1596–1608 *Nucleic Acids Research, 2012, Vol. 40, No. 4* *Published online 7 November 2011*
doi:10.1093/nar/gkr964

YY1 associates with the macrosatellite DXZ4 on the inactive X chromosome and binds with CTCF to a hypomethylated form in some male carcinomas

Shawn C. Moseley[1], Raed Rizkallah[2], Deanna C. Tremblay[1], Blair R. Anderson[3], Myra M. Hurt[2] and Brian P. Chadwick[1,*]

[1]Department of Biological Science, Florida State University, Tallahassee, FL 32306-4295, [2]Department of Biomedical Sciences, Florida State University, Tallahassee, FL 32306-2370 and [3]Department of Medicine, Duke University School of Medicine, Durham, NC 27710, USA

Received July 13, 2011; Revised October 13, 2011; Accepted October 14, 2011

ABSTRACT

DXZ4 is an X-linked macrosatellite composed of 12–100 tandemly arranged 3-kb repeat units. In females, it adopts opposite chromatin arrangements at the two alleles in response to X-chromosome inactivation. In males and on the active X chromosome, it is packaged into heterochromatin, but on the inactive X chromosome (Xi), it adopts a euchromatic conformation bound by CTCF. Here we report that the ubiquitous transcription factor YY1 associates with the euchromatic form of DXZ4 on the Xi. The binding of YY1 close to CTCF is reminiscent of that at other epigenetically regulated sequences, including sites of genomic imprinting, and at the X-inactivation centre, suggesting a common mode of action in this arrangement. As with CTCF, binding of YY1 to DXZ4 *in vitro* is not blocked by CpG methylation, yet *in vivo* both proteins are restricted to the hypomethylated form. In several male carcinoma cell lines, DXZ4 can adopt a Xi-like conformation in response to cellular transformation, characterized by CpG hypomethylation and binding of YY1 and CTCF. Analysis of a male melanoma cell line and normal skin cells from the same individual confirmed that a transition in chromatin state occurred in response to transformation.

INTRODUCTION

Now that most of the human genome sequence has been assembled (1), one of several pressing issues is to determine the role of the various DNA elements in establishing and maintaining the gene-expression profiles that underlie different cell types. Because only a relatively small portion of the genome actually codes protein (2), about half is composed of repetitive DNA (1), and a significant fraction is transcribed into non-coding RNA (3), many challenges lie ahead. Intriguing aspects of the human genome are the extensive variation among individuals in copy number (copy-number variation, CNV) (4) and the consequences associated with such diversity.

Macrosatellites are an extreme form of CNV. They are composed of individual repeat units, typically >1 kb in size, that are arranged in tandem, often spanning hundreds of kilobases (5). Some are chromosome specific, such as ZAV at chromosome 9q32 (6) and DXZ4 at chromosome Xq23 (7,8), whereas others are found on at least two chromosomes, such as D4Z4 on chromosomes 4q35 and 10q26 (9,10) or RS447 on chromosome 4p15 and 18p23 (11). The macrosatellite about which most is known is D4Z4, because of its association with onset of facioscapulohumeral muscular dystrophy (FSHD) (12). FSHD, the third most common inherited form of muscular dystrophy (OMIM 158 900), is primarily manifested as progressive muscle atrophy of the face, shoulders and upper arms and is often accompanied by gradual spread of symptoms to the lower body (13). D4Z4 is a tandem array of as many as 100 3.3-kb repeat units. Contraction of the macrosatellite to fewer than 11 repeat units on a permissive chromosomal haplotype (14) is associated with disease onset (15). Recent data implicate an altered chromatin structure for the contracted array (16–18) and stabilization of transcripts originating from the distal edge of the macrosatellite as the molecular basis of the disease (19). Collectively, these data highlight the importance of macrosatellite CNV for disease susceptibility.

Primary DNA-sequence conservation of the X-linked macrosatellite DXZ4 is restricted to higher primates (20). In humans it is composed of as few as 12 to over

*To whom correspondence should be addressed. Tel: +1 850 645 9279; Fax: +1 850 645 8447; Email: chadwick@bio.fsu.edu

100 3-kb GC-rich repeat units located at Xq23 (7,8). Like many other macrosatellites (6,21,22), DXZ4 is expressed (8,23), but unlike that of D4Z4 (24) its transcript contains no obvious conserved open reading frames. X-linkage exposes DXZ4 to the process of X-chromosome inactivation (XCI), a form of dosage compensation employed by mammals to balance the levels of X-linked gene expression in the two sexes (25). Early in female development, one of the two X chromosomes is chosen to become the future inactive X chromosome (Xi) (26). Soon thereafter most gene expression from the Xi is shut down (27), and silencing is stably maintained as the chromosome is repackaged into facultative heterochromatin (28). Although most of the X chromosome adopts this new chromatin structure, DXZ4 does not conform and instead adopts a more euchromatic organization characterized by CpG hypomethylation (7,23), dimethylation of histone H3 at lysine residue 4 (H3K4me2), acetylation of histone H3 at lysine 9 (H3K9Ac), and association with the epigenetic organizer protein CCCTC-binding factor (CTCF) (23). In contrast, DXZ4 on the active X chromosome (Xa) and male X is organized into constitutive heterochromatin characterized by CpG hypermethylation (7,23), trimethylation of histone H3 at lysine 9 (H3K9me3) (23) and association with heterochromatin protein 1 gamma (HP1g) (18). Intriguingly, DXZ4 shows several parallels with the mouse X-inactivation centre (Xic), a region of the X chromosome required for XCI (26). These include differential CpG methylation (29,30), transcription of non-coding RNAs (26), a tandem repeat sequence (30) and association with Ctcf (31). In addition, the ubiquitous zinc-finger protein Yin-Yang 1 (YY1) associates with the Xic (32).

Here, we report work showing that YY1 associated with DXZ4 specifically on the Xi alongside CTCF. Although *in vitro* CpG methylation was unable to block YY1 binding, YY1 and CTCF were restricted *in vivo* to the hypomethylated DXZ4 allele on the Xi. Furthermore, we report changes to DXZ4 chromatin on the male X as a result of cell transformation.

MATERIALS AND METHODS

Antibodies

Commercial primary antibodies used in this study were obtained from the following sources: mouse anti-CTCF (612 149), BD Biosciences. Rabbit anti-CTCF (07–729), rabbit anti-H3K4me2 (07–030) and rabbit anti-H3K9me3 (07–523), Millipore. Rabbit anti-GAPDH (sc-25778), goat anti-YY2 (Sc-47637) and rabbit anti-YY1 (sc-1703), Santa Cruz Biotechnology. Rabbit anti-YY1 (AB1, AV38301). Sigma-Aldrich. Commercial secondary antibodies used in this study, Alexa Fluor-488 goat anti-rabbit (A11001) and Alex Fluor-555 goat anti-mouse (A21422), were obtained from Invitrogen Corporation.

Cells and cell lines

Human telomerase-immortalized female retinal pigment epithelia (hTERT-RPE1), female mammary epithelia (hTERT-HME1) and male foreskin fibroblasts (hTERT-BJ1) were all originally obtained from Clontech. All are now available from the American Type Culture Collection (ATCC). Primary male skin fibroblast cells CCD1139Sk (CRL-2708) and Malme-3 (HTB-102) were obtained from ATCC. The following cell lines were all obtained from ATCC: female cervical adenocarcinoma cell line HeLa (CCL-2); Malme-3M (HTB-64), derived from a male malignant melanoma; male hepatocellular carcinoma cell line HepG2 (HB-8065); male fibrosarcoma cell line HT-1080 (CCL-121); male colorectal adenocarcinoma cell lines SW620 (CCL-227), SW480 (CCL-228), HCT-15 (CCL-225), DLD-1 (CCL-221), Caco-2 (HTB-37), HCT116 (CCL-247) and SW1116 (CCL-233). All cells were maintained according to the supplier recommendations. Tissue-culture media and supplements were purchased from Invitrogen Corporation.

Immunofluorescence and fluorescence *in situ* hybridization

Cells were grown directly on glass microscope slides before indirect immunofluorescence. They were washed with 1× phosphate-buffered saline (PBS) before fixation and extraction in fixative for 10 min (1× PBS, 0.1% Triton X-100, 3.7% formaldehyde). Cells were washed twice with 1× PBS before being blocked for 30 min (3% bovine serum albumin, BSA; 1× PBS; 0.1% Tween-20) and were then washed three times for 2 min each in 1× PBS before incubation for 60 min with the primary antibody (1:100 dilution of primary antibody in 1× PBS, 0.1% Tween-20, 1% BSA). Cells were washed as above before incubation with the secondary antibody (1:100 dilution of secondary antibody in 1× PBS, 0.1% Tween-20, 1% BSA), then washed once more and fixed as above before two washes of 2 min in 1× PBS and addition of anti-fade-containing 4′,6-diamidino-2-phenylindole (DAPI). All steps were performed at room temperature.

Indirect immunofluorescence combined with fluorescence *in situ* hybridization (FISH) was performed as above with the following additional steps before application of DAPI. Cells were dehydrated for 3 min each in 70% and 100% ethanol before being air-dried. Dried cells were denatured in 70% formamide, 2× saline–sodium citrate buffer (2× SSC; pH 7.0) for 10 min at 83°C, before dehydration in ice-cold 70% and 100% ethanol for 3 min each and air-drying. A Spectrum Orange direct-labeled DXZ4 probe was prepared by nick translation in the presence of Spectrum Orange dUTP according to the manufacturer's instructions (Abbot Molecular). The DXZ4 probe was resuspended in Hybrisol VII (MP Biomedicals) and denatured at 78°C for 5 min before being placed on ice. The probe was applied to the cells and sealed under a cover-slip with rubber cement and incubated overnight at 37°C in a humidified chamber. The following day, the cells were washed twice for 8 min each in 50% formamide, 2× SSC at 42°C, then washed once for 8 min in 2× SSC at 42°C, before being mounted in DAPI.

Imaging was performed on a DeltaVision pDV. The images were deconvolved with softWoRx 3.7.0 (Applied

Precision) and compiled with Adobe Photoshop CS2 (Adobe Systems).

Chromatin immunoprecipitation

Cells were fixed in culture media for 10 min at room temperature by addition of formaldehyde to 1% final concentration. Cross-linking was quenched for 5 min by addition of glycine to 125 mM final concentration. Cells were washed and collected with ice-cold 1× PBS supplemented with 0.1 mg/ml phenylmethylsulfonyl fluoride. Cells were resuspended at 7×10^{-6} cells/ml in lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris, pH 8.0, containing 2 μg/ml Aprotinin, 2 μg/ml Leupeptin, 1 μg/ml Pepstatin and 0.1 mg/ml phenylmethylsulfonyl fluoride). Chromatin was sheared with a Bioruptor (Diagenode) set at high power, 30 s on, 30 s off. Sonication was performed with 0.2 ml of lysate per 1.7-ml tube for 12 cycles; the bath was cooled with ice every fourth cycle. Lysate was precleared with protein-A agarose beads, before addition of primary antibody and incubation overnight at 4°C. Immune complexes were collected the next day with protein-A agarose and washed at 4°C twice for 5 min each with low wash buffer (0.1% SDS, 1% Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris, pH 8.0), once for 5 min with high wash buffer (0.1% SDS, 1% Triton X-100, 500 mM NaCl, 20 mM Tris, 8.0) and once for 5 min with TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA). Protease inhibitors were used at the same concentration as in the lysis buffer. Chromatin was eluted at room temperature with 100 mM $NaHCO_3$, 1% SDS and cross-links reversed overnight at 65°C after addition of NaCl to 0.2 M. Residual RNA was removed for 30 min at 37°C with RNase A, then protein by a 120-min incubation at 42°C with proteinase K. DNA was purified with the QIAquick PCR purification kit (Qiagen). Protein-A agarose, RNase A, proteinase K and all protease inhibitors were obtained from Roche Applied Science.

DNA immunoprecipitated with either CTCF or YY1 was assessed by PCR with DXZ4-F23 (GGACAGTCCC AAGCCACTC) and DXZ4-R26 (AGATGCTGATCCG CCATGTG). DNA immunoprecipitated with either H3K4me2 or H3K9me3 was assessed by PCR with DXZ4-F19 (GAGATGCCCATGAACTCAAG) and DXZ4-R19 (GCCAGGGGGATAGGTGTG). All oligonucleotides we used were obtained from Eurofins MWG Operon.

Electrophoresis mobility shift assays

Whole-cell extracts were prepared from HeLa S3 cells as described previously (33), with the exception that phosphatase inhibitors were not included. Oligonucleotides were generated to the following DNA sequences along with oligonucleotides to the complementary strand: DXZ4-YY1 (CGCCCCGCACATGGCGGATCAG), DXZ4-YY1-Mut (CGCCCCGCATTGGGCGGA TCAG), DXZ4-YY1L (GGAAAAAACGCCAACAGC GCCCCGCACATGGCGGATCAG). We labeled the double-stranded oligomers radioactively with T4-polynucleotide kinase (New England Biolabs) in the buffer designated by the manufacturer and (γ-^{32}P) ATP (Perkin Elmer-Cetus). HeLa whole-cell extracts were incubated with the ^{32}P-labeled double-stranded oligomers on ice for 30 min in binding buffer (10 mM Tris, pH 7.5, 50 mM NaCl, 1 mM DTT, 5% glycerol), in the presence of 1 μg dIdC and 1 μg of non-specific double-stranded DNA oligomers. The protein-DNA complexes were then separated on 4% polyacrylamide gels and fixed for 15 min (10% acetic acid, 10% methanol) before drying. Dried gels were exposed to a phosphoimager screen and scanned with a Typhoon 9410 Imager (GE Healthcare, Piscataway, NJ, USA). We assayed competition by adding cold oligomers and performed supershifts by adding specific anti-YY1, anti-YY2, or non-specific antibody (anti-GAPDH) to the binding reactions, as indicated. Using double-stranded DXZ4-YY1L as a template, we methylated the C's of CpG dinucleotides in the sequence with M.SssI according to the manufacturers' recommendations (New England Biolabs). Methylation was assayed with the methyl-sensitive restriction endonuclease HinP1I (New England Biolabs). Both methylated and non-methylated DXZ4-YY1Ls were then used in mobility-shift assays as described above. Oligonucleotides modified with 5-methyl Cytosine were synthesized directly with the same sequence as DXZ4-YY1 above; only the three C's in the CpG context were methylated for the forward and reverse sequence (Eurofins MWG Operon). Nuclear extracts from HeLa cells stably overexpressing Flag-epitope tagged YY1 (33) and purification of bacterially expressed non-tagged YY1 protein (34) have been described previously.

Bisulfite analysis

Genomic DNA was isolated from cells with the Qiagen Blood and Cell Culture DNA Midi Kit (Qiagen, Valencia, CA, USA). Bisulfite modification was performed with the Qiagen EpiTect Bisulfite Kit according to the manufacturer's instructions. We performed PCR using the primers DXZ4-Bis-F1 (CCAAACAAACTACCCAAAACC) and DXZ4-Bis-R2 (GAAGGTAGGTTAGTAAGAAGG), which amplify a 535-bp fragment of modified DXZ4. PCR products were cloned into pDrive with the Qiagen PCR cloning kit. Plasmid DNA was isolated from clones and DNA sequence was supplied by the sequencing services of Eurofins MWG Operon.

RESULTS

The Xi is the largest mass of facultative heterochromatin in a female nucleus; at interphase it is typically located at the nuclear periphery in a structure called the Barr body (35). Because many chromatin proteins and euchromatin associated histone modifications are underrepresented in or absent from the Barr body, the territory of the Xi appears as a hole in the distribution pattern obtained when immunofluorescence is performed with antibodies to any of these chromatin features (36). Because DXZ4 is packaged into euchromatin and bound by CTCF on the Xi (23), immunofluorescence for either CTCF or H3K4me2 readily detects an intense focus within the territory of the Xi, a pattern that is shared with several other

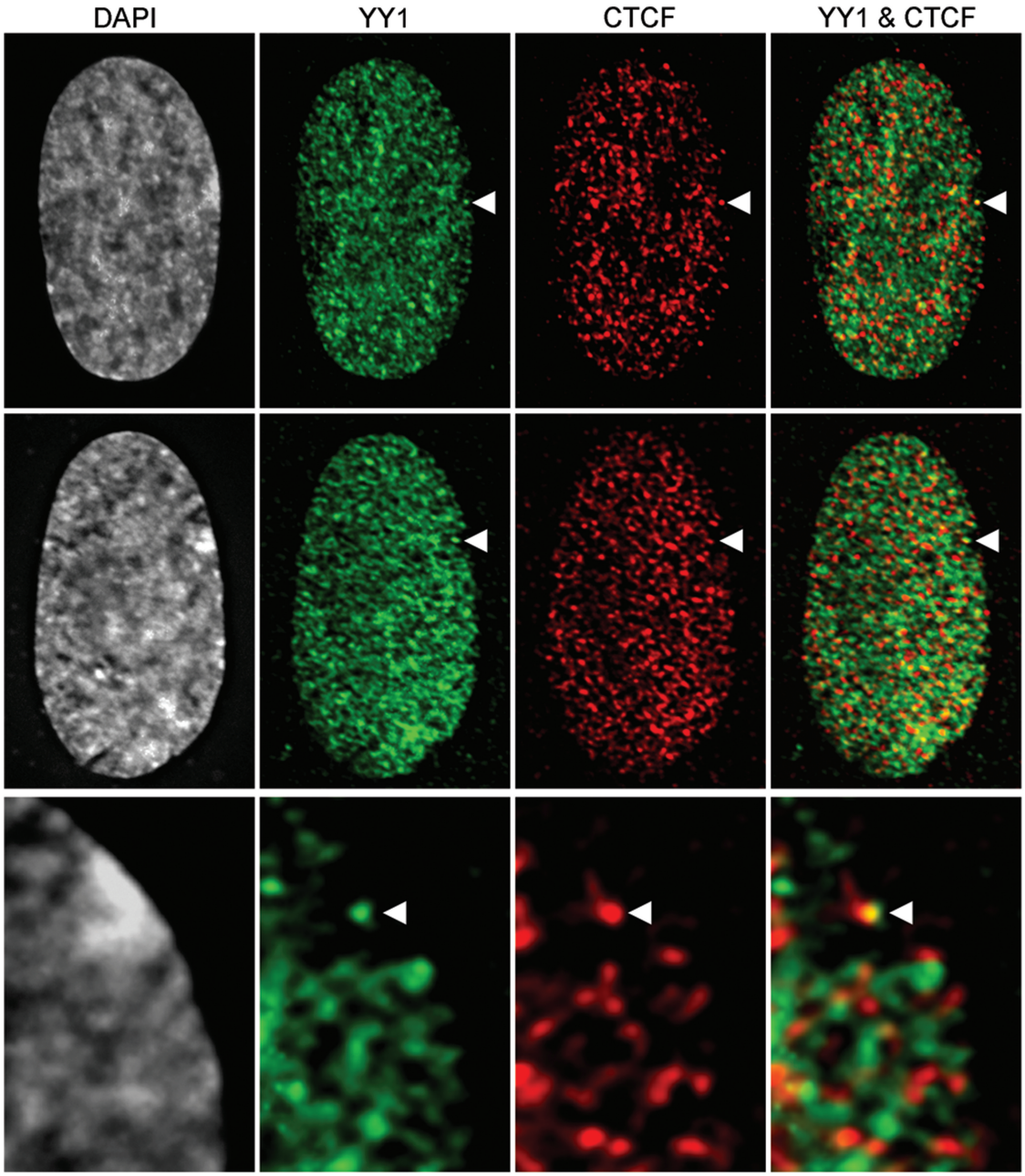

Figure 1. Examples of indirect immunofluorescence performed with anti-YY1 (H-414, sc-1703), showing the distribution of YY1 (green) and CTCF (red) in female hTERT-RPE1 nuclei on the inactive X chromosome. The 4′,6-diamidino-2-phenylindole (DAPI) images are shown in black and white, which better emphasize the densely staining Barr body. The white arrow-head points to the intense CTCF and YY1 signal within the territory of the Barr body. The bottom panel is a third example zoomed in to show more clearly that YY1 overlaps with the CTCF signal.

DXZ4-associated proteins and chromatin modifications (36).

YY1 associates with DXZ4 on the inactive X chromosome

In our indirect immunofluorescence assays, YY1 showed a general nuclear staining pattern with an obvious underrepresentation at the Xi except at a single intense focus (Figure 1). Consistent with previous data (23,36), CTCF showed a general nuclear distribution and a dot within the Barr body. Interestingly, the Xi-associated YY1 and CTCF signals overlapped.

The combination of YY1 immunofluorescence with FISH with a direct-labeled DXZ4 probe indicated that the Xi-associated YY1 signal overlapped with DXZ4 (Figure 2A), but examination of the DXZ4 signal originating from the Xa showed little to no overlap with DXZ4, suggesting that YY1 association with DXZ4 is Xi specific.

To validate the immunofluorescence and FISH data and to confirm that YY1 is actually binding to DXZ4 and not simply near it, we performed chromatin immunoprecipitation (ChIP) using anti-YY1 on several

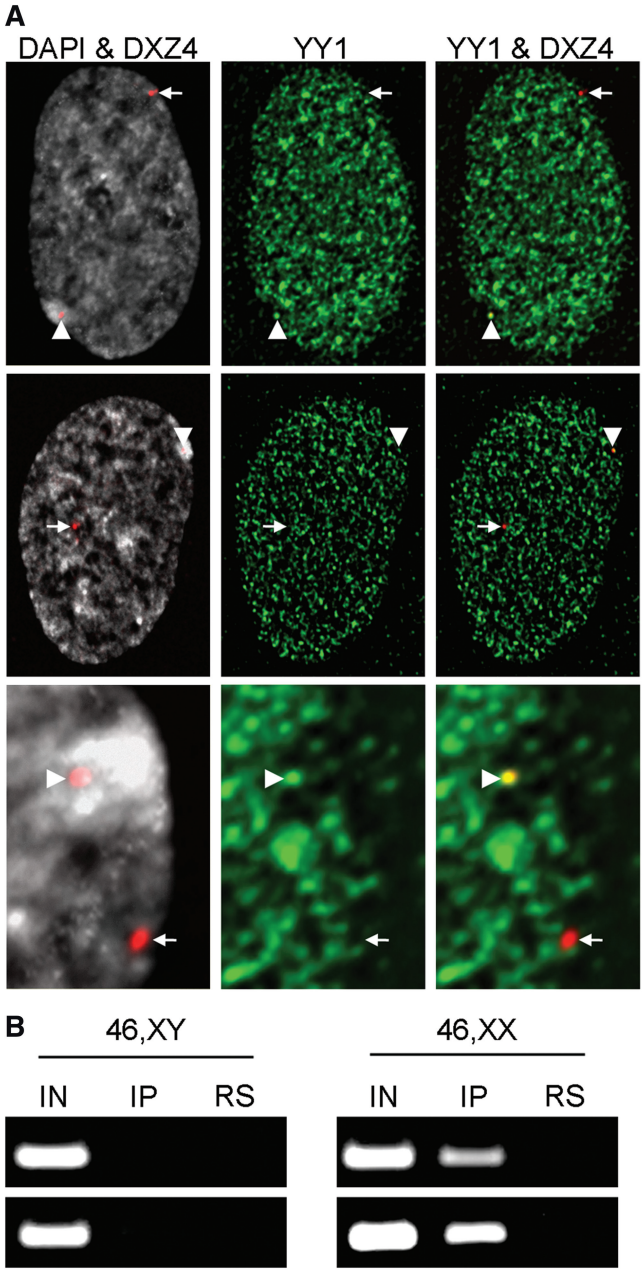

Figure 2. YY1 associates with DXZ4 on the Xi only. (**A**) Examples of female hTERT-RPE1 nuclei showing the distribution of YY1 by indirect immunofluorescence (green) combined with direct DNA fluorescence in situ hybridization with a probe for DXZ4 (red). DAPI-stained nuclei are shown in black and white, which better emphasize the Barr body (left column). The small white arrow indicates the location of DXZ4 on the active X chromosome. The white arrowhead indicates DXZ4 on the inactive X chromosome. The bottom sets of panels are a close-up of an image. Overlapping signals appear orange. (**B**) YY1 ChIP output assessed for DXZ4 association by PCR. Two male samples are shown on the left (top, hTERT-BJ1; bottom, CCD1139Sk) and two female samples on the right (top, hTERT-HME1; bottom, hTERT-RPE1). Samples assessed include the input (IN), eluted immunoprecipitation (IP) and an immunoprecipitation with non-specific rabbit serum (RS). PCR results shown use DXZ4-F23 and DXZ4-R26. Immunofluorescence and ChIP were performed with anti-YY1 (H-414, sc-1703). These results show that YY1 associates with DXZ4 on the inactive X chromosome only.

independent male and female cell cultures. When the immunoprecipitated DNA was assessed by PCR with primer sets covering subregions of a single 3-kb DXZ4 monomer (data not shown), only one primer set, encompassing a 186-bp fragment of DXZ4, consistently generated a product for the YY1 ChIP, a result that was only observed for the female samples. Figure 2B shows examples for two independent male samples (left panels) and two female samples (right panels). YY1 is therefore associated with DXZ4 on the Xi only.

YY1 binds to a histone alpha-like consensus sequence in DXZ4

When the 3-kb DNA sequence of a single DXZ4 repeating unit was examined for putative YY1-binding sites, a match was found with the 7-bp alpha element (37), a YY1-binding motif located in the coding region–activating sequence of the H2A.2 and H3.2 histone genes (38) (Figure 3A). Importantly, this sequence motif resides within the 186-bp sequence that is positive for YY1 ChIP. Alignment of the alpha element match with the DNA sequence of 60 different DXZ4 monomers from nine independent sources (8) revealed that the motif is 100% conserved (data not shown).

When double stranded DNA oligomers were generated for the 22-bp DXZ4 sequence shown in Figure 3A and labeled with ^{32}P, electrophoresis mobility shift assays (EMSA) revealed a clear shift in mobility in the presence of whole-cell extract. The shift could be specifically reduced in the presence of excess unlabeled oligomer but not when binding competition was performed with oligomers of which the core CAT sequence had been mutated (39) (Figure 3B, left four lanes). Binding in the presence of one of two independent anti-YY1 antibodies, but not with a non-specific antibody, resulted in a supershift of the labeled DNA, confirming that YY1 in the cell extract is binding to the predicted site in DXZ4 (Figure 3B, right three lanes).

In humans, a retrotransposed homolog of YY1 termed YY2 resides at Xp22.1–22.2 (40). It shares >95% amino-acid identity across the zinc-finger DNA-binding domain of YY1 (40) and recognizes similar DNA sequence motifs (41). Using a YY2 antibody, we did not see a supershift for the DXZ4 YY1 target site (Figure 3C). This result suggests that the shift that we do see is generated by YY1. Data supporting YY1 binding to DXZ4 *in vitro* are shown in Figure 3B, where an antibody raised to a portion of the amino-terminus of YY1 that is less well conserved in YY2 (AB1) shows a supershift comparable to that for antibodies raised to full-length YY1.

Binding of YY1 to its target sequence in DXZ4 is not affected by CpG methylation *in vitro*

Collectively these data support Xi-specific association of YY1 with DXZ4, at a site <80 nt from the Xi-specific CTCF-binding site (23). Like YY1, CTCF is a multi-zinc-finger DNA-binding protein with important roles in epigenetic regulation of gene expression and chromatin organization (42). At a number of genomic sites,

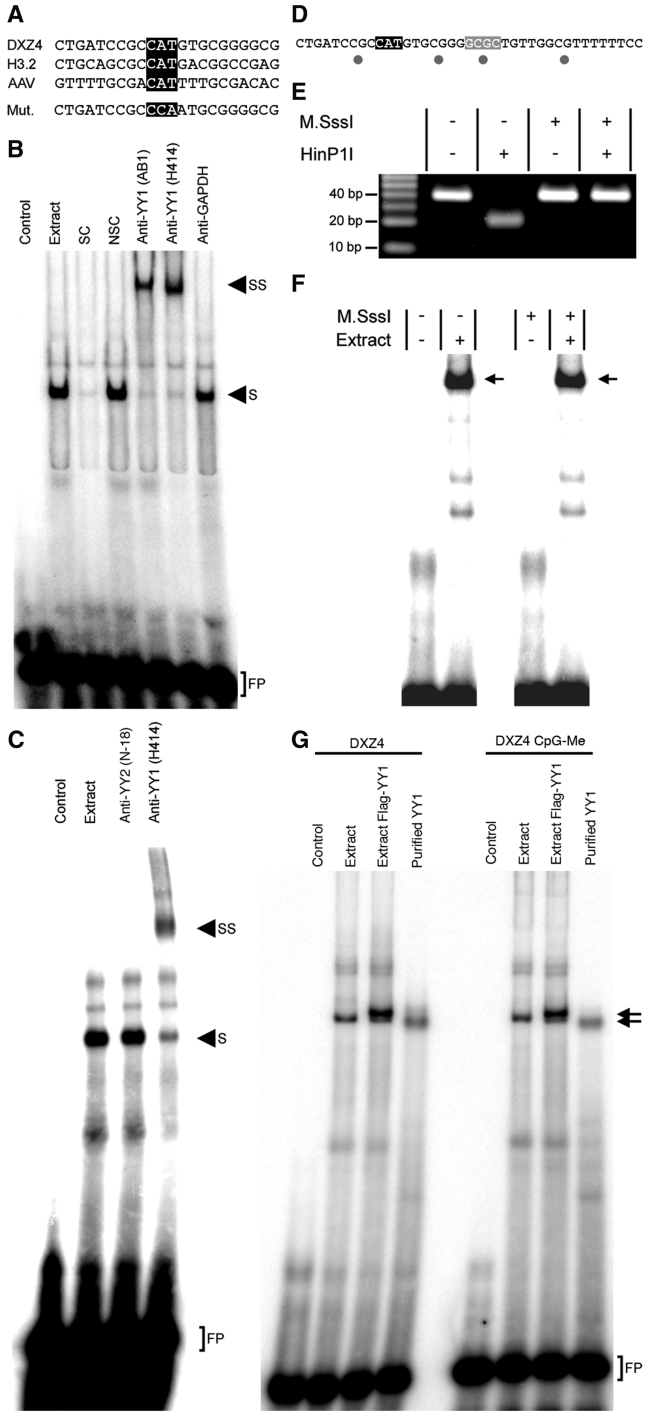

Figure 3. YY1 electrophoresis mobility-shift assays (EMSA). (**A**) Alignment of the candidate YY1-binding sequence of the macrosatellite DXZ4 (top) with two other well characterized YY1 binding consensus sequences. The core CAT element is highlighted. Below the alignments is the synthesized mutant (Mut.) sequence that no longer contains the CAT motif that was used in the EMSA. (**B**) Phosphoimager output of a dried EMSA gel. Radioactive bands appear dark. Samples loaded in each lane are labeled across the top and include specific competition (SC) and non-specific competition (NSC). Supershift analysis is shown in the last three lanes to the right with two independent YY1 antibodies and a non-specific control. The locations of the shift (S) and supershift (SS) are indicated by the arrowheads to the right of the gel. Free probe (FP) is indicated at the bottom of the gel. (**C**) Supershift analysis with anti-YY1 and anti-YY2. (**D**) Extended DNA sequence flanking the YY1 core CAT element (highlighted in black). CpG dinucleotides are highlighted by the dots below the sequence, and the location of the HinP1I site is highlighted in gray (GCGC). (**E**) Agarose gel (4.0%) analysis showing results of HinP1I digestions of the DNA sequence shown in (C), each of which was first treated or not with M.SssI, which methylates CpG dinucleotides *in vitro*. Presence or absence of either enzyme in the procedure is indicated above the corresponding lanes of the gel by the plus or minus symbol. DNA size is indicated to the right by a 10-bp ladder. (**F**) Phosphoimager output of a dried EMSA gel with the DNA sequence shown in (C). Radioactive bands appear dark. Target DNA sequence that has not (left two samples) or has been (right two samples) treated with M.SsspI before analysis is indicated above the gel by plus and minus symbols. Inclusion of whole-cell extract is indicated below. The location of the shift is indicated by the small black right-facing arrows. (**G**) Methylation insensitivity and YY1 specificity EMSA. EMSA shows shift with HeLa nuclear extract (Extract), a molecular weight shift for Flag-epitope-tagged YY1 from nuclear extract from HeLa cells overexpressing YY1-Flag (Extract Flag-YY1), and shift with purified non-tagged YY1 from bacteria (Purified YY1). Analyses are shown side by side, non-methylated double-stranded oligonucleotides to the left and CpG-methylated ones to the right. The arrows to the right of the gel indicate the Flag shift (top) and non-tagged shift (bottom).

including imprinted regions, binding of CTCF to DNA is impaired by CpG methylation (43–46). In contrast, CTCF binding to its DXZ4 target site *in vitro* is not affected by CpG methylation (23), perhaps because the CTCF site in DXZ4 does not directly contain any CpG residues. Binding of YY1 to DNA is reported to be impaired by CpG methylation at some sites (47) but not at others (48). The YY1-binding site in DXZ4 contains one CpG dinucleotide within the central 7-bp motif and several at other CpG sites nearby (Figure 3D). Previously, we have determined that these CpG sequences are largely methylated both in males and on the Xa, whereas the same sites are unmethylated on the Xi (23) (Supplementary Figure S1). We therefore sought to determine whether methylation of the YY1-binding site was sufficient to block binding and hence to restrict YY1 to the Xi.

A double-stranded 39-bp sequence encompassing the YY1-binding site and three upstream CpG dinucleotides was generated as the target for methylation sensitive EMSA (Figure 3D). One CpG dinucleotide resides within a recognition sequence (GCGC) for HinP1I, a restriction endonuclease unable to digest target sites when the central CpG is methylated. Using this feature as a measure of complete methylation, we methylated the CpG sites in the 39-bp oligomer *in vitro* using M.SssI. Methylated and unmethylated oligomers were then subjected to HinP1I digestion, which confirmed that the oligomer was fully methylated (Figure 3E). The EMSA experiment was then repeated with either the methylated or unmethylated target sequence. As can be seen in Figure 3F, methylation of the target site appears to have had no effect on YY1 binding, so like CTCF, which binds the methylated form of its DXZ4-binding site *in vitro* (23), YY1 requires more than just DNA methylation to prevent binding to DXZ4 on the Xa *in vivo*.

The clear blocking of HinP1I digestion of the methylated double-stranded oligonucleotide (Figure 3E) does not confirm efficient methylation of the other CpG dinucleotides in the DNA fragment. To complement this

analysis, we therefore synthesized a sense and anti-sense oligonucleotide, corresponding to the DXZ4 sequence shown in Figure 3A, in which the three Cs in the CpG context were methylated. Side-by-side analysis of the unmethylated and methylated double-stranded oligonucleotides revealed no impact on the ability of endogenous YY1, Flag-tagged YY1, or bacterially purified YY1 to bind to the sequence (Figure 3G). These data confirm not only that methylation of CpG does not block YY1 binding *in vitro* but also that YY1 is responsible for the EMSA as demonstrated by the shift corresponding to the Flag-tagged YY1 and a shift with purified YY1.

YY1 and CTCF bind to the hypomethylated form of DXZ4 on the active X chromosome in some male carcinomas

Changes in DNA methylation are a common feature in cancer (49), affecting both unique and repetitive components of the genome. Indeed, in various cancers, increases and decreases in DNA methylation have been reported for large tandem-repeat DNA such as D4Z4 and NBL-2 (50–52). To date, changes in CpG methylation at DXZ4 in cancer have not been investigated. Although CpG methylation does not block binding of YY1 (Figure 3E) or CTCF (23) to their DXZ4 target sites *in vitro*, both proteins are restricted *in vivo* to the hypomethylated macrosatellite on the Xi. Previously, we and others have shown that most DXZ4 CpG dinucleotides are methylated in males (7,23), so we investigated a number of male carcinoma cell lines for a reduction in DXZ4 methylation by bisulfite sequencing with the premise that any such reduction in methylation might result in altered accessibility of YY1 and CTCF to DXZ4.

Across the interval encompassing the bidirectional promoter and binding sites for YY1 and CTCF, three independent colon carcinoma cell lines (HCT15, Caco-2 and SW1116) had CpG methylation profiles indistinguishable from those of normal males (23) (Figure 4A) (Supplementary Figure S1). Notably, two CpG residues frequently appear resistant to methylation (CpG-11 and less frequently CpG-30 from the left) in normal human males (23) (Supplementary Figure S1), as well as in pig-tailed macaque DXZ4 (20). The CpG-11 resides approximately midway between the YY1 and CTCF-binding sites and CpG-30 within the defined promoter. The relevance of these largely methylation-resistant CpGs is not obvious but is intriguing.

A fourth colon-carcinoma cell line (DLD-1) also shows a typical male CpG methylation profile at DXZ4 (Figure 4B) and more importantly has a profile similar to that of HCT-15 (Figure 4A); the two cell lines are believed to be derived independently from the same cancer specimen (53,54), so any differences in methylation profiles are unlikely to result from cell-line derivation. In contrast, a fifth colon-carcinoma cell line (HCT116) and a fibrosarcoma cell line (HT-1080) show lower methylations—only 16.2 and 52.1% CpG methylation, respectively (Figure 4B).

To investigate changes to chromatin organization at DXZ4, we performed ChIP on HCT116 and HT-1080 along with DLD-1 (Figure 4C). As in male primary cells (23), the heterochromatin marker H3K9me3 was detected in all three samples. The euchromatin marker H3K4me2 could readily be detected at DXZ4 in HCT116 and HT-1080, confirming a euchromatic organization of the macrosatellite consistent with the CpG hypomethylation profile (Figure 4B). Low levels of H3K4me2 could also be detected in DLD-1, possibly because DXZ4 is expressed at varying levels in males (8,23) and this signal might therefore represent transcriptionally active DXZ4 monomers.

Transformation-associated chromatin changes at DXZ4 allow CTCF and YY1 to bind DXZ4 on the active X chromosome

Next we sought to determine what changes, if any, occur to the methylation profile and chromatin organization of DXZ4 in association with metastasis and transformation. First we determined and compared the DXZ4 CpG methylation profile of SW480, established from a primary colon adenocarcinoma, with that of SW620, derived from a metastatic lesion at a later stage of colon carcinoma in the same individual (55). In this instance the methylation profiles were very similar; overall CpG methylation was slightly higher in SW620 (Figure 4D), suggesting no dramatic alteration in DXZ4 chromatin organization. Then we examined the CpG methylation profile of DXZ4 in skin primary fibroblasts (Malme-3) and a cell line derived from a malignant melanoma from the same individual (Malme-3M). DXZ4 CpG methylation for Malme-3 was comparable to that seen in other normal male fibroblasts (Supplementary Figure S1) (79.4%), whereas a reduction in methylation was noticeable in the melanoma cell line (64.1%) (Figure 4E). ChIP, intended to explore possible chromatin changes associated with the reduced DNA methylation at DXZ4, once again revealed low levels of the euchromatic marker H3K4me2 in the primary male fibroblast cells (Figure 4F). This result was not too surprising given that >20% of CpG residues in the interval are unmethylated (Figure 4E) and that low levels of H3K4me2 could be detected in DLD-1, a male sample with higher overall CpG methylation (Figure 4C), but when CTCF was used to mark a shift from the male X/Xa arrangement toward a female Xi-like organization, CTCF could only be detected in the melanoma cell line (Figure 4F), consistent with our observations for carcinoma cell lines HT-1080 and HCT116 (Figure 4C). Furthermore, we detected binding of YY1 to DXZ4 in Malme-3M but not in Malme-3 (Figure 4F). The low YY1 signal in Malme-3M may arise because the YY1-binding region retains overall higher methylation than is seen in HCT116 and HT-1080 (compare Figure 4B with E), suggesting that CpG methylation affects YY1 binding to DXZ4 *in vivo*. Importantly, this result shows that transformation was associated with a reduction in CpG methylation and a gain of CTCF and YY1 binding at DXZ4 in this individual.

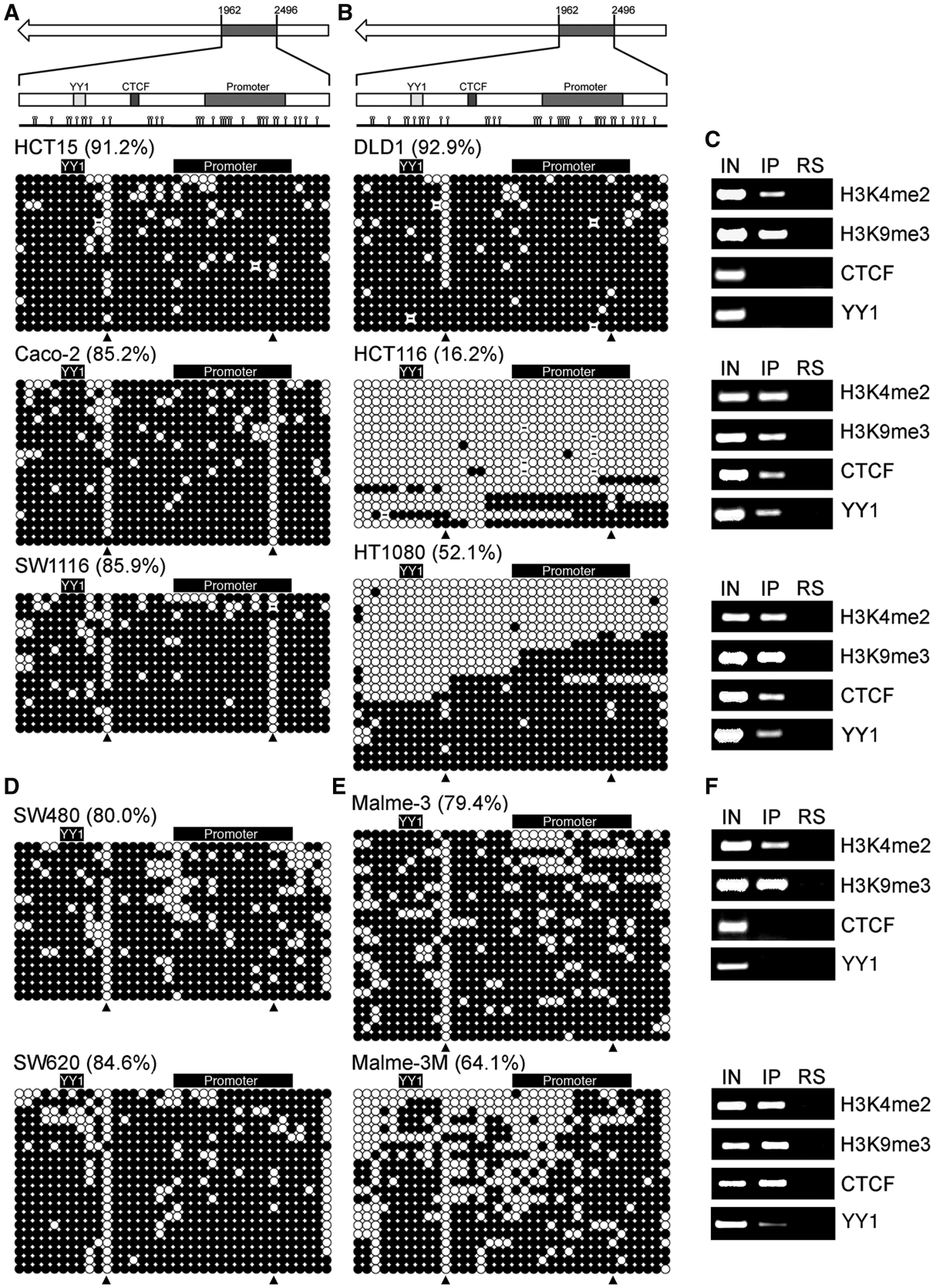

Figure 4. DXZ4 CpG methylation and chromatin structure in male carcinomas and transformation-associated changes. (**A**) Schematic representation (top) of a single 3-kb DXZ4 monomer shown right to left on the basis of the orientation on Xq. The region assessed for CpG methylation is highlighted, and the coordinates for the interval amplified with the forward and reverse primers given above the monomer. Region of interest (bottom) expanded to show the location of the promoter along with the CTCF and YY1-binding sites. The line below it shows the location of all

(continued)

Predicting DXZ4 CpG methylation profile from euchromatic chromatin signatures and CTCF association with the macrosatellite

Our data support a model in which DXZ4 on the Xa and male X is packaged into constitutive heterochromatin characterized by H3K9me3 and CpG methylation. In contrast, DXZ4 on the Xi is largely packaged into euchromatin characterized by H3K4me2 and CpG hypomethylation and is bound by YY1 and CTCF. Transformation-associated reduction in DNA methylation at the macrosatellite, however, alters the chromatin organization sufficiently to permit binding of YY1 and CTCF to male DXZ4, shifting it into a more Xi-like state. On the basis of this model, we would predict that observation of YY1 or CTCF at DXZ4 in male cells would indicate an overall hypomethylated CpG state. To test this hypothesis, we examined the chromatin profile for CTCF at DXZ4 using the publicly available Encode data from ChIP experiments combined with high-throughput sequencing (56).

Consistent with our observations (23), DXZ4 in normal female but not male cells is characterized by the presence of euchromatic markers H3K4me2 and histone H3 lysine-9 acetylation (H3K9Ac) and binding of CTCF (Figure 5A), but in the male hepatocellular carcinoma cell line HepG2, DXZ4 chromatin organization is indistinguishable from the female profiles. Analysis of the CpG methylation status of DXZ4 in HepG2 cells reveals a hypomethylated state comparable to that of HCT116 and HT-1080 (Figure 5B), supporting transformation-associated changes to DXZ4 chromatin organization on the male X chromosome.

DISCUSSION

DXZ4 is an enigmatic DNA element that adopts a chromatin organization different from that of the flanking chromosome: largely heterochromatin on the Xa and euchromatin bound by CTCF on the Xi (23). Here, we report that the zinc-finger protein YY1 associates specifically with the euchromatic form of DXZ4 alongside CTCF on the Xi. Colocalization of YY1 and CTCF has been reported at binding sites within tandem repeat DNA at autosomal imprinted loci (57) and the Xic (32). The theme that unifies these three situations is that all are classic examples of epigenetics—maintenance of two alleles in alternate chromatin states—and support a model in which chromatin marked by YY1 and CTCF in close proximity has an important role in establishing and/or maintaining an alternate chromatin organization on one chromosome.

In vitro analysis showed that both YY1 and CTCF (23) can bind methylated DXZ4 target sequences, yet *in vivo* neither is detected at such sequences. Therefore, methylation alone cannot account for the allelic exclusion of either protein. One explanation could be that *in vivo* methylated DXZ4 DNA is inaccessible. Given that the DXZ4 macrosatellite on the Xa and in males is packaged in constitutive heterochromatin characterized by H3K9me3 (23) and HP1g (18), this supposition is not unreasonable. Indeed, binding of CTCF and YY1 does closely follow the methylation profile of DXZ4, as demonstrated by the binding of both proteins to hypomethylated DXZ4 in two independent male carcinoma cell lines (HCT116 and HT-1080) (Figure 4C), and a malignant melanoma cell line, but not in primary skin cells from the same male (Figure 4F), and CTCF binding to the hypomethylated macrosatellite in HepG2 cells (Figure 5A). Interestingly, these independent male cancer-cell lines showed a shift of DXZ4 chromatin organization from a typical Xa state toward one that more closely resembled that on the Xi in females. Conceivably, this arrangement is the default for DXZ4, and in these samples it is reverting back toward this base state. If so, we would predict that DXZ4 in pluripotent cells would exist in a euchromatic arrangement, a hypothesis we are actively pursuing.

Our *in vitro* data indicate that YY1 bound to DXZ4, but given that YY2 recognizes similar DNA sequence motifs (41) and possesses a near identical DNA binding domain (40), YY2 may also be associating *in vivo* with DXZ4. Because YY1 and YY2 are not functionally redundant and can mediate antagonistic effects at target sites (58,59), possible binding of YY2 at DXZ4 warrants further investigation.

YY1 has also been shown to associate with the chromosome 4q macrosatellite D4Z4 (60), extending the parallels between this disease-associated macrosatellite and DXZ4

Figure 4. Continued
CpG dinucleotides, represented by the circles on sticks. Below the schematic map are the methylation profiles for three different male colon carcinoma cell lines. Each horizontal row of 36 circles represents the sequencing result of an independent TA clone from cloning of PCR products generated from bisulfite-modified DNA template for the corresponding cell lines. Methylated CpGs are represented by filled circles, unmethylated ones by open circles. Sequence variations at DXZ4 that result in a non-C at a consensus CpG site are represented by dashes. CpG dinucleotides contained in the YY1-binding site and promoter are indicated above each profile. The upward-facing arrowheads indicate the commonly hypomethylated CpG residues CpG-11 and CpG-30. The percentage methylation is given in brackets beside each name. (**B**) Representation as in (A) showing CpG methylation profiles for two other male colon-carcinoma cell lines and one male fibrosarcoma cell line. (**C**) ChIP analysis of DXZ4 chromatin for the male cell line shown immediately to the left in (B). Data are shown for H3K4me2, H3K9me3, CTCF, and YY1 as indicated to the right of each agarose gel image and are representative examples from three independent ChIP replicates for each cell line. Samples assessed include the input (IN), the eluted immunoprecipitation (IP), and an immunoprecipitation with non-specific rabbit serum (RS). YY1 and CTCF were assessed with DXZ4-F23 and DXZ4-R26, whereas H3K4me2 and H3K4me3 were assessed with DXZ4-F19 and DXZ4-R19. ChIP was performed with anti-YY1 (sc-1703). (**D**) Representation as in (A) showing CpG methylation profiles for a cell line derived from a primary colon adenocarcinoma (SW480) and a metastatic lesion at a later stage of colon carcinoma in the same individual (SW620). (**E**) Representation as in (A) showing CpG methylation profiles from a male normal skin fibroblast culture (top) and a malignant melanoma cell line from the same individual (bottom panel). (**F**) Analysis as in (C) showing H3K4me2, H3K9me3, CTCF and YY1 ChIP data for the samples shown immediately on the left in (E). Data are representative examples from two independent replicate ChIP experiments.

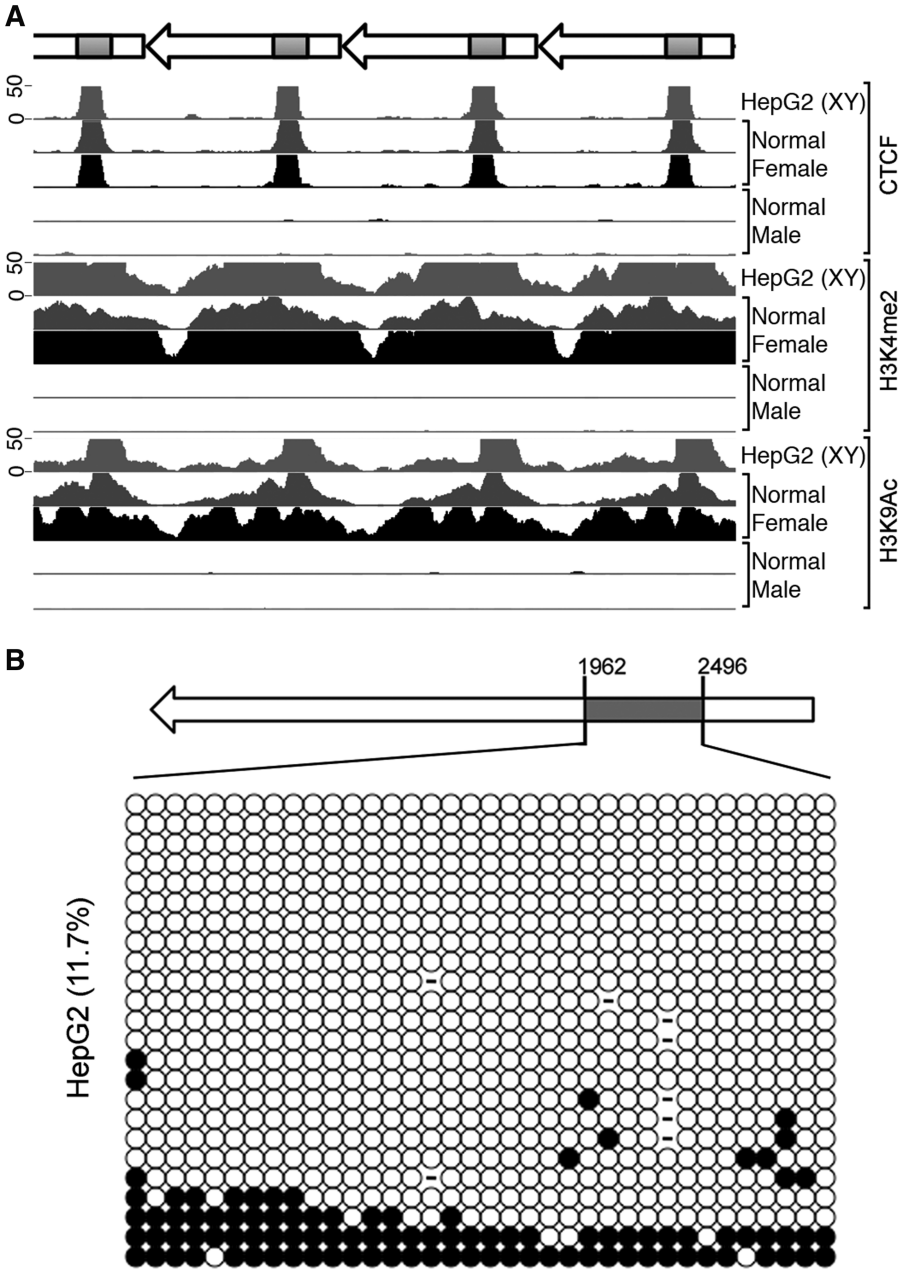

Figure 5. Correlation between CpG methylation and chromatin organization of DXZ4. (**A**) The ChIP-Seq profile for part of the DXZ4 macrosatellite [Encode histone modifications by Broad Institute ChIP Seq (64)]. The locations of individual 3-kb monomers are marked by left-facing arrows. The boxed shaded area in each arrow indicates the region containing the promoter and binding sites for CTCF and YY1. Below are the ChIP-Seq profiles for CTCF, H3K4me2, and H3K9Ac as indicated to the right. Profiles include the male hepatocellular carcinoma cell line HepG2; two independent normal female lines [top, lymphoblastoid cell line (GM12878); bottom, human mammary epithelial cells (HMEC)]; and two independent normal males [top, human skeletal-muscle myoblasts (HSMM); bottom, human umbilical-vein endothelial cells (HUVEC)]. The image was taken from the UCSC genome browser (http://genome.ucsc.edu) and shows data taken with default settings from genome build hg18, showing a 10-kb window; coordinates X; 114 881 001–114 891 000. The data shown in all tracks are the signal view and used the default scale of 0–50 as indicated to the left. (**B**) Schematic representation (top) of a single 3-kb DXZ4 monomer shown as a left-facing arrow. The region assessed for CpG methylation is indicated by the shaded region and the coordinates given. Methylation profile (bottom) of this interval for HepG2. Methylated CpGs are represented by filled circles, unmethylated ones by open circles. Sequence variations at DXZ4 that result in a non-C at a consensus CpG site are represented by dashes. The percentage methylation is given to the left of the profile.

(61), but in this instance, YY1 is associated, as part of a repressor complex (60), with the normal heterochromatic form of D4Z4, and CTCF only associates upon contraction in FSHD (17) when D4Z4 reverts toward a more euchromatic conformation (18). Whether YY1 remains associated with D4Z4 in its euchromatic contracted form remains to be determined.

Collectively, these data provide novel insight into the organization and stability of chromatin at the macrosatellite DXZ4. Further parallels can be drawn

between DXZ4 and D4Z4, as well as monoallelic chromatin states at sites of genomic imprinting and the Xic. The ability of DXZ4 to adopt two distinct chromatin states in the context of XCI (7,20,23), in the absence of the macrosatellite contraction gain of function seen for D4Z4 in FSHD (17,18), strongly suggests that DXZ4 fulfils, through epigenetic regulation, two alternate conserved (20) functions on the X chromosome—one packaged as heterochromatin on the Xa and the other in a euchromatic conformation of the Xi. Although only the XIC located at Xq13 is required for XCI in humans (62), DXZ4 may still have some role after XCI in spreading the XCI signal or facilitating organization of the Xi. Recent data from Joen and Lee (63) indicate that Yy1 possesses both DNA and RNA-binding activities and physically tethers Xist RNA to the Xi. This exciting result is a major step forward in deciphering how Xist associates exclusively with the Xi. Xist spreads in *cis* along the length of the Xi (28), so numerous other Xist entry sites defined by Xi-specific Yy1 might facilitate this process. The Xi-specific association of YY1 at DXZ4 could function to assist in the spread of XIST along the length of the chromosome.

Macrosatellites remain a perplexing and largely unexplored aspect of our genome (5). We anticipate that future studies will reveal that these DNA elements occupy an important niche in genome regulation and disease susceptibility.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Figure S1, Supplementary Reference [23].

ACKNOWLEDGEMENTS

The authors are grateful to members of the Chadwick lab for constructive discussions and evaluation of the manuscript. The authors are indebted to Anne Thistle for carefully reading and correcting the manuscript. The authors also thank the anonymous reviewers for their helpful and constructive suggestions.

FUNDING

National Institutes of Health (GM073120 to B.P.C.); Florida State University College of Medicine (MMH). Funding for open access charge: NIH R01.

Conflict of interest statement. None declared.

REFERENCES

1. Lander,E.S., Linton,L.M., Birren,B., Nusbaum,C., Zody,M.C., Baldwin,J., Devon,K., Dewar,K., Doyle,M., FitzHugh,W. *et al.* (2001) Initial sequencing and analysis of the human genome. *Nature*, **409**, 860–921.
2. International Human Genome Sequencing Consortium. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931–945.
3. Kapranov,P., Cheng,J., Dike,S., Nix,D.A., Duttagupta,R., Willingham,A.T., Stadler,P.F., Hertel,J., Hackermuller,J., Hofacker,I.L. *et al.* (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science*, **316**, 1484–1488.
4. Kidd,J.M., Cooper,G.M., Donahue,W.F., Hayden,H.S., Sampas,N., Graves,T., Hansen,N., Teague,B., Alkan,C., Antonacci,F. *et al.* (2008) Mapping and sequencing of structural variation from eight human genomes. *Nature*, **453**, 56–64.
5. Warburton,P.E., Hasson,D., Guillem,F., Lescale,C., Jin,X. and Abrusan,G. (2008) Analysis of the largest tandemly repeated DNA families in the human genome. *BMC Genomics*, **9**, 533.
6. Tremblay,D.C., Alexander,G. Jr, Moseley,S. and Chadwick,B.P. (2010) Expression, tandem repeat copy number variation and stability of four macrosatellite arrays in the human genome. *BMC Genomics*, **11**, 632.
7. Giacalone,J., Friedes,J. and Francke,U. (1992) A novel GC-rich human macrosatellite VNTR in Xq24 is differentially methylated on active and inactive X chromosomes. *Nat. Genet.*, **1**, 137–143.
8. Tremblay,D.C., Moseley,S. and Chadwick,B.P. (2011) Variation in array size, monomer composition and expression of the macrosatellite DXZ4. *PLoS One*, **6**, e18969.
9. Deidda,G., Cacurri,S., Grisanti,P., Vigneti,E., Piazzo,N. and Felicetti,L. (1995) Physical mapping evidence for a duplicated region on chromosome 10qter showing high homology with the facioscapulohumeral muscular dystrophy locus on chromosome 4qter. *Eur. J. Hum. Genet.*, **3**, 155–167.
10. Winokur,S.T., Bengtsson,U., Vargas,J.C., Wasmuth,J.J., Altherr,M.R., Weiffenbach,B. and Jacobsen,S.J. (1996) The evolutionary distribution and structural organization of the homeobox-containing repeat D4Z4 indicates a functional role for the ancestral copy in the FSHD region. *Hum. Mol. Genet.*, **5**, 1567–1575.
11. Gondo,Y., Okada,T., Matsuyama,N., Saitoh,Y., Yanagisawa,Y. and Ikeda,J.E. (1998) Human megasatellite DNA RS447: copy-number polymorphisms and interspecies conservation. *Genomics*, **54**, 39–49.
12. Wijmenga,C., Hewitt,J.E., Sandkuijl,L.A., Clark,L.N., Wright,T.J., Dauwerse,H.G., Gruter,A.M., Hofker,M.H., Moerer,P., Williamson,R. *et al.* (1992) Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. *Nat. Genet.*, **2**, 26–30.
13. Tawil,R. (2008) Facioscapulohumeral muscular dystrophy. *Neurotherapeutics*, **5**, 601–606.
14. Lemmers,R.J., Wohlgemuth,M., van der Gaag,K.J., van der Vliet,P.J., van Teijlingen,C.M., de Knijff,P., Padberg,G.W., Frants,R.R. and van der Maarel,S.M. (2007) Specific sequence variations within the 4q35 region are associated with facioscapulohumeral muscular dystrophy. *Am. J. Hum. Genet.*, **81**, 884–894.
15. van Deutekom,J.C., Wijmenga,C., van Tienhoven,E.A., Gruter,A.M., Hewitt,J.E., Padberg,G.W., van Ommen,G.J., Hofker,M.H. and Frants,R.R. (1993) FSHD associated DNA rearrangements are due to deletions of integral copies of a 3.2 kb tandemly repeated unit. *Hum. Mol. Genet.*, **2**, 2037–2042.
16. de Greef,J.C., Lemmers,R.J., van Engelen,B.G., Sacconi,S., Venance,S.L., Frants,R.R., Tawil,R. and van der Maarel,S.M. (2009) Common epigenetic changes of D4Z4 in contraction-dependent and contraction-independent FSHD. *Hum. Mutat.*, **30**, 1449–1459.
17. Ottaviani,A., Rival-Gervier,S., Boussouar,A., Foerster,A.M., Rondier,D., Sacconi,S., Desnuelle,C., Gilson,E. and Magdinier,F. (2009) The D4Z4 macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet.*, **5**, e1000394.
18. Zeng,W., de Greef,J.C., Chen,Y.Y., Chien,R., Kong,X., Gregson,H.C., Winokur,S.T., Pyle,A., Robertson,K.D., Schmiesing,J.A. *et al.* (2009) Specific loss of histone H3 lysine 9 trimethylation and HP1gamma/cohesin binding at D4Z4 repeats is associated with facioscapulohumeral dystrophy (FSHD). *PLoS Genet.*, **5**, e1000559.
19. Lemmers,R.J., van der Vliet,P.J., Klooster,R., Sacconi,S., Camano,P., Dauwerse,J.G., Snider,L., Straasheijm,K.R., Jan van Ommen,G., Padberg,G.W. *et al.* (2010) A unifying genetic model for facioscapulohumeral muscular dystrophy. *Science*, **329**, 1650–1653.

20. McLaughlin,C.R. and Chadwick,B.P. (2011) Characterization of DXZ4 conservation in primates implies important functional roles for CTCF binding, array expression and tandem repeat organization on the X chromosome. *Genome Biol.*, **12**, R37.
21. Saitoh,Y., Miyamoto,N., Okada,T., Gondo,Y., Showguchi-Miyata,J., Hadano,S. and Ikeda,J.E. (2000) The RS447 human megasatellite tandem repetitive sequence encodes a novel deubiquitinating enzyme with a functional promoter. *Genomics*, **67**, 291–300.
22. Snider,L., Asawachaicharn,A., Tyler,A.E., Geng,L.N., Petek,L.M., Maves,L., Miller,D.G., Lemmers,R.J., Winokur,S.T., Tawil,R. *et al.* (2009) RNA transcripts, miRNA-sized fragments, and proteins produced from D4Z4 units: new candidates for the pathophysiology of facioscapulohumeral dystrophy. *Hum. Mol. Genet.*, **18**, 2414–2430.
23. Chadwick,B.P. (2008) DXZ4 chromatin adopts an opposing conformation to that of the surrounding chromosome and acquires a novel inactive X-specific role involving CTCF and antisense transcripts. *Genome Res.*, **18**, 1259–1269.
24. Hewitt,J.E., Lyle,R., Clark,L.N., Valleley,E.M., Wright,T.J., Wijmenga,C., van Deutekom,J.C., Francis,F., Sharpe,P.T., Hofker,M. *et al.* (1994) Analysis of the tandem repeat locus D4Z4 associated with facioscapulohumeral muscular dystrophy. *Hum. Mol. Genet.*, **3**, 1287–1295.
25. Lyon,M.F. (1961) Gene action in the X-chromosome of the mouse (*Mus musculus* L.). *Nature*, **190**, 372–373.
26. Augui,S., Nora,E.P. and Heard,E. (2011) Regulation of X-chromosome inactivation by the X-inactivation centre. *Nat. Rev. Genet.*, **12**, 429–442.
27. Carrel,L. and Willard,H.F. (2005) X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature*, **434**, 400–404.
28. Payer,B. and Lee,J.T. (2008) X chromosome dosage compensation: how mammals keep the balance. *Annu. Rev. Genet.*, **42**, 733–772.
29. Boumil,R.M., Ogawa,Y., Sun,B.K., Huynh,K.D. and Lee,J.T. (2006) Differential methylation of Xite and CTCF sites in Tsix mirrors the pattern of X-inactivation choice in mice. *Mol. Cell. Biol.*, **26**, 2109–2117.
30. Courtier,B., Heard,E. and Avner,P. (1995) Xce haplotypes show modified methylation in a region of the active X chromosome lying 3' to Xist. *Proc. Natl Acad. Sci. USA*, **92**, 3531–3535.
31. 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.
32. Donohoe,M.E., Zhang,L.F., Xu,N., Shi,Y. and Lee,J.T. (2007) Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Mol. Cell*, **25**, 43–56.
33. Rizkallah,R. and Hurt,M.M. (2009) Regulation of the transcription factor YY1 in mitosis through phosphorylation of its DNA-binding domain. *Mol. Biol. Cell*, **20**, 4766–4776.
34. Rizkallah,R., Alexander,K.E., Kassardjian,A., Luscher,B. and Hurt,M.M. (2011) The transcription factor YY1 is a substrate for Polo-Like Kinase 1 at the G2/M transition of the cell cycle. *PLoS One*, **6**, e15928.
35. Barr,M.L. and Bertram,E.G. (1949) A morphological distinction between neurones of the male and female, and the behaviour of the nucleolar satellite during accelerated nucleoprotein synthesis. *Nature*, **163**, 676–677.
36. Chadwick,B.P. and Willard,H.F. (2003) Chromatin of the Barr body: histone and non-histone proteins associated with or excluded from the inactive X chromosome. *Hum. Mol. Genet.*, **12**, 2167–2178.
37. Bowman,T.L. and Hurt,M.M. (1995) The coding sequences of mouse H2A and H3 histone genes contains a conserved seven nucleotide element that interacts with nuclear factors and is necessary for normal expression. *Nucleic Acids Res.*, **23**, 3083–3092.
38. Eliassen,K.A., Baldwin,A., Sikorski,E.M. and Hurt,M.M. (1998) Role for a YY1-binding element in replication-dependent mouse histone gene expression. *Mol. Cell. Biol.*, **18**, 7106–7118.
39. Hyde-DeRuyscher,R.P., Jennings,E. and Shenk,T. (1995) DNA binding sites for the transcriptional activator/repressor YY1. *Nucleic Acids Res.*, **23**, 4457–4465.
40. Nguyen,N., Zhang,X., Olashaw,N. and Seto,E. (2004) Molecular cloning and functional characterization of the transcription factor YY2. *J. Biol. Chem.*, **279**, 25927–25934.
41. Kim,J.D., Faulk,C. and Kim,J. (2007) Retroposition and evolution of the DNA-binding motifs of YY1, YY2 and REX1. *Nucleic Acids Res.*, **35**, 3442–3452.
42. Filippova,G.N. (2008) Genetics and epigenetics of the multifunctional protein CTCF. *Curr. Top. Dev. Biol.*, **80**, 337–360.
43. Bell,A.C. and Felsenfeld,G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature*, **405**, 482–485.
44. 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. *Nat. Genet.*, **28**, 335–343.
45. Hark,A.T., Schoenherr,C.J., Katz,D.J., Ingram,R.S., Levorse,J.M. and Tilghman,S.M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**, 486–489.
46. Kanduri,C., Pant,V., Loukinov,D., Pugacheva,E., Qi,C.F., Wolffe,A., Ohlsson,R. and Lobanenkov,V.V. (2000) Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.*, **10**, 853–856.
47. 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.
48. Gaston,K. and Fried,M. (1995) CpG methylation has differential effects on the binding of YY1 and ETS proteins to the bi-directional promoter of the Surf-1 and Surf-2 genes. *Nucleic Acids Res.*, **23**, 901–909.
49. Kulis,M. and Esteller,M. (2010) DNA methylation and cancer. *Adv. Genet.*, **70**, 27–56.
50. Choi,S.H., Worswick,S., Byun,H.M., Shear,T., Soussa,J.C., Wolff,E.M., Douer,D., Garcia-Manero,G., Liang,G. and Yang,A.S. (2009) Changes in DNA methylation of tandem DNA repeats are different from interspersed repeats in cancer. *Int. J. Cancer*, **125**, 723–729.
51. Nishiyama,R., Qi,L., Lacey,M. and Ehrlich,M. (2005) Both hypomethylation and hypermethylation in a 0.2-kb region of a DNA repeat in cancer. *Mol. Cancer Res.*, **3**, 617–626.
52. Tsumagari,K., Qi,L., Jackson,K., Shao,C., Lacey,M., Sowden,J., Tawil,R., Vedanarayanan,V. and Ehrlich,M. (2008) Epigenetics of a tandem DNA repeat: chromatin DNaseI sensitivity and opposite methylation changes in cancers. *Nucleic Acids Res.*, **36**, 2196–2207.
53. Chen,T.R., Dorotinsky,C.S., McGuire,L.J., Macy,M.L. and Hay,R.J. (1995) DLD-1 and HCT-15 cell lines derived separately from colorectal carcinomas have totally different chromosome changes but the same genetic origin. *Cancer Genet. Cytogenet.*, **81**, 103–108.
54. Vermeulen,S.J., Chen,T.R., Speleman,F., Nollet,F., Van Roy,F.M. and Mareel,M.M. (1998) Did the four human cancer cell lines DLD-1, HCT-15, HCT-8, and HRT-18 originate from one and the same patient? *Cancer Genet Cytogenet.*, **107**, 76–79.
55. Melcher,R., Steinlein,C., Feichtinger,W., Muller,C.R., Menzel,T., Luhrs,H., Scheppach,W. and Schmid,M. (2000) Spectral karyotyping of the human colon cancer cell lines SW480 and SW620. *Cytogenet. Cell Genet.*, **88**, 145–152.
56. The ENCODE Project Consortium,T.E.P. (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol.*, **9**, e1001046.
57. Kim,J. (2008) Multiple YY1 and CTCF binding sites in imprinting control regions. *Epigenetics*, **3**, 115–118.
58. Chen,Li., Shioda,T., Coser,K.R., Lynch,M.C., Yang,C. and Schmidt,E.V. (2010) Genome-wide analysis of YY2 versus YY1 target genes. *Nucleic Acids Res.*, **38**, 4011–4026.
59. Klar,M. and Bode,J. (2005) Enhanceosome formation over the beta interferon promoter underlies a remote-control mechanism mediated by YY1 and YY2. *Mol. Cell Biol.*, **25**, 10159–10170.
60. Gabellini,D., Green,M.R. and Tupler,R. (2002) Inappropriate gene activation in FSHD: a repressor complex binds a

chromosomal repeat deleted in dystrophic muscle. *Cell*, **110**, 339–348.
61. Chadwick,B.P. (2009) Macrosatellite epigenetics: the two faces of DXZ4 and D4Z4. *Chromosoma*, **118**, 675–681.
62. Brown,C.J., Lafrenière,R.G., Powers,V.E., Sebastio,G., Ballabio,A., Pettigrew,A.L., Ledbetter,D.H., Levy,E., Craig,I.W. and Willard,H.F. (1991) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature*, **349**, 82–84.
63. Jeo,Y. and Lee,J.T. (2011) YY1 tethers Xist RNA to the inactive X nucleation center. *Cell*, **146**, 119–133.
64. Mikkelsen,T.S., Ku,M., Jaffe,D.B., Issac,B., Lieberman,E., Giannoukos,G., Alvarez,P., Brockman,W., Kim,T.K., Koche,R.P. *et al.* (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature*, **448**, 553–560.