Collection: COST Action E52 Meeting 2008 - Florence (Italy)
Evaluation of beech genetic resources for sustainable forestry
Guest Editors: Georg von Wühlisch, Raffaello Giannini

Identification and molecular characterization of LTR and LINE retrotransposable elements in *Fagus sylvatica* L.

Emiliani G, Paffetti D, Giannini R

Retrotransposable elements are important and peculiar genetic components derived from ancient retrovirus insertion inside plants genome. Their ability to move and/or replicate inside the genome is an important evolutionary force, responsible for the increase of genome size and the regulation of gene expression. Retrotransposable elements are well characterized in model or crop species like *Arabidopsis thaliana* and *Oryza sativa*, but are poorly known in forest tree species. In this paper the molecular identification of retrotransposable elements in *Fagus sylvatica* L. is reported. Two retrotransposons, belonging to the two major classes of LTR and non-LTR elements, were characterized trough a SCAR (Sequence Characterized Amplified Region) strategy. The analysis demonstrated the presence of multiple copies of retrotransposable elements inside the genome of beech, in accordance with the viral quasi-species theory of retrotransposon evolution. The cloning and sequencing of amplification products and a Cleaved Amplified Polymorphisms (CAPs) approach on the identified retrotransposons, showed a high level of diversity among the multiple copies of both elements. The identification of retrotransposable elements in forest trees represents an important step toward the understanding of mechanisms of genome evolution. Furthermore, the high polymorphism of retrotransposable elements can represent a starting point for the development of new genetic variability markers.

Keywords: ***Fagus sylvatica*, Retrotransposons, Genome evolution, Genetic diversity**

Introduction

Transposable elements (TEs) are ancient (retro)-virus insertions inside a host genome and are peculiar mobile genetic elements accounting for a large proportion of repetitive DNA regions (SanMiguel et al. 1996). TEs (Fig. 1) are classified according to their origin and structure (Craig et al. 2002, Wicker et al. 2007). Class II DNA mediated TEs are usually in low copy number as they transpose with a "cut and paste" process (Casacuberta & Santiago 2003) and are particularly important in prokaryotic and animal genetics (Craig et al. 2002, Kazazian 2004, Biémont & Vieira 2005). Miniature Inverted TEs (MITEs) are defective, non autonomous forms of DNA transposons and are the predominant TE type in or near plant genes (Feschotte et al. 2002). Retrotransposons (RNA mediated Class I TEs) are the most widespread class of eukaryotic TE, particularly relevant in shaping plant genomes (Feschotte et al. 2002, Wicker et al. 2007).

Retrotransposons can be classified in two main classes (Feschotte et al. 2002): non-LTR and LTR elements characterized, respectively, by the absence or presence of long repeated regions flanking the ORFs core. LTR retrotransposons, if still active, increase their copy number with their copy-and-paste mode of transposition. Non-LTR retrotransposons can be divided in several subclasses, such as LINE (Long INterspersed Elements) and SINE (Short INterspersed Elements - Wicker et al. 2007). Both LINEs and SINEs elements, despite the differences in structure and transposition mechanisms (Wicker et al. 2007), are ubiquitous components of eukaryotic genomes, playing a major role in their evolution.

LTR and non-LTR retrotransposons are common in plant genomes where they can be present in high copy number, ranging from 20000 to 200000 copies of *copia*-like LTR elements in barley and maize (Vicient et al. 1999, Meyers et al. 2001); the *gypsy* LTR element is present in 20000 copies in maize. There are 250000 copies of LINE elements in *Lilium* (Leeton & Smyth 1993) and SINEs are represented, for example, by 50000 copies in tobacco (Yoshioka et al. 1993). It is therefore clear that retrotransposable elements build large parts of eukaryotic genomes (Vitte et al. 2007). They represent, for example, a 20% genomic fraction in *Drosophila melanogaster* (Kapitonov & Jurka 2003), 45% in human (TIHGSC 2001) and over 80% in maize (SanMiguel et al. 1996).

The observed differences in genome size in plants are accompanied by variations in the content of LTR retrotransposons, demonstrating that such elements might be important players in the evolution of plant genomes, along with polyploidy (Flowers & Purugganan 2008). The analysis of these elements have also revealed that plant genomes underwent genome amplifications (retrotransposition) and contractions (homologous or illegitimate recombination) leading Vitte & Panaud (2005) to the definition of the increase/decrease model of genome evolution.

TEs have also an important role in gene evolution and regulation (Xiao et al. 2008). Even if they usually integrate into intergenic regions, in maize both LTR-retrotransposons and MITEs elements have been frequently found associated to genes (Wessler 1998).

Retrotransposition can directly affects gene expression by integrating into coding or regulatory regions (White et al. 1994, Durbin et al. 2001); it also acts on tissue specific alternative splicing if inserted in intron regions (Leprince et al. 2001). Deleterious effects of TEs movement and replication are controlled by the development of mechanisms limiting their mutagenic activity, being silencing, probably, the most general and effective one (Vance & Vaucheret 2002).

Since retrotransposable elements are abundant, ubiquitous and highly conserved (especially the LTR elements), they have drawn much attention for the development of genetic diversity and mapping markers; se-

Dept. of Environmental and Forestry Technologies and Sciences, University of Florence, v. S. Bonaventura 13, I-50145 Florence (Italy).

@ Giovanni Emiliani (giovanni.emiliani@unifi.it)

Received: Feb 4, 2009 - Accepted: Mar 25, 2009

Citation: Emiliani G, Paffetti D, Giannini R, 2009. Identification and molecular characterization of LTR and LINE retrotransposable elements in *Fagus sylvatica* L. iForest 2: 119-126 [online: 2009-06-10] URL: http://www.sisef.it/iforest/show.php?id=501

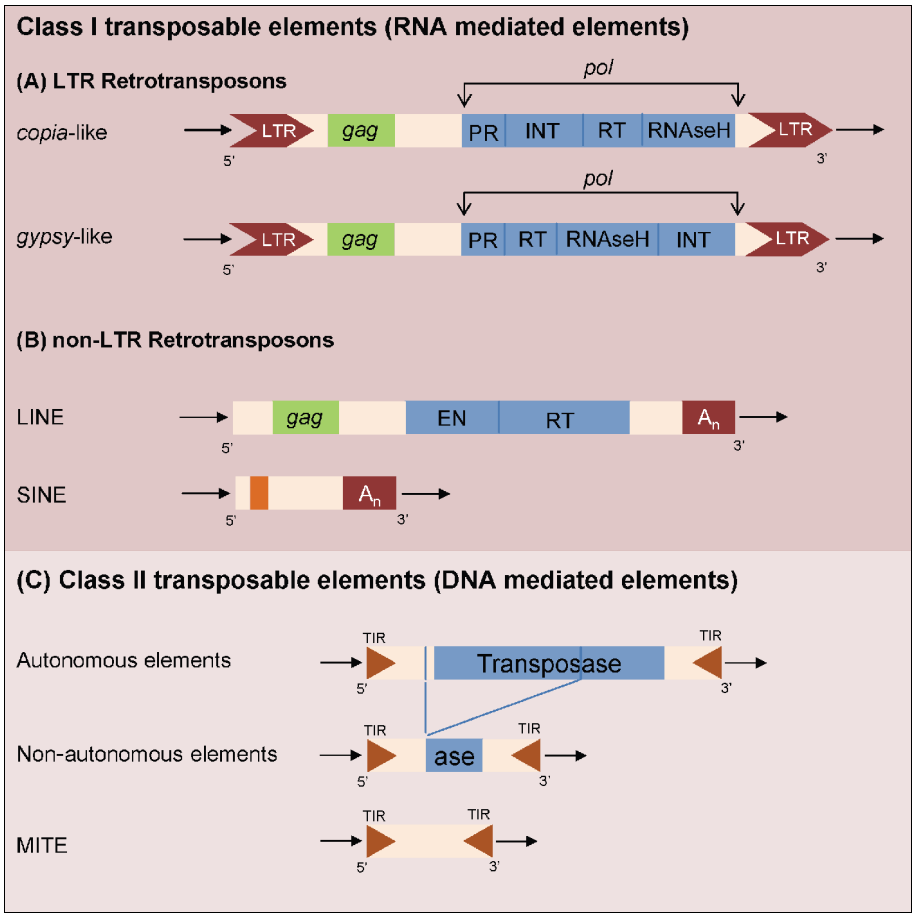

Fig. 1 - Schematic representation of the general structure of class I RNA mediated LTR TE (A), class I RNA mediated non-LTR elements (B) and class II DNA mediated elements (C) in their autonomous and non autonomous configurations. TIR, Terminal Inverted Repeats; LTR, Long Terminal Repeats; *pol* gene contains: PR, protease; EN, endonuclease; RT reverse transcriptase; INT, integrase; *gag* gene encodes a capsid like protein.

veral methods use the high level of homology of LTR regions as annealing sites for PCR based strategies such as S-SAP (Waugh et al. 1997), IRAP and REMAP (Kalendar et al. 1999). These methods have been advantageously used for variability screenings in several species (Kalendar & Schulman 2006) such as wheat (Queen et al. 2004), pea (Smykal 2006) and tomato (Tam et al. 2005). Recently retrotransposable elements have also been used for cultivar fingerprinting and certification (Venturi et al. 2006).

TEs are quite well characterized in crops, model species or agronomical important woody perennials like *Malus* (Venturi et al. 2006), *Citrus*, *Prunus* (Asins et al. 1999) and *Vitis* (Jaillon et al. 2007), whilst in forest trees they are characterized only in conifers (Kamm et al. 1996, L'Homme et al. 2000, Friesen et al. 2001).

In this paper, the molecular identification of LTR and non-LTR LINE retrotransposable elements in the genome of *Fagus sylvatica* is reported. The characterization of these peculiar genetics elements demonstrated their multi-copy presence inside beech genome. The analysis of the polymorphism shown by the identified retrotransposable elements leads also to the proposal of the development of new genetic diversity markers.

Materials and methods

Identification of retrotransposable elements in Fagus sylvatica

Retrotransposable elements of *Fagus sylvatica* were identified trough cloning and sequencing of RAPD (Random Amplified Polymorphic DNA) markers obtained during a genetic diversity screening of Italian beech populations (Emiliani et al. 2004). In such work 1000 beech trees from 30 populations of Southern Italy were analysed with 3 random 10-mer primers (1253, 1247 and RF2 - Paffetti et al. 1996), obtaining 90 reproducible markers. Only two RAPD markers (named 1253/1 and RF2/35) showed a 100% frequency in the dataset and were therefore characterized by using the SCAR (Sequence Characterized Amplified Region) strategy.

SCAR strategy

The two RAPD fragments selected (1253/1: 2000 bp and RF2/35: 680 bp) were excised from agarose gels and purified with the Qiagen Gel Extraction Kit (Qiagen, Crawley, UK), following the manufacturer's specification. DNA fragments were then cloned using the TOPO TA Cloning Kit (Invitrogen, Paisley, UK), following the manufacturer's specifications. Plasmids from positively transformed clones were extracted using the Qiagen Plasmid Extraction Kit and the inserted fragment was sequenced in both directions using an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Foster City, California).

To ensure the homology between the two cloned fragments and the RAPD markers (1253/1, RF2/35), the absence of co-migrant markers and repetitive fragments in the RAPD fingerprinting, high stringency Southern Blot analyses were performed (following the protocol reported in Paffetti et al. 1996) using the cloned fragments as probes against the RAPD profile (Fig. 2).

Sequence analysis

High quality sequences were used as queries for database searching using the BLAST tool (Altschul et al. 1997). Homologous sequences were aligned using the ClustalW software (Larkin et al. 2007). The analysis of sequence polymorphism was performed using the MEGA4 software (Tamura et al. 2007).

Analysis of retrotransposable elements diversity in different beech genotypes

The identified retrotransposable elements of beech were furthermore analysed trough cloning and sequencing of specific amplified fragments. Primers were designed using the

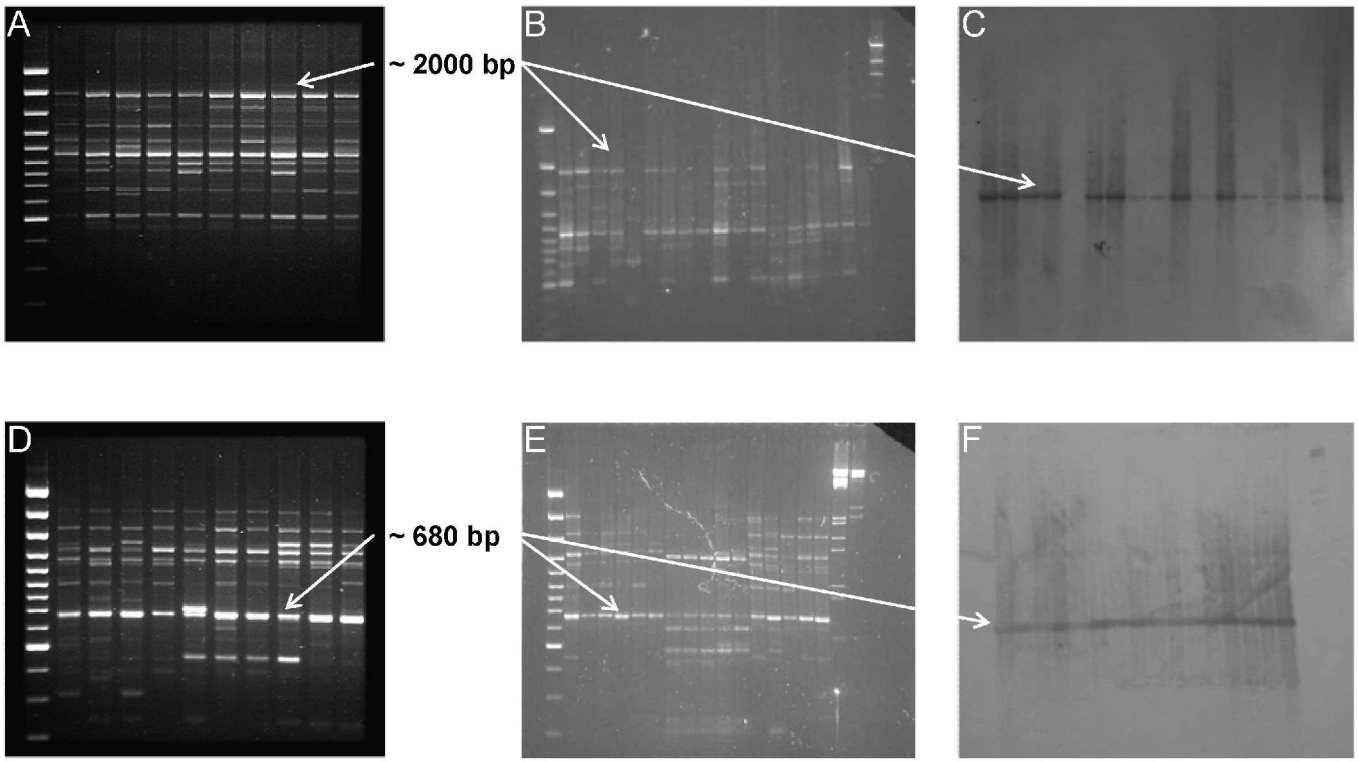

Fig. 2 - Southern Blot analysis of RAPD fragments; RAPD profiles showing the characterized fragments representing the non-LTR (A) and LTR (D) retrotransposable elements. Blotting gels (B,E) and colorimetric results (C, F) of hybridisation analysis.

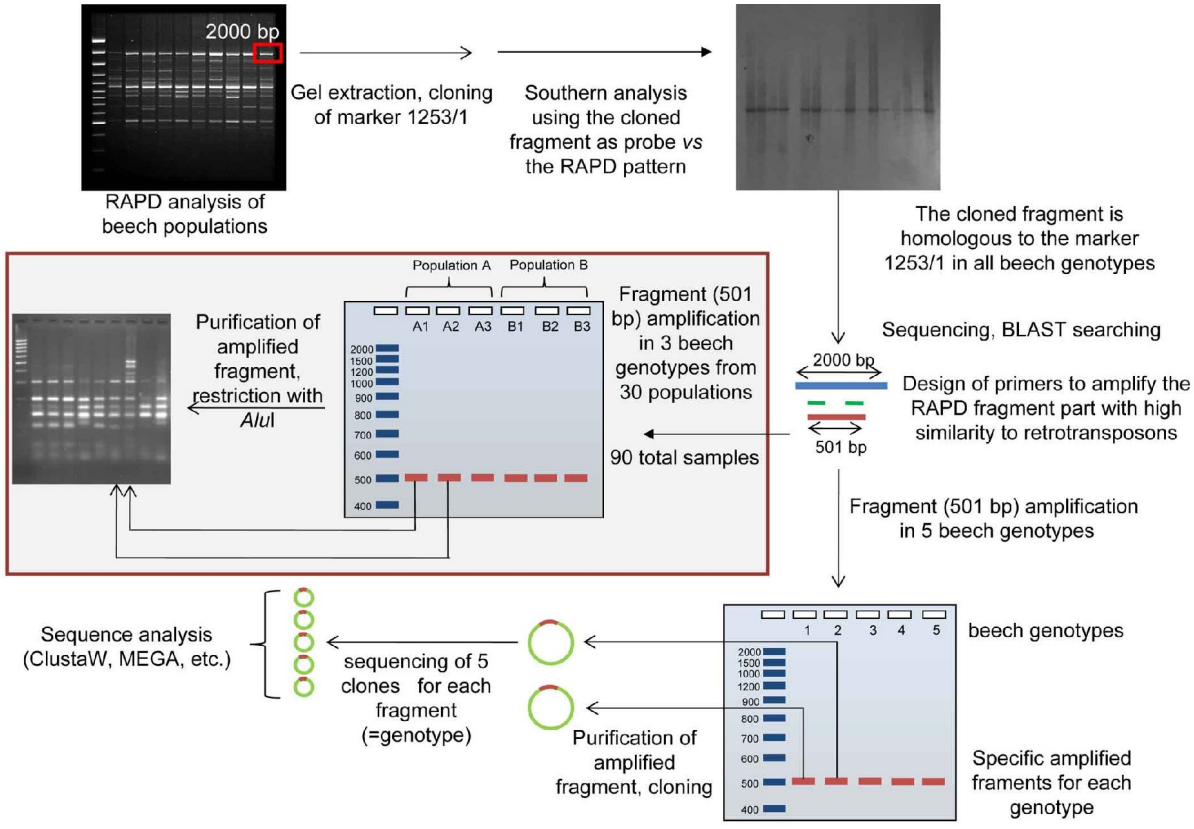

Fig. 3 - Experimental flow chart of beech retrotransposable elements characterization. The grey shaded boxed part is valid only for the non-LTR element.

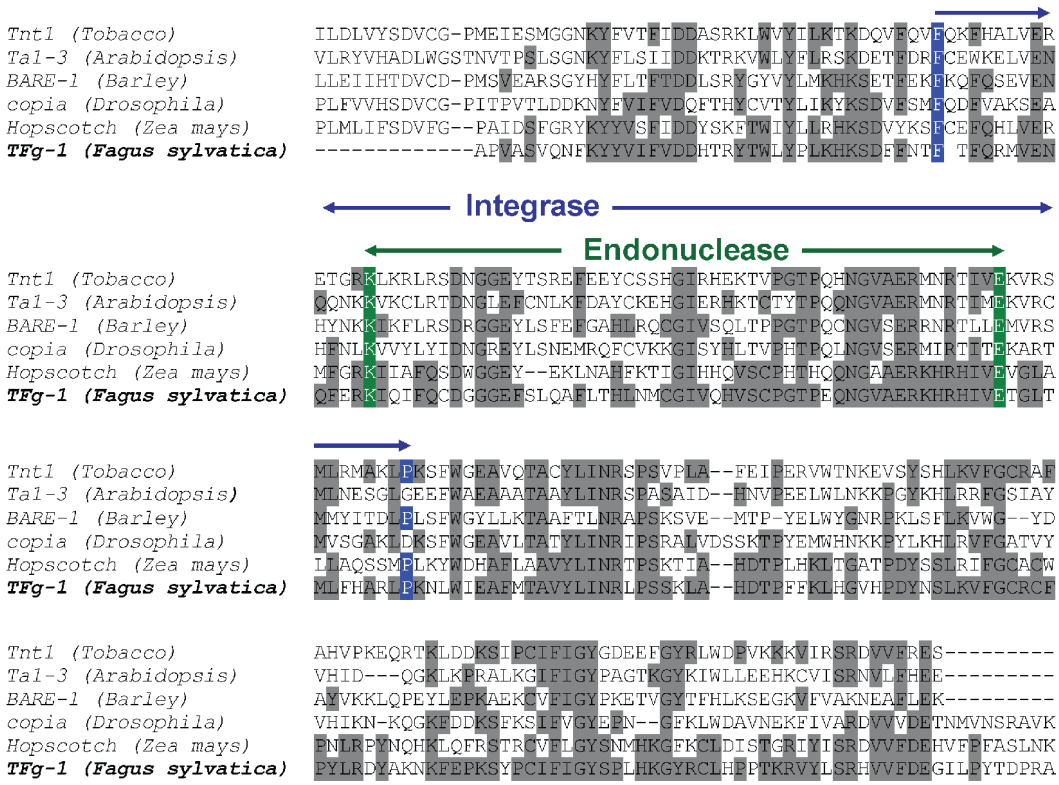

Fig. 4 - Molecular identification of beech LTR retroelement. ClustalW alignment of the original sequence form *F. sylvatica* and a selection of best hit homologous sequences obtained by database searching. The integrase and endonuclease domain of the *pol* gene are highlighted by arrows. Grey shading of aminoacids identifies site identity.

Primer3 on-line tool (Rozen & Skaletsky 2000) to amplify a region (fragments made up of 357 and 501 base pair, respectively, for the LTR and non-LTR elements) for each retrotransposon identified. Primers for the non-LTR Retrotransposable elements were: 5' - CGATTTGTCCTACCTGCTTG-3', 5'-CCGCCCTATTAGTTTGATTGG - 3'; primers for the LTR Element were: 5' - AACCGACATGGGTAACTTCG - 3', 5' - CCA CACAGTACTCGTTCAAATCT - 3'. Both fragments were amplified using a sequencing PCR cycle with an annealing temperature of 60 °C. Reactions were performed in a GeneAmp PCR System 9700 (Perkin Elmer) with 1.5 mM $MgCl_2$, 200 µM for each dNTP, 1 µM for each primer and 1 U of TaqDNA polymerase (Invitrogen - final volume 20 µl).

In order to preliminarily evaluate the diversity of retrotransposable elements in different genotypes, specific amplification fragments of a randomly selected tree from 5 beech populations (belonging to the dataset used in Emiliani et al. 2004 and selected for the high genetic diversity obtained with RAPD markers) were produced for both retroelements, using the primers and the amplification conditions reported above. The two specific amplicons of the retrotransposons were purified with the Qiagen PCR Purification Kit, cloned, sequenced and analysed as already reported. 5 clones for each genotype (for a total of 25 clones) and for each retrotransposon were sequenced (50 sequences in total - Fig. 3).

CAPs analysis of non-LTR element polymorphism

For CAPs analysis, the non-LTR reverse transcriptase fragments (501 base pair) from 3 randomly selected individuals for each of the 30 beech populations analysed by Emiliani et al. (2004) - 90 total samples - were amplified using the primers reported above. The amplification products were purified using the Qiagen PCR Purification Kit following the manufacturer's specification and digested for 4 hours with 10 U of *Alu*I restriction endonuclease (Takara, Japan). Digestion patterns were loaded on 4% (w/v) agarose gels and analysed with the PhotoCapt software (Vilber-Lourmat, France).

The whole experimental flowchart is reported as Fig. 3.

Results and Discussion

Molecular identification of retrotransposable elements in Fagus sylvatica

The SCAR strategy applied to RAPD markers allowed the molecular identification of retrotransposable elements in beech. Homology searching in databases confirmed that the cloned and sequenced RAPD marker RF2/35 is homologous to a class I LTR retroelement. As reported in Fig. 4 and Fig. 5, the identified fragment contains the integrase and endonuclease domains of the *pol* polyprotein coding gene of a LTR retrotransposon. The homology obtained between the original beech sequence and the BLAST best-hit sequences also allowed ascribing the identified retrotransposon to the *copia*-like type. The characterized fragment contains only coding regions, suggesting a possible integrity of the identified retrotransposable element.

On the contrary only a part (501 bp) of the RAPD marker 1253/1 showed high homology to characterized retrotransposons, suggesting that ORFs fragmentation occurred for this element in beech, as often reported for truncated LINE elements (Nikaido & Okada 2000). Nevertheless, the identified fragment contains three domains of the LINE non-LTR retrotransposable element reverse transcriptase (Fig. 5, Fig. 6).

The original sequences of the two identified retrotransposable elements are deposited in GeneBank under the accession number AF405557.1 and AF405555.1.

The identification of partial sequences of retrotransposable elements in broadleaf forest tree species can be useful for studies of genome evolution and for paving the way to new diversity screenings approaches already implemented in crop or model species (Vitte & Bennetzen 2006, Morgante 2006).

Polymorphism of beech retrotransposable elements

The analysis of the polymorphism of beech retrotransposable elements was carried out with two strategies: i) sequence analysis of cloned fragments and ii) CAPs analysis of amplified fragments (Fig. 3).

Through the cloning and sequencing of

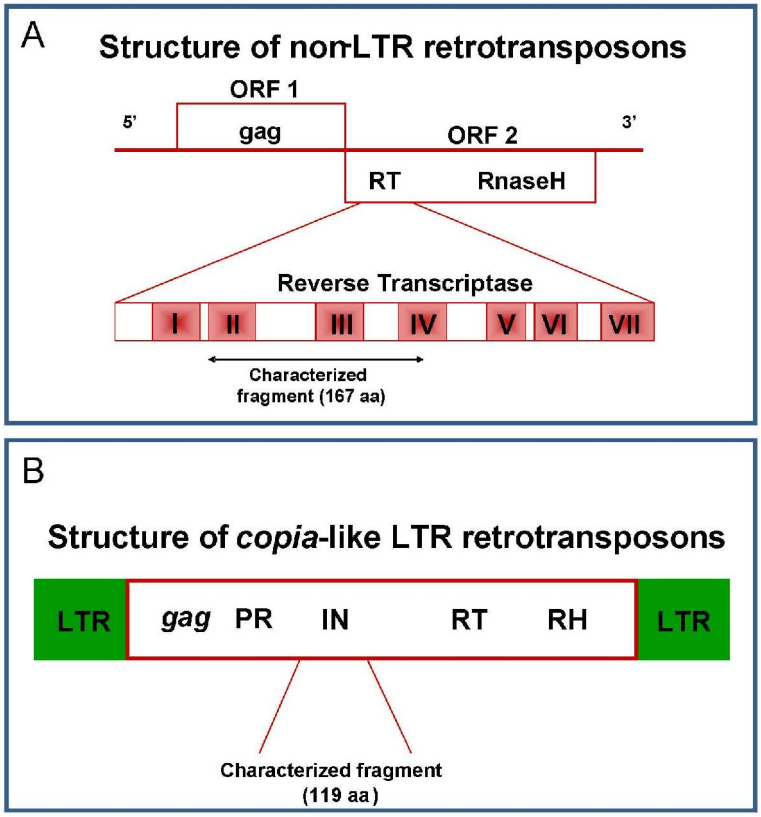

Fig. 5 - Schematic representation of non-LTR LINE (A) and *copia*-like LTR (B) retrotransposable elements. Characterized fragments are highlighted.

amplification fragment obtained with specific primers, the characterization of the retroelement diversity in different beech genotypes was possible. As reported in Tab. 1, the LTR retrotransposable element resulted polymorphic in each genotype. The level of diversity is noticeable, showing, at nucleotide level, 22 (out of 357) polymorphic sites, with 12 missense mutations. It is important to stress that all the sequenced LTR retrotransposons maintain a putatively functional protein sequence (at least for the characterized fragment) suggesting a possible active status of this class of retrotransposons in beech, as often reported for this type of retrotransposable elements in many species (Vitte & Panaud 2005). The sequences polymorphism of the 5 different clones from the same genotypes (data not shown) also allows concluding that this element forms a multi-copy family inside beech genome in agreement with the viral-quasi species theory of the evolution of retrotransposable elements (Casacuberta et al. 1997).

The experimental approach reported above also allowed determining the polymorphism of the non-LTR element. The diversity of the non-LTR retrotransposon sequences from different beech genotypes (see Supplementary Material 1) is distinctly higher than that

Fig. 6 - Molecular identification of beech non-LTR LINE retrotransposable elements. CLUSTALW alignment of the original sequence from *F. sylvatica* and a selection of best hit homologous sequences obtained by database searching. The three domains of the identified reverse transcriptase are highlighted by boxes. Grey shading of aminoacids identifies site identity.

Tab. 1 - Nucleotide and deduced protein sequence polymorphisms shown by the LTR retrotransposable elements belonging to different *F. sylvatica* genotypes. The nucleotide variations responsible of the missense mutations is highlighted in block capitals and the correspondent aminoacid variation is reported in the same column (lowermost part of the table). Dot represents site identity. Beech genotypes are reported with the name of the geographical provenance.

***F. sylvatica* genotypes**	**Nucleotide sequence variable site number (total sequence length 357 nucleotides)**																					
	39	**44**	**48**	**55**	**77**	**84**	**87**	**115**	**158**	**177**	**181**	**201**	**202**	**210**	**211**	**223**	**225**	**230**	**233**	**252**	**266**	**318**
	Variable nucleotides																					
Foresta Umbra	c	T	a	t	G	g	c	G	G	t	C	c	G	c	A	T	g	T	T	G	G	t
Mt. Etna	t	A	c	c	.	.	a	A	.	a	.	t	.	.	G	C	.	.	.	.	.	c
Madonie	.	.	.	c	A	a	.	.	.	c	.	.	.	.	G	C	a	.	.	.	.	.
Sila	t	.	c	c	.	.	a	.	A	.	G	t	A	t	G	C	.	C	A	T	A	c
Irpinia	.	.	.	c	A	a	.	.	.	c	.	.	.	.	G	C	a	.	.	.	.	.
	Protein sequence variable site number (total sequence length 119 aa)																					
	15			**26**			**39**	**53**		**61**		**68**		**71**	**75**		**77**	**78**	**84**	**89**		
	Variable aminoacids																					
Foresta Umbra	L			G			E	G		P		G		K	W		I	V	M	R		
Mt. Etna	H			.			K	.		.		.		E	R		.	.	.	.		
Madonie	.			E			.	.		.		.		E	R		.	.	.	.		
Sila	.			.			.	D		A		E		E	R		T	D	I	Q		
Irpinia	.			E			.	.		.		.		E	R		.	.	.	.		

the one reported for the LTR element, with 99 polymorphic sites out of 501 nucleotides. In this case the presence of nonsense mutation, the reported level of diversity and the abrupt ending of the coding region of the characterized fragment suggest that the identified class I TE is inactive for ORFs fragmentation and subjected to genetic drift (Dewannieux & Heidmann 2005).

The polymorphism shown by the reverse transcriptase fragment of the LINE element suggested the development of a Cleaved Amplified Polymorphism (CAP) approach in order to study the genetic diversity of the element itself and, considering its putative neutrality, the diversity of beech genotypes and populations. The restriction analysis (Tab. 2) pointed out that the specific amplicon of the reverse transcriptase constitutes, for each genotype, a family of fragments, thus demonstrating, in accordance with the cloned sequence analysis, that also the LINE element represents a multi-copy quasi-species inside each beech genome. Using on 90 beech genotypes the simple and time saving CAP approach with only one restriction enzyme, 12 alleles (*sensu lato*) were identified (Tab. 2). Each beech individual showed at least and simultaneously 4 alleles. The CAP pattern allowed recognizing 6 alleles already expected from the results of the *in silico* restriction analysis of the sequenced fragment, demonstrating the reproducibility of the approach, and 6 alleles not previously described. This also suggests that the polymorphism identified is an underestimation of the real diversity of the retrotransposon quasi-species, since the cloning strategy did not allow characterizing all the alleles in each genotype.

From the bioinformatic analysis of the available sequences is also easy to predict that the use of different restriction enzymes could produce an impressive number of markers in an easy and time-saving way. The approach proposed here represents a new method of exploiting the diversity of retrotransposable elements, for genetic variability screening which can complement already developed strategies as the S-SAP (Waugh et al. 1997), the IRAP and the REMAP (Kalendar et al. 1999) approaches.

Tab. 2 - Results of the Cleaved Amplified Polymorphism analysis on the non-LTR retrotransposable elements amplified fragment (total fragment length: 501 bp = 167 aa). Among the obtained alleles, 6 were already expected from *in silico* restriction analysis of sequenced fragments.

Allele	**Restriction fragment length (bp)**						**Comments**
1	135	300	26	7	34		Expected by sequences restriction analysis and obtained with CAPs approach
2	200	100	135	14	53		
3	200	100	33	102	14	53	
4	135	164	135	14	19	34	
5	135	201	98	26	7	34	
6	135	164	135	33	34		
7	448	19	34				Obtained with CAPs approach
8	200	135	114	53			
9	120	15	300	26	7	34	
10	135	164	147	19	34		
11	230	70	33	100	14	53	
12	135	333	34				

Conclusions

In the present work the identification of retrotransposable elements in the genome of the broadleaved forest tree species *Fagus sylvatica* is reported. Since retrotransposons are fundamental in the evolution of plants genome, gene duplication, transcription regulation and epigenetic effects, the identification of these peculiar genetic elements in beech can help the understanding of such processes in non model species, along with the increasing availability of woody perennial species genomes such as poplar (Tuskan et al. 2006) and grape (Jaillon et al. 2007). The results of the molecular characterization confirm that in beech retrotransposable elements build up multi-copy families that can be used for molecular fingerprinting.

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25: 3389-3402. - doi: 10.1093/nar/25.17.3389

Asins MJ, Monforte AJ, Mestre PF, Carbonel EA (1999). *Citrus* and *Prunus copia*-like retrotransposons. Theoretical and Applied Genetics 99: 503-510. - doi: 10.1007/s001220051263

Biémont C, Vieira C (2005). What transposable elements tell us about genome organization and evolution: the case of *Drosophila*. Cytogenetic and Genome Research 110: 25-34. - doi: 10.1159/000084935

Casacuberta JM, Santiago N (2003). Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene 311: 1-11. - doi: 10.1016/S0378-1119(03)00557-2

Casacuberta JM, Vernhettes S, Audeon C, Grandbastien MA (1997). Quasi-species in retrotransposons: a role for sequence variability in Tnt1 evolution. Genetica 100: 109-117. - doi: 10.1023/A:1018309007841

Craig NL, Craigie R, Gellert M, Lambowitz AM (2002). Mobile DNA II. American Society for Microbiology Press, Washington, DC, USA.

Dewannieux M, Heidmann T (2005). LINEs, SINEs and processed pseudogenes: parasitic strategies for genome modeling. Cytogenetic and Genome Research 110: 35-48. - doi: 10.1159/000084936

Durbin ML, Denton AL, Clegg MT (2001). Dynamics of mobile element activity in chalcone synthase loci in the common morning glory (*Ipomea purpurea*). PNAS 98 (9): 5084-5089. - doi: 10.1073/pnas.091095498

Emiliani G, Paffetti D, Vettori C, Giannini R (2004). Geographic distribution of genetic variability of *Fagus sylvatica* L. in southern Italian populations. Forest Genetics 11: 231-237.

Feschotte C, Jiang N, Wessler SR (2002). Plant transposable elements: where genetics meets genomics. Nature Reviews Genetics 3: 329-341. - doi: 10.1038/nrg793

Flowers JM, Purugganan MD (2008). The evolution of plant genomes scaling up from a population perspective. Current Opinion in Genetics & Development 18: 565-570. - doi: 10.1016/j.gde.2008.11.005

Friesen N, Brandes A, Heslop-Harrison JS (2001). Diversity, origin, and distribution of retrotransposons (*gypsy* and *copia*) in conifers. Molecular Biology and Evolution 18: 1176-1188. [online] URL: http://mbe.oxfordjournals.org/cgi/content/abstract/18/7/1176

Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, Choisne N, et al. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm *phyla*. Nature 449: 463-468. - doi: 10.1038/nature06148

Kalendar R, Grob, Regina M, Suoniemi A, Schulman A (1999). IRAP and REMAP: two new retrotransposon-based DNA fingerprinting techniques. Theoretical and Applied Genetics 98: 704-711. - doi: 10.1007/s001220051124

Kalendar R, Schulman AH (2006). IRAP and REMAP for retrotransposon-based genotyping and fingerprinting. Nature Protocols 1(5): 2478-2484. - doi: 10.1038/nprot.2006.377

Kamm A, Doudrick RL, Heslop-Harrison JS, Schmidt T (1996). The genomic and physical organization of Ty1-*copia*-like sequences as a component of large genomes in *Pinus elliottii* var. *elliottii* and other gymnosperms. PNAS 93: 2708-2713. - doi: 10.1073/pnas.93.7.2708

Kapitonov VV, Jurka J (2003). Molecular paleontology of transposable elements in the *Drosophila melanogaster* genome. PNAS 100 (11): 6569-6574. - doi: 10.1073/pnas.0732024100

Kazazian HH (2004). Mobile elements: drivers of genome evolution. Science 303: 1626-1632. - doi: 10.1126/science.1089670

L'Homme Y, Séguin A, Tremblay FM (2000). Different classes of retrotransposons in coniferous spruce species. Genome 43: 1084-1089. - doi: 10.1139/gen-43-6-1084

Larkin M, Blackshields G. Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007). ClustalW and ClustalX version 2. Bioinformatics 23 (21): 2947-2948. - doi: 10.1093/bioinformatics/btm404

Leeton PR, Smyth DR (1993). An abundant LINE-like element amplified in the genome of *Lilium speciosum*. Molecular and General Genetics 237: 97-104. - doi: 10.1007/BF00282789

Leprince AS, Grandbastien MA, Meyer C (2001). Retrotransposons of the Tnt1B family are mobile in *Nicotiana plumbaginifolia* and can induce alternative splicing of the host gene upon insertion. Plant Molecular Biology 47: 533-541. - doi: 10.1023/A:1011846910918

Meyers BC, Tingey SV, Morgante M (2001). Abundance, distribution, and transcriptional activity of repetitive elements in the maize genome. Genome Research 11: 1660-1676. - doi: 10.1101/gr.188201

Morgante M (2006). Plant genome organization and diversity: the year of the junk! Current Opinion in Biotechnology 17: 168-173. - doi: 10.1016/j.copbio.2006.03.001

Nikaido M, Okada N (2000). CetSINEs and AREs are not SINEs but are parts of cetartiodactyl L1. Mammalian Genome 11: 1123-11236. - doi: 10.1007/s003350010221

Paffetti D, Scotti C, Gnocchi S, Fancelli S, Bazzicalupo M (1996). Genetic diversity of an Italian *Rhizobium meliloti* population from different *Medicago sativa* varieties. Applied Environmental Microbiology 62: 2279-2285.

Queen RA, Gribbon BM, James C, Jack P, Flavell AJ (2004). Retrotransposon based molecular markers for linkage and genetic diversity analysis in wheat. Molecular Genetics and Genomics 271: 91-97. - doi: 10.1007/s00438-003-0960-x

Rozen S, Skaletsky HJ (2000). Primer3 on the WWW for general users and for biologist programmers. In: "Bioinformatics methods and protocols: methods in molecular biology" (Krawetz S, Misener S eds). Humana Press, Totowa, NJ, USA, pp. 365-386.

SanMiguel P, Tikhonov A, Jin Y-K, Motchoulskaia N, Zakharov D, Melake-Berhan A, Springer PS, Edwards KJ, Lee M, Avramova Z, Bennetzen JL (1996). Nested retrotransposons in the intergenic regions of the maize genome. Science 274: 765-768. - doi: 10.1126/science.274.5288.765

Smykal P (2006). Development of an efficient retrotransposon-based fingerprinting method for rapid pea variety identification. Journal of Applied Genetics 47: 221-230. [online] URL: http://jay.au.poznan.pl/JAG/pdfy/2006_Volume_47/2006_Volume_47_3-221-230.pdf

Tam SM, Mhiri C, Vogelaar A, Kerkveld M, Pearce SR Grandbastien M-A (2005). Comparative analyses of genetic diversities within tomato and pepper collections detected by retrotransposon-based SSAP, AFLP and SSR. Theoretical and Applied Genetics 110: 819-831. - doi: 10.1007/s00122-004-1837-z

Tamura K, Dudley J, Nei M, Kumar S (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24: 1596-1599. - doi: 10.1093/molbev/msm092

TIHGSC (2001). Initial sequencing and analysis of the human genome. The International Human Genome Sequencing Consortium. Nature 409 (6822): 860-921. - doi: 10.1038/35057062

Tuskan GA, DiFazio S, Jansson S, Bohlmann J, GrigorievI, Hellsten U, et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). Science 313: 1596-1604. - doi: 10.1126/science.1128691

Vance V, Vaucheret H (2002). RNA silencing in plants - defense and counterdefense. Science 292: 2277-2280. - doi: 10.1126/science.1061334

Venturi S, Dondini L, Donini P, Sansavini (2006). Retrotransposons characterization and fingerprinting of apple clones by S-SAP markers. Theoretical and Applied Genetics 112: 440-444. - doi: 10.1007/s00122-005-0143-8

Vicient CM, Suoniemi A, Anamthawat-Jonsson K, Tanskanen J, Beharav A, Nevo E, Schulman AH (1999). Retrotransposon BARE-1 and its role in genome evolution in the genus *Hordeum*. Plant Cell 11: 1769-1784. - doi: 10.1105/tpc.11.9.1769

Vitte C, Bennetzen JL (2006). Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. PNAS 103 (47): 17638-17643. - doi: 10.1073/pnas.0605618103

Vitte C, Panaud O (2005). LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenetic and Genome Research 110: 91-107. - doi: 10.1159/000084941

Vitte C, Panaud O, Quesneville H (2007). LTR retrotransposons in rice (*Oryza sativa*, L.): recent burst amplifications followed by rapid DNA loss. BMC Genomics 8: 218. - doi: 10.1186/1471-2164-8-218

Waugh R, McLean K, Flavell AJ, Pearce SR, Ku-

mar A, Thomas BBT, Powell W (1997). Genetic distribution of Bare-1-like retrotransposable elements in the barley genome revealed by sequence-specific amplification polymorphisms (S-SAP). Molecular and General Genetics 253: 687-694. - doi: 10.1007/s004380050372

Wessler SR (1998). Transposable elements associated with normal plant genes. Physiologia Plantarum 103: 581-586. - doi: 10.1034/j.1399-3054.1998.1030418.x

White SE, Habera LF, Wessler SR (1994). Retrotransposon in the flanking regions of normal plant genes: a role for copia-like elements in the evolution of gene structure and expression. PNAS 91: 11792-11796. - doi: 10.1073/pnas.91.25.11792

Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007). A unified classification system for eukaryotic transposable elements. Nature Reviews Genetics 8: 973-982. - doi: 10.1038/nrg2165

Xiao H, Jiang N, Schaffner E, Stockinger EJ, van der Knaap E (2008). A retrotransposon-mediated gene duplication underlies morphological variation of tomato fruit. Science319: 1527-1530. - doi: 10.1126/science.1153040

Yoshioka Y, Matsumoto S, Kojima S, Ohshima K, Okada N, Machida Y (1993). Molecular characterization of a short interspersed repetitive element from tobacco that exhibits sequence homology to specific tRNAs. PNAS 90: 6562-6566. - doi: 10.1073/pnas.90.14.6562