

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE BIOCÊNCIAS
DEPARTAMENTO DE GENÉTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

VARIABILIDADE MOLECULAR NA REGIÃO 3'UTR DO GENE DO
RECEPTOR DA LIPOPROTEÍNA DE BAIXA DENSIDADE – DIVERSIDADE
INTRA E INTERCONTINENTAL

ANA HELENA HELLER

Orientador: Dr. Francisco Mauro Salzano

Co-orientador: Dr. Sandro Luis Bonatto

Dissertação apresentada ao Programa de Pós-
Graduação em Genética e Biologia Molecular
da Universidade Federal do Rio Grande do Sul
para obtenção do grau de Mestre

Porto Alegre

2003

...but the real reason why we study evolutionary biology is that we are interested in origins. We want to know where we came from.

John Maynard Smith and Eörs Szathmáry

Dedico esta Dissertação a todos os povos
nativos do planeta e ao direito que possuem de
conservação de suas identidades culturais.

AGRADECIMENTOS

Ao meu orientador, Prof. Francisco M. Salzano, pela dedicação e entusiasmo com que realiza seu trabalho e que contagia a todos. Agradeço, também, por toda atenção e apoio recebidos durante esses dois anos.

Ao meu co-orientador, Sandro L. Bonatto, pela disponibilidade para demoradas discussões nos últimos momentos deste trabalho.

À Loreta B. de Freitas, por organizar e administrar nosso laboratório, proporcionando condições adequadas para o desenvolvimento dos projetos.

Aos colegas do Laboratório de Evolução Molecular, Aline, Augusto, Carlos André, Claudia, Dânae, Franceli, Jaqueline, Patrícia, Paulo, Valéria e Viviane.

À “turma dos índios”, Jaque, Claudia, Lutécia e Nelson e ao Carlos André, pelo companheirismo, pelas discussões a respeito de trabalho e assuntos extras e, principalmente, pela amizade.

À Renata Schmitt, por realizar quase metade das trabalhosas reações dos seqüenciamentos feitos para este trabalho.

Ao pessoal do Centro de Biologia Genômica e Molecular da Pontifícia Universidade Católica (PUC-RS), que sempre me recebeu com muito carinho e atenção.

Aos colegas de mestrado, André Schnorr, Adriana K. Sassi, Adriano A. Ferreira, Christiane Dresch, Clarisse P. da Silva, Edénir Palmero, Fabiana B. Kohlrausch, Jacqueline M. Cardone, Leonardo Alves Jr., Lindolfo Meirelles, Ligia Tchaika, Maisa

Michels, Marcelo F. Kern, Norma Machado, Tatiana Bressel e Tatiane C. Trigo por resistirem às críticas e apresentarem excelentes resultados no final!!

Ao PPGGBM e a todos os professores que o compõem, pela excelente qualificação aos critérios da Capes, na última avaliação realizada (1998 a 2000).

Ao Elmo Cardoso e à Ellen Mezzeck, sempre tão eficientes e dispostos a quebrar todos os galhos de última hora com a maior paciência e simpatia.

Aos meus amigos e familiares, que mesmo sem entender exatamente o que eu fazia, sempre se mostraram interessados e que ,talvez, tenham se tornado um pouco mais familiarizados com os termos “Genética e Biologia Molecular”!

À minha amiga Bethania, por nossa convivência diária durante os últimos seis anos, sempre dividindo alegrias, tristezas (e contas!). Agradeço também pelo empréstimo irrestrito do computador, que foi de extrema importância durante todos os momentos do mestrado.

Aos guris do padel, Maurício, Nelson, Paulinho e Peter por me fazerem voltar ao mundo dos “esportes com raquete” que sempre fez parte da minha vida e pela “terapia” que os nossos jogos representavam para mim.

Às gurias do “futi”, que eram tantas: valeu pelo companheirismo e pela descontração!

À Clarisse e à Fabi por dividirmos um mesmo sonho que, finalmente, estamos prestes a realizar, depois de tanto planejamento. Muita sorte para nós três!!

Ao meu cunhado , pelo apoio, amizade e interesse que sempre demonstrou em relação aos meus estudos e meus planos em geral.

Ao meu pai, meu anjo da guarda mais fiel. Tenho certeza da sua energia em todos os momentos da minha vida...

À minha mãe e à minha irmã, pela afinidade que temos, que é essencial para mim. Amo incondicionalmente vocês!

As pesquisas em desenvolvimento no laboratório de Evolução Molecular do Departamento de Genética do Instituto de Biociências de nossa universidade são financiadas pelo Programa de Apoio a Núcleos de Excelência (PRONEX), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos (FINEP), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS) e Pró-Reitoria de Pesquisa da Universidade Federal do Rio Grande do Sul (PROPESQ/UFRGS).

INDICE

	Página
Resumo	9
Abstract	10
Introdução	11
1. A genética e o estudo de populações humanas	11
1.1. DNA mitocondrial	11
1.2. Cromossomo Y	13
1.3. Cromossomo X	14
1.4. Marcadores nucleares autossômicos	15
2. Povoamento das Américas	16
2.1. Considerações gerais	16
2.2. Tipos e número de migrações	18
2.3. Origem dos grupos fundadores	20
3. Inserções <i>Alu</i>	21
4. O gene <i>LDLR</i> e o <i>Alu</i> U hipervariável	24
5. Populações estudadas	26
Objetivos	30
Artigo	31
Discussão Geral	58
Referências Bibliográficas	60

RESUMO

As origens das populações e o processo de povoamento das Américas são questões bastante controversas e, por isso, têm sido amplamente estudadas. O gene *LDLR* (low density lipoprotein receptor), que está localizado no braço curto do cromossomo 19 (p13.1-p13.3) e possui 18 éxons, bem como uma porção 3'UTR onde ocorrem dois elementos *Alu* completos (chamados de U e D) e um parcial, foi escolhido para esclarecer esses e outros problemas de evolução e variação genética humana. O objetivo geral deste trabalho foi verificar o que esta região hipervariável, especificamente na sua porção U, nos indicaria sobre a história evolutiva das populações ameríndias. Para este trabalho foram analisadas amostras de DNA de indivíduos nativos da Mongólia (n=24), da Sibéria (n=26), das Américas do Norte (n=11), Central (n=26) e do Sul (n=16), totalizando 14 populações e 103 indivíduos. Essas amostras foram amplificadas pela técnica de PCR (polymerase chain reaction), utilizando-se *primers* específicos para o segmento hipervariável U e, posteriormente, foram elas seqüenciadas automaticamente. Quatorze sítios polimórficos foram encontrados, classificáveis em sete haplótipos, sendo que o sítio 3809 apresentou uma transversão de C para G não descrita na literatura. A estimativa geral de diversidade nucleotídica (π) foi de 0.62%, considerada alta para um marcador autossômico. Verificou-se uma certa uniformidade haplotípica entre a Ásia e a América, mas com a diferenciação maior ocorrendo entre a Mongólia e a América+Sibéria. Não foi detectada diferenciação significativa entre nativos sul e centro-americanos. De um modo geral, o estudo desse marcador confirmou uma provável origem asiática para as populações ameríndias e apoiou a hipótese de uma única onda de migração no processo de povoamento pré-histórico das Américas.

ABSTRACT

Population origins and the process of the peopling of the Americas are controversial questions, and therefore have been subjected to many studies. The *LDLR* (low density lipoprotein receptor) gene, located in the short arm of chromosome 19 (p13.1-p13.3), has 18 exons, as well as, in its 3'UTR portion, two complete *Alu* elements (called U and D) and one partial, and was chosen to clarify these and other problems of human evolution and genetic variation. The general objective of this work was to verify in which ways this hypervariable region, and more specifically its U portion, would give information about the evolutionary history of Amerindian populations. With this purpose DNA samples from Mongolians (n=24), Siberians (n=26), North (n=11), Central (n=26), and South (n=16) American Natives, in a total of 14 populations and 103 individuals, were studied. These samples were amplified by the PCR (polymerase chain reaction) technique using primers that are specific for the U hypervariable segment and afterwards submitted to automatic sequencing. Fourteen polymorphic sites were found, classifiable in seven haplotypes; site 3809 presented a C to G transversion that had not been previously reported in the literature. The general estimate of nucleotide diversity (π) was 0.62%, considered high for an autosome marker. A certain haplotypic uniformity was observed between Asia and America, the greatest differentiation occurring between Mongolian vs. Siberian plus Native American populations. No significant differentiation occurred between Central vs. South American populations. In a general way the study of this marker confirmed an Asian origin for the Native American populations and supported the hypothesis of a single migration wave in the prehistoric peopling of the Americas.

INTRODUÇÃO

1. A genética e o estudo de populações humanas

Os polimorfismos de DNA, em muitos casos, representam um estado transitório, onde um alelo escolhido aleatoriamente está no caminho para a fixação ou eliminação. Na ausência de seleção, a sua variação genética será moldada pela mutação, deriva genética e fluxo gênico, dependendo fortemente de fatores demográficos. Assim, sabendo-se que a variabilidade genômica das populações atuais não depende somente dos genes, há um grande interesse nos estudos sobre a evolução dos humanos modernos e sua diversidade genética em nível mundial (Zietkiewicz et al., 1997).

1.1. DNA mitocondrial

Uma grande quantidade de trabalhos tem utilizado o DNA mitocondrial como ferramenta para estudos sobre a história evolutiva da espécie humana. Essa molécula apresenta uma taxa de mutação diversas vezes mais alta do que a do DNA nuclear simples cópia, possui modo de herança materno, ausência de recombinação e alto número de cópias por célula (Cann et al., 1987). Desta maneira, torna-se possível a determinação de divergências filogenéticas entre linhagens, sem as ambigüidades causadas pelas trocas meióticas que ocorrem entre os genes nucleares (Bailliet et al., 1994).

Quase todos os estudos de evolução humanos baseados no DNA mitocondrial têm sido feitos através de RFLP e seqüenciamento de sua região controladora. Esses dados têm sido utilizados para hipóteses sobre datações, origens, número de migrações da Ásia para as Américas, tamanho das populações migrantes, além da demografia local (Stone e Stoneking, 1998). Porém, essas regiões do DNA mitocondrial nem sempre são suficientes

para proporcionar dados conclusivos sobre eventos evolutivos e inferências filogenéticas (Ingman et al., 2000).

Os primeiros estudos genético-moleculares sobre a origem do homem moderno foram realizados a partir de dados do DNA mitocondrial. Diversas discussões, inclusive com relação aos dados arqueológicos, surgiram desde a formulação da hipótese da “Eva mitocondrial”. Essa hipótese, esquematicamente, sustentava que o DNA mitocondrial de todos os seres humanos atuais teriam tido sua origem a partir de uma única mulher que teria vivido na África, há aproximadamente 200.000 anos. A partir daí surgiu o modelo de uma origem única recente (out-of-Africa) dos seres humanos modernos, em oposição ao modelo de evolução que postulava fluxo gênico generalizado e origem multirregional para os mesmos (Lewin, 1998; Disotell, 1999).

Diversos trabalhos com esta organela (Cann et al., 1987; Vigilant et al., 1991; Ingman et al., 2000; Jorde et al., 2000) concordam na verificação de altos níveis de diversidade genética em populações africanas, quando comparadas às não-africanas. Isto poderia sugerir uma origem africana recente para o *Homo sapiens* moderno. Entretanto, esse fato também poderia ser explicado por um maior tamanho populacional ou uma expansão populacional mais antiga naquele continente (Relethford, 1998; Ingman et al., 2000). Seqüências de DNA mitocondrial indicam uma expansão de humanos modernos em torno de 40.000 anos atrás (Ingman et al., 2000), sendo essa data associada com uma certa mudança no comportamento humano como, por exemplo, a utilização de ferramentas mais avançadas e variadas e o surgimento da arte (Kaessmann et al., 2001).

Ingman et al. (2000) estimaram a idade do ancestral humano mais recente para o DNA mitocondrial, que ficou em torno de 171.500 ± 50.000 anos. Já o cálculo da separação entre africanos e não-africanos foi estimado em cerca de 52.000 ± 27.500 anos. Esses dados foram confirmados em estudo recente realizado após o seqüenciamento de 8,8 kb do DNA mitocondrial (Silva et al., 2002).

Apesar de fornecer muitas informações, o DNA mitocondrial representa um conjunto de ligação único e, por isso, pode não indicar de maneira precisa a história das populações humanas (Stoneking, 1993).

1.2. Cromossomo Y

Estudos de genética de populações e de evolução humana baseados no cromossomo Y têm aumentado durante os últimos anos, devido a descobertas de marcadores úteis nesse cromossomo. Ele apresenta características singulares, como herança paterna e ausência de recombinação na maior parte da sua extensão, o que o torna uma ferramenta poderosa para o traçado e a comparação de linhagens paternas de populações humanas, de modo complementar ao uso do DNA mitocondrial para linhagens maternas (Bosch et al., 1999).

Os trabalhos realizados utilizando-se cromossomos Y são baseados em 2 tipos de marcadores: os marcadores bialélicos (SNPs) e os microssatélites. Os marcadores bialélicos são alterações de uma única base ou pequenas inserções/deleções que ocorrem somente uma vez durante a evolução do cromossomo Y, tornando-os, assim, mais estáveis que os microssatélites (Su et al., 1999). De uma maneira geral, a taxa de evolução dos marcadores deste cromossomo é mais lenta que a do DNA mitocondrial (Lell et al., 1997).

Alguns estudos com cromossomos Y obtiveram resultados coincidentes com os do DNA mitocondrial. Entretanto, alguns dados fornecem interpretações um pouco diferentes. Segundo Karafet et al. (1999), não se deveria esperar uma concordância completa entre sistemas herdados paternalmente ou maternalmente. Isso porque há diferenças significantes quanto às histórias demográficas de homens e mulheres. Uma outra questão é a relativa à idade do ancestral comum mais recente. Baseados em dados do cromossomo Y, Pritchard et al. (1999) estimaram a idade do ancestral comum mais recente (MRCA) entre 45.000 e 90.000 anos, dependendo do modelo de mutação

adotado, além de uma expansão humana recente. Thomson et al. (2000) encontraram uma idade muito semelhante (aproximadamente 59.000 anos), mas substancialmente menor do que a estimada em estudos prévios, onde o grupo de dados era muito menor, tendo sido esta última também muito influenciada pelo modelo de coalescência utilizado.

Os microssatélites exibem variações no número de repetições de pares de base, são amplamente distribuídos pelo genoma e parecem ser seletivamente neutros (Reich e Goldstein, 1998). Eles têm sido usados com sucesso na reconstituição de relações entre populações de uma mesma espécie. Porém, ainda não está claro se eles serão úteis para grupos relacionados mais distantemente (Goldstein et al., 1995a). Alguns estudos concordam com os dados obtidos a partir do DNA mitocondrial, de que os humanos modernos teriam se originado na África e se espalhado para o resto do mundo durante os últimos 150.000 anos, corroborando a teoria “out-of-Africa” de origem do homem moderno.

Kimmel et al. (1998), Reich e Goldstein (1998) e Gonser et al. (2000) também encontraram sinais de expansão populacional nos principais grupos étnicos, confirmando resultados do DNA mitocondrial.

1.3. Cromossomo X

Mais recentemente, estudiosos da evolução de populações humanas também têm se voltado para o uso do cromossomo X como ferramenta molecular. Em uma investigação feita com o último intron do gene ZFX, por exemplo, Jaruzelska et al. (1999) sugeriram um tempo de divergência entre grupos humanos modernos e uma população ancestral comum em torno de 100.000 anos atrás. Uma subsequente expansão teria levado ao isolamento, por distância, entre grupos africanos e populações que colonizaram a Eurásia e o Novo Mundo. Harris e Hey (1999), trabalharam com o gene PDHA1 (piruvato desidrogenase E1), ligado ao X, e estimaram que a subdivisão entre africanos e não-africanos teria se dado há 200.000 anos. Isto sugere uma maior antiguidade para os

fósseis humanos mais antigos, sugerindo que a transformação para o chamado “homem moderno” teria ocorrido em uma população já subdividida. Em contrapartida, Yu e Li (2000) também trabalhando com PDHA1, não confirmaram o resultado acima citado, pois não encontraram diferenças nucleotídicas fixadas entre africanos e não-africanos, sugerindo, então, que a subdivisão entre esses povos não teria sido tão marcante.

A análise de PDHA1 também mostrou baixa variabilidade em populações não-africanas, porém, não forneceu evidências de uma expansão populacional humana significativa, contrariando dados de DNA mitocondrial e de microssatélites.

Uma outra região de interesse está localizada na porção não-codificante do DNA de Xq13.3, a qual apresenta baixas taxas de recombinação e substituição. Essas características permitem que sua história seja reconstruída com menor ambigüidade e com os efeitos de seleção minimizados, devido ao fato de ela estar situada fora dos genes (Kaessman et al., 1999). As análises revelaram alta diversidade tanto na África como na Ásia e um tempo relativo ao MRCA em torno de 535.000 ± 119.000 anos (Kaessmann et al., 1999). Como o cromossomo X tem um tamanho efetivo três vezes maior ao da mitocôndria e do cromossomo Y, espera-se que o MRCA de um loco do cromossomo X seja também três vezes maior (Ingman et al., 2000). Desta forma, essa idade para o MRCA desta região do X concorda com as estimadas para o DNA mitocondrial e o cromossomo Y (100.000-200.000 anos atrás). Através de uma análise de coalescência houve a indicação de uma expansão inicial da população humana há aproximadamente 190.000 anos atrás (Kaessmann et al., 2001).

1.4. Marcadores nucleares autossômicos

O interesse na variação da seqüência de DNA em populações humanas tem aumentado rapidamente. Porém, poucos estudos têm sido realizados considerando regiões não codificantes, especialmente em autossomos (Zhao et al., 2000).

Um dos primeiros trabalhos foi feito com o gene da β -globina (Harding et al., 1997), onde foi estimada uma idade de aproximadamente 800.000 anos para o ancestral humano comum mais recente. Apesar da baixa taxa de mutação do gene da β -globina, esse estudo também inferiu que o homem moderno teria ancestrais tanto africanos como asiáticos, datando-os em mais de 200.000 anos; além disso, sugeriu uma dispersão mais antiga, para fora da África, dos ancestrais dos humanos modernos.

Outros estudos realizados com a região de 10 Kb não codificante do cromossomo 22 (Zhao et al., 2000), região subterminal não codificante do cromossomo 16p13.3 (Alonso e Armour, 2001), região de 10 Kb do cromossomo 1 (Yu et al., 2001), gene do receptor da melanocortina (Rana et al., 1999) apresentaram baixa diversidade nucleotídica em geral. Apesar da maioria dos trabalhos apresentarem alta diversidade genética entre os africanos, ainda existem divergências quanto ao MRCA, à aceitação dos modelos “Out of Africa” ou multirregional de origem do homem moderno e à total substituição ou não das populações arcaicas asiática e europeia pela africana.

2. Povoamento das Américas

2.1. Considerações gerais

Por mais de um século, a questão do povoamento do Novo Mundo tem sido motivo de muito debate. Uma riqueza de dados genéticos, antropológicos, dentais e lingüísticos vem sendo acumulada na área de evolução humana nas Américas. Essas controvérsias dizem respeito à origem, data e ponto da primeira entrada, rotas de expansão, além do número e freqüência das migrações. Mas em um ponto a maioria dos cientistas concorda: os nativos americanos são derivados de ancestrais do Norte da Ásia, que alcançaram o Novo Mundo atravessando o Estreito de Bering, durante a última glaciação (Novick et al., 1998).

Diante de condições adversas, o homem teve que se adaptar para conseguir sobreviver. Com o tempo, a capacidade cultural humana para adaptações aumentou e, por isso, acredita-se que a entrada do homem no Novo Mundo tenha ocorrido a partir do Pleistoceno. É provável que essa entrada tenha ocorrido devido a uma interação entre o aumento das adaptações culturais e o decréscimo no rigor climático (Rogers et al., 1992). Entre as adaptações culturais relevantes destaca-se a utilização do fogo, a caça de grandes animais e a construção de moradias. Sugere-se também, que a extinção da megafauna do Hemisfério Oeste tenha ocorrido com a aparição de predadores humanos no Novo Mundo. Porém, há alguns argumentos contra este postulado (Rogers et al., 1992).

Sugere-se que a cultura Clovis representou os primeiros americanos, que teriam aproveitado o corredor livre de gelo entre as geleiras de Laurentide e Cordillera, tendo acesso ao norte da América a partir do Alasca, enquanto uma ponte de terra entre a Sibéria e o Alasca ainda estava aberta (Nettle, 1999). Haveria também outra rota, ao longo da costa do Pacífico, na América do Norte. Esta tem sido considerada como a mais significativa no processo de ocupação do Novo Mundo, devido às melhores condições ambientais então existentes.

Fortes dados lingüísticos também dão suporte à rota costeira como a principal rota de entrada do homem nas Américas. Centros de diversidade lingüística estariam concentrados na porção costeira do Pacífico e estariam praticamente ausentes nas áreas do interior (Rogers et al., 1992).

Em um estudo sobre a diversidade lingüística, Nichols (1990) sugeriu a existência de 150 padrões lingüísticos entre os indígenas americanos. A autora propõe que essas linhagens lingüísticas se ramificam a uma média constante e que, se todas essas linguagens, com exceção das famílias Na-Dené e Esquimós-Aleuta que foram introduzidas mais tarde, derivam de uma única linhagem, elas teriam começado a se ramificar há pelo menos 50.000 anos atrás, ou seja, mais cedo que a época de Clovis [11.500-11.000 AP; Nettle, 1999].

2.2. Tipos e número de migrações

O tipo e o número de migrações envolvidos no povoamento das Américas foram muito debatidos, tendo sido já propostas migrações simples e múltiplas. As migrações múltiplas foram utilizadas para tentar explicar a diversidade observada em características biológicas, culturais e lingüísticas desses nativos americanos (Rogers et al., 1992). A hipótese do modelo tripartido propõe três ondas de migração, partindo do nordeste da Ásia (Sibéria) e dando origem a três grandes grupos lingüísticos e genéticos, que seriam os Eskimo-Aleuta, Na-Dené e Ameríndio. Esta hipótese sugere que a primeira migração teria ocorrido, aproximadamente, há 15.000 anos, originando os Ameríndios (Paleoíndios), a segunda entre 15.000 e 10.000 anos atrás, originando o grupo Na-Dené e a terceira a 10.000 anos, formando os Eskimo-Aleuta. Esta hipótese recebeu certo apoio de dados dentários e genéticos (Greenberg et al., 1986). Um cenário diferente foi, porém, proposto por Szathmary (1979), que verificou uma íntima relação entre os Na-Dené e os Esquimó e, com isso, sugeriu que a diferenciação do estoque Ameríndio teria ocorrido na Beringia, após o bloqueio do corredor livre de gelo de Alberta. Este bloqueio teria isolado os povos do sul dos do norte (Bonatto e Salzano, 1997a).

Rogers et al. (1992) propuseram modelos de migrações simples e Cavalli-Sforza et al. (1994) propuseram migrações múltiplas. Existe a dúvida se as diferenças observadas entre os nativos americanos são realmente significativas e se as diferenças existentes entre essas populações são suficientes para postular-se mais de uma migração (Laughlin, 1986). A alternativa então, seria propor padrões de diversidade com base na diferenciação através de mecanismos de isolamento dentro das Américas (Rogers et al., 1992). Esse isolamento poderia ser devido às grandes massas de gelo que cobriram boa parte da América do Norte durante a última glaciação e que permitiram somente algumas áreas para habitação. Recentemente, tanto antropólogos físicos quanto arqueólogos vêm questionando seriamente o modelo de migração tripla. Eles sugeriram, com base em estudos da morfologia craniana humana, que populações não-mongolóides também

entraram no Novo Mundo. Esses achados levantaram a possibilidade de uma quarta migração, mais antiga (Dillehay, 1997; Neves et al., 1997, 1999).

Estudos com base no DNA mitocondrial inicialmente sugeriram um modelo de três ou quatro migrações (Torroni et al., 1992, 1993). Esta hipótese propunha que populações que falavam o estoque lingüístico Ameríndio teriam se originado a partir de duas migrações. A primeira, carregando os haplogrupos A, C e D do DNA mitocondrial, teria ocorrido entre 17.000 e 34.000 anos atrás, e a segunda, com portadores das seqüências de mtDNA do haplogrupo B, mais tarde. Uma terceira migração originaria os Na-Dené, postulados como mais recentes por apresentarem diversidade mais baixa do haplogrupo A do que os anteriores. Em contraste, Horai et al. (1993), usando dados de seqüências da região controladora (CR), postularam que cada haplogrupo principal representaria migrações separadas, que teriam ocorrido em torno de 14.000 a 21.000 anos atrás. Em 1994, Lorenz e Smith reafirmaram a hipótese de que os ancestrais dos Ameríndios e Na-Dené deveriam ter chegado ao Novo Mundo por meio de migrações separadas, através da verificação da presença da deleção de 9 pares de base da região V do DNA mitocondrial. Portadores desta deleção teriam aqui chegado depois da migração Ameríndia e antes da migração Na-Dené. O mesmo foi sugerido por Karafet et al. (1999) e Lell et al. (2002) em seus trabalhos com o cromossomo Y.

Wallace (1995) e Starikovskaya et al. (1998) propuseram que a primeira leva de migração à América teria ocorrido aproximadamente entre 26.000-34.000 anos, a dos portadores do haplogrupo B em torno de 12.000 e 15.000 anos e a dos Na-Dené há aproximadamente 7.000 e 9.000 anos atrás.

Estudos recentes com RFLP, a região controladora de mtDNA e o cromossomo Y sugerem, porém, uma única migração, ocorrida provavelmente entre 20.000 e 40.000 anos atrás (Bonatto e Salzano, 1997a, b; Stone e Stoneking, 1998; Santos et al., 1999; Silva et al., 2002). Este modelo postula que teria havido um aumento populacional depois da colonização da Beringia e, em seguida, uma expansão através do corredor livre de gelo de

Alberta (Bonatto e Salzano, 1997a). Um colapso nas placas de gelo Laurentide e Cordillera teria bloqueado o corredor Alberta e isolado a Beringia do resto do continente americano. Esse isolamento explicaria a íntima afinidade entre Na-Dené e Esquimó, a relação distante com os Paleoíndios, e também, o motivo dos grupos Na-Dené e Esquimó apresentarem uma baixa diversidade de DNA mitocondrial. Esta hipótese, que coloca a Beringia como o local de estabelecimento e expansão populacional, foi denominada como “out of Beringia”, em contraste com a tradicional “out of Asia”, em que a Sibéria é considerada a região geográfica de origem das populações nativas americanas e a Beringia é tida como um corredor de passagem (Bonatto e Salzano, 1997a).

Um trabalho recente, também utilizando o DNA mitocondrial, apóia a hipótese do refúgio, que supõe que a população fundadora dos Eskimo-aleuta teria se originado no refúgio Beringia/sudoeste do Alasca durante o início do período pós-glacial (~15.000 anos atrás), não concordando com a hipótese de chegada na costa do Alasca como resultado de onda de migração relativamente recente a partir do interior da Sibéria (Derbeneva et al., 2002).

2.3. Origem dos grupos fundadores

Outro aspecto bastante investigado diz respeito ao local de origem da população fundadora do Novo Mundo. Segundo Merriwether et al. (1996) e Stone e Stoneking (1998), a Mongólia representaria a melhor candidata como local de origem, porque apresenta os quatro haplogrupos de DNA mitocondrial considerados fundadores, ao contrário do que ocorre com as populações geograficamente mais próximas da Sibéria. Kolman et al. (1996) também sugeriram que os mongóis representariam um grupo central no entendimento das relações genéticas e evolucionárias entre as populações indígenas do leste asiático e do Novo Mundo, descartando os povos siberianos como possíveis fundadores por não apresentarem o haplogrupo B, apesar de sua proximidade geográfica com as Américas.

Essa hipótese foi reforçada pela descoberta do haplogrupo X do DNA mitocondrial (Derenko et al., 2001) e pela distribuição de determinados haplótipos do cromossomo Y na região central da Sibéria, na Mongólia e no Lago Baikal (Karafet et al., 1999; Santos et al., 1999; Lell et al., 2002). Além disso, evidências arqueológicas de estabelecimento humano nesta região datam de 20.000 anos atrás, ou seja, bem antes dos humanos terem alcançado o leste da Sibéria (Goebel, 1999). A presença destes haplogrupos de DNA mitocondrial em populações antigas do nordeste da Ásia confirmaria essa região como “terra natal” dos colonizadores americanos pré-históricos (Malhi et al., 2002).

O haplogrupo X representa uma pequena linhagem fundadora, que estaria restrita a grupos ameríndios do norte e, também, a determinadas populações Na-Dené (Navajo, por exemplo). Porém, para alguns autores, o fato deste haplogrupo não ter sido encontrado em asiáticos atuais, incluindo muitos siberianos, sugeriria uma origem européia para alguns nativos americanos (Brown et al., 1998; Malhi e Smith, 2002). Por outro lado, Derenko et al. (2001) detectaram a presença do haplogrupo X em populações Altaicas do sul da Sibéria, onde os outros quatro haplogrupos fundadores dos nativos americanos também estão presentes. Dessa forma, Altai é o único grupo étnico atual que apresenta todos os cinco haplogrupos nativo-americanos clássicos, constituindo-se em boa indicação para os fundadores dos nativos americanos. A confirmação da presença do haplogrupo X na América do Norte antiga foi realizada por Malhi e Smith (2002) a partir de uma amostra pré-histórica escavada no Rio Columbia (EUA), provavelmente derivada de uma população de origem siberiana. Há também sugestões de que ele teria ocorrido em populações pré-históricas da América do Sul (Lobato da Silva et al., 2001).

3. Inserções *Alu*

A inserção de fragmentos móveis de ácidos nucléicos em determinados locais do genoma pode provocar profundos efeitos na estrutura e função dos genes. Sabe-se que

somente uma pequena fração (1%) do genoma humano é necessária para codificar todas as proteínas essenciais e que boa parte dos outros 99% do DNA é formada por elementos repetitivos de diferentes famílias e subfamílias.

Dentro dessa fração de DNA repetitivo encontra-se um grupo de transposons que, como o nome indica, são capazes de se moverem de um local a outro no genoma. Uma das superfamílias de retroposons é composta de retroposons não virais, que podem ser classificados como elementos intercalados curtos (SINE) ou longos (LINE). Dentro da família SINE, as inserções *Alu* são as mais numerosas e melhor caracterizadas no genoma de primatas (Szmulewicz et al., 1998).

Os elementos *Alu* estão presentes em mais de 500.000 cópias por genoma haplóide humano e constituem aproximadamente 5% do seu total. Tiveram sua origem provavelmente a partir do gene 7SL RNA, mobilizando-se através de transcritos derivados da RNA polimerase III (Rogers, 1983). Apesar de os elementos *Alu* terem divergido, eles ainda mantêm suas seqüências aproximadamente 90% similares ao gene 7SL RNA, de 150 pares de base (Szmulewicz et al., 1998). A estrutura básica de um elemento *Alu* mostra uma organização dimérica, na qual o monômero direito está 31 nucleotídeos distante do monômero esquerdo. A metade esquerda do elemento contém um promotor típico de RNA polimerase III, que direciona o início da transcrição (Novick et al., 1995).

As seqüências *Alu* consistem em aproximadamente 280 pares de base de comprimento, são altamente metiladas, ricas em CG e contêm vários códons de parada nas direções senso e antisenso (Mighell et al., 1997). Além disso, apresentam uma seqüência rica em As entre os dois dímeros e uma cauda poli-A a 200 pares de base da extremidade 3' (Szmulewicz et al., 1998).

As seqüências *Alu* dentro do genoma humano podem ser divididas em pelo menos seis grandes subfamílias, com base em mutações diagnósticas (Jurka, 1993; Britten, 1994; Kass et al., 1995). Um dos grupos de elementos *Alu* formado mais recentemente foi

chamado de HS (human-specific) ou PV (predicted variant) e é derivado da subfamília CS. Em paralelo, a subfamília Sb2, independente da linhagem CS, também tem se expandido no genoma humano. Alguns HS e Sb2 (ou Yb8, de acordo com a nomenclatura de Batzer et al., 1996b) sofreram retroposição tão recentemente que ainda não estão totalmente fixados (Batzer et al., 1996a), ou seja, há cromossomos com e sem a inserção. Isto condiciona frequências interpopulacionais diversas. Calcula-se que a idade média de inserção *Alu* da subfamília Yb8 seja de 4,1 milhões de anos e que o elemento *Alu* do gene *LDLR* seja o membro mais antigo dessa subfamília (Hutchinson et al., 1993).

Essas subfamílias são de diferentes idades evolutivas e parecem ter se originado no genoma de primatas em diferentes épocas durante a evolução (Shaikh e Deininger, 1996). As inserções mais antigas datam em torno de 65 milhões de anos, enquanto que a inserção *Alu* HS mais antiga ocorreu, aproximadamente, há 2,8 milhões de anos; por isso, esse tipo de marcador genético permite reconstruções filogenéticas bastante precisas (Novick et al., 1995). Uma vez inseridos em locais específicos do cromossomo, a maioria dos elementos *Alu* não está sujeita à perda ou rearranjo, caracterizando-se, dessa forma, como marcadores genéticos estáveis.

Há várias razões para que essas inserções polimórficas sejam úteis para estudos de genética de populações. Entre elas, está a baixa probabilidade de perda da inserção, incorporada em um evento único na evolução humana; e a existência de uma técnica simples, com base na PCR (polymerase chain reaction), para a avaliação do genótipo individual para cada loco polimórfico. Conhecer o estado ancestral e a “direção” da mutação facilita a análise das relações populacionais, o que não ocorre com outros marcadores.

O fato do ancestral hipotético apresentar frequência zero para a presença da inserção, e a raiz das árvores evolutivas ser formada por populações africanas atuais, sugere uma origem africana para inserções polimórficas *Alu*. Isso também reforça a hipótese de uma origem africana recente para o homem moderno (Batzer et al., 1994).

A questão sobre a possível função dos *Alu* dentro do genoma tem sido alvo de muita discussão. Uma das dúvidas era se a presença dessas inserções afetaria, de alguma maneira, apenas as seqüências vizinhas ou o genoma como um todo. Eles foram inicialmente considerados como meros parasitas do genoma, sem nenhum efeito maior no organismo. Essa percepção era conhecida como a Hipótese do Gene Egoísta (Orgel e Crick, 1980). Esses elementos também eram vistos como fósseis moleculares ou inconseqüentes para a célula e para a evolução das espécies (Novick et al., 1996). Uma nova alternativa para a hipótese do DNA egoísta foi um modelo chamado de Canalização do Genoma (von Sternberg et al., 1992). Nesse processo, todas as mudanças evolutivas deveriam se adaptar ao contexto genético prevalecente, uma vez que o padrão de organização genômico já estaria estabelecido (Novick et al., 1996). Evidências moleculares da estrutura e evolução desta classe de seqüências indicam que elas propiciam uma origem dramática de variação genética e que as inserções *Alu* são responsáveis por grande número de problemas genéticos (Szmulewicz et al., 1998).

Considerando que muitos locos podem ser polimórficos, é notável o potencial dessas inserções para diagnósticos médicos e epidemiológicos, nas ciências forenses, na arqueologia molecular e na identificação humana em geral (Novick et al., 1995). Elas também atuam como organizadores, ativadores e silenciadores da expressão gênica.

4. O gene *LDLR* e o *Alu U* hipervariável

O gene *LDLR* (low density lipoprotein receptor) codifica uma proteína de M_R 95.000, possui seqüências *Alu* presentes em um éxon e também em múltiplos íntrons (Hobbs et al., 1985), está localizado no braço curto do cromossomo 19 (p13.1-p13.3) e possui 18 éxons distribuídos em 45 kb (Goldstein et al., 1995b). Ele é importante no metabolismo dos lipídios (Auwerx et al., 1986) e mutações no mesmo afetam esse processo. Os principais sintomas registrados são falha na remoção de lipoproteínas da circulação, elevado nível de colesterol no sangue e doenças cardíacas precoces (Fass et al.,

1997). Uma das doenças mais comuns causadas por defeito no funcionamento do *LDLR* é a hipercolesterolemia familiar (Brown e Goldstein, 1986; Fass et al., 1997).

Seqüências *Alu* estão presentes na maioria dos 17 íntrons e no último éxon do gene, que codifica a extremidade carboxi-terminal da proteína e também a região 3'UTR do mRNA (Hobbs et al., 1985). Este gene apresenta duas vezes mais inserções *Alu* do que a média do genoma, o que o torna instável (Lehrman et al., 1987). Dados obtidos a partir de *Southern blotting* indicam que as seqüências *Alu* podem estar presentes na região 3'UTR de chimpanzés e gorilas, mas estão ausentes em babuínos, sugerindo que sua inserção tenha ocorrido durante a evolução dos primatas (Hobbs et al., 1985).

Na porção 3'UTR de humanos, de 2,5 kb, ocorrem dois elementos *Alu* completos e um parcial (Yamamoto et al., 1984). Os dois elementos *Alu* completos desta região são chamados de U (upstream) e D (downstream). A primeira seqüência *Alu* (*Alu* U) no mRNA tem todas as características de uma seqüência *Alu* normal: é flanqueada por repetições diretas imperfeitas e por uma cauda poli-A na extremidade 3'. A segunda seqüência (*Alu* parcial) é truncada na extremidade 5' na porção monomérica à direita. A terceira seqüência (*Alu* D) não é flanqueada por repetições diretas, mas possui a cauda rica em As na extremidade 3' (Yamamoto et al., 1984). O *Alu* D e o *Alu* parcial são membros da família Sq. O *Alu* U é mais recente e pertence à família Yb8, tendo surgido após a divergência entre *Homo* e *Pan*, através de conversão gênica, sobrepondo-se ao *Alu* mais antigo (Kass et al., 1995).

Fagundes (2001) verificou na seqüência *Alu* U do gene *LDLR* a maior taxa de mutação por substituição já descrita para locos nucleares, com $\mu_{LDLR} = 0,562$ % por milhão de anos para a região *Alu* 3'UTR em geral e de $\mu_{Alu\ U} = 0,898$ % por milhão de anos para o *Alu* U em particular. Ele também observou uma alta variabilidade genética no continente asiático, ao contrário do que se esperaria, não descartando, dessa forma, uma possível participação de populações deste continente nos primeiros estágios da evolução do homem moderno.

Dois outros resultados interessantes foram obtidos por Fagundes (2001) nas comparações feitas entre ameríndios e não-ameríndios. A primeira, foi a de um valor significativamente positivo para o índice D de Tajima (1989) para a seqüência *Alu* completa, e a segunda, uma diversidade significativamente reduzida para a região *Alu U*. Estes resultados seriam compatíveis com um gargalo de garrafa (“bottleneck”) populacional no povoamento pré-histórico das Américas, como sugerido pelos dados de mtDNA (Bonatto e Salzano, 1997a, b) e do cromossomo Y (Santos et al., 1996; Bianchi et al., 1999), mas não por estudos prévios de DNA nuclear (Bevilaqua et al., 1995; Mattevi et al., 2000).

5. Populações estudadas

5.1. *Khoton*

É um grupo étnico relativamente pequeno e jovem, com aproximadamente 6.000 pessoas (0,3% da população da Mongólia). São nômades, e vivem no noroeste daquele país. Há grande interesse em estudar esta população porque não há consenso quanto à sua origem e, antropologicamente, são considerados muito relacionados a populações da Ásia Central (Munkhbat et al., 1997). Falam a linguagem chinesa Mandarim (Grimes, 1992).

5.2. *Khalkh*

É considerado o maior grupo étnico da Mongólia, com aproximadamente 1,8 milhões de pessoas, correspondendo a 80% do total de habitantes, segundo o censo de 1992 (Munkhbat et al., 1997; Katoh et al., 2002). As amostras utilizadas no presente trabalho são de indivíduos da capital, Ulaanbaatar. O grupo *Khalkh* é considerado ser o descendente direto das tribos do centro da Mongólia (Katoh et al., 2002).

5.3. Chukchi

Os Chukchi falam uma língua com o mesmo nome, classificada entre as paleoasiáticas (Grimes, 1992; Starikovskaya et al., 1998). No final do século 19, os Chukchi aumentaram consideravelmente o seu tamanho populacional e se espalharam amplamente, pressionando populações vizinhas e se espalhando por toda a costa da Península Chukchi (Starikovskaya et al., 1998). É o grupo étnico mais numeroso da península, com aproximadamente 15.000 indivíduos (Erdesz et al., 1994).

5.4. Esquimó siberiano

É um grupo étnico distribuído em diversos pontos da costa do Mar de Bering. Atualmente, existem menos de 1.500 indivíduos nativos, e é comum haver mistura com populações Chukchi que habitam a região costeira (Erdesz et al., 1994). Falam a linguagem Yupik (Grimes, 1992).

5.5. Carrier-Sekani

Os Carrier-Sekani falam uma linguagem Atabasca, do filo lingüístico Na-Dené. O seu nome peculiar deriva do fato de que antigamente uma viúva era obrigada a carregar, por três anos, as cinzas de seu marido. Sua população atual está estimada em cerca de 2.000 indivíduos. As amostras investigadas foram obtidas do bando Nadler, que pertence à Primeira Nação Carrier-Sekani (Swanton, 1952).

5.6. Bribri e Cabecar

São populações costa-riquenhas localizadas em ambos os lados da Cordilheira de Talamanca. As relações entre essas duas tribos têm sido muito intensas tanto a nível sócio-cultural e lingüístico quanto genético. O número de habitantes varia de 1.500 a

7.800 em diferentes reservas indígenas (Barrantes, 1993). Pertencem ao grupo lingüístico Talamanca (Greenberg, 1987).

5.7. *Teribe*

Os Teribe da Costa Rica (cerca de 500 indivíduos) vêm perdendo sua identidade e a população está bastante mesclada com grupos não-indígenas; alguns vivem em localidades bribri, cabecar e boruca (Barrantes, 1993). São classificados no grupo lingüístico Talamanca (Greenberg, 1987).

5.8. *Guatuso*

Historicamente o povo guatuso habitava as planícies do norte da Costa Rica, nos arredores do Rio Frio. Atualmente estão distribuídos em três pequenas localidades, com cerca de 855 pessoas. Pertencem ao grupo lingüístico Rama, apesar de existirem divergências quanto a esta classificação (Barrantes, 1993).

5.9. *Huetar*

Apesar de se constituírem em uma população pequena, bastante aculturada e com muita hibridação com grupos não indígenas, já foram o grupo majoritário na Costa Rica, localizados nas regiões central e leste. Sua classificação lingüística não é muito definida, mas, atualmente, estão agrupados com os Guatuso no grupo lingüístico Rama. Na Costa Rica vivem cerca de 800 indivíduos (Barrantes, 1993).

5.10. *Guaymi*

É o grupo ameríndio mais numeroso atualmente, com cerca de 90.000 indivíduos. Na Costa Rica existem quatro assentamentos, com aproximadamente 2.000 habitantes.

Habitam as zonas altas das cordilheiras, onde conservam um modo de vida disperso e praticam agricultura tradicional (Barrantes, 1993).

5.11. *Caingang*

A tribo Caingang era conhecida antigamente pelo nome de Guaianá. A palavra Caingang foi introduzida por Telemaco M. Borba para designar todos os índios não-Guarani do sul do Brasil (São Paulo, Paraná, Santa Catarina e Rio Grande do Sul). Os Caingang que vivem aqui teriam vindo, talvez, a pouco menos de dois séculos, provenientes de grupos do Paraná e São Paulo (Salzano, 1961). Sua linguagem foi classificada como Gê até quando Loukotka, em 1968, separou-a em um estoque independente (Salzano et al., 1980). Em torno de 15.000 indivíduos, distribuídos em 25 localidades, vivem no sul do país (Ricardo, 1991). Eles têm sido bastante investigados quanto à genética, demografia e morfologia (Salzano et al., 1997).

5.12. *Lengua*

Os Lengua, cuja língua é classificada entre os Macro-Panoas, estão distribuídos ao longo da margem esquerda do Rio Paraguai e suas comunidades localizam-se até 240 km de distância desta margem. Deve-se ter cuidado para não confundir-los com outra tribo de igual nome, porém da família Mascoi (Pericot y Garcia, 1962).

5.13. *Kuben Kran Keng*

É considerado o maior grupo Cayapó, que é classificado dentro da família lingüística Gê. Acredita-se que foram formados a partir da subdivisão dos grupos Gorotire, em 1936. Contatos mais permanentes com não-indígenas ocorreram a partir de 1952 e, ao longo dos anos, eles vêm sofrendo com um processo de aculturação, que, no entanto, não afetou de maneira marcante sua identidade cultural. Em torno de 310 indivíduos vivem próximos ao Rio Riosinho, um afluente do Rio Fresco (Salzano, 1971).

OBJETIVOS

O objetivo geral deste trabalho foi verificar o que a região hipervariável nuclear 3'UTR do gene *LDLR*, e mais especificamente a sua região U, indicaria sobre a história evolutiva das populações ameríndias e afins.

Os objetivos específicos foram:

- 1- Verificar a variabilidade da região especificada em três populações indígenas da América do Sul, seis da América Central, uma da América do Norte, duas da Sibéria e duas da Mongólia.
- 2- Comparar esses resultados com os obtidos previamente, e determinar se a sugestão de uma possível alteração populacional drástica e/ou a ação de fatores seletivos sobre esta região seria confirmada.
- 3- Levando em consideração também outros marcadores, fornecer subsídios sobre o ponto de origem e a posterior diversificação ocorrida ao longo da colonização pré-histórica das Américas.

Intra and Intercontinental Molecular Variability in the 3'UTR of the *LDLR* Gene

A.H. HELLER,¹ F.M. SALZANO,¹ R. BARRANTES,² M. KRYLOV,³ L. BENEVOLENSKAYA,³ F.C. ARNETT,⁴ B. MUNKHBAT,⁵ N. MUNKHTUVSHIN,⁵ K. TSUJI,⁶ M.V. MONSALVE,⁷ D.V. DEVINE,⁷ M.H. HUTZ,¹ F.R. CARNESE,⁸ A.S. GOICOECHEA,⁸ L.B. FREITAS,¹ AND S.L. BONATTO⁹

Artigo a ser submetido para publicação em *Human Biology*

Intra and Intercontinental Molecular Variability in the 3'UTR of the *LDLR* Gene

A.H. HELLER,¹ F.M. SALZANO,¹ R. BARRANTES,² M. KRYLOV,³ L. BENEVOLENSKAYA,³
F.C. ARNETT,⁴ B. MUNKHBAT,⁵ N. MUNKHTUVSHIN,⁵ K. TSUJI,⁶ M.V. MONSALVE,⁷ D.V.
DEVINE,⁷ M.H. HUTZ,¹ F.R. CARNESE,⁸ A.S. GOICOECHEA,⁸ L.B. FREITAS,¹ AND S.L.
BONATTO⁹

KEY WORDS: Low density lipoprotein receptor gene, Amerindians, Mongolians, Siberians

Adress to which proofs should be sent

Francisco M. Salzano

Departamento de Genética, UFRGS

Caixa Postal 15053

91501-970 Porto Alegre, RS

Brazil

E-mail: francisco.salzano@ufrgs.br

¹Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil.

²Sección de Genética Humana, Instituto Nacional de Investigaciones en Salud (INISA), Universidad de Costa Rica, San José, Costa Rica.

³Department of Epidemiology and Genetics of Rheumatic Disease, Institute of Rheumatology of the Russian Academy of Science, Moscow, Russia.

⁴University of Texas Health Science Center, Houston, TX, USA.

⁵Institute of Clinical Pathology, Ulaanbaatar, Mongolia.

⁶Department of Transplantation Immunology, School of Medicine, Tokay University, Kanagawa, Japan.

⁷Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, and Canadian Blood Services, B.C., Canada.

⁸Instituto de Ciencias Antropológicas, Facultad de Filosofía y Letras, Universidad de Buenos Aires, 1406 Buenos Aires, Argentina.

⁹Centro de Biología Genômica e Molecular, Faculdade de Biociências, Pontifícia Universidade Católica do Rio Grande do Sul, Ipiranga 6681, 90610-001 Porto Alegre, RS, Brazil.

Abstract A total of 103 individuals from two Mongolian, two Siberian, and 10 Native American populations were studied in relation to a sequence of 340 pb from the 3'UTR of the *LDLR* gene. Seven haplotypes have been determined, and B1 was the most common, accounting for about half of the sequences found. In general diversity values are quite high, about 2.5x higher than those found in other autosomal Alu sequences. Although no difference in haplotype frequencies was found between Asians and Native Americans, the greatest differentiation was found when we grouped in a single population all Native Americans plus Siberian Eskimos and Chukchi and compared them with Mongolians. This result is compatible with earlier mtDNA and Y chromosome suggestions of a single origin for the first colonizers of the continent. With this nuclear locus it was not possible to distinguish between Central and South American Natives. No evidence for selection or marked demographic changes was obtained with these data.

Genetic diversity is influenced by a large number of factors and the way they interact among themselves. Intrinsic factors are mutation rates and the constraints imposed on them by genome structure and selective pressures. Among the extrinsic agents perhaps the most important is population size, influenced in complex ways by the organic and physical environments. In humans a very important factor is culture, which can act as a buffer system between individuals and their environment or, through technological innovations, lead to bursts of population growth and migration.

The relatively recent development of molecular analyses and bioinformatics offers an unique opportunity for the investigation of human evolutionary history. Several types of scenarios have been inferred based on mitochondrial DNA, Y chromosome, X and autosomal segments. The fact that the pictures obtained were not always congruent may be due, in part, to the nature of the DNA regions investigated (whether coding or non-coding, intra or intergenic, with synonymous or nonsynonymous sites, with uni or biparental inheritance, genic or pseudogenic, with or without *Alu* insertions; selected references in Jorde et al. 2000; Alonso and Armour 2001; Chen and Li 2001).

The problem of the origin and ancient colonization of the New World can be used as an example of the different interpretations that distinct types of genetic evidence can induce. Native Americans have been extensively but unevenly studied (review in Salzano 2002), and questions related to the place(s) from where the first colonizers originated, the number of waves of migrations, the time when these events took place, and the subsequent diversification which occurred along the colonization process remain highly debated (Crawford 1998; Silva et al. 2002; Lell et al. 2002). If other kinds of evidence, besides genetics, are considered the picture becomes even more complex (Jablonski 2002).

It was therefore with great interest that we have considered a paper by Yamamoto et al (1984), which described a peculiar autosomal gene which coded the human low density lipoprotein receptor (LDLR) and contained a 3' untranslated region (3'UTR) which presented multiple copies of the *Alu* family of repetitive DNAs. The gene is located in chromosome 19, and its 3'UTR possess one partial and two complete *Alu* insertions; the latter are distinguished by their position as U (upstream) and D (downstream) (Yamamoto et al. 1984, Kass et al. 1995). Previous studies performed by our group (Fagundes 2001) indicated that the upstream region is hypervariable, and therefore appropriate for the investigation of relatively recent evolutionary events. Therefore we undertook the present investigation, whose purpose was to determine the extent of variability that could be found in this DNA region in different Asian and American populations, trying to obtain a new insight on the genetic diversification process that may have occurred during the ancient colonization of the Americas.

Subjects and Methods

Fourteen Native American, Asian and Siberian populations were studied. Their names, sample sizes, geographical location, and languages they speak are provided in Table 1 and Figure 1. Three populations from South America, six from Central America, one from North America, and four from Asia (including two from Siberia) were investigated.

The samples were collected at random from volunteers in the several communities, as a part of other genetic and epidemiological studies. They were then submit to processing in different laboratories, located in several places, as indicated in the authors' affiliations. DNA extraction was performed according to standard methods, and the material related to 12 of the 14 samples

exceptions: Caingang and Kuben Kran Keng) was sent to Porto Alegre for the present investigation. As for the two other samples, they are part of the extensive Amerindian biological specimens stored in Porto Alegre for research purposes.

The amplification of target fragments was achieved using PCR (Polymerase Chain Reaction) from genomic DNA. Specific primers (F1039: 5' ACTTCAAAGCCGTGATCGTGA 3'; R2008: 5' TGCAACAGTAACACGGCGATT 3') were used to amplify the *Alu* U plus D 950 bp segment. PCR amplification was performed in a final volume of 30 μ l. Reagent concentrations consisted of 0.2 μ M of each primer, 1x PCR buffer (10mM Tris-HCl, 50mM KCl), 0.2 mM dNTP, 1.5mM MgCl₂, 1 Unit of Taq DNA Polymerase (Invitrogen) and mili-Q water completing the final volume. One to three μ l of DNA were used.

The thermal cycling profile consisted of an initial denaturing step of 1 min at 94°C, followed by 40 cycles of 1 min at 94°C, 2 min at 65°C, and 1 min at 72°C. A 5 min 72°C extension step followed the last cycle. Afterwards, three microliters of the PCR products were electrophoresed in 2% agarose gels with 1x TBE (TrisBorato-EDTA), stained with ethidium bromide (EtBr). The amplified DNA fragments were visualized using ultraviolet light. Quantification was performed in agarose gels using a Low DNA MassTM Ladder (Gibco-BRL).

To remove unincorporated PCR primers, 6 μ l PCR products were treated with exonuclease I (0.33 μ l; 10U/ μ l; USB) and alkaline phosphatase (0.33 μ l; 1U/ μ l; USB) and incubated in a thermal cycler (MJ Research) at 37°C for 30 min, followed by enzyme inactivation at 80°C for 15 min.

The 30-50 ng purified PCR product served as template in the sequencing reactions using the sequencing primers F1039 and PS (5' ACGGAGTCTCGCTCTGTCGC 3') (0.25 μ l of each). Sequencing reactions (performed with the DYEnamic ET Dye Terminator kit, Amersham Biosciences) consisted of an initial denaturation at 94 °C for 20'' followed by incubation for 34 cycles at 94°C for 20'', plus 50°C for 15'' and 60°C for 1 min. The reactions were run on a MegaBACE 1000 Amersham Biosciences apparatus following the conditions recommended by the supplier.

The 405 bp of *Alu* U plus its adjacent region sequenced were manually checked using the CHROMAS 1.45 (www.technelysium.com.au/index.html) program. CLUSTAL X 1.81 (Thompson et al. 2001) was used for the final alignment of the sequences. A total of 340 bp (nt3603-nt3942), which could be considered with confidence for all sequences, was finally established for all the analyses. Haplotype frequency differences were tested using the SPSS v. 8 package (SPSS[®] for Windows [™], SPSS Inc.; <http://www.spss.com>) The DnaSP 3.53 (www.bio.ub.es/~julio/DNASP.html) program was used for the calculation of Tajima's (1989) D, Fu and Li's (1993) D and F, and Nei and Li's (1979) π statistics. The haplotype tree was obtained by NETWORK 3.1 (www.fluxus-engineering.com) using the median-joining algorithm (Bandelt et al. 1999). The ARLEQUIN ver. 2000 software was used to perform the Analysis of Molecular Variance (AMOVA, Schneider et al. 2000), and to obtain Ewens' (1972) θ_k nucleotide diversity. θ_k is estimated from the infinite-allele equilibrium relationship between the expected number of alleles (k), the sample size (n) and θ ($\theta=4M\mu$, or 4 times the effective population size multiplied by the mutation rate). A total of 16,000 replications was performed for the evaluation of the results. The

MDIV program was used to estimate divergence times and migration rates (Nielsen and Wakeley 2001).

Results

Table 2 gives information about the variable sites observed in the present study. Five transitions, four transversions, and five deletions were observed. Their patterns can be classified in seven haplotypes, six of which had been already identified by Fagundes (2001). Site 3809 presented however a C→G transversion in a Khoton subject that had not been previously described. Haplotype B1 was the most common in the three geographical regions considered here, accounting for about half of the sequences found. Heterozygosis for at least one site was the most prevalent (60%) pattern observed (data not shown).

A more detailed presentation of the findings is given in Table 3. The four Asian groups surveyed were more evenly sampled, and different haplotype frequencies can be discerned among them. The Khoton presented the lowest B1 (30%) and the highest C3 (75%) prevalences; the Chukchi the highest (21%) A1 frequency; while the Eskimo (together with the Khoton) had the lowest A1 (8%) and the highest B1 (75%) prevalences. Haplotype diversity (indels were not considered) was almost 2x as high in the Khalkh (0.744) as in the Eskimo (0.431). For the Central American populations the sample sizes are much smaller. However, especially remarkable are the high (90%) B1 frequency among the Guatuso, while conversely the Guaymi showed the highest (60%) C1 prevalence. Extreme values of haplotype diversity are found in the Guatuso (0.20) and Teribe (0.82). In contrast, the three South American groups present values that are not much different considering the sample sizes. The π and θ_k values vary widely among populations, but the

standard deviations are very high. Figure 2 displays the median network constructed to evaluate the phylogenetic relationships between haplotypes found in the present study. Two major clades can be distinguished (A; B+C), defined by mutations that occurred in the indicated sites. With the exception of the B3 haplotype, which appears in only one of the three regions delineated (Asia; North+Central America; South America) all the others are present in the three.

Table 4 compares the haplotype frequencies observed here with those obtained by Fagundes (2001). No differences were observed between the two Asian and two South American studies ($\chi^2=15.63$; 7 d.f.; p: 0.07: and 3.91; 5 d.f.; p: 0.61, respectively). Therefore, for the statistical analysis of the content of this table, they have been considered together. There is significant heterogeneity among continents/subcontinents as specified there ($\chi^2=123.49$; 32 d.f.; Monte Carlo estimation based on 10,000 replications: p<0.001) Residual analysis indicated that South Americans show higher (28%-37%) A1 values, as compared to North/Central Americans (8%). Asians show lower (0%-5%) C1, but higher (17%-40%) C3 prevalences than the other groups, and the same is true for the higher C2 (30%) frequency in Africans, as compared to the other samples. Also interesting is the exclusive occurrence of some haplotypes, such as A2, A3, and C3 in Asian-derived populations while the converse is true for the B2 haplotype.

Values for the diversity and neutrality tests, calculated in these same samples, are displayed in Table 5. To make the present data and those of Fagundes (2001) more comparable, the same region (nt3625-nt3942) was evaluated in the two sets. Asian and world nucleotide diversities seem to have been somewhat overestimated in Fagundes (2001) smaller samples, but the two South American estimates overlap if the standard deviations are considered. The African and European numbers are significantly higher than those of North/Central and South American Natives, but larger

samples sizes are needed to confirm this finding. The three indices that evaluate important demographic changes and/or selection (Fu and Li's 1993 D and F; Tajima's 1989 D) did not show values that significantly depart from the null hypothesis (no selection or marked demographic changes); 24 of them were positive and only three negative.

To test different types of structure and partition of variance among the populations, the AMOVA analysis was performed using three geographic and one linguistic arrangement. The results are shown in Table 6. The most general geographical clustering, which includes five groups, can be compared with the different five groups classified by language. While the intrapopulation variability numbers are quite similar (about 93%), there is slightly more variability among geographic groups (5%) than among language groups (2%-3%), with corresponding complementary contributions of the among populations within groups fraction. If just two large categories are established (Mongolians vs. non-Mongolians; Americans vs. Asians), no significant among groups variation is found, although the numbers are somewhat larger for the first (6%-8%) than for the second (negligible values) comparison.

The program MDIV was employed to test if it was possible to use these data to estimate divergence parameters such as the migration rate and divergence time between Central and South America. Its application to pooled data from Central and South America yielded a very high maximum likelihood migration rate between the two groups of populations (14.36 individuals per generation). Probably due to this fact, no maximum likelihood estimation of the divergence time between Central and South America could be obtained with this program.

Discussion

Of the nine haplotypes found so far in this DNA region, four have been found in Asian or Asian-derived populations only and just one (B2) is absent in these populations. This pattern is unexpected, as usually Africans are genetically more diverse than Eurasians (Excoffier 2002). However, this result could be explained by deviations due to the very small African and European sample sizes. On the other hand, all haplotypes, with the exception of A3 and B3, have been observed in Asian, North, Central and South America. Therefore, the origin of these alleles probably predates the American prehistoric colonization and other intercontinental migrations. These results offer some support for the suggestion that the peopling of the New World was not followed by a strong reduction in population size. There are some differences in haplotype frequencies between the populations sampled here, which may be explained by founder effects/genetic drift (processes which may occur in the generally small Amerindian populations), but larger sample sizes are necessary to confirm them. As for the diversity values, the number obtained here for South Americans (0.64) is of the same order of magnitude as that calculated for the whole world sample (0.62), and even slightly higher. These numbers are higher than the averages compiled by Yu et al. (2001) for different genomic regions (autosomal: 0.09; X-linked: 0.08; Y-linked: 0.01), although not as different (but still 2.5x higher) as those given by them for other *Alu* sequences (autosomal: 0.26; X-linked: 0.12). These results corroborate Fagundes' suggestion (2001) that this *Alu* insertion present a very high mutation rate in humans.

Fu and Li's (1993) D and F, and Tajima's (1989) D were, with three exceptions, positive, but statistically non-significant. Values for these same three neutrality statistics for nuclear loci vary considerably in published studies, with most being negative for world samples but few are

statistically significant (Ptak and Przeworski 2002). The positive values for the South and North/Central American Native populations suggest that a slight reduction in population sizes in the groups considered may have occurred in their origin.

As found in most human populations, the intrapopulation is much higher than the interpopulation variability (Ptak and Przeworski 2002). The greatest differentiation in the AMOVA analyses resulted when all Native Americans plus Siberian Eskimo and Chukchi were grouped together and compared with Mongolians. These results may be interpreted as offering support for a single origin of these populations as described for mtDNA (Bonatto and Salzano 1997) and chromosome Y data (Santos et al. 1999). The results of the divergence analysis between Central and South America suggest that, these populations could not be distinguished, possibly due to the relatively slow mutation rate at sequence level of nuclear loci.

Acknowledgments We thank Nelson J. R. Fagundes for information about some of his unpublished findings, as well for his contribution to some of the results; and Sidia M. Callegari-Jacques for statistical help and several suggestions concerning the manuscript. In Brazil, the Fundação Nacional do Índio (FUNAI) gave permission to study the Indians and provided logistic assistance. All community leaders and the subjects of the investigation were adequately informed about the aims of the study and gave their approval, which is also gratefully acknowledged. This program was approved by the Brazilian Ethics National Committee (CONEP, Resolution 123/98). Financial support was provided by Programa de Apoio a Núcleos de Excelência (PRONEX), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Financiadora de Estudos e Projetos

(FINEP), Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), and Pró-Reitoria de Pesquisa da Universidade Federal do Rio Grande do Sul (PROPESQ/UFRGS).

Literature cited

- Alonso, S., and J.A.L. Armour. 2001. A highly variable segment of human subterminal 16p reveals a history of population growth for modern humans outside Africa. *Proc. Natl. Acad. Sci. USA* 98:864-869.
- Bandelt, H-J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16:37-48.
- Bonatto, S.L., and F.M. Salzano. 1997. A single and early migration for the peopling of the Americas supported by mitochondrial sequence data. *Proc. Natl. Acad. Sci. USA* 94:1866-1871.
- Chen, F-C., and W-H. Li. 2001. Genomic divergences between humans and other hominoids and the effective size of the common ancestor of humans and chimpanzees. *Am J. Hum. Genet.* 68:444-456.
- Crawford, M.H. 1998. *The Origins of Native Americans. Evidence from Anthropological Genetics.* Cambridge, UK: Cambridge University Press.
- Excoffier, L. 2002. Human demographic history: refining the recent African origin model. *Curr. Op. Genet. Develop.* 12:675-682.

- Ewens, W.J. 1972. The sampling theory of selectively neutral alleles. *Theor. Popul. Biol.* 3:87-112.
- Fagundes, N.J.R. 2001. Variabilidade na seqüência 3'UTR do gene *LDLR* em diversas populações humanas. Master's Thesis, Federal University of Rio Grande do Sul, Porto Alegre, Brazil.
- Fu, Y-X., and W-H. Li. 1993. Statistical tests of neutrality of mutations. *Genetics* 133:693-709.
- Jablonski, N.G. 2002. *The First Americans: The Pleistocene Colonization of the New World*. San Francisco, CA: California Academy of Sciences.
- Jorde, L.B., W.S. Watkins, M.J. Bamshad et al. 2000. The distribution of human genetic diversity: A comparison of mitochondrial, autosomal and Y-chromosome data. *Am. J. Hum. Genet.* 66:979-988.
- Kass, D. H., M. A. Batzer and P. L. Deininger. 1995 Gene conversion as a secondary mechanism of short interspersed element (SINE) evolution. *Mol. Cell. Biol.* 15:19-25.
- Lell, J.T., R.I. Sukernik, Y.B. Starikovskaya et al. 2002. The dual origin and Siberian affinities of Native American Y chromosome. *Am. J. Hum. Genet.* 70:192-206.
- Nei, M., and W-H. Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76:5269-5273.

- Nielsen, R., and J.W. Wakeley. 2001. Distinguishing migration from isolation: A Markov chain Monte Carlo approach. *Genetics* 158:885-896.
- Ptak, S.E., and M. Przeworski. 2002. Evidence for population growth in humans is confounded by fine-scale population structure. *Trends Genet.*18:559-563.
- Santos, F.R. , A. Pandya, C. Tyler-Smith et al. 1999. The Central Siberian origin for Native American Y-chromosomes. *Am. J. Hum. Genet.* 64:619-628.
- Schneider, S., D. Roessli, and L. Excoffier. 2000. ARLEQUIN ver. 2000: *A Software for Population Genetics Data Analysis*. Geneva, Switzerland: Genetics and Biometry Laboratory, University of Geneva.
- Silva, W.A.Jr., S.L. Bonatto, A.J. Holanda et al. 2002. Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America. *Am. J. Hum. Genet.* 71:187-192.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- Thompson, J.D., D.G. Higgins, and T.J. Gibson. 2001. CLUSTAL X. <ftp://ftp-igbmc.u-strasbg.fr/pub/clustalx> .

Yamamoto, T., C.G. Davis, M.S. Brown et al. 1984. The human LDL receptor: A cysteine-rich protein with multiple *Alu* sequences in its mRNA. *Cell* 39:27-38.

Yu, N., Z. Zhao, Y-X. Fu, N. Sambuughin et al. 2001. Global patterns of human DNA sequence variation in a 10-kb region on chromosome 1. *Mol. Biol. Evol.* 18:214-222.

Table 1. Sample Size, Geographic Location and Linguistic Information about the Populations Considered ^a

<i>Population</i>	<i>n</i>	<i>Geographic location</i>	<i>Linguistic group</i>
Mongolians			
1. Khalkh	11	46°N – 106°E	Altaic
2. Khoton	13	45°N – 94°E	Qotong
Siberians			
3. Chukchi	14	64°N – 175°W	Chukot
4. Eskimo	12	64°N – 175°W	Yupic
North Americans			
5. Carrier-Sekani	11	52°20'N – 123°15'W	Athapaskan
Central Americans			
6. Bribri	5	9°38'N – 82°50'W	Talamanca Bribri
7. Cabecar	4	9°26'N – 83°09'W	Talamanca Cabecar
8. Teribe	4	9°20'N – 82°35'W	Talamanca Tiribi
9. Guatuso	5	10°40'N – 84°49'W	Rama Guatuso
10. Huetar	3	9°53'N – 84°14'W	Rama Huetar
11. Guaymi	5	8°13'N – 82°57'W	Guaymi
South Americans			
12. Caingang	8	27°S – 51° to 54°W	Ge-Kaingan
13. Lengua	6	23°S – 56°W	Macro-Panoan
14. Kuben Kran Keng	2	8°10'S – 58°8'W	Ge-Kaingan

a. Population numbering is the same in this table and in Figure 1

Table 2. Variability in Relation to the Reference Described for the *Alu U LDLR* Region by Yamamoto et al (1984)^a

<i>Haplotypes</i> ^b	<i>Variable sites</i>														<i>Continental counts</i> ^c		
	3 7	3 7	3 8	3 8	3 8	3 8	3 8	3 8	3 8	3 8	3 8	3 9	3 9	3 9	3 9	AS	NA
	1	5	0	0	0	1	1	2	4	5	2	2	3	6	AS	+	SA
	8	1	0	1	9	8	9	0	6	2	0	4	8	0		CA	
Reference	C	G	A	G	C	A	C	T	C	T	C	C	T	T			
A1	-	A	.	.	.	-	-	-	.	C	G	.	.	C	13	6	9
A2	-	A	.	.	.	-	-	-	.	C	G	A	.	C	9	4	2
B1	-	53	39	15
B3	-	.	.	.	G	1	0	0
C1	-	.	-	C	C	5	17	4
C2	-	.	-	C	C	.	.	C	.	2	2	1
C3	-	.	-	C	T	C	17	6	1

a. Dots indicate an identical character. Dashes indicate deletions.

b. Haplotypes A3 and B2, described by Fagundes (2001), were not observed in the present study.

c. Abbreviations: AS, Asiatics; NA+ CA, North Americans + Central Americans; SA, South Americans.

Table 3. *Alu* U *LDLR* Haplotypes Frequencies and Diversity Measures Estimated for 14 Populations

<i>Population</i>	<i>N° of chromosomes</i>	<i>Haplotype frequencies (%)</i>							<i>N° of polymorphic sites</i>	<i>Haplotype Diversity ± SD</i>	$\pi \pm SD$	$\theta_k \pm SD$
		A1	A2	B1	B3	C1	C2	C3				
Asian												
Khalkh	22	13	14	45	0	5	0	23	7	0.74 ± 0.07	2.63 ± 1.63	1.71 ± 1.36
Khoton	26	8	0	30	4	4	8	46	8	0.71 ± 0.06	2.01 ± 1.31	2.13 ± 1.51
Chukchi	28	21	11	61	0	7	0	0	6	0.59 ± 0.09	1.93 ± 1.26	1.03 ± 0.89
Eskimo	24	8	13	75	0	4	0	0	6	0.43 ± 0.12	1.70 ± 1.15	1.10 ± 0.98
North Americans												
Carrier-Sekani	22	9	9	45	0	18	5	14	8	0.76 ± 0.08	2.31 ± 1.47	2.36 ± 1.73
Central Americans												
Bribri	10	10	10	50	0	20	0	10	7	0.76 ± 0.13	2.33 ± 1.57	3.30 ± 3.44

Table 3. (cont.)

<i>Population</i>	<i>N° of chromosomes</i>	<i>Haplotype frequencies (%)</i>							<i>N° of polymorphic sites</i>	<i>Haplotype Diversity ± SD</i>	$\pi \pm SD$	$\theta_k \pm SD$
		A1	A2	B1	B3	C1	C2	C3				
Cabecar	8	0	0	62	0	25	0	13	3	0.61 ± 0.16	1.32 ± 1.04	1.25 ± 1.64
Teribe	8	25	0	37	0	25	0	13	6	0.82 ± 0.10	2.18 ± 1.53	2.50 ± 3.03
Guatuso	10	0	0	90	0	10	0	0	2	0.20 ± 0.15	0.40 ± 0.46	0.43 ± 0.69
Huetar	6	17	17	50	0	0	16	0	7	0.80 ± 0.17	3.00 ± 2.09	4.06 ± 6.15
Guaymi	10	0	0	40	0	60	0	0	2	0.53 ± 0.09	1.07 ± 0.87	0.43 ± 0.69
South Americans												
Caingang	16	31	7	50	0	0	6	6	8	0.68 ± 0.09	2.27 ± 1.47	2.09 ± 1.79
Lengua	12	25	0	42	0	33	0	0	5	0.71 ± 0.07	1.83 ± 1.27	0.93 ± 1.10
Kuben Kran Keng	4	25	25	50	0	0	0	0	5	0.83 ± 0.22	3.00 ± 2.34	3.76 ± 7.38

Table 4. *Alu* U *LDLR* Haplotype Frequencies by Continent

<i>Population</i>	<i>N° of chromosomes</i>	<i>Haplotypes (%)</i>								
		A1	A2	A3	B1	B2	B3	C1	C2	C3
African ^a	10	30	0	0	20	0	0	20	30	0
European ^a	10	50	0	0	30	10	0	0	10	0
Asian ^a	10	20	10	10	20	0	0	0	0	40
Asian ^b	100	13	9	0	53	0	1	5	2	17
North American/ Central American ^b	74	8	5	0	53	0	0	23	3	8
South American ^a	72	37	7	0	38	0	0	17	0	1
South American ^b	32	28	6	0	47	0	0	13	3	3

a. Fagundes (2001); b. Present study

Table 5. *Alu* U *LDLR* Diversity and Neutrality Tests in Continental Samples

<i>Population</i>	<i>Number of chromosomes</i>	<i>Polymorphic sites</i>	<i>N° of haplotypes</i>	$\pi \pm SD$ (%)	<i>Fu and Li's D^a</i>	<i>Fu and Li's F^a</i>	<i>Tajima's D^a</i>
African ^b	10	10 (6)	4	0.88 ± 0.01	1.35	1.49	1.28
European ^b	10	11 (7)	4	0.89 ± 0.01	-0.13	0.05	0.56
Asian ^b	10	13 (8)	5	1.08 ± 0.01	0.51	0.69	0.91
Asian ^c	100	14 (8)	7	0.66 ± 0.05	0.45	0.69	0.83
North American / Central American ^c	74	13 (7)	6	0.56 ± 0.05	1.21	1.16	0.52
South American ^b	72	12 (7)	5	0.77 ± 0.02	0.38	0.97	1.70
South American ^c	32	13 (7)	6	0.64 ± 0.06	-0.14	0.04	0.45
World ^b	102	15 (10)	8	0.84 ± 0.02	-0.04	0.35	0.92
World ^c	206	14 (8)	7	0.63 ± 0.03	0.32	0.66	0.97

a. $P > 0.10$ for all values; b. Fagundes (2001) and unpublished; c. Present study. Numbers in parentheses indicate the number of polymorphic sites if indels are not included.

Table 6. AMOVA Results. Intra and Interpopulation Variability Considering Different Hierarchical Arrangements

Comparison	Indices ^a	Variation (%)		
		Among groups	Among populations within groups	Within populations
Geographical	F_{st}	4.71	2.31	92.98
clustering 1 ^b	Φ_{st}	4.68	1.79	93.53
Geographical	F_{st}	8.09	2.91	89.00
clustering 2 ^c	Φ_{st}	6.04	3.26	90.69
Geographical	F_{st}	0.47	6.14	93.39
clustering 3 ^d	Φ_{st}	- 1.30	6.61	94.69
Classification	F_{st}	3.08	3.88	93.04
by language ^e	Φ_{st}	2.22	4.03	93.75

a. F_{st} is the classical fixation index (Nei, 1987), while Φ_{st} refers to the population pairwise genetic distances.

b. Five groups: South Americans; Central Americans; North Americans; Chukchi + Eskimo; Mongolia.

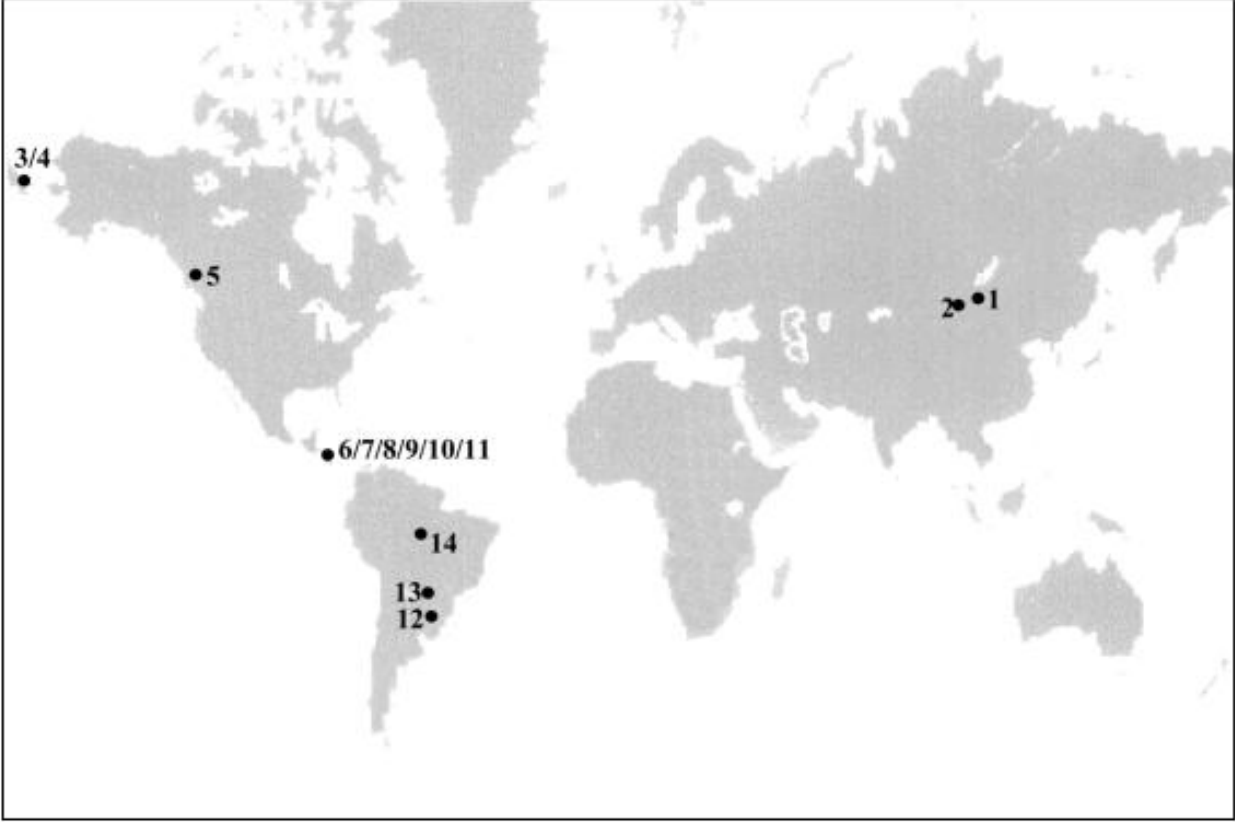
c. Two groups: North + Central + South Americans + Chukchi + Eskimo; Mongolia.

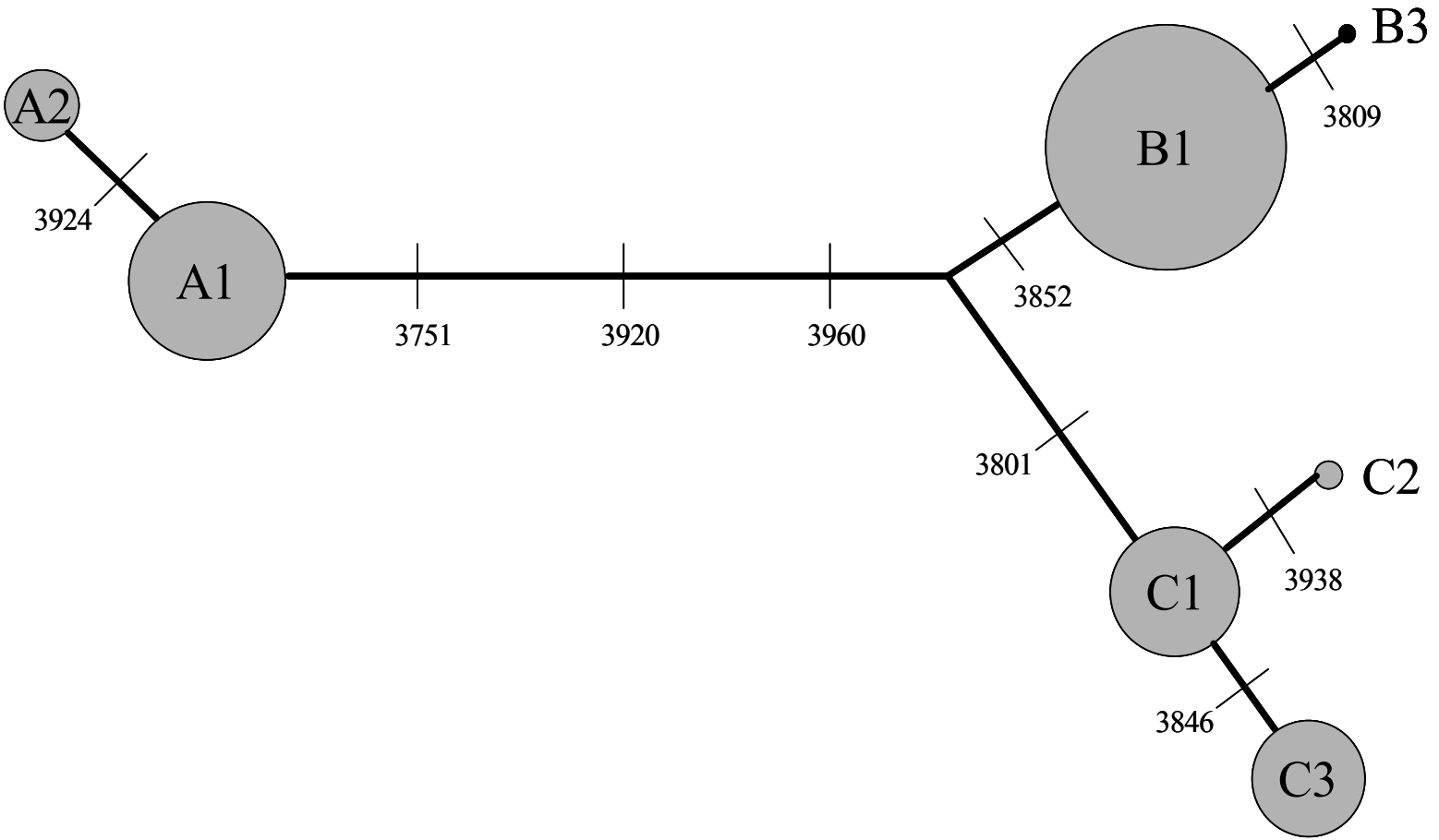
d. Two groups: Americans; Asians.

e. Five groups: Amerind; Na-Dené; Yupic; Chukot; Altaic + Qotong.

Figure 1. Geographical location of the populations sampled. Numbers refer to the populations listed in Table 1.

Figure 2. Median-joining network of the region considered. The sizes of the circles are proportional to their absolute frequencies. The two colors indicate the geographic region in which that haplotype was detected. Gray circles represent sequences observed in the three continental regions (Asia; North+Central America; South America). The black circle represents the haplotype observed in Asia only. Numbers indicate the sites in which mutations occurred, leading to the different haplotypes.





DISCUSSÃO GERAL

Dados genéticos e moleculares, tanto do DNA de indivíduos atuais como de fósseis e múmias, têm se tornado de extrema importância para estudos sobre a história evolutiva humana. Esses dados nos dão informações não só sobre as relações entre as populações, como também sobre o tamanho dessas populações no passado. Análises sobre o nível e o padrão de variação genética, combinados com outros, de origem não-genética, estão contribuindo para o entendimento da nossa evolução.

Apesar de já existir disponível uma boa quantidade de dados, eles não estão distribuídos de maneira uniforme ao longo do planeta. O DNA mitocondrial, cromossomo Y e diversos marcadores autossômicos têm sido estudados de maneira desigual, mas os resultados vêm se acumulando, podendo-se prever o fornecimento de uma boa descrição da estrutura da variação humana em populações e regiões genômicas específicas num futuro, não muito afastado (Przeworski et al 2000).

No presente estudo observou-se uma certa uniformidade haplotípica entre as populações, mas uma alta variabilidade genética intraespecífica, sendo esta a mais alta entre os marcadores autossômicos já analisados.

Diante deste objetivo de esclarecer pontos obscuros da nossa história, muitas atenções estão voltadas para um melhor entendimento sobre a origem do homem americano e questões afins como a idade e o número de migrações da Ásia para a América e o tamanho da população ancestral. Para alguns desses itens não foi possível obter-se respostas mais conclusivas, mas a variação presente no gene *LDLR* corroborou uma provável origem asiática para as populações americanas. Além disso, a pouca diferenciação observada entre os Nativos Americanos sugere que eles tenham uma idade de diversificação semelhante e compartilhem uma história comum, apoiando a hipótese de uma única onda de migração para o povoamento das Américas.

O gene *LDLR* e sua porção *Alu U* mostraram alta variabilidade nucleotídica, ao contrário do que se esperaria para um loco autossômico. Todos os haplótipos, exceto B3, foram encontrados em todos os sub-continentes, indicando formação antes das migrações intercontinentais. A região do *Alu U* mostrou desvios não significativos nos testes de neutralidade realizados. Tal resultado indica que ela pode ser um marcador importante na análise da história humana, em contraposição a regiões codificantes, que refletem principalmente a história da seqüência considerada (Zhao et al, 2000). Outra particularidade encontrada no sistema estudado foi a maior variabilidade (5% a 8%) entre grupos formados de acordo com a distribuição geográfica do que entre grupos de populações formados de acordo com a linguagem (variação entre grupos negligenciável).

Para a elucidação completa da história humana são necessários estudos de muitos marcadores, analisados de maneira uniforme ao longo dos continentes. Embora no momento este objetivo seja difícil de ser alcançado, a integração de informações sobre regiões variáveis como a aqui considerada com outras, em todos as porções do genoma, e destes dados com fatores não-genéticos (ambiente físico e biótico, aspectos sócio-culturais, no tempo e no espaço) contribuirão para esclarecer o enigma da nossa história evolutiva.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALONSO, S; ARMOUR, J.A.L. 2001. A highly variable segment of human subterminal 16p reveals a history of population growth for modern humans outside Africa. *Proc. Natl. Acad. Sci. USA* 98:864-869.
- AUWERX, J.; DE LOOF, H.; VERHOEVEN, G. 1986. *Alu* sequences in the LDL receptor messenger RNA. *Nature* 321:733-734.
- BAILLIET, G.; ROTHHAMMER, F.; CARNESSE, F.R.; BRAVI, C.M.; BIANCHI, N.O. 1994. Founder mitochondrial haplotypes in Amerindian populations. *Am. J. Hum. Genet.* 54:27-33.
- BARRANTES, R. 1993. Evolución en el trópico: los Amerindios de Costa Rica y Panamá. Editorial de la Universidad de Costa Rica, San José.
- BATZER, M.A.; STONEKING, M.; ALEGRIA-HATMAN, M.; BAZAN, H.; KASS, D.H.; SHAIKH, T.H.; NOVICK, G.E.; IOANNOU, P.A.; SCHEER, W.D.; HERRERA, R.J.; DEININGER, P. 1994. African origin of human-specific polyorphic *Alu* insertions. *Proc. Natl. Acad. Sci. USA* 91:12288-12292.
- BATZER, M.A.; ARCOT, S.S.; PHINNEY, J.W.; ALEGRIA-HARTMAN, M.; KASS, D.H.; MILLIGA, S.M.; KIMPTON, C.; GILL, P.; HOCHMEISTER, M.; IOANNOU, P.A.; HERRERA, R.J.; BOUDREAU, D.A.; SCHEER, W.D.; KEATS, B.J.B.; DEININGER, P.L.; STONEKING, M. 1996a. Genetic variation of recent *Alu* insertions in human populations. *J. Mol. Evol.* 42:22-29.

-
- BATZER, M.A.; DEININGER, P.L.; HELLMAN-BLUMBERG, U; JURKA, J.; LABUDA, D.; RUBIN, C.M.; SCHMID, C.W.; ZIETKIEWICZ, E.; ZUCKERKANDL, E. 1996b. Standardized nomenclature for *Alu* repeats. *J. Mol. Evol.* 42:3-6.
- BEVILAQUA, L.R.M.; MATTEVI, V.S.; EWALD, G.M.; SALZANO, F.M.; COIMBRA, C.E.A.; SANTOS, RV.; HUTZ, M.H. 1995. Beta-globin gene cluster haplotype distribution in five Brazilian Indian tribes. *Am. J. Phys. Anthropol.* 98:395-401.
- BIANCHI, N.O.; CATANESI, C.E.; BAILLIET, G.; MARTINEZ-MARIGNAC, V.L.; BRAVI, C.M.; VIDAL-RIOJA, L.B.; HERRERA, R.J.; LOPEZ-CAMELO, J.S. 1999. Characterization of ancestral and derived Y-chromosome haplotypes of New World native populations. *Am. J. Hum. Genet.* 63:1862-1871.
- BONATTO, S.L.; SALZANO, F.M. 1997a. A single and early migrations for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc. Natl. Acad. Sci. USA* 94:1866-1871.
- BONATTO, S.L.; SALZANO, F.M. 1997b. Diversity and age of the four major mtDNA haplogroups and their implications for the peopling of the New World. *Am. J. Hum. Genet.* 61:1413-1423.
- BOSCH, E.; CALAFELL, F.; SANTOS, F.R.; PÉREZ-LEZAUN, A; COMAS, D.; BENCHEMSI, N.; TYLER-SMITH, C.; BERTRANPETIT, J. 1999. Variation in short tandem repeats is deeply structured by genetic background on the human Y chromosome. *Am. J. Hum. Genet.* 65:1623-1638.

-
- BRITTEN, R.J. 1994. Evidence that most human *Alu* sequences were inserted in a process that ceased about 30 million years ago. *Proc. Natl. Acad. Sci. USA* 91:6148-6150.
- BROWN, M.S.; GOLDSTEIN, J.L. 1986. A receptor-mediated pathway for cholesterol homeostasis. *Science* 232:34-47.
- BROWN, M.D.; HOSSEINI, S.H.; TORRONI, A.; BANDELT, H.-J.; ALLEN, J.C.; SCHURR, T.G.; SCOZZARI, R.; CRUCIANI, F.; WALLACE, D. 1998. mtDNA haplogroup X: an ancient link between Europe/Western Asia and North America? *Am. J. Hum. Genet.* 63:1852-1861.
- CANN, R.L.; STONEKING, M.; WILSON, A. C. 1987. Mitochondrial DNA and human evolution. *Nature* 325:31-36.
- CAVALLI-SFORZA, L.L.; MENOZZI, P.; PIAZZA, A. 1994. The history and geography of human genes. Princeton University Press, Princeton.
- DERBENEVA, O. A.; SUKERNIK, R.I.; VOLODKO, N.V.; HOSSEINI, S.H.; LOTT, M.T.; WALLACE, D.C. 2002. Analysis of mitochondrial DNA diversity in the Aleuts of the Commander Islands and its implications for the genetic history of Beringia. *Am. J. Hum. Genet.* 71:415-421.
- DERENKO, M.V.; GRZYBOWSKI, T.; MALYARCHUK, B.A.; CZARNY, J.; MISCICKA-SLIWKA, D.; ZAKHAROV, I.A. 2001. The presence of mitochondrial haplogroup X in Altaians from South Siberia. *Am. J. Hum. Genet.* 69:237-241.

-
- DILLEHAY, T.D. 1997. Onde estão os remanescentes ósseos humanos do final do Pleistoceno? Problemas e perspectivas na procura dos primeiros americanos. *Revista USP*. São Paulo 34:22-33.
- DISOTELL, T.R. 1999. Origins of modern humans still look recent. *Curr. Biol.* 9:R647-R650.
- ERDESZ, S.; SHUBIN, S.V.; SHOCH, B.P.; KRYLOV, M.Y.; MYLOV, N.M.; CHEKALINA, N.A.; BENEVOLENSKAYA, L.I. 1994. Spondyloarthropathies in Circumpolar populations of Chukotcha (Eskimos and Chukchi): epidemiology and clinical characteristics. *J. Rheumatol.* 21:1101-1104.
- FAGUNDES, N.J.R. 2001. Variabilidade na seqüência 3'UTR do gene LDLR em diversas populações humanas. Dissertação de Mestrado. Curso de Pós-Graduação em Genética e Biologia Molecular, UFRGS, Porto Alegre.
- FASS, D.; BLACKLOW, S.; KIM, P.S.; BERGER, J.M. 1997. Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. *Nature* 388:691-693.
- GOEBEL, T. 1999. Pleistocene human colonization of Siberia and peopling of the Americas: an ecological approach. *Evol. Anthropol.* 8:208-227.
- GOLDSTEIN, D.B.; LINARES, A R.; CAVALLI-SFORZA, L.L.; FELDMAN, M.W. 1995a. Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* 92:6723-6727.

-
- GOLDSTEIN, J.; HOBBS, H.; BROWN, M. 1995b. Familial hypercolesterolemia. IN: SCRIVER, C.R.; BEAUDET, AL.; SLY, N.S.; VALLE, D.(eds.) The metabolic basis of inherited diseases. McGraw-Hill, New York. pp 1981-2030.
- GONSER, R.; DONNELLY, P.; NICHOLSON, G.; DI RIENZO, A. 2000. Microsatellite mutations and inferences about human demography. *Genetics* 154:1793-1807.
- GREENBERG, J.H. 1987. Language in the Americas. Stanford University Press: Stanford.
- GREENBERG, J.H.; TURNER, C.G.; ZEGURA, S.L. 1986. The settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Curr. Anthropol.* 27:477-497.
- GRIMES, B.F. 1992. Ethnologue. Languages of the world. Summer Institute of Linguistics, Dallas.
- HARDING, R. M.; FULLETON, S.M.; GRIFFITHS, R.C.; BOND, J.; COX, M.J.; SCHNEIDER, J.A; MOULIN, D.S.; CLEGG, J.B. 1997. Archaic African and Asian lineages in the genetic ancestry of modern humans. *Am. J. Hum. Genet.* 60:772-789.
- HOBBS, H.H.; LEHRMAN, M. A.; YAMAMOTO, T.; RUSSEL, D.W. 1985. Polymorphism and evolution of Alu sequences in the human low density lipoprotein receptor gene. *Proc. Natl. Acad. Sci. USA* 82:7651-7655.
- HORAI, S.; KONDO, R.; NAKAGAWA-HATTORE, Y.; HAYASHI, S.; SONODA, S.; TAJIMA, K. 1993. Peopling of the Americas, founded by four major lineages of mitochondrial DNA. *Mol. Biol. Evol.* 10: 23-47.

-
- INGMAN, M.; KAESSMAN, H.; PÄÄBO, S.; GYLLENSTEN, U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature* 408:708-713.
- JARUZELSKA, J; ZIETKIEWICZ, E.; BATZER, M.; COLE, D. E. C.; MOISAN, J-P.; SCOZZARI, R; TAVARÉ, S.; LABUDA, D. 1999. Spatial and temporal distribution of the neutral polymorphisms in the last ZFX intron: analysis of the haplotype structure and genealogy. *Genetics* 152:1091-1101.
- JORDE, L.B.; WATKINS, W.S.; BAMSHAD, M.J.; DIXON, M.E.; RICKER, C.E.; SEIELSTAD, M.T.; BATZER, M.A. 2000. The distribution of human genetic diversity: a comparison of mitochondrial, autosomal and Y-chromosome data. *Am. J. Hum. Genet.* 66:979-988.
- JURKA, J. 1993. A new subfamily of recently retroposed human *Alu* repeats. *Nucl. Ac. Res.*, 21:2252.
- KAESSMANN, H.; HEBIG, F.; von HAESELER, A.; PÄÄBO, S. 1999. DNA sequence variation in a non-coding region of low recombination on the X chromosome. *Nature* 22:78-81.
- KAESSMANN, H.; WIEBE, V.; WEISS, G.; PÄÄBO, S. 2001. Great ape DNA sequences reveal a reduced diversity and an expansion in humans. *Nature* 27:155-156.
- KARAFET, T.M.; ZEGURA, S.L.; POSUKH, O.; OSIPOVA, L.; BERGEN, A; LONG, J.; GOLDMAN, D.; KLITZ, W.; HARIHARA, S.; de KNIJFF, P.; WIEBE, V.; GRIFFITHS, R. C.; TEMPLETON, A. R.; HAMMER, M.F. 1999. Ancestral Asian source(s) of New World Y-chromosome founder haplotypes. *Am. J. Hum. Genet.* 64:817-831.
-

-
- KASS, D.H.; BATZER, M.A; DEININGER, P.L. 1995. Gene conversion as a secondary mechanism of short interspersed element (SINE) evolution. *Mol. Cell Biol.* 15:19-25.
- KATOH, T.; MANO, S.; IKUTA, T.; MUNKHBAT, B.; TOUNAI, K.; ANDO, H.; MUNKHTUVSHIN, N.; IMANISHI, T.; INOKO, H.; TAMIYA, G. 2002. Genetic isolates in East Asia: a study of linkage disequilibrium in the X chromosome. *Am J. Hum. Genet.* 71:395-400.
- KIMMEL, M.; CHAKRABORTY, R.; KING, J.P.; BAMSHAD, M.; WATKINS, W. S. 1998. Signatures of population expansion in microsatellite repeat data. *Genetics* 148:1921-1930.
- KOLMAN, C.J.; SAMBUUGHIN, N.; BERMINGHAN, E. 1996. Mitochondrial DNA analysis of Mongolian populations and implications for the origin of New World founders. *Genetics* 142:1321-1334.
- LAUGHLIN, W.S. 1986. Comments on the settlement of the Americas: a comparison of the linguistic, dental and genetic evidence. *Curr. Anthropol.* 27:489-490.
- LEHRMAN, M.A; RUSSEL, D.W.; GOLDSTEIN, J.L.; BROWN, M. S. 1987. *Alu-Alu* recombination deletes splice acceptor sites and produces secreted low density lipoprotein in a subject with familial hypercholesterolemia. *Biol. Chem.* 22:3354-3361.
- LELL, J.T.; BROWN, M.D.; SCHURR, T.G.; SUKERNIK, R.I.; STARIKOVSKAYA, Y.B.; TORRONI, A.; MOORE, L.G.; TROUP, G.M.; WALLACE, D.C. 1997. Y chromosome polymorphisms in Native American and Siberian populations: identification of Native American Y chromosome haplotypes. *Hum. Genet.* 100:536-543.

- LELL, J.T.; SUKERNIK, R.I.; STARIKOVSKAYA, Y.B.; SU, B.; JIN, L.; SCHURR, T.G.; UNDERHILL, P.A.; WALLACE, D.C. 2002. The dual origin and Siberian affinities of Native American Y chromosome. *Am. J. Hum. Biol.* 70:192-206.
- LEWIN, R. 1998. Principles of human evolution. A core textbook. Blackwell, Oxford.
- LORENZ, J.G.; SMITH, D.G. 1994. Distribution of the 9-pb mitochondrial DNA region V deletion among North American Indians. *Hum. Biol.* 66:777-788.
- LOBATO DA SILVA, D.F.; RIBEIRO-DOS-SANTOS, A.K.C.; SANTOS, S.E.B. 2001. Diversidade genética de populações humanas na Amazônia. Em: VIEIRA, I.C.G.; SILVA, J.M.C.; OREN, D.C.; D'INCAO, M.A. (orgs.) Diversidade humana e cultural na Amazônia. Museu Paraense Emílio Goeldi, Belém, pp167-193.
- MALHI, R.S.; SMITH, D.G. 2002. Haplogroup X confirmed in prehistoric North America. *Am. J. Phys. Anthrop.* 119:84-86.
- MALHI, R.S.; ESHLEMAN, J.A.; GREENBERG, J.A.; WEISS, D.A.; SHOOK, B.A.S.; KAESTLE, F.A.; LORENZ, J.G.; KEMP, B.M.; JOHNSON, J.R.; SMITH, D.G. 2002. The structure of diversity within New World mitochondrial DNA haplogroups: implications for the prehistory of North America. *Am. J. Hum. Biol.* 70:905-919.
- MATTEVI, V.S.; FIEGENBAUM, M.; SALZANO, F.M.; WEISS, K.M.; MOORE, J.; MONSALVE, M.V.; DEVINE, D.V.; HUTZ, M.H. 2000. Beta-globin gene cluster haplotypes in two North American indigenous populations. *Am. J. Phys. Anthrop.* 112:311-317.
- MIGHELL, A J.; MARKHAM, A.F.; ROBINSON, P.A. 1997. *Alu* sequences. *FEBS Letters* 417:1-5.

-
- MUNKHBAT, B.; SATO, T.; HAGIHARA, M.; SATO, K.; KIMURA, A.; MUNKHTUVSHIN, K.; TSUJI, K. 1997. Molecular analysis of HLA polymorphism in Khoton- Mongolians. *Tissue Antigens* 50:124-134.
- NETTLE, D. 1999. Linguistic diversity of the Americas can be reconciled with a recent colonization. *Proc. Natl. Acad. Sci. USA* 96:3325-3329.
- NEVES, W.; ZANINI, M.C.; MUNFORD, D.; PUCCIARELLI, H.M. 1997. O povoamento da América à luz da morfologia craniana. *Revista USP, São Paulo* 34:96-105.
- NEVES, W.A.; MUNFORD, D.; ZANINI, M.C.; PUCCIARELLI, H.M. 1999. Cranial morphological variation in South America and the colonization of the New World: towards a four migration model? *Ciênc. Cult.* 51:151-165.
- NICHOLS, J. 1990. Linguistic diversity and the first settlement of the New World. *Language* 66:475-521.
- NOVICK, G.E.; NOVICK, C.C.; YUNIS, J.; YUNIS, E.; MARTINEZ, K.; DUNCAN, G.; TROUP, G.; DEININGER, P.L.; STONEKING, M.; BATZER, M.; HERRERA, R. 1995. Polymorphic human specific *Alu* insertions as markers for human identification. *Electrophoresis* 16:1596-1601.
- NOVICK, G.E.; BATZER, M.A.; DEININGER, P.L.; HERRERA, R.J. 1996. The mobile genetic element *Alu* in the human genome. *Bioscience* 46:32-41.
- NOVICK, G.E.; NOVICK, C.C.; YUNIS, J.; YUNIS, E.; MAYOLO, P.A.; SCHEER, W.D.; DEININGER, P.L.; STONEKING, M.; YORK, D.S.; BATZER, M.A.; HERRERA, R.J. 1998. Polymorphic *Alu* insertions and the Asian origin of native american populations. *Hum. Biol.* 70:23-39.
-

-
- ORGEL, L.E.; CRICK, F.H.C. 1980. Selfish DNA: the ultimate parasite. *Nature* 284:604-607.
- PERICOT Y GARCIA, L. 1962. *America Indígena*. Salvat Editores. Spain.
- PRITCHARD, J.K.; SEIELSTAD, M.T.; PEREZ-LEZAUN, A.; FELDMAN, M.W. 1999. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol. Biol. Evol.* 16:1791-1798.
- PRZEWORSKI, M.; HUDSON, R.R.; DI RIENZO, A. 2000. Adjusting the focus on human variation. *TIG* 16:296-302.
- RANA, B.K.; HEWET-EMMETT, D.; JIN, L.; CLANG, B.H-J.; SAMBUUGHIN, N.; LIN, M.; WATKINS, S.; BAMSHAD, M.; JORDE, L.B.; RAMSAY, M.; JENKINS, T.; LI, W-H. 1999. High polymorphism at the human melanocortin 1 receptor locus. *Genetics* 151:1547-1557.
- REICH, D.; GOLDSTEIN, D. 1998. Genetic evidence for a Paleolithic human population expansion in Africa. *Proc. Natl. Acad. Sci. USA* 95:8119-8123.
- RELETHFORD, J.H. 1998. Mitochondrial DNA and ancient population growth. *Am. J. Phys. Anthrop.* 105:1-7.
- RICARDO, C. A. 1991. Povos indígenas no Brasil 1987/88/89/90. Centro Ecumênico de Documentação e Informação, São Paulo.
- ROGERS, J. 1983. Retroposons defined. *Nature* 301:460.
- ROGERS, R.A.; ROGERS, L.A.; MARTIN, L.D. 1992. How the door opened: the peopling of the New World. *Hum. Biol.* 64:281-302.
-

-
- SALZANO, F.M. 1961. Estudos genéticos e demográficos entre os índios do Rio Grande do Sul. *Boletim do Instituto de Ciências Naturais* 9:1-161.
- SALZANO, F.M. 1971. Demographic and Genetic Interrelationships among the Cayapo Indians of Brazil. *Social Biology* 18:148-157.
- SALZANO, F.M.; CALLEGARI-JACQUES, S.M.; FRANCO, M.H.L.P.; HUTZ, M.H.; WEIMER, T.A.; SILVA, R.S.; DA ROCHA, F.J. 1980. The Caingang revisited: blood genetics and anthropometry. *Am. J. Phys. Anthropol.* 53:513-524.
- SALZANO, F.M.; CALLEGARI-JACQUES, S.M.; WEIMER, T.A.; FRANCO, M.H.L.P.; HUTZ, M.H.; PETZL-ERLER, M.L. 1997. Eletrophoretic protein polymorphisms in Kaingang and Guarani indians of southern Brazil. *Am. J. Hum. Biol.* 9:505-512.
- SANTOS, F.R.; ROGRIGUEZ-DELFIN, L.; PENA, S.D.J.; MOORE, J.; WEISS, K.M. 1996. North and South Amerindians may have the same major founder Y chromosome haplotype. *Am. J. Hum. Genet.* 58:1369-1370.
- SANTOS, F.R.; PANDYA, A.; TYLER-SMITH, C.; PENA, S.D.J.; SCHANFIELD, M.; LEONARD, W. R.; OSIPOVA, L.; CRAWFORD, M.H.; MITCHELL, R.J. 1999. The central Siberian origin for Native American Y chromosome. *Am. J. Hum. Genet.* 64:619-628.
- SHAIKH, T.H.; DEININGER, P.L. 1996. The role and amplification of the HS *Alu* subfamily founder gene. *J. Mol. Evol.* 42:15-21.

SILVA; W.A.Jr.; BONATTO, S.L.; HOLANDA, A.J.; RIBEIRO-DOS-SANTOS, A.K.; PAIXÃO, B.M.; GOLDMAN, G.H.; ABE-SANDES, K.; ROGRIGUEZ-DELFIN, L.; BARBOSA, M.; PAÇÓ-LARSON, M.L.; PETZL-ERLER, M.L.; VALENTE, V.; SANTOS, S.E.B.; ZAGO, M.A. 2002. Mitochondrial genome diversity of Native Americans supports a single early entry of founder populations into America. *Am. J. Hum. Genet.* 71:187-192.

STARIKOVSKAYA, Y.B.; SUKERNIK, R.I.; SCHURR, T.G.; KOGELNIK, A.M.; WALLACE, D.C. 1998. mtDNA diversity in Chukchi and Siberian Eskimos: implications for the genetic history of ancient Beringia and the peopling of the New World. *Am. J. Hum. Genet.* 63:1473-1491.

STONE, A. STONEKING, M. 1998. mtDNA analysis of a prehistoric Oneota population: implications for the peopling of the New World. *Am. J. Hum. Genet.* 62:1153-1170.

STONEKING, M. 1993. DNA and recent human evolution. *Evol. Anthropol.* 2:60-73.

SU, B.; XIAO, J.; UNDERHILL, P.; DEKA, R.; ZHANG, W.; AKEY, J.; HUANG, W.; SHEN, D.; LU, D.; LUO, J.; CHU, J.; TAN, J.; SHEN, P.; DAVIS, R.; CAVALLISFORZA, L.; CHAKRABORTY, R.; XIONG, M.; DU, R.; OEFNER, P.; CHEN, Z.; JIN, L. 1999. Y-chromosome evidence for a northward of modern humans into eastern Asia during the last Ice Age. *Am. J. Hum. Genet.* 65:1718-1724.

SWANTON, J.R. 1952. The Indian tribes of North America. Bureau of American Ethnology, Smithsonian Institution, Washington, D.C.

-
- SZATHMARY, E.J. 1979. Blood groups of Siberians, Eskimos and subarctic and northwest coast Indians: the problem of origins and genetic relationships. In: LAUGHLIN, W.S.; HARPER A.B. (eds.) *The first Americans: origins, affinities and adaptations*. Gustav Fischer, New York, pp 185-210.
- SZMULEWICZ, M.N.; NOVICK, G.H.; HERRERA, R.J. 1998. Effects of *Alu* insertions on gene function. *Electrophoresis* 19:1260-1264.
- TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585-595.
- THOMSON, R.; PRITCHARD, J.K.; SHEN, P.; OEFNER, P.J.; FELDMAN, M.W. 2000. Recent common ancestry of human Y chromosomes: evidence from DNA sequence data. *Proc. Natl. Acad. Sci. USA* 97:7360-7365.
- TORRONI, A.; SCHURR, T.G.; YANG, C-C; SZATHMARY, E.J.E.; WILLIAMS, R.C.; SCHANFIELD, M.S.; TROUP, G.A.; KNOWLER, W.C.; LAWRENCE, D.N.; WEISS, K.M.; WALLACE, D.C. 1992. Native American mitochondrial DNA analysis indicates that the Amerind and the Nadene populations were founded by two independent migrants. *Genetics* 130:153-162.
- TORRONI, A.; SUKERNIK, R.I.; SCHURR, T.G.; STARIKOVSKAYA, Y.B.; CABELL, M.F.; CRAWFORD, M.H.; COMUZZIE, A.G.; WALLACE, D.C. 1993. mtDNA variation of aboriginal Siberians reveals distinct genetic affinities with Native Americans. *Am. J. Hum. Genet.* 53:591-608.
- VIGILANT, L.; STONEKING, M.; HARPENDING, H.; HAWKES, K.; WILSON, A. 1991. African populations and the evolution of human mitochondrial DNA. *Science* 253:1503-1507.
-

- VON STERNBERG, R.M.; NOVICK, G.E.; GAO, G.P. HERRERA, R.J. 1992. Genome canalization: the coevolution of transposable and interspersed repetitive elements with single copy DNA. *Genetica* 86:215-246.
- WALLACE, D.C. 1995. Mitochondrial DNA variation in human evolution, degenerative disease and aging. *Am. J. Hum. Genet.* 57:201-223.
- YAMAMOTO, T.; DAVIS, C.G.; BROWN, M.S.; SCHNEIDER, W.J.; CASEY, M.L.; GOLDSTEIN, J.L.; RUSSEL, D.W. 1984. The human LDL receptor: a cysteine-rich protein with multiple *Alu* sequences in its mRNA. *Cell* 39:27-38.
- YU, N.; LI, W-H. 2000. No fixed nucleotide difference between Africans and non-Africans at the pyruvate dehydrogenase E1 α -subunit locus. *Genetics* 155:1481-1483.
- YU, N.; ZHAO, Z.; FU, Y-X; SAMBUUGHIN, N.; RAMSAY, M.; JENKINS, T.; LESKINEN, E.; PATTHY, L.; JORDE, L.B.; KWEOMORI, T.; LI, W-H. 2001. Global patterns of human DNA sequence variation in a 10-kb region on chromosome 1. *Mol. Biol. Evol.* 18:214-222.
- ZHAO, Z.; JIN, L.; FU, Y-H.; RAMSAY, M.; JENKINS, T.; LESKINEN, E.; PAMILO, P.; TREXLER, M.; PATTHY, L.; JORDE, L.B.; RAMOS-ONSINS, S.; YU, N.; LI, W-H. 2000. Worldwide DNA sequence variation in a 10-kilobase noncoding region on human chromosome 22. *Proc. Natl. Acad. Sci. USA.* 97:11354-11358.
- ZIETKIEWICZ, E.; YOTOVA, V.; JARNIK, M.; KORAB-LASKOWSKA, M.; KIDD, K.K.; MODIANO, D.; SCOZZARI, R.; STONEKING, M.; TISHKOFF, S.; BATZER, M.; LABUDA, D. 1997. Nuclear DNA diversity in worldwide distributed human populations. *Gene* 205:161-171.