Mutations in retrotransposon AtCOPIA4 compromises resistance to Hyaloperonospora parasitica in Arabidopsis thaliana

Yi-Hong Wang1,2 and James T. Warren Jr.1

1School of Science, Behrend College, Penn State University, Erie, PA, USA.
2Department of Renewable Resources, University of Louisiana, Lafayette, LA, USA.

Abstract

Retrotransposons (RTEs) are a principal component of most eukaryotic genomes, representing more than 50% of the human genome (Kazazian, 2004; Lander et al., 2001) and 50%-80% of some grass genomes (Feschotte et al., 2002; SanMiguel and Bennetzen, 1998). Even in the compact genome of Arabidopsis thaliana, they account for 5.5% of the sequenced genome (Kazazian, 2004). Cellular functions of RTEs have been reported. They seem to play a role in the proliferation of cancer cells (Oricchio et al., 2007), globally interfere with the regulatory network of transcription factors p53 (Wang et al., 2007) and PSF (Song et al., 2005), and interact with the dynein light-chain which is a known component of the dynein microtubule motor (Havecker et al., 2005). In mammals, RTEs are more likely to be found in rapidly evolving gene clusters, such as those involved in defense and response to external signals, than in mRNAs of highly conserved genes involved in development, transcription, replication, cell structure and metabolism (Medstrand et al., 2005; van de Lagemaat et al., 2003). In plants, the pattern is similar. For example, Tos17 retrotransposon is preferably inserted into disease/defense-related and signal transduction (kinase) genes in the rice genome (Miyao et al., 2003). Furthermore, RTEs have been identified in disease resistance gene clusters in lettuces (Meyers et al., 1998; Michelmore and Meyers, 1998), rice (Song et al., 1995), barley (Marcel et al., 2007), the common bean (Vallejos et al., 2006), poplar (Lescot et al., 2004), and Arabidopsis (van der Biezen et al., 2002; Yi and Richards, 2007). Various RTEs have been shown to be induced by plant pathogens or elicitors in rice (Chen et al., 2007; Vergne et al., 2008), by Fusarium oxysporum in chickpea (Nimalkara et al., 2006), and by fungal elicitors in tobacco (Poueteau et al., 1994; Melayah et al., 2001). In addition, RTE Tnt1A inserted in a tobacco resistance gene cluster has been shown to drive partial transcription of the neighboring disease resistance gene TNLL1 (Hernández-Pinzón et al., 2009).

RTE coding sequences are also known to form chimeric transcripts (Kashkush et al., 2003; Peaston et al., 2004) with non-RTE mRNA sequences and chimeric transcripts displaying a different expression pattern from that of the original transcripts (Peaston et al., 2004). Chimeric resistance and retrotransposon genes may function in disease resistance. For example, L10 is a Toll/Interleukin1 receptor-nucleotide binding site-leucine-rich repeat [TIR-NBS-LRR] class of resistance gene (Lawrence et al., 1995) and a chimera of the L10 TIR domain fused with a partial tobacco retrotransposon sequence at the 3’ end has been reported. Expression of this chimera caused the same stunted phenotype produced by over-expressing full-length L6, and increased transcript abundance of a constitutive defense pro-
Hyaloperonospora parasitica (2007) reduces resistance to the downy mildew pathogen retrotransposon, 1998). et al. thus determined and indicated in Figure 1. One heterozygous plants. Position of T-DNA insertion was and sequenced to determine the exact T-DNA insertion site single PCR product from LB and RP primers was amplified 58 °C/30 s and 72 °C/2 min., followed by 72 °C for 5 min. A naturing at 94 °C for 2 min and 35 cycles of 94 °C/30 s, 58 °C/30 s, 72 °C/2 min, and the final extension of 5 min at 72 °C. All RT-PCR primers were tested for their target specificity using Seqviewer (www.arabidopsis.org). All the primers used showed desired specificity:
P1: GTAGATGTTGCCAAAACAGGTTCCTC P2: AATCCACATTGTGCCCTCCTCTT P3: TTAAGACGAAGACCTTGGAGATGGC P4: GAGGACAACAGGAGGATCAGAAA P5: TTGTTGCTCAAGGGAGAACTAAAG P6: ATGAAAACATCCGAAACAGGCAAGT UBQ1: GATCTTTGCGGAAAACATTTGGAGATGGT UBQ2: CGACCTGTCAATTTGAGAAGAGATACACAGG

To conduct pathogenicity essay, seeds were planted in soils (Metromix 360, SunGro, Canada) saturated with water and stratified at 4 °C for 48 h. Pathogenicity assays followed those described previously (Holub et al., 1994; Yoshioka et al., 2006). Briefly, 10 to 14 day-old conidio-

Figure 1 - The Arabidopsis genomic region of AtCOPIA4 (in green) and RPP4 (in red) based on Yi and Richards (2007) who have sequenced the full-length cDNA of genes in this region. Open boxes represent exons and lines between boxes represent introns in RPP4. AtCOPIA4 conserved domains are indicated above the gene. Location of T-DNA insertion is indicated for SALK_005767. Antisense sequence represented by a black line below was used for an AtCOPIA4 antisense construct. One cDNA match (AYBLZ22TR) to AtCOPIA4 is also shown. Chimeric cDNAs are drawn in red broken lines and arrows (RAFL21-45-F24 and BX842341). Two chimeric ESTs were also identified in GenBank: ES444452 and EL142415 (not shown). Affymetrix GeneChip probes for both genes are shown in blue arrows. Brown open arrows below the ends of AtCOPIA4 are the 130 bp long terminal repeats (LTRs; 9488607-9488478 and 9483894-9483755, respectively).
phores of *H. parasitica* isolate EMWA1 (kindly provided by Daniel Klessig) were collected from susceptible live plants of Nd-0 and re-suspended in cold, sterile water. The spores were vortexed for 30 s for release from the sporangia. Spore concentration was adjusted to $10^4$-$10^6$ per mL, and 1-2 μL of the spore suspension was dropped onto each cotyledon of 6 to 7 day-old plants (10 to 20 plants for each line in each replicate). The inoculated plants were covered with plastic wrap and incubated at 16 °C with 10 hour-photoperiods. At 10 to 14 days after inoculation, the number of conidiophores on each cotyledon leaf, number of cotyledon leaves with conidiophores and the total number of plants, were recorded using a dissection microscope. The experiment was replicated three times with similar trends. Both resistant (Col-0) and susceptible (Nd-0; Holub *et al.*, 1994) lines were used as pathogenicity assay controls, although only Col-0 data are shown in Figure 2 and Figure 3.

To identify Arabidopsis retrotransposons that had acquired cellular functions, we searched the genome sequences of about 1,600 annotated retrotransposon genes curated in VirtualPlants (virtualplant.bio.nyu.edu; www.virtualplant.org) for matches to ESTs. Among these, 56 had exact matches to EST sequences and intact open reading frames. Twenty two of the genes were represented in the Affymetrix Arabidopsis ATH1 GeneChip and expression of those was searched in over 3,000 GeneChips in the Genevestigator database (Zimmermann *et al.*, 2004). *AtCOPIA4* was selected because it represents a typical retrotransposon which encodes gag, integrase and reverse transcriptase proteins (Feschotte *et al.*, 2002), and it is the most highly expressed retrotransposon throughout the development stages, although generally their expression level is low due to regulation by the host. *AtCOPIA4* transcript level was found to be highest in developing leaves and flowers (Table 1). This expression pattern was also confirmed by the Massively Parallel Signature Sequencing (MPSS) mRNA signature data (Nakano *et al.*, 2006). Genes in this region have been shown to be co-expressed, probably due to local chromatin structural changes (Yi and Rich-

---

**Figure 2** - RT-PCR of Arabidopsis T-DNA insertion and antisense mutants in *AtCOPIA4*. Primers used for *AtCOPIA4* are P5 and P6, P1 and P6 for the chimeric transcript as shown in Figure 1 and P3 and P4 for RPP4. Total RNA from seedlings was used. Lines used are: Col-0-Columbia wild type; I3, I10-heterozygous and homozygous T-DNA insertion lines, respectively; A60, and A80 are antisense lines. Chimera indicates RPP4-*AtCOPIA4* chimeric mRNA.

---

**Figure 3** - Percentage of infected Arabidopsis wild type and mutant plants 10 days after inoculation of *Hyaloperonospora parasitica* EMWA1. (A) Percent of infected plants. (B) Percent of infected cotyledon leaves. (C) and (D) show an infected I10 and an uninfected wild type plant, respectively. Lines tested are: Col-0-Columbia wild type; I3, I10-heterozygous and homozygous T-DNA insertion lines, respectively; A60, and A80 are antisense lines.
ards, 2007; Zhan et al., 2006). Cluster analysis in Genevestigator also revealed that AtCOPIA4, RPP4 and At4g16880, which are three adjacent genes on chromosome 4 based on the latest genome annotation release (TIGR/AGI V8) and Yi and Richards (2007), were coexpressed under salt, cold, heat, wound, oxidative, and genotoxic conditions (data not shown). Correlation of AtCOPIA4 expression is 0.59 with At4g16880 and 0.55 with RPP4, as calculated in the ATTED-II database of Arabidopsis microarray data (Obayashi et al., 2007). This correlation between AtCOPIA4 and RPP4 is noticeable in Table 1 as well.

To elucidate the function of AtCOPIA4, a homoyzogous T-DNA insertion mutant was identified from SALK_005767 and antisense RNAi mutants were generated, as described above. Sequencing analysis indicated that T-DNA was inserted 70 bp before the start codon of AtCOPIA4 and 117 bp after the stop codon of RPP4 in SALK_005767 (Figure 1) in the LTR. RT-PCR analysis of the mutant seedlings indicates that the AtCOPIA4 transcript was undetectable in the homozygous T-DNA insertional mutant (I10) but present in the heterozygote (I3; Figure 2), indicating that transcription of AtCOPIA4 had been knocked out in the T-DNA insertion mutant. Among the two antisense mutants tested (A60, and A80), AtCOPIA4 transcript levels were undetectable in A60 and significantly reduced in A80 (Figure 2). In the mutants with no or reduced AtCOPIA4 transcript, the level of the AtCOPIA4-RPP4 chimeric transcript was also either not apparent or was at a reduced level (Figure 2). However, the abundance of RPP4 transcript was not affected in these lines, when compared to Col-0 and based on RT-PCR analysis, using primers P3 and P4, as shown in Figure 1 (Figure 2).

No noticeable morphological difference was observed between the mutants and Col-0. However, because AtCOPIA4 is located in the cluster of RPP5 class of resistance genes (van der Biezen et al., 2002; Yi and Richards, 2007), right next to RPP4 and in silico EST analysis had revealed a chimeric AtCOPIA4-RPP4 mRNA (Figure 1), we sought to evaluate the mutants for resistance to H. parasitica isolate EMWA1. Pathogenicity assays showed that on average, homozygous insertional and antisense mutants were 2 to 4 times as likely to be infected by the isolate based on percentage of infected plants, whereas heterozygous insertional mutants were as resistant to the isolate as Col-0 (Figure 3). Notwithstanding, the number of conidiophores was not significantly higher in the mutants examined 10 days after inoculation, when compared to wild type Col-0. Overall, the number of conidiophores ranges from 2 to 5 per cotyledon leaf on average for all lines and replicates and the highest number of conidiophores was 15 found in the mutants.

Table 1 - Average signal intensity of selected retrotransposon genes in different developmental stages in Arabidopsis.

| Gene     | Germinated seed | Seeding | Young rosette | Developed rosette | Bolting | Young flower | Developed flower | Flowers and siliques | Siliques |
|----------|-----------------|---------|--------------|-------------------|---------|--------------|-------------------|----------------------|---------|
| AtCOPIA4 (At4g16870) | 310       | 619     | 646         | 741               | 503     | 681          | 699               | 450                  | 460     |
| RPP4 (At4g16860)      | 80        | 432     | 1,053       | 978               | 380     | 1,291        | 771               | 332                  | 536     |
| At3g21020             | 295       | 166     | 154         | 170               | 314     | 216          | 268               | 205                  | 280     |
| At2g15510             | 125       | 145     | 134         | 147               | 112     | 193          | 195               | 124                  | 201     |
| At2g17490             | 23        | 20      | 17          | 18                | 43      | 19           | 27                | 14                   | 20      |
| ACT2 (At3g18780)      | 14,828    | 18,867  | 15,468      | 16,580            | 13,033  | 14,243       | 13,732            | 16,403               | 5,333   |
| Total arrays          | 169       | 944     | 419         | 173               | 150     | 277          | 619               | 121                  | 57      |

*Only four of the 22 RTEs are presented in the table. The RTEs are randomly selected (except AtCOPIA4) to show that AtCOPIA4 has the highest transcript abundance. RPP4 is included as a comparison for its expression pattern with that of AtCOPIA4. Actin 2 (ACT2) is included as a control. Total number of arrays (GeneChips) used to obtain the averaged signal for each stage. Data are gathered from the Genevestigator database (https://www.genevestigator.ethz.ch).
formed at the RNA level (Figure 1). Therefore, it will be of interest to see whether increasing RPP4-AtCOPIA4 chimera expression would boost resistance as conferred by RPP4 because the level of the chimerical transcript was much lower in the mutants tested (Figure 2). AtCOPIA4 expression is driven by the 130 bp LTRs flanking the coding region. Future studies should focus on how the chimeric transcript is generated with the AtCOPIA4 sequence downstream from the RPP4 TIR domain, in contrast to what has been reported in other cases.

We have shown here that knockout of an RTE compromises plant resistance to the downy mildew pathogen Hyaloperonospora parasitica EMWA1. RTEs have been shown to play a role in defense response in other eukaryotes as well. In mammals, degraded reverse transcribed RTEs can trigger defense response from the immune system (Stetson et al. 2008). Our evidence suggests that RTEs also function in defense response in plants.

Acknowledgments

We thank Ralph A. Dean of North Carolina State University, Daniel F. Voytas of the University of Minnesota, Daniel F. Klessig and Hong-Gu Kang of Boyce Thompson Institute and an anonymous reviewer for suggestions to improve the manuscript. Hyaloperonospora parasitica isolate EMWA1 was kindly provided by Daniel F. Klessig. The project was initiated at Penn State Erie. Funding for the project was provided by Penn State Erie-The Behrend College and the University of Louisiana at Lafayette.

References

Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al. (2003) Genomewide insertional mutagenesis of Arabidopsis thaliana. Science 301:653-657.

Chen F, Li Q and He Z (2007) Proteomic analysis of rice plasma membrane-associated proteins in response to chitoooligosaccharide elicitors. J Integr Plant Biol 49:863-870.

Clough SJ and Bent AF (1998) Floral dip: A simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant J 16:735-743.

Ellis JG, Lawrence GJ, Luck JE and Dodds PN (1999) Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. Plant Cell 11:495-506.

Feschotte C, Jiang N and Wessler SR (2002) Plant transposable elements: Where genetics meets genomics. Nat Rev Genet 3:329-341.

Frost D, Way H, Howles P, Luck J, Manners J, Hardham A, Finnegan J and Ellis J (2004) Tobacco transgenic for the flax rust resistance gene L expresses allele-specific activation of defense responses. Mol Plant Microbe Interact 17:224-232.

Havecker ER, Gao X and Voytas DF (2005) The Sireviruses, a plant-specific lineage of the Ty1/copia retrotransposons, interact with a family of proteins related to dynein light chain 8. Plant Physiol 139:857-868.
scripts in roots of resistant and susceptible chickpea plant (*Cicer arietinum* L.) upon *Fusarium oxysporum* infection. Physiol Mol Plant Pathol 68:176-188.

Obayashi T, Kinoshita K, Nakai K, Shibaoka M, Hayashi S, Saeki M, Shibata D, Saito K and Ohta H (2007) ATTED-II: A database of co-expressed genes and *cis* elements for identifying co-regulated gene groups in Arabidopsis. Nucleic Acids Res 35:D863-869.

Oricchio E, Sciamanna I, Beraldi R, Tolstonog GV, Schumann GG and Spadafora C (2007) Distinct roles for LINE-1 and HERV-K retroelements in cell proliferation, differentiation and tumor progression. Oncogene 26:4226-4233.

Peaston AE, Eviskov AV, Graber JH, de Vries WN, Holbrook AE, Solter D and Knowles BB (2004) Retrotransposons regulate host genes in mouse oocytes and preimplantation embryos. Dev Cell 7:597-606.

Pouteau S, Grandbastien MA and Boccara M (1994) Microbial elicitors of plant defense responses activate transcription of a retrotransposon. Plant J 5:535-542.

SanMiguel P and Bennetzen JL (1998) Evidence that a recent increase in maize genome size was caused by the massive amplification of intergene retrotransposons. Ann Bot 82:37-44.

Song WY, Wang GL, Chen LL, Kim HS, Pi LY, Holsten T, Gardner J, Wang B, Zhai WX, Zhu LH, *et al.* (1995) A receptor kinase-like protein encoded by the rice disease resistance gene, *Xa21*. Science 270:1804-1806.

Song WY, Pi LY, Wang GL, Gardner J, Holsten T and Ronald PC (1997) Evolution of the rice *Xa21* disease resistance gene family. Plant Cell 9:1279-1287.

Song X, Sun Y and Garen A (2005) Roles of PSF protein and VL30 RNA in reversible gene regulation. Proc Natl Acad Sci USA 102:12189-12193.

Stetson DB, Ko JS, Heidmann T and Medzhitov R (2008) Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell 134:587-598.

Vallejos CE, Astua-Monge G, Jones V, Pflyer TR, Sakiyama NS and Mackenzie SA (2006) Genetic and molecular characterization of the *I* locus of *Phaseolus vulgaris*. Genetics 172:1229-1242.

van de Lagemaat LN, Landry JR, Mager DL and Medstrand P (2003) Transposable elements in mammals promote regulatory variation and diversification of genes with specialized functions. Trends Genet 19:530-536.

van der Biezen EA, Freddie CT, Kahn K, Parker JE and Jones JD (2002) Arabidopsis *RPP4* is a member of the *RPP5* multi-gene family of TIR-NB-LRR genes and confers downy mildew resistance through multiple signalling components. Plant J 29:439-451.

Vergne E, Ballini E, Droc G, Tharreau D, Nottéghem JL and Morel JB (2008) RCHIPELAGO: A dedicated resource for exploiting past, present, and future genomic data on disease resistance regulation in rice. Mol Plant Microbe Interact 21:869-878.

Wang GL, Ruan DL, Song WY, Sideris S, Chen L, Pi LY, Zhang S, Zhang Z, Fauquet C, Gaut BS, *et al.* (1998) *Xa21D* encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. Plant Cell 10:765-779.

Wang T, Zeng J, Lowe CB, Sellers RG, Salama SR, Yang M, Burgess SM, Brachmann RK and Haussler D (2007) Species-specific endogenous retroviruses shape the transcriptional network of the human tumor suppressor protein p53. Proc Natl Acad Sci USA 104:18613-18618.

Yoshioka K, Moeder W, Kang HG, Kachroo P, Masmoudi K, Berkowitz G and Klessig DF (2006) The chimeric Arabidopsis CYCLIC NUCLEOTIDE-GATED ION CHANNEL 11/12 activates multiple pathogen resistance responses. Plant Cell 18:747-747.

Yi H and Richards EJ (2007) A cluster of disease resistance genes in Arabidopsis is coordinately regulated by transcriptional activation and RNA silencing. Plant Cell 19:2929-2939.

Zhan S, Horrocks J and Lukens LN (2006) Islands of co-expressed neighbouring genes in *Arabidopsis thaliana* suggest higher-order chromosome domains. Plant J 45:347-357.

Associate Editor: Márcio de Castro Silva Filho

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.