Plant Cell Rep (2011) 30:1087–1097
DOI 10.1007/s00299-011-1015-0

ORIGINAL PAPER

Genome-wide analysis of the SET DOMAIN GROUP family in Grapevine

Felipe Aquea · Andrea Vega · Tania Timmermann · María Josefina Poupin · Patricio Arce-Johnson

Received: 6 October 2010 / Revised: 7 January 2011 / Accepted: 14 January 2011 / Published online: 4 February 2011

Abstract The SET DOMAIN GROUP (SDG) proteins represent an evolutionarily-conserved family of epigenetic regulators present in eukaryotes and are putative candidates for the catalysis of lysine methylation in histones. Plant genomes analyses of this family have been performed in arabidopsis, maize, and rice and functional studies have shown that SDG genes are involved in the control of plant development. In this work, we describe the identification and structural characterization of SDG genes in the *Vitis vinifera* genome. This analysis revealed the presence of 33 putative SDG genes that can be grouped into different classes, as it has been previously described for plants. In addition to the SET domain, the proteins identified possessed other domains in the different classes. As part of our study regarding the growth and development of grapevine, we selected eight genes and their expression levels were analyzed in representative vegetative and reproductive organs of this species. The selected genes showed different patterns of expression during inflorescence and fruit development, suggesting that they participate in these processes. Furthermore, we showed that the expression of selected SDGs changes during viral infection, using as a model Grapevine Leafroll Associated Virus 3-infected symptomatic grapevine leaves and fruits. Our results suggest that developmental changes caused by this virus could be the result of alterations in SDG expression.

Keywords Grapes · *Vitis vinifera* · Chromatin remodeling · Histone methyltransferase

Abbreviations

SDG	SET DOMAIN GROUP
GLRaV-3	Grapevine Leafroll associated virus 3

Introduction

An epigenetic mechanism mediated by the modification of nucleosomal histone tails plays an important role in eukaryotic development (Ho and Crabtree 2010). Unlike histone lysine acetylation, which is generally associated with gene activation, histone methylation at specific lysine residues can lead to either gene activation or repression (Liu et al. 2010). At least six lysine residues on histone H3 (K4, K9, K27, K36, K79) and one in H4 (K20) are targeted by histone lysine methyltransferases (Liu et al. 2010). Histone methylation participates in multiple developmental processes including cell cycle regulation, heterochromatin formation, transcriptional activation and transcriptional silencing. Recently, epigenetic regulation of stress responses has been reported in plants (Chinnusamy and Zhu 2009).

A family of SET domain-containing proteins catalyzes the methylation of histone lysine residues, with the

Communicated by D. Zaitlin.

F. Aquea · P. Arce-Johnson (✉)
Departamento de Genética Molecular y Microbiología, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, P.O.Box 114-D, Santiago, Chile
e-mail: parce@bio.puc.cl

A. Vega
Departamento de Ciencias Vegetales, Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Santiago, Chile

T. Timmermann · M. J. Poupin
Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Santiago, Chile

exception of H3 lysine 79 (Feng et al. 2002). The SET domain is named by the *Drosophila melanogaster* proteins: *S*uppressor of variegation 3–9, *E*nhancer of zeste, and *T*rithorax (Jenuwein et al. 1998). Much has been learned from the biochemical characterization of the histone methyltransferase (HMT) activities of the SET domain proteins and their effects on both gene repression and gene activation (Qian and Zhou 2006; Dillon et al. 2005). However, the functions of these HMTs during eukaryotic development are still largely unclear.

Proteins containing the conserved SET domain can be found in organisms ranging from viruses to all three domains of life (Bacteria, Archaea, and Eukaryota) and the domain usually functions as part of a larger multi-domain protein. The proteins with a SET domain have been named SET DOMAIN GROUP or SDG proteins. At the moment, 2,719 SET domains present in 2,704 proteins are cataloged in the Pfam sequence alignment database, and at least 49, 34, and 31 SDG proteins are present in Arabidopsis, rice, and maize, respectively (Pontvianne et al. 2010; Ng et al. 2007). Arabidopsis SDG proteins are the best annotated and characterized. Initially, Baumbusch et al. (2001) classified 37 putative arabidopsis SDG proteins into four distinct classes: (1) enhancer of zeste [E(z)] homologs; (2) Ash1 homologs and related proteins; (3) trithorax (trx) homologs and related proteins; and (4) suppressor of variegation [Su(var)] homologs and related proteins. Later, Springer et al. (2003) classified 32 arabidopsis and 22 maize SDG proteins into five classes according to their phylogenetic relationships and domain organization. More recently, two further classes (VI and VII) have been added (Ng et al. 2007). Although much remains to be experimentally verified, in general, it appears that the resulting groupings reflect the substrate specificities of the members.

The functions of SDG proteins are regulated by protein–protein interactions that involve both intra- and intermolecular associations and are important in plant developmental processes, such as flowering time control and embryogenesis (Jarillo et al. 2009; Ahmad et al. 2010; Pontvianne et al. 2010). For example, the Arabidopsis ASH1 HOMOLOG 2 protein (ASHH2, SDG8) has been suggested to methylate H3K4 and/or H3K36 and is similar to Drosophila ASH1, a positive maintainer of gene expression, and yeast Set2, a H3K36 HMTase (Zhao et al. 2005). Mutation of the ASHH2 gene has pleiotropic developmental effects, leading to alterations in ovule and anther development, reduced dimethylation of histone H3K36 in the FLOWERING LOCUS C promoter and changes in shoot branching and carotenoid composition (Grini et al. 2009; Cazzonelli et al. 2009). The CURLY LEAF (CLF, SDG1) gene in Arabidopsis is required for stable repression of a floral homeotic gene AGAMOUS in leaves and stems (Goodrich et al. 1997). MEDEA (MEA, SDG5) encodes an Arabidopsis SET domain Polycomb protein. Inheritance of a maternal loss-of-function mea allele results in embryo abortion and prolonged endosperm production, irrespective of the genotype of the paternal allele (Grossniklaus et al. 1998). ARABIDOPSIS TRITHORAX-RELATED PROTEIN5 (ATXR5, SDG15) exhibits H3K27 monomethyltransferase activity and mutants show partial heterochromatin decondensation (Jacob et al. 2010) and ARABIDOPSIS TRITHORAX-RELATED PROTEIN7 (ATXR7, SDG25) is a putative Set1 class H3K4 methylase required for proper FLC expression (Tamada et al. 2009). In addition, ARABIDOPSIS TRITHORAX 1 (ATX1, SDG27) functions as an activator of homeotic genes, just as Trithorax does in animal systems (Alvarez-Venegas et al. 2003). On the other hand, 10 SUVH genes encode SU(VAR)3–9 homologues that control heterochromatic domains. Loss of function suppresses, whereas overexpression enhances gene silencing, causes ectopic heterochromatization and significant growth defects in Arabidopsis (Naumann et al. 2005). Recently, it has been described that SDG2, a large Arabidopsis protein member of the class III, is the major enzyme responsible for trimethylation of H3K4 and is crucial for both sporophyte and gametophyte development (Guo et al. 2010; Berr et al. 2010).

Despite this progress, many aspects of the role that SDG proteins may play in plant development and the detailed mechanism by which they regulate chromatin structure and gene activity remain unclear. Moreover, very little is known about the functions of SDG proteins in fruit tree species.

As part of our research efforts to identify genes involved in growth and development of grapevine, we present for the first time the identification of SDG proteins in the fruit tree *Vitis vinifera*. We identified 33 genes by means of in silico analysis of the grapevine genome sequence (Velasco et al. 2007; Jaillon et al. 2007). Additionally, we selected eight genes and evaluated their pattern of expression during grapevine development and viral infection using as a model Grapevine Leafroll Associated Virus 3 (GLRaV3)-infected symptomatic grapevine leaves and fruits.

Materials and methods

Protein identification

Using the *Arabidopsis thaliana* sequences of SDG proteins obtained from the plant chromatin databases (http://www.chromdb.org), we performed a search in the grapevine genome using the BLASTp algorithm in the NCBI browser (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The sequences

obtained were corroborated in the Grape Genome Browser (http://www.cns.fr/externe/GenomeBrowser/Vitis/).

Domain predictions

The protein sequences of all SDG proteins were analyzed for recognizable domains using BLAST-based NCBI conserved domain searches (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). Domains were also verified using the HMMER-based SMART Web site (http://smart.embl-heidelberg.de).

Phylogenetic analysis

Phylogenetic analyses were performed using the grape region of the SET domain identified in this work along with proteins from Arabidopsis obtained from plant chromatin databases using the MEGA4 program (Tamura et al. 2007). The evolutionary history was inferred using the Neighbour-Joining method. Only in pairwise sequence comparison, the positions containing alignment gaps and missing data were eliminated (Pairwise deletion option).

Plant material and virus detection

Different vegetative and reproductive organs were collected from grapevine plants (Cabernet-Sauvignon) growing in commercial fields in central Chile and from in vitro culture material (Aquea et al. 2010). Roots and leaves were collected from a pool of 5 in vitro plants grown in MS medium in a growth chamber adjusted to 23°C with a 16/8 h light/dark photoperiod. Fruits and flowers were collected from three plants of 6 year-old trees growing in Curacaví (Región Metropolitana), Chile. Early inflorescence clusters and developed single flowers were sampled every 2 weeks beginning at −10 weeks after anthesis (−10 WAA) until floral cap structures were detached from each flower, at 0 WAA. Grapevine berries were harvested at 4-week intervals throughout fruit development, corresponding to immature berries [E–L stage 31, 4 weeks after flowering (WAF))] and three time points within Stage III (E–L stages 35, 36, and 38), spanning from veraison to harvest (8, 12, and 16 WAF, respectively). The veraison state was considered when 30–50% of the berries in the bunch presented a clear colored phenotype. Two hundred and fifty berries (3–4 clusters) of each selected plant were collected in the aforementioned stages. The samples were collected and frozen in liquid nitrogen for RNA extraction. Healthy and GLRaV-3-infected plants were obtained from the nursery of the Agronomy Faculty, Pontificia Universidad Católica de Chile. Healthy and virus-infected plants were maintained separately in the field with similar growing conditions. The presence of 14 viruses that have a high infectious incidence in Chile was evaluated: Grapevine Virus A (GVA), Grapevine virus B (GVB), Grapevine Fanleaf Virus (GFLV), Grapevine Fleck Virus (GFkV), Tomato Ringspot Virus (ToRSV), Grapevine rootstock stem lesion associated virus (GRSLaV) and Grapevine Leaf-Roll-Associated Viruses (GLRaV) 1, 2, 3, 4, 5, 6, 7 and 8 (Vega et al. unpublished data). Viral screening was performed by RT–PCR from RNA extracted from leaf samples obtained from the medial zone of the main shoot.

RNA isolation and RT–PCR analysis

To relatively quantify the expression pattern of selected SDG genes, total RNA was isolated from all organs according to the procedure of Reid et al. (2006), using a CTAB-Spermidine extraction buffer. For cDNA synthesis, 1 μg of total RNA treated with DNAse I (RQ1, Promega) was reverse transcribed with random hexamer primers using the StrataScript® reverse transcriptase (Statagene), according to the manufacturer's instructions. The cDNA obtained was used in real time PCR assays, with three biological replicates and two technical replicates. Real-time RT–PCR was performed using the Brilliant SYBR Green QPCR Master Reagent Kit (Stratagene) and the Mx3000P detection system (Stratagene) as described in the manufacturer's manual. The primers used were: *SDG6901* 5′-AAAGATGCCTATGTGGGTCA-3′ and 5′-CCTGAAG CTGTCTTGCGTC-3′; *SDG6903* 5′-CCGTCGATGTCTT GTCTTTG-3′ and 5′-GCGCCACATG GTATATTGTC-3′; *SDG6905* 5′-TCGTGCAAACTCATCCTTTC-3′ and 5′-CC TTTGCATAGAGTTGGG-3′; *SDG6911* 5′-ACTATTCCA TTTCCGATGCC-3′ and 5′-CCACCTGATGTGAATGTT CC-3′; *SDG6925* 5′-TTAAAGAACCTG GGTGCCAT-3′ and 5′-AGAGAATGGCGGATAACGTC-3′; *SDG6926* 5′-AACAGGCTGCATAGACACTCA-3′ and 5′-TTGGAG TTCTTGTACAGTTG AGG-3′; *SDG6932* 5′-AACGGC AATGGGTTTAATTC-3′ and 5′-CCCTCTTTC ACAACT TGCAC-3′; *SDG6934* 5′-TTGAAATGTGGTGGTGCAG-3′ and 5′-GGGAAGAAGTGAGGACCAAA-3′. Amplification of the grapevine *UBIQUITIN1* gene was used for reaction normalization with the primers VvUBI-F TCT GAG GCT TCG TGG TGG TA and VvUBI-R AGG CGT GCA TAA CAT TTG CG. Amplification of a fragment of the *UBIQUITIN1* gene (99 bp; TC53702) was used for normalization, as this gene has been demonstrated to be a good housekeeping gene in grapevine (Downey et al. 2003; Poupin et al. 2007; Matus et al. 2008, 2010). Standard quantification curves with serial dilutions of PCR products were constructed for each gene to calculate amplification efficiency according to Matus et al. (2010). Ct values for *UBIQUITIN* varied not more than 1 unit between all samples analyzed for each real time experiment. Reaction specificities for each primers were tested with melt gradient

dissociation curves, electrophoresis gels and cloning and sequencing of each PCR product. The PCR mixture (25 µl) contained 2 µl of cDNA template (diluted 1:10) and 140 nM of each primer. Amplification was performed under the following conditions: 95°C for 10 min, followed by 40 cycles of 94°C, 30 s; 55°C, 40 s; and 72°C, 40 s, followed by a melt cycle from 55 to 95°C. The relative expression of each gene was gene-wise normalized using Genesis software (Sturn et al. 2002). Hierarchical clustering of gene expression data was performed using the same software.

Results

Identification and annotation of grapevine SDG genes

As part of our study regarding the growth and development of grapevine, the grape genome sequence was searched for homologues of known SDG proteins as described in the "Materials and methods". According to the ChromDB nomenclature (http://www.chromdb.org), the grapevine genes were named SDG6900. A total of 33 SDG genes with a predicted SET domain were identified in the grapevine genome (Table 1). This number is similar to those present in arabidopsis and rice (49 and 34, respectively) and can also be found in the chromatin database. Table 1 shows the Genbank accession number, the gene structural analysis (ORF length, predicted size of the encoded protein, genomic length, number of introns), the SET domain size and the chromosome location of each gene. This identification does not fully confirm the gene predictions present in the annotated grapevine genome sequences since no cDNAs were cloned and further genes could exist as the annotation of the grapevine genome improves.

The chromosome analysis revealed that grapevine SDG genes are widely distributed in 12 of the 19 chromosomes (Fig. 1). Chromosomes 2, 3, 6, 9, 13, 17 and 19 do not contain SDG genes. However, the *SDG6923* and *SDG6935* gene models have not yet been assigned to a particular chromosome. The maximum number of genes (5; 14%) was found to be localized on chromosome 16 (*SDG6912, SDG6916, SDG2917, SDG6926* and *SDG6929*).

Phylogenetic analysis of SDG proteins

To examine the phylogenetic relationships among grapevine SDG proteins and group them within the established classes, we constructed a phylogenetic tree from alignments of grapevine with arabidopsis protein sequences, using the highly-conserved SET-domain region of each sequence to perform this analysis (~150 amino acids). All grapevine SET domain protein were grouped with their arabidopsis counterparts (Fig. 2). Similar results were obtained when the rice SET domain proteins were included in the phylogenetic analysis (data not shown). Although the number of proteins grouped in each class was generally similar between grapevine and arabidopsis, some interesting exceptions could be observed. According to the latest report, the plant SDG proteins are grouped in 7 classes (Ng et al. 2007) and in our analysis, we classify the grapevine SDG proteins into these same classes (Fig. 2, Table 1). Class I includes the arabidopsis proteins CLF, EZA1/SWN and MEA. In grapevine, only homologs of CLF and EZA1/SWN (SDG6903 and SDG6905, respectively) were found, and no homologs of MEA were found in our analysis of the grapevine genome. The largest class is class V, with 12 members in grapevine. Similar numbers of members are grouped in classes II, III, IV, VI, and VII.

Conserved domains in SDG proteins

We analyzed the architecture of grapevine SDG proteins and found that proteins classified in a particular class conserved other domains in addition to the SET domain (Fig. 3). Class I also has a SANT domain (SWI3, ADA2, N-CoR and TFIIIB″ DNA-binding domains) in the N-extreme. All members of class II have a SET domain that is invariably preceded by an AWS. Furthermore, SDG6908 has a double PHD domain in the N-extreme. As shown in Fig. 3, in addition to the SET domain, Class III proteins bear several highly-conserved protein domains (PWWP, FYRC, GYF, and TUDOR) and the PHD domain. Although biochemical characterization is lacking for these members, the presence of various highly-conserved domains within this class of proteins may suggest diverse functions for these SDG proteins. The type and location of the domains present in class I and class IV proteins are similar. Class V is the largest class of proteins and is characterized by the presence of pre-SET and SET domains. Proteins from classes VI and VII lack the additional domains that can be found in other classes of SET proteins. Lack of certain domains within members of each class suggests that some annotations are incomplete (for example, SDG6928 and SDG6933).

Expression analyses of SDG genes

SDG genes have mainly been involved in the regulation of flowering time and in embryo development. In order to further associate their biological function in grapevine with flowering and other specific developmental processes, we selected eight genes and their expression levels were quantitatively analyzed in representative vegetative and reproductive organs of the grapevine. Based on known

Table 1 SDGs proteins in *Vitis vinifera*

Gene working name[a]	Class	GenBank protein ID[b]	ORF length (bp)	Protein length (aa)	SET domain size (aa)	Chromosome	Genomic length (pb)	No. of introns
SDG6901	V	CBI32864	1,434	477	90	8	2,214	5
SDG6902	III	CBI21104	3,336	1,111	128	4	31,818	12
SDG6903	I	CBI21398	2,805	934	122	7	19,389	16
SDG6904	IV	CBI37567	1,125	374	129	4	25,386	5
SDG6905	I	CBI36953	1,149	382	122	7	4,482	9
SDG6908	II	CBI36587	1,485	494	124	11	22,776	10
SDG6909	II	CBI18234	2,589	862	124	18	32,415	21
SDG6910	III	CBI39161	3,207	1,068	125	5	57,130	24
SDG6911	II	CBI18964	5,877	1,958	124	18	27,311	23
SDG6912	V	CBI38560	1,332	443	122	16	1,389	1
SDG6913	III	CBI40526	3,012	1,003	124	15	12,264	24
SDG6915	VII	CBI27360	1,350	449	241	4	19,890	8
SDG6916	V	CBI38579	1,671	556	124	16	2151	1
SDG6917	II	CBI26426	1,317	438	124	16	29,350	13
SDG6918	V	CBI23710	1,554	517	141	5	2,796	5
SDG6919	VI	CBI23159	1,512	503	195	10	1,605	1
SDG6921	VII	CBI35049	1,455	484	225	1	3,132	5
SDG6923	V	CBI29505	1,581	526	124	undetermined	2,178	4
SDG6924	V	CBI21273	2,106	701	152	14	45,152	13
SDG6925	V	CBI17591	3,948	1,315	140	4	17,823	17
SDG6926	IV	CBI23040	1,203	400	129	16	4,516	5
SDG6927	V	CBI37177	2,148	715	141	7	14,826	11
SDG6928	V	CBI22320	582	193	141	14	30,772	4
SDG6929	III	CBI23139	3,057	1,018	124	16	14,152	24
SDG6930	VI	CBI28962	1,395	464	155	12	10,040	15
SDG6931	V	CBI29255	1,458	485	152	1	13,576	10
SDG6932	III	CBI28983	6,600	2,199	141	12	16,691	25
SDG6933	V	CBI31239	555	184	139	8	5,785	1
SDG6934	V	CBI23736	1,767	588	150	5	2,019	2
SDG6935	III	CBI29431	3,384	1,127	243	undetermined	11,131	12
SDG6936	VI	CBI19071	1,443	480	244	18	44,898	13
SDG6937	VI	CBI29967	1,983	660	259	8	6,934	9
SDG6938	VI	CBI18219	1,602	533	134	18	5,716	5

[a] Named according to the nomenclature used in the chromatin database (www.chromdb.org)

[b] Accession numbers obtained from NCBI GenBank

ESTs (Childs et al. 2007, http://plantta.jcvi.org) we selected *SDG6901*, *SDG6903*, *SDG6905*, *SDG6911*, *SDG6925*, *SDG6926*, *SDG6932* and *SDG6934* for this analysis. The genes selected represent four of the seven classes identified. In all cases, the gene expression patterns differed within each class, but were conserved in all the treatments and stages studied. We detected the expression of almost all genes analyzed in vegetative and reproductive organs, with the exception of *SDG6926*, which is expressed mostly in flowers (Fig. 4). The expression of *SDG6932*, *SDG6934*, *SDG6905*, and *SDG6911* is lower in roots and higher in ripe fruits. *SDG6925* and *SDG6901* are expressed at low levels in flowers, whereas the highest expression level of *SDG6903* was detected in this organ (Fig. 4).

In order to gain further insights into the function of grapevine SDG genes, their pattern of expression was determined during inflorescence and fruit development (Figs. 5, 6). Flower development was categorized at different growth stages before anthesis (capfall, Matus et al. 2010). In our analysis, the expression of *SDG6932*, *SDG6934*, *SDG6905*, and *SDG6911* increased from the start of flowering until anthesis (Fig. 5). In contrast, *SDG6926* showed an opposite pattern; its expression was higher at the start of flowering and decreased towards anthesis. On the

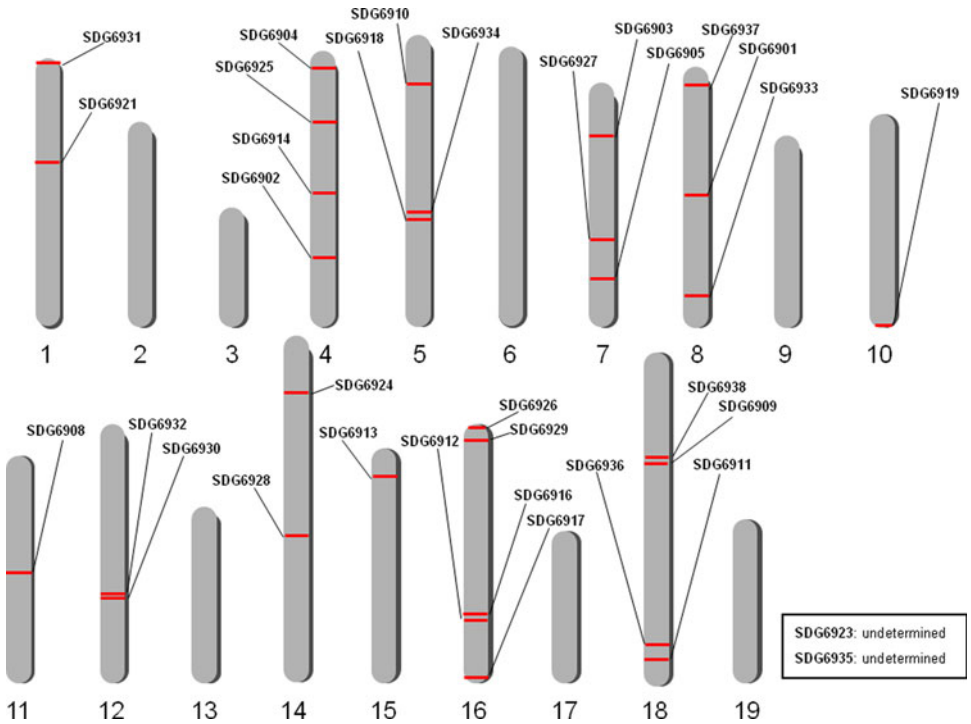

Fig. 1 Chromosomal locations of grapevine SDG genes. Chromosome numbers are indicated at the base of each chromosome

other hand, *SDG6903* and *SDG6901* had a peak of expression at −6 week and *SDG6925* at −4 week. In both cases, the expression of these genes decreased towards anthesis (Fig. 5). We continued the characterization during berry development after anthesis. The expression levels of *SDG6932, SDG6934*, *SDG6905*, *SDG6911*, *SDG6901*, and *SDG6925* increased from immature to ripe fruit, whereas *SDG6903* showed the opposite pattern (Fig. 6). *SDG6926* was not expressed during berry development (Fig. 6).

Many viruses affect grapevine cultures without inducing any resistance response, leading to the development of systemic diseases and chronic infections. Our previous work has demonstrated that viral infections induce changes in histone acetyltransferase gene expression (Espinoza et al. 2007). We evaluated the pattern of expression of selected SDG genes in leaves and berries infected with GLRaV-3 of systemically-infected plants. *SDG6926* is exclusively expressed in flowers in a healthy plant. Virus infection did not induce expression of this gene in other organs (Fig. 7). In leaves and ripening berries, *SDG6934* expression was induced by viral infection, whereas *SDG6903* expression was reduced under the same conditions. The pattern of expression of the other genes analyzed did not change after viral infection. In berries at veraison, the expression of *SDG6903* was not modified, whilst that of *SDG6932*, *SDG6934*, *SDG6905*, *SDG6911*, *SDG6901*, and *SDG6925* was induced in an infected plant. The pattern of expression is not modified significantly in infected ripe berries to except of *SDG6934* (Fig. 7).

Discussion

Lysine methylation of histones plays an essential role in diverse biological processes ranging from transcriptional regulation to heterochromatin formation (Liu et al. 2010). In plants, methylation of histones has been demonstrated to control a range of processes such as gametogenesis, embryogenesis, seed development, flowering time, branching, and floral identity (Jarillo et al. 2009; Ahmad et al. 2010; Pontvianne et al. 2010). SDG proteins represent an evolutionarily-conserved family of epigenetic regulators, which are responsible for most histone lysine methylation. The arabidopsis and rice genomes encode 49 and 34 SDG proteins, respectively (Pontvianne et al. 2010; Ng et al. 2007). In this work, we identified the SDG genes in the fruit tree, *Vitis vinifera.* The search in the grapevine genome allowed the identification of 33 genes belonging to this family of chromatin remodeling factors. This number is also similar to that found in other organisms; by searching the protein databases 58, 47 and 29 SET domain proteins were obtained from zebrafish, human, and fruit fly, respectively (Sun et al. 2008).

Previous reports have shown that plant SDG proteins fall into seven classes, according to their sequence and domain architectures, and apparently the resulting groups reflect the substrate specificities of the members of each class. In this way, class I is related with H3K27, class II with H3K36, classes III and IV with H3K4 and class V with H3K9 (Liu et al. 2010; Ng et al. 2007). Furthermore,

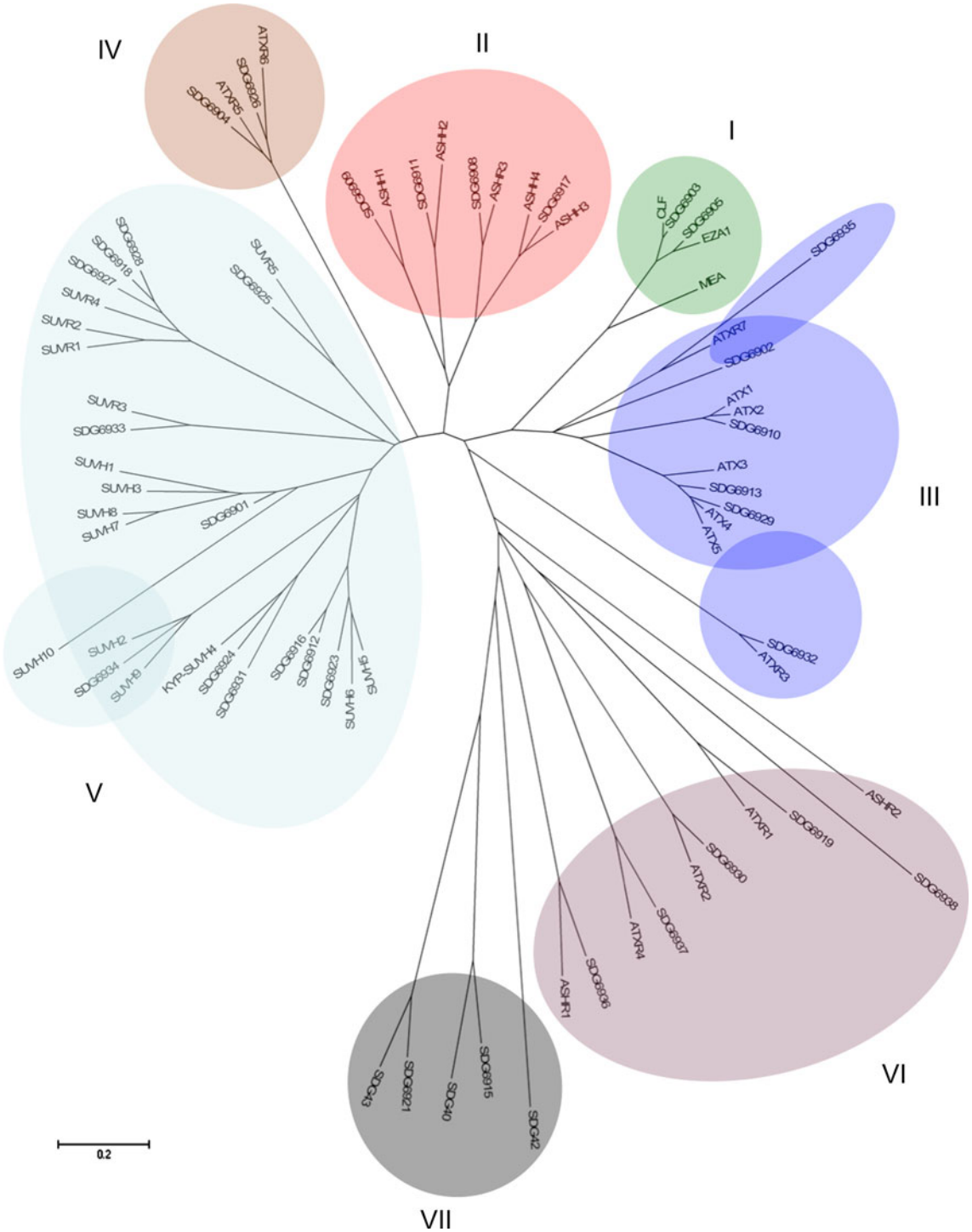

Fig. 2 Phylogenetic tree of the SDG family in grapevine and arabidopsis. The tree was constructed by the Neighbor-Joining method with MEGA program 4.0 using the conserved SET-domain region. The accession numbers of the grapevine and Arabidopsis SET domain sequences are shown in Table 1 and Springer et al. (2003), respectively

Class VI grouped proteins with either truncated or interrupted SET domains and their function has not yet been established in plants. Finally, class VII grouped SET domain proteins that methylate non-histone targets (Ng et al. 2007). The SDG proteins encoded in the grapevine genome are grouped into all of the classes identified in other species (Fig. 2).

The other domains present in many of the plant SDG proteins have been described to play a role in protein–protein interactions, indicating that the plant putative

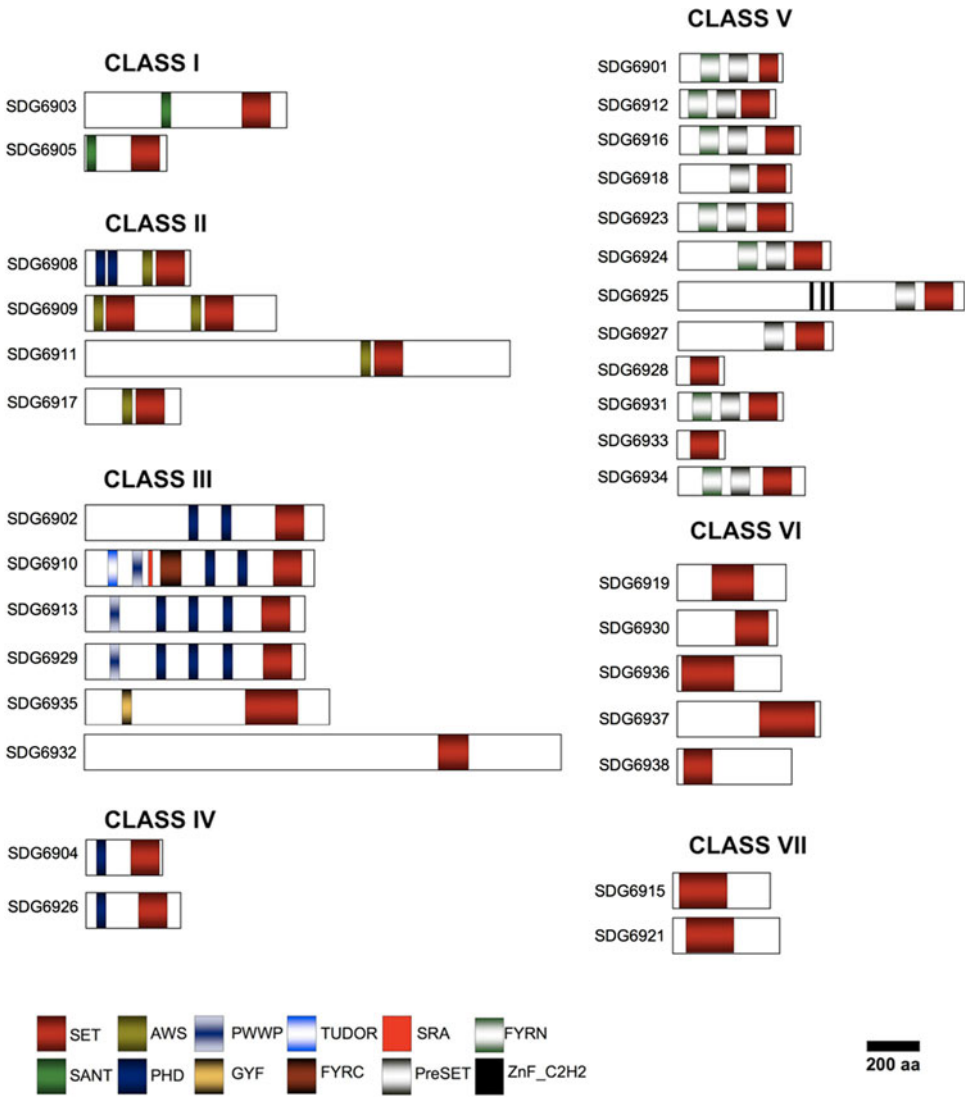

Fig. 3 Domain architecture of the different classes of grapevine SDG proteins. Domains are not drawn to scale

histone methyltransferases may act in protein complexes. The presence of PHD and PWWP domains suggest that these proteins may indeed form protein complexes. The PHD domain is a putative zinc finger that is involved in mediating protein–protein interactions (Aasland et al. 1995). The PWWP domain is also involved in mediating protein–protein interactions (Stec et al. 2000). The domains present in the N-terminal portion of SET domain proteins may therefore be important for determining interactions with other proteins.

The study of the SDG expression pattern is helpful in understanding epigenetic regulation in plant development. In this study, we found that selected SDGs were differentially expressed in root, leaf, flower, and berries (Fig. 4), which supports the assertion that expression of SDGs is related with development. In addition, the selected SDGs showed differential expression patterns during inflorescence and berry development (Figs. 5, 6). In these organs, the proliferation-differentiation cell transition, and gametogenesis and embryogenesis occur, respectively, processes that are under epigenetic control (Jarillo et al. 2009).

During pathogen infection, plants defend themselves through different signaling pathways that regulate numerous biochemical, metabolic, and molecular mechanisms to increase tolerance to adverse conditions. Recent studies have indicated that regulation of stress-responsive genes often depends on chromatin remodeling, that is, the process inducing changes in chromatin structure (Chinnusamy and Zhu 2009). In this work, we have shown for the first time that SDG gene expression is regulated in GLRaV-3-infected symptomatic leaves and fruits. It has been documented that GLRaV-3 can cause reduced plant vigor and longevity, and significant losses in both yield and quality of berries (Komar et al. 2010). Moreover, GLRaV-3 infection caused developmental problems. In advanced stages of infection, the margins of infected leaves roll downward,

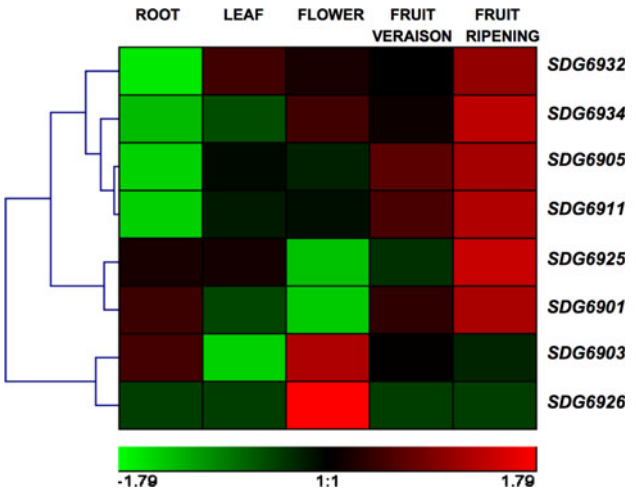

Fig. 4 Cluster analysis of selected grapevine SDG gene expression profiles in vegetative and reproductive organs. Expression analyses were performed by qRT-PCR. Transcript levels are expressed in relation to the *VvUBIQUITIN1* gene and relative gene expression data were gene-wise normalized. A *color scale*, representing signal values, is shown at the base of the figure

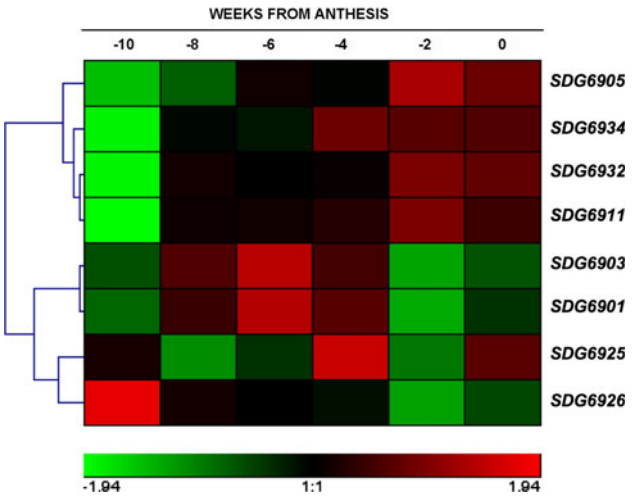

Fig. 5 Cluster analysis of selected grapevine SDG gene expression profiles during inflorescence development. Expression analyses were performed by qRT-PCR and expressed in relation to the *VvUBIQUITIN1* gene. Relative gene expression data were gene-wise normalized. A *color scale*, representing signal values, is shown at the base of the figure

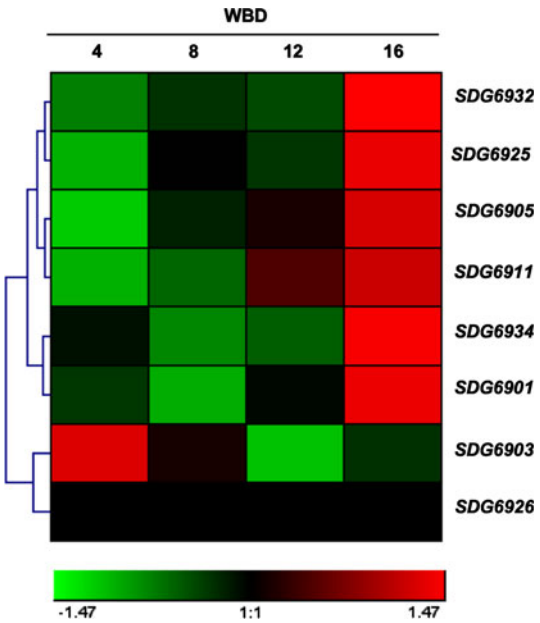

Fig. 6 Cluster analysis of selected grapevine SDG gene expression profiles during berry development. Expression analyses were performed by qRT-PCR and expressed in relation to the *VvUBIQUITIN1* gene. Relative gene expression data were gene-wise normalized. A *color scale*, representing signal values, is shown at the base of the figure. WBD (weeks berry development)

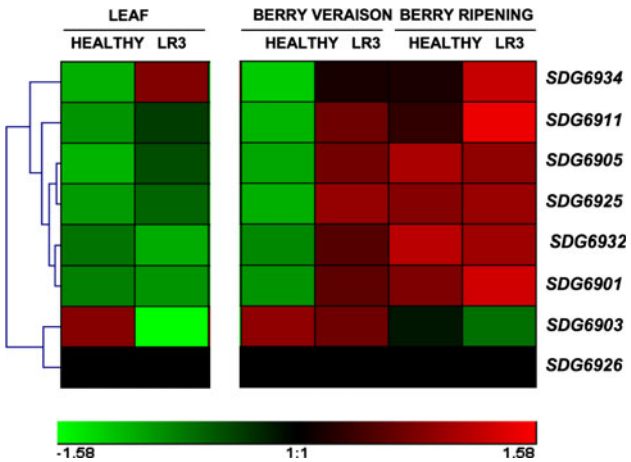

Fig. 7 Cluster analysis of selected grapevine SDG gene expression profiles in response to virus infection. Expression analyses were performed by qRT-PCR and expressed in relation to the *VvUBIQUITIN1* gene. Relative gene expression data were gene-wise normalized. *Color scale*, representing signal values, is shown at the base of the figure. LR3: GLRaV-3

developing the symptom that gives the disease its common name. Our results showed that the developmental changes caused by virus infection are associated with modifications in the expression of some SDGs. These results, together with our previous data that shows that viral infections induce changes in histone acetyltransferase gene expression (Espinoza et al. 2007), suggest that virus infection can modify the epigenetic program in the plant as a pathogenic mechanism. Further experiments could elucidate if the histone methylation pattern is modified during virus infection in grapevine.

In this paper, we have presented a genome analysis of grapevine SDG genes along with an account of their phylogenetic relationships with arabidopsis homologues and expression profiling for selected SDG genes. However, the molecular and functional characterization of the genes

described in this study and their role in grapevine development remain to be elucidated.

Acknowledgments This work was supported by the Chilean Wine Consortium 05CTE01-03, the Fruit Consortium, 07Genoma01, Fondecyt project 1100709 and Millennium Nucleus for Plant Functional Genomics (P06-009-F). F. Aquea is supported by a Postdoctoral Project "Programa Bicentenario de Ciencia y Tecnología/ CONICYT-BancoMundial" PSD74-2006. We thank Michael Handford for assistance in language support. M.J. Poupin is supported by a Postdoctoral project FONDECYT 3100040 and a CONICYT project PAI 79090016.

References

Aasland R, Gibson TJ, Stewart AF (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends Biochem Sci 20:56–59

Ahmad A, Zhang Y, Cao XF (2010) Decoding the epigenetic language of plant development. Mol Plant 3:719–728

Alvarez-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 13:627–637

Aquea F, Timmermann T, Arce-Johnson P (2010) Analysis of histone acetyltransferase and deacetylase families of *Vitis vinifera*. Plant Physiol Biochem 48:194–199

Baumbusch LO, Thorstensen T, Krauss V, Fischer A, Naumann K, Assalkhou R, Schulz I, Reuter G, Aalen RB (2001) The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionarily conserved classes. Nucleic Acids Res 29:4319–4333

Berr A, McCallum EJ, Ménard R, Meyer D, Fuchs J, Dong A, Shen WH (2010) Arabidopsis SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. Plant Cell 22:3232–3248

Cazzonelli CI, Cuttriss AJ, Cossetto SB, Pye W, Crisp P, Whelan J, Finnegan EJ, Turnbull C, Pogson BJ (2009) Regulation of carotenoid composition and shoot branching in Arabidopsis by a chromatin modifying histone methyltransferase, SDG8. Plant Cell 21:39–53

Childs KL, Hamilton JP, Zhu W, Ly E, Cheung F, Wu H, Rabinowicz PD, Town CD, Buell CR, Chan AP (2007) The TIGR plant transcript Assemblies database. Nucleic Acids Res 35:D846–D851

Chinnusamy V, Zhu JK (2009) Epigenetic regulation of stress responses in plants. Curr Opin Plant Biol 12:133–139

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

Downey M, Harvey J, Robinson S (2003) Synthesis of flavonols and expression of flavonol synthase genes in the developing grape berries of Shiraz and Chardonnay (*Vitis vinifera* L.). Aust J Grape Wine R 9:110–121

Espinoza C, Vega A, Medina C, Schlauch K, Cramer G, Arce-Johnson P (2007) Gene expression associated with compatible viral diseases in grapevine cultivars. Funct Integr Genomics 7:95–110

Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol 12:1052–1058

Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G (1997) A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature 386:44–51

Grini PE, Thorstensen T, Alm V, Vizcay-Barrena G, Windju SS, Jørstad TS, Wilson ZA, Aalen RB (2009) The ASH1 HOMOLOG 2 (ASHH2) histone H3 methyltransferase is required for ovule and anther development in Arabidopsis. PLoS One 4:e7817

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB (1998) Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. Science 280:446–450

Guo L, Yu Y, Law JA, Zhang X (2010) SET DOMAIN GROUP2 is the major histone H3 lysine 4 trimethyltransferase in Arabidopsis. Proc Natl Acad Sci USA 107:18557–18562

Ho L, Crabtree GR (2010) Chromatin remodelling during development. Nature 463:474–484

Jacob Y, Stroud H, Leblanc C, Feng S, Zhuo L, Caro E, Hassel C, Gutierrez C, Michaels SD, Jacobsen SE (2010) Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. Nature 466:987–991

Jaillon O et al (2007) French–Italian Public Consortium for grapevine genome characterization the grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. Nature 449:463–467

Jarillo JA, Piñeiro M, Cubas P, Martínez-Zapater JM (2009) Chromatin remodeling in plant development. Int J Dev Biol 53:1581–1596

Jenuwein T, Laible G, Dorn R, Reuter G (1998) SET domain proteins modulate chromatin domains in eu- and heterochromatin. Cell Mol Life Sci 54:80–93

Komar V, Vigne E, Demangeat G, Lemaire O, Fuchs M (2010) Comparative performance analysis of virus-infected *Vitis vinifera* cv. Savagnin rose grafted onto three rootstocks. Am J Enol Vitic 61:68–73

Liu C, Lu F, Cui X, Cao X (2010) Histone methylation in higher plants. Annu Rev Plant Biol 61:395–420

Matus JT, Aquea F, Arce-Johnson P (2008) Analysis of the grape MYB R2R3 subfamily reveals expanded wine quality-related clades and conserved gene structure organization across Vitis and Arabidopsis genomes. BMC Plant Biol 8:83

Matus JT, Poupin MJ, Cañón P, Bordeu E, Alcalde JA, Arce-Johnson P (2010) Isolation of WDR and bHLH genes related to flavonoid synthesis in grapevine (Vitis vinifera L). Plant Mol Biol 72:607–620

Naumann K, Fischer A, Hofmann I, Krauss V, Phalke S, Irmler K, Hause G, Aurich AC, Dorn R, Jenuwein T, Reuter G (2005) Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in Arabidopsis. EMBO J 24:1418–1429

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 1769:316–329

Pontvianne F, Blevins T, Pikaard CS (2010) Arabidopsis histone lysine methyltransferases. Adv Bot Res 53:1–22

Poupin MJ, Federici F, Medina C, Matus JT, Timmermann T, Arce-Johnson P (2007) Isolation of the three grape sub-lineages of B-class MADS-box TM6, PISTILLATA and APETALA3 genes which are differentially expressed during flower and fruit development. Gene 404:10–24

Qian C, Zhou MM (2006) SET domain protein lysine methyltransferases: structure, specificity and catalysis. Cell Mol Life Sci 63:2755–2763

Reid KE, Olsson N, Schlosser J, Peng F, Lund ST (2006) An optimized grapevine RNA isolation procedure and statistical determination of reference genes for real-time RT-PCR during berry development. BMC Plant Biol 6:27

Springer NM, Napoli CA, Selinger DA, Pandey R, Cone KC, Chandler VL, Kaeppler HF, 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

Stec I, Nagl SB, van Ommen GJ, den Dunnen JT (2000) The PWWP domain: a potential protein-protein interaction domain in nuclear proteins influencing differentiation? FEBS Lett 473:1–5

Sturn A, Quackenbush J, Trajanoski Z (2002) Genesis: cluster analysis of microarray data. Bioinformatics 18(1):207–208

Sun XJ, Xu PF, Zhou T, Hu M, Fu CT, Zhang Y, Jin Y, Chen Y, Chen SJ, Huang QH, Liu TX, Chen Z (2008) Genome-wide survey and developmental expression mapping of zebrafish SET domain-containing genes. PLoS One 3:e1499

Tamada Y, Yun JY, Woo SC, Amasino RM (2009) ARABIDOPSIS TRITHORAX-RELATED7 is required for methylation of lysine 4 of histone H3 and for transcriptional activation of FLOWERING LOCUS C. Plant Cell 21:3257–3269

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Velasco R et al (2007) A high quality draft consensus sequence of the genome of a heterozygous grapevine variety. PLoS One 2:e1326

Zhao Z, Yu Y, Meyer D, Wu C, Shen WH (2005) Prevention of early flowering by expression of FLOWERING LOCUS C requires methylation of histone H3 K36. Nat Cell Biol 7:1256–1260