# Characterization of *cis*- and *trans*-acting elements in the imprinted human *SNURF-SNRPN* locus

Sara Rodriguez-Jato[1], Robert D. Nicholls[4], Daniel J. Driscoll[2,3] and Thomas P. Yang[1,2,3,*]

[1]Department of Biochemistry and Molecular Biology, [2]Department of Pediatrics, [3]Center for Mammalian Genetics, University of Florida College of Medicine, Gainesville, FL 32610, USA and [4]Center for Neurobiology and Behavior, Department of Psychiatry, University of Pennsylvania, Philadelphia, PA 19104, USA

## ABSTRACT

**The imprinted *SNRPN* locus is a complex transcriptional unit that encodes the SNURF and SmN polypeptides as well as multiple non-coding RNAs. *SNRPN* is located within the Prader-Willi and Angelman syndrome (PWS/AS) region that contains multiple imprinted genes, which are coordinately regulated by a bipartite imprinting center (IC). The *SNRPN* 5′ region co-localizes with the PWS-IC and contains two DNase I hypersensitive sites, DHS1 at the *SNRPN* promoter, and DHS2 within intron 1, exclusively on the paternally inherited chromosome. We have examined DHS1 and DHS2 to identify *cis*- and *trans*-acting regulatory elements within the endogenous *SNRPN* 5′ region. Analysis of DHS1 by *in vivo* footprinting and chromatin immunoprecipitation identified allele-specific interaction with multiple regulatory proteins, including NRF-1, which regulates genes involved in mitochondrial and metabolic functions. DHS2 acted as an enhancer of the *SNRPN* promoter and contained a highly conserved region that showed allele-specific interaction with unphosphorylated RNA polymerase II, YY1, Sp1 and NRF-1, further suggesting a key role for NRF-1 in regulation of the *SNRPN* locus. We propose that one or more of the regulatory elements identified in this study may also contribute to PWS-IC function.**

## INTRODUCTION

*SNURF-SNRPN* (hereafter termed *SNRPN*) is a bicistronic imprinted gene on human chromosome 15 that encodes two polypeptides, the SmN splicing factor involved in RNA processing (1), and the SNURF (*SNRPN* upstream reading frame) polypeptide of unknown function (2). *SNRPN* also encodes a long (~460 kb) alternatively spliced RNA transcript that contains several families of snoRNAs (3) and extends downstream to partially overlap the *UBE3A* gene in the anti-sense orientation. The *SNRPN* promoter is associated with a CpG island that is hypermethylated on the maternally inherited allele and hypomethylated on the paternally inherited allele (4). Two alternative upstream promoters and alternatively spliced non-coding exons expressed at low levels add to the complexity of the locus (5). The gene is transcribed exclusively from the paternally inherited chromosome and shows highest levels of expression in the brain and heart (2). Furthermore, *SNRPN* is located within an imprinted gene cluster in chromosome 15 q11–q13 that is associated with the Prader-Willi syndrome (PWS) and Angelman syndrome (AS).

PWS and AS are two clinically distinct neurogenetic disorders linked to a single imprinted domain on chromosome 15 containing at least eight genes distributed across ~2 Mb [reviewed in (6)]; the similarly imprinted syntenic region in the mouse is located on chromosome 7C (6). PWS arises from loss of function of genes in this region that are expressed exclusively from the paternal chromosome, while AS arises from loss of expression or mutation of the maternally expressed *UBE3A* gene. Multiple genetic mechanisms lead to the allele-specific loss of gene expression in AS and PWS, including deletions of the entire imprinted region, uniparental disomy (UPD), and microdeletions that encompass the 5′ region of the *SNRPN* gene and/or a region upstream of *SNRPN*. High-resolution mapping of these microdeletions has led to the delineation of a bipartite imprinting control region or imprinting center (IC). All of the microdeletions associated with PWS share a 4.3 kb deleted region termed the PWS smallest region of deletion overlap (PWS-SRO), which includes the *SNRPN* promoter region, first exon, and part of the first intron (7). Similarly, the microdeletions associated with AS share a 0.8 kb AS-SRO located ~35 kb upstream from exon 1 of *SNRPN* (8). Thus, the IC is composed of two distinct functional components, the PWS-IC (including but not necessarily limited

to the PWS-SRO) that appears to be required for maintenance of the paternal epigenotype in somatic cells (9,10), and the AS-IC (including the AS-SRO), which appears to be required for establishment of the maternal epigenotype during oogenesis (11). Currently, the mechanisms of PWS-IC and AS-IC function in establishing and/or maintaining imprinted gene expression across the domain are not well understood.

The co-localization of the PWS-SRO with the *SNRPN* promoter suggests that transcription factor binding to the promoter and subsequent transcriptional activation of the *SNRPN* locus may be integral to PWS-IC function (12). Several studies have identified *cis*-acting elements in the *SNRPN* promoter region that affect promoter function in transient expression assays and regulate expression or imprinting of mouse transgenes. However, the corresponding *trans*-acting factors have not been identified and no specific transcription factors have yet been shown to bind to the endogenous *SNRPN* locus. *SNRPN* promoter function was first reported to be contained within an interval between positions −207 and +53 by using transient expression assays of reporter constructs that included an exogenous SV40 enhancer (13). Employing the same strategy (and the SV40 enhancer), the minimal *SNRPN* promoter was subsequently shown to be located between nucleotides −71 and +51, a region that contains a 7 bp element (SBE) between nucleotides −57 and −51 and an element around position +17 that act as positive regulators of transcription (14). In addition, a repressor sequence was mapped downstream of the minimal promoter in the 3′ region of exon 1 (14). The SBE coincides with one of the six sequences that are phylogenetically conserved between the 5′ flanking region of the human and mouse *SNRPN* genes (7). The SBE has also been identified as a functional promoter element in the mouse *Snrpn* gene (15). Subsequent analysis of the mouse *Snrpn* promoter in a transgene construct containing the human AS-SRO and the minimal mouse *Snrpn* promoter expressing a reporter gene identified several additional DNA sequences that appear to be involved in imprinted expression of the transgene (16). These included two *de novo* methylation signals (DNS), an allele discrimination signal (ADS) and two separate signals responsible for the maintenance of the paternal imprint (MPI1 and MPI2). The mechanisms and the associated *trans*-acting factors that mediate the function of these *cis*-acting elements remain unknown. Despite the fact that transgenic mice containing constructs with the human AS-SRO and the human (17) or the mouse (18) *SNRPN* promoters assume the correct imprinted state of the transgene, deletion of the endogenous mouse *Snrpn* promoter has no apparent effect on PWS-IC function (19), suggesting that sequences outside of the promoter may be sufficient to maintain the function of the PWS-IC in the mouse (16).

In the current study, we have used multiple strategies to identify *cis*-acting elements within the endogenous human *SNRPN* 5′ region and to determine the corresponding *trans*-acting regulatory factors that interact in an allele-specific manner with the promoter. We also have identified and analyzed *cis*-acting and allele-specific *trans*-acting elements associated with an enhancer located within the first intron of *SNRPN* just downstream of the PWS-SRO. The potential role of these regulatory elements in PWS-IC function is discussed.

## MATERIALS AND METHODS

### Cell culture

EBV-transformed lymphoblasts derived from PWS and AS patients were grown in RPMI medium supplemented with 10% fetal bovine serum (FBS) and 1% penicillin–streptomycin. Human SK-N-SH neuroblastoma cells were obtained from the American Type Culture Collection and grown in E-MEM supplemented with 10% FBS and 1% penicillin–streptomycin. Cells were grown at 37°C in 5% $CO_2$.

### DNase I treatment of permeabilized cells for mapping of hypersensitive sites and Southern blotting

Cell permeabilization and DNase I treatment was performed as described previously (20). Genomic DNA from DNase I treated cells was digested with BamHI and size-fractionated in a 0.8% agarose gel, transferred and hybridized as described previously (20). The 2.26 kb hybridization probe was isolated from plasmid pPH-B8 (21) by digestion with BamHI and EcoRI and labeled by random priming.

### Vector design

Constructs for transient expression assays were generated from the pGL3-Basic vector (Promega). Details on vector construction are provided in Supplemental Material.

### Transient transfection and luciferase reporter assays

SK-N-SH cells were transfected with firefly luciferase expression constructs using SuperFect (Qiagen) according to the manufacturer's specifications (see Supplementary Material). Cells were co-transfected with pRL-TK (Promega) which contains the *Renilla* luciferase gene. Cells were lysed 24 h post-transfection and firefly and *Renilla* luciferase activities were measured (Dual-Luciferase® Reporter Assay System, Promega) in a Sirius Luminometer V2.2. (Berthold Detection Systems). Firefly luciferase activity was normalized to *Renilla* luciferase activity. For each construct, the average and standard error of the mean [standard deviation/square root ($n$), where $n$ is the number of independent experiments] were calculated.

### Site-directed mutagenesis

Sequence-specific mutations in reporter constructs were performed using the QuikChange® XL Site-Directed Mutagenesis Kit (Stratagene) following the manufacturer's instructions. Mutated sequences were designed to include a new restriction site, and mutant clones were identified by restriction digestion and verified by sequencing. Supplementary Table S1 shows the wild-type and mutant sequences. Mutant sequences from constructs (c–g) in Figure 2G were excised with XbaI and cloned into the XbaI site of construct (c) in Figure 4B to generate the constructs (d–h) in Figure 4B.

### Dimethyl sulfate (DMS) treatment of cells and naked DNA for *in vivo* footprinting

DMS treatment of $\sim 2 \times 10^7$ lymphoblast cells and DMS treatment of purified DNA were performed as described by Hornstra and Yang (22).

### Ligation-mediated PCR (LMPCR)

LMPCR was performed essentially as described previously (20). The sequence of LMPCR primers and cycling conditions for the PCR are provided in Supplementary Table S2.

### LMPCR DNA sequencing ladders

Ladders were generated from genomic DNA as described previously (22) and were used to determine the exact position of each guanine residue in autoradiograms of LMPCR-amplified DMS-treated samples.

### Electrophoresis, transfer and hybridization of sequencing gels

LMPCR-amplified products were analyzed by size-fractionation in a DNA sequencing gel, followed by electro-transfer, hybridization and autoradiography essentially as described previously (22). Radiolabeled probes were generated as described in Supplementary Material.

### Chromatin immunoprecipitation (ChIP)

ChIP was performed essentially as described by Leach *et al.* (23) on human lymphoblasts. A total of $10^7$ cells were used in each immunoprecipitation (IP) reaction with antibodies against YY1 (Santa Cruz Biotechnology sc-1703; 5 µg), NRF-1 (kindly provided by Richard C. Scarpulla; 4 µl), Sp1 (Santa Cruz Biotechnology sc-59; 10–12.5 µg), CTCF (Upstate 06-917; 5–20 µg), unphosphorylated RNA polymerase II (Covance MMS-126R, 40–60 µg), H3 dimethyl-K4 (Abcam ab7766-50; 2 µg), acetylated H4 (Upstate 06-866; 1 µl) and H3 acetyl-K9 (Upstate 06-942; 5 µg). To account for non-specific binding, each experiment included a mock IP reaction to which no antibody was added; these reactions routinely yielded no PCR products. Immunoprecipitated DNA was analyzed by PCR using PCR conditions and primer sequences shown in Supplementary Table S3. The linearity of the amplification was verified by diluting the input DNA (i.e. the fraction of the unbound chromatin in the mock IP) by 1:2, 1:4 and 1:8. Genomic regions outside of the PWS-AS domain that are known to be associated with specific factors were analyzed as a positive control for the IP reaction with antibodies against those factors (Supplementary Figure S3A). PCR products were size-fractionated in 5% TBE (89 mM Tris, pH 8.3, 89 mM boric acid and 2 mM EDTA) polyacrylamide gels and gel bands were visualized and quantified by SyBr-green staining in a fluorescence scanner (Storm 860, Molecular Dynamics). The fraction bound was calculated as percentage of input DNA; average and standard error of the mean across experiments are shown.

## RESULTS

The location of DNase I hypersensitive (DH) sites in chromatin within and surrounding the *SNRPN* promoter were mapped by Southern blot analysis (Figure 1) and confirmed previous reports by Schweizer *et al.* (24) and Ohta *et al.* (7). Our studies identified two strong hypersensitive sites specific to the paternal chromosome: DHS1—which is located in the promoter region of the *SNPRN* gene and DHS2—which is located in the first intron roughly 1.5 kb downstream of the
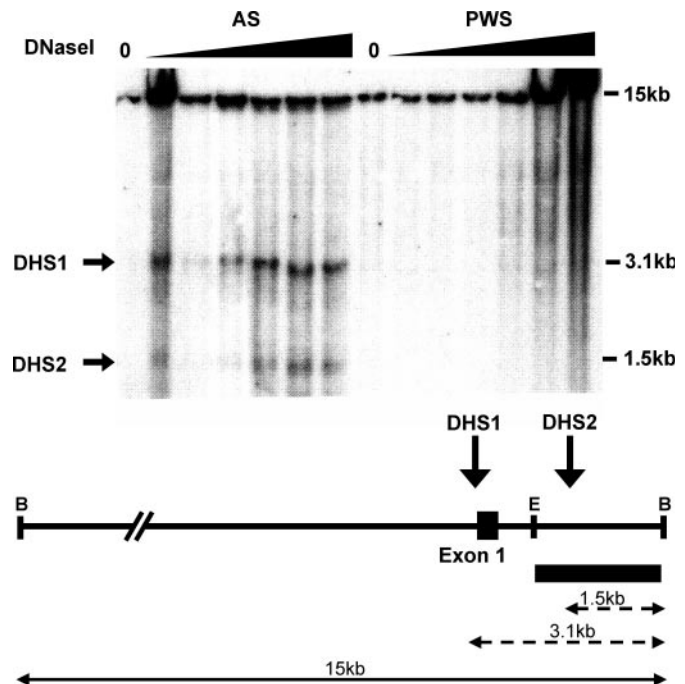


**Figure 1.** DNase I hypersensitivity of the *SNRPN* 5′ region. Lymphoblasts derived from PWS and AS patients were treated with increasing concentrations of DNase I, and DNase I hypersensitive sites were detected by Southern blot analysis. The diagram at the bottom shows the position of restriction sites, hypersensitive sub-bands, hybridization probe, and DNase I hypersensitive sites DHS1 and DHS2. The solid bar indicates the parental restriction fragment and the dashed lines indicate the subfragments generated by DNase hypersensitivity. The hatched box indicates the position of the probe. B, BamHI; E, EcoRI.

transcription initiation site. DHS2 is located just outside and downstream of the PWS-SRO (7). Because DH sites are commonly associated with *cis*-acting regulatory elements in eukaryotic chromatin, and because these DH sites were mapped within or immediately adjacent to the PWS-SRO, we used various strategies to identify and characterize both *cis*- and *trans*-acting regulatory elements within DHS1 and DHS2 of the endogenous *SNRPN* locus.

### *In vivo* footprint analysis of DHS1

To identify *cis*-acting regulatory elements associated with DHS1 within intact cells, we performed DMS *in vivo* footprint analysis of the *SNRPN* promoter region using LMPCR (22). As shown in Figure 2A, the region analyzed included a series of potential transcription factor binding sites identified by the TRANSFAC database, as well as six phylogenetically conserved sequences within the *SNRPN* promoter region (7). The paternal and maternal *SNRPN* alleles were assayed independently using cultured lymphoblasts from AS and PWS patients, respectively, carrying either UPD or deletion (Del) of the AS/PWS region. Figure 2A shows the relative position and region analyzed with each LMPCR primer set on the upper and lower strands of the endogenous *SNRPN* promoter region from position −490 to position +150.

Each region and strand of interest was analyzed in multiple DNA sequencing gels from at least two separate DMS treatments of cells, and footprints were further confirmed in two
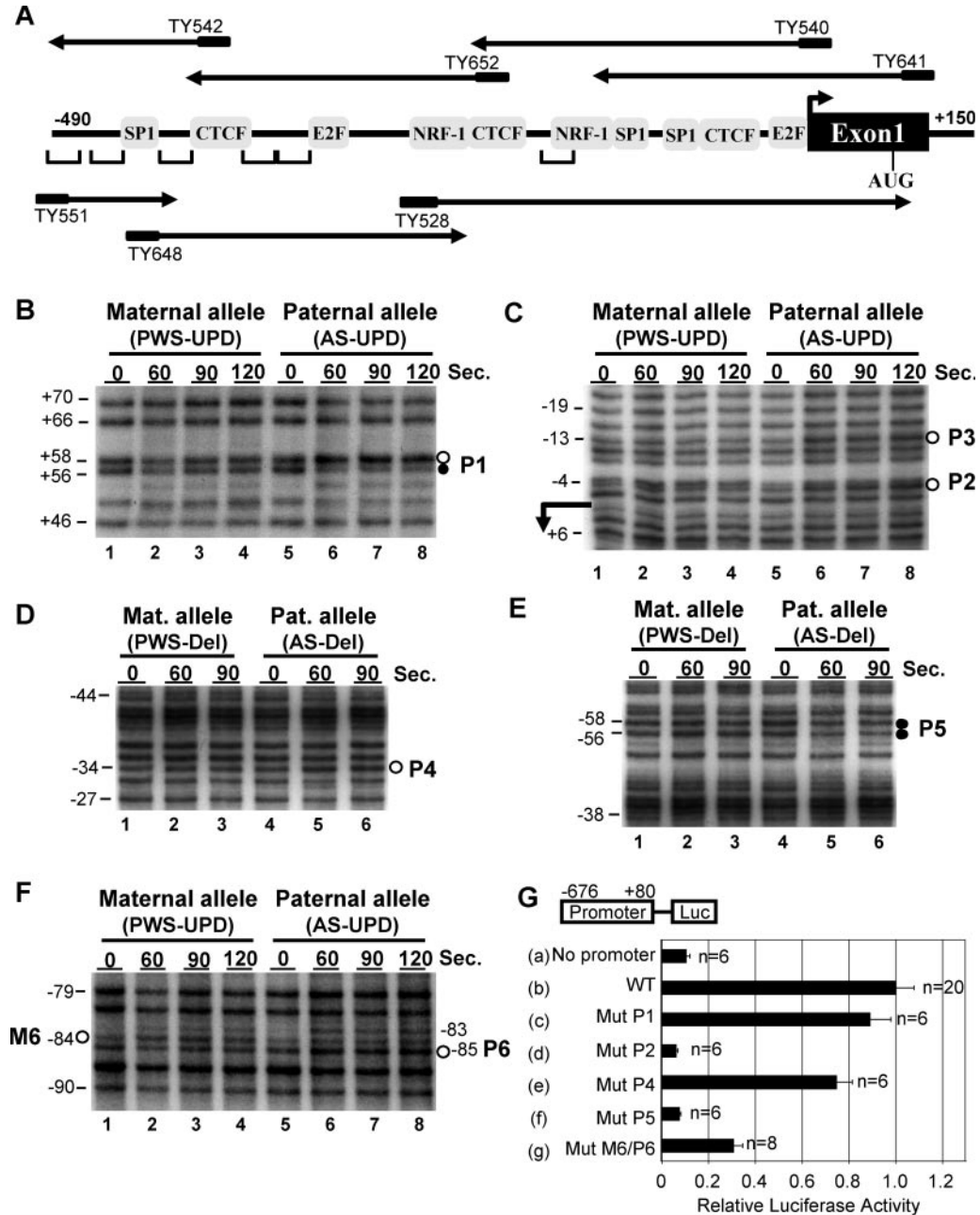
**Figure 2.** Analysis of *cis*-acting regulatory elements in the *SNRPN* 5′ region. (**A**) Diagram of the *SNRPN* 5′ region analyzed by *in vivo* footprinting. Indicated are the first exon (black box), the transcription initiation site (bent arrow), potential transcription factor binding sites (gray boxes) and sequences conserved between the human and the mouse loci (horizontal brackets) Horizontal arrows specify the position of LMPCR primer sets and the extent of the sequence analyzed with each set. Primer sets above and below the line assayed the upper and lower strands, respectively. (**B–F**) DMS *in vivo* footprint analysis of the *SNRPN* 5′region. Lymphoblasts derived from PWS and AS patients (UPD or Del, described in the text) were treated with DMS for the indicated periods of time. Time 0 s indicates purified genomic DNA treated *in vitro* with DMS. Representative sequencing gels are shown and the position of *in vivo* footprints on the paternal allele (P1–P6) and on the maternal allele (M6) are indicated with filled and open circles that represent protection or enhanced reactivity to DMS, respectively. The numbers flanking the autoradiograms indicate positions relative to the transcription initiation site. Primer set TY528 identified footprint P1 (B) and P6/M6 (F) on the lower strand; primer set TY641 detected footprints P2–P5 on the upper strand (C–E). (G) Functional analysis of *in vivo* footprints by transient expression assays. Luciferase reporter constructs including the *SNRPN* promoter carrying point mutations at *in vivo* footprints P1, P2, P4, P5 and M6/P6, were assayed by transient expression in SK-N-SH cells. *Mut M6/P6* carries point mutations for both the P6 and M6 footprints. Relative luciferase activities are shown compared with a control construct including the wild-type *SNRPN* promoter (b), which was arbitrarily assigned the value 1. Error bars represent standard error of the mean and *n* indicates the number of times each construct was assayed.

different cell lines (UPD and Del) for each parental allele. Although some of the footprints were subtle, all were highly reproducible (confirmed in 80–100% of the autoradiograms containing each site). Furthermore, transient expression assays

(see below) confirmed four of the six footprinted sites as bona fide *cis*-acting regulatory elements. Weak footprints were likely to be due to the fact that these *in vivo* footprinting assays were performed in lymphoblast cells, which express relatively
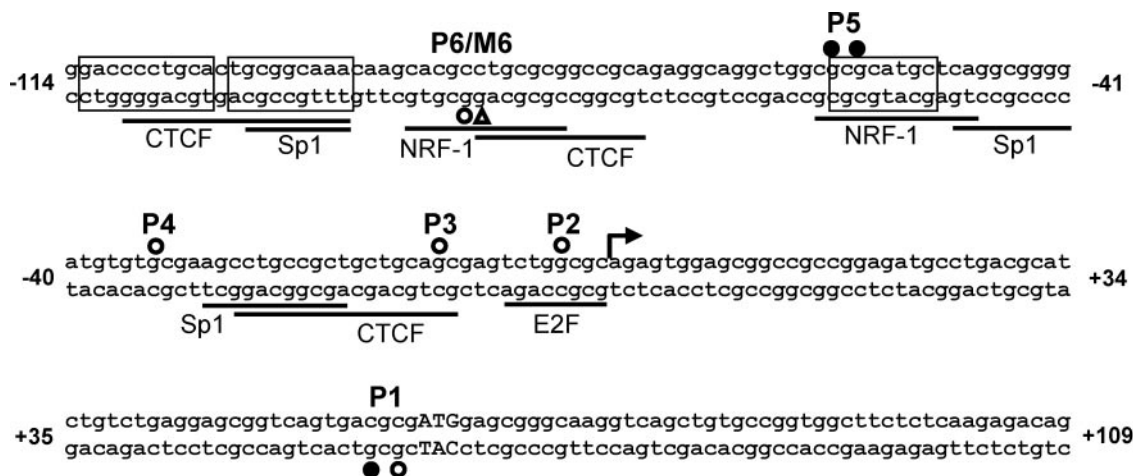
**Figure 3.** Summary of *in vivo* footprints in the *SNRPN* 5′ region. The nucleotide sequence corresponds to the human *SNRPN* 5′ region. Indicated are the positions of DMS *in vivo* footprints (circles and triangle), as well as the position of potential transcription factor binding sites (lines) and phylogenetically conserved sequences (boxes). Circles and triangles denote paternal and maternal footprints, respectively. Open and filled symbols denote enhanced and decreased DMS reactivity, respectively. Positions are indicated relative to the transcription initiation site (bent arrow).

low levels of *SNRPN* mRNA (21) and, therefore, unlikely to have a highly active *SNRPN* promoter maximally occupied by transcription factors. *In vivo* footprints were detected as bands of decreased or increased intensity relative to neighboring bands in each *in vivo*-treated sample compared with the pattern of relative band intensities in control naked DNA samples purified from the same cells. Figure 2B–F shows representative DNA sequencing gels for the *in vivo* footprint analysis; results are summarized in Figure 3.

Figure 2B shows *in vivo* footprint P1 at positions +56 and +58, which was identified with primer set TY528 on the lower strand of the paternal allele. Examination of relative band intensities revealed enhanced DMS reactivity at position +58, and protection from DMS modification at position +56 in the *in vivo*-treated samples containing the paternal allele (lanes 6–8) compared with the control naked genomic DNA samples (lanes 5). This pattern was not observed on the maternal allele (compare the naked DNA sample in lane 1 and the *in vivo*-treated samples in lanes 2–4). *In vivo* footprint P1 is located at the *SNRPN* translation initiation site within a sequence that shows weak similarity to an AP1 binding site. Furthermore, the P1 *in vivo* footprint lies within a previously reported negative *cis*-acting regulatory element mapped by transient expression assays within exon 1 of *SNRPN* (14).

Footprints P2–P5 were detected with LMPCR primer set TY641 on the upper strand exclusively on the paternal allele (Figures 2C–E). Figure 2C shows enhanced DMS reactivity at positions −4 (footprint P2) and −13 (footprint P3) in the *in vivo*-treated samples from the paternal allele (lanes 6–8) compared with the naked DNA control (lane 5) and that was not observed on the maternal allele (compare lane 1 with lanes 2–4). Footprint P2 is associated with a potential E2F binding site, and P3 lies within one of the three potential CTCF binding sites located in the *SNRPN* 5′ region (see Figure 3).

*In vivo* footprint P4, shown in Figure 2D, is detected at position −34 as a band of enhanced DMS reactivity exclusively on the paternal allele (compare lanes 5–6 with lane 4, and compare lanes 2 and 3 with lane 1). Searches of the

transcription factor database did not identify a known transcription factor binding site associated with P4. However, footprint P4 is flanked on both sides by potential Sp1 binding sites (Figure 3).

Figure 2E shows footprint P5 where two bands at positions −56 and −58 were protected from DMS modification exclusively on the paternal allele (compare lanes 5 and 6 with lanes 1 and 4, as well as lanes 2 and 3); this was confirmed in similar analysis using primer set TY540 (data not shown). Footprint P5 is associated with a potential binding site for NRF-1 (nuclear respiratory factor-1). P5 is also the only footprint that coincides with one of the six phylogenetically conserved sequences in the promoter region (Figure 3) and is included within the SBE (14).

Figure 2F shows footprints M6 and P6 that were identified with primer set TY528 on the lower strand at positions −84 and −85 on the maternal and the paternal alleles, respectively. The band at position −85 showed enhanced DMS reactivity relative to bands above and below it in samples containing only the *in vivo*-treated paternal allele (lanes 6–8) but neither in control naked DNA samples (lanes 1 and 5) nor in *in vivo*-treated samples carrying only the maternal allele (lanes 2-4). This indicates that footprint P6 is specific to the paternal allele. In contrast, the band at position −84, termed footprint M6, showed enhanced DMS reactivity only in the *in vivo*-treated samples carrying the maternal allele (lanes 2–4). Because the finding of an *in vivo* footprint on the transcriptionally repressed maternal allele was somewhat unexpected, footprint M6 was confirmed by DNase I *in vivo* footprinting (data not shown). The sequence shared by footprinted sites M6 and P6 lies within a potential NRF-1 binding site that partially overlaps with a potential CTCF binding site.

Because only altered DMS reactivity at guanine residues was used as the criterion for bona fide *in vivo* footprints, altered DMS reactivity at other nucleotides, such the adenine at position −83 in Figure 2F, was not deemed to reflect a footprint. In addition, nucleotides of apparent altered DMS reactivity that were not reproducible in at least 80% of the corresponding autoradiograms were also not considered to be

footprinted (e.g. the apparent footprint at position −79 in Figure 2F, which was not reproducible in other autoradiograms). Analysis of the *SNRPN* 5′ flanking region with the remaining four LMPCR primer sets (Figure 2A) identified no additional *in vivo* footprints.

## Functional analysis of *in vivo* footprinted elements in DHS1

To examine the potential function of the *in vivo* footprinted sites, transient expression assays were performed in which each footprinted sequence was mutated in a reporter construct containing a 756 bp fragment of the *SNRPN* promoter region (from position −676 to +80) driving the firefly luciferase gene. Figure 2G shows the relative luciferase activity of the mutant constructs compared with a wild-type control construct in human neuroblastoma SK-N-SH cells (cells which showed readily detectable *SNRPN* mRNA levels in northern blots; data not shown).

Mutations of the paternal-specific sites P2 and P5 [(d) and (f) in Figure 2G] significantly reduced the expression of the reporter gene, suggesting the association of P2 and P5 with *cis*-acting elements involved in *SNRPN* promoter function. Construct (g) in Figure 2G showed a reduction in reporter gene activity to ∼1/3 of wild-type levels upon mutation of the M6/P6 site. This reduction was most likely due to effects on interactions associated with the P6 footprint on the transcriptionally active paternal allele and not with the M6 footprint specific to the silent maternal allele.

Mutation of the sequences associated with footprinted sites P1 and P4 [(c) and (e) in Figure 2G] did not alter significantly the expression levels of the reporter. However, a potential role for footprint P1 is described below. Footprint P3 was not examined for effects on promoter function.

## Identification of an activator function associated with DHS2

To investigate the possibility that sequences within DHS2 might have an effect on *SNRPN* promoter function, a 2.2 kb (EcoRI–SmaI) genomic DNA fragment (termed 2.2-DHS2) that included DHS2 and flanking sequences was inserted into luciferase reporter constructs that included 756 bp of the *SNRPN* promoter (from positions −676 to +80) and analyzed by transient expression assays in SK-N-SH cells (Figure 4A). Constructs in which 2.2-DSH2 was cloned downstream from the luciferase gene in the forward [Figure 4A, (b)] or reverse orientations (c), and upstream of the promoter in the forward (d) or reverse (e) orientations showed an 8-, 2-, 5- and 8-fold increase in reporter activity, respectively, compared with a control construct lacking 2.2-DHS2 [(Figure 4A, (a)]. This strongly suggested that 2.2-DSH2 can act as an activator of the *SNRPN* promoter independent of orientation and position.

The activation function of 2.2-DHS2 was then examined in the context of the *SNRPN* mutations of *in vivo* footprinted sites shown in Figure 2G. We used a luciferase reporter construct that contained 756 bp of the *SNRPN* promoter as before, plus 2.2-DHS2 [Figure 4A, construct (e)]. As shown by construct (d) of Figure 4B, mutation of site P1 (which had no effect on reporter activity in the absence of the enhancer; see Figure 2G) resulted in a ∼50% increase in reporter activity in the presence
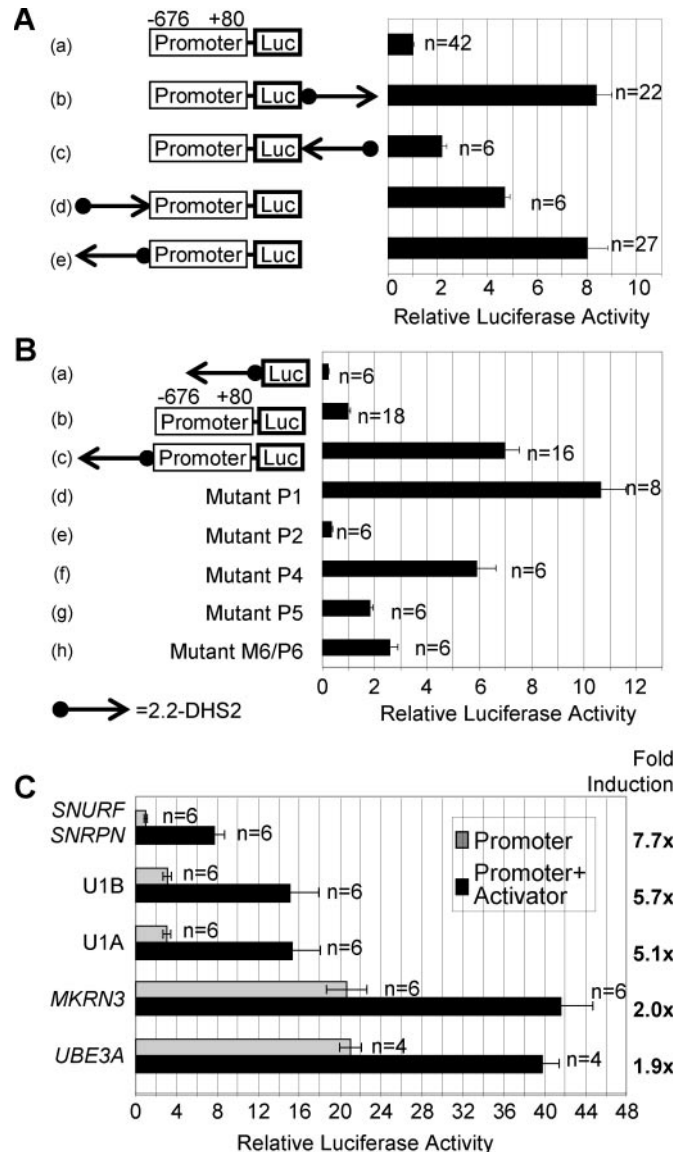


**Figure 4.** Identification of an enhancer associated with 2.2-DHS2. (**A**) Reporter constructs with 2.2-DHS2 (black arrow) at different positions and orientations with respect to the *SNRPN* promoter were assayed by transient expression assays as in Figure 2G. (**B**) Functional analysis of *in vivo* footprints in the presence of 2.2-DHS2. Construct (b) in Figure 4A was mutated at each of the *in vivo* footprints P1, P2, P4, P5 and M6/P6 and analyzed by transient expression assays as in Figure 2G. Relative luciferase activities are shown in comparison with reporter constructs containing the normal *SNRPN* promoter and/or 2.2-DHS2. (**C**) Preferential activation of the *SNRPN* promoter by 2.2-DHS2. The effect of 2.2-DHS2 on the promoter of various genes in the PWS-AS region was examined by transient expression assays as in Figure 2G. Gray bars and black bars represent the relative luciferase activity for reporter constructs with the indicated promoters and lacking 2.2-DHS2 (gray bars) or carrying 2.2-DHS2 (black bars). The construct containing the *SNRPN* promoter and no 2.2-DHS2 was arbitrarily assigned the value 1. The fold increase induced by 2.2-DHS2 with each promoter is indicated on the right.

of 2.2-DHS2 [compared with control construct (c)]. This suggested that regulatory elements associated with *in vivo* footprint P1 may be involved in negatively regulating the activation of the *SNRPN* promoter by 2.2-DHS2. Mutation of sites P2 and M6/P6 [Figure 4B, (e) and (h)] in the presence of 2.2-DHS2 had effects similar to that seen in the absence of

2.2-DHS2 (compare Figure 2G with Figure 4B). However, mutation of the footprint P5 resulted in a smaller reduction of promoter activity in the presence of 2.2-DHS2 [see Figure 4B, (c) versus (g)] compared with the absence of 2.2-DHS2 [see Figure 2G, (b) versus (f)]. Mutation of site P4 [construct (f)] had no effect on reporter activity, either in the presence [Figure 4B, (f)] or in the absence [Figure 2G, (e)] of 2.2-DHS2, which suggests that the factor(s) binding this site may play a role in functions other than direct promoter activity.

### Effect of the intronic activator on other promoters in the AS/PWS region

The human *SNRPN* locus includes several upstream promoters that are functional only from the paternal allele at low levels (5). Therefore, as shown in Figure 4C, we investigated the effect of 2.2-DHS2 on two upstream *SNRPN* promoters, U1A and U1B. Each of these promoter regions was cloned into a luciferase reporter construct that included 2.2-DHS2 positioned downstream of the reporter in the forward orientation [as in Figure 4A, construct (b)] and analyzed using transient expression assays in SK-N-SH cells as before. Both the U1A and U1B promoters were activated significantly by 2.2-DHS2 relative to a construct lacking 2.2-DHS2, though to a slightly lesser extent than the major *SNRPN* promoter itself (5.1- and 5.7-fold versus 7.7-fold, respectively). We also examined the effect of the 2.2-DHS2 sequence on promoter activity of other imprinted genes within the AS/PWS region, including the paternal *MKRN3* gene located ∼1.4 Mb upstream of *SNRPN*, and the maternal *UBE3A* gene located ∼0.5 Mb downstream of *SNRPN*. As shown in Figure 4C, both the *UBE3A* and *MKRN3* promoters by themselves showed higher levels of basal promoter activity than the *SNRPN* promoter and were both activated by the 2.2-DHS2 fragment (2.0- and 1.9-fold, respectively), though to a lesser extent than any of the *SNRPN*-related promoters (5.1- to 7.7-fold). Analysis of the SV40 promoter (that was activated up to 14-fold by 2.2-DHS2; data not shown) further corroborated data that 2.2-DHS2 can function as an activator of heterologous promoters. Furthermore, 2.2-DHS2 did not repress the maternally expressed *UBE3A* promoter in these transient expression assays, which suggests that 2.2-DHS2 is unlikely to play a direct role in repressing maternal genes on the paternally inherited chromosome.

Although 2.2-DHS2 does not activate the *SNRPN* promoter with equal efficiency in all orientations and positions (see Figure 4A), and distances (data not shown), or activate all heterologous promoters to the same degree (Figure 4C, and data not shown for the SV40 promoter), its activity in transient expression assays nonetheless meets the classical criteria for enhancer function.

### Identification of regulatory elements in 2.2-DHS2

To identify candidate *cis*-acting DNA sequences within 2.2-DHS2 that could mediate its activator function, we searched for evolutionarily conserved sequences in the human 2.2-DHS2 and the entire first intron of the mouse and rat *Snrpn* genes. As shown in Figure 5A, comparison of these sequences identified a single ∼80 bp region that showed significant conservation among these species. Sequence analysis of the

conserved sequence with the TRANSFAC database revealed highly conserved potential binding sites for the transcription factors Sp1, YY1 and NRF-1. Sp1 is a well-characterized ubiquitous transcriptional activator (25). NRF-1 regulates a variety of genes involved in mitochondrial function and energy metabolism (26), and a binding site for NRF-1 was also *in vivo* footprinted in the *SNRPN* promoter region (see Figure 3). YY1 is a ubiquitous transcription factor associated with transcriptional initiation, activation and repression [reviewed in (27)]. In addition, there are other highly conserved nucleotide sequences (from nucleotides +1302 to +1310, and from nucleotides +1333 to +1343 of the human sequence) that are not associated with known transcription factor binding sites.

Sequence analysis of the region associated with DHS2 also indicates that it contains a CpG island [(according to the criteria established by Takai and Jones (28)], which includes the 80 bp phylogenetically conserved region. This CpG island in DHS2 is distinct and separate from the CpG island at the *SNRPN* promoter. The CpG island in the promoter region has been well characterized in normal individuals and shown to be hypermethylated on the maternal allele and hypomethylated on the paternal allele (6). Analysis of the CpG island associated with DHS2 (and the 80 bp conserved sequence) also showed a similar pattern of allele-specific methylation. Digestion of genomic DNA from PWS and AS patients with methyl-sensitive restriction enzymes (HhaI to analyze CpGs at +1325 and +1375, and EaeI to analyze CpGs at +1354 and +1360) followed by PCR across the restriction site indicated that the paternal allele is hypomethylated and the maternal allele is hypermethylated (see Supplementary Figure S1).

To determine whether *cis*-acting elements within the 80 bp conserved sequence contributed to the activation function of the 2.2-DHS2 fragment, subfragments of 2.2-DHS2 containing the conserved sequence were analyzed for the activation of the *SNRPN* promoter by transient expression assays as before (Figure 5B). A 94 bp BstNI–HaeIII subfragment (Bs-Ha in Figure 5B) containing the entire 80 bp conserved sequence activated the *SNRPN* promoter 4-fold over the promoter alone. Thus, the 80 bp conserved sequence was termed the conserved activator sequence (CAS). However, the BstNI–HaeIII subfragment exhibited only about 2/3 and 1/2 of the activation activity of a fragment that included sequences flanking the CAS (the N-S subfragment in Figure 5B) and the entire 2.2-DHS2 fragment, respectively. This suggested that the CAS requires additional sequences in 2.2-DHS2 to generate the full promoter activation activity of 2.2-DHS2. However, it is unclear what these sequences might be, since we have not identified additional regions of significant sequence conservation in 2.2-DHS2 outside of the CAS and no other DNase hypersensitive sites are detected in 2.2-DHS2 on the paternal allele (24).

We also demonstrated a role for the putative YY1 binding site in the activation function associated with 2.2-DHS2. Transient expression assays yielded a 4.5-fold reduction in the activation of the *SNRPN* promoter by 2.2-DHS2, which was mutated at the putative YY1 binding site compared with wild-type 2.2-DHS2 (see Supplementary Material and Supplementary Figure S2A). Furthermore, electrophoretic mobility shift assay analyses and supershift assays with YY1 antibodies
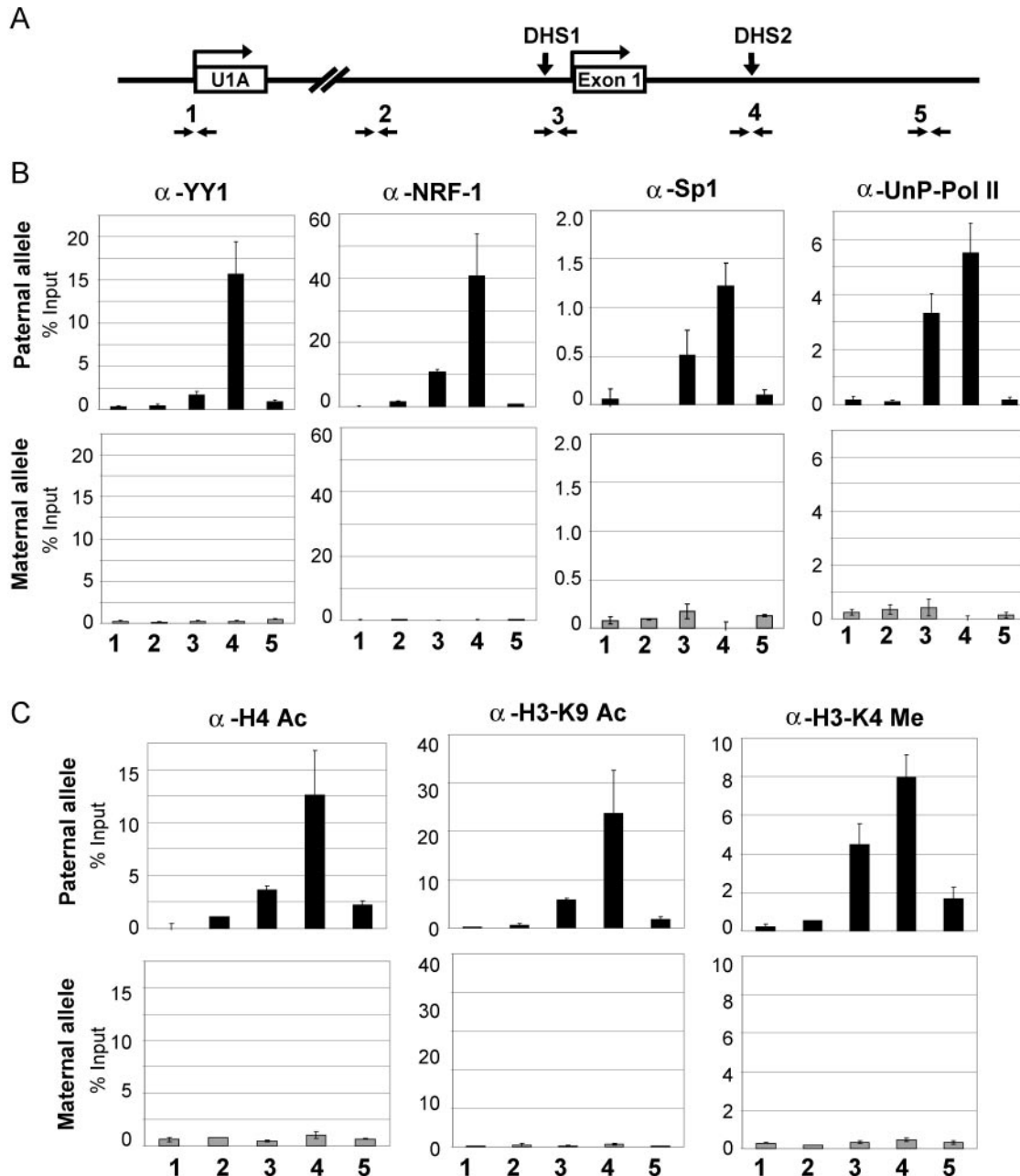
**Figure 5.** Identification and functional analysis of the CAS. (**A**) Sequence alignment between the human 2.2-DHS2 and the homologous regions of the mouse and rat *SNRPN* first intron. Numbers indicate positions with respect to the *SNRPN* transcription initiation site. Shaded boxes indicate nucleotides conserved between all three species. Potential transcription factor binding sites and restriction sites are also indicated. (**B**) Transient expression assays of activation function associated with the CAS (as in Figure 2G). A restriction map of 2.2-DHS2 is shown on the top left, and each subfragment of 2.2-DHS2 assayed in reporter constructs is depicted below by gray horizontal bars. Numbers refer to positions with respect to *SNRPN* transcription initiation site. Enzymes: B-BglII, Bs-BstNI, E-EcoRI, Ha-HaeIII, N-NotI, S-StyI and Sm-SmaI.

indicated that the potential YY1 binding site in the CAS can interact with YY1 *in vitro* (see Supplementary Figure S2B and C).

### ChIP analysis of transcription factor binding *in vivo*

ChIP analysis was used to examine the binding of Sp1, YY1, NRF-1 and CTCF to DHS1 (i.e. the *SNRPN* promoter region) and to the CAS within DHS2 *in vivo* using lymphoblasts derived from AS and PWS patients. As shown in Figure 6A, five regions in the *SNRPN* locus were assayed by ChIP analysis: (1) the U1A upstream promoter; (2) 1.3 kb upstream of the *SNRPN* promoter; (3) the *SNRPN* promoter; (4) the intronic CAS and (5) 1.3 kb downstream of the CAS. The analysis showed strong association of YY1 with the intronic CAS *in vivo* on the transcriptionally active paternal allele (i.e. in AS cells) and not on the maternal allele (i.e. in PWS cells, Figure 6B). In contrast, similar ChIP analysis of the non-imprinted glucocorticoid receptor gene, known to be bound by YY1 at an upstream regulatory region (29), showed association of YY1 in this region on both alleles (i.e. in both the AS- and PWS-derived cell lines; see Supplementary Figure S3A). Analysis of NRF-1 by ChIP assays demonstrated the association of NRF-1 with both the *SNRPN* promoter region as well as the intronic CAS (Figure 6B) specifically on the paternal allele. The association of NRF-1 with the *SNRPN* promoter is most likely to be at the potential

NRF-1 binding sites associated with paternal footprints P5 and/or P6 (Figure 3). As a positive control for NRF-1 association, ChIP analysis of the 3′ non-coding region of the non-imprinted myotonic dystrophy protein kinase gene (*DMPK*) that contains potential NRF-1 binding sites (as indicated by analysis of the TRANSFAC database; S. Rodriguez-Jato, unpublished data) detected NRF-1 binding in this region in both the AS- and PWS-derived cell lines (see Supplementary Figure S3A). The U1A upstream promoter of the *SNRPN* locus also contains a potential NRF-1 binding site (based on analysis of the TRANSFAC database); however, no association of NRF-1 with U1A was detected on either allele (Figure 6B). This is not unexpected since the upstream *SNRPN* transcripts are not expressed in blood (5) and the U1 promoters have been reported not to be nuclease hypersensitive in lymphoblast cells (24). Analysis of Sp1 association using ChIP assays showed interaction of Sp1 with both the promoter region as well as the CAS only on the paternal allele (Figure 6B).

Despite the fact that two *in vivo* footprints in the *SNRPN* promoter region were associated with potential CTCF binding sites (Figure 3), we have been unable to detect *in vivo* interactions involving CTCF on either the paternal or the maternal allele of the *SNRPN* promoter region by ChIP assays (Supplementary Figure S3B). As a positive control in these assays, we analyzed the previously reported association of CTCF with the 3′ non-coding region of the *DMPK* gene (30) and

**Figure 6.** ChIP analysis of the *SNRPN* locus. (**A**) A diagram of the region surrounding the *SNRPN* promoter. The location of primer sets used in the ChIP analysis is indicated by horizontal arrows. U1A is an upstream *SNRPN* exon. (**B** and **C**) Results of ChIP analysis. Chromatin from lymphoblasts derived from AS and PWS patients was cross-linked with formaldehyde and immunoprecipitated with antibodies against: (B) YY1, NRF-1, Sp1 and unphosphorylated RNA pol II, and (C) H4Ac, H3-K9Ac and H3-K4diMe histones. Immunoprecipitated DNA was amplified and analyzed by PCR using primers sets 1–5 and quantified by fluorescence scanning. The input DNA control sample was used to normalize the quantification of each PCR products. Average relative intensity values are shown. Error bars represent standard error of the mean.

confirmed the interaction of CTCF at this locus in both AS- and PWS-derived cell lines (Supplementary Figure S3A). Thus, it is likely that CTCF does not interact with either the maternal or paternal alleles of the *SNRPN* promoter in lymphoblast cells.

### Association of RNA polymerase II with the intronic CAS

YY1 has been shown to interact with and recruit RNA polymerase II to core promoters and activate transcription (27).

Therefore, we examined the possibility that the intronic CAS, with the possible involvement of YY1, could act as a site for recruitment of pol II to the *SNRPN* locus. Antibodies specific to the unphosphorylated form of RNA pol II were used in ChIP assays of the same *SNRPN* regions as before (Figure 6A); unphosphorylated RNA pol II is the non-processive form typically associated with transcription initiation at promoters. As shown in Figure 6B, ChIP analysis revealed strong association of unphosphorylated pol II with both the CAS and the promoter specifically on the paternal allele. Furthermore, the level

of unphosphorylated pol II on the paternal allele was higher at the CAS compared with the *SNRPN* promoter region. No significant amplification of the regions 1.3 kb upstream and downstream of the promoter and the CAS, respectively, was observed. Analysis of the U1A upstream promoter, which is inactive in blood, also did not detect association of pol II. These findings are consistent with a model of transcriptional activation in which the CAS may act to recruit unphosphorylated RNA pol II to the *SNRPN* locus (see Discussion).

### Analysis of histone modifications in DHS1 and DHS2

Allele-specific patterns of histone modification associated with the *SNRPN* promoter region have been reported previously (31,32) and consist of increased levels of H3 and H4 acetylation (H3 and H4 Ac) and H3 lysine 4 methylation (H3-K4 Me) on the paternal allele, and H3 lysine 9 methylation (H3-K9 Me) on the maternal allele. However, the elevated levels of H3-K4 diMe and H3-K9 diMe in the promoter region decrease drastically ∼200 and ∼600 bp, respectively, downstream from the transcription initiation site (32). To determine whether the downstream CAS has a distinct pattern of allele-specific histone modifications, we used ChIP analysis on AS- and PWS-derived lymphoblasts as before. We analyzed levels of H4 Ac, H3-K9 Ac and H3-K4 diMe, all of which are generally associated with transcriptional activation and are commonly elevated at the 5′ and regulatory regions of active genes (33,34). As shown in Figure 6C, the levels of all these modifications in histones associated with the transcriptionally active paternal allele are elevated at the CAS (and *SNRPN* promoter), and lower at the inactive U1A promoter and regions flanking the promoter and CAS. Significant levels of these histone modifications were not observed at any region assayed on the maternal allele. Of particular interest are the high levels of H3-K4 diMe at the CAS, given that the levels of H3-K4 diMe are known to drastically decrease 200 bp downstream from the *SNRPN* transcription initiation site (32). This suggests that the *SNRPN* promoter region and CAS are associated with two separate and distinct regions of chromatin, both characterized by histone modification patterns typical of regulatory regions in active chromatin (rather than co-localization of these two regions within one continuous stretch of chromatin having the same histone modification patterns). These results are consistent with the notion that the CAS is a regulatory region within the endogenous *SNRPN* locus.

## DISCUSSION

We have analyzed the 5′ region of the *SNRPN* gene for *cis-* and *trans*-acting regulatory elements that may mediate *SNRPN* transcription. These studies focused on the two DNase I hypersensitive sites located within the *SNRPN* promoter (DHS1) and first intron (DHS2). Both hypersensitive sites co-localize with CpG islands that are differentially methylated on the maternal and paternal alleles, a characteristic of regulatory regions associated with imprinted genes (35).

### *In vivo* footprint analysis of the *SNRPN* promoter region

DHS1 is located within the PWS-SRO and co-localizes with the *SNRPN* promoter. *In vivo* footprint analysis of the

endogenous promoter was used to identify *cis*-acting elements associated with six footprints on the paternal allele (P1–P6) and one on the maternal allele (M6). *In vivo* footprint P1 is located adjacent to the translation initiation site in exon 1 and is included within a previously described repressor sequence (14). However, our analysis showed that the repressor function appears to be limited to negatively regulating the activation of the *SNRPN* promoter by the enhancer associated with DHS2 (Figure 4B). Footprint P2, a potential E2F binding site, appears to be a *cis*-acting element essential for *SNRPN* promoter function in the presence or in the absence of the enhancer (Figures 2G and 4B). E2F consists of a family of transcription factors that participate in transcription activation and repression (36). The sequence associated with footprint P3 bears a resemblance to a potential CTCF binding site (37). CTCF has been implicated in the regulation of the imprinted *H19/IGF2* locus where it is associated with an intergenic insulator (38). However, we have been unable to demonstrate by ChIP assays that CTCF is bound in the *SNRPN* promoter region on either the paternal or maternal alleles. Therefore, *in vivo* footprint P3 is likely to be due to binding of a factor other than CTCF in lymphoblasts. It is also possible that CTCF is bound to this site in cell types other than lymphoblasts, possibly in the germ line. By transient expression assays, footprint P4 does not appear to play a direct role in *SNRPN* promoter function (Figures 2G and 4C) and does not seem to correspond to any known transcription factor binding site. However, it is located between two potential binding sites for Sp1, which our ChIP analysis showed to be associated with the *SNRPN* promoter region (Figure 6B). This suggests that footprint P4 may be the result of Sp1 interaction with one or both of the flanking Sp1 binding sites. Footprint P5 is associated with a potential NRF-1 binding site, which is phylogenetically conserved between the human and mouse *SNRPN* promoters. The interaction of NRF-1 with the paternal promoter region *in vivo* was demonstrated by ChIP (Figure 6B). In addition, footprint P5 overlaps with the previously reported SBE *cis*-acting element in the *SNRPN* promoter (14). Our analysis also found that P5 is essential for promoter function (Figure 2G). Footprints P6 on the paternal allele and M6 on the maternal allele show two distinctly different *in vivo* footprint patterns that localize to adjacent guanine nucleotides (Figure 2F). These data suggest that P6 and M6 represent interaction of different *trans*-acting factors at this sequence that perform different functions on the maternal and paternal alleles; on the maternal allele, the factor responsible for M6 may participate in the silencing of the *SNRPN* gene, while on the paternal allele, the factor representing P6 may contribute to *SNRPN* promoter function and gene activation as suggested by mutation analysis of the M6/P6 sequence (Figure 2G). The sequence associated with footprints P6 and M6 corresponds to overlapping potential NRF-1 and CTCF binding sites (Figure 3). ChIP assays in lymphoblasts confirmed the interaction of NRF-1 with the paternal allele of the *SNRPN* 5′ region (Figure 6B), but failed to detect association of CTCF with either the paternal or the maternal allele (Supplementary Figure S3B). This would suggest that footprint P6 on the paternal allele may be generated by interaction with NRF-1, and that it is unlikely that CTCF interacts with neither P6 nor M6 in these cells Therefore, it is currently unclear what factor(s) is generating the M6 footprint on the maternal allele in these lymphoblast cells.

Our *in vivo* footprint analysis of the *SNRPN* promoter included all six sequences that are phylogenetically conserved between the human and the mouse homologous regions (7). However, only one of these sequences was *in vivo* footprinted in lymphoblasts (footprint P5), which suggests that the other five phylogenetically conserved sequences might not be functional *cis*-acting elements in lymphoblasts and may be restricted to function in other somatic cell types, the germ line, and/or in early embryo development. In addition, Kantor *et al.* (16) have identified a series of sequences within the mouse *Snrpn* promoter (not including the SBE) that are involved in the correct imprinting of a transgene in mice that includes the mouse *Snrpn* promoter and the human AS-SRO. However, these mouse promoter sequences do not overlap with any of the six phylogenetically conserved sequences and are not present in the human *SNRPN* promoter. Therefore, it is not possible at this time to draw a comparison between the sequences identified by Kantor *et al.* (16) and the footprints identified in our study.

The region associated with DHS1 is located within the PWS-SRO and has been proposed to be involved in PWS-IC function (7,24) as well as in *SNRPN* promoter function. Thus, one or more of the *cis*-acting regulatory sequences associated with DHS1 identified here by *in vivo* footprinting may be essential not only for *SNRPN* expression as shown by our transient expression assays, but may also be functional elements of the PWS-IC.

### The enhancer associated with DHS2

Transient expression assays demonstrated that intronic sequences associated with DHS2 could function as an enhancer of the *SNRPN* promoter (Figure 4A) and included a highly conserved region we have termed the CAS. The CAS contains evolutionarily conserved binding sites for three known transcription factors, NRF-1, YY1 and Sp1 (Figure 5A), as well as two other highly conserved sequences that may serve as binding sites for transcription factors that have not yet been identified. One of these latter two sequences in the CAS (from nucleotides +1302 to +1310) is highly homologous to the MPI2 sequence in the mouse *Snrpn* promoter, which has been shown to be involved in maintenance of paternal imprinting of a construct in transgenic mice that contains the human AS-SRO and the mouse *Snrpn* promoter, though the corresponding DNA-binding factor for MPI2 element in the mouse promoter has not been identified (16).

In addition to activating the *SNRPN* promoter, our data have shown that the intronic enhancer can also function to activate transcription of heterologous promoters in transient expression assays, including the *MKRN3* promoter (Figure 4C). Thus, it is conceivable that the enhancer could also contribute to PWS-IC function by activating the promoters of other genes in the AS/PWS domain that are expressed only from the paternal chromosome (e.g. *MKNR3, NDN* and *MAGEL2*). This would be consistent with the model of PWS-IC function proposed by Brannan and Bartolomei (39) in which the PWS-IC acts as a positive regulator of genes expressed only from the paternal chromosome.

ChIP analysis showed that NRF-1, YY1 and Sp1 are all preferentially associated with the CAS region on the paternal allele *in vivo* (Figure 6B). Gel-shift analysis of the potential

YY1 binding site demonstrated the ability of YY1 to bind to this site *in vitro* (Supplementary Figure S2C), and transient expression assays showed that this YY1 binding site is essential for the activation function of this intronic region (Supplementary Figure S2A). YY1 is a ubiquitous factor that has been shown to act as a transcriptional initiator, activator and repressor (27), suggesting the possibility that YY1 could be involved in mediating both activation and silencing of the *SNRPN* locus, depending upon the parent-of-origin of the allele. A role for YY1 in the regulation of imprinted genes has been postulated for the imprinted PEG3 locus where YY1 appears to act as an insulator-binding protein (40). YY1 is also reported to associate with the nuclear matrix (41), suggesting a possible role for the CAS in tethering this region of the paternally inherited *SNRPN* locus to the matrix.

Similar to previous reports for the promoter region of *SNRPN* (31,32), the intronic CAS on the paternal chromosome (but not on the maternal chromosome) is associated with histone modification patterns characteristic of transcriptionally active chromatin (33) (Figure 6C). This is consistent with a role for the CAS in regulating *SNRPN* transcription *in vivo*. In addition, both YY1 and NRF-1 are known to interact with histone modifying enzymes (27,42), suggesting that these factors may have a role in establishing and/or maintaining differential patterns of histone modification on the paternal and/or maternal chromosomes.

### The role of NRF-1 in the regulation of the *SNRPN* locus

DNA sequence analysis and ChIP assays indicate that NRF-1 interacts with both the promoter region and the CAS on the paternal allele (Figure 6B). The upstream *SNRPN* promoters U1A and U1B also contain potential NRF-1 binding sites; however, no interaction of NRF-1 was detected with either allele of the U1A promoter (Figure 6B), which is not active in blood cells (5). In addition, we have identified by sequence analysis a conserved potential NRF-1 binding site in the *NDN* promoter region, which coincides with a sequence that is *in vivo* footprinted on the paternal *NDN* allele only (43). The fact that NRF-1 may be regulating at least some of the genes in the PWS/AS region is interesting because of the involvement of NRF-1 in the regulation of genes related to mitochondrial biogenesis and function, metabolism (including growth factor receptors and factors involved in glucose homeostasis), DNA replication and transcriptional regulation (26). This suggests that genes in the AS/PWS region and genes that function in metabolism and in cellular energetics may be co-regulated through the common transcriptional regulator NRF-1. This would further suggest a potential link between energy metabolism and aspects of the PWS phenotype (e.g. obesity and growth factor deficiency). However, the resting metabolic rate of PWS patients does not seem to differ from that of normal obese individuals (44).

### Potential roles of the CAS

Using antibodies against unphosphorylated pol II in ChIP assays, we found that the non-elongating form of pol II was significantly enriched at the CAS (Figure 6B), an intronic region that would normally be expected to be associated only with elongating phosphorylated pol II. This suggested that the CAS may be involved in recruiting and accumulating

the initiating form of pol II at the *SNRPN* locus. This recruitment of pol II to the CAS could be mediated by YY1, which is known to interact with and recruit pol II to promoters (27), and enhanced by Sp1 and NRF-1. Furthermore, recruitment of pol II to the CAS may serve as a mechanism for facilitating activation and transcription of the *SNRPN* gene by transfer of pol II from the CAS to the promoter and/or by increasing the local concentration of pol II in the vicinity of the *SNRPN* promoter. These mechanisms of promoter activation by distal regulatory elements have been proposed for a variety of other loci. For example, the locus control region (LCR) of the major histocompatibility complex (MHC) class II locus has been proposed to recruit pol II and transfer it to the promoter via either looping or tracking mechanisms (45). Alternatively, the LCR of the β-globin locus (46) has been postulated to recruit RNA pol II (and other *trans*-acting factors) by formation of a holocomplex or active chromatin hub (ACH) (47). The formation of an ACH by the PWS-IC to coordinately activate transcription of genes expressed from the paternal chromosome would be consistent with the model for PWS-IC function proposed by Brannan and Bartolomei (39). Thus, we speculate that the PWS-IC may function via formation of an ACH. Regulatory elements within the *SNRPN* 5′ region, including those associated with both DHS1 and DHS2 (and the CAS) may participate in the formation of a holocomplex (i.e. a chromatin hub) that interacts with and recruits regulatory regions from each of the paternally expressed genes (e.g. *MAGEL2*, *NDN* and *MKRN3*) to form an ACH specifically on the paternally inherited chromosome. This ACH would coordinately facilitate and activate transcription of the genes that are expressed exclusively on the paternal chromosome, in part by creating a highly localized region within the nucleus that is enriched in *trans*-acting positive regulators of transcription, such as histone acetyltransferases, general transcription factors and RNA polymerase II. Participation of the CAS in formation of the ACH would be consistent with our finding of unphosphorylated pol II at the CAS. As a variation of this model, it is also possible that the PWS-IC acts by facilitating recruitment of genes expressed from the paternal chromosome into transcription factories described by Osborne *et al.* (48). Both models would be consistent with the long-range interactions reported between differentially methylated regulatory regions that regulate imprinting in the *Igf2/H19* domain (49).

Participation of the CAS in formation of an ACH would also suggest a role for the CAS in PWS-IC function. The location of the CAS relative to targeted deletions of the mouse *Snrpn* gene that serve as mouse models for murine AS/PWS imprinting defects provide evidence consistent with such a role. A targeted 35 kb deletion extending 19 kb upstream and 16 kb downstream of the mouse *Snrpn* promoter region that included both the mouse CAS and *Snrpn* promoter region resulted in 100% neonatal lethality in mice when paternally inherited (50). This was accompanied by a loss PWS-IC function, including loss of paternal-specific gene expression and DNA methylation patterns of other imprinted genes in the region. These results suggested that the murine PWS-IC is contained within the 35 kb deletion. A second targeted deletion in this region, in which 0.9 kb of the *Snrpn* promoter region and exon 1 were deleted, showed no disruption of imprinted gene expression or DNA methylation patterns,

suggesting that the *Snrpn* promoter region does not, by itself, function as the PWS-IC in the mouse (19). However, a 4.8 kb deletion that included the previously deleted 0.9 kb promoter region as well as sequences ∼1.9 kb upstream and ∼2 kb downstream yielded mice with a partial imprinting defect, and 40–50% perinatal lethality when paternally inherited (19). This would suggest that sequences 1.9 kb upstream and/or 2 kb downstream of the previous 0.9 kb promoter deletion contain *cis*-acting elements that contribute to PWS-IC function in the mouse. The mouse CAS sequence is located ∼1.8 kb downstream of the *Snrpn* transcription initiation site and is, therefore, contained within the 2 kb region downstream of the promoter, suggesting that loss of the CAS could have contributed to the partial imprinting defect observed for the 4.8 kb *Snrpn* deletion. This would further suggest that the CAS may be a functional component of the PWS-IC in the mouse, and because of the high degree of sequence conservation of the CAS between the humans and rodents, the CAS may also act as a functional component of the human PWS-IC.

Because mice carrying the 4.8 kb deletion on the paternal chromosome do not exhibit the full mutant phenotype shown by mice carrying the paternally inherited 35 kb *Snrpn* deletion, it is conceivable that other regions within the 35 kb deletion (but outside of the 4.8 kb deletion) may also contribute to PWS-IC function in the mouse (16). In addition, Kantor *et al.* have proposed that components of the PWS-IC may be redundant, such that deletion of one component, e.g. the 0.9 kb *Snrpn* promoter deletion, may be compensated for by other components of the mouse PWS-IC (16). This functional redundancy in a multicomponent PWS-IC may explain why no human microdeletions that remove the CAS and not the *SNRPN* promoter have been identified to date.

## SUPPLEMENTARY MATERIAL

Supplementary Material is available at NAR Online.

## ACKNOWLEDGEMENTS

*Conflict of interest statement*. None declared.

## REFERENCES

1. Schmauss,C., Brines,M.L. and Lerner,M.R. (1992) The gene encoding the small nuclear ribonucleoprotein-associated protein N is expressed at high levels in neurons. *J. Biol. Chem.*, **267**, 8521–8529.
2. Gray,T.A., Saitoh,S. and Nicholls,R.D. (1999) An imprinted, mammalian bicistronic transcript encodes two independent proteins. *Proc. Natl Acad. Sci. USA*, **96**, 5616–5621.

3. Runte,M., Huttenhofer,A., Gross,S., Kiefmann,M., Horsthemke,B. and Buiting,K. (2001) The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. *Hum. Mol. Genet.*, **10**, 2687–2700.

4. Sutcliffe,J.S., Nakao,M., Christian,S., Orstavik,K.H., Tommerup,N., Ledbetter,D.H. and Beaudet,A.L. (1994) Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nature Genet.*, **8**, 52–58.

5. Dittrich,B., Buiting,K., Korn,B., Rickard,S., Buxton,J., Saitoh,S., Nicholls,R.D., Poustka,A., Winterpacht,A., Zabel,B. *et al.* (1996) Imprint switching on human chromosome 15 may involve alternative transcripts of the SNRPN gene. *Nature Genet.*, **14**, 163–170.

6. Nicholls,R.D. and Knepper,J.L. (2001) Genome organization, function, and imprinting in Prader-Willi and Angelman syndromes. *Annu. Rev. Genomics. Hum. Genet.*, **2**, 153–175.

7. Ohta,T., Gray,T.A., Rogan,P.K., Buiting,K., Gabriel,J.M., Saitoh,S., Muralidhar,B., Bilienska,B., Krajewska-Walasek,M., Driscoll,D.J. *et al.* (1999) Imprinting-mutation mechanisms in Prader-Willi syndrome. *Am. J. Hum. Genet.*, **64**, 397–413.

8. Buiting,K., Lich,C., Cottrell,S., Barnicoat,A. and Horsthemke,B. (1999) A 5-kb imprinting center deletion in a family with Angelman syndrome reduces the shortest region of deletion overlap to 880 bp. *Hum. Genet.*, **105**, 665–666.

9. Bielinska,B., Blaydes,S.M., Buiting,K., Yang,T., Krajewska-Walasek,M., Horsthemke,B. and Brannan,C.I. (2000) *De novo* deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. *Nature Genet.*, **25**, 74–78.

10. El-Maarri,O., Buiting,K., Peery,E.G., Kroisel,P.M., Balaban,B., Wagner,K., Urman,B., Heyd,J., Lich,C., Brannan,C.I. *et al.* (2001) Maternal methylation imprints on human chromosome 15 are established during or after fertilization. *Nature Genet.*, **27**, 341–344.

11. Perk,J., Makedonski,K., Lande,L., Cedar,H., Razin,A. and Shemer,R. (2002) The imprinting mechanism of the Prader-Willi/Angelman regional control center. *EMBO J.*, **21**, 5807–5814.

12. Nicholls,R.D., Saitoh,S. and Horsthemke,B. (1998) Imprinting in Prader-Willi and Angelman syndromes. *Trends Genet.*, **14**, 194–200.

13. Huq,A.H., Sutcliffe,J.S., Nakao,M., Shen,Y., Gibbs,R.A. and Beaudet,A.L. (1997) Sequencing and functional analysis of the SNRPN promoter: *in vitro* methylation abolishes promoter activity. *Genome Res.*, **7**, 642–648.

14. Green Finberg,Y., Kantor,B., Hershko,A.Y. and Razin,A. (2003) Characterization of the human Snrpn minimal promoter and *cis* elements within it. *Gene*, **304**, 201–206.

15. Hershko,A.Y., Finberg,Y., Kantor,B., Shemer,R. and Razin,A. (2001) The mouse Snrpn minimal promoter and its human orthologue: activity and imprinting. *Genes Cells*, **6**, 967–975.

16. Kantor,B., Makedonski,K., Green-Finberg,Y., Shemer,R. and Razin,A. (2004) Control elements within the PWS/AS imprinting box and their function in the imprinting process. *Hum. Mol. Genet.*, **13**, 751–762.

17. Kantor,B., Kaufman,Y., Makedonski,K., Razin,A. and Shemer,R. (2004) Establishing the epigenetic status of the Prader-Willi/Angelman imprinting center in the gametes and embryo. *Hum. Mol. Genet.*, **13**, 2767–2779.

18. Shemer,R., Hershko,A.Y., Perk,J., Mostoslavsky,R., Tsuberi,B., Cedar,H., Buiting,K. and Razin,A. (2000) The imprinting box of the Prader-Willi/Angelman syndrome domain. *Nature Genet.*, **26**, 440–443.

19. Bressler,J., Tsai,T.F., Wu,M.Y., Tsai,S.F., Ramirez,M.A., Armstrong,D. and Beaudet,A.L. (2001) The SNRPN promoter is not required for genomic imprinting of the Prader-Willi/Angelman domain in mice. *Nature Genet.*, **28**, 232–240.

20. Chen,C. and Yang,T.P. (2001) Nucleosomes are translationally positioned on the active allele and rotationally positioned on the inactive allele of the HPRT promoter. *Mol. Cell. Biol.*, **21**, 7682–7695.

21. Glenn,C.C., Saitoh,S., Jong,M.T., Filbrandt,M.M., Surti,U., Driscoll,D.J. and Nicholls,R.D. (1996) Gene structure, DNA methylation, and imprinted expression of the human SNRPN gene. *Am. J. Hum. Genet.*, **58**, 335–346.

22. Hornstra,I.K. and Yang,T.P. (1993) *In vivo* footprinting and genomic sequencing by ligation-mediated PCR. *Anal. Biochem.*, **213**, 179–193.

23. Leach,K.M., Nightingale,K., Igarashi,K., Levings,P.P., Engel,J.D., Becker,P.B. and Bungert,J. (2001) Reconstitution of human beta-globin locus control region hypersensitive sites in the absence of chromatin assembly. *Mol. Cell. Biol.*, **21**, 2629–2640.

24. Schweizer,J., Zynger,D. and Francke,U. (1999) *In vivo* nuclease hypersensitivity studies reveal multiple sites of parental origin-dependent differential chromatin conformation in the 150 kb SNRPN transcription unit. *Hum. Mol. Genet.*, **8**, 555–566.

25. Bouwman,P. and Philipsen,S. (2002) Regulation of the activity of Sp1-related transcription factors. *Mol. Cell. Endocrinol.*, **195**, 27–38.

26. Cam,H., Balciunaite,E., Blais,A., Spektor,A., Scarpulla,R.C., Young,R., Kluger,Y. and Dynlacht,B.D. (2004) A common set of gene regulatory networks links metabolism and growth inhibition. *Mol. Cell*, **16**, 399–411.

27. Thomas,M.J. and Seto,E. (1999) Unlocking the mechanisms of transcription factor YY1: are chromatin modifying enzymes the key? *Gene*, **236**, 197–208.

28. Takai,D. and Jones,P.A. (2002) Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl Acad. Sci. USA*, **99**, 3740–3745.

29. Breslin,M.B. and Vedeckis,W.V. (1998) The human glucocorticoid receptor promoter upstream sequences contain binding sites for the ubiquitous transcription factor, Yin Yang 1. *J. Steroid. Biochem. Mol. Biol.*, **67**, 369–381.

30. Filippova,G.N., Thienes,C.P., Penn,B.H., Cho,D.H., Hu,Y.J., Moore,J.M., Klesert,T.R., Lobanenkov,V.V. and Tapscott,S.J. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nature Genet.*, **28**, 335–343.

31. Saitoh,S. and Wada,T. (2000) Parent-of-origin specific histone acetylation and reactivation of a key imprinted gene locus in Prader-Willi syndrome. *Am. J. Hum. Genet.*, **66**, 1958–1962.

32. Xin,Z., Allis,C.D. and Wagstaff,J. (2001) Parent-specific complementary patterns of histone H3 lysine 9 and H3 lysine 4 methylation at the Prader-Willi syndrome imprinting center. *Am. J. Hum. Genet.*, **69**, 1389–1394.

33. Liang,G., Lin,J.C., Wei,V., Yoo,C., Cheng,J.C., Nguyen,C.T., Weisenberger,D.J., Egger,G., Takai,D., Gonzales,F.A. *et al.* (2004) Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc. Natl Acad. Sci. USA*, **101**, 7357–7362.

34. Valverde-Garduno,V., Guyot,B., Anguita,E., Hamlett,I., Porcher,C. and Vyas,P. (2004) Differences in the chromatin structure and *cis*-element organization of the human and mouse GATA1 loci: implications for *cis*-element identification. *Blood*, **104**, 3106–3116.

35. Mann,J.R., Szabo,P.E., Reed,M.R. and Singer-Sam,J. (2000) Methylated DNA sequences in genomic imprinting. *Crit. Rev. Eukaryot. Gene Expr.*, **10**, 241–257.

36. DeGregori,J. (2002) The genetics of the E2F family of transcription factors: shared functions and unique roles. *Biochim. Biophys. Acta*, **1602**, 131–150.

37. Chao,W., Huynh,K.D., Spencer,R.J., Davidow,L.S. and Lee,J.T. (2002) CTCF, a candidate *trans*-acting factor for X-inactivation choice. *Science*, **295**, 345–347.

38. Lewis,A. and Murrell,A. (2004) Genomic imprinting: CTCF protects the boundaries. *Curr. Biol.*, **14**, R284–R286.

39. Brannan,C.I. and Bartolomei,M.S. (1999) Mechanisms of genomic imprinting. *Curr. Opin. Genet. Dev.*, **9**, 164–170.

40. Kim,J., Kollhoff,A., Bergmann,A. and Stubbs,L. (2003) Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum. Mol. Genet.*, **12**, 233–245.

41. Bushmeyer,S.M. and Atchison,M.L. (1998) Identification of YY1 sequences necessary for association with the nuclear matrix and for transcriptional repression functions. *J. Cell. Biochem.*, **68**, 484–499.

42. Izumi,H., Ohta,R., Nagatani,G., Ise,T., Nakayama,Y., Nomoto,M. and Kohno,K. (2003) p300/CBP-associated factor (P/CAF) interacts with nuclear respiratory factor-1 to regulate the UDP-N-acetyl-alpha-D-galactosamine: polypeptide N-acetylgalactosaminyltransferase-3 gene. *Biochem. J.*, **373**, 713–722.

43. Hanel,M.L., Lau,J.C., Paradis,I., Drouin,R. and Wevrick,R. (2005) Chromatin modification of the human imprinted NDN (necdin) gene detected by *in vivo* footprinting. *J. Cell. Biochem.*, **94**, 1046–1057.

44. Goldstone,A.P., Brynes,A.E., Thomas,E.L., Bell,J.D., Frost,G., Holland,A., Ghatei,M.A. and Bloom,S.R. (2002) Resting metabolic rate, plasma leptin concentrations, leptin receptor expression, and adipose tissue measured by whole-body magnetic resonance imaging in women with Prader-Willi syndrome. *Am. J. Clin. Nutr.*, **75**, 468–475.

45. Masternak,K., Peyraud,N., Krawczyk,M., Barras,E. and Reith,W. (2003) Chromatin remodeling and extragenic transcription at the MHC class II locus control region. *Nature Immunol.*, **4**, 132–137.

46. Johnson,K.D., Christensen,H.M., Zhao,B. and Bresnick,E.H. (2001) Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol. Cell*, **8**, 465–471.

47. de Laat,W. and Grosveld,F. (2003) Spatial organization of gene expression: the active chromatin hub. *Chromosome Res.*, **11**, 447–459.

48. Osborne,C.S., Chakalova,L., Brown,K.E., Carter,D., Horton,A., Debrand,E., Goyenechea,B., Mitchell,J.A., Lopes,S., Reik,W. *et al.* (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nature Genet.*, **36**, 1065–1071.

49. Murrell,A., Heeson,S. and Reik,W. (2004) Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nature Genet.*, **36**, 889–893.

50. Yang,T., Adamson,T.E., Resnick,J.L., Leff,S., Wevrick,R., Francke,U., Jenkins,N.A., Copeland,N.G. and Brannan,C.I. (1998) A mouse model for Prader-Willi syndrome imprinting-centre mutations. *Nature Genet.*, **19**, 25–31.