

APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS NA CLASSIFICAÇÃO DOS HAPLOGRUPOS AMERÍNDIO E AFRICANO A PARTIR DA ANÁLISE DO DNA MITOCONDRIAL

LARISSA LUZ-GOMES^{1,2}, ANA CECÍLIA FEIO-DOS-SANTOS², BRUNO M. CARVALHO², JESSÉ BARROS-LOBATO², JEFFERSON M. GOMES², ÂNDREA K. C. RIBEIRO-DOS-SANTOS², ANA RÉGIA D. M. NEVES¹, ROBERTO CÉLIO LIMÃO DE OLIVEIRA³

¹UFPA/ Mestranda UFPA/CT/PPGEE- Av. Augusto Corrêa, nº01, Caixa Postal 8619. 66075.110 - Belém - Pará - Brasil

²UFPA/CCB/DEPAT/BIOINFO-LGHM - Av. Augusto Corrêa, nº01, Caixa Postal 8615. 66075-970 - Belém - Pará - Brasil

³UFPA - UFPA/CT/PPGEE/DEEC - Av. Augusto Corrêa, nº01, Caixa Postal 8619. 66075.110 - Belém - Pará - Brasil

E-mails: lariluz@yahoo.com.br, limao@ufpa.br

Abstract: This article shows the application of the Artificial Neural Networks (ANN) in the automatized classification with high precision and accuracy to the mitochondrial DNA lineages. In this classification, the diverse genetic lineages are classified and sub-classified in called groups of Haplogrupos/Sub-clados, which are representative of two ethnic groups that had participated of the formation of Brazilian population: Africans and Amerindians. The type of used net is the MPL and the rule of learning the Backpropagation. The net's inputs are equivalent to polymorphics nucleotides positions of the first hypervariable region of mtDNA, that returns outputs relative to the specific classification from each lineages

Keywords— Bioinformatics, mtDNA, Haplogrupos, Artificial Neural Networks, Backpropagation

Resumo: Este artigo mostra a aplicação de Redes Neurais Artificiais (RNA's) na classificação de linhagens de DNA mitocondrial de forma automatizada, precisa e acurada. Nesta classificação, as diversas linhagens genéticas são classificadas e sub-classificadas em grupos denominados de Haplogrupos/Sub-clados, os quais são representativos de dois grupos étnicos que participaram da formação da população brasileira: Africanos e Ameríndios. O tipo de rede utilizada é a MPL e a regra de aprendizagem a Backpropagation. As entradas da rede equivalem às posições nucleotídicas polimórficas da região hipervariável do DNA mitocondrial, as quais retornam como saída a classificação específica de cada linhagem.

Palavras-chave— Bioinformática, mtDNA, Haplogrupos, Rede Neural Artificial, Backpropagation

1 Introdução

Com o advento de técnicas de biologia molecular pôde-se entender, comparar e descrever melhor a variabilidade genética existente em diferentes populações do mundo. Este melhor entendimento decorre da análise direta de pontos polimórficos na molécula de DNA presente no conjunto gênico de uma população (Vigilante et al., 1991; Cavalli-Sforza et al., 1994, 1997; Underhill et al., 2000; Maca-Meyer et al., 2001).

Estudos Genéticos, baseados principalmente na análise do DNA mitocondrial (mtDNA) e DNA nuclear, têm produzido uma grande quantidade de dados biológicos relacionados à seqüências nucleotídicas, SNPs e microssatélites (Feio-dos-Santos et al., 2006; Palha et al., 2007; Ribeiro-dos-Santos et al., 2007; Rodrigues et al., 2007). Estes resultados fornecem base para as diferentes investigações genéticas, dentre as quais se destaca os estudos bioantropológicos sobre os mecanismos de dispersão e forma-

ção das populações humanas do mundo e em particular das populações brasileiras.

Uma das metodologias mais utilizadas nesta área de investigação tem sido o seqüenciamento da 1ª região hipervariável do mtDNA (HVS-I). A análise dessas seqüências tem possibilitado a classificação das linhagens de mtDNA em diferentes grupos, denominados de haplogrupos, os quais variam entre as distintas populações humanas que habitam os continentes do planeta (Torroni *et al.*, 1993; Bronw et al., 1998; Salas et al., 2002, 2004; Kivisild *et al.*, 2004; Černý et al., 2007).

Segundo Santos *et al.* (1996), por meio das análises de linhagens mitocondriais de uma população específica, pode-se inferir sobre parte da sua composição e estrutura genética, assim como elucidar seus processos históricos de formação e expansão.

Devido à grande importância da classificação das linhagens mitocondriais em haplogrupos para estudos bioantropológicos, observou-se a necessidade da criação de Redes Neurais Artificiais que façam a classificação de forma automatizada, uma vez que esta metodologia é praticada atualmente de

forma manual e comparativa, prejudicando a velocidade e curacidade da classificação, em consequência de um crescente volume de dados.

2 Fundamentos Biológicos

2.1 Caracterização do DNA mitocondrial

O DNA mitocondrial humano (mtDNA) é uma molécula pequena, circular, localizada no interior de mitocôndrias (organelas citoplasmáticas), composta por duas cadeias: uma interna, cadeia leve (rica em pirimidinas: C-T), e outra externa, cadeia pesada (rica em purinas: A-G). O tamanho desta molécula de DNA é de 16.569 pb, os quais foram seqüenciados por Anderson et al. (1981) e revisados por Andrews et al. (1999), sendo esta seqüência nucleotídica a referência utilizada nos estudos que envolvem esta molécula.

Algumas características particulares tornam o genoma mitocondrial uma excelente ferramenta para estudos de diversidade genética e evolução dos humanos. Entre elas citamos:

a) O mtDNA é herdado matrilinearmente, ou seja, é transmitido somente das mães para filhos e filhas, não sofrendo (em princípio) recombinação com o material paterno, portanto, representa uma herança haplóide (Giles et al., 1980);

b) Elevado número de cópias por célula (10 a 100), facilitando o acesso ao genoma (Álvarez et al., 2001);

c) A molécula apresenta uma rápida evolução. A região hipervariável evolui 25 vezes mais rápido que o DNA nuclear (Vigilant et al., 1991; Wallace, 1995; Pena et al., 2002);

d) Possui um ineficiente mecanismo de reparo de mutações (Saccone et al., 1999; Croteau et al., 1999).

2.2 Polimorfismos e Haplogrupos

O mtDNA é uma das ferramentas que tem sido amplamente utilizada para o esclarecimento da dinâmica de formação e evolução de populações humanas. A maior parte destes estudos utiliza o seqüenciamento da Alça-D, em função de sua elevada taxa de mutação (Torrioni et al., 1993; Bronw et al., 1998; Salas et al., 2002, 2004; Kivisild et al., 2004; Černý et al., 2007).

As mutações que ocorrem no mtDNA são sempre herdadas em conjunto e constituem um haplótipo. Os haplótipos, por sua vez, são reunidos em diferentes haplogrupos, os quais variam entre os grupos étnicos presentes na atualidade. Desse modo, o que se observa é uma identidade, uma marca genética (*motiff*), distinta entre grupos populacionais, resultante de uma rápida e recente evolução na molécula de mtDNA (Pena, 2002).

No Brasil, tem-se a prevalência de haplogrupos mitocondriais ameríndios, africanos e europeus em função da história de ocupação deste território (Santos et al., 1999; Pante-de-Souza et al., 1998; Alves-Silva et al., 2000). Atualmente, a classificação de cada linhagem mitocondrial, dentre os mais de 50 subtipos africanos e ameríndios propostos na literatura, é totalmente desenvolvida de forma manual. Este processo, que é de grande importância para estudos com mtDNA, demanda muito tempo e pode gerar erros de classificação.

Com o desenvolvimento da tecnologia de seqüenciamento, têm-se uma produção cada vez mais crescente de seqüências de mtDNA. Portanto, o presente trabalho tem como objetivo contribuir com um processo automatizado, preciso e acurado de classificação das inúmeras linhagens de mtDNA, encontrando-se tangente à Bioinformática, a qual reúne conhecimentos científicos de diferentes áreas, como informática, estatística, biologia e genética.

Para este fim, construiu-se uma rede neural artificial do tipo Multi-Layer Perceptron (MLP) para fazer um aplicativo denominado de Sistema de Classificação de Linhagens Mitocondriais - MITCLASS.

3 As Redes Neurais Artificiais na classificação dos Haplogrupos

3.1 Rede Neural Artificial: Descrição Sucinta

As Redes Neurais Artificiais (RNA's) são modelos computacionais não lineares, inspirados na estrutura e operação do cérebro humano. São representadas através de grafos, Figura 1, os quais caracterizam a topologia da rede.

Os neurônios são dispostos em camadas, interligados por um grande número de conexões. Essas conexões estão associadas a pesos que armazenam o conhecimento representado no modelo e ponderam as entradas recebidas por cada neurônio da rede. Há também a regra de aprendizado ou atualização que controla os pesos de cada neurônio.

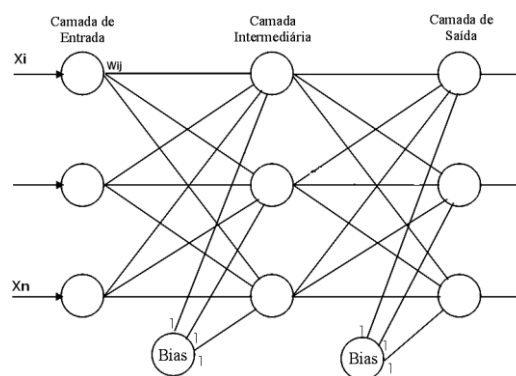


Figura 1. Grafo de uma RNA do tipo MLP.

3.2. Perceptron Múltiplas Camadas (Multi-Layer Perceptron (MLP))

O Perceptron Múltiplas Camadas (MLP) consiste em uma rede que possui n neurônios na camada de entrada, k neurônios na camada intermediária e p neurônios na camada de saída, como pode ser visto na Figura 1. Todos os neurônios estão conectados no sentido direto por meio dos pesos w_{ij} .

As entradas (x_i) são multiplicadas pelos pesos de cada conexão (w_{ij}) e propagadas para os neurônios da camada seguinte. Nestes neurônios há uma função de ativação, que irá limitar a amplitude do sinal de saída do neurônio, representada por T , que é a função não linear tangente hiperbólica, $f(z) = 1 - e^{-z} / 1 + e^{-z}$. As entradas multiplicadas pelos pesos (net) são então recebidas pela função tangente hiperbólica para produzir a saída desta camada. Os neurônios desta camada tem como equação:

$$K_j = 1 - \exp / 1 + \exp \left(- \sum_{i=1}^n w_{ij} \cdot x_i \right) \quad (1)$$

Onde K_j é o j -ésimo neurônio da camada intermediária e X_i representa a i -ésima entrada conectada a este neurônio intermediário pelo peso W_{ij} onde $i=1 \dots n$.

O resultado para o K -ésimo neurônio da camada de saída da rede será:

$$X_k = 1 - \exp / 1 + \exp \left(- \sum_{j=1}^n w_{kj} \cdot x_j \right) \quad (2)$$

Onde X_k é o k -ésimo neurônio da camada de saída e $k = h+1, \dots, p$, e W_{kj} representa os pesos das conexões entre a camada intermediária e a de saída.

Para o treinamento do MLP utilizou-se o algoritmo de aprendizagem de Retropropagação do Erro (*Backpropagation*). Neste algoritmo o aprendizado ocorre através do ajuste dos pesos. A rede aprende um conjunto pré-definido de pares de exemplos entrada/saída em ciclos de propagação/adaptação. O algoritmo de retropropagação é dividido em duas etapas.

Na primeira o vetor de treinamento é apresentado à rede e propagado por todos os nós até a camada de saída. Compara-se o valor de saída com o valor desejado, calculando-se o erro médio quadrático. O valor de ΔW é calculado e os pesos da camada de saída são atualizados.

A segunda etapa consiste na retropropagação, de forma que cada camada recalcula o valor de ΔW e atualiza seus pesos até a camada de entrada. Após atualizar todos os pesos da rede, o processo recomeça. Quando uma época (um ciclo completo) é completada, a tendência é que o erro quadrático médio diminua, até alcançar um mínimo pré-estabelecido, momento no qual a rede converge.

4 Modelagem da Rede Neural Artificial

Em populações brasileiras, miscigenadas do ponto de vista biológico, encontramos diferentes linhagens mitocôndrias de diferentes grupos étnicos. Inicialmente tratamos as entradas de forma sistemática e mapeada, no sentido de alcançar resultados relacionados não somente a uma classificação genética, mas também a uma classificação étnico-geográfica.

Em relação aos haplogrupos mitocondriais ameríndios, têm-se a prevalência de cinco principais grupos (A, B, C, D e X) (Torrioni *et al.*, 1993; Bronw *et al.*, 1998). Por outro lado, dentre os grupos mitocondriais genéticos de origem africana, têm-se até o presente momento seis grandes haplogrupos (L0 à L5). Todavia, neste universo de variabilidade genética, a literatura específica propõe ainda uma sistemática de classificação em até 61 diferentes subgrupos, denominados de sub-clados (Salas *et al.*, 2002, 2004; Kivisild *et al.*, 2004; Černý *et al.*, 2007).

Em relação à modelagem, inicialmente foi feito o mapeamento qualitativo e quantitativo das entradas a serem utilizadas na rede, para que a mesma possua uma alta precisão e acurácia de classificação. Para isso, os dados de entrada referentes aos 400 pontos polimórficos, que compõe a HVS-I do mtDNA, em associação com SNPs (*Single Nucleotide Polimorphism*) dispersos ao longo da molécula, variaram no intervalo de -1 a 1. Para os treinos, testes e validações utilizou-se o programa Matlab na versão 6.5.

Após esta etapa, foram feitas duas redes, a primeira, referente à classificação dos haplogrupos ameríndios, foi modelada com 16 neurônios na camada de entrada, 16 na camada intermediária e 5 na camada de saída, os quais equivalem aos 5 diferentes haplogrupos. A segunda rede, referente aos haplogrupos/sub-clados africanos, foi modelada com 82 neurônios de entrada, 82 na camada intermediária e 61 na camada de saída, os quais representam toda variabilidade genética proposta para este grupo étnico.

Em relação ao treinamento de cada rede, utilizou-se amostras de mtDNA cadastradas e disponibilizadas no LGHMitdatabase (banco de dados nacional de seqüências de mtDNA, do Laboratório de Genética Humana e Médica da Universidade Federal do Pará, elaborado pela equipe de bioinformática do laboratório e disponibilizado no site: www.lghm.ufpa.br/bioinformatica). Estas amostras foram selecionadas de forma a refletir toda a variabilidade genética a qual se deseja classificar pelo MITCLASS.

O presente trabalho direcionou a construção das redes neurais para a classificação de haplogrupos ameríndios e africanos, visto que esta escolha possibilita inferir sobre a capacidade de discernimento de uma RNA entre seqüências nucleotídicas

com significativa variação entre os seus pontos polimórficos (grupo de mtDNA de ancestralidade africana), assim como seqüências com pouca variação entre seus pontos polimórficos (grupo de mtDNA de ancestralidade ameríndios).

5 Análise dos Resultados

Individualmente, cada rede apresentou treinamento satisfatório em função do alcance do pequeno erro médio quadrático de 10^{-3} estipulado na construção da rede (Figura 3(a) e (b)).

Adicionalmente, as redes, como sistema de classificação de seqüências de mtDNA, apresentaram resultados excelentes em relação a sua precisão e acurácia, pois obtiveram 100% de êxito, como pode ser observado nas tabelas 1(a), 1(b), 1(c) e 1(d), tabelas estas, que evidenciam algumas das saídas dos neurônios obtidos após validações das redes. Assim como, nas figuras 2(a) e 2(b) que mostram os gráficos com os dados de saída obtidos antes do treino (cor azul), dados de saída obtidos durante o treino (cor verde) e os dados de saída obtidos após o treino da rede, cor vermelha, mostrando o aprendizado da rede.

Os treinos foram feitos com 70% do total dos dados e os testes e validações com os 30% restantes. É importante ressaltar que as redes de classificação ameríndia e africana foram testadas com 104 e 210 amostras respectivamente.

Haplogrupo	Amostras Testadas			
	A	A	B	B
A	0.99033	0.99181	-0.010184	0.0066143
B	0.019615	-0.0081131	0.98037	0.98247
C	0.024809	0.027926	-0.00049744	0.041297
D	0.027103	0.0040528	0.018634	0.015265

Tabela 1(a). Imagens das respostas, de algumas das 104 amostras de Ameríndio, obtidas após testes feitos pela rede de Ameríndio. Precisão de 100%.

Haplogrupo	Amostras Testadas			
	C	C	D	D
A	-0.0062673	0.036842	0.02702	0.024065
B	-0.0094023	-0.00046159	0.020751	0.014381
C	0.99173	0.98611	-0.017785	0.0032764
D	0.022459	0.0065963	0.99503	0.99517

Tabela 1(b). Imagens das respostas, de algumas das 104 amostras de Ameríndio, obtidas após testes feitos pela rede de Ameríndio. Precisão de 100%.

Haplogrupo	Amostras Testadas			
	L0	L0	L0a	L0a
L0	0.9758	0.97579	-0.0039894	-0.006453
L0a	-0.0025497	-0.0035715	0.8783	0.87701
L0a1	-0.042845	-0.043732	-0.0015347	0.0012684
L0a1a	0.019364	0.022048	0.0057044	0.0066317

Tabela 1(c). Imagens de partes das respostas, de algumas das 989 amostras de Africano, obtidas após testes feitos pela rede de Africano. Precisão de 100%.

Haplogrupo	Amostras Testadas			
	L0a1	L0a1	L0a1a	L0a1a
L0	0.0073885	0.0061824	-0.0003767	-0.0042484
L0a	0.013561	0.011819	-0.0051648	-0.0073062
L0a1	0.89119	0.89126	0.0062745	0.0074568
L0a1a	0.006497	0.0063533	0.93733	0.93669

Tabela 1(d). Imagens de partes das respostas, de algumas das 989 amostras de Africano, obtidas após testes feitos pela rede de Africano. Precisão de 100%.

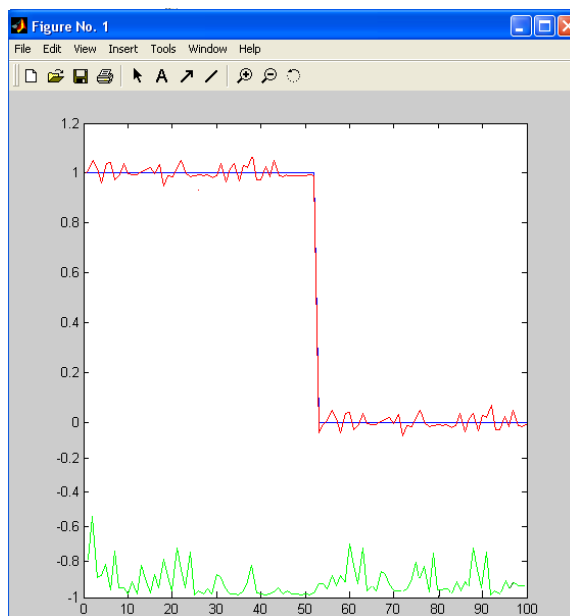


Figura 2(a). Gráfico dos dados de saída de Ameríndio, antes do treino (cor azul), dados durante o treino (cor verde) e depois do treino, validação (cor vermelha).

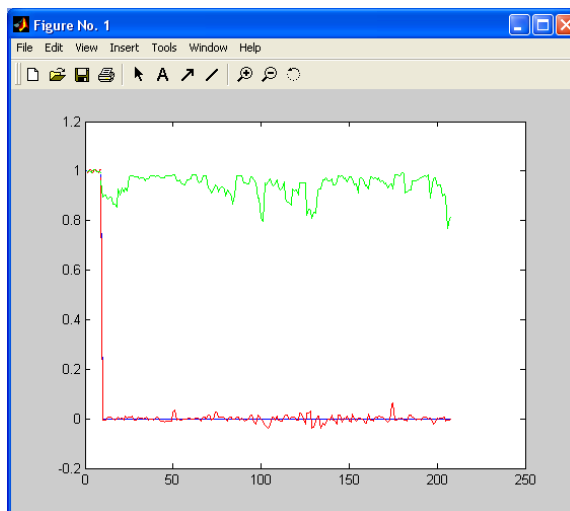


Figura 2(b). Gráfico dos dados de saída de Africano, antes do treino (cor azul), dados durante o treino (cor verde) e depois do treino (cor vermelha).

Os resultados positivos no uso de RNA's para a classificação de seqüências de DNA, mais especificamente o mtDNA que foi utilizado no presente estudo, devem estar diretamente relacionados com as escolhas acertadas dos pontos polimórficos da região HVS-I e SNPs e o ajuste funcional dessas informações para a camada de neurônios de entrada das redes.

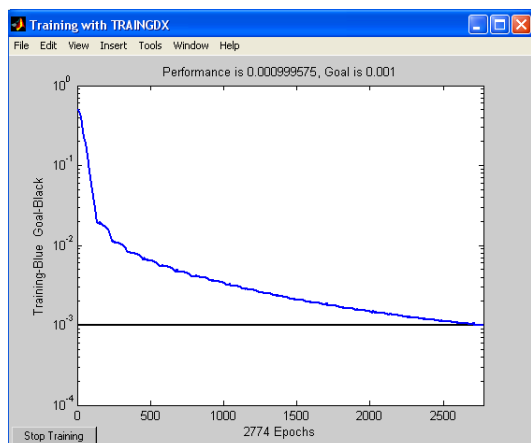


Figura 3(a). Gráfico da taxa de aprendizado da RNA para classificação do haplogrupo ameríndio.

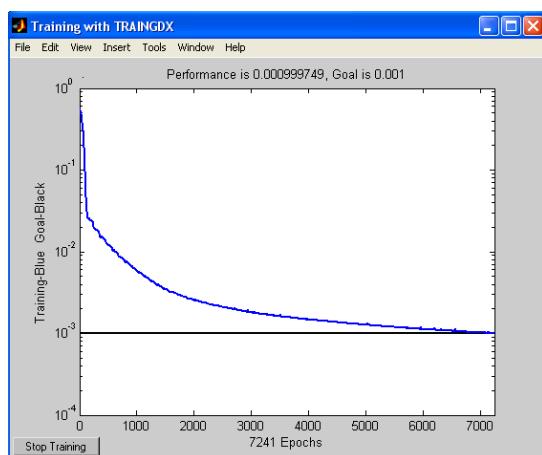


Figura 3(b). Gráfico da taxa de aprendizado da RNA para classificação do haplogrupo africano.

6 Conclusões

A partir dos resultados obtidos neste trabalho multidisciplinar, observa-se que a RNA do tipo MLP, com algoritmo de aprendizagem Backpropagation, realiza com precisão e acurácia a classificação automatizada de todos os 66 diferentes haplogrupos e subclados de linhagens mitocondriais de origem ameríndia e africana. Verifica-se ainda que as redes possuem, neste caso, a capacidade de associar e generalizar padrões desconhecidos mesmo com um número reduzido de amostras representativas de cada grupo genético no decorrer dos treinos.

Como perspectiva, o MITCLASS será implementado como um aplicativo web no site do LGH-MitDatabase da UFPA para que o usuário possa fazer a classificação de forma automática e on-line de qualquer linhagem mitocondrial, visando facilitar a classificação das seqüências dos usuários e desta forma ampliar os conhecimentos biológicos e antropológicos por meio da análise das seqüências de mtDNA.

Referências Bibliográficas

- Álvarez, J.C., Entrala, C. Lorente, J.A., Fernández-Rosado, F.J., Martínez-Spín, E., Rodríguez, E. Villanueva, E. *Análisis de ADN. Mitocondrial. Forénsica*, 1: 40-59, 2001.
- Alves-Silva, J., Santos, M.S., Guimarães, P.E.M., Ferreira A.C.S., Bandelti, H.J. Pena, S.D.J., Prado, V.F. *The Ancestry of Brazilian mtDNA Lineages*. American Journal of Human Genetics, 67: 444-461, 2000.
- Anderson, S., Bankier, A.T., Barrell, B.G., De Bruijn, M.H.L., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Rose, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R., Young, I.G. *Sequence and Organization of the Human Mitochondrial Genome*. Nature, 290: 457-467, 1981.
- Andrews, R.M., Kubacka, I., Chinnery, P.F., Lightowler, R.N. Turnbull, D.M., Howell, N. *Reanalysis and Revision of the Cambridge Reference Sequence for the Human Mitochondrial DNA*. Nature Genetics, 23: 147, 1999.
- Barreto, Jorge Muniz. *Inteligência Artificial no limiar do século XXI*. 2ed. Florianópolis: Duplic, 1999. 296p.
- Brown, M.D. Hosseini, S.H., Torroni, A. Bandelt, H.J., Allen, J.C., Schurr, T.G., Scozzari, R., Cruciani, F., Wallace, D.C. *mtDNA haplogroup X: An Ancient Link Between Europe/Western Asia and North America?* American Journal of Human Genetics, 63: 1852-1861, 1998.
- Cavalli-Sforza, L.L.; Menozzi, P.; Piazza, A. *The History and Geography of Human Genes*. Princeton, New Jersey, Princeton University Press, 1994. 518p
- Cavalli-Sforza, L.L. (1997) *Genes, peoples, and languages*. Proceeding of the National Academy of Science USA, 94:7719 – 7724
- Cerny V, Salas A, Hajek M, Zaloudkova M, Brdicka R. *A Bidirectional Corridor in the Sahel-Sudan Belt and the Distinctive Features of the Chad Basin Populations: A History Revealed by the Mitochondrial DNA Genome*. Ann Hum Genet. 2007 Jan 17.
- Croteau, D.L., Stierum, R.H., Bohr, V.A., *Mitochondrial DNA Repair Pathways*. Mutation Research, 434: 149-159, 1999.
- Feio-dos-Santos, Ana Cecília, Carvalho, Bruno Maia, Santos, Sidney Emanuel Batista e Ribeiro-dos-Santos, Ândrea Kely Campos. *Nucleotide variability of HV-I in admixed population of the Brazilian Amazon Region*. Forensic Science International, Volume 164, Issues 2-3, 20 December 2006, Pages 276-277.
- Giles, R.E., Blanc, H., Cann, H.M., Wallace, D.C. *Inheritance of Human Mitochondrial DNA*.

- Proceeding of the National Academy of Science USA, 77: 6715-6719, 1980.
- Haykin, Simon, *Neural Network: A Comprehensive Foundation*. MacMillan College Publishing Co. New York, 1994.
- Haykin, Simon. *Redes Neurais - Princípios e Práticas*. 2ed. Tradução: Paulo Martins Engel. Porto Alegre: Bookman. 2001.
- Kivisild, T., Reidla, M.; Metspalu, E.; Rosa, A.; Brehm, A.; Pennarun, E.; Parik, J.; Geberhiwot, T.; Usanga, E.; Villems, R. *Ethiopian mitochondrial DNA heritage: tracking gene flow across and around the gate of tears*. American Journal of Human Genetics, 75: 752 – 770. 2004.
- Kovács, Zsolt László. *Redes neurais artificiais: fundamentos e aplicações*. 2ed. rev. aum. São Paulo: Collegium Cognitio, 1996.
- Maca-Meyer, N., Gonzáles, A. M., Larruga, J.M., Flores, C., Cabrera, V.M. *Major Genomic Mitochondrial lineages delineate early human expansions*. BMC Genetics, 2:13, 2001.
- Palha, Teresinha de Jesus Brabo Ferreira, Rodrigues, Elzemar Martins Ribeiro e Santos, Sidney Emanuel Batista. *Y-chromosomal STR haplotypes in a population from the Amazon region*. Brazil. Forensic Science International, Volume 166, Issues 2-3, 2 March 2007, Pages 233-239.
- Pena, S.D.J. *Retrato Molecular do Brasil*, Versão 2001. In: *Homo Brasilis: aspectos genéticos, lingüísticos e históricos e socioantropológicos da formação do povo brasileiro*. Pena, S.D.J. Ribeirão Preto, São Paulo. FUNPEC, 2002. p. 11 - 28.
- Pante-De-Souza, G.; Mousinho-Ribeiro, R.C.; Santos, E.J.M.; Zago, M.A.; Guerreiro, J.F. *Origin of the hemoglobin S gene in a northern Brazilian population: the combined effects of slave trade and internal migration*. Genetics and Molecular Biology, 21:427 – 430, 1998.
- Ribeiro-dos-Santos, Ândrea Kely Campos, Carvalho, Bruno Maia, Feio-dos-Santos, Ana Cecília e Santos, Sidney Emanuel Batista. *Nucleotide variability of HV-I in Afro-descendants populations of the Brazilian Amazon Region*. Forensic Science International, Volume 167, Issue 1, 22 March 2007, Pages 77-80.
- Rodrigues, EMR; Palha TJBF and Santos SEB. *Allele frequencies data and statistic parameters for 13 STR loci in a population of the Brazilian Amazon Region*. Forensic Science International, Volume 168, Issues 2-3, 24 May 2007, Pages 244-247.
- Saccone, C., De Giorgi, C., Gissi, C., Pesole, G., Reyes, A. *evolutionary Genomics in Metazoa: the Mitochondrial DNA as a Model System*. Gene, 238: 195-209, 1999.
- Salas, A., Richards, M., De La Fe, T., Lareu, M.V., Sobrino, B., Sánchez-Diz, P., Macaulay, V., Carracedo, A. *The making of the African mtDNA Landscape*. American Journal of Human Genetics, 71: 1082-1111, 2002.
- Salas, A., Richards, M., Lareu, M.V., Scozzari, R., Coppa, A., Torroni, A., Macaulay, V., Carracedo, A. *The African Diaspora: Mitochondrial DNA and the Atlantic Slave Trade*. American Journal of Human Genetics, 74: 000-000, 2004.
- Santos, S.E.B., Ribeiro-Dos-Santos, A.K.C., Meyer, D., Zago, M.A. *Multiple Founder haplotypes of mitochondrial DNA in Amerindians revealed by RFLP and sequencing*. Annals of Human Genetics, 60: 305-319, 1996b.
- Santos, S.E.B., Rodrigues, J.D., Ribeiro-Dos-Santos, A.K.C., Zago, A. *Differential Contribution of Indigenous Men and Women to the Formation of an Urban Population in the Amazon Region as Revealed by mtDNA and y-DNA*. American Journal Physical Anthropology, 109: 175-180, 1999.
- Torroni, A., Schurr, T.G., Cabell, M.F., Brown, W.M., Nell, J.V., Larsen, M., Smith, D.G., Vullo, C.M. & Wallace, D.C. *Asian Affinities and Continental Radiation of the four Founding Native Amerindians mtDNAs*. American Journal of Human Genetics, 53: 563-590. 1993.
- Underhill, P.A.; Shen, P.; Lin, A.A.; Jin, L. Passarino, G.; Yang, W.H.; Kauffman, E.; Bonne-Tamir, B.; Bertranpetit, J.; Francalacci, P.; Ibrahim, M.; Jenkins, T.; Kidd, J.R.; Mehdi, S.Q.; Seielstad, M.T.; Wells, R.S.; Piazza, A.; Davis, R.W.; Feldman, M.W.; Cavalli-Sforza, L.L.; Oefner, P.J. *Y Chromosome sequence variation and the history of the human population*. Nature Genetics, 26: 358-361 (2000)
- Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K., Wilson, A.C. *African Populations and the Evolution of Human Mitochondrial DNA*. Science, 253: 1503-1507, 1991.
- Wallace, D. C. *Mitochondrial DNA Variation in Human Evolution, Degenerative Disease, and Aging*. American Journal of Human Genetics, 57: 201-223, 1995.