Characterization of a partial sequence of the P-subunit of the glycine decarboxylase gene from four coffea genotypes

Ana M. García[1], Yamel López[2], Néstor M. Riaño[3,*]

[1]Clínica Cardiovascular Santa María, calle 78b No. 75-21 Medellín, Colombia.
[2]Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia - Palmira, Colombia.
[3]Fisiología Vegetal, Centro Nacional de Investigaciones de Café- CENICAFE, Chinchiná (Caldas) Colombia.
*Author for correspondence; fax: 57-68-504723, email: nestorm.riano@cafedecolombia.com

ABSTRACT

A 427 bp fragment of DNA isolated from leaf tissue of four coffee genotypes (*Coffea arabica* cv. Caturra, *Coffea canephora*, *Coffea congensis* and *Coffea eugenioides*) was amplified by PCR, using heterologous primers synthesized following the nucleotide sequence encoding a part of P-subunit of glycine decarboxylase from pea *Pisum sativum*. Amplified fragments were cloned and their nucleotide sequence was determined. Sequence analysis showed 79% homology between the gene sequence of the P-subunit of glycine decarboxylase from *P. sativum* and the four *Coffea* genotypes. The amino acid sequence of cloned fragments showed 81% homology between *P. sativum* and the four *Coffea* genotypes, whereas between *Flaveria pringlei* and *Coffea* such homology was 85%. Comparison of nucleotide sequences for the P-subunit gene from the four Coffea genotypes showed 98-100% and 99-100% homology at amino acid level, except for a substitution of glutamine by glutamic acid (position 59) in *C. canephora* and *C. congensis*.

Keywords: *Coffea* genotypes, glycine decarboxylase P-subunit, nucleotide sequence, aminoacid substitution

Caracterización de una secuencia parcial de la subunidad P de la glicina decarboxilasa de cuatro genotipos de café

COMPENDIO

Se amplificó un fragmento de 427 pb de ADN aislado a partir de tejido foliar de cuatro genotipos de café (Coffea arabica cv Caturra, Coffea canephora, Coffea congensis y Coffea eugeniodes) mediante PCR y usando cebadores heterólogos sintetizados con base en la secuencia de nucleótidos que codifica para una parte de la subunidad P de la glicina decarboxilasa de arveja Pisum sativum. Se clonaron los fragmentos amplificados y se les determinó la secuencia de nucleótidos. El análisis de la secuencia de nucleótidos mostró una homología del 79% entre la secuencia de la subunidad P de la glicina decarboxilasa de P. sativum y los cuatro genotipos de Coffea. La secuencia de aminoácidos codificada en los fragmentos clonados mostraron un 81% de homología entre P. sativum y los cuatro genotipos de Coffea, mientras que entre Flaveria pringlei y Coffea la homología fue del 85%. La comparación de la secuencia de nucleótidos del gene de la subunidad P de los cuatro genotipos de Coffea mostró una homología del 98 y 99% a nivel de los aminoácidos, con la excepción de una glutamina por un ácido glutámico en la posición 59 tanto en C. canephora como en C. congensis.

Palabras claves: *genotipos de café, subunidad - glicina decarboxilasa, secuencia nucleóticos, substitución aminoácidos*

Introduction

A limiting factor of photosynthesis is photorespiration, a process by which the plant decarboxylates photosynthetic products in presence of O_2, releasing CO_2 and producing serine from glycine, with a high waste of energy and without any apparent benefit for the plant. In this process, a great number of enzymes distributed in chloroplasts, mitochondria, and peroxisomes are involved. Most of such enzymes are encoded by nuclear DNA and are carried to the corresponding organelles after translation, as in the case of glycine decarboxylase, which is in the mitochondria and is responsible for CO_2 release by glycine decarboxylation. This enzyme consists of four soluble protein subunits located in the mitochondrial matrix: P-subunit, which decarboxylates glycine and catalyzes CO_2 exchange; H-subunit, which acts as a methylamine carrier and contains lipoic acid as an acceptor for electrons from glycine; T-subunit, which transfers aminomethyl radical to tetrahydrofolic acid (THF) with NH_3 release and formation of $N^5 N^{10}$ methylene-tetrahydrofolate; and L-subunit, which catalyzes electron transfer from reduced lipoic acid of H-subunit to NAD^+, to form NADH (Canvin 1990, Oliver 1994, Whelan and Glaser, 1997).

The P-subunit is a pyridoxal 5 phosphate (PLP) dependent carboxylase, and is considered the true glycine decarboxylase enzyme. The mature protein from *P. sativum*, composed of 971 amino acids is encoded by a small multigene family, is a 105-kDa homodimer. It has a presequence of 86 amino acids, which is removed after entering the mitochondrial matrix. The homology in gene sequence and amino acid sequence of this subunit among different plant and animal species are 60-76% and 84% (Kume et al., 1991, Turner et al., 1992, Kopriva and Bauwe 1994). In coffee, there is no available information on this enzyme. Decker and Tio (1959) found photorespiration in coffee and its importance for crop production was established after an estimation of 50% loss of harvest potential through this pathway. Following this work Zelitch (1971) and several other workers have studied photorespiration and showed the detrimental effect on C_3 crop production. This study performed a partial analysis of the DNA sequence of the gene encoding a fragment of the P-subunit of glycine decarboxylase from different *Coffea* genotypes. Analysis of individual sequences and restriction patterns was included in order to find homologies and variations that could represent functional and/or evolutionary differences.

Materials and Methods

Plant Materials

Young leaves from one-year-old plants of *Coffea arabica* cv. Caturra, *Coffea canephora*, *Coffea congensis* and *Coffea eugenioides* plants grown outdoors from the collection of the Centro Nacional de Investigaciones de Café, CENICAFE (Chinchiná, Caldas, Colombia), were used.

DNA Purification by molecular exclusion chromatography.

The molecular exclusion chromatography technique in superfine Sephacryl S-1000 (Pharmacia, Biotech Inc.) was used, with elution buffer 0.01 M Tris-HCl pH 8.0, 0.2 M NaCl, 0.5 mM EDTA, and a flow-rate of 1.0 ml min^{-1} in order to eliminate phenolic metabolites and other compounds from the sample (García, *non published data*).

Polymerase Chain Reaction (PCR)

The reaction was performed using heterologous primers Gly1: AGACGCCATAACTCAGCAACAC and Gly2: CCGCAGCAGCAGTACCTTCATC, constructed following highly homologous regions from DNA sequences of the P-subunit of glycine decarboxylase from *P. sativum*. Primers were designed using the Laser Gene Program (Software System, DNASTAR Inc. Madison, WI, USA).

For the PCR, 0.75 U *Taq* DNA Polymerase and 1X of the enzyme buffer purchased from Amersham, 1mM of $MgCl_2$, 0.2 mM of dNTPs, 150 ng of Gly1 and Gly2 primers and 100 ng of total DNA were used. The reaction cycle was performed using a Perkin Elmer DNA Thermal Cycler 480, with one cycle of initial denaturation (5 min at 95°C), 35 cycles of 30 s at 94°C, followed by 30 s at 60°C (annealing), 1 min at 72°C (extension) and a final incubation cycle of 7 min at 72°C (Ausubel et al., 1994).

Cloning of Amplification Products

Invitrogen A-T Cloning Kit and Amersham pMOS*Blue* Blunt Ended Cloning Kit were used, which use pCRII

and pMOS*Blue* as vectors, respectively. *E. coli* strain DH5a (Bethesda Res. Lab.) was transformed using these vectors and transformants were selected using IPTG and X-Gal in solid LB medium with ampicillin (100 µg ml^{-1}) as described (Sambrook et al 1989). Plasmids were extracted by Mini-prep technique using protocols described (Sambrook et al., 1989). Presence of the inserts was determined by digestion with *Eco*RI for pCRII and *Eco*RI-*Hind*III for pMOS*Blue*. The restriction products were evaluated by electrophoresis in 1.5 % agarose gel, stained with ethidium bromide (Sambrook et al., 1989; Ausubel et al., 1994).

Sequencing of Cloned Products

The primers for pMOS*Blue* inserts were forward T7 and reverse M13, and for pCRII, T7 and U19 (Invitrogen). DNA samples for sequencing were isolated from culture of an individual colony of each genotype in 2XYT medium and DNA was isolated with Promega Wizard Plus Minipreps. Manual sequencing was performed using Amersham ThermoSequenase kit, and cycle sequencing with [a-^{35}S] dATP (10µCi mL^{-1}) as a label in a Perkin Elmer 940 thermocycler. Automated sequencing was done in the Applied Biosystems 373 DNA Sequencer using the Perkin Elmer ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction Kit. Sequences were analyzed using Generunner (DNASTAR Inc., Madison, WI), Clustal W, version 1.6 and Blast (NCBI-Genbank) Programs.

Results and discussion

DNA Isolation

DNA (A_{260}/A_{280} = 1.9) from leaf tissue from different coffee genotypes was isolated using the technique by Mettler (1987) and Chaparro (1993), modified using molecular exclusion chromatography on Sephacryl S-1000 (Pharmacia). This modification allowed isolation of clean high molecular weight genomic DNA free from phenolics and RNA (García, *non published data*).

Polymerase Chain Reaction (PCR)

DNA from leaf tissue of *C. arabica* cv. Caturra, *C. canephora*, *C. congensis*, *C. eugenioides* and *P. sativum* was amplified. Electrophoresis of amplification products shows a band of approximately 450 bp in all genotypes (***Figure 1***). In *C. arabica* cv. Caturra there is another band corresponding to amplification products of approximately 700 bp, whereas in *C. canephora* and *C. congensis* these same two bands are present in addition to a band of approximately 1500 bp. This suggests that in all genotypes an apparently similar size fragment (approximately 450 bp) is amplified. Contrary to what was observed in *Coffea* genotypes, amplification of DNA from *P. sativum* only showed one band as expected, since the primers used were specific for pea. The additional bands from amplification products of coffee DNA could be eliminated by increasing annealing stringency in PCR. However, under these conditions, no products could be amplified from DNA of *C. arabica* cv. Caturra, *C. eugenioides* and *P. sativum*. Therefore, conditions were not modified and cloning was made with the amplification products obtained.

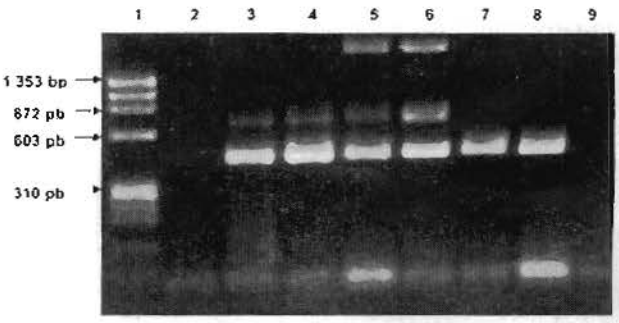

Figure 1. *Electrophoresis of DNA fragments isolated from all genotypes, amplified by PCR. Lane1: Molecular weight marker 1 cut with HindIII; lanes 2 and 9 negative controls (without DNA); lanes 3 and 4: Coffea arabica cv. Caturra; lane 5: Coffea canephora; lane 6: Coffea congensis, lane 7: Coffea eugenioides; lane 8: Pisum sativum. Electrophoresis was performed in a low melting point 1.5% agarose gel*

In order to determine validity of amplified fragments, restriction analysis for the amplification product of *P. sativum* with enzyme *Hae*III was performed. Two fragments were obtained (***Figure 2***), the first one of approximately 100 bp and the second one of 330 bp. Such values correspond to the reported sizes for this species (Turner 1992). This result suggests that the amplification product corresponds to the expected DNA fragment.

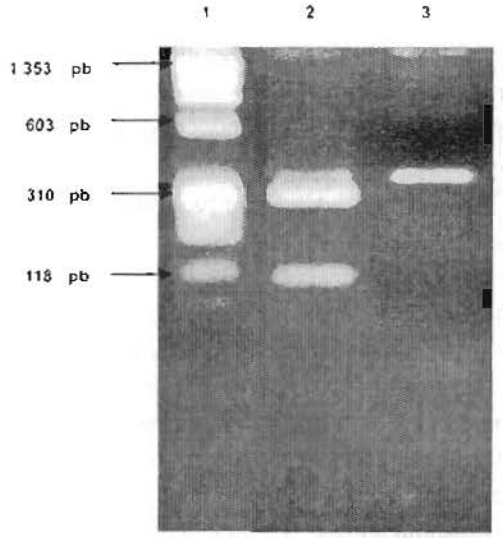

Figure 2. Electrophoresis of the DNA fragment of Pisum sativum treated with restriction enzyme HaeIII and compared to intact fragment (lane 1. ɸX174 HaeIII; lane 2. Fragment 437HaeIII; lane 3. Fragment 437). Electrophoresis conditions were as in Fig. 1.

Cloning

Transformants for *C. eugenioides* and *C. congensis* were obtained in vector pMOS*Blue*, and for *C. arabica* cv. Caturra and *C. canephora* in vector pCRII, all transformed into *E. coli* strain DH5a. Amplification of cloned products allowed obtaining probes of the gene for the P-subunit of glycine decarboxylase in the four mentioned species of *Coffea*.

Sequence Analysis

The nucleic acid sequences (***Figure 3***) for *C. arabica* cv. Caturra, *C. canephora*, *C. congensis* and *C. eugenioides* were translated to amino acids in the six possible reading frames using Generunner program. For all genotypes, the only open reading frame obtained was from reading frame 1, which was chosen for the subsequent analyses. Chemical nature of peptides of the glycine decarboxilase from *Coffea* genus was clearly polar and show 100% homology among all genotypes (***Table 1***). Individual amino acid analysis shows leucine as the most prevailing one, followed by serine, alanine, and threonine. There is a very low level of cysteine (0.7%) in coffee, which was consistent through all the sequences (***Table 2***).

Nucleotide analysis of these fragments shows a low GC content in all of them and there is no significant difference among genotypes (41.08% ± 0.015). Low GC content is characteristic of plant genomes, never above 50% (***Table 3***). In other plants, GC content is around 40%, with a minimum value of 38.7 for Kiwi and a maximum of 46.3% for *Chondrius crispus* (Marie et al., 1993). Results show that sequences are highly conserved in all studied genotypes.

The result of DNA restriction analysis of the different genotypes is shown in ***Table 4***. Two different restriction pattern groups were observed: group 1 corresponds to *C. arabica* cv. Caturra, *C. canephora*, *C. congensis*, and *C. eugenioides*, and group 2 corresponds to *P. sativum*. Both have unique cutting sites which give them their identity: (*Bcl*I, *Mob*II, *Nsi*I, *Nsp*I, and *Xho*II) for group 1, and (*Acc*II, *Eco*RV, *Hae*III, *Hap*II, *Msp*II, and *Taq*I) for group 2. Both groups share restriction sites for *Acc*I, *Hin*fI and *Rsa*I, showing conserved sequences located in the same sites. Restriction sites for *Acy*II, *Dde*I, *Dnp*I, and *Nla*III are also partially shared. An important difference in restriction sites for group 1 is indicated in ***Table 4***, corresponding to base 426. Restriction site for *Pst*I appears in *C. arabica* cv. Caturra, where a one base change has occurred.

In the nucleic acid sequence, high similarity among the four genotypes is shown, but when comparing in more detail two minor changes in the sequences may be observed. Between *C. arabica* cv. Caturra and the other three genotypes there is a transition in base 423, where C is changed for T (***Figure 4A***). There is a transversion in base 175, where C changes for G (***Figure 4B***). These changes indicate a subtle polymorphism in the sequences, but the high degree of homology confirms the fact that they are very conserved sequences.

Since sequences of all coffee genotypes were practically identical, *C. arabica* cv. Caturra was chosen to represent the other genotypes in the comparisons made with fragment sequences that have been published. Blast analysis of nucleic acids showed that the sequence of *C. arabica* L. cv. Caturra was 79% homologous with the *P. sativum* sequence (***Figure 5***). The pea sequence started at base 377 and ended at 803, the exact sites where the sequences for designing the primer were taken from, which corresponds to the same fragment of glycine decarboxylase described *P. sativum*.

With the aim of determining the degree of homology, comparisons between *C. arabica* cv. Caturra and *P. sativum* and *Flaveria pringlei* were also carried out.

A. *Coffea arabica* cv. Caturra

```
   1  agacgccata actcagcaac acctgaagaa caaatcaaaa tggttgagaa ttgtggattc
  61  cccagtcttg attcgcttat tgacgccact gtgcctaaat ctatcaggct tgatggtatg
 121  acgtttagta agtttgatga gggattaact gaggctcaaa tgattgatca catgcaaaag
 181  ttagcatcta agaacaaagt ttttaagtca tatattggaa tgggatacta taatacottt
 241  gttccgcctg ttatttttgag gaatctcctg gaaaatcctg cttggtatac tcagtatact
 301  cctatcagg ctgagatttc gcagggacgt cttgaatccc tgctgaatta tcagaccatg
 361  attgcggatc ttactggatt gcctatgtct aatgcatctt tactagatga aggtactgct
 421  gccgcgg
```

B. *Coffea canephora*

```
   1  agacgccata actcagcaac acctgaagaa caaatcaaaa tggttgagaa ttgtggattc
  61  cccagtcttg attcgcttat tgacgccact gtgcctaaat ctatcaggct tgatggtatg
 121  acgtttagta agtttgatga gggattaact gaggctcaaa tgattgatca catggaaaag
 181  ttagcatcta agaacaaagt ttttaagtca tatattggaa tgggatacta taatacottt
 241  gttccgcctg ttatttttgag gaatctcctg gaaaatcctg cttggtatac tcagtatact
 301  cctatcagg ctgagatttc gcagggacgt cttgaatccc tgctgaatta tcagaccatg
 361  attgcggatc ttactggatt gcctatgtct aatgcatctt tactagatga aggtactgct
 421  gctgcgg
```

C. *Coffea congensis*

```
   1  agacgcgata actcagcaac acctgaagaa caaatcaaaa tggttgagaa ttgtggattc
  61  cccagtcttg attcgcttat tgacgccact gtgcctaaat ctatcaggct tgatggtatg
 121  acgtttagta agtttgatga gggattaact gaggctcaaa tgattgatca catggaaaag
 181  ttagcatcta agaacaaagt ttttaagtca tatattggaa tgggatacta taatacottt
 241  gttccgcctg ttatttttgag gaatctcctg gaaaatcctg cttggtatac tcagtatact
 301  cctatcagg ctgagatttc gcagggacgt cttgaatccc tgctgaatta tcagaccatg
 361  attgcggatc ttactggatt gcctatgtct aatgcatctt tactagatga aggtactgct
 421  gctgcgg
```

D. *Coffea eugenioides*

```
   1  agacgccata actcagcaac acctgaagaa caaatcaaaa tggttgagaa ttgtggattc
  61  cccagtcttg attcgcttat tgacgccact gtgcctaaat ctatcaggct tgatggtatg
 121  acgtttagta agtttgatga gggattaact gaggctcaaa tgattgatca catgcaaaag
 181  ttagcatcta agaacaaagt ttttaagtca tatattggaa tgggatacta taatacottt
 241  gttccgcctg ttatttttgag gaatctcctg gaaaatcctg cttggtatac tcagtatact
 301  cctatcagg ctgagatttc gcagggacgt cttgaatccc tgctgaatta tcagaccatg
 361  attgcggatc ttactggatt gcctatgtct aatgcatctt tactagatga aggtactgct
 421  gctgcgg
```

Figure 3. *Sequence of PCR fragments of the four coffee genotypes. A: Coffea arabica cv. Caturra (GeneBank accession number AF042072), B: Coffea canephora (GeneBank accession number AF043097), C: Coffea congensis (GeneBank accession number AF043098), D: Coffea eugenioides (GeneBank accession number AF043099).*

Homology percentages are practically equal: 79% *C. arabica* cv. Caturra vs. *P. sativum*, 77% *C. arabica* cv. Caturra vs. *F. pringlei*, 78% *P. sativum* vs. *F. pringlei*, which indicates the level of association among them. Analysis of the amino acid sequence shows that *C. arabica* cv. Caturra fragment has 83% homology with the sequence of *P. sativum*, even higher than that found for nucleic acids, which confirms homology of the cloned sequence with that of glycine decarboxylase described for *P. sativum* (***Figure 6***).

A comparison of *C. arabica* cv. Caturra, *P. sativum*, and *F. pringlei* sequences shows that homologies of sequences are similar and high: 83% *C. arabica* cv. Caturra vs. *P. sativum*, 85% *C. arabica* cv. Caturra vs. *F. pringlei* and 86% *P. sativum* vs. *F. pringlei*, confirming the degree of conservation of this fragment and evolutionary closeness of genotypes. The Clustal program showed there were very few differences among nucleic acids within the coffee group (98% matching), which suggests this 2% corresponds to positions in which nucleotides have changed according to the sequence results. All genotypes share similar homology to *P. sativum* (79% matching). These results correlate very well to those obtained by blast analysis (***Table 5***).

Table 1. Percentage of amino acids of glycine decarboxylase P-subunit from coffee genotypes and Pisum sativum.

Amino Acid	Genotype				
	Coffea arabica cv. Caturra	*Coffea canephora*	*Coffea eugenioides*	*Coffea congensis*	*Pisum sativum*
Ala A	7.75	7.75	7.75	7.75	7.04
Arg R	3.52	3.52	3.52	3.52	3.52
Asn N	5.63	5.63	5.63	5.63	5.63
Asp D	4.93	4.93	4.93	4.93	2.57
Cys C	0.70	0.70	0.70	0.70	0.70
Glu E	6.34	6.34	6.34	6.34	6.34
Gln Q	4.93	4.93	4.93	4.93	4.23
Gly G	5.63	5.63	5.63	5.63	6.34
His H	1.41	1.41	1.41	1.41	2.11
Ile I	5.63	5.63	5.63	5.63	4.94
Leu L	10.56	10.56	10.56	10.56	9.15
Lys K	4.93	4.93	4.93	4.93	6.34
Met M	4.93	4.93	4.93	4.93	5.63
Phe F	3.52	3.52	3.52	3.52	4.23
Pro P	5.63	5.63	5.63	5.63	4.93
Ser S	7.75	7.75	7.75	7.75	7.04
Thr T	7.04	7.04	7.04	7.04	8.45
Trp W	0.70	0.70	0.70	0.70	0.70
Tyr Y	4.93	4.93	4.93	4.93	5.69
Val V	3.52	3.52	3.52	3.52	4.42

Table 2. Grouping of aminoacids according to chemical characteristics from coffee genotypes and Pisum sativum.

Genotype	A.A. Type (No. of A.A.)			
	Acidic	Basic	Polar	Non Polar
Coffea arabica, cv. Caturra	18	20	74	68
Coffea canephora	18	20	74	68
Coffea congensis	18	20	74	68
Coffea eugenioides	18	20	74	68
Pisum sativum	20	19	75	67

Clustal analysis for amino acids was also performed and matching percentages obtained. Results were similar to those for nucleotide sequences: no significant differences within the coffee group were found (99%) and all were equally different to *P. sativum* (83%) (***Table 6***). One amino acid substitution, corresponding to the change in amino acid 157, was found in the coffee group (***Figure 4B***). Codon GAA, which corresponds to glutamic acid, was found for genotypes *C. arabica* cv. Caturra and *C. eugenioides*. Codon CAA, which corresponds to glutamine, was found for genotypes *C. congensis* and *C. canephora*. In spite of the fact that these amino acids are very related, this change should have some implications on the functioning and structure of the enzyme.

Amplification of fragments of the glycine decarboxylase gene by PCR, using primers designed from homologous sequences of other species, was possible. Probes of the gene of the P-subunit of glycine decarboxylase from four different coffee species have been obtained, cloned and are available for further use. High homology among the four sequences of the glycine decarboxylase gene P-subunit in the *Coffea* genus was established. There is only one important variation which causes an amino acid change in the sequence in this region of the gene (one glutamic acid for one glutamine, in genotypes *C. arabica* cv. Caturra, *C. eugenioides* and *C. congensis*, *C. canephora*, respectively); which could contribute to explain the different photorespiratory rates found for these genotypes (Mosquera, 1995).

High homology of the nucleic acid and amino acid sequences encoded by these fragments in relation to other plants (*P. sativum* and *F. pringlei*, 83% and 85% homology for amino acid sequences) shows that such homology goes beyond the limits of the species. This is the first report of cloning a partial DNA fragment for P-subunit of glycine decarboxylase from coffee and this will allow isolation of the full-length cDNA clones for this gene from coffee for further analysis.

Table 3. *Percentage of nucleic acids of glycine decarboxilase P-unit gene from coffee genotypes and Pisum sativum.*

	Nucleic acids (%)				
Genotype	G	C	A	T	%GC
Coffea arabica cv. Caturra	21.55	19.43	28.10	30.92	40.98
Coffea canephora	21.78	19.20	28.10	30.92	40.98
Coffea congensis	21.78	19.20	28.10	30.92	40.98
Coffea eugenioides	21.55	19.66	28.10	30.69	41.21
Pisum sativum	22.25	18.74	29.74	29.27	40.97

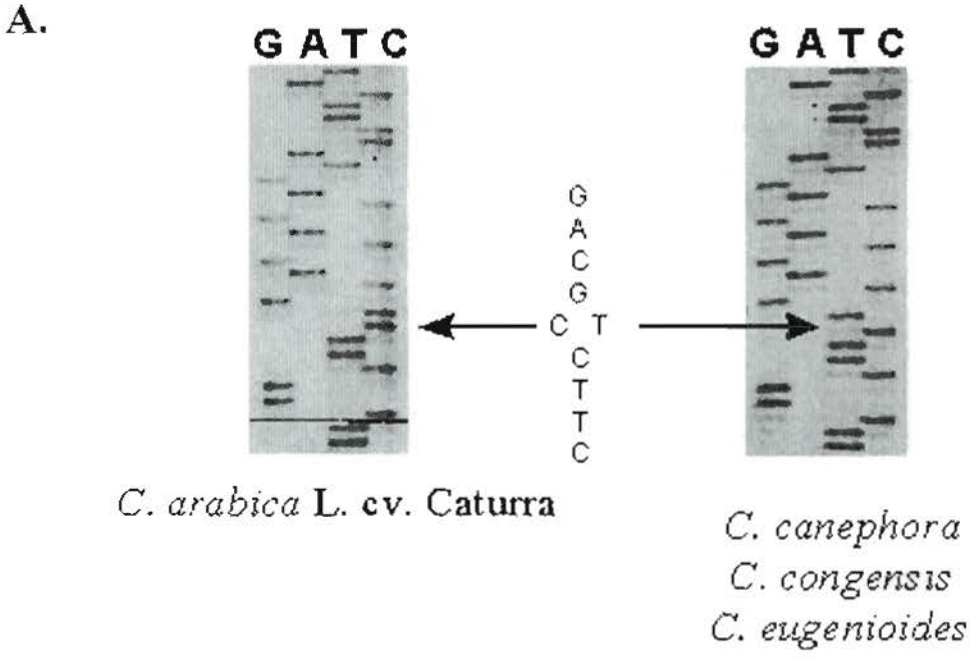

Figure 4. *Analysis of the sequences of the different genotypes and their differences. A. Difference in base 423, changes C-T B. Difference in base 175, change of C-G.*

```
emb|X59773|PSGDPP Pisum sativum mRNA for P protein, a part of glycine
           cleavage complex,
           Length = 3404.
           Plus Strand HSPs:

 Score = 1334 (368.6 bits), Expect = 3.3e-103, P = 3.3e-103
 Identities = 338/427 (79%), Positives = 338/427 (79%), Strand = Plus / Plus

CAT:     1 AGACGCCATAACTCAGCAACACCTGAAGAACAAATCAAAATGGTTGAGAATTGTGGATTC 60
           ||||||||||||||||||||||| || |||||||  |||||||  |||  |||  |||
PEA:   377 AGACGCCATAACTCAGCAACACCGGATGAACAAACGAAAATGGCTGAATCGGTTGGTTTC 436

CAT:    61 CCCAGTCTTGATTCGCTTATTGACGCCACTGTGCCTAAATCTATCAGGCTTGATGGTATG 120
             |  |||||||| ||  |||| || || |||||  |||||  ||  ||  |  |  |||
PEA:   437 GATACGCTTGATTCACTCGTTGATGCAACCGTGCCGAAATCCATTCGGTTGAAGGAAATG 496

CAT:   121 ACGTTTAGTAAGTTTGATGAGGGATTAACTGAGGCTCAAATGATTGATCACATGCAAAAG 180
           | ||| | ||| |||||||  |||||  || || | ||||||||||||  ||||||  |  |
PEA:   497 AAGTTCAATAAATTTGATGGTGGATTGACAGAAGGTCAAATGATTGAGCACATGAAGGAT 556

CAT:   181 TTAGCATCTAAGAACAAAGTTTTTAAGTCATATATTGGAATGGGATACTATAATACCTTT 240
           |||||  ||  ||  |||||  |||||  ||  ||  |  ||||||  ||||||||||||||  ||    |
PEA:   557 TTAGCTTCGAAAAACAAGGTTTTCAAATCGTTTATTGGTATGGGATACTATAACACTCAT 616

CAT:   241 GTTCCGCCTGTTATTTTGAGGAATCTCCTGGAAAATCCTGCTTGGTATACTCAGTATACT 300
           |||||  |||||  ||||||||||||  ||  ||||  |||||||||||||||||  ||||||||
PEA:   617 GTTCCACCTGTGATTTTGAGGAACATCATGGAGAATCCTGCTTGGTATACACAGTATACA 676

CAT:   301 CCCTATCAGGCTGAGATTTCGCAGGGACGTCTTGAATCCCTGCTGAATTATCAGACCATG 360
           ||  |||||  |||||||  ||  ||  ||  ||||||||||  ||  ||||||  ||||||||||
PEA:   677 CCTTATCAAGCTGAGATATCTCAAGGCCGTCTTGAATCTTTGTTGAATTTTCAGACCATG 736

CAT:   361 ATTGCGGATCTTACTGGATTGCCTATGTCTAATGCATCTTTACTAGATGAAGGTACTGCT 420
           |||  |  |||||  |||||  ||||||||||  |||||  ||  ||  ||  ||||||||||||||||
PEA:   737 ATTACTGATCTCACTGGTTTGCCTATGTCAAATGCTTCATTGCTTGATGAAGGTACTGCT 796

CAT:   421 GCCGCAG 427
           ||  ||||
PEA:   797 GCTGCGG 803
```

Figure 5. *Homology analysis of nucleic acid sequences of C arabica cv. Caturra (CAT) vs. Pisum sativum (PEA).*

```
sp|P26969|GCSP_PEA GLYCINE DEHYDROGENASE (DECARBOXYLATING) PRECURSOR
           (GLYCINE DECARBOXYLASE) (GLYCINE CLEAVAGE SYSTEM P-PROTEIN)
           pir||A42109 glycine dehydrogenase (decarboxylating) (EC 1.4.4.2)
           component P precursor - garden pea gi|20741 (X59773) component of
           aminomethyltransferase [Pisum sativum]. Length = 1057

 Score = 623 (284.1 bits), Expect = 3.2e-79, P = 3.2e-79
 Identities = 119/142 (83%), Positives = 130/142 (91%)

CAT:     1 RRHNSATPEEQIKMVENCGFPSLDSLIDATVPKSIRLDGMTFSKFDEGLTEAQMIDHMQK 60
           RRHNSATP+EQ KM E+ GF +LDSL+DATVPKSIRL  M F+KFD GLTE QMI+HM+
PEA:   100 RRHNSATPDEQTKMAESVGFDTLDSLVDATVPKSIRLKEMKFNKFDGGLTEGQMIEHMKD 159

CAT:    61 LASKNKVFKSYIGMGYYNTFVPPVILRNLLENPAWYTQYTPYQAEISQGRLESLLNYQTM 120
           LASKNKVFKS+IGMGYYNT VPPVILRN++ENPAWYTQYTPYQAEISQGRLESLLN+QTM
PEA:   160 LASKNKVFKSFIGMGYYNTHVPPVILRNIMENPAWYTQYTPYQAEISQGRLESLLNFQTM 219

CAT:   121 IADLTGLPMSNASLLDEGTAAA 142
           I DLTGLPMSNASLLDEGTAAA
PEA:   220 ITDLTGLPMSNASLLDEGTAAA 241
```

Figure 6 *Homology analysis of amino acid sequence of Coffea arabica cv. Caturra (CAT) vs Pisum sativum (PEA)*

```
PEA     AGACGCCATAACTCAGCAACACCGGATGAACAAACGAAAATGGCTGAATCGGTTGGTTTC
CAT     AGACGCCATAACTCAGCAACACCTGAAGAACAAATCAAAATGGTTGAGAATTGTGGATTC
CAN     AGACGCCATAACTCAGCAACACCTGAAGAACAAATCAAAATGGTTGAGAATTGTGGATTC
CON     AGACGCCATAACTCAGCAACACCTGAAGAACAAATCAAAATGGTTGAGAATTGTGGATTC
EUG     AGACGCCATAACTCAGCAACACCTGAAGAACAAATCAAAATGGTTGAGAATTGTGGATTC
        ************************ ** ******* ******* *** ..  *** ***

PEA     GATACGCTTGATTCACTCGTTGATGCAACCGTGCCGAAATCCATTCGGTTGAAGGAAATG
CAT     CCCAGTCTTGATTCGCTTATTGACGCCACTGTGCCTAAATCTATCAGGCTTGATGGTATG
CAN     CCCAGTCTTGATTCGCTTATTGACGCCACTGTGCCTAAATCTATCAGGCTTGATGGTATG
CON     CCCAGTCTTGATTCGCTTATTGACGCCACTGTGCCTAAATCTATCAGGCTTGATGGTATG
EUG     CCCAGTCTTGATTCGCTTATTGACGCCACTGTGCCTAAATCTATCAGGCTTGATGGTATG
          ..* ******** ** ** **** ** ** ***** ***** ** ** *  * * ***

PEA     AAGTTCAATAAATTTGATGGTGGATTGACAGAAGGTCAAATGATTGAGCACATGAAGGAT
CAT     ACGTTTAGTAAGTTTGATGAGGGATTAACTGAGGCTCAAATGATTGATCACATGCAAAAG
CAN     ACGTTTAGTAAGTTTGATGAGGGATTAACTGAGGCTCAAATGATTGATCACATGGAAAAG
CON     ACGTTTAGTAAGTTTGATGAGGGATTAACTGAGGCTCAAATGATTGATCACATGGAAAAG
EUG     ACGTTTAGTAAGTTTGATGAGGGATTAACTGAGGCTCAAATGATTGATCACATGCAAAAG
        *.***.* *** ******* ***** ** ** * ********** ****** *  *

PEA     TTAGCTTCGAAAAACAAGGTTTTCAAATCGTTTATTGGTATGGGATACTATAACACTCAT
CAT     TTAGCATCTAAGAACAAAGTTTTTAAGTCATATATTGGAATGGGATACTATAATACCTTT
CAN     TTAGCATCTAAGAACAAAGTTTTTAAGTCATATATTGGAATGGGATACTATAATACCTTT
CON     TTAGCATCTAAGAACAAAGTTTTTAAGTCATATATTGGAATGGGATACTATAATACCTTT
EUG     TTAGCATCTAAGAACAAAGTTTTTAAGTCATATATTGGAATGGGATACTATAATACCTTT
        ***** ** ** ***** ***** ** ** * ****** ************** ** . *

PEA     GTTCCACCTGTGATTTTGAGGAACATCATGGAGAATCCTGCTTGGTATACACAGTATACA
CAT     GTTCCGCCTGTTATTTTGAGGAATCTCCTGGAAAATCCTGCTTGGTATACTCAGTATACT
CAN     GTTCCGCCTGTTATTTTGAGGAATCTCCTGGAAAATCCTGCTTGGTATACTCAGTATACT
CON     GTTCCGCCTGTTATTTTGAGGAATCTCCTGGAAAATCCTGCTTGGTATACTCAGTATACT
EUG     GTTCCGCCTGTTATTTTGAGGAATCTCCTGGAAAATCCTGCTTGGTATACTCAGTATACT
        ***** ***** ********** .. ** **** ***************** ********.

PEA     CCTTATCAAGCTGAGATATCTCAAGGCCGTCTTGAATCTTTGTTGAATTTTCAGACCATG
CAT     CCCTATCAGGCTGAGATTTCGCAGGGACGTCTTGAATCCCTGCTGAATTATCAGACCATG
CAN     CCCTATCAGGCTGAGATTTCGCAGGGACGTCTTGAATCCCTGCTGAATTATCAGACCATG
CON     CCCTATCAGGCTGAGATTTCGCAGGGACGTCTTGAATCCCTGCTGAATTATCAGACCATG
EUG     CCCTATCAGGCTGAGATTTCGCAGGGACGTCTTGAATCCCTGCTGAATTATCAGACCATG
        ** ***** ******** ** ** ** ********** .. ** ****** **********

PEA     ATTACTGATCTCACTGGTTTGCCTATGTCAAATGCTTCATTGCTTGATGAAGGTACTGCT
CAT     ATTGCGGATCTTACTGGATTGCCTATGTCTAATGCATCTTTACTAGATGAAGGTACTGCT
CAN     ATTGCGGATCTTACTGGATTGCCTATGTCTAATGCATCTTTACTAGATGAAGGTACTGCT
CON     ATTGCGGATCTTACTGGATTGCCTATGTCTAATGCATCTTTACTAGATGAAGGTACTGCT
EUG     ATTGCGGATCTTACTGGATTGCCTATGTCTAATGCATCTTTACTAGATGAAGGTACTGCT
        *** * ***** ***** ********** ***** ** ** ** ***************


PEA     GCTGCGG
CAT     GCCGCAG
CAN     GCTGCGG
CON     GCTGCGG
EUG     GCTGCGG
        ** ****
```

Figure 7. Clustal analysis of nucleic acids for all genotypes. PEA: Pisum sativum. CAT. Coffea arabica cv. Caturra. CAN: Coffea canephora, CON: Coffea congensis, EUG: Coffea eugenioides.

Table 4. *Restriction analysis from studied coffee genotypes and Pisum sativum.*

Restriction Enzyme	Genotype				
	Coffea arabica Cv. Caturra	*Coffea canephora*	*Coffea congensis*	*Coffea eugenioides*	*Pisum sativum*
AccI	286, 295	286, 295	286, 295	286, 295	286, 295
AccII	-	-	-	-	170
AcyI	3, 83, 327	3, 83, 327	3, 83, 327	3, 83, 327	3
AluI	-	-	-	-	184, 310
BclI	165	165	165	165	-
Csp45I	-	-	-	-	187
	12,149,188	12,149,188	12,149,188	12,149,188	
DdeI	290, 311	290, 311	290, 311	290, 311	12, 311
DpnI	167, 368	167, 368	167, 368	167, 368	368
EcoRV	-	-	-	-	317
HaeIII	-	-	-	-	326
HapII	-	-	-	-	22
HgaI	11, 91	11, 91	11,91	11, 91	11
HincII	-	-	-	-	-
HinfI	56, 70	56, 70	56, 70	56, 70	46, 70, 237,
	261, 334	261, 334	261, 334	261, 334	334
MboI	165, 366	165, 366	165, 366	165, 366	366
MboII	37	37	37	37	-
MspI	-	-	-	-	22
NlaIII	174, 360	174, 360	174, 360	174, 360	174, 241
					270, 360
NsiI	396	396	396	396	-
NspI	174	174	174	174	-
PstI	426	-	-	-	-
RsaI	414	414	141	414	414
TaqI	-	-	-	-	59, 187
XhoII	366	366	366	366	-
	GROUP 1				GROUP 2

Table 5. *Homology percentages of the nucleic acid sequence among genotypes according to Clustal analysis.*

Genotype	*Coffea canephora*	*Coffea congensis*	*Coffea eugenioides*	*Pisum sativum*
Coffea arabica cv. Caturra	98	98	98	79
Coffea canephora	-	100	98	79
Coffea congensis	-	-	98	79
Coffea eugenioides	-	-	-	79

Table 6. *Homology percentages of nucleic acids among genotypes according to Clustal analysis*

Genotype	*Coffea canephora*	*Coffea congensis*	*Coffea eugenioides*	*Pisum sativum*
Coffea arabica cv. Caturra	99	99	89	83
Coffea canephora	-	100	98	83
Coffea congensis		-	98	83
Coffea eugenioides			-	83

Acknowledgements

This research was supported by Grant 2251-07-002-93 of Colombian Institute for the Science and Technology Development "Francisco José de Caldas" - COLCIENCIAS and National Federation of Coffee Growers of Colombia. We are grateful to Dr. A. Gaitán (Cenicafé) for his assistance in the design of primers and the sequencing of the first clone obtained, Dr. L.A. Murillo (Immunology Institute, Universidad Nacional de Colombia, Bogotá, Colombia) for the assistance in the sequencing of DNA fragments. We also thank to Dr. G. Podila (Michigan Tech. University) for critical reading of the manuscript.

Bibliography

Ausubel, F.M.; Brent, R.; Kingston, R.E.; Moor, D.D.; Seidman, J.A.; Smith, J.G.; Struhl, A. 1994. Current protocols in molecular biology Vol. 1 and 2, John Willey & Sons, Massachusetts.

Canvin, D.T. 1990. Photorespiration and CO_2 concentrating mechanisms, *In*: Plant Physiology, Biochemistry and Molecular Biology. Longman Scientific & Technical, New York. pp. 253 - 273.

Chaparro, K. 1993. Contribución al estudio del genoma del cafeto. Construcción de una biblioteca genómica de *Coffea arabica* L. variedad Caturra, Tesis, Universidad Nacional de Colombia, Bogotá. pp. 26-28.

Decker, J.P.; Tio, M.A. 1959. Photosynthetic surges in coffee seedlings. J. Agr. Univ. Puerto Rico. 43: 50 – 55.

Kopriva, S.; Bauwe, H. 1994. P-protein of Glycine Decarboxilase from *Flaveria pringlei*. Plant Physiol. 104: 1077 - 1078.

Kume, A.; Koyata, H.; Sakakibara, T.; Ishiguro, Y.; Kure, S.; Hiraga, K. 1991. The Glycine decarboxylase cleavage system. J. Biol. Chem. 266: 3323 - 3329.

Marie, D.; Brown, C. 1993. A cytometric exercise in plant DNA histograms with 2C values for 70 species. Biol Cell. 78: 41 - 51.

Mettler, I.J. 1987. A simple and rapid method for minipreparation of DNA from tissue cultured plant cells. Plant Mol. Biol. Rep. 5: 346 - 318.

Mosquera, L.P. 1995. Magnitud de la fotorrespiración y la respiración mitocondrial en hojas de diferentes cultivares de *Coffea arabica* L., Tesis, Universidad del Cauca, Popayán, Colombia. pp. 44 - 57.

Oliver, D.J. 1994. The Glycine decarboxylase complex from plant mitochondria. Ann. Rev. Plant Physiol. 45: 323 - 337.

Sambroock, J.; Fritsch, E.F.; Maniatis, T. 1989. Molecular Cloning: A laboratory manual. Cold Spring Harbor Laboratory. New York. pp. 28 – 54.

Turner, S.R.; Ireland, R.; Rawsthorne, S. 1992. Cloning and characterization of the P-subunit of Glycine decarboxylase from pea (*Pisum sativum*). J. Biol. Chem. 267: 5355 - 5360.

Whelan, J.; Glaser, E. 1997. Protein import into plant mitochondria. Plant Mol. Biol. 33: 771 - 789.

Zelitch, I. 1971. Photosynthesis, Photorespiration and Plant Productivity. Academic Press, New York. pp. 1 – 154.