

**Marília de Oliveira Scliar**

**GENÉTICA DE POPULAÇÕES E INFERÊNCIAS ESTATÍSTICAS  
SOBRE A HISTÓRIA DEMOGRÁFICA DE POPULAÇÕES  
NATIVAS SUL AMERICANAS**

Tese apresentada ao curso de Doutorado do Departamento de Biologia Geral do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Doutor em Genética.

Orientador: Prof. Dr Eduardo Martin Tarazona Santos

Universidade Federal de Minas Gerais  
Departamento de Biologia Geral  
Instituto de Ciências Biológicas  
Belo Horizonte  
Janeiro 2012

## RESUMO

No presente trabalho, geramos, pela primeira vez, dados de ressequenciamento de 10 regiões autossômicas não codificantes (~20000 pb) em 10 indivíduos da maior população andina, os Quechuas, e em 10 indivíduos de uma população Machiguenga (comunidade Shima) localizada na região conhecida como Selva Alta peruana. Esta é uma região de transição geográfica entre os Andes e a Amazônia, e por isso, pensamos que suas populações poderiam ser bons alvos na tentativa de elucidar possíveis direções de migração e povoamento entre o Oeste e o Leste do continente. Nós utilizamos duas metodologias bayesianas para inferir os parâmetros de um modelo de isolamento com migração entre populações nativas americanas e entre populações nativas americanas e asiáticas. Uma das metodologias é baseada no cálculo da verossimilhança utilizando simulações MCMC para aproximar uma probabilidade a *posteriori* (programa IM). A outra é o método ABC (Approximate Bayesian Computation, que utilizamos apenas para o modelo de isolamento com migração entre as populações Quechua e Shima) que ao invés de calcular a verossimilhança de um dado modelo, aproxima a probabilidade a *posteriori* através da comparação de estatísticas sumárias estimadas para os dados observados com as estatísticas sumárias computadas para dados simulados de um determinado modelo (no caso, o modelo de isolamento com migração). Nossos resultados identificaram uma alta diversidade na população Quechua, comparável com a de populações europeias, e uma pequena diversidade nos Shimaas, sendo que esta parece ser uma subamostra da diversidade dos Quechuas. Além disso, estimamos que quando da divergência entre as populações Quechua e Shima, 95% da população ancestral deu origem aos Quechuas, que se manteve até o presente com um grande tamanho efetivo populacional. Estimamos um tempo de divergência menor que cinco mil anos entre os Quechuas e os Shimaas, que pode se dizer surpreendente, por se tratar de populações identificadas com duas tradições completamente diferentes, a cultura amazônica associada aos Shimaas, e a cultura andina associada aos Quechuas. Nossas estimativas, entre populações nativas americanas e asiáticas, corroboraram as estimativas anteriores feitas com outros marcadores genéticos, mostrando que nossa população Quechua, na qual subsiste grande parte da variabilidade genética das populações nativas americanas, é uma excelente população para se estudar processos evolutivos relativos não só às populações andinas, mas também a todo povoamento da América.

## ABSTRACT

In this study, we generated for the first time, resequencing data for 10 autosomal noncoding regions (~ 20,000 bp) for 10 individuals of the larger Andean population, the Quechua, and for 10 individuals from the Machiguenga population (Shimaa community) located in the Selva Alta from Peru. This is a geographical region of transition between the Andes and Amazon, and therefore, we think that their populations could be good targets to elucidate possible directions of migration and settlement between the West and the East of the continent. We used two Bayesian methods to infer the parameters of an isolation model with migration among Native American populations and among Native American and Asian populations. One of the methods is likelihood-based and uses Markov Chain Monte Carlo simulations to approximate the posterior probability (IM program). The other, is the ABC method (Approximate Bayesian Computation, we only use this method for the isolation model with migration between Shimaa and Quechua populations) that instead of calculating the likelihood of a given model, it approximates the posterior probability by comparing summary statistics estimated for the observed data with the summary statistics computed for simulated data on a particular model (in this case the isolation model with migration). Our results indicate a high diversity in the Quechua population, comparable with that of European populations, and a little diversity in Shimaa, and this seems to be a subset of the diversity of Quechua. Additionally, we estimated that, when the divergence between Shimaas and Quechuas occurred, 95% of the ancestral population gave rise to the Quechua, who remained until the present with a large effective population size. We estimate a small divergence time between the Quechua and Shimaa, you could say surprising, because the two populations are identified with two completely different traditions, the Shimaas are associated with the Amazonian culture, and the Quechuas are associate with the Andean culture. Our estimates among Asian and Native American populations corroborate previous estimates using other genetic markers, showing that our population Quechua, in which remains much of the genetic variability of Native American populations, is an excellent population to study evolutionary processes concerning not only to Andean populations, but also the entire population of America.