Architectural and functional diversity of Polycomb group Response Elements in Drosophila

J. Lesley Brown and Judith A. Kassiss

Program on Genomics of Differentiation, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892-2785
Running title: characteristics of *engrailed* PREs

Key words: Polycomb, PREs, transcriptional repression

Corresponding author:

Judith A. Kassis

6 Center Dr. MSC 2785

Bldg. 6B, Rm. 3B-331

Bethesda, MD 20892-2785

Phone: 301-496-7879

e-mail: jkassis@mail.nih.gov
ABSTRACT

Polycomb group response elements (PREs) play an essential role in gene regulation by the Polycomb group (PcG) repressor proteins in Drosophila. PREs are required for the recruitment and maintenance of repression by the PcG proteins. PREs are made up of binding sites for multiple DNA binding proteins but it is still unclear what combination(s) of binding sites is required for PRE activity. Here we compare the binding sites and activities of two closely linked yet separable PREs of the Drosophila engrailed (en) gene, PRE1 and PRE2. PRE1 and PRE2 both contain binding sites for multiple PRE-DNA binding proteins but the number, arrangement, and spacing of the sites differs between the two PREs. These differences have functional consequences. PRE1 and PRE2 both mediate pairing-sensitive silencing of mini-white, a functional assay for PcG repression, however, PRE1 requires two binding sites for Pleiohomeotic (Pho), whereas PRE2 requires only one Pho binding site for this activity. Further, for full pairing-sensitive silencing activity, PRE1 requires an AT-rich region not found in PRE2. These two PREs behave differently in a PRE embryonic and larval reporter construct inserted at an identical location in the genome. Our data illustrate the diversity of architecture and function of PREs.

INTRODUCTION

The Polycomb Group (PcG) and Trithorax group (TrxG) proteins are key regulators of genomic programming and differentiation in multi-cellular organisms. In Drosophila, PcG proteins work as multi-protein complexes recruited to specific regions in the DNA (PREs, Polycomb group Response Elements), to heritably repress target gene
expression through post-translational covalent modifications of histones and modulation of chromatin structure. Three major complexes have been extensively discussed (PRC1 and PRC2, Polycomb Repressive Complex 1 and 2 and Pho-RC) but as more different types of tissue, developmental stages and purification techniques are used a number of other complexes have been identified (PR-DUB (Scheuermann et al. 2010); dRAF; Pcl-PRC2 (reviewed in Simon and Kingston 2009; Kerppola 2009; Müller and Verrijzer 2009).

Of the characterized Drosophila PcG complexes, only Pho-RC contains a DNA binding PcG protein (Pho (Pleiohomeotic) or Phol (Pho-like)) that binds directly to PREs (Brown et al. 1998, 2003). Pho-RC consists of Pho or Phol and dSfmbt (Klymenko et al. 2006; Grimm et al. 2009). Scm (Sex combs on midleg) has genetically and physically been shown to interact with dSfmbt (Grimm et al. 2009) and it has been proposed that Scm may interact with an as yet unidentified DNA binding protein and cooperate with Pho-RC to recruit PRC1 and PRC2 (Wang et al. 2010). PRC2 is comprised of E(z) (Enhancer of zeste), Esc (Extra sex combs), Su(z)12 (Suppressor of zeste 12) and Nurf55. PRC2 tri-methylates H3K27 through the histone methyltransferase activity of the E(z) SET domain to set a repressive mark (Cao et al. 2002; Czermin et al. 2002; Müller et al. 2002). A variant of PRC2, Pcl-PRC2, is required for high levels of H3K27me3 (Nekrasov et al. 2007). PRC1 core complex is comprised of Ph (Polyhomeotic), Psc (Posterior sex combs), Pc (Polycomb) and Sce (Sex combs extra) also known as dRing (Shao et al. 1999; Fritsch et al. 2003). dRing/Sce has ubiquitin ligase function (Wang et al. 2004). dRaf, is a variant of PRC1 that has demethylase activity and removes an active mark
from histone H3K36me2 and replaces it with repressive ubiquitylation at H2A-K118 (Lagarou et al. 2008). PR-DUB has deubiquitinase activity in vitro (Scheuermann et al. 2010). Multiple orthologs of the PcG genes are present and play major roles in differentiation, stem cell maintenance, and cancer in mammals (reviewed in Gieni and Hendzel 2009).

Central to PcG repression is the recruitment of PcG proteins to the PRE (reviewed in Müller and Kassis 2006; Schuettengruber, et al. 2007; Kassis and Brown 2013). PREs from a number of genes have been closely examined, including four PREs from the engrailed (en) and invected (inv) genes (Americo et al. 2002; Cunningham et al. 2010), the bxd, iab2, and iab7 (also known as Fab-7) PREs from the bithorax complex (BX-C; Hagstrom et al. 1997; Shimell et al. 2000; Mishra et al. 2001; Busturia et al. 2001; Dejardin et al. 2005) and an even-skipped (eve) PRE (Fujioka et al. 2008). PREs are made up of binding sites for many different proteins including Pho/Phol (Brown et al. 1998; 2003; Fritsch et al. 1999), Spps (Sp1 factor for pairing-sensitive silencing, Brown and Kassis 2010), GAGA Factor (GAF), and Dsp1. Grainy head (Grh), and Zeste have also been implicated in PRE function (reviewed in Müller and Kassis 2006; Schuettengruber, et al. 2007; Kassis and Brown 2013). Genome-wide studies show that not all of these factors are bound to every PRE although Pho/Phol seems to be a key component of most PREs. Genome-wide studies on Spps have not yet been reported.

Despite knowing a number of factors that bind directly to PREs, it is not possible to identify PREs based on DNA sequence alone. PRE predictions based on clustering of DNA binding sites within a given region only identify about 10-15% of PREs identified by genome-wide ChIP experiments with components of the PcG complexes (Ringrose et
al. 2003; Fiedler and Rehmsmeier 2006; Schuettengruber et al. 2009). Some PRE-binding proteins are still not identified (Americo et al. 2002). It may be that there is no one size fits all formula for PREs. Differences in DNA sequence and transcription factor binding site composition might be important for different types of PREs and allow PcG activity to respond to different cell type specific cues. Genome-wide ChIP has shown differences in the binding patterns of PcG proteins/complexes in different cells types (Negre et al. 2006; Schwartz et al. 2006; Tolhuis et al. 2006; Oktaba et al. 2008; Kwong et al. 2008; Schuettengruber et al. 2009). Initially thought to be a simple off/on switch, regulation by the PcG proteins has the ability to dynamically respond to changing needs during development.

*en* is a segment polarity gene that is a target of PcG repression (Moazed and O’Farrell, 1992). The regulatory sequences of *en* span over 70kb. Embryonically, *en* is expressed in a complex manner, in stripes, in the peripheral and central nervous system, the fat body and portions of the head. In larvae, *en* is expressed in the posterior compartments of imaginal discs and in a subset of cells of the nervous system. A 2kb region extending from -2.4kb to -395bp upstream of the *en* transcription start site contains two PREs (PRE1 and PRE2) (DeVido et al. 2008; see Figure 1). This 2kb piece of DNA has PRE, pairing-sensitive silencing (PSS), and homing ability (Kassis 1994, 2002; DeVido et al. 2008; Cheng et al. 2012). In addition to silencing, these PREs can act with distant enhancers to facilitate transcriptional activation (DeVido et al. 2008; Kwon et al. 2009). Here we focus our efforts on analyzing different aspects of the *en* PREs and determining what DNA sequences are needed to constitute specifically an *en* PRE. Our data show there are differences in the DNA sequence requirements of the two closely
linked PREs in the pairing-sensitive silencing assay. Further, these two PREs behave differently in embryonic and larval functional assays. Our data refute the idea that all PREs are the same, an important point in understanding recruitment and functioning of the PcG system of repression at such complex developmental target genes.

**MATERIALS AND METHODS**

**P-element transgene constructs:** These constructs were made as described in Americo *et al.* 2002. Mutations of individual sites within the PREs were generated by PCR with mutated oligonucleotides. The oligonucleotides incorporated EcoR1 and BamH1 restriction endonuclease site ends to facilitate cloning. The PCR products were cloned into the EcoR1-BamH1 sites of *pCaSpeR* (Pirrotta 1988). The clones were confirmed by sequencing. Transgenic fly lines were generated by Genetic Services Inc.

**Cloning and injection of φC31 constructs:** PRE1 or PRE2 DNA was amplified using oligonucleotides with XbaI restriction site ends and cloned into the XbaI site of the PRE300-∗bx-D-Ubx-Z plasmid (Fujioka *et al.* 2008) in place of the *eve* PRE. The PREs were cloned in the same orientation with respect to the *Ubx* promoter as they are with respect to the *en* promoter at the endogenous *en* locus. RMCE was performed as described by Bateman *et al.* 2006 into the attP site located at 52D (injections by Genetic Services Inc). The integrity of the RMCE products were confirmed by PCR and the constructs used here were all inserted in the same orientation to minimize variability.

**Gel mobility shift assays:** These were performed as described in Americo *et al.* 2002 except they were carried out using 3μl of commercially prepared SL2 extracts from Active Motif in place of embryonic nuclear extracts or with full length recombinant Pho protein synthesized in vitro using the TNT-coupled transcription/translation system from
Promega. The sequences of the Pho 1-4 probes are as follows; Pho 1
AAAGGCAGCCATTTC; Pho 2 CACATGGCCATCTTTTC; Pho 3
GCCAGCCATTGTGTCA; Pho 4; GTCAGCCATTTAAAGTC

Searching known PREs for predicted protein binding sites: The en, eve and
iab7/Fab-7 PREs were searched for the presence of consensus binding sites using

Cross-linked Chromatin Immunoprecipitation (X-ChIP): Imaginal discs along with
the central nervous system, mouth hooks and some anterior cuticle were dissected from
3\textsuperscript{rd} instar larvae (20 larvae per sample) and immediately placed in Schneider’s medium
(Invitrogen), on ice. The disc sets were fixed in 2\% formaldehyde (Ted Pella Inc, Redding, CA) fixing solution (50mM Hepes pH 7.6, 100mM NaCl, 0.1mM EDTA,
0.5mM EGTA) for 15 minutes, then rinsed in stop solution (PBS, 0.01\% Triton X-100,
0.125M Glycine) for 10 minutes, followed by 2 x 10 minute washes with wash solution
(50mM Tris-HCl pH 8.0, 10mM EDTA, 0.5mM EGTA, 0.25\% Triton X-100). Fixed and
washed discs were stored at -80 degrees Celsius in storage solution (10mM Tris-HCl pH
8.0, 1mM EDTA, 0.5mM EGTA). The storage buffer was replaced with 300\(\mu\)l of action
buffer (50mM Tris pH 7.4, 1mM EDTA), supplemented with Complete Protease
Inhibitor Cocktail (Roche) in a 1.5ml microcentrifuge tube, and sonicated in BioRuptor
(Diagenode, Denville, NJ) for 30 seconds on/ 30 seconds off 20 cycles at high power.
The sonication resulted in chromatin fragments tightly concentrating at 200 base pairs,
with a diminishing smear up to 1500 base pairs. Remaining insoluble material was
removed by centrifugation and the chromatin supernatant was transferred to a new tube.
1\% of total volume was saved from each sample for input reactions. ChIP was performed
with 1:200 dilution of the required antibody (Grh antibody, Kim and McGinnis 2011; H3K27me3 antibody, Millipore) and the Protein A agarose Chromatin Immunoprecipitation Assay Kit from Millipore according to the manufacturers instructions. Crosslinks were reversed by incubating at 65° C for 4 hours with 100mM NaCl. All samples were then purified with standard phenol/chloroform extraction. DNA samples were ethanol precipitated overnight, washed with 75% ethanol, and resuspended in 100μl of water.

**qPCR analysis of X-ChIP:** ChIP samples were analyzed by qPCR using a Lightcycler 480 Real-Time PCR System (Roche Applied Science) and Lightcycler 480 DNA SYBR Green I Master Mix (Roche Applied Science), according to manufacturer instructions. All samples are given as percent input. Each samples was analyzed in triplicate.

**Primer sets:**

**PRE1 F:** TGAACAGCTCAGATGCATAAATTG  
**PRE1 R:** CAGACTGGAAGGTGCGTTTC  
**PRE2 F:** GCTTATGAAAAGTGTCTGTG  
**PRE2 R:** GGGGCTTGTTAGGCAGCAAT  
En control F: CGCCTTAAGGTGAGATTCAGTT  
En control R: GGCGGTGTCAATATTITGGT  
**PRE_D F:** CGAAATGCTACTGCTCTCTA  
**PRE_D R:** GCGTAGTCTTTATCTGTATCT  
Ubx control F: CCAGCATAAAACCGAAAGGA  
Ubx control R: CGCCAAACATTGAGATAG  
**PRE/bxd for PRE1 F:** GCATCTGCGCTGTATCA
PRE/bxd for PRE2 F: CGGTAACGCCCGGTAG
PRE/bxd R: TCGAGCAGCCCTGTG
Bxd/Ubx F: GCCACATTCCAGTTTCGACTTGC
Bxd/Ubx R: AGTTCTTGCTACATCATGCAGTTATTG
LacZ F: CAGCCGTTTGCCGTCTGAATTTGA
LacZ R: TGGAAATCGCTGATTTGCGTGGTC

RESULTS

Analysis of predicted protein binding sites in en PREs: PREs were first identified as fragments of DNA from the BX-C that could mediate repression by the PcG genes in embryonic reporter constructs (Müller and Bienz 1991; Simon et al. 1993). The ability of a fragment of DNA to mediate PcG-repression in a reporter construct is known as the PRE-maintenance assay. PREs act in transgenes to mediate the phenomenon of pairing-sensitive silencing (PSS) of the mini-white gene (Kassis et al. 1991; Kassis 1994; Kassis 2002). While there is some controversy as to whether all fragments of DNA that mediate PSS are PREs (Kassis 2002), in our studies of PREs from the en/inv region we have found a 100% correlation of fragments that demonstrate PRE activity in the maintenance assay and fragments that give PSS (Americo et al. 2002; Brown et al. 2005; Cunningham et al. 2010). Here, we assay PSS of mini-white in transgenic lines carrying PREs with mutated binding sites to assess if a given binding site might play a role in PRE activity. This is a P-element based system with random integration into the genome. The ability to mediate PSS relies to some extent on the site of insertion and only around 60% of insertion sites give PSS with the wild type PRE.
To date, four well defined binding sites in PRE2 have been shown to be required for its PSS activity; these are binding sites for the proteins Pho/Phol, Spps, GAF and Dsp1 (Brown et al. 1998, 2003; Americo et al. 2002; Decoville et al. 2001; Dejardin et al. 2005). Chromatin-immunoprecipitation studies show that these 4 proteins localize to the en PREs as well as to many other PREs in the genome (Klymenko et al. 2006; Brown and Kassis 2010; Schwartz et al. 2006; Schuettengruber et al. 2009). Further, PSS is disrupted in pho, Dsp1, and Spps mutants, showing that these proteins are required for the PSS activity of PRE2 (Brown et al. 1998; Dejardin et al. 2005; Brown and Kassis 2010).

Here we define other binding sites required for PSS activity of PRE2 and compare the results with sequences required for PRE1 activity. Figure 1 shows the locations of PRE1 and PRE2 in relation to the en transcription start site and the locations and numbers of putative binding sites for proteins implicated in PRE activity. We included in our analysis the consensus sequences for the Grh and Zeste proteins as both have been implicated in the activity of some PREs (Tuckfield et al. 2002; Blastak et al. 2006; Saurin et al. 2001; Hur et al. 2002; Mulholland et al. 2004).

**Identification of Site A:** In previous studies we identified a protein-binding site within an 18bp region of PRE2 (AGAGGAGTGAACTGC), (Americo et al. 2002 and Fig. 2). Mutation of the central region of this site (AGTGA to TACCT) caused a complete loss of PSS (Americo et al. 2002). Closer examination of this 18bp sequence showed it contains a match to the consensus binding sequence for the Zeste protein (the sequence GGAGTGA matches the Zeste consensus site BGAGTGV) (see Fig. 2A). We used gel mobility shift analysis and competition with oligonucleotides to determine the critical
nucleotides for binding to the 18bp oligonucleotide. Our data showed there was binding
to a sequence (Site A) bound by an unknown protein (Protein A) that overlaps the Zeste
consensus sequence (Figure 2). The Protein A gel mobility shift was not competed by an
oligonucleotide containing known Zeste binding sites (CGAGTGAGTGTGAGTG; Figure 2B). Further, mutation of the Zeste consensus sequence within the 18bp sequence
(Mut Z1) still competed the gel shift, whereas mutation of Site A disrupted the
competition, indicating these were the critical nucleotides for binding (Figure 2C). We
introduced a mutation that disrupts Site A in PRE2 and assayed this mini-white reporter
construct in transgenic flies for the ability to give PSS (Fig. 2D). When Site A was
mutated, PSS was reduced from 66% of the lines to 13% of the lines (comparable to what
we see with mutation of either the Pho or the Spps sites, Americo et al. 2002; Brown et
al. 2005). We have identified the potential core-binding site of protein A to be
GAACAG. This is where the mutations of the binding site are centered and, in addition,
another oligonucleotide containing GAACAG was able to compete the site A gel
mobility shift (data not shown). We are currently trying to identify the protein that binds
to Site A.

**Zeste/Fs(1)h binding sites:** The Zeste consensus binding site CGAGTG has also been
shown to bind the protein Fs(1)h (female sterile (1) homeotic; Chang et al. 2007). PRE2
has two matches to this consensus sequence, the one next to Site A and one in the middle
of the fragment (Z2, CGAGTG). Mutation of both matches to the Zeste consensus
sequence reduced PSS to 19%. Mutation of the adjacent Zeste and A sites together gave
no PSS, more severe than either of the mutated sites alone. Thus, both of these consensus
sequences contribute to the PSS activity of PRE2. Note that in our gel mobility shift
assays we do not detect a shift due to a factor binding to the Zeste consensus sequence (data not shown). The protein that binds this site may not be present in the S2 extracts used for these experiments or may require cooperative interaction of more than one binding site as has been shown for Zeste binding at other regulatory regions (Benson and Pirrotta 1988).

**Zeste mutants do not affect PSS by PRE2:** Zeste has been implicated in PRE activity due to its association with PRC1 (Saurin *et al.* 2001; Hur *et al.* 2002; Mulholland *et al.* 2003) but it has also been linked to trxG-mediated activation (Dejardin and Cavalli 2004). Genome-wide ChIP-chip studies show a limited overlap of Zeste protein binding with PcG protein binding in embryos but Zeste was bound to the *en* PREs during embryogenesis (Moses *et al.* 2006). We tested whether a mutation in zeste (*z*) affected PSS by PRE2; in heterozygous and homozygous *z* mutants there was no change in pairing-sensitive silencing of mini-white in flies carrying PRE2-mini-white constructs (data not shown). Therefore, Zeste is not the protein working at these sites in later development. It is possible that Fs(1)h is the protein that acts here (Chang *et al.* 2007).

**Grh:** Grh has been described as binding to PREs and to cooperatively interact with Pho. The binding site for Grh identified in the *iab-7* PRE was reported to be TGTTTTTTTT (Blastyák *et al.* 2006). However, other groups report a Grh consensus sequence as WCHGGTT (where W is A or T and H is not G) (Venkateson *et al.* 2003; Almeida and Bray, 2005). We did not find any match to the TGTTTTTTTT sequence within *en* PRE2 however, there is a potential match to WCHGGTT towards the end of PRE2 (Figure 1). The sequence is ACCGGCT, the one position that does not match is underlined. This sequence lies in a region that we have deleted from PRE2 in previous studies and still
saw PSS (139bp PRE, shown in Figure 2D, Americo et al. 2002). The 139bp PRE was functional but not as robust as the entire 181bp region. There is a potential match to the Grh consensus in PRE1 (see below). Chromatin immunoprecipitation with an antibody against Grh showed binding of Grh to each of these PREs in tissue derived from 3rd instar larvae (Fig. 2E). Grh is the only PRE binding protein identified so far that does not have ubiquitous expression.

**Other factors:** We previously showed specific binding of two other proteins (B and C) to sites within PRE2 but we did not know if they were needed for PSS activity (Americo et al. 2002). Mutation of one of these sites (site C) does not impact PSS (data not shown). The binding site for protein B has been narrowed down to a 12bp region (TGCCGCTATATG) but mutations of this site have not yet been tested in the PSS assay. The positions of these two binding sites relative to the other protein binding sites in PRE2 are shown in Figure 1A.

Some genome-wide ChIP experiments showed enrichment at PREs for TTG repeats and others found an enrichment of GT repeats (Ringrose et al. 2003; Schuettengruber et al. 2009). Deletion of the GTGT sequences in the PRE of the vestigial (vg) gene led to a reduction of silencing activity of this fragment suggesting that this sequence may play a role in PcG repression (Okulski et al. 2011). There is only one TGTGT sequence in PRE2 that lies in the region of the 181bp PRE that can be deleted without affecting PSS. There are 3 sequences in PRE1 that show GT repeats (GTGTGT or GTGTG). These sites have not been tested within the context of PRE1. To date, the protein that binds to GT repeats has not been identified.
**PRE1**: PRE1 originally referred to a 2kb region from -2.4kb to -576bp upstream of the *en* transcription start site. This region acts as a PRE, mediates PSS, and contributes to homing (Kassis 1994, DeVido *et al.* 2008, Cheng *et al.* 2012). Here we wanted to define the minimal fragment for PRE1. A 441bp region (-1944 to -1503) encompasses a region showing the most clustering of binding sites associated with PSS/PRE activity (DeVido *et al.* 2008). In a previous study this fragment gave PSS of mini-white in 3/5 lines but constructs with the region -1944 to -1671bp (273bp, PRE1.1) or with -1672 to -1503 (169bp, PRE 1.2) did not (Kassis 1994, Fig. 1A). We generated additional lines of PRE1 and PRE1.1 and found that 12/22 PRE1 lines gave PSS while 0/19 PRE1.1 lines gave PSS. From here on we use PRE1 to refer to the 441bp region from -1944 to -1503bp upstream of the *en* transcription start site.

**PRE1 and PRE2 differ in their requirements for the number of Pho binding sites for PSS**: Mohd-Sarip *et al.* (2002, 2005, 2006) proposed a model requiring two Pho sites for in vivo chromatin association of Pho and the Polycomb Core Complex (PCC) with the *iab*-7 and *bxd* PREs. They proposed that sequences around the core consensus Pho binding site contribute on one side to enhancing PHO binding and on the other side to cooperative binding of Pho and the PRC1 core complex. The consensus sequence they derived for Pho sites is **CRGCCATYDTRGD** (R is A or G; D is A, T, or G). CR is called the U sequence and is thought to enhance Pho binding. RGD is the consensus for the site that is required for association of PCC. When RGD was mutated, both the cooperative binding of Pho with PCC and PSS mediated by the *iab*-7 PRE were abolished (Mohd-Sarip *et al.* 2005).
We have previously shown that mutation of a Pho binding site (Pho1 in Fig. 3) greatly diminishes PSS by PRE2 (Brown et al. 1998). PRE2 has a second potential Pho binding site (Pho2), on the 5’ end of the PRE (Fig. 3). However, when this site is used as a competitor in a gel shift assay with in vitro transcribed/translated full length Pho bound to the Pho1 site there is very little competition (Figure 3B). Conversely, when the labeled Pho2 oligo was used in the gel mobility shift assay, no specific gel shift was detected (data not shown). Thus, this second Pho site is at best a very weak binding site for Pho compared to the Pho1 site. However, in the context of the PRE, perhaps a weak site can bind Pho in combination with a high affinity-binding site.

We mutated the Pho2 binding site in the context of PRE2 in a transformation vector carrying the mini-white gene and looked for PSS. PSS was unaffected in lines where the Pho1 site was mutated (88% PSS for mutated Pho1 compared to 66% for wild-type PRE2, Fig. 3C). Thus, PRE2 does not require two Pho sites for PSS activity.

There are two predicted Pho sites in PRE1 and both compete as well as the Pho1 site in a gel mobility shift assay using in vitro transcribed/translated full length Pho protein (Fig. 3B). We note that Pho1, 3 and 4 all give good matches to the U site, whereas Pho2 does not (Fig. 3A). This data supports the view that the U site enhances Pho binding. In contrast, none of the Pho sites contain a perfect match to the RGD sequence required for recruitment of PCC in vitro (Mohd-Sarip et al. 2005). We introduced mutations that disrupt Pho binding into the Pho3 and Pho4 sites individually and together in PRE1. Disruption of either the Pho3 or Pho4 binding sites decreased the PSS activity of the fragment (Fig. 3D). When both Pho sites were mutated we saw no PSS. From this, we conclude that the PRE1 requires two Pho binding sites for full
function. Mutation of either site leads to only partial loss of activity. Similarly, mutation of the single Pho site in PRE2 does not completely abrogate function. We suggest that in PRE2 another DNA binding protein plays the role of the second Pho site for PcG complex recruitment.

**An AT-rich region is required for full activity of PRE1:** PRE1.1 contains binding sites for all known factors required for the activity of PRE2, but it has only one of the two Pho sites required for PRE1 activity (Fig. 1, 3). This led to the prediction that adding back the piece of DNA that contains the 2\(^{\text{nd}}\) Pho site to PRE1.1 might restore full PSS activity to PRE1.1. Surprisingly, adding back the second Pho site only gives PSS in 11% of the lines, no better than having a single Pho site (Fig. 4A). Clearly something else is required for the PSS activity of PRE1. We then added back more pieces of DNA from this fragment to see what else was required to recover full PSS activity.

The sequences added back contain consensus-binding sites for a number of factors believed to be involved in PRE activity (Sp1, Site A, Grh, Zeste site) (Fig. 4A). As these binding sites are added back, PSS increased from 11% (with the 2\(^{\text{nd}}\) Pho site) to 21% but full PSS activity was not restored until we included the entire PRE1 fragment (54%). The last sequence added back did not contain any matches to the binding sites previously identified at the Drosophila PREs. This region of DNA is extremely AT rich (72%) and is perhaps needed for structural reasons such as the ability of the DNA to bend and facilitate long-range interactions (Figure 4B). The sequence contains 3 repeats of ATTTAT. Gel mobility shift experiments show that at least one factor specifically binds to this sequence (Fig. 4C).
Studies of evolutionary conservation of transcription factor binding sites in the 12 sequenced species of Drosophila has proven to be a very valuable tool in the study of enhancers however it has not been as useful in the study of PREs. PREs seem to have built in plasticity in the arrangement of potential binding sites making them very hard to identify when using a best fit analysis based on homology over large regions (Dellino et al. 2002; Hauenschild et al. 2008). Comparison of the en 2.4kb region encompassing PRE1 and PRE2 over the 12 Drosophila species showed little conservation between all 12 species. However, when the equivalent region of each species is searched in GenePalette for consensus sequences for the factors we have discussed here many of the sites are present. The Pho and GAF/PSQ sites are the most conserved features. Kassis et al. (1989) compared the upstream region of the en gene between D. melanogaster and D. virilis. The corresponding region for PRE2 in D. virilis was identified and shown to give PSS (Kassis 1994) implying that this element most likely contains all the protein-binding sites necessary for PRE function. Using GenePalette we searched the D. virilis PRE2 sequence for the binding sites that we have studied here. The results are shown in Figure 5A. Between D. melanogaster and D. virilis there is conservation of the Pho and overlapping Dsp1 site, SiteA, SiteB, Spps and GAF/PSQ. There is conservation of a sequence that we have designated site B (GCCGCT) that gives a specific gel shift with Drosophila embryonic extracts (Americo et al. 2002). Not conserved are the Zeste consensus binding sites again implying that Zeste is most likely not the protein that binds to these sites. We extended this analysis to look for conservation of PRE1, although we note that this DNA fragment has not been tested for PRE activity. In PRE1 the order and spacing of the predicted binding sites is not conserved, although many of the binding
sites are present. Further, there is an AT rich region at the end of the predicted D. virilis PRE1 including multiple copies of the ATTTAT sequence (Figure 5B). Without further mutational analysis, it is not possible to distinguish between a requirement for general AT richness and a specific protein binding DNA sequence.

**The activity of PRE1 and PRE2 differs in a bxd-Ubx-lacZ reporter construct:** We have previously shown that PRE2 can act as a PRE-maintenance element in an embryonic bxd-Ubx-lacZ P-element reporter construct (Americo et al. 2002). Further, both PRE2 and PRE1 (contained within a 2kb fragment) acted as PREs in an embryonic en-lacZ reporter construct but PRE1 had greater activity in this context (DeVido et al. 2008). In order to directly compare the activities of PRE1 and PRE2, they were tested in a bxd-Ubx-lacZ reporter inserted into the genome using φC31 recombinase-mediated cassette exchange (RMCE), (Bateman et al. 2006; Groth et al. 2004) incorporated into the attP insertion site located at 52D (see Bateman et al. 2006). The recombinants generated by this system do not carry the mini-white gene; therefore PSS cannot be assessed at this location. This vector and insertion site was previously used to demonstrate the embryonic maintenance activity of the eve PRE (Fujioka et al. 2008).

Based on genome-wide studies of H3K27me3 in embryos (Schuettengruber et al. 2009), the 52D insertion site does not lie in an H3K27me3 domain. Embryos with vector alone, PRE1-vector, and PRE2-vector were stained with an antibody against β-galactosidase (β-gal). In the absence of a PRE, bxd-Ubx-lacZ is expressed in parasegments 6-13 early in development, but later becomes derepressed and is expressed throughout the entire embryo (Fig. 6A). PRE1, but not PRE2, was able to maintain the restricted expression of bxd-Ubx-lacZ in this chromosomal insertion site (Fig. 5A). That PRE2 did not maintain
repression was somewhat surprising, since we previously showed that it was able to maintain the restricted pattern of the same bxd-Ubx-lacZ reporter gene at 9 out of 12 chromosomal insertion sites in a P-element based vector (Brown et al. 2005). With RCME, constructs were inserted in the genome in either orientation and similar results were obtained from both orientations. This shows that there is a difference in how PRE1 and PRE2 work in embryos at the 52D insertion site.

We next examined the lacZ expression in the bxd-Ubx-lacZ larvae, with and without the PREs (Fig. 6A). In larvae with vector alone, β-gal activity was detected in the salivary glands and in a subset of cells in the brain. This pattern of expression was not due to the insertion site since a similar pattern was observed using the same vector inserted into an attP site at 95E (Fujioka et al. 2008; data not shown). Addition of PRE2 into the vector abrogated expression of β-gal in the salivary glands but not in the brain of the transgenic larvae. Addition of PRE1 to the vector abrogated β-gal expression in the salivary gland and greatly reduced the number of cells in the brain that express β-gal. In order to investigate whether the repression of β-gal activity was due to the action of the PcG proteins, we assayed for the presence of H3K27me3 on the transgene by chromatin immunoprecipitation followed by qPCR. H3K27me3 is a signature of PcG activity and Ubx PRED is included as a positive control. In larvae with vector alone, H3K27me3 is not associated with the lacZ gene in either the brain or the salivary glands. In contrast, salivary glands from flies with vector plus PRE1 or PRE2 have H3K27me3 associated with the lacZ gene. The fact that PRE2 can act as a PRE in the salivary glands, but not in the brain or embryo suggests that either its activity is tissue specific, or that the large number of closely aligned transgenes present in the polytene chromosomes potentiated
the activity of PRE2 allowing it to overcome the activity of the transcriptional activators acting on this construct.

**DISCUSSION**

**PRE architecture:** PREs are made up of binding sites for many different proteins, some known and others unknown. Our detailed analysis of PRE2 suggests that at least six DNA binding proteins are required for its PSS activity including Pho/Phol, Dsp1, Spps, GAF/Psq, Zeste/Fs(1)h, and Protein A. PRE1 contains presumptive binding sites for all of those proteins as well as an AT-rich region not found in PRE2. Further, the spacing and order of sites differs between these two PREs (Fig. 7). We examined the minimal eve and iab-7/Fab-7 PREs for the presence of these sites (Fig. 7). We note that not all sites were found in the eve and iab-7/Fab-7 PREs. For example, we could not find a Dsp1 consensus sequence in the eve PRE and Dsp1 is not bound to this locus in embryos as shown by chromatin immunoprecipitation (Schuettengruber *et al.* 2009). A comparison of the order and arrangement of the confirmed and putative binding sites of en PRE1, en PRE2, eve PRE, and Fab7/iab-7 PRE shows a large variation in the order and spacing of the binding sites between the different PREs. This variation may reflect critical differences that are integral to how a specific PRE functions at its endogenous locus however, this is hard to assess given that even within a given PRE the number, order, and spacing of predicted Dsp1, Pho, GAF, and Zeste sites was found to diverge rapidly within PREs present at conserved positions in various Drosophila species (Hauenschild *et al.* 2008). The functional consequences of these differences have not been explored but it does point to an overall plasticity in the system.
This rapid evolution may in part result from redundancy in PRE activity. In the *en* gene, PRE1 and PRE2 are contained within a 1.6kb region and may act together to recruit PcG complexes in their normal context. However, both of these PREs are dispensable for viability in the laboratory (Cheng *et al.* 2012). The *en/inv* gene complex has two other major PREs and a number of minor ones that can apparently take over PRE1 and PRE2 functions. This apparent redundancy in PRE activity may allow for rapid evolution and increase the diversity of these regulatory elements.

**Functional differences between PREs:** Both PRE1 and PRE2 mediate pairing-sensitive silencing of the mini-*white* gene but behave differently in the *bxd-Ubx-lacZ* reporter complex inserted by RCME at an *attP* site at 52D. While both PRE1 and PRE2 could silence the *bxd-Ubx-lacZ* gene in salivary gland nuclei via the PcG proteins (as measured by H3K27me3 association with the *lacZ* gene), only PRE1 was active in embryos and larval brain. Why this difference? We suggest that PREs evolve in the context of other regulatory DNA and, while retaining overlapping functions, they may adapt to work better with different enhancers. So, while core PRE DNA binding proteins may remain relatively constant between PREs, some PREs may require additional factors to overcome the activity of specific enhancers.

Differences in PRE activities in transgenes have been previously described (Horard *et al.* 2000; Park *et al.* 2012; Okulski *et al.* 2011). In one study, multimerized sub-fragments of the *bxd* PRE showed different behaviors in a mini-*white* repression assay (Horard *et al.* 2000). In another, different PREs from the *Psc/Su(z)2* complex exhibited different silencing activities on an imaginal disk reporter construct (Park *et al.* 2012). Both these studies were done in P-element based vectors and thus the activities of
the reporter constructs were not examined in the same insertion sites. Since PREs are notoriously sensitive to the effects of flanking DNA, it is not easy to determine whether the differences observed were a reflection of intrinsic differences in the activity of the PREs or to the different insertion sites of the reporter genes. In order to avoid this problem, Okulski et al. (2011) examined the activity of a PRE from the vestigial (vg) gene with activity of the iab-7/Fab-7 PRE inserted in the genome at 4 different insertion sites via \( \Phi C31 \) integration. This careful study showed that these two PREs exhibited different degrees of mini-white repression at different insertion sites and behaved differently with respect to pairing-sensitive silencing at the same insertion sites. The problem with this study was that the investigators used relatively large fragments of DNA (1.6kb). Thus it is likely that other regulatory sequences were contained within the PRE fragments, and it is not known how additional regulatory sequences modify PRE activity. Even in the 181bp DNA fragment that we call PRE2, there is another regulatory activity, a promoter-tethering element, which may affect its ability to act as a PRE in some chromosomal locations (Kwon et al. 2009). In previous studies from our group, when a 2.6kb piece of \( en \) DNA that included PRE1 and PRE2 and flanking DNA was included in a \( bx d-Ubx-lacZ \) P-element based transgene, the \( en \) fragment almost completely silenced the entire \( lacZ \) expression pattern in embryos (Americo et al. 2002). We suspect that regulatory DNA within that fragment was able to silence the \( bx d \) enhancer, and that this activity might be separable from the PRE activity per se.

**Concluding remarks:** From this work and the work of many others we conclude that PREs are a diverse group of related elements that have overlapping but not identical activities. PREs vary in both size and binding site composition requiring a number of
different DNA binding proteins to function collaboratively to recruit PcG complexes and introduce repressive histone modifications. The diverse PRE architecture could influence exactly which PcG protein complexes get recruited and what types of enhancers they are able to silence or repress.

ACKNOWLEDGEMENTS

We thank Miki Fujioka for the bxd-Ubx-lacZ RCME vector and for transgenic lines; Bill McGinnis for the generous gift of Grh antibodies; and Emma Hornick for construction of the bxd-Ubx-lacZ reporter constructs. Thanks to Yuzhong Cheng, Payal Ray, and Sandip De for critical reading of this manuscript. This work was supported by the Intramural Research Program of the Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health.

REFERENCES


Hagstrom, K., M. Müller, and P. Schedl, 1997 A Polycomb and GAGA dependent silencer adjoins the Fab-7 boundary in the Drosophila bithorax complex. Genetics 146: 1365-1380.

Hauenschild, A., L. Ringrose, C. Altmutter, R. Paro, and M. Rehmsmeier, 2008


Schuettengruber, B., D. Chourrout, M. Veroort, B. Leblanc, and G. Cavalli, 2007


Comparative analysis of chromatin binding by sex Comb on Midleg (SCM) and other

Wang, H., L. Wang, L., H. Erdjument-Bromage, M. Vidal, P. Tempst, R. S. Jones, and
Y. Zhang, 2004 Role of histone H2A ubiquitination in Polycomb silencing. Nature 431:
873–878.

**FIGURE LEGENDS**

**Figure 1. Predicted binding sites in PRE1 and PRE 2** (A) Shows a schematic drawing
of the genomic region of the *en* locus showing the *en* transcription unit and the two
upstream PREs, PRE1 and PRE2 (DeVido *et al.* 2008). Each PRE is shown as an
enlargement with predicted binding sites/consensus sequences represented by color-
coded bars using the following consensus sequences: Pho-GCCAT; GAF/Psq-GAGAG;
Zeste-BGAGTGV or YGAGYG (Mohrmann *et al.* 2002; Ringrose *et al.* 2003) (B=C,T,
or G; V=A, C, or G; Y=C or T); Sp1/Klf-RRGGYG (R=A or G)(Shields and Yang
1998); Dsp1-GAAAA (Dejardin *et al.* 2005), Grh-TGTTTTTT or WCHGGTT (W=A or
T; H=A, T, or C) (Blastyak *et al.* 2006; Venkateson *et al.* 2003; Almeida and Bray 2005);
Site A-GAACNG. (N=A, C, T, or G). PRE1 (-1944 to -1503) when subdivided into PRE
1.1 or PRE 1.2 leads to loss of PSS with either fragment. The relative positions of sites B
and C in PRE2 are shown (from Americo *et al.* 1992). We did not include sites B or C in
our binding site analysis since mutation of site C did not affect PSS and site B has not
been fully defined or tested in a functional assay. (B) Shows a comparison of the number of predicted consensus binding sites for each factor in PRE1, PRE1.1 and PRE 2.

Figure 2. Defining Site A and ChIP with αGrh (A) Shows the sequence of the Z1/A 18bp annealed oligonucleotide used in the gel mobility shift assays with Schneider cell extract (SL2) shown in B and C and the mutant sites tested: Zeste site (Mut Z1), Site A (Mut A) and the Zeste/siteA double mutant (Mut A+Z1). The mutated bases are shown for each site, and altered bases were made in the context of the 18bp oligo sequence (B) Gel mobility shift assay using SL2 extract on the Z1/A oligo with 50X competition with unlabeled non-specific competitor (Sp1/KLF site, TCACGCGGCGTTACCGAG; Brown et al. 2005); the wild type Z1/A oligo; and an oligo known to bind the Zeste protein (CGAGTGGATGTGACTG). The known Zeste binding site does not compete. C. Gel mobility shift assay on the Z1/A oligo showing competition with the Z1/A oligo with a mutated site A or a mutated Zeste site. N.Sp (non-specific) is competition with the Sp1/KLF binding site oligo and WT represents the unmutated Z1/A oligo. (D) Effect of binding site mutations in PRE2 on PSS activity. The location of the Z1/A oligo is boxed. PSS is given as the number of lines and as a percentage. The 139bp construct is a minimal PRE2 sequence that was previously described in Americo et al. (2002). The effect of mutating Z1 or Z2 individually has not been tested. (E) Grh/ChIP at en PRE1 and PRE2 in imaginal disk and CNS material isolated from 3rd instar larvae. Grh ChIP was carried out 3 times with two different Grh antibodies (rabbit or guinea pig, Kim and McGinnis 2011). The graph shown represents two independent experiments using the guinea pig Grh. A schematic on the right shows the approximate position of each oligonucleotide set. ChIP values are plotted as percent input and error bars are shown.
Figure 3. Defining the PHO site requirements of PRE1 and PRE2. (A) Sequences of Pho1-4 lined up with the extended Pho consensus binding sequence proposed by Mohd-Sarip (2005). The bases that are part of the extended consensus sequence are given in red and are underlined in the consensus site. The bases in bold are the matches to the core PHO consensus binding site. Sites that are marked with an asterisk show strong Pho binding. (B) Shows gel mobility shift analysis of in vitro transcribed/translated full length Pho protein binding to the Pho1 site with the inclusion of each of the oligonucleotides in competition. The first panel shows binding in the presence of 100 X unlabeled Pho1 Oligo, Mut Pho1 oligo and Pho2 oligo. The second panel is full-length PHO protein binding to the Pho1 oligo in competition with a non-specific oligo and Pho3 and Pho4 respectively. (C) and (D) Schematic of the predicted binding factor sites in PRE1 and PRE2 respectively. The boxed areas highlight the Pho sites and the absence of a Pho site in a construct represents mutation of this binding site. The numbers and percentages of lines that give PSS are shown.

Figure 4. PRE1 requires an AT rich region for full activity. (A) Shows the set of constructs that were made of PRE1. Increasing amounts of PRE1.2 sequence were added back to PRE1.1 and the levels of PSS were observed for each. The predicted binding sites are shown for each construct. (B) The sequence of the AT-rich sequence of PRE1 that is denoted by the dotted line in A above. When this sequence is added to PRE1.1Z, full PSS activity is restored. The boxed sequences are repeats of ATTTAT the line above the sequence shows the oligo used for the gel mobility shift assays in C below. The asterisks show the bases that are mutated in a mutant oligo used in competition of the gel mobility
shift below. The indicated bases were changed to GCAG. (C) Shows gel mobility shift analysis of SL2 extract with an oligo spanning one of the ATTTAT repeats. The gel shift is competed by the specific oligo but not by an oligo in which part of the ATTTAT sequence has been mutated. Oligos in which the ATTTAT is mutated in the first AT and the 2 preceding bases no longer competes the gel shift (data not shown).

Figure 5. Comparison of predicted binding sites between D.melanogaster and D.virilis at PRE1 and PRE2. (A) Shows the predicted binding site comparison between PRE1 and 2 of D. melanogaster and D.virilis. See the Figure 1 legend for the consensus binding sequences used. The asterisks denote sites conserved in PRE2. Note that the D.virilis PRE1 fragment has not been tested for PRE activity. (B) Shows the comparison of the A/T-rich sequence at the end of PRE1 between D. melanogaster and D. virilis. The arrows delineate the AT-rich sequence of PRE1, extra bases on either side are included. Spaces are introduced in order to get the best line up of sequence. The D. virilis sequence contains several ATTTAT sequences (highlighted in red) and the bold area shows the most AT-rich part of the sequence: 72% (83/113 bases) and 77% (111/144 bases) A/T richness for D. melanogaster and D. virilis respectively. Identical sequences are boxed.

Figure 6. Activity of PRE1 and PRE2 in the Ubx-lacZ reporter construct inserted at 52D in embryos and larval tissues. (A) Anti-ß-gal-stained embryos, salivary glands and brains from transgenic lines carrying a bxd-Ubx-lacZ reporter gene inserted at 52D. Staining is shown for transgenic lines carrying vector alone, vector + PRE2 and vector + PRE1. The embryos shown are stage 14 (anterior left, dorsal up). The double-headed
arrow highlights the repression of lacZ expression in late stage embryos carrying the PRE1 construct but not the vector alone or the construct carrying PRE2. Salivary gland staining is seen only with the vector and not with the constructs carrying PRE1 or PRE2. Brain staining is seen in a specific pattern that is reduced in the flies carrying PRE1 but not in flies carrying PRE2. (B) ChIP with anti-H3K27 on 3rd instar larval salivary glands of transgenic lines carrying the bxd-Ubx-lacZ reporter gene inserted at 52D (see Bateman et al. 2006). Results are shown for vector only, vector + PRE2 and vector + PRE1. qPCR was carried out with transgene specific primers spanning the PRE/bxd junction, the Ubx/lacZ junction and within the lacZ gene. Primers from the endogenous PRE of the Ubx locus (PRED) were used as a positive control in each of the transgenic lines. Values are given as percent input and error bars represent the variation between experiments.

**Figure 7. Comparison of the order and spatial arrangement of consensus factor binding sites in en, eve and iab-7/Fab7 PREs**

The sequences of en PRE1 and PRE2, the minimal eve PRE (PRE300-3’ (158bp); Fujioka et al. 2008), and the 219bp iab/Fab7 PRE (Dejardin et al. 2005) were analyzed in GenePalette using the consensus sequences described in the Fig.1 legend. The distribution and order of the predicted binding sites are shown. All fragments shown mediate PSS.
### A

The diagram shows the transcription factor binding sites in the PRE1 and PRE2 regions. The binding sites are indicated by different colors and labels:

- **Pho**
- **Sp1**
- **GAF/Psq**
- **A**
- **B**
- **C**
- **Zeste**
- **Dsp1**
- **grh**

### B

The table summarizes the number of binding sites for different transcription factors:

<table>
<thead>
<tr>
<th>Binding Site</th>
<th>PRE2</th>
<th>PRE1</th>
<th>PRE1.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHO/PHOL</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>GAGAG</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Sp1/KLF</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Zeste</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Dsp1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Site A</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Grh</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
### A

(Mut A + Z1)

TACCT

(Mut Z1)

CTGT  CTAG (Mut A)

AGAGGGAGTGAAACAGTGCT

### B

N. Sp.  WT  zeste

### C

N. Sp.  WT  Mut A  Mut Z1

### D

<table>
<thead>
<tr>
<th></th>
<th>Z1</th>
<th>Z2</th>
<th>PSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE2</td>
<td></td>
<td></td>
<td>14/21 66</td>
</tr>
<tr>
<td>Mut A + Z1</td>
<td></td>
<td></td>
<td>0/11  0</td>
</tr>
<tr>
<td>Mut Z1/Z2</td>
<td></td>
<td></td>
<td>5/27 19</td>
</tr>
<tr>
<td>Mut A</td>
<td></td>
<td></td>
<td>4/30 13</td>
</tr>
<tr>
<td>139bp</td>
<td></td>
<td></td>
<td>13/22 59</td>
</tr>
</tbody>
</table>

- Pho  - Sp1  - GAF/Psq  - Zeste  - A  - grh  - Dsp1

### E

Grh at en PREs

(-2.4kb)

PRE1  PRE2  +1

en

% input

1  0.8  0.6  0.4  0.2  0.0

1  2  3
**A**

- **CRGCCAT**YDTRGD
- *Pho1** CAGCCAT**TTTCCCT**
- **Pho2** ATGCCAT**CTCTTT**
- *Pho3** CAGCCAT**TGTTGT**
- *Pho4** CAGCCAT**TAAAG**

**B**

![Image of electrophoresis gels showing Pho1, Mut Pho1, Pho2, non Sp., Pho #3, Pho #4](image)

**C**

<table>
<thead>
<tr>
<th></th>
<th>Pho2</th>
<th>Pho1</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mut Pho1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mut Pho2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PSS</th>
<th># lines</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>14/21</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>1/11</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>9/11</td>
<td>81</td>
<td></td>
</tr>
</tbody>
</table>

**D**

<table>
<thead>
<tr>
<th></th>
<th>Pho3</th>
<th>Pho4</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mut Pho3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mut Pho4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mut Pho3+4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>PSS</th>
<th># lines</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>12/22</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>3/20</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>2/25</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>0/21</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

- Pho
- Sp1
- GAF/Psq
- Zeste
- A
- grh
- Dsp1
A

PRE1.1
PRE1.1X
PRE1.1Y
PRE1.1Z
PRE1

# lines  %
0/19     0  
2/18     11 
3/19     16 
5/23     21 
12/22    54 

B

**
ATTTCAATTTATGATCTGCTTTTAAATATTATTTATTTATTATTATTTATT
TAAAGTTAAATAACGTAGACTCGACAAGTATATATAAAATTTAATTTAAT
CGATTCAAGTCCATTGATAAC
GCTAAGTTCAGGTAAACTATG

C

Non Sp. Mut oligo WT oligo

[Image of gel]