A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells.

Angéline Eymery, Béatrice Horard, Michèle El Atifi-Borel, Geneviève Fourel, François R. Berger, Anne-Laure Vitte, Arnaud van den Broeck, Elisabeth Brambilla, Alexandra Fournier, Mary B. Callanan, et al.

A transcriptomic analysis of human centromeric and pericentric sequences in normal and tumor cells

Angéline Eymery[1], Béatrice Horard[2], Michèle El Atifi-Borel[3], Geneviève Fourel[2], François Berger[3], Anne-Laure Vitte[1], Arnaud Van den Broeck[1,4], Elisabeth Brambilla[1,4], Alexandra Fournier[1], Mary Callanan[1], Sylvie Gazzeri[1,4], Saadi Khochbin[1], Sophie Rousseaux[1], Eric Gilson[2] and Claire Vourc'h[1,*]

[1]INSERM, U823, Université Joseph Fourier, Institut Albert Bonniot, La Tronche F-38706, Grenoble, [2]Laboratoire de Biologie Moléculaire de la Cellule, UMR 5239, CNRS, ENS-LYON, Université LYON 1/HCL, 165, Chemin du Grand Revoyet, 69495 Pierre Bénite, [3]INSERM, U836, Université Joseph Fourier, Institut de Neurosciences, La Tronche, F-38700, Grenoble and [4]CHU de Grenoble, La Tronche F-38706

ABSTRACT

Although there is now evidence that the expression of centromeric (CT) and pericentric (PCT) sequences are key players in major genomic functions, their transcriptional status in human cells is still poorly known. The main reason for this lack of data is the complexity and high level of polymorphism of these repeated sequences, which hampers straightforward analyses by available transcriptomic approaches. Here a transcriptomic macro-array dedicated to the analysis of CT and PCT expression is developed and validated in heat-shocked (HS) HeLa cells. For the first time, the expression status of CT and PCT sequences is analyzed in a series of normal and cancer human cells and tissues demonstrating that they are repressed in all normal tissues except in the testis, where PCT transcripts are found. Moreover, PCT sequences are specifically expressed in HS cells in a Heat-Shock Factor 1 (HSF1)-dependent fashion, and we show here that another independent pathway, involving DNA hypo-methylation, can also trigger their expression. Interestingly, CT and PCT were found illegitimately expressed in somatic cancer samples, whereas PCT were repressed in testis cancer, suggesting that the expression of CT and PCT sequences may represent a good indicator of epigenetic deregulations occurring in response to environmental changes or in cell transformation.

INTRODUCTION

Since its first description by Emile Heitz in 1928, heterochromatin has mainly been portrayed as a transcriptionally inactive condensed nuclear compartment, inaccessible to transcription factors. The discovery that specific PCT transcripts play an essential role in the formation and maintenance of heterochromatin in fission yeast has considerably changed this view (1–3).

Repetitive DNA is a common feature of centromeric (CT) and pericentromeric (PCT) regions, often grouped under the concept of 'CT heterochromatin'. Although CT and PCT sequences are spatially related, their structure and function are clearly distinct. While CT regions are enriched in the histone variant CenH3 and bind CT specific proteins, PCT regions are enriched in the epigenetic repressive marks H3K9me3 and HP1. During mitosis, whereas CT regions play a direct role in spindle attachment, PCT regions ensure sister chromatids cohesion (4). CT and PCT regions also differ in the nature of their repetitive sequences. CT sequences, also named alpha satellites or alphoids, consist of repetitions of an AT rich motif of 171 pb, whereas PCT sequences mainly contain satellite II and III sequences, both enriched in GGAAT motif (5–7).

The existence of conserved features within the general organization of CT and PCT regions in yeast, mouse and human argues in favor of a conserved role for CT and PCT transcripts across species. Indeed, evidence is accumulating that these sequences are transcriptionally competent in mammalian cells in diverse biological contexts (7).

In human cells, the most dramatic example of transcriptional activation of PCT repeats is that occurring in

*To whom correspondence should be addressed. Tel: + 33 4 76 54 94 70; Fax: + 33 4 76 54 95 95; Email: claire.vourch@ujf-grenoble.fr

response to cell stress (8,9). Indeed, we and others have shown that, during heat-shock, the hyperphosphorylated form of Heat-Shock Factor 1 (HSF1) binds to the 9q12 locus, enriched in satellite III sequences, forming nuclear structure also known as nuclear stress bodies or nSBs (10). Within nSBs, these transcripts remain associated with the 9q12 region from which they originate (8). The functional implication of this observation is still unknown. Moreover, the possibility that other sites of non-coding repeated sequences could be actively transcribed during this process and/or under other physiological or pathological circumstances remains to be investigated. These questions are particularly difficult to address in human cells due to the repetitive nature and high degree of polymorphism of these sequences, which represent major drawbacks to the qualitative and quantitative analysis of their expression.

In order to circumvent this problem, we have designed a transcriptomic macro-array allowing a quantitative analysis of CT and PCT sequences expression in human cells, taking into account their sequence specificities as well as the orientation of the generated transcripts. Using this macro-array, we studied the expression of the different types of CT and PCT sequences during heat-induced stress. The presence and levels of the transcripts were also compared between normal and tumor cells, in a wide range of tissues. The impact of known chromatin remodelers on the expression of CT and PCT sequences was also investigated. Altogether, our data strengthen the idea that CT and PCT transcripts represent good indicators of the cell response to environmental changes and broad alterations of the epigenome. These transcripts may also represent additional markers for the detection of cancers and related diseases.

MATERIALS AND METHODS

Cell culture and treatments

Cell lines. HeLa cells, from cervix adenocarcinoma, and A431 cells, from vulva epidermoid carcinoma, were purchased from ATCC (American Type Culture Collection, VA, USA). IMR90 primary fibroblasts from embryonic lung were purchased from Coriell Institute (Camden NJ, USA). HCT116 cells, from colon epithelial carcinoma, and HSF1 non-expressing HeLa cells were obtained from Dr Bert Volgenstein (John Hopkins University, USA) and Lea Sistonen (University of Turku, Finland), respectively.

Stress conditions. Heat-shock experiments were performed by immersion of the culture flasks in a warm water bath. Unless mentioned, heat-shock was performed for 1 h at 43°C (cancer cells) or 45°C (normal cells). Recovery experiments following heat-shock were performed in an incubator at 37°C.

Drug treatments. Incubation with 5-Azacytidine (Aza-C) (Sigma) was performed for a period of 72 h, at a concentration of 5 µM. Aza-C was renewed every 24 h. Trichostatin A (TSA) (Sigma) and butyrate (BUT) were used at a concentration of 330 and 10 mM, respectively, with incubation time of 2 h (TSA) and 7 h (BUT).

SiRNA. Two SiRNA against Dicer were used: DicerA 5′UUUGUUGCGAGGCUGAUUC3′ and DicerB 5′UCUAUUAGCACCUUGAUGU3′ (11). Double transfections (DicerA + DicerB) were performed in HeLa cells with oligofectamine (Invitrogen) according to manufacturer's instructions.

RNA

For all cell lines, RNA extraction was performed with TRI-Reagent (Sigma) according to the manufacturer's instructions and RNA quality was assessed by the integrity of rRNA bands following gel electrophoresis. Total RNA was then quantified by UV spectrophotometry.

RNA pools from adrenal gland, bone marrow, whole brain, fetal brain, fetal liver, heart, kidney, liver, whole lung, placenta, prostate, salivary gland, skeletal muscle, testis, thymus, thyroid gland, trachea, uterus, colon and spinal cord were purchased from Clontech (Human Total RNA Master Panel II, 636643).

The set of RNA from human normal/tumor tissues from the same patient were obtained from Ambion. ovary: AM7256 (papillary cyctadenocarcinoma, 32-year-old female patient); liver: AM7244 (hepatoblastoma, 3-month-old male patient); lung: AM7224 (squamous cell carcinoma stage 2A, 60-year-old male patient); testis AM7260 (seminoma stage 1, 43-year-old male patient). Tumor tissue and normal lung parenchyma taken at distance from the bulk of the tumor from the same patient were obtained from the biological resource center (CRB) of CHU A.Michallon (Grenoble) (12). RNA from the lymph node biopsies of patients diagnosed with diffuse large B cell lymphoma wase obtained from the Grenoble hospital lymphoma collection. Normal spleen B cells were used as a control. All samples were obtained following informed consent and in accordance with institutional ethical guidelines.

RepChip

Oligonucleotides. Oligonucleotides (listed in Supplementary Table 2) were obtained from Invitrogen.

Macro array support. The combination of nylon array with radiolabeled probes is a valuable tool to detect low-expressed RNA due to the very high absorption capacity of nylon and to the sensitivity of radioactive labeling (13). Oligonucleotides of about 50-mer at 100 nM were loaded in four 96-well plates and colored with Bromophenol Blue in order to detect the presence of potential spotting default. Oligonucleotides were printed in duplicate onto the nylon membrane (Hybond N+, Amersham) by GMS 417 Arrayer automatic system. After spotting, targets were crossed-link to nylon membrane with short UV irradiation at 260 nm.

RNA labeling and hybridization. Three micrograms of total RNA were reverse-transcribed using both hexanucleotides (Roche) and dT25 (Invitrogen Life

Technologies) priming and α^{33}P-dCTP (Amersham or Perkin-Elmer) labeling. RNA annealing was performed with 8 µg/µl of dT25 during 8 min at 70°C and progressively cooled to 42°C. The reverse transcription was carried out in a reaction mixture containing 40U RNasin (RNasin Ribonuclease Inhibitor, Promega), 30 µCi α^{33}P-dCTP, 0.4 mM unlabeled dATP, dTTP and dGTP, 2.4 µM dCTP (all from Promega), 1× reverse transcription buffer (Promega) and 400 U Reverse Transcriptase (Promega), during 2 h at 42°C. Radiolabeled cDNA were purified using an exclusion column (YM-30, Millipore). Reverse transcription efficiency was checked by radioactive counting of 1 µl of the reaction on Ready CapTM in a Spectrometer (Beckman). Template RNA was removed by treatment at 68°C for 30 min with 10% SDS, 0.5M EDTA and 3M NaOH. Neutralization was done by adding 1 M Tris–HCl and 3 N HCl. After denaturation at 100°C for 5 min, the probe was incubated with the pre-hybridized nylon membrane in hybridization buffer (5× SSC, 5× Denhardt's, 12% BSA) for 5 days at 50°C. Membranes were then washed for 2 h at 45°C in 5× SSC and 0.5% SDS. Dry membranes were exposed onto a phosphorimaging plate in a Fujifilm BAS cassette for 24 h then scanned with a radio-imager BAS 5000 (Fuji). Radio-imagerFLA-8000, (Fuji) was used for the analysis of CT and PCT in lung tissues.

Signal analysis. The intensity of the different signals was quantified with the radio-imager (BAS 5000, Fuji) and corrected with Image Gauge software (Fuji) so that any data with intensity lower than the background plus 3 SD was excluded. The inter-assays normalization for each group is indicated in the text. Results are presented as box-plots with the thick line representing the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom circles the maximum and minimum values, respectively. *P*-values were calculated using a paired *T*-test.

Methylated DNA immunoprecipitation (MeDIP) assay

Genomic DNA was prepared using DNeasy Tissue Kit (Qiagen) according to the manufacturer's instructions. The quality of the DNA was controlled by gel analysis and spectrometry. For MeDIP experiments sonicated DNA was denatured and immunoprecipitated with an anti 5-methyl-cytosine antibody (Megabase Research Products) (14). Control immunoprecipitation was performed with anti IgG antibody (Sigma). Real-time PCR was then performed with specific primer pairs. The relative enrichment of target sequences after immunoprecipitation was calculated by comparing the number of cycles between MeDIP-enriched samples and input DNA for sequences of interest and a control sequence.

Immunofluorescence and western blot

Immunofluorescence. HeLa cells were grown on glass slides. The heat-shock experiment was performed by immersion of petri dishes containing the slides in a warm-bath at 43°C for 1 h. Immunofluorescence was performed on formaldehyde-fixed cells as previously described (15). The working dilution for the rabbit anti-HSF1 antibody (Stressgen) and for anti-HP1beta (Euromedex) was 1 : 500. Immunostaining of lung cancer tissues with an anti-H3K27me3 antibody (Upstate) was done on frozen sections as previously described (12) with a working dilution of 1 : 500.

Western blot. HeLa were HS for 1 h at 43°C. Protein extraction was realized in 8 M Urea. The western blot analysis of extracts was performed on 8% acrylamide gels for HSF1 detection. The working dilution for the rabbit anti-HSF1 antibody and the mouse anti-Dicer (Abcam) antibody were respectively 1 : 5000 and 1 : 100.

RT–PCR and qPCR

RT–PCR. One microgram of total RNA was reverse transcribed thanks to Super Script III first strand synthesis system for PCR (Invitrogen) according to manufacturer's instructions. PCR was performed in presence of Taq polymerase. Primers for Dicer are 5′CATGGATAGTG GGATGTAC3′ and 5′CTACTTCCAGAGTGACTC TG3′ (11). The amplification profile was 94°C for 30 s, 55°C for 30 s, 72°C for 30 s.

qPCR. qPCr reactions were performed using Brilliant SYBR Green qPCR MasterMix (Stratagene) on an M×3005p cycler (Stratagene).

RESULTS

Design of an oligonucleotide array aiming to detect CT and PCT specific transcripts

Human CT and PCT sequences display a higher level of polymorphism than in the mouse, hampering the efficiency and interpretation of PCR-based approaches for the analysis of their expression.

We have designed oligonucleotides aiming to represent all CT and PCT sequences present in the human genome. Several approaches were undertaken with two main objectives. First, our aim was to design a set of probes as representative as possible of all human CT and PCT sequences in order to detect the corresponding transcripts if present. Second, we also designed a set of chromosome specific probes enabling the detection of specific patterns of expression in different environmental or cellular contexts.

Based on FISH analyses, the similarities and divergence among PCT and CT human sequences allows the design of particular CT and PCT probes, which can detect either all of the human chromosomes or specific chromosomes, reflecting the heterogeneity and complexity of the composition of these sequences. In this general context, the design of oligonucleotides aiming to represent all CT and PCT sequences present in the human genome was performed using different strategies.

Chromosome specific CT sequences were obtained by multiple alignments of chromosome specific sequences with the 'Multalin DNA sequences' software as described

in ref. (14). The use of multiple related sequences minimizes the risk of missing specific CT or PCT transcripts, which was our major concern. Three oligonucleotides specific for chromosome 7, 10 and 16 respective CT regions (CT7′, CT10′ and CT16′) were added to the set of CT chromosome specific sequences. These sequences correspond to regions spanning preferential sites of integration of latent HIV virus in infected cells (16). CT7′ and CT10′ sequences were also present in human EST database (Supplementary Table 1). Sequences corresponding to two non-overlapping regions of the 170-bp alphoid sequence (CTcons1 and 3) described by Vissel and Choo (6) were also added together with a sequence restricted to the cenpB motif (CTcons2) (17). These probes correspond to consensus sequences of alphoid sequences.

An extensive search for the presence of CT and PCT expression in EST databases was undertaken using the various genomic CT and PCT probes available in the literature. This analysis suggested that these sequences display a tissue-specific expression.

All PCT sequences described in the literature were blasted against EST libraries. Probes specific for chromosome 1 (PCT1) (18), chromosome 9 (PCT9b) (19), chromosome 14 (PCT14) (20) and chromosome 16 (PCT16a) (19) were designed, when possible within regions of homologies with EST sequences but within regions devoid of long stretches of GGAAT repeat. A well-characterized sat III genomic sequence of unknown chromosome origin (5) was also used (PCTncs). For chromosome 9, an oligonucleotide was also designed (PCT9a), spanning a region of lesser homology with EST sequences with the idea that differences of frequency within EST databases might reveal distinct patterns of expression. For chromosome 16, an oligonucleotide containing a duplication of a consensus sequence specific for chromosome 16 and available in the literature was also used (PCT16b) (19). The specificity of three of these probes was tested by fluorescent in situ analysis confirming the specificity of PCT1, PCT9a while PCT16b recognized both chromosomes 1 and 16 (data not shown). Four consensus sequences were also designed and added to the list of PCT specific oligonucleotides. One is a succession of GGAAT motifs (PCTcons3), one corresponds to a sat III specific sequence already described in the literature (PCTcons2) (21). Two other consensus sequences PCTcons1 and PCTcons4, were obtained by multiple alignment of a set of sat II and sat III sequences (14).

In order to globally assess the transcriptional activity of CT and PCT sequences, oligonucleotides for CT and PCT sequences were designed in both orientations, which were arbitrarily termed 'sense' and 'antisense'. For CT sequences, the term 'sense' refers to the A/T rich transcripts in the same orientation as the consensus alphoid sequence of 170 bp, displayed in the paper by Vissel and Choo (6). For PCT transcripts the term 'sense' refers to G-rich transcripts.

The sequences of the selected set of CT- and PCT-specific oligonucleotides are given in Supplementary Table 2 together with control oligonucleotides corresponding to Actin B, histone H4, 18S and 28S ribosomal genes and hsp genes (hsp40, hsp70).

Characterization of the PCT and CT sequences expression in response to heat-shock

Since heat-shock is known to induce the expression of PCT sequences, primarily at the 9q12 locus (8,9) our objective was to perform a quantitative analysis of PCT and CT transcripts in both orientations. Indeed, while the transcriptional activation of chromosome 9 specific PCT sequences has been clearly documented in the heat-shock response, the transcriptional capacity of other PCT sequences and of CT sequences has not been yet fully addressed (8,9).

The expression of CT and PCT sequences was measured in HeLa cells over a continuous heat-shock exposure of 30 min to 1 h, as well as in HeLa cells submitted to a 1 h of continuous heat-shock followed by a recovery period from 1 to 6 h.

As shown in Figure 1a, no stress-dependent variation of signal intensity was observed for the control genes actin B, histone H4 and rRNA. In contrast, an accumulation of transcripts was detected with the whole set of oligonucleotides specific for PCT sequences with a peak of accumulation observed after 1 h of recovery following heat-shock. This pattern of expression, closely correlated with that of hsp genes, indicated that the whole set of chromosome specific PCT sequences behaved similarly during heat-shock. No expression of any of the CT sequences was detected upon heat-shock in HeLa cells, demonstrating that PCT sequences are globally and specifically activated during heat-shock, and that the respective expression of CT and PCT sequences are submitted to independent pathways and control mechanisms.

The respective levels of expression of sense and antisense PCT transcripts in non-heat-shocked (NHS) and heat-shocked (HS) cells were then compared (Figure 1). The corresponding values are represented as box plots in Figure 1b. Interestingly, the respective levels of induction of sense and antisense PCT transcripts during heat-shock were different, since they were respectively increased 44.9 and 11.6 times. The possibility that antisense transcripts, which are less abundant, could represent artifacts of the reverse transcription procedure (22) was examined using 9q12 specific PCT transcripts. The products of reverse transcription of sense or antisense 9q12 specific transcripts generated *in vitro*, were hybridized to sense and antisense oligonucleotides. As expected, reverse transcripts obtained from sense 9q12 specific PCT transcripts only hybridized to sense oligonucleotides while reverse-transcripts generated from antisense transcripts only hybridized to antisense oligonucleotides, indicating that the expression of PCT sequences in an antisense orientation actually occurs in the cell, and is not the mere consequence of the experimental procedure (data not shown). The rate of induction of sense PCT sequences that we find upon heat-shock is in agreement with previous experiments quantifying radiolabeled reverse chromosome 9 specific transcripts from NHS and HS HeLa cells (8), using the same cells heat-shock conditions.

HSF1 is known to be a key determinant in the heat-shock response (8). The question of whether PCT
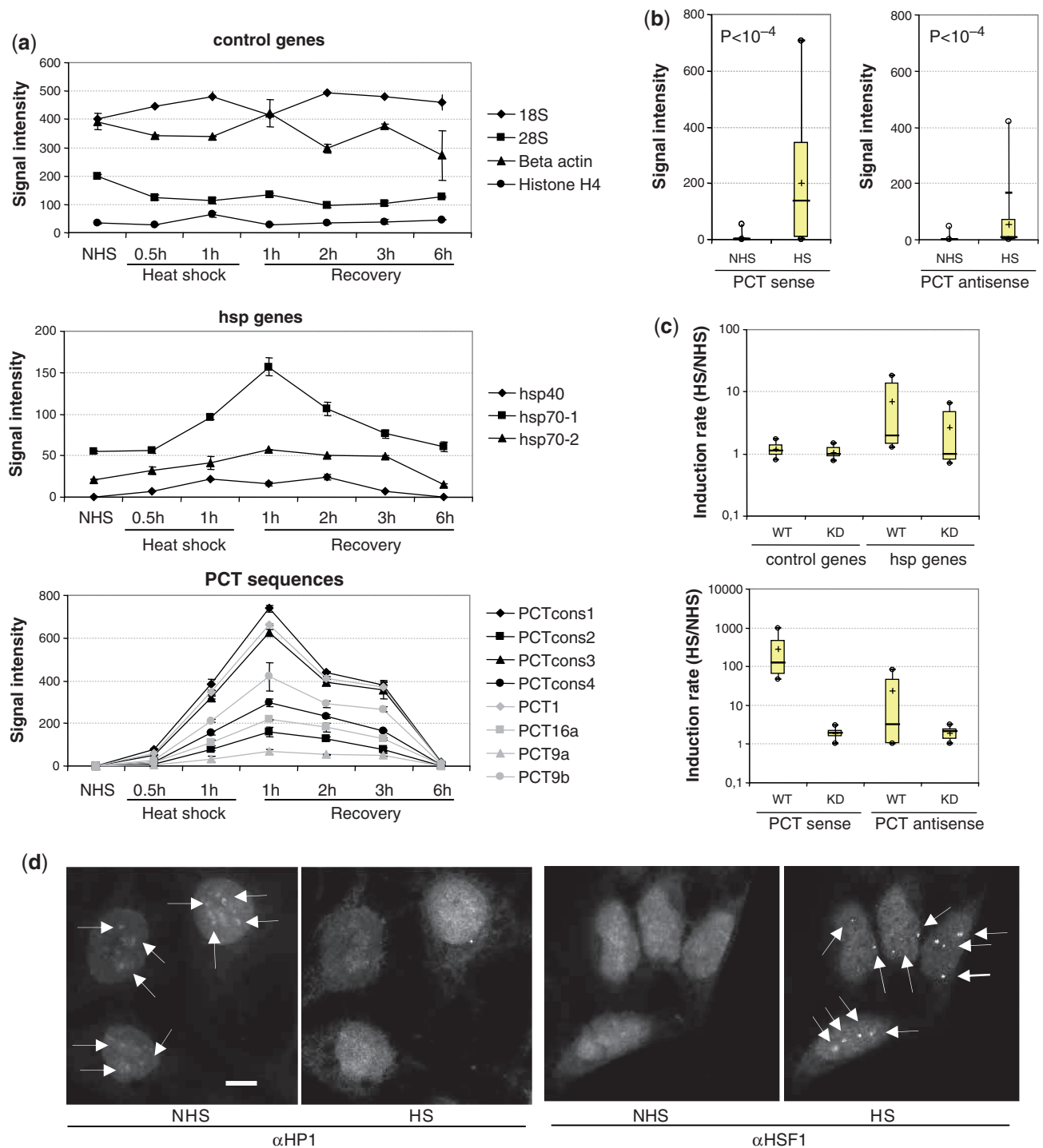
Figure 1. Kinetics of PCT sequence expression during heat-shock response in HeLa cells, detected by RepChip. (**a**) Expression of sense PCT transcripts in the course of the heat-shock response. The expression of PCT sequences was evaluated in NHS HeLa cells, in HeLa cells over a continuous heat-shock exposure of 30–60 min, as well as in HeLa cells submitted to a 1 h of continuous heat-shock followed by a recovery period of 1–6 h. A peak of accumulation of hsp and PCT specific sense transcripts was observed after 1 h of recovery following heat-shock. No heat-shock dependent variation of expression was observed for control genes. (**b**) Sense and antisense PCT sequences expression in NHS and HS cells. Raw values of signal intensities corresponding to sense and antisense PCT sequences expression in NHS HeLa cells and in HeLa cells submitted to a 1 h of heat-shock (HS) are displayed as box plots. The *P*-values were calculated using a paired *T*-test. The induction rates of sense and antisense PCT sequences upon stress are respectively 44.9- and 11.6-fold indicating that PCT transcripts in sense and antisense orientations are more abundant in HS cells than in NHS cells. (**c**) Impact of HSF1 on PCT sequence expression. PCT sequence expression induction was determined in HSF1 expressing (WT) and non-expressing (KD) HeLa cells. The values of signal intensities corresponding to sense and antisense PCT sequences expression in HS versus NHS HeLa cells are displayed as ratios (HS/NHS). Heat-induced accumulation of sense and antisense PCT transcripts and of hsp transcripts is severely impaired in KD HeLa cells. (**d**) Heat-shock is accompanied by a nuclear redistribution of HP1ß. In NHS cells, HP1ß displays a non-homogeneous nuclear distribution and is enriched in pericentric heterochromatic regions (white arrows). In HS cells [assessed by the presence of HSF1 nSBS (white arrows)], HP1ß displays a homogeneous nuclear distribution. Bar: 5 μm.
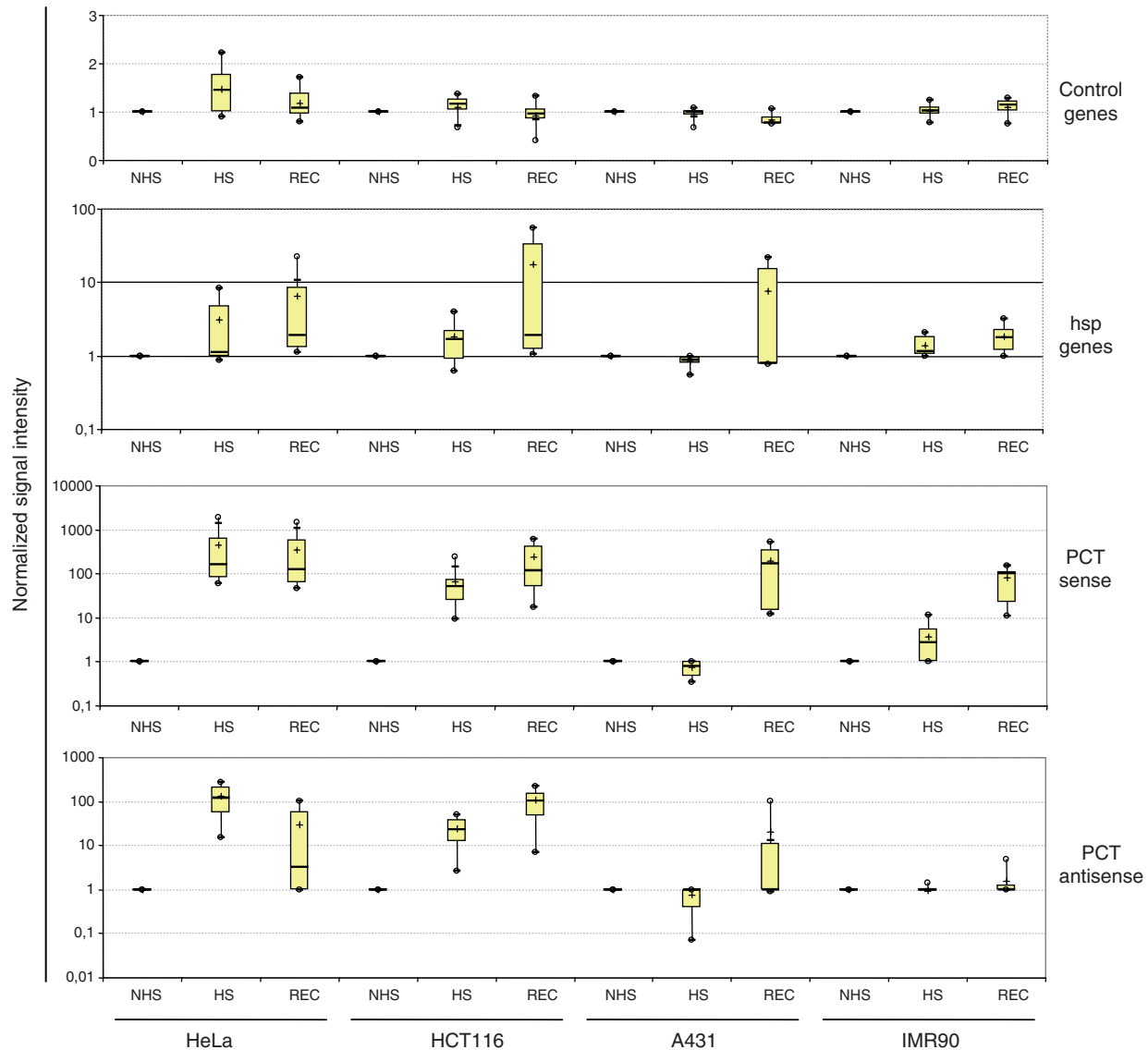
Figure 2. PCT sequence expression upon stress is a common feature of normal and cancer cell lines. The amount of sense and antisense PCT transcripts was analyzed in HeLa, HCT116, A431 and IMR90 cell lines submitted or not to a 1 h heat-shock with or without a period of recovery of 3 h following heat-shock. For each condition and cell type, the relative levels of expression of PCT sequences and of genes were normalized using the corresponding values in NHS cells. Heat-induced accumulation of PCT transcripts in both orientations is observed in all cell lines with marked differences in the amplitude and kinetics of accumulation. REC: recovery. The thick line represents the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom circles the maximum and minimum values, respectively.

expression is dependent on the HSF1 pathway was therefore addressed. Comparative analysis of PCT sequences expression in HS HeLa cells expressing HSF1 (WT) or knocked down for HSF1 (KD) was performed. As shown in Figure 1c, heat-induced accumulation of PCT transcripts was severely impaired in KD HeLa cells. This demonstrates that HSF1 is necessary for the heat-shock induced expression of PCT transcripts. It is likely that HSF1 dependent transcription of PCT sequences also involves chromatin-remodeling events at these regions. In HS cells, chromosome 9 specific PCT regions are not enriched in H3K9me3. Interestingly we found that heat-shock also causes a nuclear redistribution

of HP1ß, a well-known marker of pericentric heterochromatin (23) (Figure 1d).

So far, the only cellular model for which a heat-induced expression of PCT sequences has been formally demonstrated was cervical carcinoma HeLa cells. The capacity of other cell types to express PCT sequences, and possibly CT sequences, was therefore also assayed in colon cancer (HCT116) and vulva carcinoma (A431) cancer cell lines as well as in primary IMR90 fibroblasts. These cells were submitted (or not) to a 1 h heat-shock followed (or not) by a 3 h recovery at 37°C (Figure 2). Our results show that the heat-induced expression of PCT sequences is not restricted to HeLa cells, but

Table 1. Expression of CT and PCT sequences in normal tissues

	Immunity system		Central nervous system and effector muscles			Endocrine system		Digestive system		
	Thymus	Bone marrow	Brain	Spinal cord	Skeletal muscle	Tyroid gland	Adrenal gland	Salivary gland	Liver	Colon
PCT	No	No	No	No	No	No	No	No	No	No
CT	No	No	No	No	No	No	No	No	No	No

	Cardiovascular system	Reproductive system				Respiratory system		Embryo and placenta		
	Heart	Testis	Prostate	Uterus	Ovary	Trachae	Lung	Liver	Brain	Placenta
PCT	No	Yes	No	No	No	No	No	No	No	No
CT	No	No	No	No	Yes	No	No	Yes	No	Yes

CT and PCT sequences expression was analyzed by the RepChip approach in 21 different tissues. As indicated by highlighted characters, the expression of PCT sequences was only detected in testis, in an antisense orientation, while an expression of CT sequences was detected in embryonic liver, placenta and ovary. In these tissues, CT7′ and CT10′ specific transcripts were respectively expressed in both sense and antisense orientations.

represents a common feature of all cell lines studied here, including normal and cancer cells. In these different cell lines, all PCT sequences displayed a profile of expression similar to that displayed in Figure 1a, with no variation in the relative level of intensity values obtained with the different oligonucleotides (data not shown). We also demonstrate that the kinetics of accumulation of PCT transcripts is different between cell types. Indeed, the peak of PCT transcripts accumulation was observed after 3 h of recovery in A431 and IMR90 cells, and only after 1 h of heat-shock in HeLa and HCT116 cells. In all analyzed cell lines, the kinetics of accumulation of antisense PCT transcripts was found to follow the same pattern as that of sense PCT transcripts. The difference in the kinetics of PCT accumulation in HS A431 and IMR90 cell lines on one hand, and in HeLa and HCT116 cell lines on the other hand, may reflect differences of cell sensitivity to heat-shock. Supporting this hypothesis, in both cases, the kinetics of PCT transcripts accumulation was similar to that of hsp specific transcripts.

CT sequences were not found transcribed in IMR90, HeLa or HCT116 cells (data not shown). Interestingly, a sporadic expression of transcripts corresponding to three CT sequences (CTcons3, CT10′ and CT16) was detected in a sense orientation in both NHS and HS A431 cells (data not shown). The fact that the transcription of CT sequences are not induced by heat-shock indicates that the control of CT and PCT sequences relies on different mechanisms and reinforces the idea that PCT sequences fulfill specific functions in HS cells (7,24).

These results show that the expression of all PCT sequences is induced by heat-shock through an HSF1 dependent pathway and occurs mainly in a sense orientation, whereas most CT sequences remain silent. Moreover, we show that the heat-induced expression of PCT sequences occurs in all tested normal and cancer cell lines. We then addressed the question of the expression of CT and PCT sequences *in vivo*, in normal differentiated tissues, as well as in tumors.

Expression of CT and PCT sequences in normal and cancer tissues

To assess the transcriptional capacities of CT and PCT sequences *in vivo*, their expression was analyzed in a panel of RNA fractions from different normal tissues. RNA fractions from adult tissues of different origins (immunity system, central nervous system, muscle, endocrine system, digestive system, cardiovascular system, reproductive system, respiratory system) were analyzed together with tissues from fetal and extraembryonic origin (liver, brain and placenta) (Table 1).

Most of the examined normal tissues showed no expression of CT sequences. However, transcripts corresponding to two different CT sequences, CT7′ in a sense orientation and CT10′ in an antisense orientation, were detected in ovary, placenta and fetal liver (data not shown). Interestingly our *in silico* analysis show that these two CT sequences were found to give the highest numbers of blast hits among the whole set of CT sequences (see Supplementary Table 1) suggesting that they indeed display a specific expression compared to other CT sequences.

PCT sequences were not expressed in any of the normal somatic tissues, but, surprisingly, antisense PCT transcripts were detected in the testis (Figure 3). This observation was confirmed in three testis samples from different individuals (data not shown). Since no accumulation of hsp-specific transcripts was observed in testis, in this case, the transcription of PCT sequences does not seem to be related to the heat-shock pathway.

We then wished to compare the expression of CT and PCT sequences between normal and cancer cells originated from the same patient, using pairs of normal and cancer tissues from testis, liver, ovary and lung each obtained from one patient. Differences in the general pattern of CT or PCT sequences were clearly detected in two of the pairs.

Interestingly, a loss of antisense PCT transcripts was observed in testis cancer when compared to normal testis showing that, in this particular case, the tumor phenotype is associated with a down-regulation of PCT
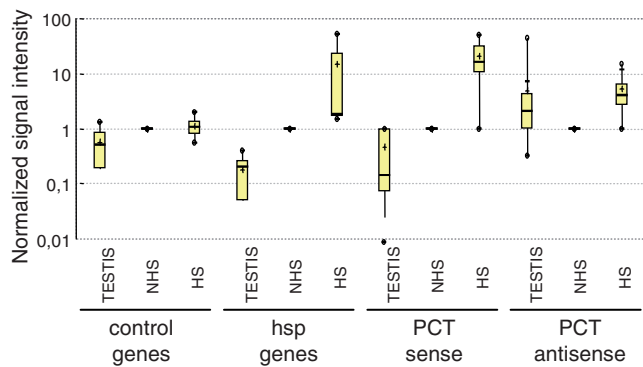
Figure 3. Comparative analysis of PCT sequences expression in HeLa cells and testis. The intensity values obtained for control genes, HS genes and PCT sequences expression in testis, in NHS and HS HeLa cells are presented as box plots. For each condition, the relative levels of expression of PCT sequences and of control genes were normalized using the corresponding values in NHS HeLa cells. In contrast to HS HeLa cells, where PCT transcripts are detected in both orientations, PCT sequences in testis are mostly expressed in an antisense orientation. The accumulation of PCT transcripts in testis does not seem to involve the heat-shock signaling pathway since the intensity values corresponding to the heat-shock genes remained at a lower level than in HS HeLa cells. The thick line represents the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively and the top and bottom circles the maximum and minimum values, respectively.

sequences expression (Figure 4a). A striking difference between normal and cancer tissues was also observed in lung where an accumulation of sense and antisense CT (CTcons1s, CTcons2s, CTcons3s, CTcons3as, CT8 and CT16) and PCT transcripts was detected in the tumor sample. The absence of induction of hsp genes, confirmed by qPCR analyses, suggests that, in this case again, the heat-shock pathway is not involved (Figure 4b). CT specific transcripts were also detected in normal (CTcons3as, CT7′s and CT10′as) and cancer (CTcons3as, CT7′s and CT10′as, CT21s) liver tissues as well as in normal ovary tissue (CT7′ and CT10′). The presence of CT transcripts in normal liver might represent a specific feature of early differentiating liver cells, since the samples representing the normal and cancer liver tissues were obtained from a three months old baby and since, as mentioned above, the expression of CT sequences was also observed in normal fetal liver cells. The importance of CT and PCT transcripts expression in lung cancer tissue prompted us to extend the analysis of lung cancer tissues to a larger number of tumor tissues samples for which normal tissues corresponding to the same patients were also available (12). The result of this analysis is displayed in Figure 5 showing a major and global expression of CT and PCT sequences in both orientations in tumor tissues. The list of CT sequences
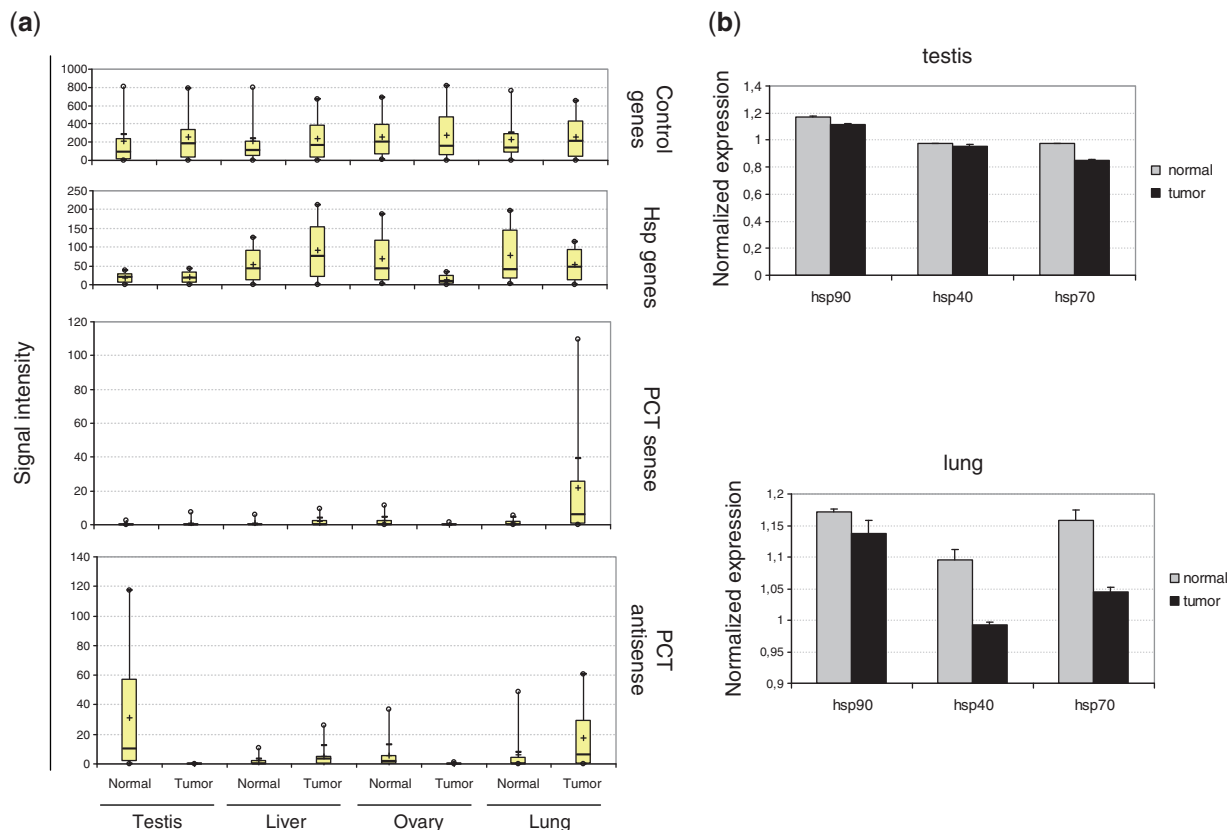
(a)

(b)

Figure 4. Variations in CT and PCT sequences expression in normal and tumor tissues. (**a**) Expression of PCT transcripts in testis, liver, ovary and lung normal and tumor tissues. For each tissue, the RNA fractions originating from normal and tumor tissues from the same patient were analyzed. Signal intensities corresponding to control and hsp genes and to PCT sequences in a sense and antisense orientations are represented as box plots. An accumulation of antisense PCT sequences is observed in normal testis. PCT transcripts in both orientations are abundant in tumor lung tissue. The thick line represents the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom circles the maximum and minimum values, respectively. (**b**) hsp genes expression in testis and lung normal and tumor tissues. Hsp genes expression were analysed by qPCR. No variation of hsp genes expression was detected between normal and tumor testis tissues. No induction of hsp genes expression was observed in tumor lung tissue compared to normal lung tissue.
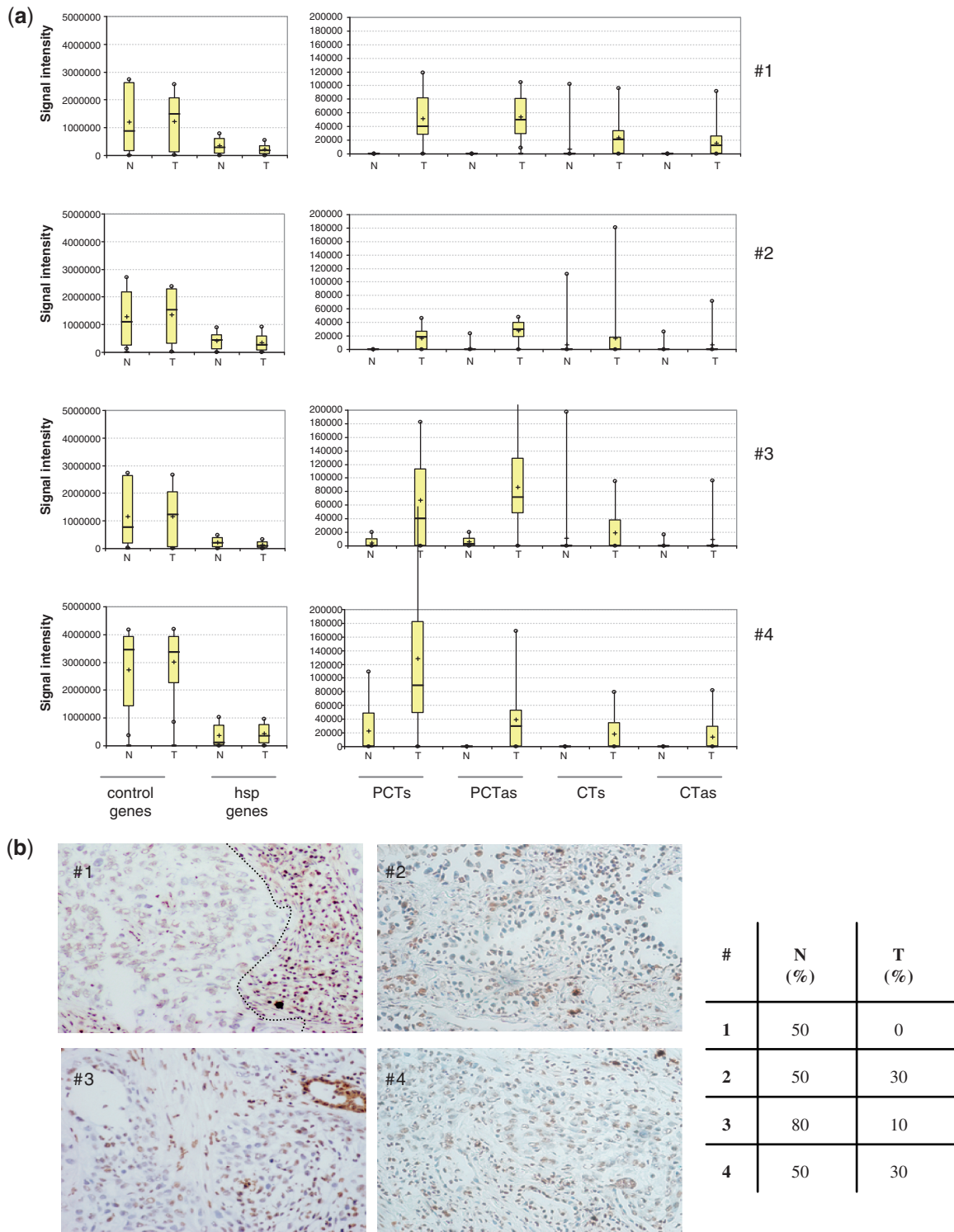
Figure 5. Variations in CT and PCT sequences expression in a series of normal and tumor lung tissues. (**a**) Analysis of the expression of CT and PCT in four normal (N) and corresponding tumor lung tissues (T). The signal intensities corresponding to control and hsp genes and to CT and PCT sequences are represented as box plots. An accumulation of sense and antisense CT and PCT sequences in both orientations is observed in tumor lung tissues. N: normal, T: tumoral, #: patient identification. The thick line represents the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom circles the maximum and minimum values, respectively. (**b**) Global loss of H3K27me3 specific labeling in lung tumors. Immunohistochemistry analysis of lung cancer tissues from Patients 1, 2, 3 and 4 was performed on frozen sections using anti-H3K27me3 (immunoperoxidase labeling ×200). Nuclear staining of trimethylated H3K27 is observed in non-tumor cells while a decrease of H3K27me3 labeling is observed in tumor cells as compared to normal cells: in one case (Patient 1) the tumor limit is indicated by a dotted line between tumor cells (left) and adjacent normal cells (right). For each patient analyzed, the percentage of H3K27me3 positive cells respectively found in normal (N) and tumor tissues (T) is given in the table.

Table 2. Expression of CT sequences in normal and tumor lung tissues

Patient no.	Normal tissue	Tumor tissue
1	21s	cons1s, cons1as, cons2s, cons3s, cons3as, 1as, 2s, 2as, 3s, 3as, 4s, 4as, 7′s, 7′as, 8s, 8as, 9s, 9as, 10′s, 10′as, 12s, 12as, 16s, 16as, 16′s, 16′as, 17as, 18s, 20s, 20as, 21s, 22s, 22as
2	21s	cons1s, cons1as, cons2s, cons3s, cons3as, 1as, 2s, 2as, 3s, 9s, 10′s, 10′as, 12s, 13s, 13as, 16as, 16′s, 16′as, 21s
3	21s	cons1s, cons1as, cons2s, cons3s, cons3as, cons4s, cons4as, 2s, 2as, 4s, 7′s, 7′as, 10′s, 10′as, 16s, 16as, 21s
4		cons1s, cons1as, cons3s, cons3as, 2s, 2as, 3s, 8s, 8as, 9s, 9as, 10′s, 12s, 16s, 16as, 16′s, 17as, 18s, 18as, 20s, 20as

List of oligonucleotides which detect CT specific transcripts in normal and tumor lung tissues. In all cases, an expression of PCT sequences in both orientations was also detected.

detected in the four tumor samples analyzed is given in Table 2. No induction of hsp genes expression was detected by transcriptomic (Figure 5a) and qPCR (data not shown) approaches in cancer tissues, suggesting that the heat-shock pathway is not involved in the expression of CT and PCT sequences. The above results suggest that a de-regulation of CT and PCT sequences can occur in cancers, leading to the abnormal expression of these sequences in some somatic tumors, or to the aberrant repression of PCT sequences in testis cancer.

The possibility that a de-regulation of CT and PCT sequences could represent a specific features of certain types of cancer, such as lung cancer for which CT and PCT transcripts were detected in five out of five samples, is suggested by an additional study that we also performed on B cell lymphoma. In this case, CT and PCT sequences in both orientations were only detected in three out of ten patients, independent of the heat-shock pathway (data not shown). Interestingly, in lung tumor cells, a global loss of H3K27me3, an epigenetic mark of heterochromatin (25), was detected by immunohistochemistry in all four lung tumor tissues analyzed compared to non-tumoral tissues (Figure 5b).

Impact of chromatin remodelers on the expression of PCT and CT sequences

To get an insight into the mechanisms involved in the transcriptional control of PCT and CT sequences, we next sought to determine the impact of two major epigenetic landmarks on the expression of CT and PCT sequences.

The impact of histone acetylation was assessed using Trichostatin A (TSA), a potent inhibitor of class I HDAC and butyrate (BUT), a specific inhibitor of class I and II HDAC but HDAC6. The effect of these HDACs inhibitors on the expression of CT and PCT sequences was monitored in HeLa cells. No significant increase of constitutive expression of PCT (Figure 6a) or CT sequences (Supplementary Figure 1) was observed in cells treated with TSA or butyrate, supporting that global changes in histone acetylation levels are not sufficient to induce an activation of satellite sequences expression.

We then explored the effect of a drug-induced DNA demethylation on the expression of CT and PCT sequences in HeLa cells treated with 5-Aza-Cytidine (Aza-C), a potent demethylating agent. The hypomethylated status of PCT sequences in Aza-C treated HeLa cells

was confirmed by MeDIP (Figure 6b). As shown in Figure 6a, an accumulation of PCT specific transcripts was observed in Aza-C treated HeLa cells. All chromosome specific PCT sequences were expressed. The relative levels of expression of the different probes were similar to those found in HS cells (data not shown). The absence of CT specific transcripts (Supplementary Figure 1) in these cells despite demethylation of these sequences (Figure 6b), strengthens the hypothesis made above, that CT and PCT sequences are not regulated by the same pathways. The expression of CT and PCT sequences was also analyzed in HCT116 cells knock-out for dnmt3b (3bKO), a DNA methyltransferase involved in *de novo* DNA methylation, as well as in HCT116 cells double knock-out (DKO) for dnmt3b and dnmt1, involved in DNA methylation maintenance. In the parental HCT116 cell line, no expression of CT and PCT sequences was detected (Supplementary Figure 1). Despite a lower level of DNA methylation at PCT sequences in both 3bKO and DKO cell lines (14), no de-repression of PCT sequences expression was observed when compared to the parental HCT116 cells (Figure 6a). This apparent discrepancy between the respective effect of drug-induced and genetically-induced DNA demethylation suggests the existence of other mechanisms repressing PCT sequences, despite DNA hypomethylation in KO cells.

The transcriptional activation of PCT sequences in Aza-C treated cells could involve the heat-shock pathway. In order to test this hypothesis, the level of hsp genes transcripts was quantified by qPCR while the presence of the hyperphoshorylated/DNA binding competent form of HSF1 was assessed by western blot analysis and *in situ* approaches. Upon stress, HeLa cells displayed reduced electrophoretic mobility and active HSF1 was present within nSBs in HeLa cells (Figure 6d). In Aza-C treated HeLa cells, no accumulation of hsp specific transcripts (Figure 6c) and no formation of nSBs enriched in active HSF1 was observed (Figure 6d), showing that the accumulation of PCT transcripts in Aza-C treated cells is HSF1 independent.

Since hypomethylation of PCT sequences allows an accumulation of PCT transcripts in Aza-C treated cells, we next assayed whether accumulation of PCT specific transcripts in HS cells conversely involve demethylation of these sequences. This hypothesis was examined using MeDIP with DNA from NHS and HS HeLa cells. As shown in Figure 6e, no variation of DNA methylation

of PCT sequences was observed between NHS and HS cells suggesting that demethylation of PCT sequences is not a prerequisite to their expression and that the expression of PCT sequences in response to DNA hypomethylation and heat-shock relies on two independent pathways.

Impact of Dicer on CT and PCT sequences expression

In fission yeast *Schizosaccharomyces pombe*, small-sized PCT transcripts, generated by Dicer, through targeting by the RITS complex, play a role in both the structure and the function of heterochromatin regions (26,27). Thus, by analogy to what occurs in the yeast model, long transcripts detected in human cells may represent precursors of small dsRNA (7). There is no direct evidence in the literature that 21–25-nt-long dsRNA are generated from CT and PCT sequences in human cells. We have not ourselves been able to visualize 21–25-nt-long CT or PCT transcripts in RNA fractions from NHS Hela cells enriched in small-sized molecules (<200 nt). In the human cells we have studied, if they exist, small-sized RNA from CT and PCT sequences are scarce. The possibility that heat-shock could alter the activity of the Dicer machinery and, as a consequence, induce an accumulation of PCT specific transcripts, was nonetheless explored in Hela cells treated with siRNA targeting Dicer. The efficiency of the siRNA procedure was determined at a transcriptional and post-transcriptional level (Figure 7a). The absence of accumulation of PCT transcripts in Dicer deficient HeLa cells therefore suggests that the accumulation of PCT transcripts is not a mere consequence of an alteration of the Dicer machinery. The absence of accumulation of PCT (Figure 7b) or CT (data not shown) specific transcripts in DICER deficient HeLa cells also suggests that, in Hela cells, and possibly more generally in human cells, an alteration of the DICER machinery does not result in the accumulation of PCT and CT transcripts as it does in the chicken/human somatic hybrid cell line (28).

DISCUSSION

The expression of centromeric (CT) and pericentric (PCT) sequences has been shown to be involved in major genomic functions in lower eukaryotes. We and others have also shown that satellite sequences are actively transcribed under stress in several mammalian species including human (7) but the functional significance of this expression is still poorly understood. From our previous data, we can deduce that at least some of the PCT sequences remain associated to the region from which they originate (8). This could also be the case of the other PCT sequences, but has not been proven. Our objective was to explore the extent of expression of satellite sequences in human cells. However, in human cells, this question is particularly difficult to address because of the repetitive nature and high degree of polymorphism of these sequences. A macro-array specifically dedicated to study the expression of human satellite sequences was therefore designed here. It contains a set of oligonucleotides

representative of the large spectrum of CT and PCT specific repeats, which allowed an exhaustive analysis of the expression of satellite sequences in different physiological and pathological contexts. In order to optimize the sensitivity of detection, we chose to use nylon membranes and radioactive labeling rather then glass supports and fluorescent labeling. Accordingly, we show here that this approach allows the detection of transcripts, even expressed at low levels, with no need for PCR amplifications.

Using this tool we show that PCT sequences behave as a coherent entity and that chromosome 9 specific PCT sequences are not the only ones to be expressed. This observation suggests that, although primarily forming at the 9q12 locus, secondary sites of HSF1 accumulation, which are detected on other chromosomes enriched in satellite II or satellite IIII sequences including chromosome 1 (S.Fritah *et al.*, in preparation), are actively transcribed. The fact that in HS cells the expression of all PCT sequences is HSF1 dependent and correlated with that of hsp genes also suggests that their expression is finely controlled and coordinated. Our comparison of the kinetics of expression of PCT sequences and hsp genes between normal and cancer cell lines suggests that, as for hsp genes expression, differences in the kinetics of PCT sequences expression reveals differences of cell sensitivity to heat-shock.

For the first time, the expression of satellite sequences was explored in a large set of human tissues, revealing that in normal adult somatic cells these sequences are not expressed. Interestingly, in adult testis, PCT sequences are expressed in an antisense orientation and our results suggest that this expression is not correlated to a general activation of the heat-shock response. Recent published data have shown that Y chromosome specific PCT sequences in an antisense orientation are involved in transplicing events (29). We show here that other PCT are also specifically expressed in the testis. The fact that, among the large panel of tested tissues, testis is the only one in which an expression of PCT sequences was observed could be related to the specific reprogramming of male germinal cells during spermatogenesis (30). In some somatic cancers, we found that these sequences are abnormally expressed, whereas in testis tumor PCT sequences become aberrantly repressed. Interestingly this pattern of expression is similar to that of several testis-specific genes, which are normally exclusively expressed in the testis, repressed in normal somatic tissues, but sporadically and aberrantly expressed in some somatic cancers (31).

In cancer, the expression of CT and PCT sequences could reflect a global and major epigenetic deregulation. Indeed, in lung cancer tissue and to a lesser extend, in lymphoma, in contrast to the expression observed in HS cells or in testis, PCT sequences are expressed in both orientations, suggesting a non-regulated process. Similarly the PCT expression observed in Aza-C treated cells, occurs in both orientations, suggesting that DNA methylation is, at least partly, associated with the repression of PCT sequences in adult somatic cells.
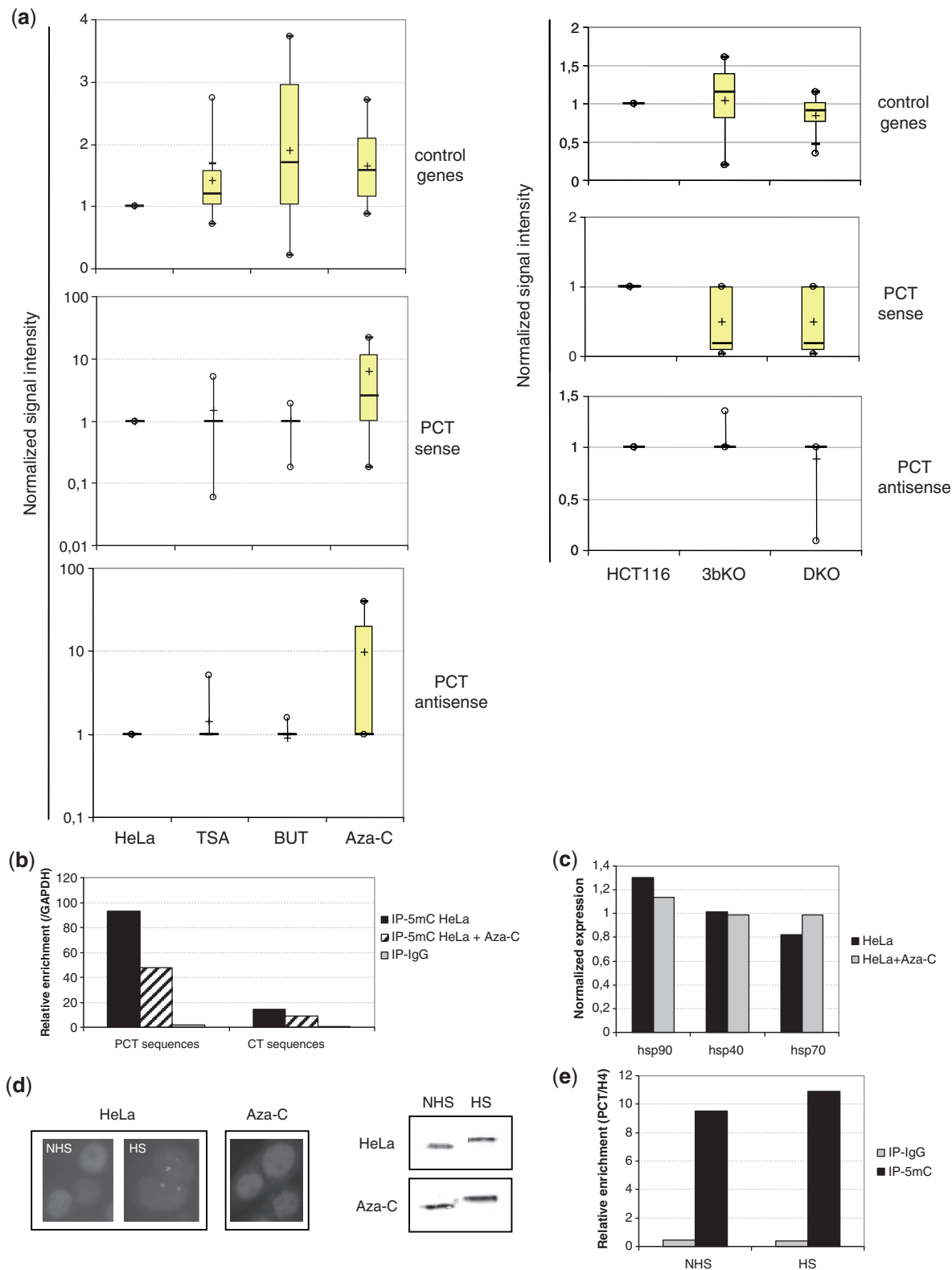
Figure 6. Epigenetic control of PCT sequences expression. (**a**) Drug-induced DNA demethylation can induce PCT sequences expression. Comparative analysis of PCT sequences expression was performed in NHS HeLa cells treated with Trichostatin A (TSA), butyrate (BUT) or 5-Azacytidin (Aza-C). Expression of PCT sequences was also analyzed in HCT116 cells deficient for dnmt3b (3bKO) and for both dnmt1 and dnmt3b (DKO) genes. The relative levels of expression of PCT sequences and of control genes were normalized using the corresponding values in NHS HeLa cells (left) or in normal HCT116 cells (right). PCT sequences expression is only detected in Aza-C treated cells (in both orientations). The thick line represents the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom circles the maximum and minimum values, respectively. (**b**) PCT sequences are hypomethylated in Aza-C treated cells. MeDIP analyses of
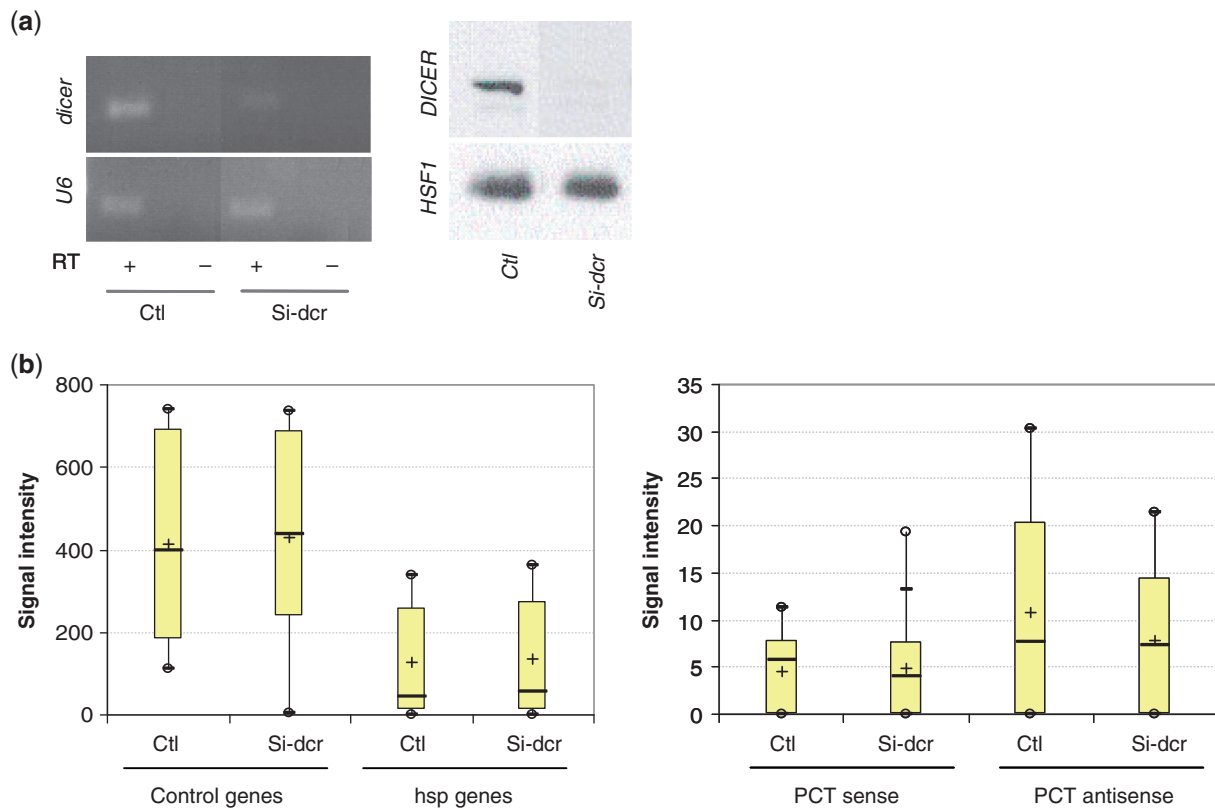
Figure 7. The knock down of Dicer (dcr) does not induce PCT sequences expression. (**a**) siRNA efficiency to generate a knock down of Dicer expression was validated by RT-PCR (left) and western blot analysis (right) on siRNA treated (Si-dcr) and control (Ctl) cells. U6 snRNA expression was used as a control in RT-PCT analysis while anti HSF1 was used as a control in western blot analysis. Dicer mRNA and Dicer protein rates were greatly decreased in Si-dcr cells. (**b**) Impact of Dicer knock-down on PCT sequences expression. Signal intensities corresponding to control genes, hsp genes and PCT sequences expression in control (Ctl) and Si-dcr HeLa cells are displayed as box plots. No significant difference in the expression of any of the sequences was observed between Ctl and Si-dcr treated cells. The thick line represents the median, the cross the mean, the top and bottom boundaries of the boxes corresponding to the 75 and 25 percentiles, respectively, and the top and bottom circles the maximum and minimum values, respectively.

It has been suggested that PCT transcripts could play a role in stabilizing heterochromatic structures at PCT regions (7). Increased transcription of PCT sequences in Aza-C treated cells and of CT and PCT sequences in cancer cells is therefore probably associated to the desta-bilization of heterochromatin in these cells, However, our data in HCT116 cells deficient for dnmt1 and dnmt3b suggest that DNA hypomethylation per se is not always associated to a de-repression of PCT sequences expression, in agreement with observations in mouse ES cells (32,33), suggesting that adaptative mechanisms prob-ably also exist to maintain a transcriptional repression of PCT sequences when these sequences are demethylated.

The existence of at least two independent pathways regulating the expression of PCT sequences, one involving HSF1 and the other involving DNA hypomethylation, indicates that expression of these sequences relies on vari-ous mechanisms involved in the control of either CT or PCT or both sequences expression. A recent publication by the group of G. Biamonti clearly shows, for example,

PCT and CT sequences were performed in Aza-C treated cells. The hypomethylated status of CT and PCT sequences in HCT116 dnmt mutated cells, described in ref. (14) is not represented herein. The histograms indicate the relative enrichment in 5-methyl-cytosine of target sequences after normalization to a reference promoteur (GAPDH promoter). IP-IgG: immunoprecipitation with non-specific IgG, IP-5mC: immunoprecipitation with antibody specific for 5-methyl-cytosine. (**c**) The expression of PCT sequences in Aza-C treated cells is not correlated to the accumulation of hsp specific transcripts. Hsp gene (hsp 90, hsp40 and hsp70) expression was evaluated by qPCR in Aza-C treated cells and in control HeLa cells (HeLa). Results were normalized, thanks to three control genes (GAPDH, U6 and histone H4). No induction of hsp genes was detected in these cells. (**d**) The expression of PCT sequences in Aza-C treated cells is HSF1 independent. The status of HSF1 activation in NHS and HS HeLa cells, Aza-C treated HeLa cells was determined by western blot analysis (right) and immunofluorescence with an antibody against HSF1 (left). Active HSF1 which displays a lower electrophoretic mobility was only detected in HS cells, but not in NHS Aza-C treated cells. Similarly, nuclear foci formed of active HSF1 and which form primarily form at the 9q12 locus were only present in HeLa HS cells. (**e**) PCT sequences expression upon stress is not associated with DNA demethylation. MeDIP analysis of PCT sequences was performed in NHS and HS HeLa cells. The histograms indicate the relative enrichment of PCT sequence after normalization to histone H4 promoter. No variation of DNA methylation of PCT sequences was observed between NHS and HS cells. IP-IgG: immunoprecipitation with non-specific IgG, IP-5mC: immunoprecipitation with antibody specific for 5-methyl-cytosine.

that another transcription factor, tonEBP, involved in the response to hypertonic stress, also activates PCT sequences expression (34). A loss of heterochromatic specific epigenetic markers could also favor an expression of these sequences. This idea is sustained by the absence of H3K9me3 in the pericentric regions of chromosome 9 (8,9) as well as delocalization of heterochromatin protein 1 (HP1) from PCT regions in HS Hela cells (Figure 1d).

Interestingly global epigenome alterations have been reported in cancer. For instance Fraga *et al.* (35) have reported a global change of histone marks in a large panel of cancer cells, including a loss of H4K20 methylation, a repressive mark normally enriched at PCT sequences. They also correlated this loss with DNA hypomethylation at PCT sequences. Moreover, we also reported a global loss of the repressive epigenetic mark H4K20me3 in lung cancer (12). All the four tumor lung tissues that we have examined by immunohistochemistry display a loss of H3K27me3 when compared to non-tumoral tissues. Whether global alterations of epigenetic marks can be correlated with the expression of CT or PCT sequences in cancers remains to be explored.

Our study reveals for the first time that CT and PCT sequences can be expressed in normal and cancer human cells and highlights the complexity of the mechanisms involved in the regulation of these sequences. Moreover, this work suggests that the expression of CT and PCT sequences may represent a good indicator of large-scale epigenetic modifications occurring in response to environmental conditions and/or in different contexts of cell differentiation and cell transformation.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

Conflict of interest statement. None declared.

REFERENCES

1. Grewal,S.I. and Elgin,S.C. (2007) Transcription and RNA interference in the formation of heterochromatin. *Nature*, **447**, 399–406.
2. Volpe,T.A., Kidner,C., Hall,I.M., Teng,G., Grewal,S.I. and Martienssen,R.A. (2002) Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. *Science*, **297**, 1833–1837.
3. Zaratiegui,M., Irvine,D.V. and Martienssen,R.A. (2007) Noncoding RNAs and gene silencing. *Cell*, **128**, 763–776.
4. Amor,D.J., Kalitsis,P., Sumer,H. and Choo,K.H. (2004) Building the centromere: from foundation proteins to 3D organization. *Trends Cell Biol.*, **14**, 359–368.
5. Prosser,J., Frommer,M., Paul,C. and Vincent,P.C. (1986) Sequence relationships of three human satellite DNAs. *J. Mol. Biol.*, **187**, 145–155.
6. Vissel,B. and Choo,K.H. (1987) Human alpha satellite DNA—consensus sequence and conserved regions. *Nucleic Acids Res.*, **15**, 6751–6752.
7. Eymery,A., Callanan,M. and Vourc'h,C. (2009) The secret message of heterochromatin: new insights into the mechanisms and function of centromeric and pericentric repeat sequence transcription. *Int. J. Dev. Biol.*, **53**, 259–268.
8. Jolly,C., Metz,A., Govin,J., Vigneron,M., Turner,B.M., Khochbin,S. and Vourc'h,C. (2004) Stress-induced transcription of satellite III repeats. *J. Cell Biol.*, **164**, 25–33.
9. Rizzi,N., Denegri,M., Chiodi,I., Corioni,M., Valgardsdottir,R., Cobianchi,F., Riva,S. and Biamonti,G. (2004) Transcriptional activation of a constitutive heterochromatic domain of the human genome in response to heat shock. *Mol. Biol. Cell*, **15**, 543–551.
10. Biamonti,G. (2004) Nuclear stress bodies: a heterochromatin affair? *Nat. Rev. Mol. Cell Biol.*, **5**, 493–498.
11. Chendrimada,T.P., Gregory,R.I., Kumaraswamy,E., Norman,J., Cooch,N., Nishikura,K. and Shiekhattar,R. (2005) TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature*, **436**, 740–744.
12. Van den Broek,A., Brambilla,E., Moro-Sibilot,D., Lantuejoul,S., Brambilla,C., Eymin,B., Khochbin,S. and Gazzeri,S. (2008) Loss of Histone H4K20 trimethylation occurs in Preneoplasia and influences prognosis of non small cell lung cancer. *Clin. Cancer Res.*, **14**, 7237–7245.
13. Bertucci,F., Bernard,K., Loriod,B., Chang,Y.C., Granjeaud,S., Birnbaum,D., Nguyen,C., Peck,K. and Jordan,B.R. (1999) Sensitivity issues in DNA array-based expression measurements and performance of nylon microarrays for small samples. *Hum. Mol. Genet.*, **8**, 1715–1722.
14. Horard,B., Eymery,A., Fourel,G., Vassetzky,N., Puechberty,J., Roizes,G., Lebrigand,K., Barbry,P., Laugraud,A., Gautier,C. *et al.* (2009) Global analysis of DNA methylation and transcription of human repetitive sequences. *Epigenetics*, **5**, 339–350.
15. Jolly,C., Konecny,L., Grady,D.L., Kutskova,Y.A., Cotto,J.J., Morimoto,R.I. and Vourc'h,C. (2002) In vivo binding of active heat shock transcription factor 1 to human chromosome 9 heterochromatin during stress. *J. Cell Biol.*, **156**, 775–781.
16. Jordan,A., Bisgrove,D. and Verdin,E. (2003) HIV reproducibly establishes a latent infection after acute infection of T cells in vitro. *EMBO J.*, **15**, 1868–1877.

17. Matera,A.G. and Ward,D.C. (1992) Oligonucleotide probes for the analysis of specific repetitive DNA sequences by fluorescence in situ hybridization. *Hum. Mol. Genet.*, **1**, 535–539.

18. Cooke,H.J. and Hindley,J. (1979) Cloning of human satellite III DNA: different components are on different chromosomes. *Nucleic Acids Res.*, **6**, 3177–3197.

19. Moyzis,R.K., Albright,K.L, Bartholdi,M.F., Cram,L.S., Deaven,L.L., Hildebrand,C.E., Joste,N.E., Longmire,J.L., Meyne,J. and Schwarzacher-Robinson,T. (1987) Human chromosome-specific repetitive DNA sequences: novel markers for genetic analysis. *Chromosoma*, **95**, 375–386.

20. Choo,K.H., Earle,E., Vissel,B. and Kalitsis,P.A. (1992) Chromosome 14-specific human satellite III DNA subfamily that shows variable presence on different chromosomes 14. *Am. J. Hum. Genet.*, **50**, 706–716.

21. Tagarro,I., Fernández-Peralta,A.M. and González-Aguilera,J.J. (1994) Chromosomal localization of human satellites 2 and 3 by a FISH method using oligonucleotides as probes. *Hum. Genet.*, **93**, 383–388.

22. Perocchi,F., Xu,Z., Clauder-Munster,S. and Steinmetz,L.M. (2007) Antisense artifacts in transcriptome microarray experiments are resolved by actinomycin D. *Nucleic Acids Res.*, **35**, e128.

23. Fanti,L. and Pimpinelli,S. (2008) HP1: a functionally multifaceted protein. *Curr. Opin. Genet. Dev.*, **18**, 169–174.

24. Jolly,C. and Lakhotia,S.C. (2006) Human sat III and Drosophila hsr omega transcripts: a common paradigm for regulation of nuclear RNA processing in stressed cells. *Nucleic Acids Res.*, **34**, 5508–5514.

25. Trojer,P. and Reinberg,D. (2007) Facultative heterochromatin: is there a distinctive molecular signature? *Molecular Cell*, **28**, 1–13.

26. Verdel,A., Jia,S., Gerber,S., Sugiyama,T., Gygi,S., Grewal,S.I. and Moazed,D. (2004) RNAi-mediated targeting of heterochromatin by the RITS complex. *Science*, **303**, 672–676.

27. Buhler,M. and Moazed,D. (2007) Transcription and RNAi in heterochromatic gene silencing. *Nat. Struct. Mol. Biol.*, **14**, 1041–1048.

28. Fukagawa,T., Nogami,M., Yoshikawa,, M. Ikeno,M., Okazaki,T., Takami,Y., Nakayama,T. and Oshimura,M. (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat. Cell Biol.*, **6**, 784–791.

29. Jehan,Z., Vallinayagam,S., Tiwari,S., Pradhan,S., Singh,L., Suresh,A., Reddy,H.M., Ahuja,Y.R. and Jesudasan,R.A. (2007) Novel noncoding RNA from human Y distal heterochromatic block (Yq12) generates testis-specific chimeric CDC2L2. *Genome Res.*, **17**, 433–440.

30. Govin,J., Escoffier,E., Rousseaux,S., Kuhn,L., Ferro,M., Thévenon,J., Catena,R., Davidson,I., Garin,J., Khochbin,S. and Caron,C. (2007) Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J. Cell Biol.*, **176**, 283–294.

31. Simpson,A.J., Caballero,O.L., Jungbluth,A., Chen,Y.T. and Old,L.J (2005) Cancer/testis antigens, gametogenesis and cancer. *Nat. Rev. Cancer*, **5**, 615–625.

32. Lehnertz,B., Ueda,Y., Derijck,A.A., Braunschweig,U., Perez-Burgos,L., Kubicek,S., Chen,T., Li,E., Jenuwein,T. and Peters,A.H. (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.*, **13**, 1192–1200.

33. Martens,J.H., O'Sullivan,R.J., Braunschweig,U., Opravil,S., Radolf,M., Steinlein,P. and Jenuwein,T. (2005) The profile of repeat-associated histone lysine methylation states in the mouse epigenome. *EMBO J.*, **24**, 800–812.

34. Valgardsdottir,R., Chiodi,I., Giordano,M., Rossi,A., Bazzini,S., Ghigna,C., Riva,S. and Biamonti,G. (2008) Transcription of Satellite III non-coding RNAs is a general stress response in human cells. *Nucleic Acids Res.*, **36**, 423–434.

35. Fraga,M.F., Ballestar,E., Villar-Garea,A., Boix-Chornet,M., Espada,J., Schotta,G., Bonaldi,T., Haydon,C., Ropero,S., Petrie,K. *et al.* (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat. Genet.*, **37**, 391–400.