Plant Omics Journal POJ

POJ 4(2):95-99 (2011) ISSN:1836-3644

Phylogenetic analysis of SET domain in trithorax SlTX1 of *Solanum lycopersicum*

Monther T. Sadder[1,2*], Abdullah Alsadon[3], Mohammad Al-Thamra[2], Adel M. Zakri[3], Abdullah Al-Doss[3]

[1]Center of Excellence in Biotechnology Research, King Saud University, PO Box 2455, Riyadh 11451, Saudi Arabia
[2]Plant Biotechnology Lab, Faculty of Agriculture, University of Jordan, Amman, 11942, Jordan
[3]Department of Plant Production, College of Food and Agricultural Sciences, King Saud University, PO Box 2460, Riyadh 11451, Saudi Arabia

***Correspondent author: msadder@ksu.edu.sa**

Abstract

Trithorax gene family is a major player in epigenetic activities. This family of proteins carries the SET domain (SET: Su (var)3-9, Enhancer-of-zeste and Trithorax), which has methyltransferase activity. The mission of this work aims to study the SET domain of a novel tomato (*Solanum lycopersicum*) trithorax (SlTX1). *Arabidopsis* trithorax (ATX1) was used to fish out homologous sequences from tomato genome. The predicted SlTX1 sequence was used to design overlapping primers to clone the SET domain. The clone was sequnces and deposited in the GenBank. After multiple alignment with homologs along the plant kingdom, the sequences were analyzed using Neighbor-Joining method and bootstrapped 500 times. The generated phylogenetic tree revealed two major clades; higher plants clade and moss clade. Consequently, the first clade sharply separated monocots from dicots into two clusters. The monocot clade showed tighter grouping of maize with sorghum than with rice. On the other hand, the dicot clade was further subdivided into smaller subclades. The SET domain of tomato SlTX1 was subclustered in one subclade with the grape ortholog. The current investigation illustrates the distinctiveness evolution of the SET domain among monocot and dicot plants. These attractive functional features would be valuable in future plant phylogeny.

Keywords: tomato, SET domain, trithorax, phylogenetics.
Abbreviations:SET: SET: Su(var)3-9, Enhancer-of-zeste and Trithorax, SlTX1: *Solanum lycopersicum* trithorax number one. ATX1: *Arabidopsis* trithorax, PcG/TrxG: Polycomb group/Trithorax group.

Introduction

Epigenetics include several factors related to chromatin modification and thereby have a huge impact on gene accessibility and activity (Grimaud et al., 2006; Kiefer, 2007; Kouzarides, 2007). Epigenetic machinery is an integrated function of counteracting Polycomb group/Trithorax group (PcG/TrxG) gene families. PcG and TrxG genes are responsible to cause chemical modifications at histone tails (Avarmova, 2002), such modifications can be read by another family of factors that can recognize them. These modifications are considered the second language of the genetic code (Kouzarides, 2007; Bird, 2007).One of the major modifications to histone tails is the methylation by the functional domain "SET" (Su(var)3-9, Enhancer-of-zeste, Trithorax), which exists in major counteracting epigenetic controlling factors (Álvarez-Venegas and Avramova, 2002).PcG/TrxG related epigenetic effects were first discovered in animal systems and were thought to be limited to animals. PcG/TrxG gene families were linked to genome plasticity and adaptability, in centromeric- and telomeric-silencing, in cell-cycle regulation, in senescence, disease and cancer (Sparmann and van Lohuizen, 2006). PcG/TrxG gene families were discovered recently in related plants, where they were shown to control major developmental processes (Zhang, 2008; Pien et al., 2008). Moreover, these PcG/TrxG gene families and their epigenetic influence in response to adaptation and responses to biotic and abiotic stresses are just being investigated as a new area that can lead to a comprehensive understating of plant adaptation to stress (Li et al., 2008; Álvarez-Venegas et al., 2007). Plants responses to biotic and abiotic stresses (drought) are increasingly being investigated, where they are thought to be related to on/off expression of certain genes (Álvarez-Venegas et al., 2006b). This would include sugar transporter genes expressed under salt stress (Charkazi et al., 2010). The environmental sensing machinery involved signals that are by cells to altered gene expression. The lipid ligand, phosphatidyl-inositol 5-phosphate (PtdIns5*P*) was found to regulate ATX1 activity and thereby affecting expression of certain responsive genes (Álvarez-Venegas et al., 2006a). However, only few plant trithorax genes have been cloned and investigated.The aim of this work is to clone and sequnce SET domain of the putative tomato trithorax homolog SlTX1, followed by phylogenetic analysis using related plant species.

Materials and methods

Tomato cDNA

Tomato (*Solanum lycopersicum* L.) seeds were germinated on filter paper in dishes. Seedlings were transferred into pots in the greenhouse and watered as necessary. Total genomic DNA, as control, was isolated from young tomato leaves

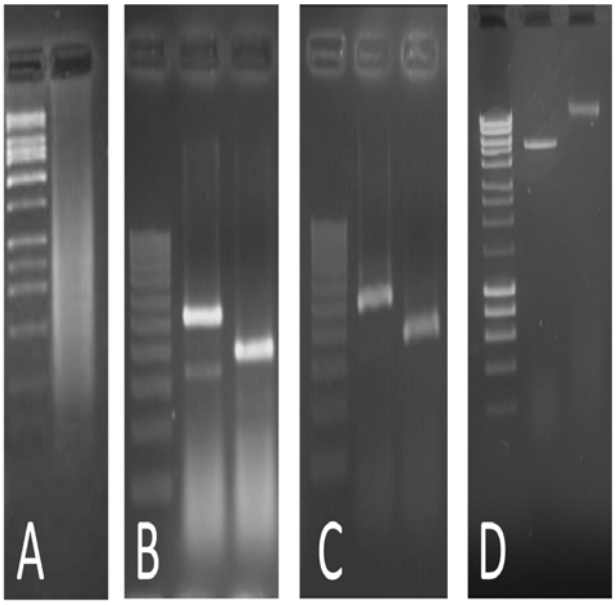

Fig 1. Samples analyzed using gel electrophoresis . (A) DNA ladder (lane 1), total RNA (lane 2). (B) RT-PCR using flower bud RNA: DNA ladder (lane 1), SET domain (lane 2), SET-post domain (lane 3). (C) RT-PCR using leaf RNA: DNA ladder (lane 1), SET domain (lane 2), SET-post domain (lane 3). pET28a digestion: DNA ladder (lane 1), native plasmid (lane 2), digested plasmid (lane 3).

using CTAB method (Hoisington et al. 1994). Total RNA was extracted from frozen leaves and flower buds using TRIzol reagent following the manufacturer's instructions (Invitrogen, USA). First strand cDNA was synthesized by reverse transcriptase (Promega, USA) based on total RNA.

Primers

Trithorax sequence from Arabidopsis, rice, grape and maize were retrieved from the Genbank database (Benson et al., 2005) to be used as probes to fish tomato homologs (SGN, 2010). Sequences were downloaded into VectorNTI software for analysis and primer design. The putative tomato SlTX1 was scanned for related domains using the Simple Modular Architecture Research Tool (SMART) (SMART, 2010). One forward primer (F07 SLX: 5'– caaagactagccttcggtaaatcc – 3'), and two reverse primers (R07 SLX: 5'– cctgtaatcatatgtcagttcttcc – 3') and (R08 SLX: 5'– ttctccttcccaatctattaac – 3') were designed. Primers F07 SLX and R07 SLX were used to amplify the SET domain, while primers F07 SLX and R08 SLX were used to amplify the SET domain including the post-SET feature available in SlTX1.Restriction recognition sites were added to clone the amplicons; (*Nco*I 5' end and *Not*I at 3' end).

Cloning

Designed extended primers were used to amplify candidate gene using thermalcycler with high fidelity *Pfu* Turbo DNA polymerase (Stratagene, USA). The thermocycler program starts with 5 min at 95^{o}C, followed by 35 cycles amplifications cycles (1 min at 95^{o}C, 30 sec at 50^{o}C, 1.5 min at 72^{o}C), and final extension step for 10 min at 72^{o}C. DNA fragments were purified with gel extraction kit (Qiagen, Germany). Used restriction enzymes, alkaline phosphatase and T4 DNA ligase were from New England Biolabs (USA). The purified cDNAs were cloned into linearized pET28a expression plasmid (Novagen, USA). Luria Bertani (LB) media were used for cultures (Davis et al, 1994). Plasmid isolation was carried out as described by Sambrook et al. (1989). *E. coli* strain DH5α competent cells were prepared using $CaCl_2$ method and transformed with heat shock technique (Sambrook et al., 1989). Gel electrophoresis (0.7-2% (w/v) agarose) was carried out as described (Sambrook et al., 1989). Gels were stained with ethidium bromide and photographed under UV light. For cloning, SYBR Gold dye (Invitrogen, USA) and safe blue light were used. Several DNA ladders were used including the 1 Kb Hi-Lo^{TM} DNA molecular weight marker (Minnesota Molecular Inc.).

SlTX1 SET domain

For sequencing, plasmid mini-preps were isolated (Omega biotek). Positive clones were verified with PCR using cloning primers. Inserts were sequenced using both T7p and T7t primers available in the backbone of the pET28a expression plasmid (Novagen, USA).

Phylogenetics

The identified SET domain of SlTX1 was used to retrieve similar domain in other plant species using blastp at the Genbank (Benson et al., 2005). The threshold limit was set to 50% identities. In addition, repeated entries with the same sequence from Arabidopsis, grapes, and moss were excluded. Consequently, the following sequences were selected from the thirteen plant species were XP_002301643 (*Populus trichocarpa*), XP_002268621 (*Vitis vinifera*), XP_002527758 (*Ricinus communis*), XP_002320433 (*Populus trichocarpa*), NP_172074 (*Arabidopsis thaliana*), AAF29390 (*Arabidopsis thaliana*), NP_850170 (*Arabidopsis thaliana*), ACR35362 (*Zea mays*), NP_001062635 (*Oryza sativa japonica*), XP_002459928 (*Sorghum bicolor*), XP_001767466 (*Physcomitrella patens* subsp. *patens*), XP_001780587 (*Physcomitrella patens* subsp. *patens*), and XP_001766115 (*Physcomitrella patens* subsp. *patens*). All available domains were subjected to ClustalX multiple sequence alignment. The phylogenetic analysis was performed by bootstrapping the aligned domains 500 times and exposing them to Neighbor-Joining method in MEGA software (Kumar et al., 2004).

Results

Total RNA was isolated and tested using gel electrophoresis (Fig. 1.A). After several attempts and temperature fine tuning, the predicted SET domain of the tomato trithorax homolog SlTX1 was amplified using RT-PCR (Fig. 1). The domain was expressed in leaves (Fig. 1.B) and flower buds (Fig. 1.C), however, it was relatively expressed more prominently in flower buds. The plasmid pET28a was linearized with *Nco*I and *Not*I restriction enzymes (Fig. 1.D) and used for ligation. The sequenced clones carrying the SET domain of SlTX1 gave 100% match to the predicted sequence. The SET domain of SlTX1 was deposited in the Genbank (Benson et al., 2005) under the accession number HQ426963.1. The tomato trithorax homolog SlTX1 (Fig. 2), was scanned for related domains using Simple Modular Architecture Research Tool (SMART) (SMART, 2010). All

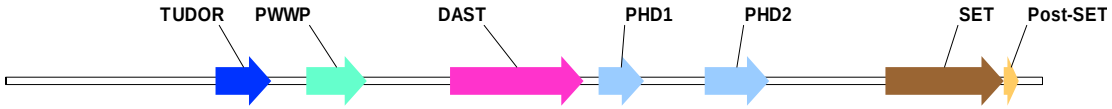

Fig 2. SMART domain prediction of SlTX1 (TUDOR, domain of unknown function present in several RNA-binding proteins; PWWP, domain with conserved PWWP motif; DAST, domain associated with SET in trithorax; PHD1 and PHD2, plant homeodomain; SET, SET [Su(var)3-9, Enhancer-of-zeste, Trithorax] domian; post-SET).

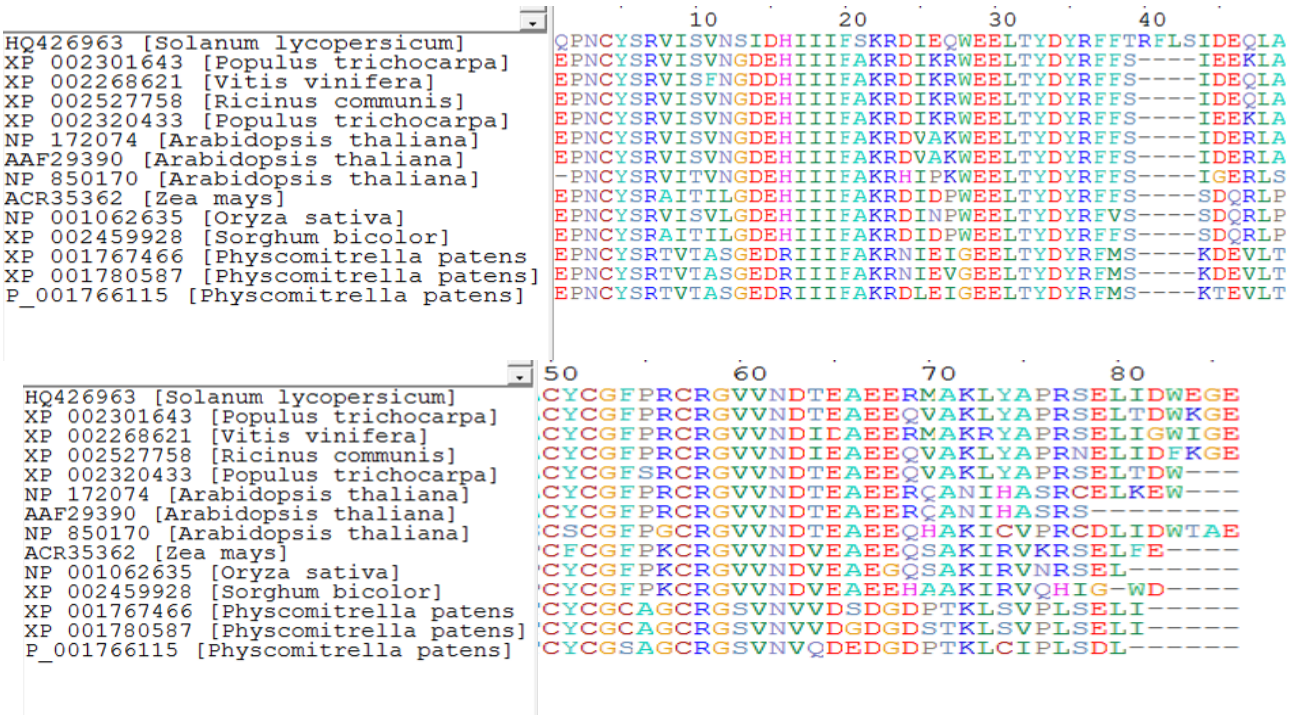

Fig 3. Multiple alignment for SET domain of plant trithorax genes.

the aligned SET sequences contained the motif CXCX4C, featured with three conserved cysteine residues, which are essential for its histone lysine methyltransferase activity (Fig. 3). The SET domain based analysis revealed two major clades; the higher plants clade and the moss clade (Fig. 4). Furthermore, the first clade sharply separated monocots from dicots into two clusters. In the monocot clade, maize was strongly clustered first with sorghum homolog (93%) then with rice homology (96%). No clustering of Arabidopsis was detected with monocots, indicating the distinctive structure of SET domain in monocots as compared to dicots. On the other hand, the dicot clade was further subdivided into smaller subclades. The SET domain of tomato SlTX1 was subclustered in one subclade with the grape ortholog. In the monocot clade, maize was strongly clustered first with sorghum homolog (93%) then with rice homology (96%). No clustering of Arabidopsis was detected with monocots, indicating the distinctive structure of SET.

Discussion

Our results showed SlTX1 expression to be allocated in flower buds more than leaves. Our earlier investigation with trithorx homolog in Arabidopsis showed differential expression of the gene using reporter gene driven by their promoters (Saleh et al., 2008). A recently developed *in vitro* flowering and fruiting technique for tomato, would help in large sacl functional analysis of flower related genes (Mamidala and Nanna, 2009). It is however, important to emphasis the nature of gene duplications in Arabidopsis, with at least five homologs. This would make it more difficult to functionally analyze individual genes using knockout or knockdown tool (Álvarez-Venegas et al., 2006a). Consequently, the unique tomato gene would be expected to cover an array of complex functions as related to a set of paralogs in Arabidopsis. This would be tweaked by spatial and temporal interactions with signal molecules and other regulatory proteins. Further investigations still needed to clarify this behaviour in tomato.The SET domains of thirteen trithorax proteins from eight species, along with the novel SlTX1 cloned in this study, were subjected to phylogenetic analysis. The SET domain based analysis revealed two major clades; the higher plants clade and the moss clade. This indicates a major evolutionary division at the epigenetic control units. Furthermore, the first clade sharply separated monocots from dicots into two clusters. The resulted two major subclades based on structural similarities, could imply a functional differentiations between the two groups. Similar structural variation was recorded between maize and Arabidopsis SET domains (Springer et al., 2003). The SlTX1 was moderately clustered (53%) in the first subclade (dicots) together with homologs from poplar, caster bean, grape, and two genes from Arabidopsis. While the second subcalde clustered only monocots (maize, rise and sorghum). This

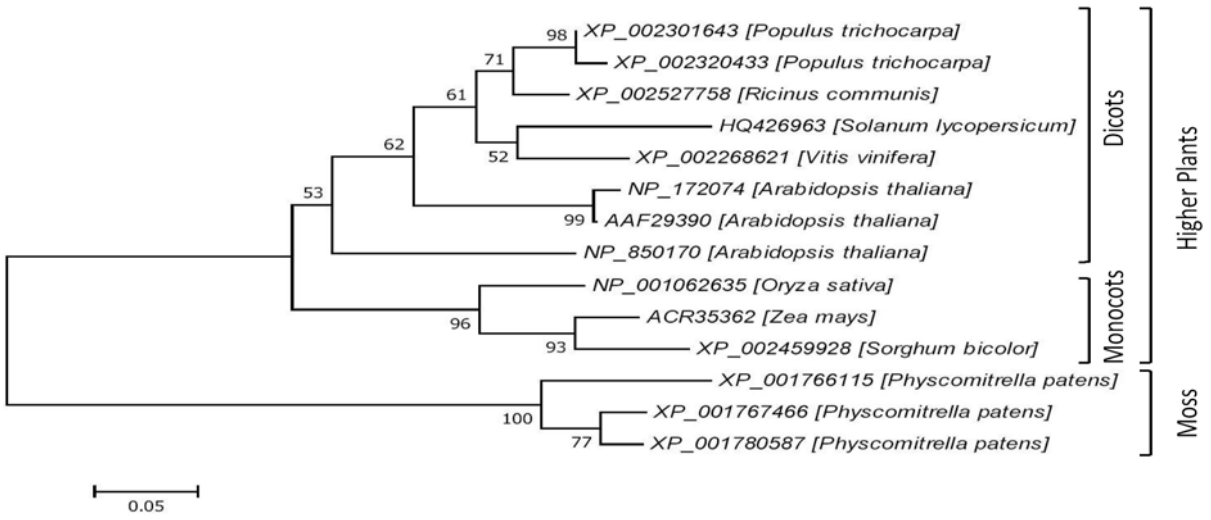

Fig 4. Phylogenetic analysis of SET domain of plant trithorax genes showing a consensus tree. Nodes are showing percentage of bootstrap values.

would imply a functional specialization between monocot and dicot plants (Dillon et al., 2005; Ng et al., 2007; Pontvianne et al., 2010). Although, some proteins in the first clade represent paralogs from the same species, e.g. Arabidopsis and poplar. The previously cloned trithorax genes from Arabidopsis (Álvarez-Venegas et al., 2003) were clustered separately from SlTX1 SET, indicating a potential divergent methylation activity (Saleh et al., 2008), which still need to be investigated using in vitro assay. This study demonstrates the distinctiveness evolution of the SET domain among monocot and dicot plants. These attractive functional features would be valuable in future plant phylogeny.

Acknowledgement

The authors acknowledge the generous financial support by Center of Excellence in Biotechnology Research grant (CEBROR-1430/1), King Saud University and Ministry of Higher Education, Saudi Arabia.

References

Álvarez-Venegas R, Al Abdallat A, Guo M, Alfano JP, Avramova Z (2007) Epigenetic control of a transcription factor at the cross section of two antagonistic pathways. Epigenetics 3:106-113.

Álvarez-Venegas R, Avramova Z (2002) SET-domain proteins of the Su(var)3-9, E(z) and Trithorax families. Gene 285:25–37.

Álvarez-Venegas R, Pien S, Sadder M, Witmer X, Grossniklaus U, Avramova Z (2003) ATX-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. Curr Biol 13:627-37.

Álvarez-Venegas R, Sadder M, Hlavacka A, Baluška F, Xia Y, Lu G, Firsov A, Sarath G, Moriyama H, Dubrovsky J, Avramova Z (2006a) The *Arabidopsis* homolog of trithorax, ATX1, binds phosphoinositide 5-phosphate and the two regulate a common set of target genes. Proc Natl Acad Sci 103:6049-6054.

Álvarez-Venegas R, Xia Y, Lu G, Avramova Z (2006b) Phosphoinositide 5-phosphate and phosphoinositide 4-phosphate trigger distinct specific responses of Arabidopsis genes. Plant Signal Behav 1(3):140-151

Avramova Z (2002) Heterochromatin in animals and plants. similarities and differences. Plant Physiol 129:40–49.

Bird A (2007) Perceptions of epigenetics. Nature 447:396-8.

Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL (2005). GenBank. Nucleic Acids Research 33:D34–D38.

Charkazi F, Ramezanpour SS, Soltanloo H (2010) Expression pattern of two sugar transporter genes (*SuT4* and *SuT5*) under salt stress in wheat. Plant Omics Journal 3(6):194-198.

Davis LD, Kuehl MW, Battey JF (1994) Basic methods in molecular biology. second edition. Paramount Publishing Business And Professional Group, USA. pp.47-53.

Dillon SC, Zhang X, Trievel RC, Cheng X (2005) The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol 6:227. doi:10.1186/gb-2005-6-8-227.

Grimaud C, Negre N, Cavalli G (2006) From genetics to epigenetics: the tale of Polycomb group and trithorax group genes. Chromosome Res 14:363-75.

Hoisington D, Khairallah M, González-de-León D (1994) Laboratory protocols: CIMMYT applied molecular genetics laboratory, CIMMYT, Mexico.

Kiefer JC (2007) Epigenetics in development. Dev Dyn 236:1144-56.

Kouzarides T (2007) Chromatin modifications and their function. Cell 128:693-705.

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings Bioinf 5:150-163.

Li W, Wang X, He K, Ma Y, Su N, He H, Stolc V, Tongprasit W, Jin W, Jiang J, Terzaghi W, Li S, Deng XW (2008) High-resolution mapping of epigenetic modifications of the rice genome uncovers interplay between DNA methylation, histone methylation, and gene expression. Plant Cell 20:259-276.

Mamidala P, Nanna RS (2009) Efficient *in vitro* plant regeneration, flowering and fruiting of dwarf Tomato cv. Micro-Msk. Plant Omics Journal 2(3):98-102.

Ng DW, Wang T, Chandrasekharan MB, Aramayo R, Kertbundit S, Hall TC (2007) Plant SET domain-containing proteins: structure, function and regulation. Biochim Biophys Acta 1769(5-6):316-329.

Pien S, Fleury D, Mylne JS, Crevillen P, Inzé D, Avramova Z., Dean C, Grossniklaus U (2008) Unraveling *trithorax* function in plants: ATX1 dynamically regulates the activation of *FLC* via histone 3 lysine 4 tri-methylation. Plant Cell 20:580-588.

Pontvianne F, Todd B, Craig SP (2010). Arabidopsis histone lysine methyltransferases. Adv Bot Res 53:1-22.

Saleh A, Alvarez-Venegas R, Yielmaz M, Li O, Hou Q, Sadder M, Al-Abdallat A, Xia Y, Lu G, Ladunga I, Avramova Z (2008) The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. Plant Cell 20(3):568-579.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual 2nd ed. Cold Spring Harbor: Cold Spring Harbor Laboratory Press. SGN (2010) <http://www.sgn.cornell.edu>.

SMART (2010) <http://smart.embl-heidelberg.de/>.

Sparmann A, van Lohuizen M (2006) Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer 6:846-856.

Springer NM, Napoli CA, Selinger DA, Pandey R., Cone KC, Chandler VL, Kaeppler HF, and Kaeppler SM (2003) Comparative analysis of SET domain proteins in maize and arabidopsis reveals multiple duplications preceding the divergence of monocots and dicots. Plant Physiol 132:907–925.

Zhang X (2008) The epigenetic landscape of plants. Science 320:489-492.