PEST SEQUENCES IN *ARABIDOPSIS THALIANA* TRANSCRIPTION FACTORS

Israel Maruri-Lopez, Maria de Luz Guerrero-Gonzalez and Juan Francisco Jimenez-Bremont
Division de Biologia Molecular, Instituto Potosino de Investigacion Cientifica y Tecnologica AC. Camino a la Presa de San Jose 2055, CP 78216 San Luis Potosí, México. jbremont@ipicyt.edu.mx

Key words: PEST, WRKY, and unstructured regions

Introduction. Initially was considered that only repressor proteins carried out the transcription repression, but an alternative transcriptional control may involve the destruction of the factors that promote this (Muratani and Tansey, 2003). Although each transcriptional factor (TF) is eventually degraded as part of their natural turnover, TF stability is a process strictly regulated. It is know that the 26S proteasome plays a central role in the degradation of cytosolic and nuclear proteins (Yao and Ndoja, 2012).
Through different stimuli TFs and their regulators are degraded by the 26S proteasome, such as the mammalian IkBα, p53, yeast Gcn4, and plant NPR1, DREB2A (Kornitzer et al. 1994; Whitmarsh and Davis, 2000; Sakuma et al. 2006; Spoel et al. 2009). These proteins do not have target genes in common and any phylogenetic relationship. However, these proteins share regions rich in amino acids such as proline (P), glutamic acid (E), serine (S), threonine (T), and to a lesser extent aspartic acid (D). These regions are known as PEST sequences, which reduce drastically the protein half-life. The PEST regions are hydrophilic, and may be present as unstructured regions in the proteins (Rechsteiner and Rogers, 1996). Unstructured regions are extremely sensitive to proteolysis (Tompa, 2002).
In this study we present a bioinformatic analysis of the PEST sequences in all families of *Arabidopsis thaliana* TF. In detail the PEST sequences distribution in WRKY family was examined and compared with their *A. lyrata* orthologous, and its correlation with disordered regions.

Methods. The sequences of *A. thaliana* TFs were obtained in the AGRIS database (arabidopsis.med.ohio-state.edu/). The WRKY family classification was made according to Wang et al. 2011, and the *A. lyrata* orthologous were found by BLAST in Phytozome database (www.phytozome.net/). PEST sequences were identified using the PEST-find algorithm (emboss.bioinformatics.nl/cgi-bin/emboss/epestfind), and the disorder regions were predicted by PreDisorder program (casp.rnet.missouri.edu/predisorder.html).

Results. We analyzed 1808 sequences of *A. thaliana* TFs, which 61.44% have PEST regions with positive score. Based on the presence of conserved domains *A. thaliana* TFs are grouped into 50 families, where only the CCAAT-DR1 family lacks of sequences with positive PEST score (Fig.1).
In the same way, we examined the distribution of PEST regions in a specific TF family. The TF WRKY group is involved in plant senescence, biotic and abiotic responses; WRKY family is composed of more than 70 members, which can be divided into three groups on the basis of both the number of WRKY domains and the features of their zinc-finger-like motif (Wang et al. 2011).
Of the total WRKYs proteins 68% have positive PEST sequences. We found that in WRKY family, the PEST regions are distributed along the polypeptide chain, mainly located within disordered regions and outside the DNA binding domain. The same PEST scores and distribution are conserved in the *A. lyrata* orthologous proteins (Fig.2).

Fig.1. Percentage of positive PEST regions in WRKY transcription factor family

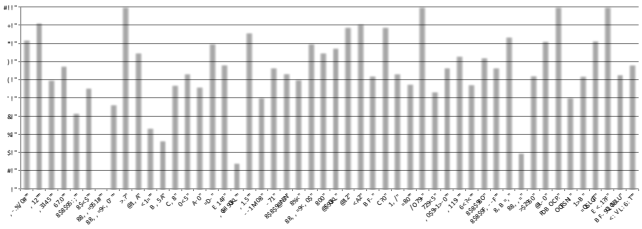
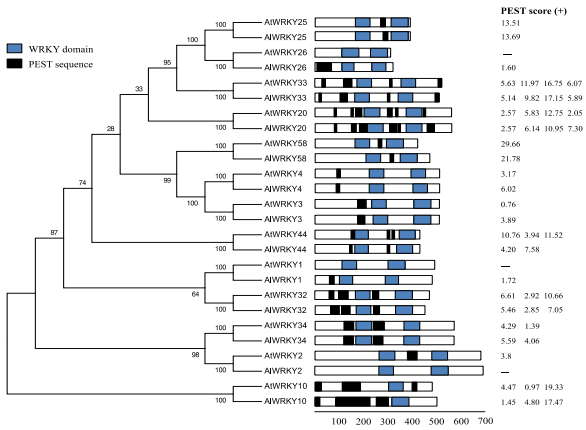


Fig. 2. Positive PEST regions in Arabidopsis WRKY group I



Conclusion. Our bioinformatics analysis suggests a high presence of potential PEST sequences. These degradation sequences could be involved in the regulation of plant transcription factors.

References.

1. Muratani M, Tansey WP. (2003). *Nat Rev Mol Cell Biol*. 4(3):192-201.
2. Yao T, Ndoja A. (2012). *Semin Cell Dev Biol*. 23(5):523-529.
3. Kornitzer D, Raboy B, Kulka RG, Fink GR. (1994). *EMBO J*. 13(24):6021-30.
4. Whitmarsh AJ and Davis RJ. (2000). *Cell Mol Life Sci*. 57 1172-83.
5. Sakuma Y, Maruyama K, Osakabe Y, Qin F, Seki M, Shinozaki K, Yamaguchi-Shinozaki K. (2006). *Plant Cell*. 18(5):1292-309.
6. Spoel SH, Mou Z, Tada Y, Spivey NW, Genschik P, Dong X. (2009). *Cell*. 137:860-872
7. Rechsteiner M, Rogers SW. (1996). *Trends Biochem Sci*. 21:267-271.
8. Tompa P. (2002). *Trends Biochem Sci* 27(10):527-533
9. Wang Q, Wang M, Zhang X, Hao B, Kaushik SK, Pan Y. (2011). *Genetica*. 139:973-983.