

Notas Científicas

Evidências genético-arqueológicas sobre a origem do feijão comum no Brasil

Fábio de Oliveira Freitas⁽¹⁾

⁽¹⁾Embrapa Recursos Genéticos e Biotecnologia, Caixa Postal 02372, CEP 70770-900 Brasília, DF. E-mail: fabiof@cenargen.embrapa.br

Resumo – Neste trabalho se discute a origem do feijão comum, *Phaseolus vulgaris* L. Amostras modernas e arqueológicas foram analisadas geneticamente, utilizando-se seqüências da proteína faseolina (*Phs*). A amostra arqueológica foi encontrada em uma caverna no Norte de Minas Gerais. Os resultados evidenciam que esta amostra se relaciona mais com as variedades de feijão encontrados no Norte da América do Sul e México, o que sugere influências culturais remotas entre aquelas regiões e Minas Gerais. Além disto, deve ter havido um único evento de domesticação, com local provável entre o Norte da América do Sul e o México.

Termos para indexação: *Phaseolus vulgaris*, evolução, faseolina, migração humana, domesticação, redes de alelos.

Genetic-archaeological evidences about the origin of common bean in Brazil

Abstract – This work discusses the origin of common bean, *Phaseolus vulgaris* L. Modern and archaeological samples were genetically analyzed, using sequences of phaseolin (*Phs*). The archaeological sample was found in a cave in northern Minas Gerais State. Our results showed that this sample is close to those found in Northern South America and Mexico, indicating cultural influences in the past, between those regions and Minas Gerais. Besides, there must have been a single domestication event, probably between Northern South America and Mexico.

Index terms: *Phaseolus vulgaris*, evolution, phaseolin, human migration, domestication, allele networks.

A origem evolutiva do gênero *Phaseolus* e sua diversificação primária ocorreram nas Américas (Vavilov, 1931, citado por Debouck, 1991), mas o local exato onde isto se deu é ainda motivo de controvérsia (Gepts & Debouck, 1991).

Populações selvagens de feijão crescem, atualmente, desde o Norte do México até o Norte da Argentina, em altitudes entre 500 e 2.000 m, e não são encontradas naturalmente no Brasil (Debouck, 1986).

Vestígios arqueológicos da espécie cultivada chegam a idades próximas de 10.000 anos (Gepts & Debouck, 1991). A ampla área de ocorrência de populações selvagens da espécie é um dos fatores que permitiram o surgimento de diversas raças locais, embora também seja uma das causas da dificuldade de localização exata dos locais de domesticação desta cultura.

Pesquisas moleculares, que têm como alvo principal o gene *Phs*, codificador da proteína faseolina, são

atualmente a ferramenta mais utilizada em estudos evolutivos sobre o feijão. Pelo menos dez tipos dessa proteína já foram encontrados em variedades cultivadas e populações selvagens de feijão (Gepts et al., 1986), com uma alta correlação entre o tipo e o local geográfico de origem dos materiais, principalmente em relação ao material selvagem (Tabela 1).

Dados recentes sugerem que as variedades atuais de feijão são o resultado de múltiplos eventos de domesticação, com dois centros primários, um na América Central e o outro ao Sul dos Andes (Sul do Peru, Bolívia, Norte da Argentina). Um terceiro centro é ainda sugerido na região da Colômbia (Debouck, 1986; Gepts & Debouck, 1991).

No Brasil, a falta de estudos com amostras arqueológicas locais de feijão dificulta a reconstituição da história regional desta espécie, como, por exemplo, quais tipos foram introduzidos, quando, por onde, por quais grupos humanos, entre outros questionamentos.

Este trabalho teve como objetivo definir a espécie à qual pertence a amostra arqueológica encontrada em uma caverna, em Januária, MG, e compará-la geneticamente às amostras de diferentes regiões das Américas, cada qual com tipos distintos da proteína faseolina, a fim de se determinar a provável origem do material arqueológico, considerando as relações culturais das antigas populações humanas de Januária e a evolução desta espécie.

O material arqueológico do estudo é oriundo da caverna Lapa do Boquete, na região do Vale do Peruaçu, Município de Januária, no norte do Estado de Minas Gerais, a qual possui vestígios de ocupação humana de 10.000 anos atrás (Prous et al., 1984). Foi obtido DNA para análise a partir de uma única semente.

Foram analisadas dez amostras modernas de feijão (*Phaseolus vulgaris*), oriundas do CIAT, na Colômbia, com tipos distintos de faseolina, previamente conhecidos (Tabela 2). Sequências de *P. vulgaris* e *P. lunatus*, de bancos de dados (National Center for Biotechnology Information, 1999), foram ainda acrescentados.

Os ácidos nucléicos foram obtidos da extração por CTAB, com uma purificação secundária (Allaby et al., 1997). Duas regiões alvos do gene *Phs* foram estudadas. A primeira região (PCR1) varia entre 245 e 278 pb e abrange o 6° exon e parte do flanco 3'. Os primers utilizados foram [5' CTGTGGATCCACGTGTTGGGGC TTACGTTT 3'], a montante e [5' CTGTGGATCCAAG AAGTGAGATGGAGCTCAG 3'], a jusante.

A segunda região alvo do *Phs* (PCR2), corresponde a um fragmento que varia entre 245 e 260 pb, abrange todo o 4° exon deste gene e parte do 4° intron. Os primers utilizados foram [5' CTGTGGATCCATAGA GCAAATTCGAGGAGATC 3'], a montante, e [5' CTGTGGATCCATGGTTTTTCT TTGTATTT 3'], a jusante. A escolha de duas regiões foi feita porque a região PCR1 permite que se defina a espécie da amostra e a região PCR2 é mais propícia para se visualizar a diversidade genético-geográfica entre os diferentes tipos desta proteína.

Os parâmetros utilizados nos ciclos da PCR foram: 93,5°C, 2 min e 30 s; 60°C, 1 min; 74°C, 1 min; 93,5°C, 1 min (40 ciclos); 60°C, 1 min; 74°C, 8 min. Os seguimentos amplificados foram restringidos com *Bam*HI, clonados no vírus M13mp18 e, então, sequenciados com o sequenciador de DNA ABI 377. As sequências foram alinhadas com CLUSTALW (Thompson et al., 1994) e ajustadas a olho.

Múltiplas sequências foram obtidas de cada amostra para remover erros de polimerase. As sequências foram comparadas pela técnica de “network” (Allaby & Brown, 2001). Todos os cuidados, em relação a se evitar a contaminação do material arqueológico por DNA moderno, foram tomados, a fim de se assegurar a autenticidade do DNA da amostra arqueológica (Freitas et al., 2003). A fim de validar as sequências arqueológicas, só foram utilizadas sequências que tivessem pelo menos quatro clones.

A idade da amostra arqueológica é de 251±39 AP (anos antes do presente; ano base 1950), e remonta aos anos de 1660 a 1738 de nossa era, portanto após a chegada dos europeus. Entretanto, como a chegada e influência dos colonizadores não se fizeram presentes na região de Januária, até o início do século XVIII (Vasconcelos, 1974), esta amostra é um provável representante do material que os índios da região cultivavam desde tempos mais remotos, o que é corroborado por outras amostras pré-coloniais ali descobertas (Prous et al., 1984; Freitas et al., 2003).

Com a análise genética da região PCR1, da amostra arqueológica de Januária, foi possível identificá-la como pertencente à espécie *P. vulgaris*. Além disto, nesta região se constatou que o tipo da proteína faseolina, presente na amostra arqueológica, é do tipo α e não β , em razão da presença da repetição direta (27 bp) na posição 1.318 do gene, classificação esta diferente em relação aos tipos apresentados na Tabela 1 (Kami & Gepts, 1994).

Na análise dos dados da região PCR2, as sequências foram comparadas por meio de um gráfico de “network” (Figura 1), que estima a divergência evolutiva entre elas. Foram utilizadas sequências de *P. lunatus* como um

Tabela 1. Tipo de faseolina em populações de feijão selvagem e domesticado e sua origem geográfica (adaptado de Gepts & Debouck, 1991).

Região	Tipo de faseolina	
	Material selvagem	Material domesticado
América Central e México	“S”; “M”	“S” (92%); “T” (8%)
Colômbia	“B”; “CH”	“S” (64%) “T” (26%) “C” (7%) “B” (3%)
Andes (exceto a Colômbia)	“T”	“T” (50) “S” (17%) “A” (1%) “H” (1%) “P” “I”

Tabela 2. Amostras modernas e arqueológicas utilizadas nas análises, seu número de identificação e local onde foram coletadas⁽¹⁾.

Sigla ⁽²⁾	Número CIAT	Localização	Tipo de faseolina
8	G 19890	Salta, Argentina	T
9	G 19895	Tucuman, Argentina	J
10	G 23463	Cundinamarca, Colômbia	B
11	G 23589	Apurimac, Peru	H
12	G 24423	Cundinamarca, Colômbia	CH
13	G 23583	Piura, Peru	I
14	G 23458	Cuzcu, Peru	P
15	G 23576	Cuzcu, Peru	C
16	G 11034	Durango, México	M
17	G 12935	Jalisco, México	S
A	-	Januária, MG, Brasil	-

⁽¹⁾Para análise final incluíram-se seqüências obtidas do banco mundial (National Center for Biotechnology Information, 1999). ⁽²⁾Para cada amostra podem existir diferentes alelos.

“out group”, o que indicou que as seqüências mais primitivas de *P. vulgaris* pertencem à amostra B 10-2, da Colômbia. Isto sugere que os primeiros alelos da espécie *P. vulgaris* deveriam ser mais próximos dos que são hoje encontrados naquela amostra, o que não significa que ali tenha se dado a origem da espécie. Assim, quanto mais afastados os alelos se encontrarem do alelo-base (B 10-2), mais novos eles tendem a ser (Allaby & Brown, 2001), como, por exemplo, o alelo T 8-1 (Figura 1).

Na porção centro-esquerda do gráfico observam-se diversas seqüências/alelos interligados, alguns compartilhados por diversos indivíduos, o que indica um grande fluxo gênico entre as populações de feijão que continham estes alelos e/ou retenção de polimorfismos ancestrais. Esta região do gráfico foi denominada tronco.

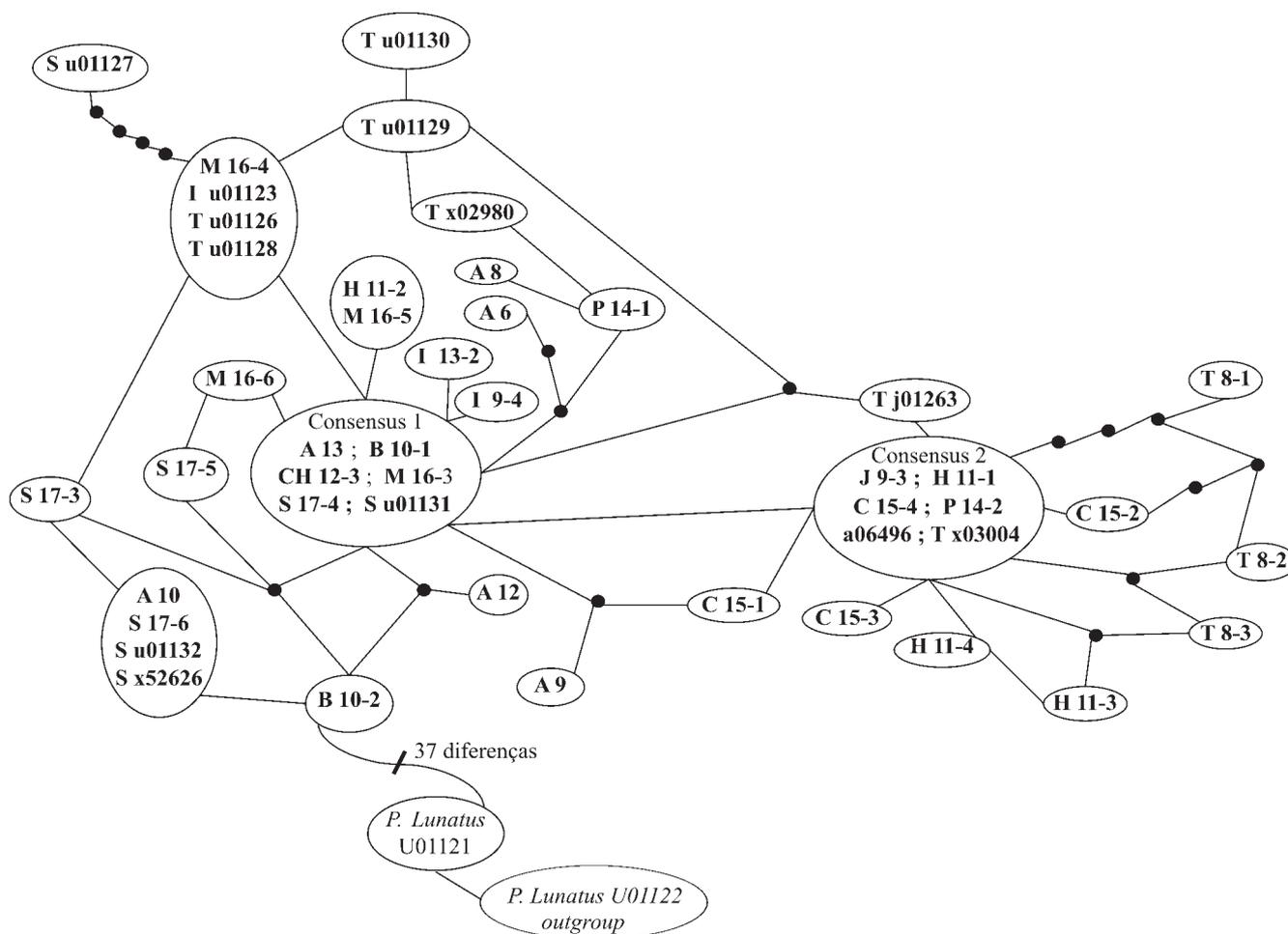


Figura 1. Gráfico de similaridade tipo “network”, que mostra a relação evolutiva entre as seqüências (alelos) de cada amostra. Cada amostra é identificada pelo tipo da proteína faseolina (primeira letra em maiúsculo), seguida de seu número e respectivo alelo, representados dentro dos círculos brancos. Os círculos pretos cheios são possíveis alelos que não foram encontrados nas amostras do estudo. A porção vertical centro-esquerda chama-se tronco e a porção horizontal direita chama-se ramificação. O alelo indicado como Consensus sugere ser aquele a partir do qual grande parte dos alelos divergiu.

Diferentemente da região tronco, na porção direita da Figura 1 se encontra uma série de seqüências que formam uma ramificação, cujos alelos são, em termos evolutivos, mais recentes. O fluxo gênico entre estas duas regiões é menor do que entre os alelos de cada um dos grupos, fazendo com que as novas mutações fiquem confinadas a seus respectivos grupos.

Quando se inclui o fator geográfico, verifica-se que todas as amostras do México, Colômbia, Equador e Norte do Peru (faseolinas tipo M, S, B, CH e I) se encontram exclusivamente na região tronco, enquanto que na região ramificada estão presentes apenas amostras do Sul do Peru e da Argentina (tipos C, H, T, P e J). Este fato sugere que existia um fluxo gênico entre as populações de feijão, das regiões do México até o Norte da América do Sul (Colômbia e Equador), pois são muito mais relacionados e formam um grupo à parte daquele outro formado pelos alelos encontrados ao Sul do Peru. O fator cultural-humano deve ter tido forte influência neste aspecto, pois pelas migrações humanas e suas relações inter-étnicas, espécies e variedades que a princípio estariam isoladas por grandes distâncias puderam ficar lado a lado, aumentando a chance de fluxo gênico. Grupamentos humanos mais isolados acabaram por propiciar o surgimento de plantas com características particulares, compartilhadas por um menor número de populações. Assim, a presença de dois grandes grupos de populações de alelos de feijão reflete, provavelmente, a existência de pelo menos dois grandes grupos culturais humanos, que mantinham um maior contato intragrupos do que intergrupos, como já evidenciado em outros trabalhos (Freitas et al., 2003).

Em vista da análise dos dados, propõe-se que a origem do feijão deva ser circunscrita a um único evento, pois pelo gráfico, as amostras do grupo Sul (ramificado) derivam das do Norte (tronco), e mostram um efeito evolutivo de deriva gênica, com um centro de origem e múltiplos centros de diversidade, contrariamente as hipóteses de múltiplos centros de origem.

Em relação à análise da seqüência da faseolina tipo I, Kami et al. (1995), em estudos com isoenzimas, relatam que o tipo I encontrado no Equador apresentou um padrão intermediário entre todos os outros tipos de faseolina encontrados no grupo México, Colômbia e dos tipos restantes da Região Andina. Esses autores sugerem que, ou este tipo I é muito antigo, o qual originou todos os outros tipos de faseolina, ou este padrão intermediário deve-se à mistura entre tipos destas duas regiões. Os dados do presente estudo reforçam a segunda hipótese.

Na amostra arqueológica, foram observados seis haplótipos distintos (note-se que se trata de uma família de multigenes e, por isto, pode haver mais de dois alelos por amostra). Dois dos alelos da amostra arqueológica são idênticos aos encontrados em amostras do grupo da Região Norte (alelos A-10 e A-13). Os outros quatro alelos de Januária são exclusivos. Mesmo assim, pode-se notar que eles ocorrem ao redor da região tronco, como é o caso da amostra A-12, que possui apenas uma mutação divergente, portanto também mais ligado aos tipos encontrados na Região Norte.

A ausência de alelos da amostra de Januária na região ramificada do gráfico e a ocorrência de alelos comuns ou muito próximos de tipos da região geográfica Norte sugerem que a amostra arqueológica de Januária tenha origem em populações do extremo Norte da América do Sul e/ou da América Central, com pouca influência da Região dos Andes Centrais e Meridional. Isto contribui para a reconstituição de rotas migratórias e de contato de populações humanas, visto que as espécies e variedades de plantas cultivadas foram difundidas pelo homem pré-histórico.

Agradecimentos

Ao Dr. André Prous e sua equipe, do Museu da Universidade Federal de Minas Gerais, pelo empréstimo da amostra arqueológica; ao Dr. Robin Allaby e Terrence Brown, pela ajuda laboratorial e na discussão dos dados; ao Dr. Debouck, pelas amostras modernas de feijão do CIAT.

Referências

- ALLABY, R.G.; BROWN, T.A. Network analysis provides insights into the evolution of 5S rDNA arrays in *Triticum* and *Aegilops*. **Genetics**, v.157, p.1331-1341, 2001.
- ALLABY, R.G.; O'DONOGHUE, K.; SALLARES, R.; JONES, M.K.; BROWN, T.A. Evidence for the survival of ancient DNA in charred wheat seeds from European archaeological sites. **Ancient Biomolecules**, v.1, p.119-129, 1997.
- DEBOUCK, D.G. Primary diversification of *Phaseolus* in the Americas: three centers? **Plant Genetic Resources Newsletter**, v.67, p.2-8, 1986.
- DEBOUCK, D.G. Systematics and morphology. In: SCHOONHOVEN, A. van; VOYSEST, O. (Ed.). **Common beans: research for crop improvement**. Cali: CIAT, 1991. p.55-118.
- FREITAS, F.O.; BENDEL, G.; ALLABY, R.G.; BROWN, T.A. DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into

- South America. **Journal of Archaeological Science**, v.30, p.901-908, 2003.
- GEPTS, P.; DEBOUCK, D.G. Origin, domestication, and evolution of the common bean (*Phaseolus vulgaris*). In: SCHOONHOVEN, A. van; VOYSEST, O. (Ed.). **Common beans: research for crop improvement**. Cali: CIAT, 1991. p.7-53.
- GEPTS, P.; OSBORN, T.C.; RASHKA, K.; BLISS, F.A. Phaseolin-protein variability in wild forms and landraces of the common bean (*Phaseolus vulgaris*): evidence for multiple centers of domestication. **Economic Botany**, v.40, p.451-468, 1986.
- KAMI, J.A.; GEPTS, P. Phaseolin nucleotide sequence diversity in *Phaseolus*. I. Intraspecific diversity in *Phaseolus vulgaris*. **Genome**, v.37, p.751-757, 1994.
- KAMI, J.A.; VELASQUEZ, V.B.; DEBOUCK, D.G.; GEPTS, P. Identification of presumed ancestral DNA sequences of phaseolin in *Phaseolus vulgaris*. **Proceedings of National Academy of Sciences**, v.92, p.1101-1104, 1995.
- NATIONAL CENTER FOR BIOTECHNOLOGY INFORMATION. **Entrez nucleotide**. Disponível em: <<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=nucleotide&val=403593>>. Acesso em: 20 mar. 1999.
- PROUS, A.; JUNQUEIRA, P.A.; MALTA, I.M. Arqueologia do alto médio São Francisco. Região de Januária e Montalvânia. **Revista de Arqueologia**, v.2, p.59-72, 1984.
- THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v.22, p.4673-4680, 1994.
- VASCONCELOS, D. **História antiga de Minas Gerais**. Belo Horizonte: Itatiaia, 1974. 283p.

Recebido em 18 de agosto de 2005 e aprovado em 25 de abril de 2006