

Universidade Federal do Rio Grande do Sul
Instituto de Biociências
Programa de Pós Graduação em Genética e Biologia Molecular

VARIAÇÕES GENÉTICAS NA VIA DE *TP53* EM NATIVOS AMERICANOS:
SINAIS DE ADAPTAÇÕES AO CLIMA

Vanessa Cristina Jacovas

Dissertação submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Mestre em Genética e Biologia Molecular.

Orientadora: Prof^a Dr^a Maria Cátira Bortolini

Coorientadora: Dr^a Virginia Ramallo

Porto Alegre

Março 2015

Este trabalho foi realizado e desenvolvido do Laboratório de Evolução Humana e Molecular (LEHM) do departamento de genética do Instituto de Biociências da Universidade Federal do Rio Grande do Sul (UFRGS) com recursos do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e FAPERGS (Fundação de Amparo à Pesquisa do Rio Grande do Sul)

AGRADECIMENTOS

Aos meus pais, pelo amor e apoio incondicional em todos os momentos da minha vida. Por todo o esforço que vocês fizeram pra que nós, eu e o mano, tivéssemos a oportunidade de ser quem nós quiséssemos ser.

Ao meu irmão, pelo carinho e amizade, e também por ter baixado o volume do *video game* pra eu poder me concentrar!

À professora Maria Cátira, pela confiança em mim depositada antes mesmo de me conhecer, pelos conhecimentos transmitidos diariamente, pela orientação, incentivo, e principalmente por ser para todos nós, uma referência de pesquisadora e mulher.

À Virginia Ramallo por toda ajuda incondicional, pelas palavras de força, e por estar sempre disponível para explicar uma, duas, ou quantas vezes fossem necessárias.

Ao professor Salzano, pelo exemplo de excelente profissional, humildade e pelo incentivo diário.

Aos amigos do Laboratório de Evolução Humana e Molecular pelo apoio, amizade e parceria que cresce a cada dia que passa.

E por fim, mas não menos importante, obrigada Rafael, pela amizade, parceria, paciência e carinho. Por ter me dado a mão e caminhar ao meu lado na vida. Juntos alcançaremos todos os nossos sonhos!

LISTA DE ABREVIATURAS

3' UTR: *3' Untranslated region* (Região 3' não traduzida)

Arg: Arginina

CNPq: Conselho Nacional de Desenvolvimento Científico e

Tecnológico CONEP: Conselho Nacional de Ética e Pesquisa.

DNA: *deoxyribonucleic acid* (Ácido desoxirribonucleico)

HAUSP: Gene *Herpesvirus-associated ubiquitin-specific protease*

HAUSP: Proteína *Herpesvirus-associated ubiquitin-specific protease*

iRNA: RNA de interferência

LIF: Gene *Leukemia-inhibitory factor*

LIF: Proteína *Leukemia-inhibitory factor*

MDM2: Gene *Mouse double minute 2 homolog*

MDM2: Proteína *Mouse double minute 2 homolog*

MDM4: Gene *Mouse double minute 4 homolog*

MDM4: Proteína *Mouse double minute 4*

homolog miRNA: microRNA

NES: *Nuclear export signal* (Sinal de exportação nuclear)

NLS: *Nuclear localization signal* (Sinal de exportação

nuclear) p53: Proteína supressora tumoral p53

pb: pares de base.

PCR: *Polymerase Chain Reaction* (Reação em Cadeia da Polimerase). Pro: Prolina

SNP: *Single Nucleotide Polymorphism* (Polimorfismo de Nucleotídeo Único).

TP53: *Tumor protein p53*

UFRGS: Universidade Federal do Rio Grande do

Sul. *USP7*: Proteína *Ubiquitin-specific protease 7*

USP7: Gene *Ubiquitin-specific protease*

7 UV: Ultravioleta

A, C, G e T: Bases nitrogenadas adenina, citosina, guanina e timina respectivamente.

RESUMO

Considerando o papel fundamental da proteína p53 na homeostase celular, buscamos no presente trabalho correlacionar variações comuns em genes ligados a rede clássica de *TP53* com variáveis ambientais e geográficas em distintas regiões habitadas por nativos americanos.

Um total de 282 indivíduos (12 populações) caracterizados como nativo ou como tendo um nível elevado (>90%) de ascendência indígena americana foram estudados em relação aos cinco polimorfismos de nucleotídeo único de genes da via *TP53* (*TP53*, rs1042522; *MDM2*, rs2279744 ; *MDM4*, rs1563828; *USP7*, rs1529916 e *LIF*, rs929271).

Além disso, os dados da literatura foram obtidos para 100 pessoas de cinco outras populações nativas americanas. As amostras foram classificadas como vivendo em alta altitude ($\geq 2,500$ m) ou em terras baixas ($< 2,500$ m). Alelos *USP7-G*, *LIF-T*, e *MDM2-309-T* mostraram indícios significativos de que eles podem ter um papel protetor em relação às variáveis ambientais adversas relacionadas com a alta altitude.

Estes dados fornecem informações adicionais sobre um processo fisiológico chave responsável pela colonização humana bem sucedida das terras mais altas de nosso continente: os Andes.

Palavras-chave: Nativos Americanos, Adaptação à Altas Altitude, Genes da via *TP53*.

ABSTRACT

Considering the fundamental role of the protein p53 in the cellular homeostasis, we seek in the present study to correlate common variations in genes linked to classical network of *TP53* with environmental and geographical variables found in different regions inhabited by Native Americans.

A total of 282 individuals (12 populations) characterized as Native or as having a high level (>90%) of Native American ancestry were studied in relation to five single nucleotide polymorphisms of genes of the *TP53* pathway (*TP53*, rs1042522; *MDM2*, rs2279744; *MDM4*, rs1563828; *USP7*, rs1529916; and *LIF*, rs929271).

In addition, data from the literature was obtained for 100 persons from five other Native American populations. The samples were classified as living in high altitude ($\geq 2,500$ m) or in lowlands ($< 2,500$ m). Alleles *USP7-G*, *LIF-T*, and *MDM2-309-T* showed significant indications that they may have a protective role in relation to harsh environmental variables related to high altitude.

These data provide additional information about a key physiological process responsible for the successful human colonization of the Andes.

Key words: Native Americans, High altitude adaptation, *TP53* pathway genes.

SUMÁRIO

CAPÍTULO 1. INTRODUÇÃO	8
<i>1.1 Considerações gerais</i>	9
<i>1.2 O gene TP53</i>	9
<i>1.3 Rede clássica de TP53</i>	14
<i>1.4 O gene MDM2</i>	15
<i>1.5 O gene MDM4</i>	19
<i>1.6 O gene USP7</i>	21
<i>1.7 O gene LIF</i>	24
<i>1.8 Populações Nativo-americanas</i>	26
<i>1.9 Adaptações genéticas mediadas pelo clima</i>	27
CAPÍTULO 2. OBJETIVOS	30
<i>2.1 Objetivo Geral</i>	31
<i>2.2 Objetivos Específicos</i>	31
CAPÍTULO 3. RESULTADOS EM FORMATO DE ARTIGO	32
<i>Genetic variation in the TP53 pathway in Native Americans: signals of climate adaptation</i>	
CAPÍTULO 4. DISCUSSÃO	67
CAPÍTULO 5. CONCLUSÕES	71
REFERÊNCIAS BIBLIOGRÁFICAS	73

CAPÍTULO 1
INTRODUÇÃO

1.1 Considerações gerais:

Desde a migração do *Homo sapiens* para fora da África, há cerca de 50-100 mil anos atrás, os humanos modernos tem colonizado uma ampla gama de ambientes, e adaptações comportamentais e biológicas têm sido necessárias para o sucesso desta extraordinária trajetória.

Considerando a variabilidade dos ambientes e ecossistemas dentro do continente americano, é razoável supor que parte dela possa ter sido moldada por processos evolutivos direcionais, tal como a seleção natural Darwiniana ou positiva.

Variações em genes humanos têm sido associadas com prováveis pressões seletivas, incluindo dentro do continente Americano (Hünemeier *et al.*, 2012). Surpreendentemente, o gene *TP53*, cuja proteína traduzida, p53, é considerada "guardiã do genoma", nunca foi estudado num contexto evolutivo, considerando nativos americanos. p53 é mantido em níveis basais em células não estressadas, devido ao seu curto tempo de meia-vida, cerca de 30 minutos. No entanto, frente à danos ao DNA como aqueles produzidos pela hipóxia e radiação ultravioleta, os níveis de p53 na célula são aumentados para que haja uma resposta celular adequada.

Considerando o papel fundamental da proteína p53 na homeostase celular, buscamos encontrar variações em genes ligados a rede clássica de p53, seus principais reguladores, a fim de que possamos correlacionar essas modificações genéticas com os ecossistemas distintos habitados por nativos americanos.

1.2 O gene *TP53*:

A proteína p53 foi descrita em 1979 (Lane e Crawford, 1979; Linzer e Levine, 1979). Os autores determinaram que quando p53 estava associada a antígenos T em células transformadas e infectadas com vírus símio 40 (SV40), esta se encontrava expressa em elevados níveis celulares em diversos tipos de células tumorais. Assim, foi sugerido que o gene codificador dessa proteína, *TP53* (*tumor protein p53*, OMIN191170), fosse um oncogene, ou seja, que ele estaria relacionado com o surgimento de tumores (Lane e

Crawford, 1979; Linzer e Levine, 1979). Foi apenas dez anos mais tarde que se verificou que essa propriedade era resultado de formas mutantes do gene, enquanto que o alelo selvagem tinha a capacidade de suprimir a transformação de um tumor, tanto *in vitro* quanto *in vivo* (Finlay *et al.*, 1989). Em 1992, Donehower e colegas descreveram que camundongos *knockout* para o gene *TP53* desenvolviam-se normalmente, embora tenham maior propensão a desenvolver tumores (Donehower *et al.*, 1992). Assim, *TP53* foi definitivamente identificado como um gene supressor tumoral, responsável principalmente pela manutenção da integridade genômica.

TP53 está localizado no braço curto do cromossomo 17 (17p13.1) abrangendo uma região de 20kb, possui 11 éxons, sendo o primeiro não codificador. A proteína p53, codificada por *TP53*, é constituída por 393 aminoácidos, e desempenha um papel fundamental coordenando respostas a estressores genotóxicos a fim de manter a estabilidade genômica (Linzer *et al.*, 1979; McBride *et al.*, 1986).

A proteína p53 apresenta quatro regiões com funções distintas, chamadas domínios da proteína (Figura 1). O primeiro é chamado de domínio N-terminal de transativação, está localizado na extremidade amino-terminal. Está compreendido entre os aminoácidos 1 e 44 e é a região da proteína que interage com a maquinaria de transcrição de genes regulados positivamente por p53. A seguir está um domínio rico em prolinas (58-101) que está envolvido na interação proteína-proteína e na função apoptótica de p53. A região central é o domínio de ligação ao DNA, entre os aminoácidos (102-292), que possibilitam a ligação de p53 em sítios específicos do DNA. Essa região é altamente conservada evolutivamente, e é onde ocorre o maior número de mutações associadas à patologias como o câncer. Na extremidade carboxi-terminal (C-terminal) (325-356) encontramos o domínio de oligomerização, que é responsável pela dimerização e posterior tetramerização da proteína. Há ainda sinais de exportação nuclear (NES - *Nuclear export signal*) e sinais de localização nuclear (NLS - *Nuclear localization signal*) (Laptenko e Prives 2006; Brosh e Rotter 2009).

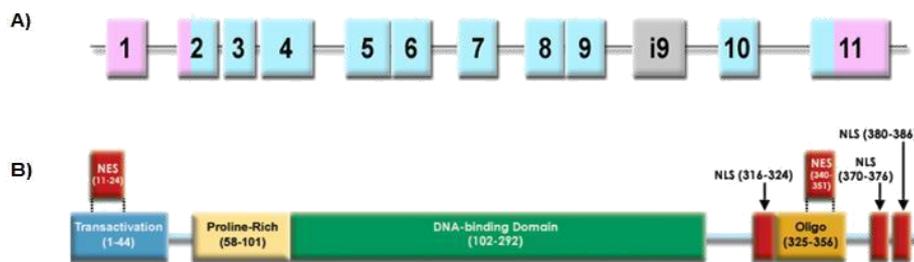


Figura 1. A) Distribuição dos 11 éxons do gene *TP53*, a cor rosa demonstra regiões não traduzidas, a cor azul a região codificadora, e a cor cinza denota um éxon dentro de um íntron. B) Respective domínios da proteína p53. Adaptado de *p53knowledgebase* (<http://p53.bii.a-star.edu.sg/aboutp53/>)

A quantidade de p53 nas células é determinada principalmente pela taxa na qual ela é degradada, ao invés da taxa na qual é produzida. Ela é constitutivamente expressa em todos os tipos de célula, mas em células normais ela não se acumula e é mantida em níveis basais devido à sua rápida degradação via proteossomo (Nag *et al.*, 2013; Chao, 2015). No entanto, em resposta a diferentes sinais celulares, ocorre um acúmulo de p53, além de modificações pós-transcricionais na proteína, que vão gerar resposta a tais estímulos (Chen, 2012). Por monitorar a integridade do genoma e atuar como um sensor de danos no DNA auxiliando no recrutamento de moléculas do sistema de reparo, o termo “guardião do genoma” tem sido frequentemente atribuído à p53. Esses sinais incluem agentes prejudiciais ao DNA (estresse genotóxico, como a radiação UV), ativação constitutiva de cascatas de sinalização de crescimento (estresse oncogênico), bem como outros tipos de estresse, tais como depleção de ribonucleotídeos ou hipóxia (Figura 2) (Ljungman, 2000).



Figura 2. Mecanismos que podem ativar p53 e as possíveis respostas celulares desencadeadas.

O dano ao DNA desencadeia rotas celulares sinalizadoras que garantem que a p53 se acumula no núcleo através de diversos mecanismos pós-traducionais que incluem fosforilação, ubiquitinação, acetilação metilação, sumorilação e nedilação, bem como a regulação por miRNAs, além da interação com outras proteínas (Brooks e Gu, 2003; Le *et al.*, 2009). Entre essas modificações, a ubiquitinação, fosforilação e acetilação são as mais amplamente estudadas. Elas estão envolvidas na regulação de p53 e envolvem 3 passos para a sua ativação: (1) estabilização de p53, (2) ligação ao DNA e (3) ativação transcricional. A ubiquitinação contribui de forma significativa para o controle da estabilidade de p53, e esta modificação é mediada principalmente pela oncoproteína MDM2 (*Mouse double minute 2 homolog*) seu principal regulador negativo (Shi e Gu 2012).

Mutações pontuais, bem como perda de alelos no gene *TP53*, são verificados em quase todos os tipos de câncer. Sendo que mutações são encontradas neste gene em cerca de 50% dos tipos de cânceres humanos. As mutações são diversificadas na sua posição e natureza, podendo promover a alteração nos aminoácidos da proteína p53, alterando sua função, podendo causar fenótipos de câncer agressivos (Garritano *et al.*, 2013). Um banco de dados de todas as mutações publicados é mantida na *International Agency for Research on Cancer* (<http://www-p53.iarc.fr/p53database.htm>) onde cerca de 30000 mutações, dentre elas somáticas e germinativas, já foram descritas.

Dentre todas as mutações descritas até hoje no gene *TP53*, aproximadamente 85% são mutações de ponto, e dentre as mutações de ponto, cerca de 86% são do tipo *missense*, ou seja, acarreta a troca do aminoácido a ser codificado. A maior parte delas ocorre em regiões exônicas (preferencialmente éxons 5, 6, 7 e 8) localizadas no domínio de ligação ao DNA. Nesta região, destacam-se alguns códons que encontram-se mais frequentemente mutados. Eles abrigam cerca de 30% de todas as mutações conhecidas no gene: R175, G245, R248, R249, R273 e R282, os chamados "*hotspots*" (*IARC Database*).

Diversos polimorfismos foram descritos na sequência do gene *TP53* em diferentes populações humanas. As frequências alélicas destes polimorfismos podem divergir de maneira marcante entre continentes e/ou populações, o que sugere que eles podem estar sujeitos à seleção natural.

Cerca de 90% desses polimorfismos são encontrados em regiões não codificadoras, sendo que há apenas 11 polimorfismos não sinônimos descritos na região codificadora do gene (*IARC Database*).

Uma das variantes mais bem estudadas no gene *TP53* é o SNP (*Single Nucleotide Polymorphism*) c.215C>G, que consiste na troca de uma Prolina (Pro: alelo C) por uma arginina (Arg: alelo G) p.Pro72Arg, no códon 72, éxon 4, rs1042522 (Matlashewski *et al.*, 1987). A isoforma Pro72 é também observada em outros primatas, incluindo chimpanzé, enquanto que Arg72 é presente apenas em humanos, o que nos mostra que alelo C (Pro 72) corresponde ao alelo ancestral.

Há um consenso atual de que a isoforma Arg72 seja mais eficiente em induzir a apoptose. Dumont e colaboradores (2003) demonstraram que células contendo a isoforma Arg72 tem maior habilidade em direcionar p53 para a membrana mitocondrial, assim, após a liberação do citocromo c, há a ativação da cascata de caspases culminando no processo apoptótico. Além disso, descreveram que esta isoforma interage de forma mais eficiente ao principal regulador negativo de p53, MDM2, sendo mais eficiente em realizar a exportação nuclear de p53 quando comparada a isoforma Pro72 de p53 (Dumont *et al.*, 2003). Em contraste, Pro72 parece ter uma capacidade elevada de induzir a parada do ciclo celular na fase G1, devido ao fato dessa isoforma interagir mais eficientemente com elementos da maquinaria transcricional. Isso demonstra diferenças significantes em como o polimorfismo Pro72Arg afeta a atividade biológica de p53 (Pin e Banks 2004). No entanto, ainda não é bem compreendido como que essas diferenças funcionais entre Arg72 e Pro72 possam se comportar em diferentes tipos celulares e qual a relevância dessa diferença *in vivo* (Whibley *et al.*, 2009).

Beckman e colaboradores (1994) foram os primeiros a demonstrar uma diferença nas distribuições alélicas das variantes Arg72 e Pro72 de p53. Eles observaram pela primeira vez a diferença nas frequências do alelo Pro72 entre uma população nigeriana (Africanos) e uma população sueca (Europa Ocidental), que foram de 17 e 63%, respectivamente, e observaram diferenças nas frequências desses alelos entre as populações que vivem na mesma latitude geográfica. Assim surgiu a hipótese de que a distribuição do mesmo seria dependente da latitude (Beckman *et al.*, 1994 e Sjalander *et al.*, 1996). Dada à formação de um cline em relação às frequências alélicas do polimorfismo Pro72Arg, pode ser sugerida a hipótese de que a distribuição geográfica das

frequências alélicas seja coordenada e moldada por pressões seletivas devido a variáveis ambientais como a exposição aos raios ultravioleta.

Shi *et al.* (2009) descreveram uma correlação entre a frequência dos alelos desse polimorfismo com variáveis ambientais em populações do Leste Asiático que habitam diferentes regiões latitudinais. Os autores verificaram uma clara dependência de latitude para a ocorrência do alelo Arg72: quanto maior a latitude, maior era a frequência do mesmo ($p < 0,01$). A fim de avaliar se esta distribuição latitude-dependente do alelo Arg72 era devido à variação média de exposição à radiação UV que as populações que vivem em diferentes latitudes estariam sujeitas, Shi e colaboradores (2009) coletaram dados da média de intensidade de radiação UV no Centro Nacional de Informação Meteorológica da China para todas as regiões onde havia amostras coletadas. Como resultado, este grupo de pesquisadores observou que, a frequência do alelo Arg72 não estava correlacionada com a intensidade média de radiação UV ($p = 0,619$), sugerindo que a variação latitude-dependente dos alelos do locus do códon 72 de p53 não era devido a essa condição do ambiente.

No mesmo trabalho, Shi e colegas verificaram outro parâmetro ambiental importante: a temperatura, que também varia significativamente com a latitude. Diferentemente do primeiro caso, os autores observaram uma forte associação de temperaturas mais baixas com uma alta frequência do alelo Arg72 ($p < 0,01$), sugerindo assim que há uma pressão seletiva positiva para a ocorrência em altas frequências deste alelo em climas frios Shi *et al.* (2009).

1.3 Rede clássica de TP53:

Para ocorrer uma resposta transcricional apropriada de p53, outros genes nessa rede devem estar com suas atividades normalizadas e funcionando adequadamente. Dentre esses podem ser citados alguns, cujos produtos são reguladores críticos da proteína p53:

MDM2 (*Mouse double minute 2 homolog*, OMIM 164785), *MDM4* (*Mouse double minute 4 homolog*, OMIM 602704) *USP7* (*Ubiquitin-specific protease 7*, OMIM 602519) também

conhecida como *HAUSP* (*Herpesvirus-associated ubiquitin-specific protease*) e tendo sua regulação modulada via p53, *LIF* (*Leukemia-inhibitory factor*, OMIM 159540); Figura 3.

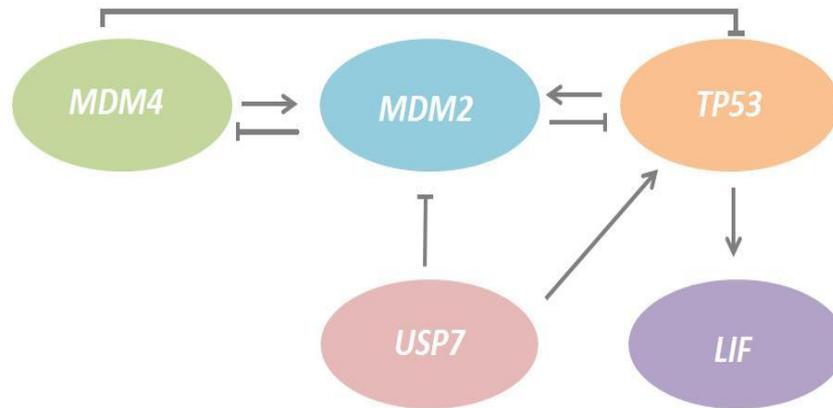


Figura 3. Rede clássica dos genes que interagem com *TP53*; (Adaptado de Kang *et al.*, 2009).

1.4 O gene *MDM2*:

O gene *MDM2* (*Murine Double Minute Clone 2*, OMIM 164785) foi originalmente identificado a partir da purificação de cromossomos acêntricos em células espontaneamente transformada Balb/c 3T3. Quando um clone genômico de *MDM2* foi amplificado em células de roedores, observou-se um alto potencial tumorigênico, o que levou a primeira sugestão de que este fosse um oncogene (Iwakuma e Lozano 2003; Nag *et al.*, 2013)

Nessa mesma época, havia se descoberto que o gene *TP53* não era um oncogene, e sim um gene supressor tumoral, assim, pesquisas com a proteína p53 estavam surgindo paralelamente em diversos laboratórios. Foi em um desses estudos, através do isolamento de proteínas acopladas a p53 com a justificativa de que tais proteínas podem regular a atividade p53, que se observou uma fosfoproteína de 90 kDa que co-imunoprecipitou junto com p53. Essa proteína acabou por ser o produto do gene *MDM2*, MDM2. Além da descrição, neste mesmo trabalho já iniciaram-se estudos da interação p53- MDM2, e foi observado que quando MDM2 estava super expresso a propensão à cânceres era maior.

Assim, após diversos experimentos, pela primeira vez foi descrito o complexo de interação MDM2-p53, e que MDM2 pode inibir a transativação de p53. Em estudo subsequente foi demonstrado a regulação negativa de MDM2 sobre p53. Essas observações levam a hipótese de que a superexpressão de *MDM2* pode servir como mecanismo molecular pelo qual a célula pode inativar p53, permitindo que o câncer se desenvolva, o que foi posteriormente confirmado (Momand *et al.*, 1992).

Estudos genéticos *in vivo* fortaleceram a ideia de que MDM2 seja regulador de p53. Ratos *knockout* para *MDM2* mostram letalidade embrionária, enquanto que a concomitante deleção de ambos os alelos (em *MDM2* e *TP53*) salva os embriões, indicando que p53 é letal na ausência do seu regulador negativo, MDM2, durante o desenvolvimento embrionário desses ratos (Montes *et al.*, 2005).

Em humanos, o gene *MDM2* está presente no braço longo do cromossomo 12 (12q13-q14) e apresenta dois *enhancer* que regulam sua expressão. A primeira região está localizada na região 5' do primeiro éxon e preferencialmente regula os níveis basais de MDM2 em uma célula não submetida a estresse. A segunda região está no primeiro íntron e contém uma sequência de DNA que aumenta a expressão de MDM2 em resposta a diferentes estímulos. (Bond *et al.*, 2004). MDM2 é uma proteína E3 ubiquitina-ligase membro da família Ring finger, constituída de 491 aminoácidos, com aproximadamente 56kDa, cuja principal função é regular os níveis e a atividade de p53 (Zhang *et al.*, 2014; Bohlman e Manfredi, 2014).

As proteínas p53 e MDM2 possuem feedback auto-regulado, assim p53 regula positivamente MDM2, estimulando a transcrição do gene. A proteína MDM2, por sua vez, regula negativamente p53 (Kohn e Pommier, 2005 e Nag *et al.*, 2013). MDM2 também contém uma sequência sinal que é similar a um sinal de exportação de diversas proteínas virais. Assim, depois de se ligar a p53 ele induz a sua exportação nuclear (Tao e Levine 1999). No citoplasma MDM2 funciona como uma E3-ubiquitina-ligase e p53 se torna alvo da degradação pelo proteossomo (Haupt *et al.*, 1997).

MDM2 possui quatro domínios funcionais independentes (Figura 4): 1) um domínio N-terminal (em lilás) que é onde p53 se liga (indo até o domínio ácido, em azul) aminoácidos 19-220; 2) um domínio ácido central, que se liga a inúmeras proteínas; 3)

domínio zinc finger; 4) domínio ring finger responsável pela atividade E3 da ubiquitina-ligase de MDM2 (Nag *et al.*, 2013).

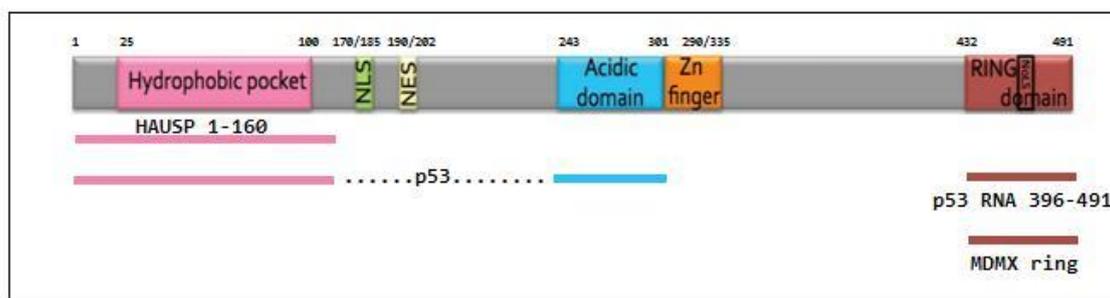


Figura 4. Representação esquemática dos domínios de MDM2. Algumas das proteínas que interagem com MDM2 estão demonstradas logo abaixo do domínio no qual interagem, representadas da mesma cor. (Adaptado Fähræus & Olivares-Illana, 2013)

Foi descrito um polimorfismo (transversão de Timina para Guanina) localizado no promotor/enhancer intrônico de resposta a p53, chamado SNP309T>G (rs2279744), que cria um sítio para o fator de transcrição *Sp1* (OMIM 189906) aumentando a sua afinidade pelo mesmo. Como resultado, homozigotos para o alelo G expressam mais RNA mensageiro e, conseqüentemente, mais proteína MDM2 é produzida quando comparado aos homozigotos para o alelo T. Com maior quantidade de MDM2 há uma atenuação na expressão de p53 celular, o que diminuiria a resposta da célula ao estresse, resultando em uma maior taxa de mutação, devido ao processo de reparo ao DNA prejudicado, redução da apoptose e senescência, aumentando assim o risco da tumorigênese (Bond *et al.*, 2004; Alazzouzi *et al.*, 2007).

Estudos têm demonstrado que os alelos 309-T e 309-G estão diferencialmente distribuídos em populações humanas (Millikan *et al.*, 2006). O alelo 309-G se encontra em frequência mais elevada em descendentes de europeus (43%) do que em afro-americanos 10% (Atwal *et al.*, 2007). A provável ausência de 309-G na África torna a hipótese da distribuição em gradiente através de diferentes latitudes também possível nesse caso.

A hipótese postulada é que a mutação Timina para Guanina teria ocorrido mais recentemente na história evolutiva humana, particularmente depois da saída do homem moderno da África. Devido ao fato de que 309-G levar a uma menor frequência de apoptose quando comparado ao alelo 309-T, a presença de G poderia compensar uma

maior frequência de apoptose trazida pela p53-Arg, quando os humanos modernos se moveram para fora da África.

No estudo de Shi *et al.*, (2009), além do SNP do códon 72 do gene *TP53*, os autores também investigaram o SNP309T>G. Eles notaram que o polimorfismo SNP309 de *MDM2* estava associado com a média de radiação UV, um fator ambiental dependente de latitude. A força da radiação UV mostrou uma significativa correlação negativa com a frequência do genótipo GG do SNP309 ($p < 0,01$). Assim, a alta frequência do genótipo GG estaria associada com a baixa intensidade de radiação UV.

A proteína p53 também atua na apoptose, uma via do organismo que elimina células danificadas durante o desenvolvimento e assim auxilia na manutenção da integridade do genoma. Deste modo, a perda funcional de p53 também leva à desregulação deste processo, que pode ocasionar diferentes problemas incluindo infertilidade. Huang *et al.*, (2012) avaliaram variações no gene *TP53* e *MDM2* (principal regulador de *TP53*, ver abaixo) em homens inférteis da população chinesa e observaram que o polimorfismo 309T>G de *MDM2* esta diretamente associado com um alto risco para infertilidade masculina.

O alelo G de *MDM2* em homozigose a uma elevada expressão da proteína (Alazzouzi *et al.*, 2007). Isto faz com que os níveis celulares de p53 diminuam, inibindo a via apoptótica de TP53. Assim, embora p53 Pro72Arg não tenha sido diretamente a causa da infertilidade nesses homens, ficou demonstrado que a combinação de alelos em genes desta rede afetam diretamente os níveis de p53.

Devido ao intenso grau de interação entre *TP53* e *MDM2* o estudo de polimorfismos em ambos os genes em diferentes populações ao redor do mundo ajudaria a elucidar suas interações e delinear um cenário de adaptação a diferentes climas.

1.5 O gene *MDM4*:

O gene *MDM4* (*Mouse Double Minute 4 Homolog*, OMIM 602704), codifica a proteína *MDM4* que, similarmente ao *MDM2*, atua no controle dos níveis de *p53*, diretamente, ou através da interação com *MDM2* (Zhang *et al.*, 2014). *MDM4* foi originalmente isolada a partir de uma biblioteca de expressão de cDNA de camundongos, e posteriormente descrita como uma nova proteína que interage com *p53*. (Shvarts *et al.*, 1996). O gene ortólogo humano, *MDM4*, ou *MDMX*, foi identificado um ano mais tarde, e está localizado no braço longo do cromossomo 1, na região 1q32.1 (Shvarts *et al.*, 1997).

A comparação entre as sequências de *MDM4* e *MDM2* revela uma conservação muito forte no domínio N-terminal de ambas as proteínas, justamente o domínio de ligação à *p53* (Figura 5). Outra região comum bem conservada entra ambas as proteínas é o domínio ring finger localizado na região carboxi-terminal de cada proteína. Sua integridade é essencial para heterodimerização *MDM4*- *MDM2* (Sharp *et al.*, 1999).

As regiões centrais da *MDM2* e *MDM4* não mostram nenhuma similaridade significativa, no entanto ambas as regiões são ricas em resíduos ácidos. Por causa da sua similaridade com *MDM2* e sua habilidade de inibir a transcrição de *p53*, foi hipotetizado que *MDM4* também atuasse como um regulador negativo de *p53*.

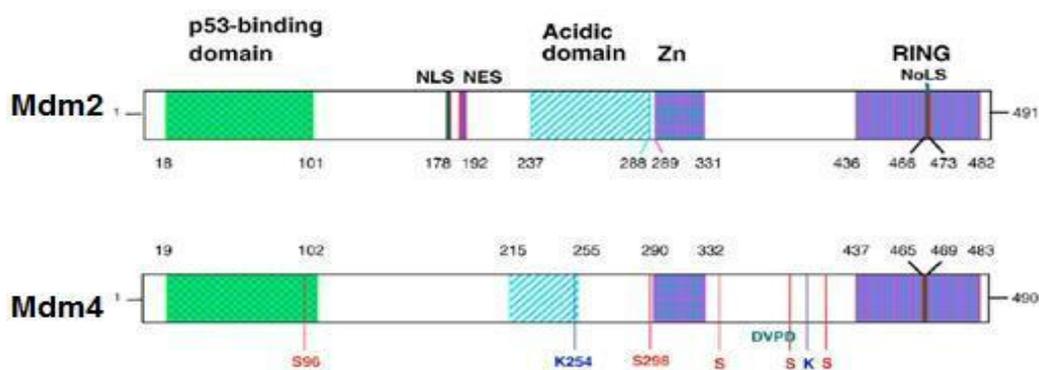


Figura 5. Comparação entre as estruturas primárias de *MDM2* e *MDM4*. Note que diversos domínios são conservados, dentre eles vale destacar o domínio de ligação com *p53*, zinc finger (Zn) e RING finger - contendo o sinal de localização nucleolar de localização (NoLS). A porcentagem de identidade compartilhada entre esses domínios é, respectivamente, 53.6%, 41,9% e 53.2%. Adaptado de Marine *et al.*, 2006.

De fato essas duas proteínas estão extremamente correlacionadas *in vivo*. A hetero-oligomerização entre MDM4 e MDM2 é descrita como muito mais estável do que a oligomerização de cada proteína. (Marine, 2007). O heterodímero formado por essas duas proteínas tem uma alta capacidade de ubiquitinação, causando então a degradação de p53 (Pei *et al.*, 2012). Além disso, MDM4 pode ainda interagir como MDM2 e inibir a sua degradação. (Pan e Chen 2003).

O fato de, nem MDM2, nem MDM4, poderem compensar a atividade um do outro *in vivo* para inibir p53, sugerem que ambos desempenham uma função crítica e não sobreposta na regulação de p53 (Wade *et al.*, 2010). No entanto, devido a letalidade embrionária precoce de indivíduos com mutações *MDM2-null* associada com *MDM4-null*, dificulta a elucidação de como se dá exatamente as contribuições fisiológicas de cada proteína. No entanto, estudos têm demonstrado que MDM2 funciona como uma E3 ubiquitina ligase que cataliza a ubiquitinação de p53, enquanto que MDM4 funciona principalmente se ligando a p53 e inibindo a ativação transcricional de p53 (Huang *et al.*, 2011).

A proteína MDM4 foi encontrada sob elevada expressão em diversas classes de tumores humanos (Li *et al.*, 2014; Fan *et al.*, 2014). Além disso, estudos sobre sua estrutura têm indicado a presença de diversos SNPs, e evidenciado que eles podem influenciar na sua interação com p53. Quando MDM4 se encontra em níveis elevados na célula, ocorre a diminuição da atividade de p53, o que resulta em uma capacidade diminuída de reparos ao DNA, aumentando a susceptibilidade a tumores.

Um polimorfismo bastante estudado neste gene, é a alteração de uma Citosina para timina (C>T) no 10º intron de *MDM4* (rs1563828). Um estudo analisando a correlação deste SNP com indivíduos com câncer de mama associou o genótipo TT do gene com desenvolvimento precoce de câncer de mama. Mulheres homocigotas para o alelo C foram diagnosticadas com média de 56 anos de idade, enquanto que heterocigotas e homocigotas para o alelo derivado foram diagnosticadas em média 4 ou 5 anos mais cedo respectivamente (Atwal *et al.*, 2009).

Outro estudo também com câncer de mama encontrou resultados que corroboram os de Atwal e colaboradores 2009, onde portadores do genótipo TT desenvolveram câncer

de mama cerca de 2 anos mais cedo quando comparados com indivíduos CC/CT (Song *et al.*, 2012).

Similarmente, este mesmo polimorfismo foi encontrado associado com idade precoce do diagnóstico de câncer de nasofaringe. A média de idade de diagnóstico em indivíduos com o genótipo CC foi de 48.2 anos, enquanto que indivíduos TT foi de 39.3 e heterozigotos CT apresentaram idade intermediária, 45.5 anos. Neste caso, portadores do genótipo TT foram diagnosticados cerca de 9 anos mais cedo (Zhang *et al.*, 2012).

Outro SNP no gene de *MDM4* (rs4245739), juntamente com a variante polimórfica de p53 Pro72Arg, foi relacionado a portadores de câncer de mama de etnia chinesa portadores de câncer de mama. A alteração funcional de A>C na região 3' não traduzida do gene, cria um sítio alvo do miR-191 que resulta em expressão diminuída de *MDM4* (Liu *et al.*, 2013). Recentemente outro estudo com esse mesmo polimorfismo realizado em população chinesa, revelou associação significativa do alelo C de *MDM4* com indivíduos portadores de Linfoma Não-Hodgkin (Fan *et al.*, 2014).

Já Terzian *et al.* (2007) demonstraram que semelhante a *MDM2*, a haploinsuficiência de *MDM4* leva à atividade aumentada de p53 exibindo um aumento da sensibilidade à danos do DNA, o que levaria a uma diminuição da transformação celular e desenvolvimento do tumor. Estes dados sugerem que, a mudança de apenas um par de base neste gene pode alterar os níveis de *MDM4* o suficiente para afetar a via de p53 e, portanto, o câncer em seres humanos.

Estes achados nos mostram que similarmente à *MDM2*, *MDM4* também tem um papel regulatório importante na via p53, sendo que variantes em *MDM4* podem estar sob forte pressão de seletiva.

1.6 O gene *USP7*:

Outro regulador imprescindível na rota da regulação de p53 é o gene *USP7* (*Ubiquitin-specific protease 7*, OMIM 602519) também conhecido como *HAUSP* (*Herpes virus-associated ubiquitin-specific protease*). Ele está localizado no braço curto do

cromossomo 16, na região 16p13.3, e codifica a proteína USP7, que atua principalmente na estabilização de MDM2, MDM4 e p53. *USP7* possui 31 éxons, e um transcrito com comprimento total de 4013 pares de bases (Hu *et al.*, 2006).

A proteína USP7 pertence à família das proteases ubiquitina específicas (UBP), um conjunto de enzimas que são capazes de promover a desubiquitinação, reconhecendo e clivando as moléculas de ubiquitina das proteínas, livrando-as assim da degradação. Ela foi identificada pela primeira vez como um fator celular associada ao vírus do herpes, e posteriormente foi descrito que ela atua também estabilizando p53 por meio da ubiquitina ligase E3. USP7 possui 1102 aminoácidos e o seu peso molecular é de aproximadamente 135kDa (Li *et al.*, 2004).

Embora USP7 atue em diversos substratos (p53, MDM2, FOXO4 e MDM4), o mais estudado é a ubiquitina ligase E3 de p53 e MDM2 (Li *et al.*, 2002). Em estudos de super-expressão Li e colaboradores mostraram que p53 pode estar sendo estabilizada através da deubiquitinação promovida por Usp7, sugerindo portanto que Usp7 possa atuar como uma proteína supressora tumoral.

Estas três proteínas, p53, MDM2 e USP7, tem um equilíbrio íntimo e delicado que mantém os níveis adequados da proteína p53. A desubiquitinação da proteína p53 tem sido reconhecida como um passo importante para regular muitos processos celulares, e USP7 tem um papel fundamental nesse processo, antagonizando a degradação da p53 por meio de mecanismos diretos e indiretos (Brooks *et al.*, 2007). Estudos com iRNA tem mostrado que, de fato, a redução de níveis endógenos de USP7 desestabiliza a concentração de p53, fazendo com que aumentem seus níveis celulares. Demonstrou-se também que USP7 interage fisicamente com MDM2 de maneira independente à p53. Ela pode deubiquitar MDM2, *in vivo* e *in vitro*, e este mecanismo é requerido na estabilização de MDM2 em células normais (Li *et al.*, 2004)

Estudos estruturais e bioquímicos revelam que p53 e MDM2 competem pelo mesmo sítio de ligação ao domínio N-terminal de USP7. MDM2 liga-se a USP7 com maior afinidade do que p53, e sendo assim, em células com crescimento normal MDM2 é o substrato preferencial de USP7 (Hu *et al.*, 2006).

No entanto, em células contendo DNA danificado, a fosforilação de MDM2 e MDM4 diminui essa afinidade entre USP7 e p53. Assim, em condições de estresse, p53 se

torna o substrato preferido de USP7, evitando sua degradação, aumentando diretamente os níveis de p53 na célula a fim de permitir uma resposta adequada de p53 frente aos danos no DNA. Indiretamente MDM2 é liberado para ser ubiquitinado e degradado (Meulmeester et al., 2005; Meulmeester et al., 2005b).

Uma importante variante no gene *USP7* consiste na alteração de uma guanina por uma adenina (G>A) no íntron 25 (rs1529916). Kang *et al.* (2009) descreveram a associação do alelo A com a infertilidade em mulheres e, além disso, demonstraram haver fortes pressões de seleção influenciando a frequência desse alelo. A frequência do alelo A é de 33% nos americanos caucasianos, e apenas 16% em americanos descendentes de africanos, e curiosamente, há muitos diferentes haplótipos para o alelo G, no entanto apenas um haplótipo foi encontrado em populações caucasianas.

Um estudo associou este mesmo polimorfismo ao prognóstico de câncer de próstata. O alelo derivado (A) apareceu conferido um risco aumentado de desenvolver um quadro mais grave de câncer da próstata (de acordo com a classificação de risco D'Amico). Portadores do genótipo AA parecem ter chance cerca de 1,5 vez maior de desenvolver câncer de próstata de risco intermediário quando comparado ao genótipo GG. Isto pode ter ocorrido através da influência de USP7 na rede p53 e também através de desbalanço na sinalização hormonal (Sun *et al.*, 2010).

Além disso, similarmente ao MDM4, a quantidade aumentada de USP7 parece predizer um pior prognóstico em carcinoma de pulmão de células escamosas e carcinoma de pulmão de grandes células (Zhao *et al.*, 2014).

Todas essas informações nos sugerem que o alelo G é o alelo ancestral e que, durante os efeitos de gargalo de garrafa que a população humana sofreu, poderia ter sido selecionado em diferentes regiões do globo o alelo derivado, A, já que apresenta distribuições diferentes entre populações humanas. Além disso, podemos sugerir que alterações neste gene afetam diretamente a capacidade da célula de responder a agentes estressores ao DNA, o que pode deixar a célula mais susceptível.

1.7 O gene *LIF*

A *LIF* (*Leukemia inhibitory factor*) é uma citocina pleiotrófica expressa em diversos tipos celulares, sendo que diversas células expressam receptores de *LIF*, entre elas neurônios, adipócitos, hepatócitos, osteoblastos, entre outros (Hilton, 1992). *LIF* é uma glicoproteína e pode ser secretada em uma ampla gama de formas com pesos moleculares distintos, podendo variar desde 38 à 67kDa dependendo do seu grau de glicosilação (Trouillas *et al.*, 2009).

LIF foi descrita pela primeira vez em 1969 por Ichikawa, que descreveu sua atividade biológica como sendo capaz de inibir a proliferação e induzir a diferenciação de macrófagos de uma linhagem de células de leucemia mielóide (Ichikawa, 1969).

Estudos adicionais tem mostrado *LIF* como um fator de diferenciação das células hematopoiéticas, além de diversas outras funções que tem sido sugeridas para essa proteína, entre elas: inibição da diferenciação em células estaminais embrionárias, estimulação da diferenciação e efeitos de sobrevivência em células neuronais, bem como na indução da fase aguda em células hepáticas e outros tipos celulares do sistema imune (Metcalf, 1992).

Diversos artigos tem descrito a regulação de *LIF* via p53, mediando a implantação blastocística sob condições fisiológicas normais, além de variações no gene *LIF* (OMIM 159540) já terem sido descritos associados a alguns tipos de câncer (Lass *et al.*, 2001; Yu *et al.*, 2014).

A principal característica de *LIF* está correlacionada com a potencialização da formação blastocística de embriões humanos através da modulação a diferenciação trofoblástica. No período de implantação do blastocisto, o gene *LIF* é transcrito em maior quantidade nas glândulas endometriais, assim *LIF* é secretado do útero tornando-o receptivo para o blastocisto (Stewart *et al.*, 1992). Por isso recentemente o *LIF* tem sido extensamente estudado e relacionado falhas na implantação do embrião e infertilidade.

É descrito que alelo Pro72 induz a ativação transcricional dos genes alvos de p53, incluindo *LIF*, em comparação ao alelo Arg de p53. Os níveis de expressão de *LIF* são duas vezes maiores nas células com o alelo de Pro72, em comparação com o alelo de

Arg72, o que pode originar a diminuição da taxa de implantação e da fertilidade. Esses dados sugerem fortemente que p53 desempenha uma função importante na implantação blastocística através da regulação do gene LIF (ver revisão em Paskulin *et al.*, 2012).

Kay *et al.* (2006), por exemplo, mostraram que a frequência do alelo Pro72 é significativamente maior em mulheres que tiveram falhas na implantação de embriões após fertilização *in vitro* (FIV) do que naquelas que não apresentaram esse tipo de problema. Hu *et al.*, (2007), por sua vez, observaram que ratos fêmeas que tinham p53 inativado apresentam uma expressão reduzida de LIF no útero, e como esperado, uma redução da taxa de reprodução devida à deficiência na implantação nas fêmeas. Administrando LIF para as ratas deficientes em p53 no 4º dia de gestação aumentou significativamente a taxa de gestação e tamanho da ninhada, consequências diretas da melhoria de implantação do blastocisto. Estes resultados demonstram que a inativação de p53 e diminui os níveis de função de LIF uterino, corroborando, assim, a ideia de que p53 tem papel na reprodução, através da regulação do LIF.

A importância de variantes de LIF na fertilidade humana também foi investigada em um estudo de associação do SNP c.1414T>G localizado na região 3' UTR do gene (rs929271) realizado por Kang *et al.*, (2009). O alelo G deste SNP é encontrado em uma frequência de 30% em americanos de origem europeia e não existe em afro-americanos, o que levou os autores a sugerir que esse alelo surgiu mais tarde na evolução, depois da migração para fora da África.

Além disso, os autores estudaram mulheres com problemas de fertilidade de uma clínica de fertilização *in vitro*, juntamente com um grupo controle, o que demonstrou que o códon 72 de p53 tem um impacto significativo sob os níveis de LIF, o qual poderia por sua vez resultar em diferentes taxas de implantação do embrião e levar a diferentes desfechos na gravidez humana. Os autores observaram que a taxa de implantação é significativamente menor em pacientes jovens, com idade inferior a 35 anos, homocigotos para Pro72 (19%) comparadas com pacientes que possuíam pelo menos um alelo Arg72 (42%).

Mais uma vez é demonstrada a provável ação de seleção positiva em um gene ligado a via do p53, tendo em vista que a reprodução é um processo chave na evolução biológica.

1.8 As populações nativo americanas:

A chegada dos primeiros grupos populacionais à América é um assunto repleto de novas descobertas, hipóteses e divergências que atingem a curiosidade de antropólogos, geneticistas, paleontólogos, arqueólogos dentre outros. Contudo, apesar do desacordo que pode haver entre as diferentes visões, é fascinante pensar sobre a desafiadora tarefa que o *Homo sapiens* moderno teria enfrentado ao se lançar em terras completamente desconhecidas.

Os nativos americanos têm sido alvos de diversos estudos, em diferentes linhas de pesquisa. Há décadas tem se buscado resolver questões que vão desde o povoamento do continente até a busca por explicações para as características particulares encontradas nos povos que passaram a habitar a América.

Resumidamente, pode-se sugerir que a Beríngia foi colonizada durante o último máximo glacial, provavelmente há cerca de 21 mil anos atrás, por uma população vinda da Ásia, e que eles tenham permanecido na região por pelo menos 5 mil anos, tempo suficiente para o surgimento e fixação de variantes genéticas autóctones que podem ser observadas tanto no DNA mitocondrial (Fagundes *et al.*, 2008), como no cromossomo Y (Bortolini *et al.*, 2003) e DNA genômico (Schroeder *et al.*, 2007). Devido a permanência nesta região, houve uma expansão populacional na Beríngia, o que acredita-se que tenha ocorrido concomitantemente com a imigração para o continente americano. A colonização efetiva do continente americano se deu provavelmente pela região costeira do oceano Pacífico, e é descrito na literatura que o fluxo gênico no Círculo Polar Ártico se mantém relativamente constante (González-José *et al.*, 2008; Bortolini *et al.*, 2014).

Frente a migração dessas populações que colonizaram nosso continente, podemos inferir que estes indivíduos tiveram que se adaptar rapidamente ao novo ambiente. De fato, considera-se que muito da variabilidade genética intra e interpopulacional seja resultado dos eventos demográficos, e é esperado que uma fração desta variação observada seja decorrente das pressões seletivas vivenciadas pelos humanos, visto que durante a migração para fora do continente africano encontraram ampla variedade de ambientes (Bamshad e Wooding, 2003; Harris e Meyer, 2006)

No entanto, quando observamos as populações nativo-americanas, os processos microevolutivos causuísticos como a deriva genética e efeito fundador, por exemplo, devem ser considerados, uma vez que causam uma flutuação aleatória na frequência dos alelos. Tais fenômenos se mostram extremamente importante nas populações ameríndias devido ao seu pequeno tamanho populacional, além do considerável isolamento genético por motivos geográficos e sócio-culturais (Salzano, 2009). No entanto, diversos estudos têm conseguido demonstrar sinais de seleção positiva em povos nativos americanos (Hünemeier *et al.*, 2012; Foll *et al.*, 2014; Ko *et al.*, 2014).

1.9 Adaptações genéticas mediadas pelo clima:

Desde a migração do *Homo sapiens* para fora do continente africano, há cerca de 50-100 mil anos atrás, os humanos modernos tem colonizado uma ampla gama de ambientes. Adaptações comportamentais e biológicas têm contribuído para o sucesso desta extraordinária trajetória. Vários exemplos de adaptações para uma vida bem sucedida tanto dentro quanto fora da África tem sido descritos (Meier *et al.*, 1991; Tobias *et al.*, 1994; Bersaglieri *et al.*, 2004; Potts, 2005; Hancock *et al.*, 2008 e Hünemeier *et al.*, 2012), dentre os quais cita-se: a capacidade das populações humanas de sobreviver ao frio extremo, grande amplitude térmica, clima árido, alto índice de radiação ultravioleta, baixa concentração de oxigênio dentre outras (Beall *et al.*, 2007; Foll *et al.*, 2014; Rademaker *et al.*, 2014).

Mesmo quando consideramos somente o continente americano, podemos observar uma extensa diversidade de ecossistemas, e, conseqüentemente uma população muito diversificada, que habita igualmente esses diversos ambientes com características geológicas e ecológicas distintas, como o clima extremamente frio, tropical e subtropical que caracteriza o continente, juntamente com variações de altitude e índices de umidade extremos; o deserto do Atacama, no Chile, por exemplo, é um dos lugares mais secos do mundo (Salzano & Sans, 2014).

As populações do altiplano andino sul americano, por exemplo, vivem sob condições climáticas extremas, com altitude média acima de 2500 metros, e temperaturas

baixas na maior parte do ano, com média de temperatura anual de 15°C. Apesar disto, essa região está ocupada há aproximadamente 12.800 anos antes do presente (Rademaker *et al.*, 2014). Para que isto seja possível, diversas características têm sido associadas ao fenótipo adaptado à altitude Andina, dentre elas, observam-se: aumento da capacidade pulmonar, relativa tolerância a hipoxia, e quantidade maior de glóbulos vermelhos. Entretanto, a elucidação completa da natureza dos processos biológicos relacionados a adaptação individual e populacional para a aquisição de tais fenótipos ainda permanece desconhecido.

Os padrões de diversidade genética encontrados nas populações humanas constituem uma base importante para muitas áreas de pesquisa em genética humana. Mais notavelmente, eles fornecem uma fonte inestimável de dados para inferências sobre nossa história evolutiva. Além disso, tais variações genéticas auxiliam na interpretação de estudos que buscam genes e suas variantes que conferem maior suscetibilidade a doenças.

Exemplos de seleção positiva em populações humanas vêm sido descritas em genes da via de p53: Shi *et al.*, (2009) genotiparam o SNP309 de *MDM2* e Pro72Arg de *TP53* em 67 populações do leste asiático em uma ampla gama de condições ambientais. Eles concluíram que o alelo G de *TP53*, que codifica Arg, foi positivamente associado com baixas temperaturas no inverno, e o alelo G do SNP309 de *MDM2* correlacionou-se negativamente com os índices de radiação ultravioleta; além disso, eles propuseram que as relações observadas entre as frequências alélicas e variáveis ambientais refletiam a ação de diferentes pressões seletivas. Sucheston *et al.*, (2011) genotiparam esses mesmos SNPs em 971 indivíduos de 52 populações de todo o mundo. No entanto, eles não observaram correlações significativas das variantes polimórfica em estudo com as diferentes latitudes e com variáveis ambientais.

Os resultados divergentes destes trabalhos destacam a importância de avaliar a evidência de assinaturas de seleção no contexto evolutivo dos diferentes grupos de populações humanas, resultado especialmente interessante se avaliarmos a variabilidade dentro do continente americano. Este território foi provavelmente um dos últimos a ser ocupado pelo ser humano, mantendo-se isolado do resto do mundo desde o fim da última era glacial até o século XV.

Dadas estas características históricas únicas, as populações nativas americanas apresentam-se como especialmente instigantes para estudos de variabilidade genética.

Como antecedentes, podemos citar Gaspar *et al.*, (2002), que investigaram vários SNPs, incluindo SNP Pro72Arg de *TP53*, em populações nativo-americanas, sendo que o alelo derivado (G) foi descrito em frequências muito altas. Os autores não conseguiram distinguir, entretanto, se estas altas frequências poderiam ser interpretadas como resultado da ação da deriva genética ou da seleção natural.

Considerando o exposto acima e nos itens anteriores podemos nos perguntar se há diferenças entre as frequências alélicas dos polimorfismos da via de p53 entre os nativos americanos que vivem em diferentes regiões geográficas, incluindo ambientes extremos como o altiplano andino e a floresta tropical, e também quais seriam as razões para a existência e manutenção de tal padrão.

CAPÍTULO 2

OBJETIVOS

2.1 Objetivo Geral

Investigar a presença de variantes alélicas de genes relacionados à via de *TP53* em populações ameríndias sul americanas, oriundas de ecossistemas distintos, para inferir se a variação encontrada nos *loci* sob investigação pode ser explicada pela ação da seleção natural, correlacionando também com diferentes variáveis climáticas.

2.2 Objetivos Específicos

- ✓ Estimar a variabilidade dos genes *TP53*, *MDM2*, *MDM4*, *USP7* e *LIF* através da investigação dos SNPs: rs1042522 (*TP53*), rs929271 (*LIF*), rs2279744 (*MDM2*), rs1563828 (*MDM4*), rs15229916 (*USP7*);
- ✓ Verificar se a distribuição dos alelos desvia-se do esperado considerando a teoria neutra de evolução molecular;
- ✓ Correlacionar diferenças nas frequências genotípicas de diferentes populações com as variáveis geográficas, ambientais e climáticas.
- ✓ Montar um cenário evolutivo que explique os achados considerando a rede envolvendo os genes investigados.

CAPÍTULO 3

Genetic variation in the TP53 pathway in Native Americans: signals of climate adaptation

Genetic variation in the TP53 pathway in Native Americans: signals of climate adaptation

Vanessa C. Jacovas¹; Orlando Pérez²; Soledad de Azevedo²; Gabriel de S. Macedo¹; José Raul Sandoval³; Alberto Salazar-Granara³; Mercedes Villena⁴, Jean-Michel Dugoujon⁵; Rafael Bisso-Machado¹; Maria Luiza Petzl-Erler⁶; Francisco Mauro Salzano¹; Patrícia Ashton-Prolla¹; Virginia Ramallo²; Maria Cátira Bortolini^{1*}.

1Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brasil;

2Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Puerto Madryn, Argentina;

3Facultad de Medicina Humana, Universidad de San Martín de Porres (USMP) Lima, Peru

4Instituto Boliviano de Biología de Altura (IBBA), Universidad Mayor de San Andrés, La Paz, Bolivia

5Anthropologie Moléculaire et Imagerie de Synthèse, CNRS UMR 5288, Université Paul Sabatier Toulouse III, Toulouse 31000, France.

6Laboratório de Genética Molecular, Departamento de Genética, Universidade Federal do Paraná, Curitiba, PR, Brasil

Manuscrito a ser submetido a revista de circulação internacional.

***Corresponding author:**

Maria Cátira Bortolini, PhD

Departamento de Genética, UFRGS - Campus do Vale

Prédio 43323 - Sala 126 - CEP: 91501-970

Av. Bento Gonçalves, 9500. Porto Alegre, RS, Brasil;

e-mail: maria.bortolini@ufrgs.br

ABSTRACT

A total of 282 individuals (12 populations) characterized as Native or as having a high level (>90%) of Native American ancestry were studied in relation to five single nucleotide polymorphisms of genes of the *TP53* pathway (*TP53*, rs1042522; *MDM2*, rs2279744; *MDM4*, rs1563828; *USP7*, rs1529916; and *LIF*, rs929271). In addition, data from the literature was obtained for 100 persons from five other Native American populations. The samples were classified as living in high altitude ($\geq 2,500$ m) or in lowlands ($< 2,500$ m). Alleles *USP7-G*, *LIF-T*, and *MDM2-309-T* showed significant indications that they may have a protective role in relation to harsh environmental variables related to high altitude. These data provide additional information about a key physiological process responsible for the successful human colonization of the Andes.

Key words: Native Americans, High altitude adaptation, TP53 pathway genes.

Introduction

The product of the *TP53* gene is a transcription factor (p53) that activates or represses a large number of target genes, regulating many cellular functions of extreme importance, such as cell cycle, metabolism, DNA repair, senescence, and apoptosis. This factor is therefore essential for maintaining the genome integrity (Botcheva, 2014). In humans, p53 has 393 amino acids and *TP53* is located in the short arm of chromosome 17 (Linzer and Levine, 1979).

The amount of p53 in cells is primarily determined by the rate in which it is degraded, rather than the rate in which it is produced. The factor is constitutively expressed in all cell types, but does not accumulate in nonstressed cells, since it is maintained at baseline levels due to its rapid proteasomal degradation via ubiquitination (Nag *et al.*, 2013; Chao, 2015). On the other hand, the p53 level increases in response to various stress signals such as UV radiation, low oxygen concentrations (hypoxia) and temperature exposures (Ljungman, 2000; Latonen *et al.*, 2001; Sermeus and Michiels, 2011; Chen, 2012).

There are many polymorphisms described for *TP53*, but a C→G non-synonymous substitution (rs1042522: c.215C>G, p. Pro72Arg; Matlashewski *et al.*, 1987) that leads to the Pro→Arg amino acid change at codon 72 is the most widely studied. This polymorphism is important in the susceptibility to develop cancer, since the ancestral *p53-Pro* allele is less active than derived *p53-Arg* to induce apoptosis, among other characteristics (Dumont *et al.*, 2003; Whibley *et al.*, 2009).

For proper p53 transcription activity other genes must also function properly, among these *MDM2* (*Mouse double minute 2 homolog*, OMIM 164785), *MDM4* (*Mouse double minute 4 homolog*, OMIM 602704) and *USP7* (*Ubiquitin-specific protease 7*,

OMIM 602519) also known as *HAUSP* (*Herpesvirus-associated ubiquitin-specific protease*). Their products are critical for p53 function. The *LIF* (*Leukemia-inhibitory factor*, OMIM 159540) is regulated by p53. These five genes constitute the classical *TP53* network (Kang *et al.*, 2009).

MDM2 is an E3 ubiquitin-protein ligase, a cellular inhibitor of p53 (Momand *et al.*, 1992; Marine *et al.*, 2006; Chao, 2015). It ubiquitylates p53 and leads to its proteasomal degradation. MDM2 is also a transcriptional target for p53, and that tightly regulates it. This auto-regulatory loop ensures the precise regulation of both p53 and MDM2 protein levels, keeping low levels of p53 in unstressed cells (Kohn and Pommier, 2005; Atwal *et al.*, 2007; Nag *et al.*, 2013).

The most studied polymorphism in the *MDM2* gene (rs2279744: c.14+309T>G) is located in its internal promoter. A single-nucleotide change from T→G increases the affinity of a *MDM2* sequence to the *Sp1* transcription factor (*Sp1*; OMIM 189906). Thus, homozygotes for the G derived allele express more MDM2 than homozygotes TT (Bond *et al.*, 2004; Alazzouzi *et al.*, 2007). With higher amounts of MDM2 there is a decrease in p53, reducing the response of the cell to stress. This condition leads to problems which appear are related to DNA repair, decreased apoptosis, and senescence (Bond *et al.*, 2004). Some studies have also demonstrated that these alleles have different distributions in human populations (Millikan *et al.*, 2006; Atwal *et al.*, 2007). For instance, derived allele *MDM2-309-G* has higher frequency in European/Asian than African populations (average values: ~0.35/~0.70 and ~0.03, respectively; Suscheston *et al.*, 2011; Shi *et al.*, 2009; 1000 Genome Project: <http://www.1000genomes.org/>) and it could compensate for higher apoptotic frequencies brought by derived allele p53-72Arg in Eurasians (~0.56; Suscheston *et al.*, 2011; Shi *et al.*, 2009; 1000 Genome Project: <http://www.1000genomes.org/>), suggesting an adaptive condition (Atwal *et al.*, 2007).

The higher amounts of MDM2, associated to GG genotype, lead to lower levels of p53 and impaired response of the cellular stress, such as DNA repair defective, decreased induction of apoptosis and senescence (Bond *et al.*, 2004).

The MDM4 protein encoded by the *MDM4* gene, acts as a negative binding regulator of p53, inhibiting its transcriptional activation (Shvarts *et al.*, 1996, 1997; Liu *et al.*, 2013). MDM2 and MDM4 form heterodimers that have high capacity of ubiquitination, thus causing the degradation of p53 (Pei *et al.*, 2012). Deletion of either *MDM2* or *MDM4* induces p53-dependent early embryonic lethality in an animal model (Xiong, 2013). The AA genotype for the single-nucleotide *MDM4* polymorphism (rs1563828:g.204547449A>G) was associated with increased risk for breast cancer (Song *et al.*, 2012).

Another important regulator of *TP53* is the USP7 protein, encoded by the gene *USP7*, which deubiquitilates p53 and protects it from proteasome degradation (Shan *et al.*, 2008). The *USP7* gene has a G → A transition in intron 25 (rs1529916: g.8897333G>A), and derived allele A has been associated with endometriosis, female infertility, and prostate cancer (Kang *et al.*, 2009; Sun *et al.*, 2010).

Finally, LIF is a cytokine expressed in different cell types, and its main function is to strengthen the blastocystic training of human embryos. In the blastocyst implantation period, the *LIF* gene is transcribed in high quantities in the endometrium, making the uterus receptive to the blastocyst implantation. Evidence show that allele G (T→G transversion at the 3' UTR region of the gene; rs929271: g.30242237T>G) is associated with female infertility (Kang *et al.*, 2009). The literature also provides information that the LIF expression level is 2x lower in cells bearing the *p53-72Pro* allele as compared to *p53-72Arg*, which can lead to the decrease of the implantation and fertility rate.

In summary, several studies have strongly suggested that p53 signaling pathway polymorphisms play an important role in blastocystic implantation and recurrent pregnancy loss (Paskulin *et al.*, 2012; Fraga *et al.*, 2014).

The genetic variability observed in contemporary human populations, as well as the function changes associated with the above-mentioned polymorphisms, allow to infer that a simple neutral model of mutation and drift is insufficient to explain the pattern found. Thus, positive selection associated to adaptation of *Homo sapiens* through different ecosystems has been suggested. For example, the *p53-72Arg* allele (rs1042522) is more common in Europeans than in Africans, and a hypothesis was made that its distribution is dependent on latitude and is maintained by natural selection (Beckman *et al.*, 1994; Sjölander *et al.*, 1996).

Shi *et al.* (2009), on the other hand, showed that winter temperature and UV radiation were significantly correlated with the *TP53* (rs1042522) and *MDM2* (rs2279744) allele distributions in East Asian populations, indicating possible adaptation to distinct environments.

America was the last continent occupied by humans at pre-colonial times. González-José *et al.* (2008) and Bortolini *et al.* (2014) suggested that an initial major dispersal began after 21,000 years before present and that the biological and cultural characteristics of the first Americans that emerged during a standstill period in Beringia were reshaped by recurrent trans-Beringian/circum-Arctic gene flow and important local population dynamics. For instance, Native Americans have experienced dramatic episodes of genetic drift, and successive bottleneck events during the migration across the continent. Furthermore, signals of positive natural selection associated to autochthonous environmental and cultural conditions have also been described (Hünemeier *et al.*, 2012a, b; Foll *et al.*, 2014).

Based on the findings discussed above, our hypothesis is that the allele distributions of the classical *TP53* pathway genes in Native American populations may be reflecting adaptation, besides demographic and/or random events. To test this hypothesis, we genotyped the five above-mentioned SNPs in 282 unrelated individuals and compared the results to a large number of climate-related environmental variables, such as altitude, temperature and seasonal mean UV radiation. Additional data regarding two of these SNPs (*TP53*-rs1042522 and *MDM2*-rs2279744) were compiled from the literature for more extensive population analysis.

Subjects and Methods

Samples and ethical procedures

Five SNPs (rs929271, rs1042522, rs1563828, rs2279744, rs1529916) were genotyped in 282 volunteers characterized as Native or as having large (>90%; Sandoval *et al.*, 2013) Native American ancestry. The populations where they live are located in different ecoregions, namely: (a) high altitudes in the Andes, Aymara (N=18) and Quechua (N=17) speakers from Bolivia; Quechua or Aymara speakers from Peru: Chivay (N=18), Cabanaconde (N=17), Yanke (N=10); Taquile (N=43), Amantani (N=29), Anapia (N=15), and Uros (N=22) all located in the Andean area, on Lake Titicaca islands or in its vicinity; and (b) lowland populations, a Native Amazonian population living in Andoas (N=61), North Peru, and Guaraní Amerindians from Brazil (Tupian speakers from two partialities: Ñandeva, N=16; and Kaiowa, N=16). Details about these populations can be found in Tsuneto *et al.* (2003); Marrero *et al.* (2007) Gayà-Vidal *et al.* (2010); and Sandoval *et al.*

(2013). To facilitate the presentation of the results and discussion, from here on all communities will be referred as Native American.

Ethical approval for the use of these samples was provided by the National Ethics Committee (Resolution No. 123/98 CONEP), for the Brazilian tribes; and by the Ethics Committee of Universidad San Martín de Porres, Lima, Peru (Peruvian samples) and Instituto Boliviano de Biología de Altura, La Paz, Bolivia (Bolivian samples). Written informed consent, or, verbal informed consent (illiterate persons) were obtained individually from tribal participants, and they were obtained according to the Declaration of Helsinki. Ethics committees approved these procedures, besides the use of these samples for evolutionary and anthropological studies.

We also included additional data from the literature, in a total of 100 individuals from five Amerindian populations: Surui and Karitiana (Brazil), Piapoco (Colombia), Maya and Pima (Mexico) (for more details of these samples, see <http://www.cephb.fr/HGDP-CEPH-Panel/>).

Data collection

Genomic DNA was obtained from saliva, whole blood or plasma, using the QIAamp DNA extraction MiniKit (Qiagen) in accordance to the manufacturer's instructions. Genotyping of *TP53*-rs1042522, *MDM4*-rs1563828, *USP7*-rs1529916, *LIF*-rs929271 and *MDM2*-rs2279744 was performed by the allelic discrimination method using the TaqMan Genotyping Assay (Applied Biosystems, USA), assay numbers: C__2403545_10 (*TP53*), C__9493064_10 (*MDM4*), C__9688119_1_ (*USP7*), and C__7545904_10 (*LIF*). The *MDM2* rs2279744 genotyping was obtained by a customized (assay-by-design) assay using probes FAM-TCCCGCGCCGCAG and VIC-

CTCCCGCGCCGAAG; forward primer, 5'- CGGGAGTTCAGGGTAAAGGT-3', and reverse primer 5'-ACAGGCACCTGCGATCATC-3'.

PCR reactions were conducted in 48-well plates, each reaction containing 10ng of genomic DNA, 2x TaqMan® genotyping MasterMix (Applied Biosystems, USA), specific probes for each SNP (40x), and ultra pure water for a final reaction volume of 10 µl. The PCR conditions were as follows: 10 minutes at 95°C, followed by 45 cycles of 75 seconds each (95°C for 15 s and 63°C for 60 s). *MDM2* rs2279744 genotyping was also conducted in 48-well plates, each reaction containing: 10 ng of genomic DNA, 2x TaqMan® genotyping MasterMix (Applied Biosystems, USA), 5µM of each primer and probe, and water to reach a final volume of 10 µl. *MDM2* PCR conditions were as follows: 2 min at 50 °C, 10 min at 95°C, followed by 45 cycles of 75 s (95°C for 15 s and 60°C for 60 s). All reactions were performed in an Illumina Eco Real-Time PCR System, and the results were analyzed using a v5.0 software associated to that system.

All laboratory procedures were performed at the Laboratory of Human and Molecular Evolution, Federal University of Rio Grande do Sul.

Analysis of molecular variance (AMOVA; Weir *et al.*, 1984; Weir, 1996; Excoffier *et al.*, 1992) was performed using Arlequin 3.5.1.227. A *p* value equal to or less than 0.05 was considered statistically significant for these tests.

Allele frequencies were tested for possible correlations with three sets of environmental conditions: 1) geographic: altitude, latitude, and longitude; 2) annual and seasonal mean UV radiation, and 3) Nineteen climate-related variables [annual mean temperature, mean diurnal range, isothermality (the ratio of mean diurnal range to temperature annual range), temperature seasonality (standard deviation x 100 of average annual daily temperatures), maximum temperature of warmest month, minimum temperature of coldest month, temperature annual range, mean temperature of wettest

quarter, mean temperature of driest quarter, mean temperature of warmest quarter, mean temperature of coldest quarter, annual precipitation, precipitation of wettest month, precipitation in the driest month, precipitation seasonality (coefficient of variation of average annual daily precipitations), precipitation of wettest quarter, precipitation of driest quarter, precipitation of warmest quarter and precipitation of coldest quarter]. Statistical analyses were performed using population data rather than those of the 382 unrelated individuals of our sample, since it is a more conservative approach. Shi *et al.* (2009) used the same approach to investigate the correlation of *TP53* (rs1042522) and *MDM2* (rs2279744) allele distributions with environmental conditions in East Asian populations.

Environmental data for each population (Table S1) were collected from The SoDa Service and WorldClim, (<http://www.soda-is.com/> and <http://www.worldclim.org/>, respectively; last access: December 19, 2014).

A principal component analysis (PCA) was performed to convert the nineteen possibly correlated bioclimatic variables into a set of linearly uncorrelated values.

To evaluate the relationship between SNP allele frequencies and the three data sets of environmental conditions, Spearman's rho correlation coefficient, with two tailed tests for *p*-values (< 0.05) was performed using SPSS v20.0.

Determining the association between genetic polymorphisms (SNPs) and environmental/geographical variables is not a trivial task. That is partly difficult because the distribution of one SNP affecting one trait can be dependent on the effect of another SNP. Traditional statistical methods are not suited to investigate the gene-gene interaction, when it involves more than two genes and moderate and small population sample sizes. Thus, in the present study we performed Logistic regression (SPSS v20.0) and Multifactor Dimensionality Reduction (MDR, v3.0.2; Moore *et al.*, 2006) analyses were carried out to detect environmental conditions with main effects, as well as potential gene-

gene interactions. MDR is a non-parametric genetic model (it assumes no particular inheritance pattern), which combine multi-locus genotypes to create a single variable and detect gene-gene (epistasis) and/or gene-environment interactions (Ritchie *et al.*, 2001). For these two last analyses, individual genotypic data is required.

Multiple regression and MDR analysis using two altitude categories were performed ($\geq 2,500$ meters and $<2,500$; Moore, 2001).

Results

Table 1 shows the derived allele frequency for each SNP investigated. Wide variation was observed in some allele frequencies. For instance, the frequency of derived *MDM2-309-G* is about five times higher in the Guaraní Ñandeva than in the Guaraní Kaiowa, probably reflecting the action of genetic drift since the split of these two Guaraní partialities which occurred less than two thousand years ago (Marrero *et al.*, 2007).

Environmental conditions for all populations (present study and literature) were compiled in Table S1. Unfortunately, for the HGDP-CEPH-panel samples, only data from the *TP53* (rs1042522) and *MDM2* (rs2279744) polymorphisms were available (Sucheston *et al.*, 2011). Thus, we performed our analysis considering two sets of data: (a) First, with the five SNPs (rs1042522, rs2279744, rs1529916, rs1563828, and rs929271) and the twelve South American populations for which original data was obtained for the present study; and (b) Second, including all above-mentioned populations plus the other five for which *TP53* rs1042522 and *MDM2* rs2279744 data were available (Sucheston *et al.*, 2011).

The AMOVA analysis, using both data sets (Table 1), indicates that homogeneity and population structure can be seen in both categories highland and lowland. For

instance, population structure measured by *Fst* statistics (*i.e.* the among-group component of genetic variance) is observed in the two groups considering *TP53* rs1042522 (*Fst* = 0.068 and 0.054, for highland and lowland, respectively), while no significant difference among the allele distributions was detected for *MDM4* rs1563828 locus. Only the *Fst* value observed for *LIF* rs929271 in highland group (11.8%) is similar to the average estimated across the genome (12%; Altshuler *et al.*, 2005).

When the analyses were performed using the first data set, the first principal component (PC1) accounted for 73% of the total variance, comprising the bioclimatic variables Bio 1, Bio 2, and Bio 5 to Bio 19.

Spearman's rho correlation tests performed among the PC components, geographic and UV data showed that *USP7* allele *G* (rs1529916) has a strong and significant correlation with PC1 ($\rho = -0.74$, $p = 0.01$), with annual mean of ultraviolet irradiance ($\rho = 0.76$ $p = 0.01$) and winter mean of ultraviolet irradiance ($\rho = 0.68$ $p = 0.04$). Allele *LIF T* has also a strong significant correlation with annual mean of ultraviolet irradiance ($\rho = 0.69$, $p = 0.03$) and with PC1 ($\rho = -0.66$, $p = 0.03$). Allele ancestral *MDM2 T* has a strong correlation ($\rho = 0.60$) with winter mean of ultraviolet irradiance, but after the Bonferroni correction the significance was lost. No significant correlation between the *TP53* and *MDM4* alleles and environmental conditions was detected with the first data set.

When we expanded our analysis to a larger number of populations (second data set), PC1 now represents 59% of total variance comprising the sixteen bioclimatic variables Bio 1, Bio 2, Bio 5, Bio 6, and Bio 8 to Bio 19. The Second Principal Component (PC2) represents 23% of the variance, and comprise the remaining Bio 3, Bio 4, and Bio 7, all connected with climatic changes by seasonality.

Spearman's rho correlation tests performed among the PC components and environmental conditions show that the ancestral *MDM2-309-T* allele has a very strong and

significant correlation with winter mean UV irradiance ($\rho = 0.82$, $p = 0.00013$), a strong correlation with altitude ($\rho = 0.67$, $p = 0.009$), and strong and moderate correlations with PC1 ($\rho = -0.61$, $p = 0.01$) and PC2 ($\rho = -0.56$, $p = 0.03$).

Additionally, we performed a bivariate logistic regression analysis, to search for possible associations between the *MDM2* and *TP53* genotypes and the two altitude categories ($\geq 2,500$ meters and $<2,500$). The *MDM2*-TT genotype was found to be 6 times more frequent in the highlands ($p < 0,001$). A confounding factor could be admixture with Europeans, which is more important in Andean than in the lowland populations considered here (Wang *et al.*, 2007; Sandoval *et al.*, 2013; Verdu *et al.*, 2014). However, any effect of admixture would be in the opposite direction, since *MDM2*-G frequencies are relatively high in Spaniards (0.39; 1000 Genome Project: <http://www.1000genomes.org/>).

TP53 is not associated with altitude when considered in isolation and no main effect is detected in the MDR analysis (Figures S1A and B). However, *MDM2*-309-TT has a clear major implication in the adaptation to high altitudes. Note that the largest number of people living in the highlands are carriers of the *MDM2*-309-TT genotype (Figure S2). Additionally, a strong synergistic interaction between it and *TP53* genotypes was detected (Figure S1-A and B). These data indicate that although *MDM2*-TT frequency may reflect the effect of selective pressure, its magnitude is modified by the status of the *TP53* genotypes.

Discussion

More than sixty thousand scientific studies have been published in the last 30 years on the role of the *TP53* network genes, as well as of their variants, in the human susceptibility to cancer and other pathological conditions. This simple observation contrasts with the rarity of studies considering their evolutionary context, indispensable to

explain their distributions in human populations, which often cannot be simply understood as a result of stochastic processes. Our goal in the present paper was to help fill this gap, providing information about five polymorphisms of the classical *TP53* network in Native American and associated populations.

Our analysis with data set (a), which included original information of 5 SNPs in 12 populations, suggest the well-known role of genetic drift in those groups, illustrated by the wide difference of derived *MDM2-309-G* frequencies between the two Guaraní partialities.

However, other instigating results could be associated to Native American adaptation to environmental conditions. Alleles *USP7-G* (rs1529916) and *LIF-T* (rs929271) were strongly correlated with ultraviolet irradiance, as well as temperature and precipitation variables comprising the First Principal Component, PC1. Checking the variables with the highest representation in the PC1 components (> .90) it is possible to see that in regions where the annual mean temperature, minimum temperature of coldest month, mean temperature of driest quarter, mean temperature of coldest quarter, and precipitation are low, the presence of ancestral G and T alleles are significantly higher. In other words, ancestral alleles *USP7-G* and *LIF-T* are more present in environments of low temperature, arid climate, wide temperature range during the day, and high UV radiation, typical of high altitudes. It should be noted that the derived alleles of these SNPs (*USP7-A* and *LIF-G*) have been related to higher cancer susceptibility, infertility, and endometriosis (Kang *et al.*, 2009; Sun *et al.*, 2010), so that alleles *USP7-G* and *LIF-T* could be considered as protecting against both the consequences of harsh environmental stresses and the occurrence of these pathological conditions.

Human populations living at high altitude are likely to have developed specific adaptations to support not only these harsh conditions described above, but also low oxygen concentrations (Foll *et al.*, 2014). Monge (1948) proposed that hypoxia could

reduce fertility in humans; however, more recent studies (see Vitzthum, 2013) have shown that the reproductive functioning of high altitude natives is adapted to the hypoxia and other extreme conditions. Our results with *USP7* (rs1529916) and *LIF* (rs929271) polymorphisms could be connected with the adaptation of the successful reproduction of the Andes natives.

When the second set of data (b) was used, ancestral allele *MDM2-309-T* was shown to be strongly correlated with winter mean UV radiation, altitude and PC1. The highest factors influencing PC1 (> 0.90) are annual mean temperature, minimum temperature of coldest month, minimum temperature of coldest quarter, and annual precipitation. Allele *T* is significantly more present in communities located at high altitudes with extreme environmental conditions such as high UV radiation and dry and cold climate. In addition, the bivariate logistic regression analysis showed that the chance of finding *MDM2-309-TT* individuals living in the highlands is six times greater than in lowlands. *MDM2-309-TT* homozygotes express a normal quantity of MDM2, keeping an adequate level of p53 (Alazzouzi *et al.*, 2007).

Our findings are in agreement with those of Shi *et al.* (2009), since they showed that low levels of UV are significantly correlated with *MDM2-GG* in Han Chinese. These authors suggested that *MDM2-GG* is selected in areas of low UV activity (at high altitudes, the thinner atmosphere will filter less UV radiation, consequently for every 1000 m. increase in altitude, the UV radiation level will increase ~12%; ~~Hong Kong Observatory,~~ available from http://www.weather.gov.hk/radiation/tidbit/201012/uv_e.htm).

As already mentioned, Andean natives have successfully adapted to environments with low oxygen concentration. One gene recruited for adaptation to hypoxia is *EPAS1* [Endothelial PAS domain-containing protein 1, also known as *HIF-2 α* , Hypoxia-inducible factor - alpha 2 (OMIM 603349)], which acts by preventing the toxicity promoted by this

condition. This gene plays an important role in the classical, as well as expanded *TP53* network. For instance, an alpha subunit of this protein regulates p53 activity, including its stability against damage, and its nuclear export through interactions with MDM2 (Chen *et al.*, 2003). Foll *et al.*, (2014) confirmed the presence of a signal of positive selection for *EPAS1* in both Tibetans and Andeans. Furthermore, several studies have revealed the role of p53 and its regulates in physiological and metabolic process resulting of the low oxygen concentration (Alarcón *et al.*, 1999; Vousden and Ryan, 2009; Sermeus and Michiels, 2011). Recently, Eichstaedt *et al.* (2014) studied a native population living in the Argentinean Andes (Colla) and identified signatures of positive selection in genes involved in cellular hypoxia, including *TP53*. Noteworthy, that hypoxia induces p53 accumulation through MDM2 down-regulation (Alarcón *et al.*, 1999).

As whole, these results reinforce our suggestion that the individuals with *MDM2-309-TT* genotype present an appropriate cell responses to the environmental stresses to which populations of high altitude are submitted.

It is known that, under neutral model conditions, South Amerindians living in lowlands present higher levels of population structure when compared to those seen in Andean native communities (Tarazona-Santos *et al.*, 2001; Wang *et al.*, 2007). However, not all *Fst* values calculated here are as expected (Table 1). Positive selection perturbs patterns of genetic variation relative to what is expected under a standard neutral model (Biswas and Akey, 2006). Additionally, it is possible to see that some derived alleles such as *MDM2-309-G* has high frequencies in putative ancestral Asian populations (0.57-0.82; Suscheston *et al.*, 2011; Shi *et al.*, 2009; 1000 Genome Project: <http://www.1000genomes.org/>), but in Andeans its distribution is surprisingly low (average value: ~0.13). An excess of low and/or high frequency of derived alleles can also be considered a signal of positive selection (Biswas and Akey, 2006).

In summary, our results are showing that frequencies of functional polymorphisms of the classical *TP53* network such as *USP7-G* (rs1529916), *LIF-T* (rs929271), and *MDM2-309*, as well as to the synergistic interaction among them and *TP53* common alleles are correlated with responses to the high altitude biological stress, at least in Andean populations. The distribution of the classical *TP53* pathway alleles in Native American populations could therefore be under selective pressure. Sucheston *et al.* (2011) investigated 52 worldwide populations from the HGDP-CEPH-panel for the *MDM2-309* polymorphisms and found no significant association with climate variables. However, the Native American sample investigated by them was more restricted than those of the present study, and this may be the cause for the divergent results.

Finally, government surveys in Peru indicate that the rate of gestational and postpartum complications in Aymara regions is lower than the national average (1.8% and 5% respectively; http://www.dge.gob.pe/publicaciones/pub_asis/asis26.pdf: page 165; <http://www.dge.gob.pe/portal/docs/intsan/asis2012.pdf>: page 76). This same official source indicates also some instigating differences in the cancer incidences between lowland localities and some regions situated at high altitude (for example in the Puno state, where the Anapia community is located; http://www.dge.gob.pe/portal/docs/asis_cancer.pdfhttp://www.dge.gob.pe/portal/docs/asis_cancer.pdf: page 64). However, only future studies can accurately relate our evolutionary findings with those related to the health of the contemporary Andean populations.

Acknowledgments: We are very grateful to the individuals who donated the samples analyzed here and to the Fundação Nacional do Índio (Brazil) for logistic support. We thank René Vasquez for his assistance in the sample collection in Bolivia and Sidia Maria Callegari Jacques for suggestions in an early version of this manuscript.

References

Alarcón R, Koumenis C, Geyer RK, Maki CG, Giaccia AJ (1999) Hypoxia induces p53 accumulation through MDM2 down-regulation and inhibition of E6-mediated degradation. *Cancer Res* 59:6046-51.

Alazzouzi H, Suriano G, Guerra A, Plaja A, Espín E, Armengol M, Alhopuro P, Velho S, Shinomura Y, González-Aguilera JJ, Yamamoto H, Aaltonen LA, Moreno V, Capellà G, Peinado MA, Seruca R, Arango D, Schwartz S Jr (2007) Tumour selection advantage of non-dominant negative P53 mutations in homozygotic MDM2-SNP309 colorectal cancer cells. *J Med Genet* 44:75-80.

Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P, International HapMap Consortium (2005) A haplotype map of the human genome. *Nature*. 437:1299–1320.

Atwal GS, Bond GL, Metsuyanin S, Papa M, Friedman E, Distelman-Menachem T, Ben Asher E, Lancet D, Ross DA, Sninsky J, White TJ, Levine AJ, Yarden R (2007) Haplotype structure and selection of the MDM2 oncogene in humans. *Proc Natl Acad Sci USA* 104:4524-9.

Beckman G, Birgander R, Sjölander A, Saha N, Holmberg PA, Kivelä A, Beckman L (1994) Is p53 polymorphism maintained by natural selection? *Hum Hered* 44:266-70.

Biswas S, Akey JM (2006) Genomic insights into positive selection. *Trends Genet.* 22:437-446.

Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P, Onel K, Yip L, Hwang SJ, Strong LC, Lozano G, Levine AJ (2004) A single nucleotide polymorphism in the MDM2 promoter attenuates the p53 tumor suppressor pathway and accelerates tumor formation in humans. *Cell* 119:591-602.

Bortolini MC, González-José R, Bonatto SL, Santos FR (2014) Reconciling pre-Columbian settlement hypotheses requires integrative, multidisciplinary, and model-bound approaches. *Proc Natl Acad Sci USA* 111:E213-4.

Botcheva K (2014) p53 binding to human genome: crowd control navigation in chromatin context. *Front Genet* 5:447.

Chao CC (2015) Mechanisms of p53 degradation. *Clin Chim Acta* 438:139-47.

Chen J (2012) The roles of MDM2 and MDMX phosphorylation in stress signaling to p53. *Genes Cancer* 3:274-82.

Chen D, Li M, Luo J, Gu W (2003) Direct interactions between HIF-1 alpha and Mdm2 modulate p53 function. *J Biol Chem* 278:13595-8.

Dumont P, Leu JI, Della Pietra AC 3rd, George DL, Murphy M (2003) The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet* 33:357-65.

Eichstaedt CA, Antão T, Pagani L, Cardona A, Kivisild T, Mormina M (2014) The Andean adaptive toolkit to counteract high altitude maladaptation: genome-wide and phenotypic analysis of the Collas. *PLoS One* 9(3):e93314.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Foll M, Gaggiotti OE, Daub JT, Vatsiou A, Excoffier L (2014) Widespread signals of convergent adaptation to high altitude in Asia and America. *Am J Hum Genet* 95:394-407.

Fraga LR, Dutra CG, Boquett JA, Vianna FS, Gonçalves RO, Paskulin DD, Costa OL, Ashton-Prolla P, Sanseverino MT, Schuler-Faccini L (2014) p53 signaling pathway polymorphisms associated to recurrent pregnancy loss. *Mol Biol Rep* 41:1871-7.

Foll M, Gaggiotti OE, Daub JT, Vatsiou A, Excoffier L (2014) Widespread signals of convergent adaptation to high altitude in Asia and America. *Am J Hum Genet* 95:394-407.

Gayà-Vidal M, Dugoujon J-M, Esteban E, Athanasiadis G, Rodriguez A, Villena M, Vasquez R, Moral P (2010) Autosomal and X chromosome *Alu* insertions in Bolivian Aymaras and Quechuas: two languages and one gene pool. *Am J Hum Biol* 22:154-62.

González-José R, Bortolini MC, Santos FR, Bonatto SL (2008) The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. *Am J Phys Anthropol* 137:175-87.

Human Genome Diversity Project (HGDP) and Centre d'Étude du Polymorphisme Humain (CEPH) cell line panel. Available from http://www.cephb.fr/en/hgdp_panel.php, last accessed October, 2014.

Hünemeier T, Gómez-Valdés J, Ballesteros-Romero M, Azevedo S, Martínez-Abadías N, Esparza M, Sjøvold T, Bonatto SL, Salzano FM, Bortolini MC, González-José R (2012a) Cultural diversification promotes rapid phenotypic evolution in Xavante Indians. *Proc Natl Acad Sci USA* 109:73-77.

Hünemeier T, Amorim CE, Azevedo S, Contini V, Acuña-Alonzo V, Rothhammer F, Dugoujon JM, Mazières S, Barrantes R, Villarreal-Molina MT, Paixão-Côrtes VR, Salzano FM, Canizales-Quinteros S, Ruiz-Linares A, Bortolini MC (2012b) Evolutionary responses to a constructed niche: ancient Mesoamericans as a model of gene-culture coevolution. *PLoS One* 7:e38862.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy ME, Rebbeck TR, Rosenwaks Z, Levine AJ, Hu W (2009) Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. *Proc Natl Acad Sci USA* 106:9761-6.

Kohn KW, Pommier Y (2005) Molecular interaction map of the p53 and Mdm2 logic elements, which control the Off-On switch of p53 in response to DNA damage. *Biochem Biophys Res Commun* 331:816-27.

Latonen L, Taya Y, Laiho M (2001). UV-radiation induces dose-dependent regulation of p53 response and modulates p53-HDM2 interaction in human fibroblasts. *Oncogene* 20:6784-6793.

Linzer DI, Levine AJ (1979) Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 17:43-52.

Liu J, Tang X, Li M, Lu C, Shi J, Zhou L, Yuan Q, Yang M (2013) Functional MDM4 rs4245739 genetic variant, alone and in combination with P53 Arg72Pro polymorphism, contributes to breast cancer susceptibility. *Breast Cancer Res Treat* 140:151-7.

Ljungman M (2000) Dial 9-1-1 for p53: mechanisms of p53 activation by cellular stress. *Neoplasia* 2:208-25.

Marine JC, Francoz S, Maetens M, Wahl G, Toledo F, Lozano G (2006) Keeping p53 in check: essential and synergistic functions of Mdm2 and Mdm4. *Cell Death Differ* 13:927-34.

Marrero AR, Silva-Junior WA, Bravi CM, Hutz MH, Petzl-Erler ML, Ruiz-Linares A, Salzano FM, Bortolini MC (2007) Demographic and evolutionary trajectories of the Guarani and Kaingang natives of Brazil. *Am J Phys Anthropol* 132:301-10.

Matlashewski GJ, Tuck S, Pim D, Lamb P, Schneider J, Crawford LV (1987) Primary structure polymorphism at amino acid residue 72 of human p53. *Mol Cell Biol* 7:961-3.

Millikan RC, Heard K, Winkel S, Hill EJ, Heard K, Massa B, Mayes L, Williams P, Holston R, Conway K, Edmiston S, de Cotret AR (2006) No association between the MDM2 -309 T/G promoter polymorphism and breast cancer in African-Americans or Whites. *Cancer Epidemiol Biomarkers Prev* 15:175-7.

Momand J, Zambetti GP, Olson DC, George D, Levine AJ (1992) The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* 69:1237-45.

Monge C (1948) *Acclimatization in the Andes*. Baltimore: Johns Hopkins University Press.

Moore LG (2001) Human genetic adaptation to high altitude. *High Alt Med Biol*. 2:257-79.

Moore JH, Gilbert JC, Tsai CT, Chiang FT, Holden T, Barney N, White BC (2006) A flexible computational framework for detecting, characterizing, and interpreting statistical patterns of epistasis in genetic studies of human disease susceptibility. *J Theor Biol* 241:252-61.

Nag S, Qin J, Srivenugopal KS, Wang M, Zhang R (2013) The MDM2-p53 pathway revisited. *J Biomed Res* 27:254-71.

OEGE - Online Encyclopedia for Genetic Epidemiology studies. Available from <http://www.oege.org/software/hwe-mr-calc.shtml>, last accessed January, 2015.

Paskulin DD, Paixão-Côrtes VR, Hainaut P, Bortolini MC, Ashton-Prolla P (2012) The TP53 fertility network. *Genet Mol Biol* 35(4, suppl):939-46.

Pei D, Zhang Y, Zheng J (2012) Regulation of p53: a collaboration between Mdm2 and Mdmx. *Oncotarget* 3:228-35.

Rademaker K, Hodgins G, Moore K, Zarrillo S, Miller C, Bromley GR, Leach P, Reid DA, Álvarez WY, Sandweiss DH (2014) Paleoindian settlement of the high-altitude Peruvian Andes. *Science* 346:466-9.

Sandoval JR, Salazar-Granara A, Acosta O, Castillo-Herrera W, Fujita R, Pena SDJ, Santos FR (2013) Tracing the genomic ancestry of Peruvians reveals a major legacy of pre-Columbian ancestors. *J Hum Genet* 1-8.

Sermeus A, Michiels C (2011) Reciprocal influence of the p53 and the hypoxic pathways. *Cell Death Dis* 2:e164.

Shan J, Brooks C, Kon N, Li M, Gu W (2008) Dissecting roles of ubiquitination in the p53 pathway. *Ernst Schering Found Symp Proc* 1:127-36.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, Li Y, Su B, Lu X (2009) Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. *Am J Hum Genet* 84:534-41.

Shvarts A, Steegenga WT, Riteco N, van Laar T, Dekker P, Bazuine M, van Ham RC, van der Houven van Oordt W, Hateboer G, van der Eb AJ, Jochemsen AG (1996) MDMX: a novel p53-binding protein with some functional properties of MDM2. *EMBO* 15:5349-57.

Shvarts A, Bazuine M, Dekker P, Ramos YF, Steegenga WT, Merckx G, van Ham RC, van der Houven van Oordt W, van der Eb AJ, Jochemsen AG (1997) Isolation and identification of the human homolog of a new p53-binding protein, Mdmx. *Genomics* 43:34-42.

Själänder A, Birgander R, Saha N, Beckman L, Beckman G (1996) p53 polymorphisms and haplotypes show distinct differences between major ethnic groups. *Hum Hered* 46:41-8

Song CG, Fu FM, Wu XY, Wang C, Shao ZM (2012) Correlation of polymorphism rs1563828 in MDM4 gene with breast cancer risk and onset age. *Zhonghua Wai Ke Za Zhi* 50:53-6.

Sucheston L, Witonsky DB, Hastings D, Yildiz O, Clark VJ, Di Rienzo A, Onel K (2011) Natural selection and functional genetic variation in the p53 pathway. *Hum Mol Gene* 20:1502-8.

Sun T, Lee GS, Oh WK, Pomerantz M, Yang M, Xie W, Freedman ML, Kantoff PW (2010) Single-nucleotide polymorphisms in p53 pathway and aggressiveness of prostate cancer in a Caucasian population. *Clin Cancer Res* 16:5244-51.

Tarazona-Santos E, Carvalho-Silva DR, Pettener D, Luiselli D, De Stefano GF, et al. (2001) Genetic differentiation in South Amerindians is related to environmental and cultural diversity: evidence from the Y chromosome. *Am J Hum Genet* 68: 1485–1496.

Tsuneto LT, Probst CM, Hutz MH, Salzano FM, Rodriguez-Delfin LA, Zago MA, Hill K, Hurtado, AM, Ribeiro-dos-Santos AKC, Petzl-Erler ML (2003) HLA class II diversity in seven Amerindian populations. Clues about the origins of the Aché. *Tiss Antig* 62:512-26.

Verdu P, Pemberton TJ, Laurent R, Kemp BM, Gonzalez-Oliver A, Gorodezky C, Hughes CE, Shattuck MR, Petzelt B, Mitchell J, Harry H, William T, Worl R, Cybulski JS, Rosenberg NA, Malhi RS (2014) Patterns of admixture and population structure in Native populations of Northwest North America. *PLoS Genet* 10:e1004530.

Vitzthum VJ (2013) Fifty fertile years: anthropologists' studies of reproduction in high altitude natives. *Am J Hum Biol* 25:179-89.

Vousden KH, Ryan KM. (2009) p53 and metabolism. *Nat Rev Cancer* 9:691-700.

Wang S, Lewis Jr. CM, Jakobsson M, Ramachandran S, Ray N, Bedoya G, Rojas W, Parra MV, Molina JA, Gallo C, Mazzotti G, Poletti G, Hill K, Hurtado AM, Labuda D, Klitz W, Barrantes R, Bortolini MC, Salzano FM, Petzl-Erler ML, Tsuneto LT, Llop E, Rothhammer F, Excoffier L, Feldman MW, Rosenberg NA, Ruiz-Linares A (2007) Genetic variation and population structure in Native Americans. *PLoS Genetics*, 3:2049-2067.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.

Weir BS (1996) The second National Research Council report on forensic DNA evidence. *Am J Hum Genet* 59: 497-500.

Whibley C, Pharoah PD, Hollstein M (2009) p53 polymorphisms: cancer implications. *Nat Rev Cancer* 9:95-107.

WorldClim, available from <http://www.worldclim.org/>, last accessed December 19, 2014.

Xiong S (2013) Mouse models of Mdm2 and Mdm4 and their clinical implications. *Chin J Cancer* 32:371-5.

Internet resources

<http://www.cephb.fr/HGDP-CEPH-Panel/>, last accessed October, 2014.

<http://www.soda-is.com/>, last accessed December 19, 2014.

<http://www.worldclim.org/>, last accessed December 19, 2014.

<http://www.oege.org/software/hwe-mr-calc.shtml>, last accessed January, 2015.

http://http://www.weather.gov.hk/radiation/tidbit/201012/uv_e.htm, last accessed

January, 2015.

Table 1. Derived¹ allele frequencies in the samples studied, classified by altitude:

Populations (n)	<i>TP53 G</i>	<i>MDM2 G</i>	<i>MDM4 G</i>	<i>USP7 A</i>	<i>LIF G</i>	Reference
	<i>rs1042522</i>	<i>rs2279744</i>	<i>rs1563828</i>	<i>rs1529916</i>	<i>rs929271</i>	
Highlands ($\geq 2,500$ m)						
Amantani (29)	0.86	0.14	0.60	0.40	0.78	This study
Anapia (15)	0.77	0.13	0.57	0.3	0.47	This study
Cabanaconde (17)	0.88	0.12	0.74	0.09	0.32	This study
Chivay (18)	0.83	0.14	0.67	0	0.36	This study
Taquile (43)	0.92	0.06	0.63	0.23	0.44	This study
Uros (22)	0.93	0.18	0.84	0.05	0.27	This study
Yanke (10)	0.65	0.05	0.80	0.10	0.45	This study
Aymara (16-18) ³	0.78	0.21	0.50	0.14	0.12	This study
Quechua (15-17) ³	0.53	0.21	0.66	0.24	0.40	This study
<i>F_{st}</i>	0.068 (<i>p</i> = 0.017)	-0.020 (<i>p</i> = 0.801)	-0.006 (<i>p</i> = 0.506)	0.068 (<i>p</i> = 0.013)	0.118 (<i>p</i> < 0.000)	
Lowlands (< 2,500 m)						
Andoas (61)	0.74	0.15	0.53	0.39	0.59	This study
Guarani Kaiowa (16)	0.94	0.06	0.66	0.34	0.56	This study
Guarani Ñandeva (15-16) ³	0.87	0.33	0.72	0.57	0.63	This study
Karitiana (23-24) ³	0.62	0.59	ND ²	ND ²	ND ²	Sucheston <i>et al.</i> , (2011)
Maya (23-21) ³	0.83	0.64	ND ²	ND ²	ND ²	Sucheston <i>et al.</i> , (2011)
Piapoco/Curripaco (12-13) ³	0.92	0.81	ND ²	ND ²	ND ²	Sucheston <i>et al.</i> , (2011)
Pima (21-24) ³	0.65	0.69	ND ²	ND ²	ND ²	Sucheston <i>et al.</i> , (2011)
Surui (17-20) ³	1	0.32	ND ²	ND ²	ND ²	Sucheston <i>et al.</i> , (2011)
<i>F_{st}</i>	0.054 (<i>p</i> = 0.028)	0.274 (<i>p</i> < 0.000)	0.029 (<i>p</i> = 0.200)	0.020 (<i>p</i> = 0.243)	-0.0439 (<i>p</i> = 1.000)	

¹ Defined in comparison to the chimpanzee sequence.

² No data available

³ The number of individuals vary according with the investigated locus.



Figure 1: Map of America, showing the 17 populations included in the present study.

Supplementary Information:

Table S1. Climatic variables that were considered in this study

Characteristics considered ¹	Populations																
	<i>Aymara</i>	<i>Quechua</i>	<i>G. Ñandeva</i>	<i>G. Katowá</i>	<i>Amantani</i>	<i>Andoas</i>	<i>Anapia</i>	<i>Cabanaconde</i>	<i>Chivay</i>	<i>Taqtile</i>	<i>Uros</i>	<i>Yanke</i>	<i>Karitana</i>	<i>Maya</i>	<i>Pima</i>	<i>Surui</i>	<i>Piapoco</i>
Altitude	4312	2656	319	456	3827	235	4366	3441	3820	3825	3818	3820	86	3	361	303	117
Latitude	-19.25	-14.5	-23.8	-23.1	-15.65	-2.9	-16.31	-15.62	-15.63	-15.76	-15.74	-15.64	-8.75	20.22	33.17	-8.4	3.0
Longitude	-69.08	-69.0	-54.5	-55.2	-69.71	-76.4	-68.85	-71.98	-71.98	-69.68	-69.93	-71.95	-63.84	-90.47	-11.87	-54.9	-68.0
UV irradiance 1	230	202	212	210	212	206	212	248	224	216	214	226	180	216	246	180	162
UV irradiance 2	210	198	150	154	206	209	204	241	214	206	206	214	194	184	183	176	172
UV irradiance 3	252	205	212	210	220	203	219	219	235	220	221	238	165	246	309	161	153
Bio 1	32	166	218	221	83	257	86	132	83	84	85	93	260	267	212	253	267
Bio 2	181	113	117	118	114	99	137	171	183	115	131	182	103	102	179	135	104
Bio 3	71	72	57	61	70	86	65	79	73	71	71	74	71	62	46	69	83
Bio 4	2664	1112	2748	2595	1245	495	1920	1070	1632	1216	1370	1540	631	2060	7809	523	586

Bio 5	145	236	314	309	155	316	174	232	191	156	167	200	329	347	410	354	337
Bio 6	-109	81	112	116	-7	202	-36	16	-57	-4	-16	-43	185	183	24	161	212
Bio 7	254	155	202	193	162	114	210	216	248	160	183	243	144	164	386	248	260
Bio 8	61	173	237	238	93	253	104	142	99	94	96	109	259	282	308	248	260
Bio 9	16	148	186	186	64	262	58	119	64	66	64	70	258	264	247	251	272
Bio 10	61	177	249	250	94	262	105	142	99	95	98	109	268	288	313	259	274
Bio 11	-7	148	179	185	64	250	58	116	58	66	64	70	252	237	114	247	260
Bio 12	194	1668	1593	1566	964	2552	728	289	453	1134	743	430	2124	867	206	2415	3192
Bio 13	79	276	194	191	210	260	161	75	107	249	158	106	310	177	33	385	459
Bio 14	0	25	66	57	6	165	7	1	1	7	3	1	21	11	3	9	95
Bio 15	159	62	29	33	84	13	81	107	105	83	89	107	61	79	48	69	46
Bio 16	179	773	553	555	539	728	400	205	307	632	435	295	921	455	77	1083	1266
Bio 17	1	101	246	213	27	573	28	7	9	34	16	8	100	36	13	49	363
Bio 18	179	537	482	498	420	573	333	205	268	298	318	260	319	199	61	343	377
Bio 19	1	101	322	266	27	646	28	9	9	34	16	8	353	62	62	1082	1233

¹Altitude in meters; Latitude in degrees (positive values, north, negative values, south); Longitude: in degrees west of Greenwich; UV irradiance 1: Annual mean of ultraviolet irradiance; UV irradiance 2: Mean of ultraviolet irradiance in the coldest semester; UV irradiance 3: Mean of ultraviolet irradiance in the warmest semester; Bio = Bioclimatic variables: Bio 1: Annual mean temperature; Bio 2: Mean diurnal range; Bio 3: Isothermality (the ratio of mean diurnal range to temperature annual range x 100); Bio 4: Temperature seasonality (standard deviation x 100 of average annual daily temperatures); Bio 5: Maximum temperature of warmest month; Bio 6: Minimum temperature of coldest month; Bio 7: Temperature annual range; Bio 8: Mean temperature of wettest quarter; Bio 9: Mean temperature of driest quarter; Bio 10: Mean temperature of warmest quarter; Bio 11: Mean temperature of coldest quarter; Bio 12: Annual precipitation; Bio 13: Precipitation of wettest month; Bio 14: Precipitation in the driest month; Bio 15: Precipitation seasonality (coefficient of variation of average annual daily precipitations); Bio 16: Precipitation of wettest quarter; Bio 17: Precipitation of driest quarter; Bio 18: Precipitation of warmest quarter; Bio 19: Precipitation of coldest quarter. Maximum, average, and minimum temperatures are in degrees Celsius x 10. Precipitation numbers are in mm, and irradiance are in W/m².

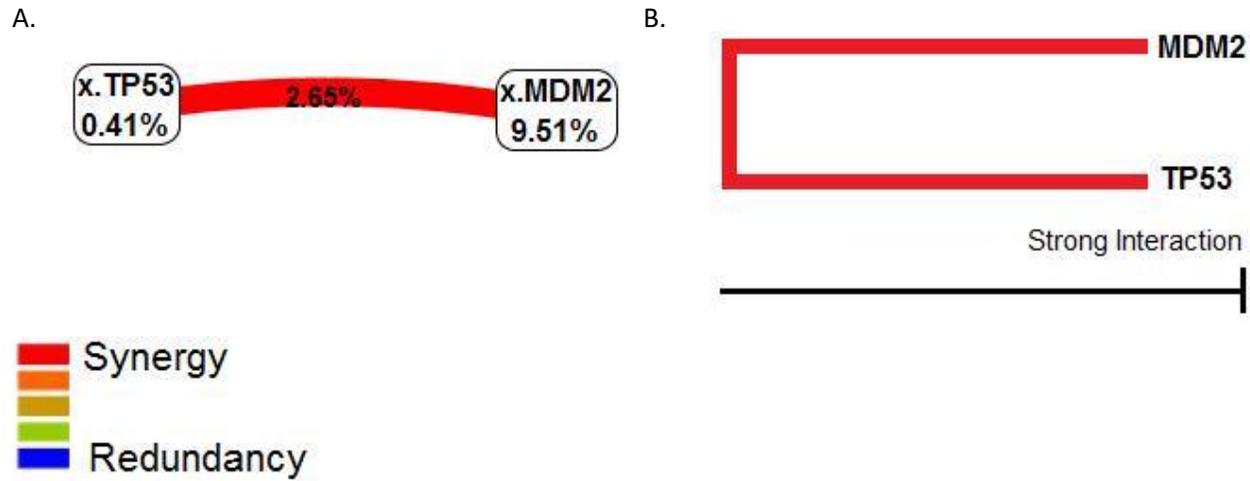


Figure S1. Interaction entropy map. (A) Percentage reported explains the reduction of the whole uncertainty of the data, or conversely, the information gain of a given SNP. Synergistic interaction between *TP53* rs1042522 and *MDM2* rs2279744 is displayed by the red line (2.65% of information gain). The independent main effect was relatively large for *MDM2* (9.51%) and non-significant for *TP53* (0.41%); and (B) Interaction dendrogram for *TP53* and *MDM2*.

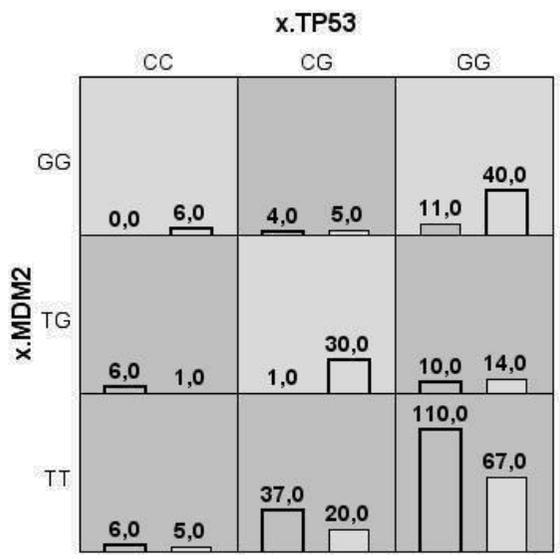


Figure S2. The best model for the gene-environment interaction between *TP53* and *MDM2* for altitude adaptation. The left columns in the histograms within each genotype category represent the absolute number of individuals who inhabit the highlands ($\geq 2,500$ m), whereas the right columns show the numbers of individuals who inhabit the lowlands ($< 2,500$ m).

CAPÍTULO 4

DISCUSSÃO

A discussão específica e completa referente aos resultados obtidos neste trabalho encontram-se no manuscrito a ser submetido em revista científica de circulação internacional (ver capítulo 3). Sendo assim, neste capítulo somente aspectos mais gerais referentes ao tema serão abordados. No entanto alguma sobreposição é inevitável.

Desde que o gene *TP53* foi descoberto, em 1979, diversos estudos tem sido conduzidos a fim de, inicialmente, desvendar sua função e regulação, e mais recentemente esclarecer o papel de variações neste gene na susceptibilidade à patologias humanas. Dentre as patologias associadas à alterações no gene *TP53* podemos citar câncer, síndrome de Li Fraumeni e desordens gestacionais (Fraga *et al.*, 2014; Paskulin *et al.*, 2012; Palmero *et al.*, 2010).

Milhares de estudos científicos têm sido publicados nos últimos anos sobre o papel de *TP53* e genes que pertencem a sua rede clássica de interações. No entanto, quando nos referimos a estudos em um contexto evolutivo, raríssimos estudos são encontrados. Tendo isso em vista, nosso principal objetivo aqui é ajudar a preencher essa lacuna, fornecendo informações acerca de *TP53* e mais 4 genes cujas proteínas codificadas interagem intimamente com *TP53* em populações nativo americanas.

As análises apresentadas aqui com o conjunto de dados (a), que incluiu informações originais de 5 SNPs em 12 populações nativas americanas, revela o papel bem conhecido da deriva genética nesses grupos, ilustrado por grande diferença de frequências alélicas *MDM2-309-G* entre os duas tribos Guarani. No entanto, outros resultados podem estar associados com adaptações dos andinos nativos a altas altitudes. Os alelos *USP7-G* (rs1529916) e *LIF-T* (rs929271) foram fortemente correlacionada com irradiação ultravioleta, bem como as variáveis de temperatura. Em outras palavras, os alelos ancestrais *USP7-G* e *LIF-T* estão mais presentes nos ambientes estressores (baixa temperatura, clima árido, grande amplitude térmica e alto índices de radiação UV) típico de altitudes elevadas.

Os alelos derivados desses SNPs, *USP-A* e *LIF-G*, têm sido associados a diversas patologias humanas. O alelo A de *USP7* foi associado a um risco aumentado de desenvolver câncer da próstata de alto risco ou intermediário (classificação de risco D'Amico). Portadores do genótipo AA apresentaram chance cerca de 1,5 vez maior de desenvolver câncer de próstata de risco intermediário quando comparado ao genótipo GG (Sun *et al.*, 2010). Já o alelo G de *LIF* tem sido associado com mulheres jovens

(abaixo de 35 anos) que tiveram falhas na fertilização *in vitro* (Sun *et al.*, 2010; Kang *et al.*, 2009). Sendo assim podemos considerar que estes alelos ancestrais estejam sendo mantidos em frequências elevadas como fatores protetores para as populações de terras altas.

Anteriormente acreditava-se que as populações humanas que vivem em grandes altitudes submetidas a baixas concentrações de oxigênio teriam a sua fertilidade reduzida (Monge, 1948). Entretanto estudos mais recentes (Vitzthum, 2013) têm mostrado que o funcionamento reprodutivo dos nativos de grande altitude é na verdade adaptado à hipóxia e outras condições extremas. Esse dado é suportado por registros do governo do Peru, que indicam uma taxa média de complicações gestacionais e pós-parto menor em regiões Aymara quando comparada a média nacional (1.8% e 5% respectivamente) http://www.dge.gob.pe/publicaciones/pub_asis/asis26.pdf: page 165; <http://www.dge.gob.pe/portal/docs/intsan/asis2012.pdf>: page 76).

Esses resultados nos fazem acreditar que as distribuições alélicas encontradas nos polimorfismos nos genes *USP7* (rs1529916) e *LIF* (rs929271) possam estar conectados com adaptações que levaram ao sucesso reprodutivo dos nativos em ambientes hostis, como aquele encontrado nos Andes.

Quando foi utilizado o segundo conjunto de dados (b: 17 populações nativas apenas para os genes *MDM2* e *TP53*) foi encontrado uma forte correlação do alelo T de *MDM2* com índice da radiação ultravioleta, altitude e médias na temperatura. Ainda, o alelo T estava significativamente mais presentes em comunidades localizadas em altas altitudes.

O alelo derivado de *MDM2*, G, quando em homozigose apresenta uma elevada taxa de mRNA, e conseqüentemente mais proteína é traduzida. Como a regulação de *MDM2* e p53 é negativa, aumentando a concentração de *MDM2* automaticamente os níveis de p53 diminuem, o que deixará a célula mais suscetível a danos no DNA (Alazzouzi *et al.*, 2007). Desse modo, nossos achados fazem sentido do ponto de vista biológico.

Além disso, os resultados apresentados aqui estão de acordo com Shi *et al.* (2009), uma vez que eles mostraram que os baixos níveis de UV são significativamente correlacionadas com o genótipo GG de *MDM2* em populações chinesas. Os autores

sugeriram que tal genótipo é selecionado em áreas de baixa atividade de UV para evitar efeitos indesejáveis da p53 (em altitudes mais elevadas, a atmosfera é mais fina e irá filtrar menos radiação UV, observe na figura 6. No entanto, diferentemente de Shi *et al.*, que encontraram associação de p53 Arg com a latitude, nós não observamos correlação entre a latitude e esse alelo de p53.

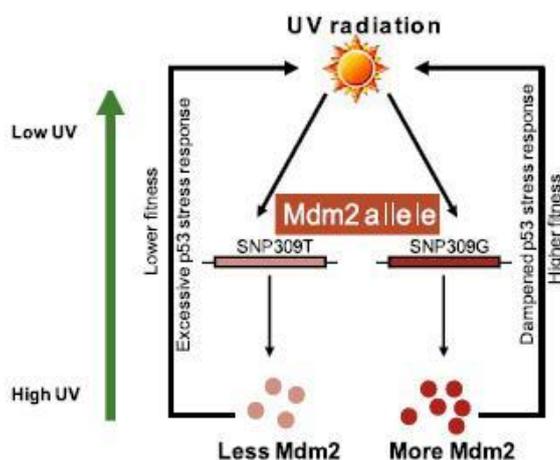


Figura 6. Correlação funcional de MDM2, sugerindo que o alelo G de *MDM2* seria capaz de evitar a superexpressão de p53, permitindo então que ela funcione adequadamente como um supressor tumoral.

Já Sucheston *et al.*, (2011) investigaram 52 populações em todo o mundo a partir do painel de CEPH-HGDP o, para os mesmos polimorfismos nos genes *TP53* e *MDM2*, e não encontraram nenhuma correlação significativa com variáveis climáticas. No entanto, a amostra desse estudo é muito heterogênea, utilizando populações de diversos continentes, e a amostra nativa americana investigada é muito mais restrita do que a do presente estudo, o que pode ser a causa dos resultados divergentes.

Apesar de dois estudos prévios (Shi *et al.*, 2009 e Sucheston *et al.*, 2011) já terem analisado as variações no genes *TP53* e *MDM2* sob um contexto evolutivo, este foi o primeiro estudo a expandir a rede clássica de *TP53* para um total de cinco genes. Além disso, nunca esse conjunto de marcadores tinha sido investigados em nativos americanos. Desse modo foi possível pela primeira vez detectar que a rede de p53 possa estar envolvida na adaptação em altas altitudes, pelo menos para populações andinas.

CAPÍTULO 5
CONCLUSÕES

Uma rede de p53 bem regulada é crucial para manter a integridade do genoma. Vários polimorfismos nesta via têm sido descritos e associados com as mais diversas patologias, no entanto, quando observado sob ponto de vista evolutivo, as diferentes frequências de alelos ao longo das populações humanas podem ser interpretadas como sinal de seleção natural.

O presente estudo identifica potenciais assinaturas de adaptação de alta altitude em populações andinas na rede de *TP53*.

- O alelo G de *USP7* tem uma correlação negativa forte com índices de temperatura, e positiva forte com média anual de irradiação ultravioleta e média irradiação UV no inverno;

- AleloT de *LIF* tem também uma forte correlação negativa forte com índices de temperatura, e positiva com a média anual de irradiação ultravioleta;

- Alelo MDM2-309-T tem uma correlação muito forte com características com inverno, com alta irradiação UV, altitude e índices de temperatura;

A alta frequência dos alelos ancestrais desses polimorfismos nos genes *USP7*, *LIF* e *MDM2* em populações andinas submetidas a um ambiente de estresse celular pode estar indicando que estes alelos, em certas circunstâncias, estariam protegendo essas populações. Esses achados corroboram a sugestão de seleção natural atuando na via de *TP53*.

REFERÊNCIAS BIBLIOGRÁFICAS

Alazzouzi H, Suriano G, Guerra A, Plaja A, Espín E, Armengol M, Alhopuro P, Velho S, Shinomura Y, González-Aguilera JJ, Yamamoto H, Aaltonen LA, Moreno V, Capellà G, Peinado MA, Seruca R, Arango D, Schwartz S Jr (2007) Tumour selection advantage of non-dominant negative P53 mutations in homozygotic MDM2-SNP309 colorectal cancer cells. *J Med Genet* 44:75-80.

Atwal GS, Bond GL, Metsuyanin S, Papa M, Friedman E, Distelman-Menachem T, Ben Asher E, Lancet D, Ross DA, Sninsky J, et al. (2007) Haplotype structure and selection of the MDM2 oncogene in humans. *Proc Natl Acad Sci USA* 104:4524–4529.

Atwal GS, Kirchhoff T, Bond EE, Monagna M, Menin C, Bortorelle R, Scaini MC, Bartel F, Böhnke A, Pempe C, et al. (2009) Altered tumor formation and evolutionary selection of genetic variants in the human MDM4 oncogene. *Proc Natl Acad Sci USA* 106:10236–10241.

Bamshad M e Wooding SP (2003) Signatures of natural selection in the human genome. *Nat Rev Genet* 4(2):99-111.

Beall CM (2007) Two routes to functional adaptation: Tibetan and Andean high-altitude natives. *Proc Natl Acad Sci U S A*, 1:8655-60.

Beckman G, Birgander R, Sjölander A, Saha N, Holmberg PA, Kivelä A e Beckman L (1994) Is p53 polymorphism maintained by natural selection? *Hum Hered* 44:266–270.

Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE e Hirschhorn JN (2004) Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet* 74:1111-1120.

Bohlman S, Manfredi JJ (2014) p53-independent effects of Mdm2. *Subcell Biochem* 85:235-46.

Bond GL, Hu W, Bond EE, Robins H, Lutzker SG, Arva NC, Bargonetti J, Bartel F, Taubert H, Wuerl P, Onel K, Yip L, Hwang SJ, Strong LC, Lozano G, Levine AJ (2004) A Single nucleotide polymorphism in the MDM2 promoter attenuates the

p53 tumor suppressor pathway and accelerates tumor formation in humans 119(5):591-602.

Bortolini MC, Salzano FM, Thomas MG, Stuart S, Nasanen SPK, Bau CHD, Hutz MH, Layrisse Z, Petzl-Erler ML, Tsuneto LT *et al.*, (2003) Y-Chromosome evidence for differing ancient demographic histories in the Americas. *Am J Hum Genet* 73(3): 524–539.

Bortolini MC, González-José R, Bonatto SL, Santos FR (2014) Reconciling pre-Columbian settlement hypotheses requires integrative, multidisciplinary, and model-bound approaches. *Proc Natl Acad Sci USA* 111:E213-4.

Brooks CL, Gu W (2003) Ubiquitination, phosphorylation and acetylation: the molecular basis for p53 regulation. *Curr Opin Cell Biol* 15(2):164-71.

Brooks CL, Li M, Hu M, Shi Y, Gu W (2007) The p53-Mdm2-HAUSP complex is involved in p53 stabilization by HAUSP. *Oncogene* 26(51):7262-6.

Brosh R, Rotter V (2009) When mutants gain new powers: news from the mutant p53 field. *Nat Rev Cancer* 9(10):701-13.

Cahilly-Snyder L, Yang-Feng T, Francke U, George DL (1987) Molecular analysis and chromosomal mapping of amplified genes isolated from a transformed mouse 3T3 cell line. *Somat Cell Mol Gene* 13(3):235-44.

Chao CC (2015) Mechanisms of p53 degradation. *Clin Chim Acta* 438:139-47.

Chen J (2012) The Roles of MDM2 and MDMX Phosphorylation in Stress Signaling to p53. *Genes Cancer* 3(3-4):274-82

Cho Y, Gorina S, Jeffrey PD, Pavletich NP (1994) Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*; 265:346–55.

Donehower LA, Harvey M, Slagle BL, McArthur MJ, Montgomery CA Jr, Butel JS, Bradley A (1992) Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 356(6366):215-21.

Dumont P, Leu JI, Della Pietra AC 3rd, George DL, Murphy M (2003) The codon 72 polymorphic variants of p53 have markedly different apoptotic potential. *Nat Genet* 33(3):357-65.

Fagundes NJ, Kanitz R, Bonatto SL (2008) A reevaluation of the Native American mtDNA genome diversity and its bearing on the models of early colonization of Beringia. *PLoS One* 3(9):e3157.

Fakharzadeh SS, Trusko SP, George DL (1991) Tumorigenic potential associated with enhanced expression of a gene that is amplified in a mouse tumor cell line. *EMBO J* 10(6):1565-9.

Fan C, Wei J, Yuan C, Wang X, Jiang C, Zhou C, Yang M. The functional TP53 rs1042522 and MDM4 rs4245739 genetic variants contribute to Non-Hodgkin lymphoma risk. *PLoS One*. 2014 Sep 9;9(9):e107047.

Finlay CA, Hinds PW e Levine AJ (1989) The p53 proto-oncogene can act as a suppressor of transformation. *Cell* 1989, 57:1083–1093.

Foll M, Gaggiotti OE, Daub JT, Vatsiou A, Excoffier L (2014) Widespread signals of convergent adaptation to high altitude in Asia and America. *Am J Hum Genet* 95:394-407.

Fraga LR, Dutra CG, Boquett JA, Vianna FS, Gonçalves RO, Paskulin DD, Costa OL, Ashton-Prolla P, Sanseverino MT, Schuler-Faccini L (2014) p53 signaling pathway polymorphisms associated to recurrent pregnancy loss. *Mol Biol Rep* 41:1871-1877.

Garritano S, Inga A, Gemignani F e Landi S (2013) More targets, more pathways and more clues for mutant p53. *Oncogenesis* 1;2:e54.

Gaspar PA, Hutz MH, Salzano FM, Hill K, Hurtado AM, Petzl-Erier ML, Tsuneto LT e Weimer TA (2002) Polymorphisms of CYP1A1, CYP2E1, GSTM1, GSTT1, e TP53 genes in Amerindians. *Am J Phys Ant* 119:249-256.

González-José R, Bortolini MC, Santos FR, Bonatto SL (2008) The peopling of America: craniofacial shape variation on a continental scale and its interpretation from an interdisciplinary view. *Am J Phys Anthropol* 137:175-87.

Greenblatt MS, Bennett WP, Hollstein M, Harris CC. (1994) Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res.*; 54: 4855–78. MEDLINE

Hancock AM, Witonsky DB, Gordon AS, Eshel G, Pritchard JK, Coop G e Di Rienzo A (2008) Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genet* 4:e32.

Harris EE e Meyer D (2006) The molecular signature of selection underlying human adaptations. *Am J Phys Anthropol.* 2006;Suppl 43:89-130.

Haupt, Y., Maya, R., Kazaz, A. & Oren, M (1997) Mdm2 promotes the rapid degradation of p53. *Nature* 387, 296–299.

Hu M, Gu L, Li M, Jeffrey PD, Gu W e Shi Y (2006) Structural basis of competitive recognition of p53 and MDM2 by HAUSP/USP7: implications for the regulation of the p53-MDM2 pathway. *PLoS Biol* 4:228-239.

Hu W, Feng Z, Teresky AK and Levine AJ (2007) p53 regulates maternal reproduction through LIF. *Nature* 450:721-724.

Huang C, Liu W, Ji GX, Gu AH, Qu JH, Song L, Wang XR (2012) Genetic variants in TP53 and MDM2 associated with male infertility in Chinese population. *Asian J Androl* 14(5):691-4.

Huang, L, Yan, Z, Liao, X, Li, Y, Yang, J, Wang, ZG, Zuo, Y, Kawai, H, Shadfan, M, Ganapathy, S & Yuan, ZM (2011) The p53 inhibitors MDM2/MDMX complex is required for control of p53 activity in vivo. *Proc Natl Acad Sci U S A* 108: 12001-12006.

Hünemeier T, Amorim CEG, Azevedo S, Contini V, Acuna-Alonzo V, Rothhammer F, Dugoujon JM, Mazières S, Barrantes R, Villareal-Molina MT, et al. (2012) Evolutionary responses to a constructed niche: Ancient Meso americans as a model of gene-culture coevolution. *PLoS One* 7(6): e38862.

Ichikawa Y (1969) Differentiation of a cell line of myeloid leukemia. *J Cell Physiol*; 74: 223-34.

Iwakuma T e Lozano G (2003) MDM2, an introduction. *Mol Cancer Res* 1:993-1000.

Jones SN, Roe AE, Donehower LA, Bradley A (1995) Rescue of embryonic lethality in Mdm2-deficient mice by absence of p53. *Nature* 378: 206-8.

Kang HJ, Feng Z, Sun Y, Atwal G, Murphy WE, Rebbeck T, Rosenwaks Z, Levine AJ e Hu W (2009) Single-nucleotide polymorphisms in the p53 pathway regulate fertility in humans. *Proc Natl Acad Sci USA* 106:9761-6.

Kay C, Jeyendran RS e Coulam CB (2006) p53 tumour suppressor gene polymorphism is associated with recurrent implantation failure. *Reprod Biomed Online* 13:492–496.

Ko A, Cantor RM, Weissglas-Volkov D, Nikkola E, Reddy PM, Sinsheimer JS, Pasaniuc B, Brown R, Alvarez M, Rodriguez A *et al.*, (2014) Amerindian-specific regions under positive selection harbour new lipid variants in Latinos. *Nat Commun* 5:3983.

Kohn KW e Pommier Y (2005) Molecular interaction map of the p53 and Mdm2 logic elements, which control the Off-On switch of p53 in response to DNA damage. *Biochem Biophys Res Commun* 331:816-27.

Lane DP (1992) Cancer. p53, guardian of the genome. *Nature* 358(6381):15-6.

Lane DP e Crawford LV (1979) T antigen is bound to a host protein in SV40-transformed cells. *Nature* 278:261–263.

Laptenko O, Prives C (2006) Transcriptional regulation by p53: one protein, many possibilities. *Cell Death Differ* 13(6):951-61.

Lass A, Weiser W, Munafo A, Loumaye E (2001) Leukemia inhibitory factor in human reproduction. *Fertil Steril* 76(6):1091-6.

Le MT, Teh C, Shyh-Chang N, Xie H, Zhou B, Korzh V, Lodish HF, Lim B (2009) MicroRNA-125b is a novel negative regulator of p53. *Genes Dev* 23: 862–876

Levine AJ, Hu W e Feng Z (2006) The P53 pathway: what questions remain to be explored? *Cell Death Differ* 13:1027-1036.

Li L, Tan Y, Chen X, Xu Z, Yang S, Ren F, Guo H, Wang X, Chen Y, Li G, Wang H (2014) MDM4 overexpressed in acute myeloid leukemia patients with complex karyotype and wild-type TP53. *PLoS One* 9(11):e113088.

Li M, Brooks CL, Kon N, Gu W (2004) A dynamic role of HAUSP in the p53-Mdm2 pathway. *Mol Cell* 13(6):879-86.

Li M, Chen D, Shiloh A, Luo J, Nikolaev AY, Qin J e Gu W (2002) Deubiquitination of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* 416:648–653.

Linzer DI e Levine AJ (1979) Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 17:43–52.

Liu J, Tang X, Li M, Lu C, Shi J, Zhou L, Yuan Q e Yang M (2013) Functional MDM4 rs4245739 genetic variant, alone and in combination with P53 Arg72Pro23 polymorphism, contributes to breast cancer susceptibility. *Breast Cancer Res Treat* 140:151-7.

Ljungman M (2000) Dial 9-1-1 for p53: mechanisms of p53 activation by cellular stress. *Neoplasia* 2:208-25.

Marine JC, Dyer MA e Jochemsen AG (2007) MDMX: from bench to bedside. *J Cell Sci* 120:371-378.

Marine JC, Francoz S, Maetens M, Wahl G, Toledo F e Lozano G (2006) Keeping p53 in check: essential and synergistic functions of Mdm2 and Mdm4. *Cell Death Differ* 13:927–934.

Matlashewski GJ, Tuck S, Pim D, Lamb P, Schneider J e Crawford LV (1987) Primary structure polymorphism at amino acid residue 72 of human p53. *Mol Cell Biol* 7:961-3.

McBride OW, Merry D e Givol D (1986) The gene of human p53 cellular tumor antigen is located on chromosome 17 short arm (17p13). *Proc Natl Acad Sci USA* 83:130-134.

Meier DE, Luckey MM, Wallenstein S, Clemens TL, Orwoll ES e Waslien CI (1991) Calcium, vitamin D, and parathyroid hormone status in young white and black women: association with racial differences in bone mass. *J Clin Endocrinol Metab* 72:703-710.

Metcalf D (1992) Leukemia inhibitory factor a puzzling polyfunctional regulator. *Growth Factors*. 7:169-73.

Meulmeester E, Maurice MM, Boutell C, Teunisse AF, Ovaa H, Abraham TE, Dirks RW e Jochemsen AG (2005) Loss of HAUSP-mediated deubiquitination contributes to DNA damage induced destabilization of Hdmx and Hdm2. *Mol Cell* 18:565-576.

Meulmeester E, Pereg Y, Shiloh Y e Jochemsen AG (2005b) ATM-mediated phosphorylations inhibit Mdmx/Mdm2 stabilization by HAUSP in favor of p53 activation. *Cell Cycle* 4:1166-1170.

Millikan RC, Heard K, Winkel S, Hill EJ, Heard K, Massa B, Mayes L, Williams P, Holston R, Conway K, et al. (2006) No association between the MDM2 -309 T/G promoter polymorphism and breast cancer in African-Americans or Whites. *Cancer Epidemiol Biomarkers Prev* 15:175-7.

Momand J, Zambetti GP, Olson DC, George D e Levine AJ (1992) The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* 69:1237-1245.

Monge C (1948) *Acclimatization in the Andes*. Baltimore: Johns Hopkins University Press.

Montes de Oca Luna R, Wagner DS, Lozano G (1995) Rescue of early embryonic lethality in mdm2- deficient mice by deletion of p53. *Nature* 378:203-6.

Nag S, Qin J, Srivenugopal KS, Wang M e Zhang R (2013) The MDM2-p53 pathway revisited. *J Biomed Res* 27:254-71

Olivier M, Eeles R, Hollstein M, Khan MA, Harris CC, Hainaut P. (2002) The IARC TP53 database: new online mutation analysis and recommendations to users. *Hum Mutat* 19: 607–14.

Palmero EI, Achatz MI, Ashton-Prolla P, Olivier M, Hainaut P (2010) Tumor protein 53 mutations and inherited cancer: beyond Li-Fraumeni syndrome. *Curr Opin Oncol* 22(1):64-9.

Pan Y e Chen J (2003) MDM2 promotes ubiquitination and degradation of MDMX. *Mol Cell Biol* 23(15):5113-21.

Paskulin DD, Paixão-Côrtes VR, Hainaut P, Bortolini MC, Ashton-Prolla P (2012) The TP53 fertility network. *Genet Mol Biol* 35(4, suppl):939-46.

Pei D, Zhang Y, Zheng J (2012) Regulation of p53: a collaboration between Mdm2 and Mdmx. *Oncotarget* 3:228-35.

Pim D e Banks L (2004) p53 polymorphic variants at codon 72 exert different effects on cell cycle progression. *Int J Cancer* 108(2):196-9.

Potts JT (2005) Parathyroid hormone: past and present. *J Endocrinol* 187:311-325.

Rademaker K, Hodgins G, Moore K, Zarrillo S, Miller C, Bromley GR, Leach P, Reid DA, Álvarez WY, Sandweiss DH (2014) Paleoindian settlement of the high-altitude Peruvian Andes. *Science* 346:466-9.

Rideout WM 3rd, Coetzee GA, Olumi AF, Jones PA (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 249(4974):1288-90.

Salzano FM (2006) The Prehistoric Colonization of the Americas: Evidence and Models. *Evo Edu Outreach* 4:199–204

Salzano FM (2009) The fissio-fusion concept. *Curr Anthropol* 50:959.

Salzano FM e Sans M (2014) Interethnic admixture and the evolution of Latin American populations. *Genet Mol Biol.* 37(1 Suppl):151-70.

Schroeder KB, Schurr TG, Long JC, Rosenberg NA, Crawford MH, Tarskaia LA, Osipova LP, Zhadanov SI, Smith DG (2007) A private allele ubiquitous in the Americas. *Biol Lett* 3(2):218-23.

Sharp DA, Kratowicz SA, Sank MJ e George DL (1999) Stabilization of the MDM2 oncoprotein by interaction with the structurally related MDMX protein. *J Biol Chem* 274:38189-38196.

Shi D, Gu W (2012) Dual Roles of MDM2 in the Regulation of p53: ubiquitination dependent and ubiquitination independent mechanisms of MDM2 repression of p53 activity. *Genes Cancer*(3-4):240-8.

Shi H, Tan SJ, Zhong H, Hu W, Levine A, Xiao CJ, Peng Y, Qi XB, Shou WH, Ma RL, et al. (2009) Winter temperature and UV are tightly linked to genetic changes in the p53 tumor suppressor pathway in Eastern Asia. *Am J Hum Genet* 84:534–541.

Shvarts A, Bazuine M, Dekker P, Ramos YF, Steegenga WT, Merckx G, van Ham RC, van der Houven van Oordt W, van der Eb AJ e Jochemsen AG (1997) Isolation and identification of the human homolog of a new p53-binding protein, Mdmx. *Genomics* 43:34-42.

Shvarts A, Steegenga WT, Riteco N, van Laar T, Dekker P, Bazuine M, van Ham RC, van der Houven van Oordt W, Hateboer G, van der Eb AJ, et al. (1996) MDMX: a novel p53-binding protein with some functional properties of MDM2. *EMBO J* 15:5349-5357.

Sjalander A, Birgander R, Saha N, Beckman L e Beckman G (1996) p53 polymorphisms and haplotypes show distinct differences between major ethnic groups. *Hum Hered* 46:41–48.

Song CG, Fu FM, Wu XY, Wang C, Shao ZM (2012) Correlation of polymorphism rs156382 in MDM4 gene with breast cancer risk and onset age. *Zhonghua Wai Ke Za Zhi* 50(1):53-6.

Stewart CL, Kaspar P, Brunet LJ, Bhatt H, Gadi I, Köntgen F e Abbondanzo SJ (1992) Blastocyst implantation depends on maternal expression of leukaemia inhibitory factor. *Nature* 359:76-9.

Sucheston L, Witonsky DB, Hastings D, Yildiz O, Clark VJ, Di Rienzo A e Onel K (2011) Natural selection and functional genetic variation in the p53 pathway. *Hum Mol Genet* 20:1502-8.

Sun T, Lee GS, Oh WK, Pomerantz M, Yang M, Xie W, Freedman ML, Kantoff PW (2010) Single-nucleotide polymorphisms in p53 pathway and aggressiveness of prostate cancer in a Caucasian population. *Clin Cancer Res* 16:5244-51.

Tao, W. & Levine, A. J (1999) Nucleocytoplasmic shuttling of oncoprotein Hdm2 is required for Hdm2-mediated degradation of p53. *Proc Natl Acad Sci USA* 96, 3077–3080.

Terzian T, Wang Y, Van Pelt CS, Box NF, Travis EL, Lozano G (2007) Haploinsufficiency of Mdm2 and Mdm4 in tumorigenesis and development. *Mol Cell Biol* 27(15):5479-85.

Thomas M, Kalita A, Labrecque S, Pim D, Banks L, Matlashewski G (1999) Two polymorphic variants of wild-type p53 differ biochemically and biologically. *Mol Cell Biol* 19(2):1092-100.

Tobias JH, Cook DG, Chambers TJ e Dalzell N (1994) A comparison of bone mineral density between Caucasian, Asian and Afro-Caribbean women. *Clin Sci (Lond)* 87:587-591.

Trouillas M, Saucourt C, Guillotin B, Gauthereau X, Taupin JL, Moreau JF, Boeuf H (2009) The LIF cytokine: towards adulthood. *Eur Cytokine Netw* 20(2):51-62.

Vitzthum VJ (2013) Fifty fertile years: anthropologists' studies of reproduction in high altitude natives. *Am J Hum Biol* 25:179-89.

Wade M, Wang YV, Wahl GM (2010) The p53 orchestra: Mdm2 and Mdmx set the tone. *Trends Cell Biol* 20: 299–309

Whibley C, Pharoah PD, Hollstein M (2009) p53 polymorphisms: cancer implications. *Nat Rev Cancer* 9(2):95-107.

Yu H, Yue X, Zhao Y, Li X, Wu L, Zhang C, Liu Z, Lin K, Xu-Monette ZY, Young KH, Liu J, Shen Z, Feng Z, Hu W (2014) LIF negatively regulates tumour-suppressor p53 through Stat3/ID1/MDM2 in colorectal cancers. *Nat Commun* 5:5218.

Zhang Q, Zeng SX, Lu H (2014) Targeting p53-MDM2-MDMX loop for cancer therapy. *Subcell Biochem* 85:281-319.

Zhang YW, Guan J, Zhang Y, Qiu YR, Chen LH (2012) Role of an MDM4 polymorphism in the early age of onset of nasopharyngeal carcinoma. *Oncol Lett* 3(5):1115-1118.

Zhao GY, Lin ZW, Lu CL, Gu J, Yuan YF, Xu FK, Liu RH, Ge D, Ding JY (2014) USP7 overexpression predicts a poor prognosis in lung squamous cell carcinoma and large cell carcinoma. *Tumour Biol*.

Internet resources

<http://www.cephb.fr/HGDP-CEPH-Panel>, last accessed October, 2014.

http://www.dge.gob.pe/publicaciones/pub_asis/asis26, last accessed February 2015.

<http://www.dge.gob.pe/portal/docs/intsan/asis2012>, last accessed February 2015.

<http://www-p53.iarc.fr/p53database.htm>, last accessed January 2015.

