

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**A GENÔMICA MITOCONDRIAL APLICADA AO  
PROBLEMA DO POVOAMENTO PRÉ-COLOMBIANO  
DAS AMÉRICAS**

**RICARDO KANITZ**

**Orientadora: Loreta Brandão de Freitas**

**Co-orientador: Sandro Luis Bonatto**

Trabalho apresentado como um dos  
requisitos para obtenção do grau de  
Bacharel no Curso de Ciências  
Biológicas Ênfase Molecular, Celular  
e Funcional.

Porto Alegre, Dezembro de 2005.

UFRGS - BIBLIOTECA  
INST. BIOCÊNCIAS

## **AGRADECIMENTOS**

Aos meus pais, pelo carinho, amor e apoio dado em toda a minha vida.

À minha irmã que completa a família.

À minha namorada, Aline pela compreensão e interesse por este trabalho e por todo o amor a mim dedicado.

Ao Nelson, pelo exemplo de pessoa e pesquisador que representa para mim e principalmente pelo grande amigo que é.

Ao Sandro, meu orientador, por me acolher já há mais de dois anos e meio em seu laboratório e estar sempre disposto a ajudar e principalmente discutir os assuntos do mitogenoma.

À Loreta, pela ajuda ao me indicar ao laboratório Genoma já no terceiro semestre da faculdade e por todo o suporte dado para a realização do meu bacharelado.

Aos meus duplamente colegas e grandíssimos amigos Manoel e Luana.

Às meninas que ajudaram, e muito, no desenvolvimento deste projeto: Roberta, Ana Carolina, Marina e Luana novamente.

Ao Ronaldo, pela ajuda desenvolvendo scripts e programas que muito facilitaram as nossas vidas na realização deste e de outros trabalhos.

A todo o incontável pessoal do laboratório Genoma – aqueles que por lá passaram e aqueles que lá estão – pelo companheirismo, assim como, pelos bolos, tortas, chocolates e cafés quase todos acompanhados por discussões de altíssima importância sobre ciência, política, futebol e a vida alheia.

Aos meus grandes amigos que fazem parte do componente ambiental daquilo que hoje eu chamo de “eu”.

## ÍNDICE

Resumo	3
Introdução	5
Histórico dos estudos sobre o povoamento das Américas	5
Estudos baseados no DNA mitocondrial (mtDNA)	7
O mtDNA como marcador em estudos evolutivos humanos	9
Objetivos	11
Materiais e Métodos	13
Amostras de DNA	13
Sobre as seqüências geradas	15
Análises dos dados	16
Resultados e Discussão	19
Conclusões	31
Referências Bibliográficas	32

## RESUMO

Abordagens genéticas com seqüências de DNA mitocondrial (mtDNA) têm sido usadas de forma bem sucedida na resolução de diversos problemas históricos da evolução humana. Um dos problemas mais estudados é o povoamento pré-colombiano do continente americano. Algumas das dúvidas que ainda pairam sobre o assunto são (i) o número e a (ii) datação dos eventos de migração assim como o (iii) tamanho da população fundadora. Procurar uma resposta para os dois primeiros problemas foi o objetivo do nosso trabalho. Para tanto, maximizamos a retirada de informação do mtDNA pelo seqüenciamento completo da região codificante dessa molécula.

As amostras foram obtidas de indivíduos indígenas de todos os cinco haplogrupos mitocondriais nativos (A, B, C, D e X) em diversas tribos por todo o continente. O número total de genomas mitocondriais de nativos americanos analisados é de 122 sendo que 58 dessas seqüências foram obtidas pelo nosso grupo. Para fins de comparação, todas as seqüências de não-americanos pertencentes aos haplogrupos A-D e X disponíveis na literatura também foram colocadas nas análises. Verificamos a existência de marcadores genéticos exclusivos no mtDNA para nativos americanos em todos os haplogrupos exceto para o C. Esses dados então sugerem um único haplótipo fundador para cada haplogrupo. As estatísticas populacionais obtidas de diversidade nucleotídica e expansão populacional (D de Tajima) para os haplogrupos A-D foram bastante semelhantes. A diversidade variou em torno 0,045% e o D de Tajima mostrou-se significativamente negativo sugerindo expansão populacional para todos. Para o



haplogrupo X, encontramos diferenças nesses dois parâmetros: o valor de diversidade encontrado foi de 0,0305% e o teste de neutralidade não diferiu de forma significativa de zero não sugerindo expansão populacional.

Para explicar esses resultados, temos duas hipóteses: (i) trata-se de uma segunda e mais recente entrada no continente trazendo somente esse haplogrupo ou então (ii) o que vemos é algum efeito inerente ao processo de colonização no continente em que seria possível que este haplogrupo ocorresse em baixa frequência na população fundadora e não seguiu o ritmo de expansão dos outros haplogrupos. Para ajudar a decidir entre essas duas hipóteses, análises com métodos de máxima verossimilhança, árvores linearizadas e *median joining* (parâmetro  $\rho$ ) foram feitas para estimar as idades de expansão de cada haplogrupo. Nessas análises verificamos não haver diferenças significantes entre as idades dos haplogrupos A-D e do haplogrupo X. Todas elas resultam em torno de 23 mil anos atrás com a taxa de uma mutação a cada 5.140 anos retirada da literatura. Com esses dados, o povoamento das Américas teria ocorrido em um único evento trazendo simultaneamente todos os haplogrupos indígenas antes do último máximo glacial a partir da Beríngia provavelmente por uma rota costeira.

## INTRODUÇÃO

### **Histórico dos estudos sobre o povoamento das Américas:**

Até há pouco tempo, a hipótese mais aceita acerca do povoamento do continente americano, e que ainda se encontra nos livros-texto (Kelly, 2003), é a apresentada pelo modelo "*Clovis First*". Esse modelo se baseia nos primeiros achados arqueológicos nos EUA nos complexos de Clovis (nos EUA continental) e Nenama (no território do Alasca) (Kelly, 2003). Nessa hipótese, os primeiros americanos teriam vindo da Ásia através da Beríngia – istmo que ligava Ásia e América durante o pleistoceno. Esse evento teria acontecido há cerca de 14 mil anos (11.800 anos-radiocarbono) que é a idade mais antiga de vestígios no Complexo Nenama. Na América continental, a mais antiga evidência humana, segundo esse modelo, é justamente o sítio de Clovis com 13 mil anos (10.900 anos-radiocarbono). Esse modelo apresenta a datação do povoamento das Américas posterior ao último máximo glacial (LGM – *Last Glacial Maximum*) – de 21 até 16,5 mil anos atrás (Hall *et al.*, 2004). Portanto é cabível inferir que, nos moldes desse modelo, o povoamento da América continental poderia ter se dado através do corredor livre de gelo de Alberta que estaria localizado nesse período entre as duas grandes geleiras que cobriram boa parte da América do Norte durante a última glaciação.

O modelo de "*Clovis First*", entretanto, tem enfrentado várias contestações (Schurr, 2004) vindas de outras evidências arqueológicas como as descobertas de sítios que podem ser mais antigos na própria América do Norte e na América do

Sul. Como exemplo, tem-se o sítio de Monte Verde no Chile datado em 14,5 mil anos (12.200 anos-radiocarbono) (Meltzer, 1997).

Uma importante abordagem multidisciplinar, já envolvendo genética, para um melhor entendimento do processo de povoamento das Américas, foi feita por Greenberg *et al.* em 1986. Nesse trabalho, os autores abordaram o problema conciliando dados genéticos (grupos sanguíneos e protéicos), lingüísticos e dentais e apresentaram um modelo em que os nativos americanos estariam separados em três grandes grupos lingüísticos – Ameríndios, Na-Dené e Esquimó-Aleuta – correspondendo cada um a um evento migratório (FIGURA 1). A primeira migração teria trazido os Ameríndios que se distribuíram por quase toda a América seguidos pelos Na-Dené e Esquimós-Aleutas, (FIGURA 1). Em relação ao modelo “*Clovis First*”, essa hipótese não difere quanto às datações: todas elas após o LGM; ou seja, recentes. Apesar de ter tido uma boa aceitação inicial, esse modelo acabou não encontrando suporte em estudos posteriores, tanto de cunho genético quanto lingüístico (Diamond, 1990). Em contraponto aos modelos de migrações recentes, Rogers *et al.* (1992) e Prous (1995) apresentaram modelos de entradas antigas, em torno de 25 mil anos atrás, no continente. Os achados de Meltzer (1997), com 14,5 mil anos, e Neves *et al.* (1999) e Roosevelt *et al.* (1996), contemporâneos a Clovis, na América do Sul reforçam a idéia de uma entrada pré LGM. Além disso, o modelo apresentado por Neves *et al.* em 1999, baseado em caracteres morfológicos diz que, independentemente de quantas migrações sejam consideradas em estudos com amostras atuais, há mais uma mais antiga que teria trazido indivíduos de morfologia distinta da mongolóide dos nativos atuais (não diferenciada) que teriam se extinguido ou, mais provavelmente, sido incorporados

aos seus sucessores. Esse modelo não encontra suporte na maioria dos trabalhos envolvendo marcadores genéticos.

Ainda ultimamente a relação dos caracteres morfológicos com genéticos tem sido discutida e é possível que a tão destacada morfologia não diferenciada apresentada por Neves *et al.* (1999) seja parte de um gradiente de variação morfológica presente nos primeiros habitantes do continente e que ainda pode ser verificado em algumas das populações remanescentes atualmente (Rolando Gonzáles-José, com. pess.). Há ainda várias propostas menos consideradas de modelos envolvendo até migrações vindas da Europa (Brown *et al.*, 1998) e Polinésia (Cann, 1994). O fato é que as questões sobre o número de ondas migratórias e a idade do povoamento do continente americano ainda estão em aberto e a utilização de ferramentas como marcadores genéticos pode se mostrar bastante útil.

#### **Estudos baseados no DNA mitocondrial (mtDNA):**

Desde Schurr *et al.* (1990), o mtDNA tem mostrado utilidade na resolução de alguns problemas relacionados ao povoamento das Américas. Naquele trabalho, os autores identificaram, por RFLP, os quatro grandes haplogrupos – denominados A, D, C e D – ocorrentes no continente americano e viram que esses mesmos haplogrupos estavam presentes na Ásia. Esses dados ressaltam a origem dos nativos americanos naquele continente. A existência desses haplogrupos foi confirmada por Horai *et al.* (1993) que propuseram um modelo em que quatro eventos independentes teriam carregado para as Américas os quatro haplogrupos entre 21 e 14 mil anos atrás. O quinto haplogrupo mitocondrial



ocorrente no continente, o haplogrupo X, foi descrito por Forster *et al.* (1996) e confirmado por Brown *et al.* (1998). Nesse estudo, foi proposto que a presença deste haplogrupo, praticamente ausente em populações asiáticas e mais comum na Europa, seria evidência de uma migração via Oceano Atlântico de populações protocaucasóides vindas da Europa. A frequência do haplogrupo X é bastante pequena e restrita à América do Norte (FIGURA 1). A variação da frequência dos cinco haplogrupos na América (FIGURA 1) se deve a efeitos de deriva genética pós-povoamento ocorrentes já dentro do continente pouco afetando a diversidade total da população continental.

Baseando-se no trabalho de Greenberg *et al.* (1986) e na aparente divergência entre Ameríndios e Esquimó-Aleuta/Na-Dené, Ballinger *et al.* (1992) e Wallace e Torroni (1992) propuseram um modelo de duas e três migrações, respectivamente, em que a mais antiga delas teria ocorrido há cerca de 30 mil anos. Por outro lado, Merriwheter *et al.* (1995) e Bonatto e Salzano (1997a,b) propuseram a ocorrência de um único evento de povoamento da Beríngia. Segundo Bonatto e Salzano (1997a,b), utilizando como marcador as regiões hipervariáveis do mtDNA, esse evento teria ocorrido há cerca de 30 mil anos com uma expansão na Beríngia que teria povoado o continente como um todo. A diferenciação entre os grandes grupos lingüísticos teria ocorrido posteriormente já nas Américas durante o máximo glacial devido ao isolamento geográfico causado pelas geleiras. Esse modelo de uma única e antiga expansão na Beríngia – chamado “*Out of Beringia*” – encontrou suporte em estudos posteriores realizados por Mahli *et al.* (2002) e Silva *et al.* (2003) e também é compatível com as

datações dos diversos e antigos sítios arqueológicos sul-americanos apresentados acima.

### **O mtDNA como marcador em estudos evolutivos humanos:**

O DNA mitocondrial tem sido utilizado amplamente em diversos estudos evolutivos, não só de humanos, como visto acima, mas de variadas espécies de eucariotos (p. ex.: Hugall *et al.*, 2002; Beheregaray *et al.*, 2003; Savolainen *et al.*, 2004). De maneira geral, regiões mais restritas da molécula possuem informação suficiente para os fins dos estudos como os exemplificados acima. Para obter boa resolução de marcadores moleculares, têm-se duas possibilidades: (i) utilizar seqüências mais curtas e de evolução rápida ou (ii) seqüências maiores de evolução relativamente mais lenta e potencialmente constante. O mtDNA humano pode ser dividido em duas grandes regiões funcionais: a região codificante para os variados genes mitocondriais e a região controladora da expressão dos genes onde há também o centro de origem de replicação da molécula. A primeira é caracterizada por ter uma evolução rápida se comparada com o DNA nuclear e mais lenta e, o que é mais importante, mais constante quando comparada à segunda região (a controladora) (Ingman *et al.*, 2000). Portanto, ambas as regiões podem ser utilizadas em estudos evolutivos dependendo do propósito destes. Isto é, onde é importante obter bastante informação a partir de segmentos curtos de DNA, a região controladora e suas subdivisões encontram grande utilidade; onde o foco do estudo se baseia no pressuposto de uma evolução constante ("clocklike"), a região codificadora e seus genes têm grande aplicabilidade. O presente estudo, por tratar de um evento evolutivo recente dentro de uma curta

história evolutiva comparada com outras espécies de mamíferos (a do *Homo sapiens*), fez uso do máximo de informação com evolução constante retirável do mtDNA seqüenciando toda a extensão do genoma mitocondrial (16.569 pares de base) dando ênfase à sua região codificadora (15.444 pb).

Há especulações em relação à validade do uso do mtDNA humano como marcador pois há suspeitas de que possa haver recombinação entre moléculas de linhagens distintas causando sério ruído em análises evolutivas. A recombinação entre mtDNAs ocorre em músculos esqueléticos, como verificado por Zsurka *et al.* (2004) e Kraytsberg *et al.* (2005). A recombinação entre linhagens diferentes de mtDNA dar-se-ia em indivíduos portadores de heteroplasmias, nos quais ocorre a herança de mtDNA também a partir do pai (Zsurka *et al.*, 2004). Esses autores inferem que possa haver complicações evolutivas nesses fenômenos, entretanto a ocorrência deles, até agora só foi verificada no tecido muscular esquelético que não é herdado pelos filhos dos recombinantes. Portanto, não foi verificado nenhum caso em que mtDNAs recombinados pudessem ser passados de uma geração a outra. Sendo assim, não há nenhuma implicação no uso do mtDNA com marcador em estudos evolutivos evidenciada nos eventos de recombinação até agora demonstrados para humanos. A hipótese de que a recombinação entre diferentes linhagens de mtDNA não ocorre ou, pelo menos, é muito rara a ponto de não afetar estudos evolutivos encontra respaldo nos próprios estudos filogenéticos e filogeográficos de populações humanas nos quais ainda não se verificou sinal algum de recombinação de linhagens mitocondriais. Portanto, a utilização do DNA mitocondrial em estudos evolutivos e populacionais em *Homo sapiens* é válida e não encontra problemas causados por ruído de recombinação entre linhagens.

## Objetivos:

Com a finalidade de obter estimativas mais precisas sobre a idade de eventos de povoamento do continente americano, este trabalho utilizou como marcador todo o genoma mitocondrial. Além do marcador mais preciso, foram, pela primeira vez, acrescentadas às análises amostras de indígenas pertencentes ao haplogrupo X. Esse haplogrupo teve sua história comparada à dos demais haplogrupos permitindo, assim, uma estimativa mais completa do possível número de eventos relacionados ao povoamento das Américas.

A partir de uma inferência da variação demográfica das seqüências de nativos americanos, pôde-se estimar um tamanho para a população pioneira nas Américas. A partir dos dados obtidos, então, possíveis cenários de povoamento do continente foram discutidos.



FIGURA 1: Mapa das Américas com as rotas de migração e as estimativas de tamanho populacional para as populações pioneiras, com base em dados genéticos de indivíduos nativos das Américas. O mapa mostra as rotas de migração e as estimativas de tamanho populacional para as populações pioneiras, com base em dados genéticos de indivíduos nativos das Américas. O mapa mostra as rotas de migração e as estimativas de tamanho populacional para as populações pioneiras, com base em dados genéticos de indivíduos nativos das Américas.



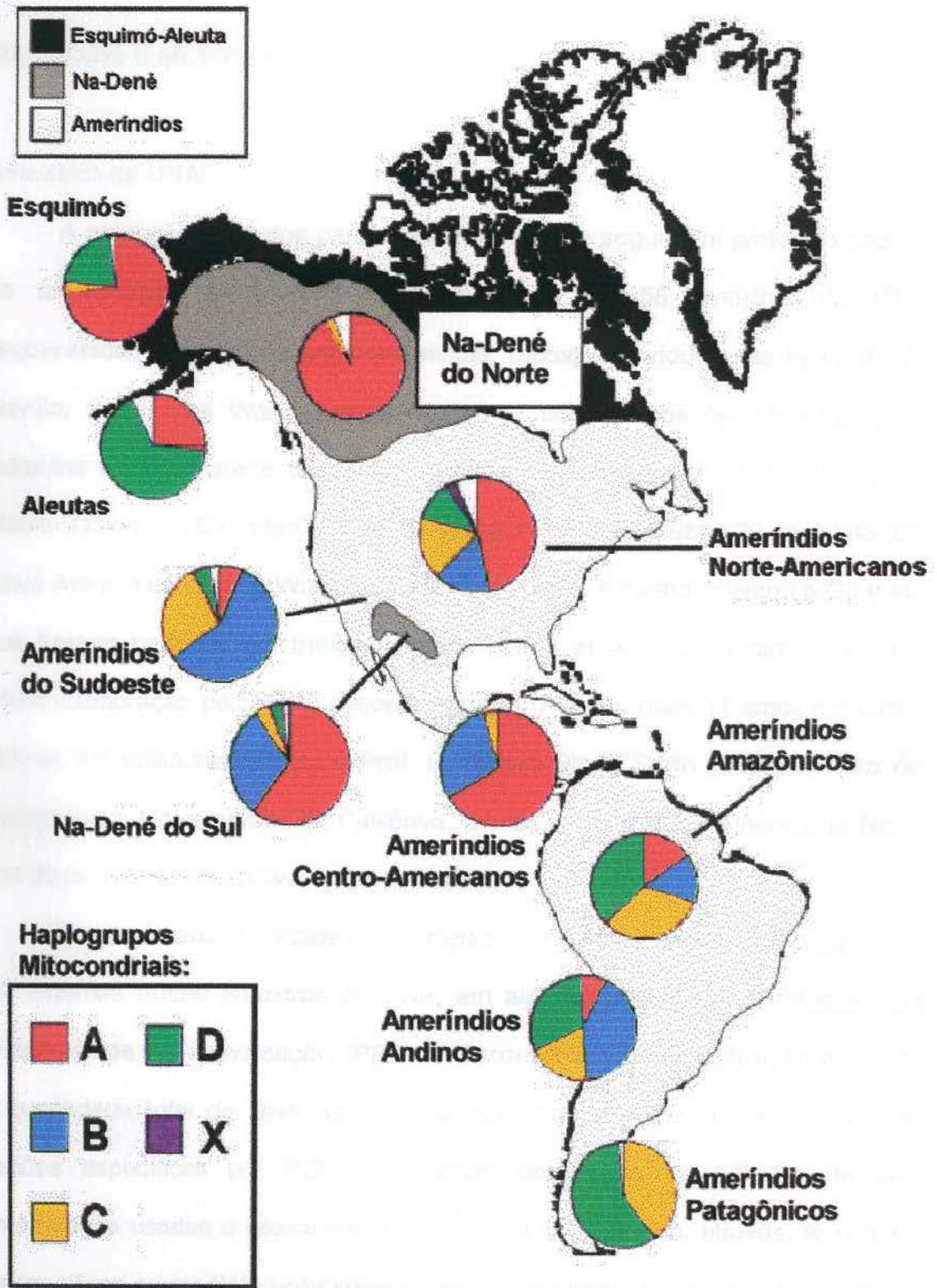


FIGURA 1: Mapa das Américas com as divisões lingüísticas propostas por Joseph Greenberg em 1986 (legenda superior) e proporções da distribuição dos haplogrupos de mtDNA em diferentes grandes populações nativas dentro do continente (legenda inferior). Faixas brancas nos gráficos representam ruído de miscigenação nas populações aborígenes ou seqüências que sofreram reversões de mutações marcadoras (adaptado de Schurr, 2004).

## MATERIAIS E MÉTODOS

### Amostras de DNA:

A obtenção de dados para a posterior análise seguiu um protocolo padrão de amplificação de DNA e seqüenciamento. As 58 amostras de DNA seqüenciadas tiveram três origens distintas. Dezoito indivíduos das tribos Ache, Gavião, Suruí, Wai Wai, Xavante e Zoró já tinham suas amostras de DNA extraídas e faziam parte do banco de DNA montado pelo Prof. Dr. Francisco Mauro Salzano no Departamento de Genética da UFRGS. Outras 29 amostras das tribos Arara, Yanomama, Waiampi, Tyrio, Poturujara, Katuena, Yayapo e Guarani, que fizeram parte de um trabalho anterior (Silva *et al.* 2003) foram fornecidas como colaboração pelo autor daquele trabalho. Por fim, mais 11 amostras foram obtidas em colaboração com o Prof. Dr. David Glenn Smith (Departamento de Antropologia, Universidade da Califórnia, Davis) de nativos da América do Norte das tribos Jemez, Chippewa, Pomo e Sioux.

Devido à baixa qualidade pela antiguidade da amostra ou pouca quantidade em algumas outras amostras de DNA, em alguns casos foram adotados dois protocolos de pré-amplificação: IPEP ou *GenomiPhi*. A pré-amplificação aumenta a quantidade total de DNA na amostra facilitando a posterior amplificação de regiões específicas por PCR. O princípio de ambos os métodos de pré-amplificação usados é basicamente o mesmo e consiste em, através de *primers* inespecíficos e uma DNA polimerase especial, amplificar aleatoriamente o material genético através de uma reação de amplificação. Os resultados das pré-

amplificações foram satisfatórias e resolveram boa parte dos problemas relacionados à baixa qualidade ou quantidade de DNA.

Para as reações de polimerização em cadeia (*Polimerase Chain Reaction: PCR*), foram usados 24 pares de *primers* conforme Rieder *et al.* (1998) cobrindo todo o genoma mitocondrial. O sucesso das reações foi verificado por eletroforese em gel de agarose 1% com tampão Tris-Borato-EDTA (TBE) e corante Brometo de Etídio (EtBr). Para a purificação do DNA pós PCR, primeiramente adotou-se um protocolo enzimático utilizando Fosfatase Alcalina e Exonuclease I. Posteriormente, outro protocolo se mostrou mais eficiente e barato, utilizando PEG8000, e acabou sendo adotado como padrão.

O seqüenciamento das amostras amplificadas foi feito utilizando os mesmos *primers* utilizados para a PCR. As reações foram feitas conforme as instruções fornecidas pelo fabricante do seqüenciador automático utilizado – MegaBACE 1000 (General Eletrics Health Care). Em alguns casos, para melhor resolução dos cromatogramas, fez-se uso de um seqüenciador automático ABI3100 (Applied Biosystems) em colaboração com o laboratório do Prof. Dr. Giancarlo Pasquali (Departamento de Biotecnologia, UFRGS). Outros problemas referentes a polimorfismos mais localizados puderam também, em alguns casos, ser resolvidos fazendo uso de reações com algumas enzimas de restrição específicas e conferência por eletroforese em gel de agarose.

O processo de obtenção das seqüências contou com o auxílio, em diferentes tempos, de Robert Eckert, Ana Carolina Silva Valls, Marina Ochoa Favarini e Luana Cardoso-Silva que ajudaram, também, a resolver diversos problemas nesse processo.

### **Sobre as seqüências geradas:**

Os cromatogramas gerados após os seqüenciamentos foram montados nos programas Phred, Phrap e Consed ([www.phred.org](http://www.phred.org)). As montagens foram feitas para cada indivíduo resultando, no fim, na seqüência de toda a molécula do mtDNA para as suas duas fitas. As montagens foram revisadas visualmente para todos os indivíduos para verificar possíveis erros ou faltas de leitura para algumas regiões. Qualquer problema dessa ordem foi imediatamente resolvido pelo re-sequenciamento da região do mtDNA em questão daquele indivíduo a partir de outra reação de amplificação. Uma vez montadas as seqüências completas do mtDNA de todos os indivíduos, essas foram alinhadas e comparadas umas com as outras e com outras seqüências retiradas da literatura incluindo a seqüência de referência de Cambridge corrigida (Anderson *et al.*, 1981; Andrews *et al.*, 1999). Todos os sítios variáveis foram então anotados através de um *script* computacional criado pelo nosso grupo e verificados novamente nos cromatogramas de cada indivíduo. Uma vez feita essa verificação, aplicou-se a metodologia sugerida por Bandelt *et al.* (2002) e Salas *et al.* (2005) para verificar a existência de mutações fantasma. Mutações fantasma são aquelas que surgem por possíveis erros inerentes aos processos de obtenção da seqüência como a química do seqüenciador, da PCR ou mais freqüentemente da leitura do cromatograma pelos programas de montagem das seqüências. Havendo sítios candidatos a ter esse tipo de mutação, fazia-se novamente a verificação das seqüências diretamente nos seus cromatogramas e, caso a dúvida não pudesse ser sanada, refazia-se o processo de amplificação e seqüenciamento daquela região do genoma mitocondrial para o indivíduo em questão. Findada a



conferência da qualidade das seqüências geradas, elas eram novamente alinhadas e estavam finalmente prontas para que se procedessem as análises estatísticas.

### **Análises dos dados:**

Às 58 seqüências geradas pelo nosso grupo, foram acrescentadas outras 64 de nativos americanos disponíveis em Herrnstadt *et al.* (2002), Mishmar *et al.* (2003), Ingman *et al.* (2000), Maca-Meyer *et al.* (2001) e Reidla *et al.* (2003) e ainda 169 seqüências de não nativos americanos (Herrnstadt *et al.* 2002; Mishmar, 2003; Ingman *et al.*, 2000; Maca-Meyer *et al.*, 2001; Tanaka *et al.*, 2004; Kong *et al.*, 2003; Starikovskaya *et al.*, 2005; Finnilä *et al.*, 2001; Levin *et al.*, 1999; Reidla *et al.*, 2003 e Derbeneva *et al.*, 2002) pertencentes aos mesmos haplogrupos dos nativos americanos. Somam-se então 291 seqüências utilizadas na realização deste trabalho sendo 122 de nativos americanos das quais 33 pertencem ao haplogrupo A; 32, ao B; 24 ao C; 21 ao D; e 12 ao haplogrupo X. É importante ressaltar que, devido à sua inconstância na taxa evolutiva (Ingman *et al.*, 2000), a região controladora do mtDNA não foi utilizada nas análises exceto para a verificação de sítios marcadores exclusivos de nativos americanos em cada um dos cinco haplogrupos (ver resultados). Todas as demais análises conduzidas neste trabalho se basearam nas seqüências da região codificante do mtDNA que possui uma taxa mais lenta e constante de evolução (Ingman *et al.*, 2000).

Para o alinhamento manual das seqüências a fim de buscar homologia entre elas, fez-se uso do programa BioEdit Sequence Alignment Editor 6.0.7 ([www.mbio.ncsu.edu/BioEdit/bioedit.html](http://www.mbio.ncsu.edu/BioEdit/bioedit.html)). Para análises preliminares de

diversidades nucleotídica e haplotípica e verificação de possíveis sítios com mutações específicas para nativos americanos, dois programas foram utilizados: Mega3.1 ([www.megasoftware.net](http://www.megasoftware.net)) e DnaSP4.0 (Rozas e Rozas, 1999). Para a geração das árvores filogenéticas usou-se o programa Mega3.1. Para as redes de haplótipos e estimativa do parâmetro *Rho*, o programa utilizado foi o Network 4.1.1.0 ([www.fluxus-engineering.com/sharenet.htm](http://www.fluxus-engineering.com/sharenet.htm)).

Para as estimativas de idade de expansão de cada haplogrupo a partir da sua diversidade nucleotídica ( $\pi$ ), a taxa de divergência do DNA mitocondrial humano foi estimada a partir da divergência entre humanos e chimpanzés que ocorreu em torno de 6 milhões anos atrás (Glazko e Nei, 2003). Essa taxa é simplesmente a razão entre distância genética média entre as duas espécies e o tempo de divergência entre cada uma das linhagens e seu ancestral comum. Ou seja:  $K = d_{hp} / 2 \cdot t_{divergência}$ . Onde:  $K$  - taxa de divergência;  $d_{hp}$  - distância Homo-Pan;  $t_{divergência}$  - tempo de divergência de cada uma das espécies (6 milhões de anos). O tempo de divergência de cada haplogrupo foi estimado a partir da mesma fórmula considerando a distância genética média dentro do haplogrupo (diversidade nucleotídica) e o valor da taxa de divergência já obtida. Já a taxa de substituição de nucleotídeos (diferente da taxa de divergência entre seqüências pois se refere à evolução em uma só linhagem no tempo) de  $1,26 \cdot 10^{-8}$  substituições/sítio/ano utilizada nas estimativas de idade do ancestral comum mais recente (TMRCA do inglês: The Most Recent Common Ancestor) de cada haplogrupo nativo foi retirada de Mishmar *et al.* (2003). As árvores linearizadas foram estimadas no programa Mega3.1. Por fim, as análises baseadas em

inferência bayesiana da variação do tamanho populacional dos nativos americanos e do tempo de coalescência de cada haplogrupo mitocondrial nativo foram feitas no pacote *Bayesian Evolutionary Analysis Sampling Trees* (BEAST) ([evolve.zoo.ox.ac.uk/beast](http://evolve.zoo.ox.ac.uk/beast)). Maiores detalhes estão descritos adiante juntamente com os resultados obtidos com cada método.

## RESULTADOS E DISCUSSÃO

Foram verificadas mutações do mtDNA específicas de nativos americanos para os haplogrupos A, B, D e X (TABELA 1), considerando somente as mutações fixadas em todos os nativos americanos e ausentes em populações não-nativas. Esse resultado corrobora a hipótese de haver um efeito gargalo-de-garrafa no evento de povoamento do continente americano, pois um dos resultados de uma grande restrição do tamanho populacional é a redução da diversidade genética, podendo ocorrer até a fixação de somente um alelo. Para todos os haplogrupos, exceto para o C, verifica-se a fixação de um único haplótipo fundador da diversidade genética atual do haplogrupo nas Américas. Esse efeito pode ser visualizado nas redes de haplótipos geradas a partir das seqüências da região codificante do mtDNA (FIGURA 2) dos haplogrupos A, B, D e X e também na estimativa da variação do tamanho populacional a partir das seqüências de todos os nativos americanos pelo *Skyline Plot* (FIGURA 5).

**TABELA 1:** Sítios com transições específicas de nativos americanos. \*Mutações da região controladora: não consideradas na construção das redes de haplótipos (FIGURA 2).

Haplogrupo	Mutações Marcadoras de Nativos Americanos
A	135*, 8027, 12007, 16111*
B	3547, 4977, 6473, 11177
C	-
D	2092, 16325*
X	8913, 12397, 14502

Para que se possam obter estimativas da idade de divergência de cada haplogrupo nativo americano a partir das diversidades nucleotídicas calculadas ( $\pi$ )

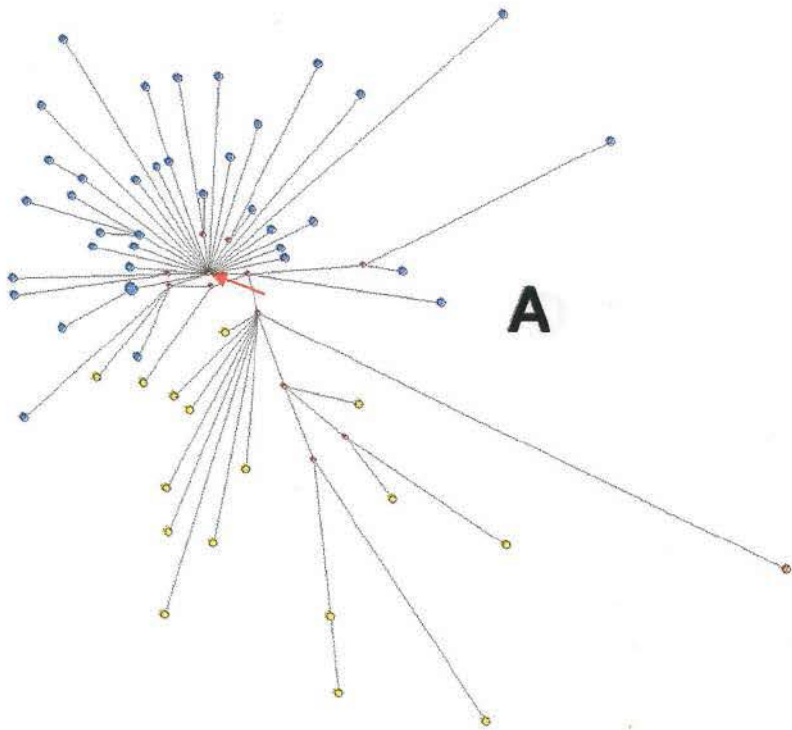


(TABELA 2), alguns pressupostos são necessários: (i) deve ter havido um efeito gargalo-de-garrafa redutor da diversidade genética; (ii) após a fixação dessa baixa diversidade na população, deve haver uma forte expansão populacional minimizando a perda de novos haplótipos por deriva; e (iii) a taxa de mutação do marcador deve ser constante – o que é uma característica da região controladora do mtDNA (Ingman *et al.*, 2000). Sendo esses pressupostos sustentados, pode-se então inferir que a diversidade genética encontrada atualmente para cada haplogrupo nos nativos americanos seja diretamente relacionada ao tempo de expansão desses haplogrupos no continente através da constante taxa de divergência (K).

Verificou-se, então, efeito gargalo-de-garrafa para os haplogrupos nativos pela fixação de alelos específicos das Américas na maioria deles, mas o sinal de expansão verificado pelo teste de neutralidade de Tajima ( $D_{Taj}$ ) (Tajima, 1989) – para marcadores de evolução neutra como o mtDNA – não foi significativo para o haplogrupo X (TABELA 2) o que impossibilitou estimar a sua idade de expansão dentro do continente. As idades estimadas por essa metodologia estão colocadas na TABELA 2 e na FIGURA 4.

**TABELA 2:** Valores do teste de neutralidade de Tajima ( $D_{Taj}$ ), de diversidade nucleotídica ( $\pi$ ) com erro padrão e estimativa de idade de cada haplogrupo através da diversidade nucleotídica.

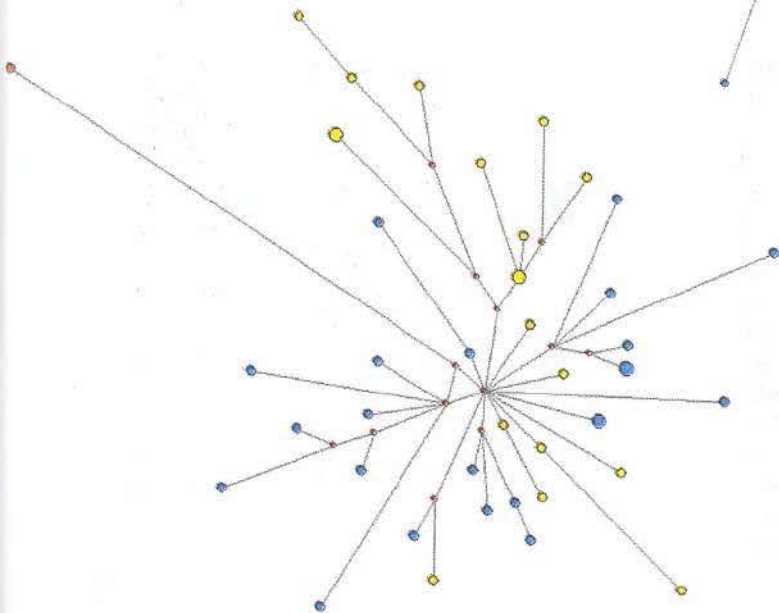
Haplogrupo	$D_{Taj}$	$\pi$	EP( $\pi$ )	Idade
A	-2,74190 <sup>P&lt;0,001</sup>	0,000494	0,000046	22.641 (+/- 2.704)
B	-2,71862 <sup>P&lt;0,001</sup>	0,000571	0,000072	26.170 (+/- 3.299)
C	-2,10182 <sup>P&lt;0,05</sup>	0,000500	0,000091	22.916 (+/- 4.170)
D	-2,25647 <sup>P&lt;0,01</sup>	0,000450	0,000067	20.624 (+/- 3.070)
X	-1,27728 <sup>NS</sup>	0,000305	0,000086	-



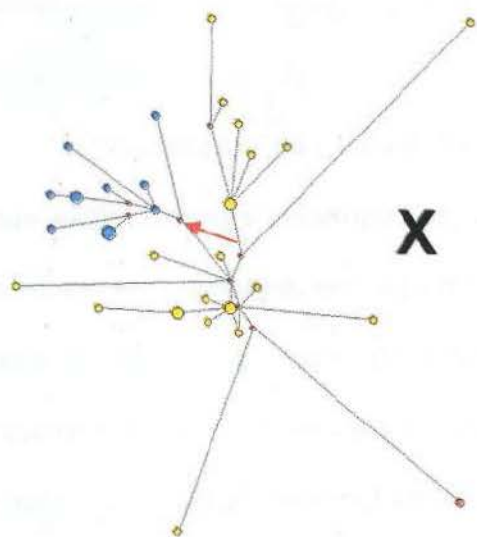
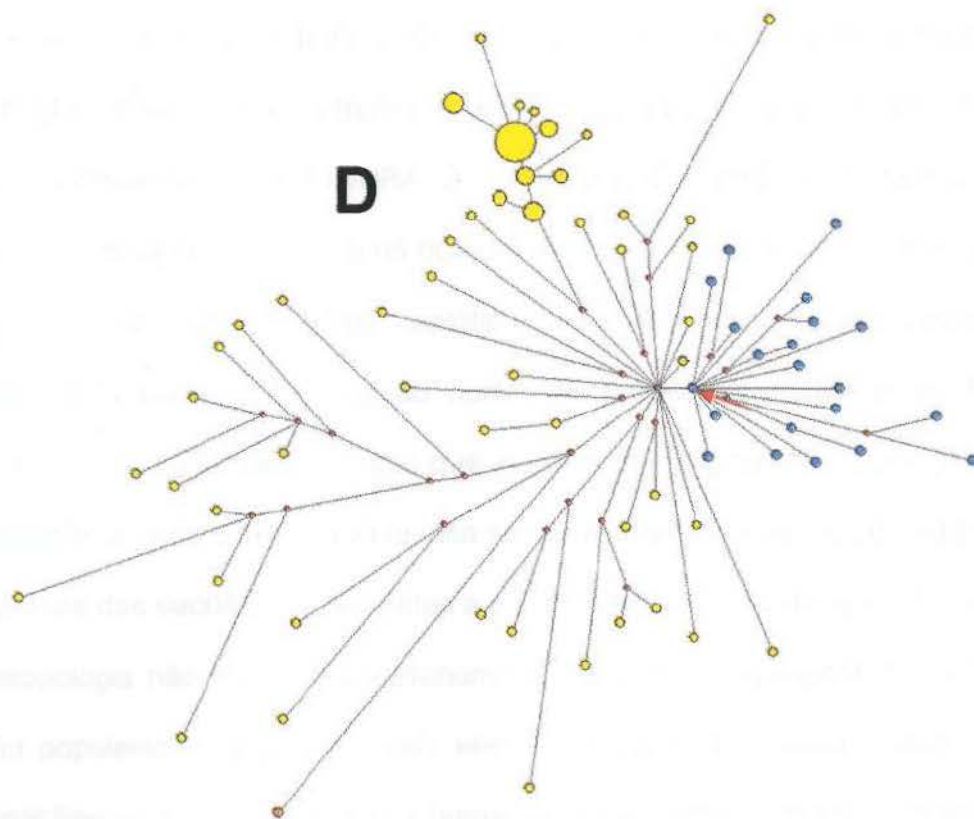
**A**



**B**



**C**



**FIGURA 2:** Redes de haplótipos construídas por *median-joining* (Bandelt *et al.*, 1999) a partir das seqüências da região codificadora do mtDNA. A, B, C, D e X são os haplogrupos. Haplótipos marcados em azul são os nativos americanos. Em amarelo, nativos da Ásia ou Europa (haplogrupo X). Os haplótipos em vermelho são haplótipos perdidos ou não amostrados e os haplótipos marcados em laranja em todas as árvores são a seqüência de referência de Cambridge corrigida (rCRS) (Andrews *et al.*, 1999). Os haplogrupos A, B, D e X possuem um haplótipo fundador ancestral de todas seqüências de nativos americanos (marcado com a seta). O haplogrupo C não possui um haplótipo fundador para a região analisada.

Pela impossibilidade de calcular uma idade para o haplogrupo X e para verificar as idades estimadas para os demais haplogrupos, outros métodos foram aplicados. O primeiro e mais simples deles foi realizado a partir das redes de haplótipos apresentadas na FIGURA 2: o método do parâmetro *Rho*. Essa estimativa se baseia simplesmente no número médio de mutações que separam o haplótipo dado com ancestral dos haplótipos dele derivados. Esse parâmetro, multiplicado pela taxa de substituição nucleotídica, retirada de Mishmar *et al.* (2003) (de  $1,26 \cdot 10^{-8}$  mutações/sítio/ano que, no genoma mitocondrial, corresponde a uma mutação a cada 5.140 anos) resulta em uma estimativa da média de tempo da divergência das seqüências derivadas a partir da ancestral. Estimar a idade por essa metodologia não implica necessariamente ter como pressuposto uma forte expansão populacional, mas possíveis efeitos de deriva em algum haplogrupo com menor freqüência, ou seja, menor tamanho efetivo, mesmo em expansão não seriam filtrados porque a perda de diversidade pode correr da mesma maneira podendo gerar resultados espúrios. Os resultados obtidos por essa estimativa estão colocados na TABELA 3 e FIGURA 4.

Metodologias não baseadas na diversidade interna de cada haplogrupo, e que exigem menos pressupostos, foram também aplicadas no problema. A mais simples delas consistiu em, a partir de árvores linearizadas, chegar em uma idade para o TMRCA de cada um dos cinco haplogrupos aborígenes da América. A hipótese de que as seqüências utilizadas comportam-se de maneira "clocklike" (como um relógio molecular) não foi refutada em teste aplicado no PAML ([abacus.gene.ucl.ac.uk/software/paml.html](http://abacus.gene.ucl.ac.uk/software/paml.html)). A árvore mostrada na FIGURA 3 foi construída com os 122 nativos americanos no programa Mega3.1 por *Neighbor-*



*Joining* Tamura-Nei com correção gama ( $\alpha=0,12$  estimado no TreePuzzle 5.0 ([www.nsc.liu.se/software/biology/puzzle5](http://www.nsc.liu.se/software/biology/puzzle5))) – modelo conforme estimado pelo programa Modeltest nos testes hierárquico e Akaike ([darwin.uvigo.es/software/modeltest.html](http://darwin.uvigo.es/software/modeltest.html)) – e a taxa do relógio molecular foi mesma (1/5.140 mutações/ano). As idades obtidas mostraram-se bastante próximas e variaram entre 20 e 25 mil anos (TABELA 3, FIGURA 3 e FIGURA 4).

Para obter estimativas mais refinadas de idade, fez-se uso de metodologias probabilísticas. A idade do ancestral comum mais recente (TMRCA) para cada haplogrupo dos nativos americanos foi estimada por inferência bayesiana com 100 milhões de Cadeias de Markov Monte Carlo (MCMC). Juntamente com as estimativas de idade do TMRCA de cada haplogrupo (TABELA 3 e FIGURA 4), uma curva de variação demográfica baseada em amostragem de filogenias (*SkyLine Plot*) (FIGURA 5) foi inferida também com 100 milhões de MCMC no pacote BEAST. Na FIGURA 5 pode-se observar claramente um efeito gargalo-de-garrafa começado em torno de 27 mil anos atrás tendo o tamanho populacional mínimo de cerca de 18 mil indivíduos há 21 mil anos. A partir dessa data, que coincide com as demais datações de idade para os haplogrupos independentemente, começa um crescimento populacional bastante acentuado chegando a cerca de quatro milhões de indivíduos. Pode-se inferir que o tamanho da população fundadora do continente americano tivesse esse tamanho mínimo na variação demográfica estimado em torno de 18 mil pessoas (entre 6.916 e 75.264 com  $\alpha=0,05$ ) e então se expandiu no continente aumentando em número.

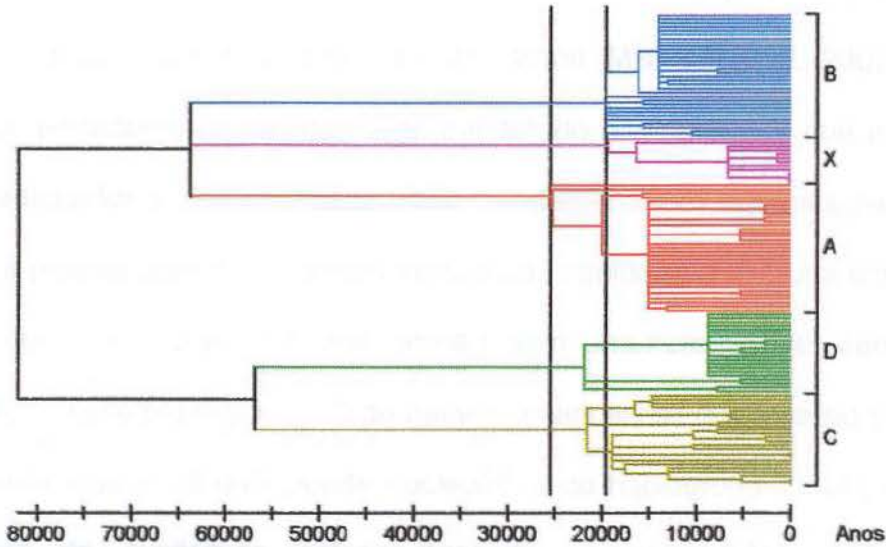
As estimativas de datações para a entrada de cada haplogrupo no continente americano através dos diferentes métodos estão dispostas de maneira



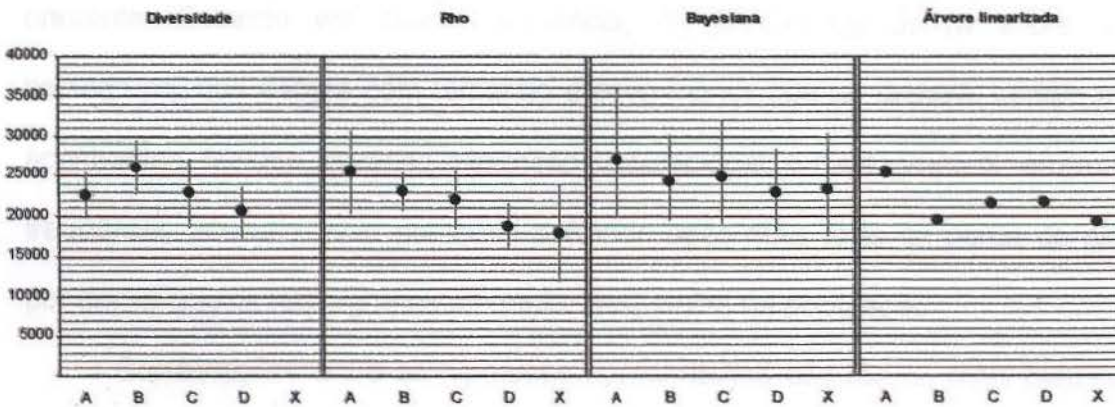
comparativa na TABELA 3 e na FIGURA 4. Não se verificam diferenças significativas entre as idades de cada haplogrupo em nenhuma das metodologias aplicadas. Em especial, o haplogrupo X, cuja diversidade nucleotídica mostrou-se significativamente menor que as dos demais haplogrupos (TABELA 2); em metodologias mais robustas, teve sua idade de expansão estimada bastante próxima à dos haplogrupos A, B, C e D. Essa menor diversidade encontrada naquele haplogrupo poderia estar relacionada a duas explicações: (i) esse haplogrupo poderia ter chegado em uma segunda leva migratória encontrando um continente já povoado no qual não conseguiu se expandir e ficou restrito à América do Norte em baixa frequência não desenvolvendo grande diversidade como consequência. Entretanto, se o haplogrupo X tivesse uma entrada mais recente nas Américas seria improvável que o seu ancestral comum mais recente (haplótipo fundador nas Américas) tivesse a mesma idade estimada dos demais haplogrupos que teriam chegado antes dele. Portanto, a hipótese que sustentaria a idéia de que o (ii) povoamento das Américas se deu em um grande evento de tomada do continente constituindo a composição genética atualmente encontrada com todos os cinco haplogrupos (inclusive o X) encontra mais suporte.

**TABELA 3:** Estimativas de idade por haplogrupo através de diferentes metodologias (as estimativas baseadas em diversidade estão colocadas na TABELA 2). O intervalo de confiança apresentado baseia-se em um  $\alpha=0,05$ .

Haplogrupo	Rho	Intervalo (Rho)	Bayesiana	Intervalo (Bayes.)	Árvore Lin.
A	25.561	30.579-20.543	27.110	35.780-20.000	25.357
B	23.121	25.403-20.839	24.340	30.060-19.600	19.484
C	22.124	25.615-18.633	24.790	31.740-19.300	21.557
D	18.839	21.530-16.148	23.030	28.280-18.220	21.778
X	17.983	23.947-12.019	23.370	30.340-17.700	19.252



**FIGURA 3:** Árvore filogenética estimada por *Neighbor-Joining* Tamura-Nei com correção gama ( $\alpha=0,12$ ) linearizada com relógio molecular com taxa de substituição de nucleotídeos de  $1,26 \cdot 10^{-8}$  mutações/sítio/ano. A, B, C, D e X representam os haplogrupos cujos TMRCA's tem suas idades estimadas conforme eixo Y em anos.



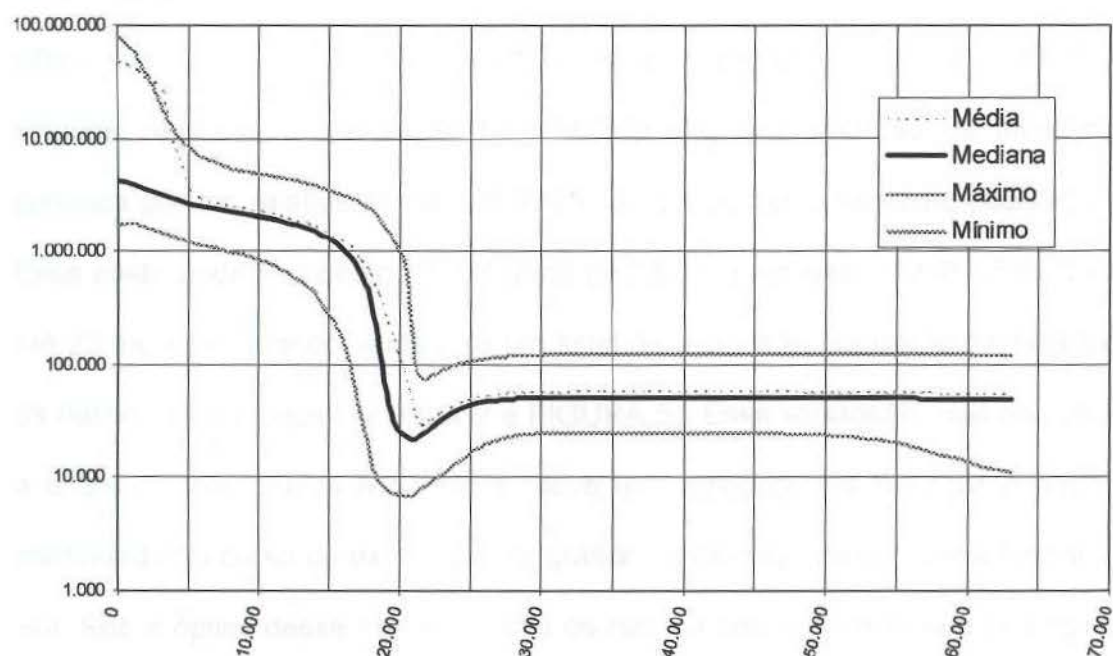
**FIGURA 4:** Gráficos comparativos entre estimativas de idade pelos diferentes métodos. Na ordem: (i) estimativas de idade de expansão de cada haplogrupo nas Américas por diversidade nucleotídica, (ii) parâmetro *Rho*, (iii) estimativa da idade dos coalescentes dos haplogrupos nativos americanos por inferência bayesiana com MCMC e (iv) pela árvore linearizada com relógio molecular (ver FIGURA 3). O eixo Y representa a idade em anos atrás; o X, o nome dos haplogrupos em cada método de estimativa.

Para explicar as peculiaridades do haplogrupo X nas Américas (suas baixas diversidade, frequência e distribuição geográfica e fraco sinal de expansão populacional), podem-se apresentar várias hipóteses. Entre elas, estariam (i) a



ação do acaso que teria levado esse haplogrupo a não se expandir no ritmo dos demais ou (ii) possíveis pressões de seleção natural (Mishmar *et al.*, 2003) contra o indivíduos portadores carregadores do mtDNA do haplogrupo X que poderiam não estar adaptados ao frio ou então sujeitos a apresentarem alguma enfermidade associada à presença de mutações específicas do haplogrupo X. Outra explicação plausível seria (iii) uma possível amostragem insuficiente de seqüências pertencentes a esse haplogrupo (12 de nativos americanos) que teria sido incapaz de captar a totalidade de diversidade nucleotídica do haplogrupo. Uma (iv) última possibilidade de explicação estaria baseada na possibilidade de, já no povoamento do continente, o haplogrupo X estar em baixa freqüência na população – como, aliás, ocorre na Europa onde esse haplogrupo também está presente. Estando em baixa freqüência, os efeitos de deriva sobre esse haplogrupo, que estaria com tamanho efetivo menor que os demais, seriam mais acentuados. Mesmo havendo expansão populacional, o haplogrupo X, em menor freqüência, passaria mais gerações sofrendo ação mais forte de perda de alelos por deriva que os haplogrupos em maior freqüência na população.

Combinações entre as hipóteses apresentadas são tão ou mais plausíveis do que elas isoladamente. É possível, por exemplo, que a freqüência do haplogrupo X fosse e ainda seja baixa nas populações em que ele ocorre por seleção natural contra mutações típicas dele. A ação do acaso na baixa freqüência desse haplogrupo na população fundadora também deve ser levada em conta se consideramos que efeitos de deriva genética tiveram e têm papel importantíssimo na evolução do *Homo sapiens* e na sua distribuição pelo Mundo.



**FIGURA 5:** Visualização gráfica do *SkyLine Plot* bayesiano com 100.000.000 de MCMC. Eixo X representa o tempo do presente para o passado; Y, o tamanho da população. Curvas representam a correlação dos eixos conforme a legenda.

A datação obtida para a expansão dos nativos americanos no continente, 23 mil anos na média, é antiga se comparada com as primeiras estimativas propostas no modelo "*Clovis First*" e, de certa forma, concordante com os achados na América do Sul e com outros trabalhos utilizando o mtDNA como marcador (Bonatto e Salzano, 1997a,b). O problema ainda não resolvido é onde se deu essa expansão populacional. À primeira vista, seria possível considerar que essa expansão poderia ter se dado somente na América continental, ao sul das grandes geleiras. Mas a existência dos grupos lingüísticos Na-Dené e Esquimó-Aleuta, ao norte das geleiras, estreitamente relacionados aos outros nativos americanos (Szathmary *et al.*, 1993) não sustenta essa idéia. Há evidências de que aquelas populações tenham participado da mesma expansão no continente sofrida pelos

Ameríndios (Bonatto e Salzano, 1997a). É mais plausível, a partir desses dados, inferir que a expansão dos nativos americanos tenha começado na Beríngia. Naquele território, a população teria sofrido efeito de redução de diversidade genética por ter, relativamente aos asiáticos, um pequeno tamanho populacional. Esse efeito pode ter começado há cerca de 28 mil anos (ver FIGURA 5) e durado até 23 mil anos quando verifica-se um sinal de expansão populacional para todos os nativos americanos (TABELA 2 e FIGURA 5). Essa expansão, que deu origem a todos os três grupos lingüísticos, deve ter começado na Beríngia e pode ter continuado no curso da expansão geográfica continente adentro até a América do Sul. Sob a óptica desse modelo, todos os nativos americanos teriam se originado na Ásia, se expandido em tamanho populacional ainda na Beríngia e, então, de forma quase contígua, podem ter se espalhado pelo continente como um todo.

As duas possíveis rotas da Beríngia para o sul das geleiras são a (i) bastante conhecida rota pelo corredor livre de gelo e (ii) a possível rota costeira pelo Pacífico desviando das geleiras. Segundo Mandryk *et al.* (2001), o corredor livre de gelo permaneceu intransponível ou inexistente entre 30 e 11,5 mil anos atrás. Pelas datas apresentadas no modelo "*Clovis First*" uma travessia por aquela região seria plausível, mas, dadas todas as evidências de diversas origens já apresentadas contestando as datas excessivamente recentes desse modelo, aquela rota não se mostra viável. Além disso, uma rota pela costa já na Beríngia seria preferível por humanos devido a questões climáticas. O interior da Beríngia permaneceu por muito tempo bastante seco e com escassez de recursos (Hall *et al.*, 2004). Portanto, a rota que foi seguida na expansão dos Ameríndios para o sul das geleiras a partir da Beríngia, pelas datações obtidas, deve ter sido pela costa



do Pacífico onde há diversas evidências de regiões climaticamente agradáveis e ricas em recursos (Mandryk *et al.*, 2001). Apesar dos evidentes problemas de se obter amostras arqueológicas de uma antiga região de costa agora submersa, há evidências de humanos naquela região há pelo menos 10,3 mil anos (Mandryk *et al.*, 2001). Havendo mais estudos concentrados em achar evidências humanas na antiga linha de costa das Américas, é bastante provável que mais vestígios de populações ali viventes há mais tempo sejam encontrados.

## CONCLUSÕES

O modelo de povoamento das Américas que poder ser apresentado a partir dos dados aqui obtidos combinados com evidências retiradas da literatura é derivado do modelo apresentado por Bonatto e Salzano (1997a,b). Naqueles trabalhos, os autores mostram evidências de uma expansão mais antiga (~30mil anos atrás) do que a encontrada neste trabalho. A estimativa de idade neste trabalho, além de se basear em métodos computacionalmente mais refinados, é feita com um marcador reconhecidamente mais confiável (Ingman *et al.*, 2000). A expansão das populações que deram origem aos nativos americanos pré-colombianos deve ter se dado, então, na Beríngia há cerca de 23 mil anos e a rota tomada pelas populações que deram origem aos Ameríndios ao sul das geleiras deve ter sido, por exclusão, uma rota costeira envolvendo estabelecimento ao longo da linha da costa da América do Norte com o Pacífico (Mandryk *et al.*, 2001).

## REFERÊNCIAS BIBLIOGRÁFICAS:

- Anderson S, Bankier AT, Barrel BG, De Bruijn MHL, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJH, Staden R e Young IG (1981) Sequence and organization of the human mitochondrial genome. *Nature* 290: 457-465.
- Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, e Howell N (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nature Genetics* 23: 147.
- Ballinger SW, Schurr TG, Torroni A, Gan Y-Y, Hodge JA, Hassan K, Chen KH e Wallace DC (1992) Southeast Asian mitochondrial DNA analysis reveals genetic continuity of ancient Mongoloid migrations. *Genetics* 130: 139-152.
- Bandelt HJ, Quintana-Murci L, Salas A e Macaulay V. (2002) The Fingerprint of Phantom Mutations in Mitochondrial DNA Data. *Am. J. Hum. Genet* 71(5):1150-1160.
- Bandelt H-J, Forster P e Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16: 37-48.
- Beheregaray LB, Ciofi C, Geist D, Gibbs JP, Caccone A e Powell JR. (2003) Genes record a prehistoric volcano eruption in the Galapagos. *Science* 302(5642): 75.
- Bonatto SL e Salzano FM (1997a) A single and early migration for the peopling of the Americas supported by mitochondrial DNA sequence data. *Proc Natl Acad Sci USA* 94: 1866-1871.

- Bonatto SL e Salzano FM (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.
- Brown MD, Hosseini SH, Torroni A, Bandelt H-J, Allen JC, Schurr TG, Scozzari R, Cruciani F e Wallace DC (1998) mtDNA haplogroup X: an ancient link between Europe/western Asia and North America? *Am J Hum Genet* 63: 1852-1861.
- Cann RL (1994) mtDNA and Native Americans: a southern perspective. *Am J Hum Genet* 55: 7-11.
- Derbeneva OA, Sukernik RI, Volodko NV, Hosseini SH, Lott MT e Wallace DC (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.
- Diamond JM (1990) The speak of the Americas. *Nature* 344: 589-90.
- Finnilä S, Lehtonen MS e Majamaa K (2001) Phylogenetic network for European mtDNA. *Am J Hum Genet* 68: 1475-1484.
- Forster P, Harding R, Torroni A e Bandelt H-J (1996) Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet* 59: 935-945.
- Glazko GV e Nei M (2003) Estimation os Divergence Times for Major Lineages of Primate Species. *Mol. Biol. Evol.* 20(3): 424-434.
- Greenberg JH, Turner II CG e Zegura SL (1986) The settlement of the Americas: a comparison of the linguistic dental and genetic evidence. *Curr Anthrop* 27: 477-497.



- Hall R, Roy D e Boling D (2004) Pleistocene Migration Routes Into the Américas: Human Biological Adaptations and Enviromental Constraints. *Evolutionary Anthropology* 13: 132-144.
- Hermstadt C, Elson JL, Fahy E, Preston G, Turnbull DM, Anderson C, Ghosh SS, Olefsky JM, Beal MF, Davis RE e Howell N. (2002) Reduced-median-network analysis of complete mitochondrial DNA coding-region sequences for the major African, Asian, and European haplogroups. *Am J Hum Genet* 70(5):1152-1171.
- Horai S, Kondo R, Nakagawa-Hattori Y, Hayashi S, Sonoda S e Tajima K (1993) Peopling of the Americas founded by four major lineages of mitochondrial DNA. *Mol Biol Evol* 10: 23-47.
- Hugall A, Moritz C, Moussalli A e Stanistic J (2002) Reconciling peledistribution models and comparative phylogeography in the Wet Tropics rainforests land snail *Gnarosophia bellendenkerensis* (Brazier 1875). *Proc Natl Acad Sci USA* 99(9): 6112-6117.
- Ingman M, Kaessmann H, Pääbo S e Gyllensten U (2000) Mitochondrial genome variation and the origin of modern Humans. *Nature* 408: 708-713.
- Kelly RL (2003) Maybe we do know when people first came to North America; and does it mean if we do? *Quaternary International* 109-110: 133-145.
- Kong QP, Yao YG, Sun C, Bandelt HJ, Zhu CL e Zhang YP (2003) Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences. *Mol Biol Evol* 20(4): 671-676.
- Kraytsberg Y, Schwartz M, Brown TA, Ebralidse K, Kunz WS, Clayton DA, Vissing J e Khrapko K (2004) Recombination of Human Mitochondrial DNA. *Science* 304: 981.



- Levin BC, Cheng H e Reeder DJ (1999) A human mitochondrial DNA standard reference material for quality control in forensic identification, medical diagnosis, and mutation detection. *Genomics* 55: 135–146.
- Maca-Meyer N, González AM, Larruga JM, Flores C e Cabrera, VM (2001) Major genomic mitochondrial lineages delineate early human expansions. *BMC Genetics* 2: 13.
- Mahli RS, Eshleman JA, Greenberg JA, Weiss DA, Shook BAS, Kastle FA, Lorenz JG, Kemp BM, Johnson JR e Glenn Smith D (2002) The structure of diversity within New World mitochondrial DNA haplogroups implications for the prehistory of North America. *Am J Hum Genet* 70: 905-919.
- Mandryk CAS, Josenhans H, Fedje DW e Mathewes RW (2001) Late Quaternary paleoenvironments of Northwestern North America: implications for inland versus coastal migration routes. *Quaternary Science Reviews* 20: 301-314.
- Meltzer DJ (1997) Monte Verde and the Pleistocene peopling of the Americas. *Science* 276: 754-755.
- Merriwether DA, Rothhammer F e Ferrel RE (1995) Distributions of the four founding lineage haplotypes in Native Americans suggest a single wave of migration for the New World. *Am J Phys Anthropol* 98: 411-430.
- Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD, Sukernik RI, Olckers A e Wallace DC (2003) Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA* 100(1): 171-176.

- Neves WA, Munford D, Zanini MC e Pucciarelli HC (1999) Cranial morphological variation and the colonization of the New World: towards a four migration model? *Ciência e Cultura* 51: 151-165.
- Prous A (1995) Archeological analysis of the oldest settlements on the Americas. *Braz J Genet* 18: 689-699.
- Reidla M, Kivisild T, Metspalu E, Kaldma K, Tambets K, Tolk H, Parik J, Loogväli E, Derenko M, Malyarchuk B, Bermisheva M, Zhadanov S, Pennarun E, Gubina M, Golubenko M, Damba L, Fedorova S, Gusar V, Grechanina E, Mikerezi I, Moisan J, Chaventré A, Khusnutdinova E, Osipova L, Stepanov V, Voevoda M, Achilli A, Rengo C, Rickards O, De Stefano GF, Papiha S, Beckman L, Janicijevic B, Rudan P, Anagnou N, Michalodimitrakis E, Koziel S, Usanga E, Geberhiwot T, Hermstadt C, Howell N, Torroni A e Villems R (2003) Origin and Diffusion of mtDNA Haplogroup X. *Am. J. Hum. Genet.* 73: 1178–1190.
- Rieder MJ, Taylor SI, Tobe VO e Nickerson DA (1998) Automating the identification of DNA variations using quality-based fluorescence re-sequencing: analysis of the human mitochondrial genome. *Nucleic Acids Res* 26: 967-973.
- Rogers CA, Rogers LA e Martin CP (1992) How the door opened: the peopling of the New World. *Hum Biol* 64: 281-302.
- Roosevelt AC, Lima da Costa M, Lopes Machado C, Michab M, Merciek N, Valladas H, Feathers J, Barnett W, Inazio da Silveira M, Henderson A, Silva J, Chemoff B, Reese DS, Holman JA, Toth N e Schick K (1996) Paleoindian cave dwellers in the Amazon: the peopling of the Americas. *Science* 272: 373-384.

- Rozas J e Rozas R (1999) DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* 15: 174-175.
- Salas A, Carracedo A, Macaulay V, Richards M e Bandelt HJ (2005) A practical guide do mitochondrial DNA error prevention in clinical, forensic and population genetics. *Biochemical and Biophysical Research Communications* 335: 891-99.
- Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E e Lundenberg J (2004) A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc Natl Acad Sci USA* 101(33): 12387-12390.
- Schurr TG, Ballinger SW, Gan Y-Y, Hodge JA, Merriwether DA, Lawrence DN, Knowler WC, Weiss KM e Wallace DC (1990) Amerindian mitochondrial DNAs have rare Asian mutations at high frequencies suggesting they derived from four primary maternal lineages. *Am J Hum Genet* 46: 613-623.
- Silva WA, Bonatto SL, Holanda AJ, Ribeiro-dos-Santos AKC, Paixão BM, Goldman GH, Abe-Sandes K, Rodriguez-Delfin L, Barbosa M, Paçó-Larson ML, Petzl-Erler ML, Valente V, Santos SEB e Zago MA (2003) Correction: Mitochondrial DNA variation in Amerindians. *Am J Hum Genet* 72: 1346-1349.
- Starikovskaya EB, Sukernik RI, Derbeneva OA, Volodko NV, Ruiz-Pesini E, Torroni A, Brown MD, Lott MT, Hosseini SH, Huoponen K e Wallace DC (2005) Mitochondrial DNA Diversity in Indegenous Populations of the Southern Extent of Siberia, and the Origins of Native American Haplogroups. *Ann Hum Gen* 69: 67-89.
- Szathmary EJE (1993) MtDNA and the peopling of the Americas. *Am J Hum Genet* 53: 793-799.



- Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123: 585-595.
- Tanaka M, Cabrera VM, González AM, Larruga JM, Takeyasu T, Fuku N, Guo L, Hirose R, Fujita Y, Kurata M, Shinoda K, Umetsu K, Yamada Y, Oshida Y, Sato Y, Hattori N, Mizuno Y, Arai Y, Hirose N, Ohta S, Ogawa O, Tanaka Y, Kawamori R, Shamoto-Nagai M, Maruyama W, Shimokata H, Suzuki R e Shimodaira H (2004) Mitochondrial Genome Variation in Eastern Asia and the Peopling of Japan. *Genome Research* 14: 1832-1850.
- Wallace DC e Torroni A (1992) American Indian prehistory as written in the mitochondrial DNA: a review. *Hum Biol* 64: 403-416.
- Zsurka G, Kraytsberg Y, Kudina, T, Komblum C, Elger CE, Khrapko K e Kunz WS (2005) *Nature Genetics* 37(8): 873-877.