VERSITA

CELLULAR & MOLECULAR BIOLOGY LETTERS
Volume 12 (2007) pp 120 - 126
http://www.cmbl.org.pl

DOI: 10.2478/s11658-006-0053-z

Received: 16 May 2006
Revised form accepted: 06 October 2006
Published online: 22 November 2006

Short communication

SIRE-1, A PUTATIVE PLANT RETROVIRUS IS CLOSELY RELATED TO A LEGUME TY1-COPIA RETROTRANSPOSON FAMILY

STEPHEN R. PEARCE*
[1]School of Life-Sciences, Department of Biology and Environmental Science, University of Sussex, Falmer, Brighton, BN1 9QG, United Kingdom

Abstract: SIRE-1 is a potential soybean retrovirus which has a gene order similar to Ty1-*copia* retrotransposons but also contains an envelope-like open reading frame (ORF), which is characteristic of retroviruses. PCR and Southern analysis reveals that SIRE-1 is closely related to a legume-wide family of envelope-lacking Ty1-*copia* group retrotransposons which suggests that SIRE-1 was formed by the recent acquisition of an envelope gene by a Ty1-*copia* retrotransposon.

Key words: *Glycine max,* Retroelement, RNAseH

INTRODUCTION

Ty1-*copia* retrotransposons are mobile genetic elements that are abundant in plants, are closely related to the retroviruses, but have a different polyprotein (pol) gene order to that of retroviruses and Ty3-*gypsy* elements [1]. Retroviruses also have a characteristic envelope gene (env) which enables the virus to be infectious and defines their host range [2]. Infectious retroviruses although common in animals, have not been identified in plants although several examples of plant retrotransposons with env-like genes have been reported [3-6]. The soybean retroelement SIRE-1 is highly unusual as it contains an envelope gene characteristic of retroviruses but unlike most other envelope-containing retroelements which have an internal structural organization similar to Ty3-*gypsy* elements, SIRE-1 exhibits the gene order of Ty1-*copia* retrotransposons [7]. Extensive sequence analysis shows that all characterized SIRE-1 insertions

* E -mail: S.R.Pearce@sussex.ac.uk

Abbreviations used: env – envelope; orf – open reading frame; pol – polyprotein

are relatively recent, suggesting that this element is a relatively new feature of the soybean genome [8]. Despite extensive sequence analysis the origins of SIRE-1 remain unclear but two major possibilities exist. Firstly that SIRE-1 was introduced as a complete sequence into the soybean genome by infection or secondly that SIRE-1 was formed by the insertion of an envelope gene into a pre-existing soybean retroelement.

This investigation uses Ty1-*copia* specific primers to isolate a population of SIRE-1 related sequences from soybean, pea and broad bean. Analysis of the RNAseH sequences and Southern hybridization data show that soybean SIRE-1 sequences form a single envelope-containing lineage within a closely related and widespread population of envelope-lacking legume Ty1-*copia* retrotransposons which are present in pea, broad bean and possibly all other legume species. These findings strongly suggest that SIRE-1 was formed by the acquisition of the envelope gene by a conventional soybean Ty1-*copia* retrotransposon.

MATERIALS AND METHODS

Plant materials

Seeds of soybean (*Glycine max* cv. Williams) were donated by Prof. R. Shoemaker, Iowa state University USA. Pea, *Pisum sativum* JI15 was from the John Innes collection, John Innes Centre, Norwich UK. Broad bean (*Vicia faba* cv. troy, was from the SCRI collection, SCRI, Dundee UK). Surface sterilized seeds were germinated and DNA isolated by "qiaquick plant mini kit" (Qiagen).

Generation of RNAseH LTR sequences and sequence analysis

RNAseH-LTR sequences were amplified by a modification of the S-SAP technique [9]. The second nested primer stage was omitted which enabled RNAseH sequences to be recovered from the upstream primer site (RTKHID). S-SAP products were subcloned into pCR2.1-TOPO vector (Invitrogen) by TOPO TA cloning, and the recombinant plasmids transformed into "Top10 one shot" competent cells (Invitrogen). Plasmid inserts were isolated by PCR using M13 forward and M13 reverse primers, and purified by PCR purification kit (Qiagen). Purified inserts were sent to the Advanced Biotechnology Centre, Charing Cross Hospital, London, UK for sequence determination. DNA sequences were compared and phylogenetic trees generated using CLUSTALW [10].

Southern analysis

Southern analysis was carried out using 3µg of *EcoRI* digested genomic DNA per lane. Slot blot transfer was carried out as [11] but with 3µg DNA per slot. Probes were labeled and detected using AlkPhos direct labeling and CDP-Star (Amersham). SIRE-1 envelope (AY205610 5870-6282) and RNAseH (AY205610 5428-5842) probes were generated by PCR from soybean DNA using primer pairs E1/E2 and R1/R2 respectively.

E1 = 5'-CGTTAATAGCGCGTTCTCTACTGGG-3'
E2 = 5'-GCGGATCTTCTTATCCATCTCCG-3'
R1 = 5'-GGTGGATGTTTCTATTTGGG-3'
R2 = 5'-GCTATAAATCCTCTAGCAGAC-3'
Probe composition was verified by subcloning and sequencing (data not shown).

RESULTS

Isolation and sequence analysis of SIRE-1 family sequences

Retrotransposon sequences originating within the RNAseH and continuing approximately 200bp downstream, and including the 5' end of the SIRE-1 env gene (Fig. 1) were generated by a modification of the S-SAP technique [9] from soybean, broad bean and pea. Sequence analysis reveals a heterogeneous population of Ty1-*copia* group retrotransposons closely related to SIRE-1 (Fig. 1). All soybean RNAseH sequences share high levels of sequence similarity with the SIRE-1 sequence with Tgm5 being the most closely related. The sequences

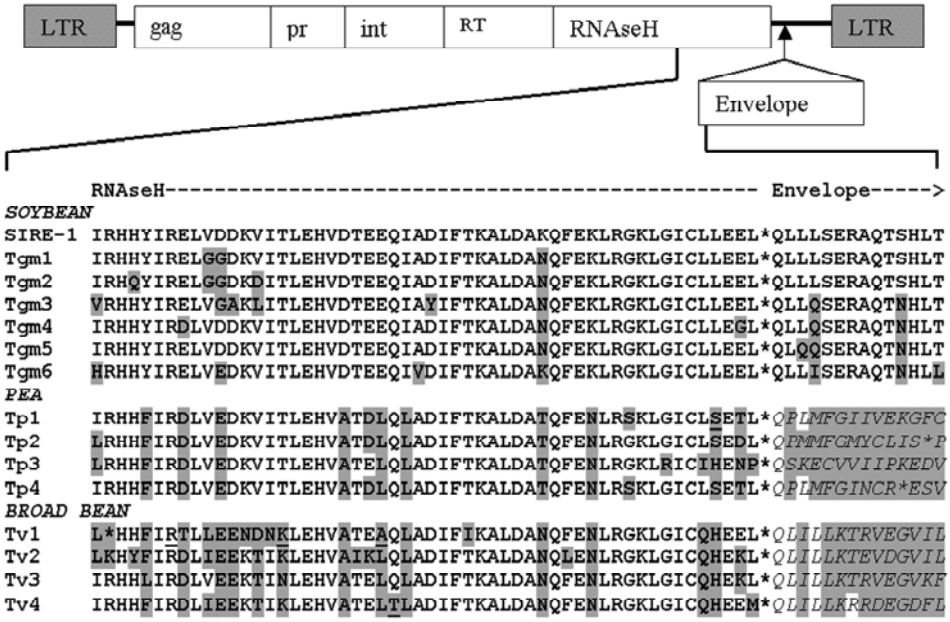

Fig. 1. Retrotransposon RNAseH sequences from soybean *(Tgm* and SIRE1-13), pea (*Tp)* and broad bean (*Tv*). Shading shows *differences* to SIRE1-13. Frameshifts are indicated by underlining. Italics represent peptide sequence *not* representing env open reading frame (EMBL accession numbers are AM408574-AM408587).

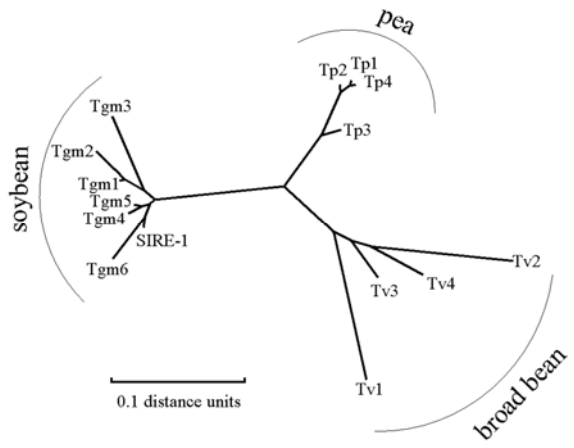

Fig. 2. Unrooted phylogenetic analysis of RNAseH peptide sequences from soybean (*Tgm*1-6 and SIRE1-13), broad bean (*Tv*1-4) and pea (*Tp*1-4).

from pea (Tp1-4) and broad bean (Tv1-4) are closely related to SIRE-1, but further phylogenetic analysis shows them to be clustered along species lines (Fig. 2). As is typical for populations of conventional Ty1-*copia* retrotransposons, each species population is heterogeneous, with the soybean population exhibiting 12.5%, pea 7.5% and broad bean 24% intra-species divergence. This data suggests that these sequences have been present in each species for long enough for them to have evolved differently in each of these species since they diverged from a common ancestor (Fig. 2) which is typical of vertically transmitted retrotransposon populations.

When the sequence corresponding to the 5' end of the envelope gene is scrutinized we see that the gene is present in all of the soybean sequences, but is absent in each of the pea and broad bean sequences (Fig. 3). As well as the lack of the envelope gene, the pea and broad bean sequences in many cases also contain consensus polypurine tracts and LTR termini ($Pu_{(n>5)}T(T)G$; [9]) which are characteristic of conventional envelope-lacking Ty1-*copia* retrotransposons (indicated by the black shading in Fig. 3).

Due to the limited number of sequences recovered in each species and the possibility that the PCR primers used may have preferentially generated envelope containing sequences from soybean and envelope-lacking sequences from pea and broad bean Southern blots were used to investigate the distribution and composition of SIRE-1 relatives in these species (Fig. 4). A SIRE-1 RNAseH probe (panel A and C) hybridizes clearly to the soybean, pea and broad bean lanes indicating that close relatives of SIRE-1 are present in each of these species. The envelope gene (Fig. 4 B, D) in contrast shows strong hybridization to the soybean DNA (panel B) but no hybridization is detected in pea or broad bean confirming that SIRE-1 relatives are present throughout the legumes, but the SIRE-1envelope gene is confined to soybean.

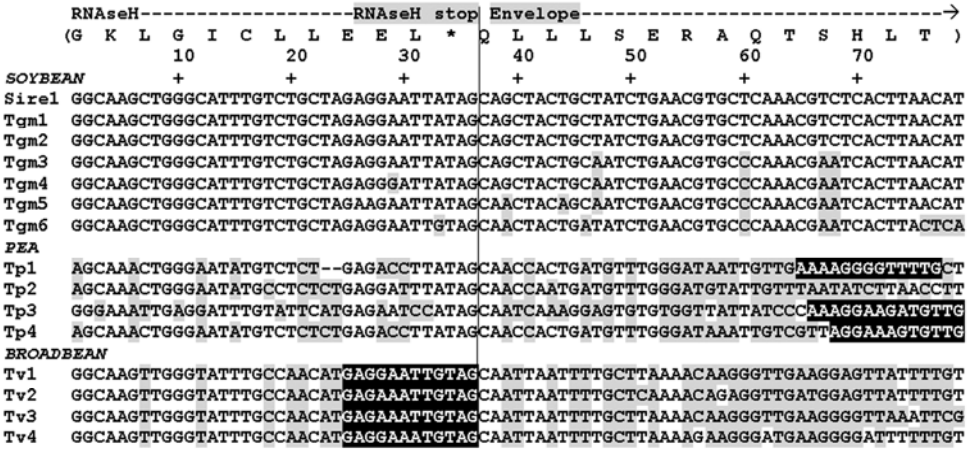

Fig. 3. Retrotransposon RNAseH 3' terminus downstream nucleotide sequences from soybean (SIRE1-13: AY205612 and *Tgm* 1-6), pea (*Tp*1-4*)* and broad bean (*Tv*1-4) retroelements. Shading indicates sequence *differences* to SIRE1-13. Vertical line indicates the 3' terminus of the RNAseH gene. The peptide sequence of the SIRE-1 envelope gene is indicated. Black boxes indicate consensus polypurine tract and putative LTR termini.

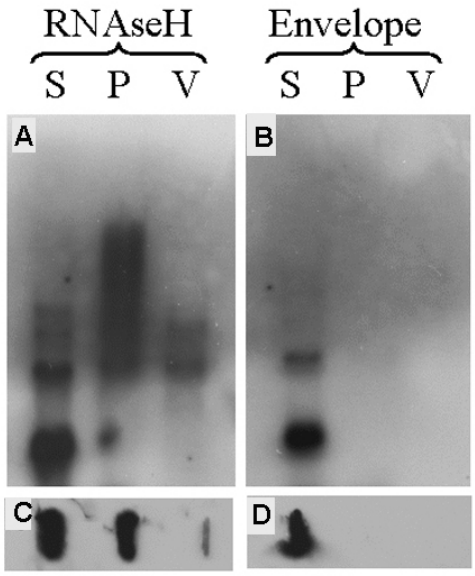

Fig. 4. Southern hybridisation to soybean (S), pea (P) and broad bean (V) DNA. A – SIRE-1 RNAseH probe, B – SIRE-1 envelope probe. Panels C and D are slot blots of the same DNA and probes as the genomic blots in the upper panels.

DISCUSSION

Previous studies have shown that SIRE-1 is a highly unusual retroelement as it has a gene order similar to Ty1-*copia* retrotransposons but has an envelope gene characteristic of retroviruses [12, 13]. This and the relative homogeneity of the SIRE-1 population in soybean has lead to the suggestion that SIRE-1 is a fairly recent acquisition to the soybean genome [14]. The findings presented here show that close relatives of SIRE-1 are present in other legume species the majority of which have characteristics of Ty1-*copia* retrotransposons including consensus polypurine tracts and LTR terminal sequences which are consistent with their lack of envelope genes.

The high level of sequence similarity within the SIRE-1 RNAseH gene at the peptide level suggests that these elements have a common evolutionary history. Their clustering along species lines suggests they have evolved differently in each of the species making it unlikely that the *complete* SIRE-1 sequence entered soybean from a distant source by recent infection. The presence of the envelope gene *only* in soybean, suggests that either the ancestral legume SIRE-1 element had an envelope gene, which was then lost from the pea and broad bean sequences, or alternatively that the ancestral element lacked the envelope gene, which was later acquired by the soybean lineage. If the envelope gene was present in the ancestral element it is unclear why the loss has occurred efficiently in the pea and broad bean lineages but not in soybean.

Although it is possible that the PCR-based techniques used in this study have preferentially isolated envelope-lacking sequences in pea and broad bean this is unlikely. The primer sequences used are able to successfully isolate a wide range of retrotransposon sequences in diverse species [9] and have been shown in this study to readily identify envelope-containing SIRE-1 sequences. More importantly the sequence data is corroborated by Southern hybridization data which shows that SIRE-1 relatives are present in pea and broad bean but the populations importantly lack the envelope gene, which is confined to the soybean lineage.

The presence of heterogeneous populations of envelope-lacking retrotransposons which are related to SIRE-1 in other legume species compared to the homogeneity of the envelope-containing SIRE-1 element population found only in soybean makes it extremely likely that the *complete* SIRE-1 element did not enter the soybean genome from a remote host but that this element was formed by the acquisition of an envelope gene by a regular soybean Ty1-*copia* retrotransposon.

Although it may be more likely that infectious retroviruses may revert to non-infectious retrotransposons via the loss of their envelope genes, this study shows an example of a retrotransposon converting to a potential retrovirus by gaining an envelope gene. Considering the extremely high copy number of retroelements in eukaryotic genomes, this process may play a more significant role in novel retrovirus formation than was previously considered.

REFERENCES

1. Kumar, A. and Bennetzen, J.L. Plant retrotransposons. **Ann. Rev. Genetics** 33 **(**1999) 479-532.
2. Wilson, C., Reitz, M.S., Okayama, H. and Eiden, M.V. Formation of infectious hybrid virions with gibbon ape leukemia virus and human T-cell leukemia virus retroviral envelope glycoproteins and the gag and pol proteins of Moloney murine leukemia virus. **J. Virol.** 63 (1989) 2374–2378.
3. Peterson-Burch B.D., Wright, D.A., Laten H.M. and Voytas D.F. Retroviruses in plants? **Trends Genet.** 16 (2000) 151-152.
4. Vicient C.M., Kalendar, R. and Schulman A.H. Envelope-class retrovirus-like elements are widespread, transcribed and spliced, and insertionally polymorphic in plants. **Genome Res.** 11 (2001) 2041-2049.
5. Wright, D.A. and Voytas, D.F. *Athila4* of *Arabidopsis* and *Calypso* of soybean define a lineage of endogenous plant retroviruses. **Genome Res.** 12 (2002) 122-131.
6. Neumann, P., Pozarkova, D., Koblizkova, A. and Macas, J. PIGY, a new plant envelope class LTR retrotransposon. **Mol. Genet. Genomics** 273 (2005) 43-53.
7. Laten, H.M. Phylogenetic evidence for Ty1-*copia* –like endogenous retroviruses in plant genomes. **Genetica** 107 (1999) 87-93.
8. Laten, H.M. Havecker, E.R., Farmer, L.M. and Voytas, D.F. SIRE1, a endogenous retrovirus family from *Glycine max*, is highly homogeneous and evolutionarily young. **Mol. Biol. Evol.** 20 (2003) 1222-1230.
9. Pearce S.R., Stuart-Rogers C., Knox, M.R., Kumar, A., Ellis, T.H.N and Flavell A.J. Rapid isolation of plant Ty1-*copia* group retrotransposon LTR sequences for molecular marker studies **Plant J.** 19 (1999) 711-717.
10. Higgins, D.G. and Sharp, P.M. CLUSTAL - A package for performing multiple sequence alignment on a microcomputer. **Gene** 73 **(**1998) 237-244.
11. Pearce, S.R., Harrison, G., Li, D., Heslop-Harrison, J.S., Kumar, A. and Flavell, A.J. The Ty1-*copia* group retrotransposons in *Vicia* species: Copy number, sequence heterogeneity and chromosomal localisation. **Mol. Gen. Genet.** 250 (1996) 305-315.
12. Malik, H.S., Hinikoff, S. and Eikbush, T.H. Poised for contagion: Evolutionary origins of the infectious abilities of invertebrate retroviruses. **Genome Res.** 10 **(**2000) 1307-1318.
13. Boeke J.D., Eikbush T.H., Sandmeyer S.B. and Voytas D.F. Pseudoviridae. In: **Virus Taxonomy: Eighth report of the international committee on taxonomy of viruses.** (Fauquet, C.M. Ed,) Academic press, New York (2004) 663-672.
14. Peterson-Burch, B.D. and Voytas, D.F. Genes of the pseudoviridae (Ty1/*copia* retrotransposons) **Mol. Biol. Evol.** 19 **(**2002) 1832-1845.