RAPD FRAGMENT SEQUENCE ANALYSIS IN COMMON BEAN

Pañeda A. , Español Y. , Rodríguez C., Giráldez R.
Área de Genética Dpto. Biología Funcional University of Oviedo, Spain.

INTRODUCTION

Phaseolus vulgaris is one of the smallest genomes in legumes (0'66pg/haploid genome) with approximately 637Mbp (Arumuganathan & Earle, 1991). Little information was available about the *P. vulgaris* genetic map until the development of new molecular markers in the 90´s that allowed an increase in the number of markers and maps. In recent years, the more useful tools for breeding programs are SCAR markers, due to their reproducibility and specificity. So the objectives of this study were to obtain a new core map based on SCARs markers and determine if there is any specific association of the polymorphic RAPD sequences with cloned genes or intergenic sequences.

MATERIAL AND METHODS

DNA sequences from polymorphic RAPD markers uniformly located in the 11 bean linkage groups were analyzed. This selection was based in a map developed by our group (Méndez de Vigo, 2001). The general molecular characteristic of the markers selected are present in Table 1: name of the Operon RAPD, map location, amplification pattern for the parents of our F2 mapping population, and the molecular weight estimated at first by comparison with a ladder and then by sequencing.

The selected fragments were cloned and the nucleotide sequence was determined on both strains. The PCR fragments were analysed with a ABI PRISM 3100 Genetic Analyzer (Applied Biosystems) and the homology with Gen Bank sequences was performed using BLAST analysis. It is worth mentioning that the Gen Bank contains only 1,495 nucleotide entries and 760 protein entries for *Phaseolus vulgaris*, so most of the homologies will be related with other species**.** In recent years, the number of legumes being sequencing is increasing so we expect that the number of sequences and species analyzed will increase and the homologies studied should be better. We have just concentrated in plant sequences, so any other homologies are not shown.

Table 1. Summary table of the RAPD markers analyzed

		Amplification pattern			
RAPD	**LG**	**A25**	**A252**	**MW_1**	**MW_2**
OAS15	4B	0	1	250	264
OD13	2D	0	1	625	633
OE02	6G	1	0	490	496
OE04	8F	0	1	700	724
OE15	9K	0	1	525	551
OF04	2D	1	0	750	735
OF10	1H	0	1	500	479
OG14	10I	1	0	700	671
OG16	3C	1	0	375	374
OH05	7A	0	1	750	745
OH08	10I	1	0	950	968
OH18	5E	1	0	1000	*
OY08	11J	0	1	350	361

LG: linkage group, MW1: Molecular weight estimated by comparison with molecular weight marker, MW2: Molecular weight estimated by sequencing, *no estimation

RESULTS

Present work displays the study of 13 sequences, from which only 5 exhibit an E value under or equal to 10^{-13}, a significant value determined by other homology studies. There are also 8 sequences that exhibit E values bigger than 0,05**.**

One of the best values for E was obtained for the $OE02^{490}$ RAPD sequence, with an arcelin protein from *Phaseolus vulgaris*. This protein is part of the lectin family, which is coded by a genetic family mapped to the B4 linkage group (Freyre *et al*., 1998). There are other loci for these genetic families, that have not been mapped yet (Sales *et al*., 2000; Kelly *et al*., 2003).

Our marker is located in B6 linkage group in our map (Méndez de Vigo, 2001), which suggest that one of the other loci for the lectin family is B6 linkage group.

The OF10^{500} fragment was described by Young *et al* in 1997, as a molecular marker linked from *Co-1* at 1.4-1.9 cM (anthracnose resistance gene). Our group has already mapped it in B1 linkage group (Méndez de Vigo, 2001) and has developed the SCAR marker that would be very useful in all the breeding programs involving this gene. Nevertheless, we found that the sequence structure (Figure 1) has 22 repeats of a 13 bp sequence in which the primer is included (5'CCAAGCTTCCATA 3'). A shorter sequence of 197 bp without any repeats was selected to design the new 24 bp SCAR-primers that were specific for one band. No significant homology was found for this fragment, although it is supposed to be related with anthracnose resistance genes.

Figure 1. Complete sequence of OF10^{500} fragment. It is indicated with bold letters the RAPD primers and with italic letters the 22 sequence repeats

5'YGGAAGCTTGGGGAACACGAGCTCGAAATGAATAGGCTTGTTGTCCAAGAAAGTGAGGAAAAACA
CAACAAGGGCATAACACTCTAAGCTTGCAATCACAAACAACAACAAGACTCAAGTGGAAGTGATGAA
GACACTATGAGTTTGTTGTCAAGAAAATTCAACAAGTTCTTGAAGAAGAAAAG*CCAAGCTTCC*ATA*CCA*
*AGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*
ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCA*
*AGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*
ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA*CCAAGCTTCC*ATA**CCAAGCTCCY3'**

The OD13^{625} sequence presented a good homology with the endopolygalacturonase inhibitor protein gene (PGIP) from *P. vulgaris* with a very significant E value. The location of this marker (Freyre *et al*., 1998) with the design of a RFLP from the gene sequence (Toubart *et al*., 1992) is also in chromosome B2 where our OD13^{625} sequence is mapped (Méndez de Vigo, 2001). Therefore we can assume that this sequence is part of the gene that codify for the PGIPs.

Another good homology (E=2x10-88) was observed for the OG14^{700} sequence. In this case results of BLAST showed that the homology is not for the complete sequence of mARN C subunit vacuolar H+-ATPase, because there is a 350 bp fragment in the OG14^{700} that is not present in other plant species. This small fragment could correspond to an intron sequence, and therefore is not present in the processed mARN of other species.
Only most relevant results for homology are present here, but there are quite a number of other sequences that have good homology with retroelements. These mobile genetic elements are present in plant genomic DNA in a high percentage (>50% corn genome and around 90% wheat genome) and it is also considered that these elements have significantly contributed to the evolution of structure and genome function (McCarthy *et al*., 2002). About the question: Could it make sense to relate polymorphic sequences with intergenic or conserved sequences and monomorphic with transcription sequences? We are unable to concluded any theory with these few sequences, but at least we can see that about 50% of them are sequences related with retroelements, 30% code transcriptional genes and finally a small percentage is not characterized.

REFERENCES

Arumuganathan K and Earle ED. 1991. Nuclear DNA content of some important plant species. Plant. Mol. Bio. Rep., 9: 208-218.

Freyre R, Skroch PW, Geffroy V, Adam-Blondon A-F, Shirmohamadali A, Johnson WC, Llaca V, Nodari RO, Pereira PA, Tsai SM, Thome J, Dron M, Nienhus J, Vallejos CE and Gepts P. 1998. Towards an integrated linkage map of common bean IV. Development of a core linkage map and alignment of RFLP maps. Theor Appl Genet, 97: 847-856

Kelly JD, Gepts P, Miklas PN and Coyne DP. 2003. Tagging and mapping of genes and QTL and molecular marker-assisted selection for traits of economic importance in bean and cowpea. Field Crops Research, 82: 135-154

McCarthy EM, Liu J, Lizhi G and McDonald JF. 2002. Long terminal repeat retrotransposons of Oryza sativa. Genome Biology, 3(10): 1–11

Méndez de Vigo B. 2001. Mapa genético de *Phaseolus vulgaris* L. y resistencia a antracnosis en faba granja asturiana. Tesis Doctoral, Universidad de Oviedo

Sales MP, Gerhardt IR, Grossi-de-Sá MF and Xavier-Filho J. 2000. Do Legume Storage Proteins Play a Role in Defending Seeds against Bruchids? Plant Physiology, 124(2): 515–522

Toubart P, Desiderio A, Salvi G, Cervone F, Daroda L and De Lorenzo G. Cloning and characterization of the gene encoding the endopolygalacturonase-inhibiting protein (PGIP) of *Phaseolus vulgaris* L. 1992. Plant J., 2(3):367-373