beta$ -cell autoimmunity and to remaining  $\beta$ -cell function. *Hum Immunol.* 2009; 70: 447-451.

Motsinger-Reif AA, Antas PRZ, Oki NO, Levy S, Holland SM and Sterling TR. Polymorphisms in IL-1 $\beta$ , vitamin D receptor *FokI*, and Toll-like receptor 2 are associated with extrapulmonary tuberculosis. 2010. *BMC Med Genet* [Online]. 11: 37. [Data de acesso 20.06.2010]. Disponível em <http://www.biomedcentral.com/1471-2350/11/37>.

Nairz M, Theurl I, Ludwiczek S, Theurl M, Mair SM, Fritsche G and Weiss G. The co-ordinated regulation of iron homeostasis in murine macrophages limits the availability of iron for intracellular *Salmonella typhimurium*. *Cell Microbiol.* 2007; 9: 2126-2140.

Nejentsev S, Godfrey L, Snook H, Rance H, Nutland S, Walker NM, Lam AC, Guja C, Ionescu-Tirgoviste C, Undlien DE, Rønningen KS, Tuomilehto-Wolf E, Tuomilehto J, Newport MJ, Clayton DG and Todd JA. Comparative high-resolution analysis of linkage disequilibrium and tag single nucleotide polymorphisms between populations in the vitamin D receptor gene. *Hum Mol Genet.* 2004; 13: 1633-1639.

Noguchi T, Inoue H and Tanaka T. The M1- and M2- type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. *J Biol Chem.* 1986; 261: 13807-13812.

Noguchi T, Yamada K, Inoue H, Matsuda T and Tanaka T. The L- and R-type isozymes of rat pyruvate kinase are produced from a single gene by use of different promoters. *J Biol Chem.* 1987; 262: 14366-14371.

Pacheco AG and Moraes MO. Genetic polymorphisms of infectious diseases in case-control studies. *Dis Markers*. 2009; 27: 173-186.

Pan X-M, Li D-R, Yang L, Wang E-Y, Chen T-Y, Liu Y-J, Liu M and Liao Z-G. No association between vitamin D receptor polymorphisms and coronary artery disease in a Chinese population. *DNA Cell Biol*. 2009; 28: 521–525.

Parra FC, Amado RC, Lambertucci JR, Rocha J, Antunes CM and Pena SDJ. Color and genomic ancestry in Brazilians. *Proc Natl Acad Sci USA*. 2003; 100: 177-182.

Pereira L, Gusmão L, Alves C, Amorim A and Prata MJ. Bantu and European Y-lineages in sub-Saharan Africa. *Ann Hum Genet*. 2002; 66: 369-378.

Pereira L, Gusmão L, Prata MJ, Mota P, Trovoada MJ and Amorim A. Detection of additional structural variation at the FES/FPS system and population data from S. Tomé e Príncipe and North Portugal. *Int J Legal Med*. 1999; 112: 204-206.

Pereira L, Macaulay V, Torroni A, Scozzari R, Prata MJ and Amorim A. Prehistoric and historic traces in the mtDNA of Mozambique: insights into the Bantu expansions and the slave trade. *Ann Hum Genet*. 2001; 65: 439-458.

Pissard S, Max-Audit I, Skopinski L, Vasson A, Vivien P, Bimet C, Goossens M, Galacteros F and Wajcman H. Pyruvate kinase deficiency in France: a 3-year study reveals 27 new mutations. *Br J Haematol*. 2006; 133: 683-689.

Preston RJ. Molecular epidemiology: potential impacts on the assessment of public health. *Mutat Res*. 2003; 543: 121-124.

Rezende VB, Barbosa Jr. F, Montenegro MF, Sandrim VC, Gerlach RF and Tanus-Santos JE. An interethnic comparison of the distribution of vitamin D receptor genotypes and haplotypes. *Clin Chim Acta*. 2007; 384: 155-159.

Risch N and Merikangas K. The future of genetic studies of complex human diseases. *Science*. 1996; 273: 1516-1517.

Roy MF, Riendeau N, Bédard C, Hélie P, Min-Oo G, Turcotte K, Gros P, Canonne-Hergaux F and Malo D. Pyruvate kinase deficiency confers susceptibility to *Salmonella typhimurium* infection in mice. *J Exp Med*. 2007; 204: 2949-2961.

Ruwende C, HKhoo SC, Snow RW, Yates SNR, Kwiatkowski D, Gupta S, Warn P, Allsopp CEM, Gilbert SC, Peschu N, Newbold CI, Greenwood BM, Marsh K and Hill AVS. Natural selection of hemi and heterozygotes for G6PD deficiency in Africa by resistance to severe malaria. *Nature*. 1995; 376: 246-249.

Saijo T, Ito M, Takeda E, Mahbulul Huq AHM, Naito E, IYokota I, Sone T, Pike J W and Kuroda Y. A unique mutation in the vitamin D receptor gene in three Japanese patients with vitamin D-dependent rickets type II - Utility of single-strand conformation polymorphism analysis for heterozygous carrier detection. *Am J Hum Genet*. 1991; 49: 668-673.

Satoh H, Tani K, Yoshida MC, Sasaki M, Miwa S and Fujii H. The human liver-type pyruvate kinase (PKL) gene is on chromosome 1 at band q21. *Cytogenet Cell Genet*. 1988; 47: 132-133 (abstract).

Schuster SC, Miller W, Ratan A, Tomsho LP, Giardine B, Kasson LR, Harris RS, Petersen DC, Zhao F, Qi J, Alkan C, Kidd JM, Sun Y, Drautz DI, Bouffard P, Muzny DM, Reid JG, Nazareth LV, Wang Q, Burhans R, Riemer C, Wittekindt NE, Moorjani P, Tindall EA, Danko CG, Teo WS, Buboltz AM, Zhang Z, Ma Q, Oosthuysen A, Steenkamp AW, Oostuisen H, Venter P, Gajewski J, Zhang Y, Pugh BF, KD Makova, Nekrutenko A, Mardis ER, Patterson N, Pringle TH, Chiaromonte F, Mullikin JC, Eichler EE, Hardison RC, Gibbs RA, Harkins TT and Hayes VM. Complete Khoisan and Bantu genomes from southern Africa. *Nature*. 2010; 463: 943-947.

Serra C (coord). *História de Moçambique: Parte I - Primeiras Sociedades sedentárias e impacto dos mercadores, 200/300- 1885; Parte II - Agressão imperialista, 1886-1930. . 2.ª ed.* Maputo: Livraria Universitária, Universidade Eduardo Mondlane; 2000.

Silva MCG, Alta Guiné e Maranhão: tráfico atlântico e rotas comerciais na segunda metade do século XVIII, in 4<sup>o</sup> Encontro Escravidão e Liberdade no Brasil Meridional, 2009, Curitiba.

Takegawa S, Fujii H and Miwa S. Change of pyruvate kinase isozymes from M2- to L-type during development of the red cell. *Br J Haematol*. 1983; 54: 467-474.

Tani K, Fujii H, Nagata S and Miwa S. Human liver type pyruvate kinase: complete aminoacid sequence and the expression in mammalian cells. *Proc Natl Acad Sci USA*. 1988; 85: 1792-1795.

Tiret L, Poirier O, Nicaud V, Barboux S, Herrmann S-M, Perret C, Raoux S, Francomme C, Lebard G, Trégouët D and Cambien F. Heterogeneity of linkage disequilibrium in human genes has implications for association studies of common diseases. *Hum Mol Genet*. 2002; 11: 419-429.

Tishkoff SA and Williams SM. Genetic analysis of African populations: Human evolution and complex disease. *Nat Rev Genet.* 2002; 3: 611-621.

Trovoada MJ, Alves C, Gusmão L, Abade A, Amorim A and Prata MJ. Evidence for population sub-structuring in São Tomé e Príncipe as inferred from Y-chromosome STR analysis. *Ann Hum Genet.* 2001; 65: 271-283.

Trovoada MJ, Pereira L, Gusmão L, Abade A, Amorim A and Prata MJ. Pattern of mtDNA variation in three populations from São Tomé e Príncipe. *Ann Hum Genet.* 2004; 68: 40-54.

Trovoada MJ, Tavares L, Gusmão L, Alves C, Abade A, Amorim A and Prata MJ. Dissecting the genetic history of São Tomé e Príncipe: a new window from Y-chromosome biallelic markers. *Ann Hum Genet.* 2007; 71: 77-85.

Uitterlinden AG, Fang Y, Van Meurs JBJ, Pols HAP and Van Leeuwen JPTM. Genetics and biology of vitamin D receptor polymorphisms. *Gene.* 2004; 338: 143-156.

Valdivielso JM and Fernandez E. Vitamin D receptor polymorphisms and diseases. *Clin Chim Acta.* 2006; 371: 1-12.

van Wijk R, van Solinge WW, Nerlov C, Beutler E, Gelbart T, Rijksen G and Nielsen FC. Disruption of a novel regulatory element in the erythroid-specific promoter of the human *PKLR* gene causes severe pyruvate kinase deficiency. *Blood.* 2003; 101: 1596-1602.

Verstuyf A, Carmeliet G, Bouillon R and Mathieu C. Vitamin D: a pleiotropic hormone. 2010. *Kidney Int* [Online]. [Data de acesso 09.06.2010]. Disponível em <http://www.nature.com/ki/journal/vaop/ncurrent/pdf/ki201017a.pdf>.

Whitfield GK, Selznick SH, Haussler CA, Hsieh JC, Galligan MA, Jurutka PW, Thompson PD, Lee SM, Zerwekh JE and Haussler MR. Vitamin D receptors from patients with resistance to 1,25-dihydroxyvitamin D<sub>3</sub>: point mutations confer reduced transactivation in response to ligand and impaired interaction with the retinoid X receptor heterodimeric partner. *Mol Endocrinol.* 1996 10: 1617-1631.

WHO. Organização Mundial de Saúde. Global Tuberculosis Control 2009. Epidemiology, Strategy, Financing. Geneva: 2009a. [Data de acesso: 10.09.2010]. Disponível em [http://whqlibdoc.who.int/publications/2009/9789241598866\\_eng.pdf](http://whqlibdoc.who.int/publications/2009/9789241598866_eng.pdf).

WHO. Organização Mundial da Saúde. Global tuberculosis control - A Short update to the 2009 report. Geneva: 2009b. [Data de acesso: 15.05.2010]. Disponível em [http://whqlibdoc.who.int/publications/2009/9789241598866\\_eng.pdf](http://whqlibdoc.who.int/publications/2009/9789241598866_eng.pdf).

WHO. Organização Mundial da Saúde. World malaria report 2009. Geneva: 2009c. [Data de acesso: 15.05.2010]. Disponível em [http://whqlibdoc.who.int/publications/2009/9789241563901\\_eng.pdf](http://whqlibdoc.who.int/publications/2009/9789241563901_eng.pdf).

Wilkinson RJ, Llewelyn M, Toossi Z, Patel P, Pasvol G, Lalvani A, Wright D, Latif M and Davidson RN. Influence of vitamin D deficiency and vitamin D receptor polymorphisms on tuberculosis among Gujarati Asians in west London: a case-control study. *Lancet*. 2000; 355:

Zanella A, Bianchi P, Baronciani L, Zappa M, Bredi E, Vercellati C, Alfinito F, Pelissero G and Sirchia G. Molecular characterization of *PK-LR* gene in pyruvate kinase deficient Italian patients. *Blood*. 1997; 89: 3847-3852.

Zanella A, Fermo E, Bianchi P and Valentini G. Red cel pyruvate kinase deficiency: molecular and clinical aspects. *Br J Haematol*. 2005; 130: 11-25.

Zhang H-Q, Deng A, Guo C-F, Wang Y-X, Chen L-Q, Wang Y-F, Wu J-H and Liu J-Y. Association between FokI polymorphism in vitamin D receptor gene and susceptibility to spinal tuberculosis in Chinese Han population. *Arch Med Res*. 2010; 41: 46-49.

Zmuda JM, Cauley JA, Danielson ME and Ferrell RE. Vitamin D receptor and aromatase gene interaction and bone Mass in older African-American women. *Metabolism*. 2003; 52: 521-523.

