SEED STORAGE PROTEINS IN COFFEE

SANDRA M. T. BAÚ[1], PAULO MAZZAFERA[2] AND LUIZ G. SANTORO[3]

Departamento de Fisiologia Vegetal, Instituto de Biologia, Caixa Postal 6109, Universidade Estadual de Campinas,13.083-970, Campinas, SP, Brazil.

ABSTRACT - It has been reported that *Coffea arabica* seeds contain as the main reserve protein, a legumin-like protein, constituted of two subunits, α and β, of approximately 35 and 20 kDa. In this work the seed proteins of several coffee species and varieties were investigated by SDS-PAGE and gel filtration. No differences were observed in the electrophoretic profiles among varieties of *C. arabica*, however, marked differences were observed among species, or even among individuals of some species. In general, the molecular weight of the subunits α and β accounted for a monomer of 48 to 62 kDa. However, native molecular weight obtained by gel filtration showed that for most of the species there is association of 6 of such proteins, in a hexamer. The most marked difference was observed for *C. canephora* and *C. racemosa*. The former clearly showing isoforms of the subunits, and the later showing absence of the β subunit. The influence of proteases in this observations is discussed.

ADDITIONAL INDEX TERMS: *Coffea*, legumin, endosperm

PROTEÍNAS DE RESERVA DE SEMENTES DE CAFÉ

RESUMO – Sementes de *Coffea arabica* possuem como principal proteína de reserva uma legumina, constituída por duas subunidades, α e β, com aproximadamente 35 e 20 kDa. No presente trabalho, proteínas das sementes de várias espécies e variedades de café foram estudadas, utilizando-se SDS-PAGE e filtração em coluna. Não se observaram diferenças entre os perfis eletroforéticos de variedades de *C. arabica*; porém, diferenças foram observadas entre espécies ou mesmo entre indivíduos de uma mesma espécie. De modo geral, nos vários materiais estudados, o peso molecular das subunidades α e β indicaram monômeros variando de 48 a 62 kDa. Entretanto, o peso molecular obtido por filtração em coluna mostrou haver na maioria dos casos a associação de 6 monômeros, formando um hexâmero. A diferença mais marcante foi observada em sementes de *C. canephora* e *C. racemosa*, em que a primeira apresentou isoformas das subunidades, e a última, ausência da subunidade β. A possível influência de proteases nessas variações é discutida.

TERMOS ADICIONAIS PARA INDEXAÇÃO: *Coffea*, legumina, endosperma

---

Received: 20/10/2000 – Accepted: 22/3/2001

1. Bióloga, Departamento de Fisiologia Vegetal, IB,UNICAMP.
2. Professor Adjunto, Doutor, Departamento de Fisiologia Vegetal, IB, UNICAMP – Autor para correspondência (e-mail: pmazza@obelix.unicamp.br)
3. Professor Assistente, Doutor., Departamento de Fisiologia Vegetal, IB, UNICAMP

INTRODUCTION

The average protein content of coffee beans is 10% (Clifford 1985) and despite several indications of the importance of these proteins in the beverage quality (Amorim *et al*., 1975, Arnold and Ludwig 1996, Melo and Amorim 1975) little is known about coffee seed proteins. Most studies have been carried out with *Coffee arabica*, the most cultivated coffee species, with the aim to establish a correlation between water soluble proteins and coffee quality (Amorim and Amorim 1977, Amorim and Josephson 1975, Bade and Stegemann 1982, Centi-Grossi *et al*., 1969). These reports showed that two bands were predominant in SDS-PAGE electrophoresis. Luthe (1992) suggested that they were the α and β subunits of a legumin-like protein.

Recently, Acuña *et al*. (1999) confirmed that seeds of *C. arabica* variety Colombia have a 11S globulin with an apparent molecular weight (M_r) of 55 kDa, consisting of two subunits of 33 and 24 kDa. Primers designed from the amino acids sequence were used to generate two full length cDNAs, which showed that the coffee legumin had a characteristic putative legumin box. In addition, similar banding pattern was observed when a cDNA probe was used to hybridize with digested DNA from four other coffee species.

A more detailed biochemical characterization of the coffee seed legumin was carried out by Rogers *et al*. (1999) using 2D-electrophoresis and amino acid microsequencing. They used *C. arabica* variety Caturra. Several abundant polypeptides were sequenced and identified as the α and β components of a 11S protein, suggesting that the protein was present as different isoforms. The alignment of these polypeptides with the sequences obtained by Acuña *et al*. (1999) showed 23 amino acid sequence differences, supporting the presence of isoforms in *C. arabica*. Preliminary electrophoretic results also indicated a secondary family of 11S proteins in Robusta, a variety of *Coffea canephora*. The similarity with other 11S proteins led Rogers *et al*. (1999) to suggest that in coffee this protein has a storage function. Indeed, there is a decrease of the α and β subunits during germination of seeds of *C. arabica* (Shimizu and Mazzafera, 2000).

In this paper the 11S storage proteins of several coffee species and varieties were investigated.

MATERIAL AND METHODS

Ripe fruits were collected from coffee trees growing at the Experimental Station of Agronomic Institute, at Campinas, Brazil. The fruits were manually depulped and dried at room temperature to 10% water content. The parchment was removed and the seeds (50 g) were finely ground in a knife mill and in a ball mill. The resulting powder was defatted in a Soxhlet with hexane during 12 h and used for protein extraction. Proteins were extracted 20 min at 4°C from 0.5 g of seed material in a mortar with pestle using 10 ml of 100 mmol.L^{-1} Na-BO_3 buffer, pH 8, 50 mmol.L^{-1} sodium diethyldithiocarbamate (DIECA), 50 mmol.L^{-1} ethylenediaminetetraacetic acid (EDTA), 300 mmol.L^{-1} NaCl, 114 mmol.L^{-1} ascorbate, 0.5 g polyvinylpolypyrrolidone (PVPP). The supernatant recovered by centrifugation (39,200 x *g*, 20 min) was combined with the supernatant from a second extraction (10 ml buffer) of the pellet, and the protein concentration determined (Bradford 1976). The extract was saturated to 80% with $(NH_4)_2SO_4$ and the pellet recovered by centrifugation was dissolved in a small volume of distilled H_2O and dialyzed against distilled H_2O for 24 h at 4°C. Proteins in the dialyzed extract were lyophilized and stored in freezer -20°C.

The apparent molecular weight (M_r) of the proteins were determined on a Sephacryl S-300 column (83.2 x 2.6 cm, Pharmacia). Lyophilized proteins were dissolved (4 mg.mL^{-1}) in 100 mmol.L^{-1} Na-Pi buffer, pH 7, and applied in to the column. 50 mmol.L^{-1} Tris-HCl, pH 8, 50 mmol.L^{-1} KCl, 10% glycerol was used as elution buffer at a flow rate of 0.8 mL.min^{-1}. Fractions were collected every 6 min and analyzed for protein. Aldolase (158 kDa), catalase (232 kDa), ferritin (440 kDa)

and thyroglobulin (669 kDa) (Pharmacia) were used to calibrate the column.

For SDS-PAGE analysis, lyophilized proteins were dissolved (4 $g.L^{-1}$) in 100 $mmol.L^{-1}$ Na-P buffer, pH 7, 4% SDS and 2% β-mercaptoethanol, denatured at 95°C for 10 min and separated in 17% polyacrylamide discontinuous gel (Laemmli 1970). Proteins were stained with Coomassie Blue G-250 and the band intensities were analysed by densitometry.

RESULTS AND DISCUSSION

SDS-PAGE of seed proteins of several *C. arabica* cultivars showed no marked difference in the banding profile (Figure 1). Catuaí (lane 3) and Caturra (lane 6) are dwarf varieties, Maragogipe (lane 2) has large beans, Cera (lane 5) has yellow seeds, and Mundo Novo (lane 4) is a mutation of the variety Bourbon, which descends from the Typica variety (lane 1). The latter variety has been used as a reference in coffee genetic analysis because its botanical characteristics closely resemble the description of the first plants introduced in Europe from Ethiopian forests. It was the first variety introduced in South America, originating from seeds of an individual *C. arabica* tree growing at the Amsterdam Botanic Garden. Descriptions of the origin and genetic background of these varieties are presented by Carvalho and collaborators (Carvalho 1958, 1993, Carvalho *et al*. 1952, 1985, 1991a, 1991b). Based on 5 electrophoretic profiles of Typica extracts, the apparent M_r of the two main bands were 37.6 ± 3.3 and 20.9 ± 1.4 kDa.

For each coffee variety or species, seeds were collected from three plants and differences in the protein profile were observed among individuals of *C. eugenioides* and *C. canephora* variety Kouillou (also known in Brazil as Conillon). *C. eugenioides* plant 2 (Figure 2 lane 16) showed a clear distinction of the α subunit as three bands, while plants 1 and 3 (lanes 2 and 15) showed more intense staining of the heavier band of these three bands. Protein profiles of the plants 70/9 (lane 7) and 70/14 (lane 17) of *C. canephora* variety Kouillou were different of plant 70/10 (lane 18) regarding the α subunit. Plant 70/10 showed a protein profile similar to variety Typica (lanes 1, 5, 8 and 19). Two bands in the position of the β subunit were detected in Kouillou 70/9. As suggested by Rogers *et al*. (1999), these bands may represent isoforms of the coffee legumin. However, proof that hese bands are subunits of coffee legumin can only be assured by purification and sequencing.

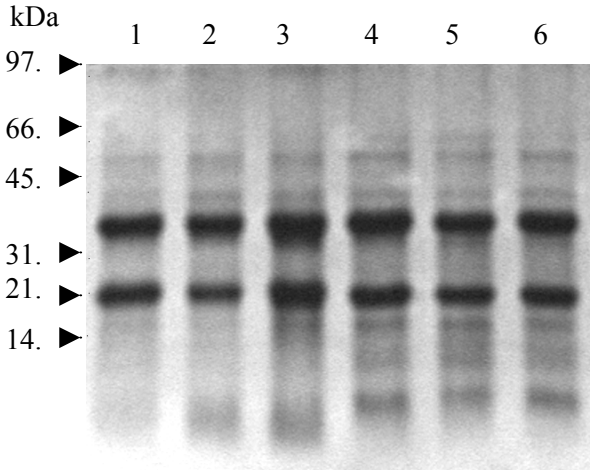

FIGURE 1 - SDS-PAGE of seed proteins of *C. arabica* varieties: lane 1 – Typica; lane 2 – Maragogipe; lane 3 – Catuaí; lane 4 – Mundo Novo; lane 5 – Cera; lane 6 - Caturra. M_r markers: phosphorylase b (97.4 kDa), BSA (66.2 kDa), eggalbumin (45 kDa), carbonic anhidrase (31 kDa), trypsin inhibitor (21.5 kDa), lysozyme (14.4 kDa).

Significant variation in the protein profile was also observed among different coffee species (Figure 2). *C. bukobensis* (lane 4) are genetically related to the *C. canephora* variety Kouillou. As observed for Kouillou 70/9, this species showed two isoforms of the subunit α. Similar banding patterns were also observed for *C. liberica* (lane 9), *C. dewevrei* (lane 10) and *C. klainii* (lane 13) which are also closely related (Clifford *et al*., 1989).

C. salvatrix also presented a α subunit apparently constituted of three bands (lane 12). Compared to the other species, *C. stenophylla* presented a heavier α subunit and a lighter β subunit

(lane 11). *C.racemosa* showed the most marked banding pattern, since the β subunit was not observed (Figure 2, lane 3), although a strong band was detected running at the electrophoresis front. *C. kapakata* (lane 6) showed similar banding to *C. arabica*.

The soluble protein content of coffee seeds varied from 4.7% in *C. racemosa* to 9.1% in *C. canephora* variety Kouillou (Table 1). We are unaware of any report in the literature on protein content covering several coffee species as presented here. For those species reported in the literature, total protein content is obtained by multiplying the total N content, obtained by the Kjeldahl method, by 6.25 (Clifford 1985). *C. arabica* and *C. canephora* variety Robusta in general have a total protein content in the range 10-14%.

Native M_r of coffee proteins was obtained by gel filtration on a Sephacryl S300 column. For most of the studied species a main protein peak was observed in the elution profiles and SDS - PAGE of fractions containing the highest protein concentration revealed similar banding with those observed in Figures 1 and 2.

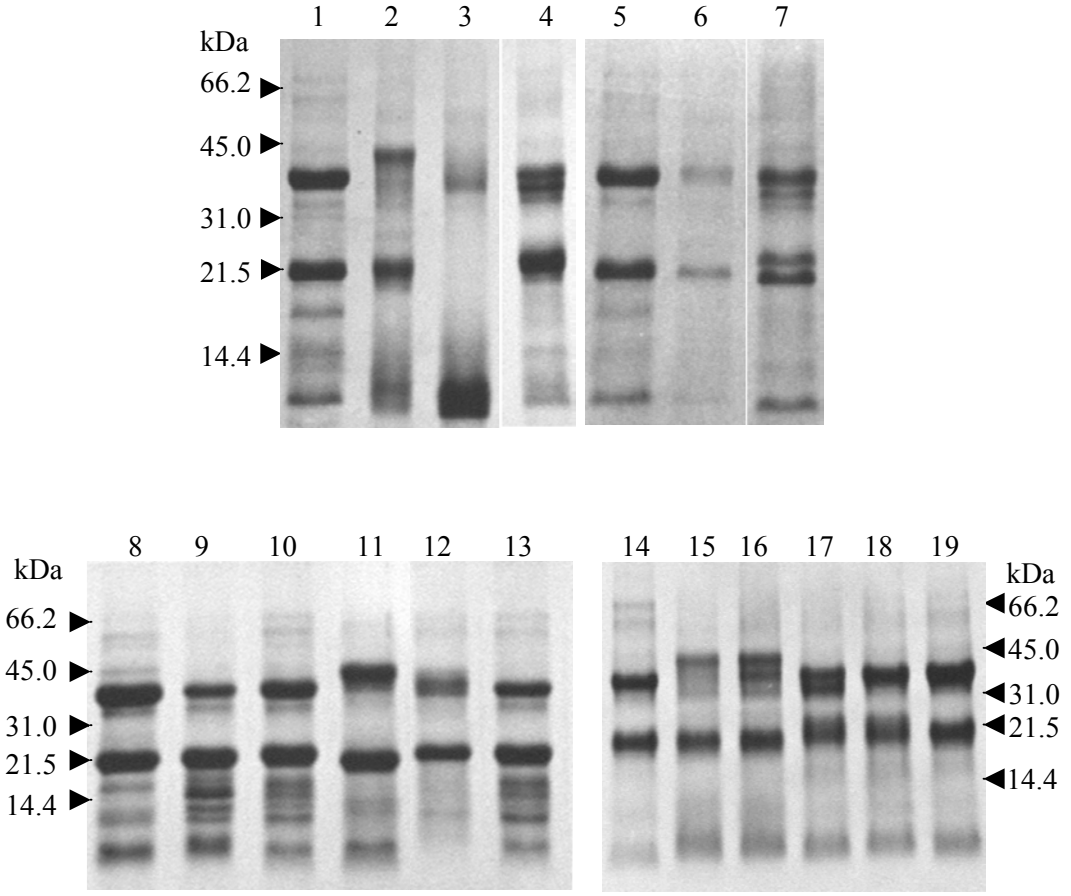

FIGURE 2 - SDS-PAGE of seed protein of coffee species: *C. arabica* var. Typica (lanes 1, 5, 8, 19); *C. eugenioides* (plant 1 – lane 2, plant 2 – lane 16, plant 3 – lane 15); *C. racemosa* (lane 3); *C. bukobensis* (lane 4); *C. kapakata* (lane 6); *C. canephora* var. Kouillou (70/9 – lane 7, 70/10 – lane 18, 70/14 – lane 17); *C. liberica* (lane 9); *C. dewevrei* (lane 10); *C. stenophylla* (lane 11); *C. salvatrix* (lane 12); *C. klainii* (lane 13); *C. laurenti* (lane 14). M_r markers: see Figure 1 legend.

TABLE 1 - Protein content, apparent native M_r determined on a Sephacryl S300 column, ratio of the scanned areas of α and β subunits, and estimation of the number of monomers associated by complexes of 11S coffee proteins.

Coffee specie / variety	Protein (% dry wt)	M_r (kDa)	α / β ratio	Number of monomers
C. arabica var Typica	7.8 ± 0.4	189	1.08	3
C. canephora var Kouillou plant 70/9	7.7 ± 0.4	349	0.94	6
C. dewevrei	8.1 ± 0.5	471	0.93	9
C. stenophylla	4.9 ± 0.2	342	0.67	6
C. klainii	5.9 ± 0.3	470	0.61	9
C. liberica	9.1 ± 0.8	273	0.44	6
C. eugenioides plant 3	6.6 ± 0.2	220	0.54	4
C. kapakata	8.4 ± 0.7	339	0.83	5
C. laurentii	6.5 ± 0.7	252	0.71	5
C. bukobensis	6.9 ± 0.2	349	0.87	7
C. racemosa	4.7 ± 0.4	87	---	2

A typical profile is shown in Figure 3A, obtained for the *C. arabica* variety Typica. An exception was *C. racemosa* (Figure 3B) which showed several small peaks before a main peak corresponding to an apparent mass of 87 kDa. This was observed for the three *C. racemosa* plants. Although the seeds were not stored more than one week between harvesting and defating in a Soxhlet, which would inactivate proteases, it is possible that proteolysis occurred. Ludwig *et al*. (2000) detected five to seven proteases in seeds of *C. arabica* and *C. canephora*. The former species showed two more active bands while in *C. canephora* there was predominance of one band of activity. In *C. arabica* the same bands of activity were found in protein extracts from freshly collected samples as well as in stored samples. In the present study, although the possibility that proteases have degraded the coffee legumin, it is curious that such intense degradation was observed only with the β subunit and only happened with *C. racemosa*. In addition, the fraction from the filtration column corresponding to the 87 kDa peak gave a similar profile in SDS-PAGE. Rogers *et al*. (1999) also discussed the presence of small fragments (12 and 16 kDa) in 2D electrophoresis, which were identified as fragments of the 11S coffee protein, as possible products of the cleavage by specific proteases of α subunit.

The SDS-PAGE gels were scanned and the ratio between the areas corresponding to those bands that were considered as α and β subunits were calculated. For those species with more than one band per subunit, the sum was used. In general, the values were close to 1, suggesting a 1:1 proportion between the subunits, although for some species, such as *C. liberica* and *C. eugenioides* plant 3, these values were lower (Table 1). In these cases we should to consider that degradation of subunits by specific proteases may have changed the staining intensity of some bands.

In general, the M_r of the α and β subunits in the species quantified for a monomer varying from 48 to 62 kDa. Dividing the native M_r from the Sephacryl column by the M_r of the monomers (for those with more than one band per subunit, a mean

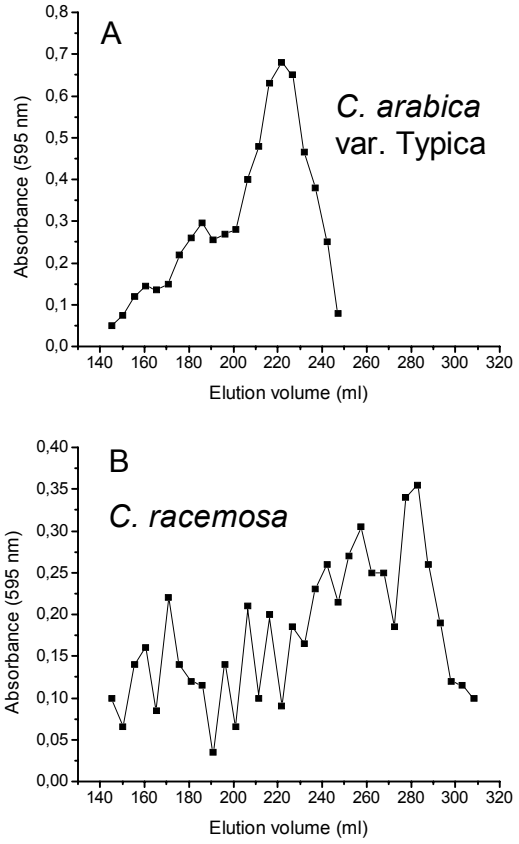

FIGURE 3 - Elution profile of seed protein extracts of *C. arabica* variety Typica (A) and *C. racemosa* (B) from a Sephacryl S300 column.

value was calculated), it was estimated that the coffee legumin in the different coffee species would be present as an hexamer (Table 1). Exceptions were *C. klainii* and *C. dewevrei*, and *C. racemosa*, with 9, 9 and 2 protein copies. However, it is interesting that non-denaturing electrophoretic gels of *C. arabica* and/or *C. canephora* variety Robusta extracts showed apparent M_r of approximately 55-65 kDa (Amorim and Josephson 1975, Acuña *et al*., 1999, Bade and Stegemann 1982, Rogers *et al*. 1999). Although we have not run native PAGE, it seems that for most of the species studied here the hexamer detected in the Sephacryl column was due to a folding of six protein units. The variations observed for *C. klainii* and *C. dewevrei* might be a different number of folded subunits.

Amorim and Josephson (1975) detected under nonreducing electrophoresis that most of the coffee proteins had M_r 150 and 170 kDa, as well as proteins of 66 kDa were also stained. Rogers *et al*. (1999) observed in 2D electophoresis of coffee proteins that a 80 kDa polypeptide was a tetramer of the β subunit. A good example of how proteins can fold differently is the work of Yamada *et al*. (1981) with the legumin arachin in peanut. Trimers and hexamers were observed by folding of monomers (constituted by α and β subunits linked by disulphide bonds) due to formation of non covalent bonds.

Our data show that legumin like proteins are the main storage protein in several coffee species. The marked differences observed among species and even in plants of the same species, indicating therefore the presence of isoforms, suggest that they are present in *Coffea* as a gene family. Rogers *et al*. (1999) came to the same conclusion based on the detection of multiple pI isoforms in protein extracts from *C. arabica* and *C. canephora* proteins. These authors also cloned a full-length cDNA for one of the isoforms, and in contradiction to evidence from the protein electrophoresis, they could not detect multiples copies of the gene by Southern blotting experiments. On the other hand, although Acuña *et al*. (1999) have shown strong hybridization of a cDNA probe with a single fragment of genomic DNA from *C. arabica*, they also detected other fragments when hybridization was carried out at low stringency.

ACKNOWLEDGEMENTS

S.T.B. and P.M. thanks CNPq-Brazil for M.Sc. and research fellowships, respectively. The authors wish to thank Prof. Ladaslav Sodek for helpful comments and for correcting the English version of the manuscript.

REFERENCES

ACUÑA, R.; BASSÜNER, R.; BEILINSON, V.; CORTINA, H.; CADENA-GÓMEZ, G.; MONTES, V. & NIELSEN, N.C. Coffee seeds contain 11S storage proteins. **Physiologia Plantarum**, 105: 122-131, 1999.

AMORIM, H.V. & AMORIM, V.L. Coffee enzymes and coffee quality. In. Ory, R.L. & St Angelo, A.J. (Eds.) **Enzymes in Food and Beverages Processing**, ACS Synposium Series, 47., American Chemical Society, 1977. 27-56.

AMORIM, H.V. & JOSEPHSON, R.V. Water-soluble protein- and non-protein components of Brazilian green coffee beans**. Journal of Food Science**, 40: 1179-1185, 1975.

AMORIM, H.V.; TEIXEIRA, A.A.; MELO, M.; CRUZ, V.F. & MALAVOLTA, E. Chemistry of Brazilian green coffee and the quality of the beverage. 4. electrophoresis of soluble proteins in agar-gel and its interaction with chlorogenic acids. **Turrialba**, 25: 18-25, 1975.

ARNOLD, U. & LUDWIG, E. Analysis of free amino acids in green coffee beans II. changes of the amino acid content in arabica coffees in connection with post-harvest model treatment. **Zeitschrift fur Lebensmittel-Untersuchung und-Forschung**, 203: 379-384, 1996.

BADE, H. & STEGEMANN, H. Protein patterns of coffee beans. Characterization by one- and two-dimensional electrophoresis. **Journal of Agronomy and Crop Science**, 51: 89-98, 1982.

BRADFORD, M.N. A rapide and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. **Analytical Biochemistry**, 72: 248-254, 1976.

CARVALHO, A. Advances in coffee production technology. Recent Advances in our knowledge of coffee trees. 2. Genetics. **Coffea Tea Industry Flavor Field**, 81: 30-36, 1958.

CARVALHO, A. Histórico do desenvolvimento do cultivo do café no Brasil. **Documentos IAC** 34: 1-7, 1993.

CARVALHO, A.; KRUG, C.A.; MENDES, J.E.T.; ANTUNES FILHO, H.; MORAES, H.; ALOISI SOBRINHO, J.; MORAES, M.V. & ROCHA, T.R. Melhoramento do cafeeiro. IV. Café Mundo Novo. **Bragantia,** 12: 97-129, 1952.

CARVALHO, A.; MEDINA FILHO, H.P.; FAZUOLI, L.C.; GUERREIRO FILHO, O. & LIMA, M.M.A. Aspectos genéticos do cafeeiro. **Revista Brasileira de Genética**, 14: 135-183, 1991a.

CARVALHO, A.; MEDINA FILHO, H.P.; GUERREIRO FILHO, O. & FAZUOLI, L.C. Somatic mutations in *Coffea arabica* L. **Ciência Cultura**, 43: 32-36, 1991b.

CENTI-GROSSI, M.; TASSI-MICCO, C. & SILANO, V. Albumin fractionation of green coffee seed varieties by acrilamide gel-electrophoresis. **Phytochemistry** 8: 1749-1751, 1969.

CLIFFORD, M.N. Chemical and physical aspects of green coffee and coffee products. In. Clifford, M.N. & Wilson, K.C. (Eds.) **Coffee: Botany, Biochemistry and Production of Beans and Beverage**, Westport, AVI Publishing Company, Inc., 1985. 305-374

CLIFFORD, M.N.; WILLIAMS, T. & BRIDSON, D. Chlorogenic acids and caffeine as possible taxonomic criteria in *Coffea* and *Psilanthus*. **Phytochemistry**, 28: 829-838, 1989.

LAEMMLI, U.K. Cleavage of structural protein during the assembly of head of bacteriophage T_4. **Nature,** 227: 680-684, 1970.

LUDWIG, E.; LIPKE, U.; RACZEK, U. & JÄGER, A. Investigations of peptides and proteases in green coffee beans. **European Food Research Technology**, 211: 111-116, 2000.

LUTHE, D.S. Electrophoretic analysis of seed proteins in the Dycotiledoneae. **Plant Molecular Biology Reporter,** 10: 254-262, 1992.

MELO, M. & AMORIM, H.V. Chemistry of Brazilian green coffee and the quality of the beverage. VI. U.V. and visible spectral analysis and chlorogenic acids content on TCA soluble buffer extracts. **Turrialba**, 25: 243-248, 1975.

ROGERS, W.J.; BÉZARD, G.; DESHAYES, A.; MEYER, I.; PÉTIARD, V. & MARRACCINI, P. Biochemical and molecular characterization and expression of the 11S-type storage protein from *Coffea arabica* endosperm. **Plant Physiology and Biochemistry**, 37:261-272, 1999.

SHIMIZU, M.M. & MAZZAFERA, P. Compositional changes of proteins and amino acids in germinating coffee seeds. **Brazilian Archives of Biology and Technology**, 43: 259-265, 2000.

YAMADA, T.; AIBARA, S. & MORITA, Y. Reconstitution of arachin from isolated subunits. **Agricultural and Biological Chemistry** 45: 1243-1250, 1981.