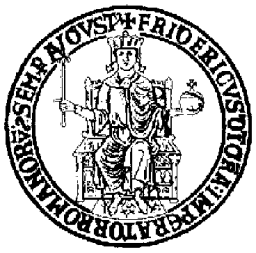
UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II



Dottorato di ricerca in
Patologia e Fisiopatologia Molecolare

XIX ciclo - 2003–2007

Coordinatore: Ch.mo Prof. V. E. Avvedimento

"Identification and characterization of CBX7 interacting proteins"

Tutor
Ch.mo Prof. Alfredo Fusco

Candidata
Dott.ssa Antonella Federico

ORIGINAL PUBLICATIONS

Pallante P, Federico A, Berlingieri MT, Bianco M, Ferraro A, Esposito F, Monti M, Pierantoni GM, Iaccarino A, Russo M, Forzati F, Fedele M, Leone V, Sacchetti S, Pucci P, Troncone G, Santoro M, Fusco A.

Critical role of the block of CBX7 expression in the progression step of thyroid carcinogenesis. (submitted for publication)

Visconti R, Federico A, Coppola V, Pentimalli F, Berlingieri MT, Pallante P, Kruhoffer M, Orntoft TF, Fusco A.

Transcriptional profile of Ki-Ras-induced transformation of thyroid cells.

Cancer Invest. 2007 Jun;25(4):256-66.

Pierantoni GM, Battista S, Pentimalli F, Fedele M, Visone R, Federico A, Santoro M, Viglietto G, Fusco A.

A truncated HMGA1 gene induces proliferation of the 3T3-L1 pre-adipocytic cells: a model of human lipomas.

Carcinogenesis. 2003 Dec;24(12):1861-9.

ABSTRACT

To identify the genes involved in the process of thyroid carcinogenesis we have analysed the gene expression profiles of six thyroid carcinoma cell lines compared with normal human thyroid primary cell culture using an Affymetrix HG_U95Av2 oligonucleotide array representing ~10.500 unique genes. This study led to the identification of genes whose expression was up- or down-regulated in the carcinoma cell lines compared with the primary cell culture of normal thyroid origin.

We identified 510 up-regulated transcripts and 320 down-regulated transcripts with a fold change higher than 10 in all of the carcinoma cell lines. In particular our interest was focused on the genes whose expression was drastically (at least 10 fold) up- or down-regulated in all the cell lines, assuming that the genes altered in their expression in all of the thyroid carcinoma cell lines might represent candidate genes involved in thyroid cell transformation.

Among the down-regulated genes, we directed our attention to Chromobox homolog 7 (CBX7) that resulted down-regulated about 20 fold in all the carcinoma cell lines analysed.

CBX7 is a chromobox family protein, probably involved in the compaction of heterochromatin. Expression analyses performed using several techniques such as RT-PCR and immunohistochemistry showed that the down-regulation of this gene strongly correlated with the onset of a highly malignant phenotype of thyroid carcinomas.

Thus, to better elucidate the function of CBX7 in carcinogenesis, we settled on the identification of novel CBX7 biological partners.

In the cells, many proteins display their biological functions through the rapid and transient association within large protein complexes. The association of an unknown protein with partners belonging to a specific protein complex involved in a particular process could inform on its biological function. Furthermore, a detailed description of the cellular signalling pathways might greatly benefit from the elucidation of protein–protein interactions in the cell.

For this aim we have performed a proteomic assay that led us to the identification of some interesting proteins that can modulate the activity of CBX7 to E-cadherin, an important mediator of the epithelial-mesenchymal transition (EMT).

Loss of E-cadherin expression seems to be crucial involved in EMT, and E-cadherin is therefore emerging as one of the caretakers of the epithelial phenotype. Downregulation of E-cadherin (or mutations that occur in cancer) has several important consequences that are of direct relevance to EMT.

The ability of CBX7 to positively regulate E-cadherin expression by interacting with HDAC2 and inhibiting its activity on the E-cadherin promoter would account for the correlation of the loss of CBX7 and E-cadherin expression with a highly malignant phenotype in thyroid cancer patients.

We have also identified another protein, Metastasis associated protein 3 (MTA3) that interact with CBX7 and could regulate E-cadherin expression, as already described in breast cancer.

Finally, another pathway that might be important in the regulation of activity of CBX7 is the arginine methylation by PRMT1 and PRMT5.

Thus, the many important interacting proteins of CBX7 and the pathways in which they are involved strongly suggest that CBX7 can be considerated a very important regulator of thyroid malignant transformation.

1 INTRODUCTION

1.1 The thyroid gland

The thyroid gland is the largest endocrine organ in humans (1). It is located in the neck region, on the anterior surface of the trachea, and is formed by two distinct cell types, the follicular cells (TFC) and the parafollicular or C cells (2). The TFC constitute the most numerous cell population and form the thyroid follicles, spherical structures that store and release the thyroid hormones (3). The C cells are located between follicles, mostly in a parafollicular position (Figure 1).

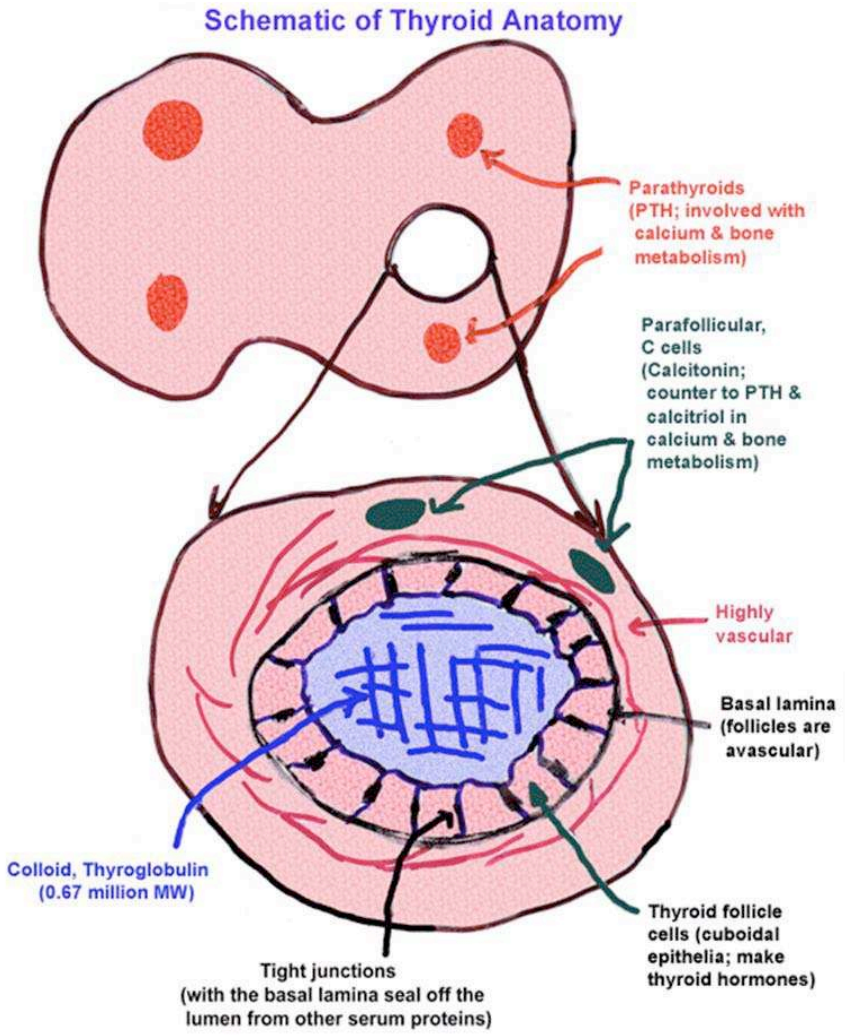


Figure 1 The general architecture of the thyroid gland
Figure shows the general architecture of the thyroid gland with the parafollicular cells, interspersed in small groups among the follicles in the intermediate part of the thyroid lobes.

These two cell types take origin from two different embryological structures: the thyroid anlage for the TFC and the ultimobranchial bodies for the C cells. The thyroid anlage is an area enclosing a small group of endodermal cells, and it is located on the midline of the embryonic mouth cavity in its posterior part.

The ultimobranchial bodies are a pair of transient embryonic structures derived from the fourth pharyngeal pouch and located symmetrically on the sides of the developing neck: C cell precursors migrate from the neural crest (4) in these structures (5). The cells of the thyroid anlage and the ultimobranchial bodies migrate from their respective sites of origin and ultimately merge in the definitive thyroid gland. Once merged, the thyroid anlage and the ultimobranchial bodies disappear and their cells disperse in the adult thyroid gland. At this point the cells from the anlage continue to organize the thyroid follicles, whereas the C cells scatter within the interfollicular space. Interestingly, in some animals the ultimobranchial structures remain distinct from the rest of the thyroid gland (6).

In the adult human, the gland has a butterfly shape and has a weight of 15–25 g. It is comprised of aggregates (lobules) of spherical follicles that are filled with colloid. The follicles range in size from 50–500 μm and are lined by cuboidal-to-flat follicular epithelial cells. The main functions of the thyroid gland are synthesis, storage and secretion of thyroid hormones, L-triiodothyronine (T3) and L-thyroxine (T4) (1). This mechanism is under the control of the hypothalamic-pituitary axis with negative feedback by the thyroid hormones: thyrotropin releasing hormone (TRH), secreted from the hypothalamus, stimulates the release of thyroid-stimulating hormone (TSH) from the anterior pituitary gland. TSH, then, is able to stimulate the follicular cells to synthesize and secrete

thyroid hormones. Disruption of thyroid hormone homeostasis results in hypothyroidism, goiter and in childhood cretinism (1).

The parafollicular C-cells secrete calcitonin, which is important for the bone formation. Its secretion is stimulated by elevate calcium concentration in the serum (7,8).

1.2 Thyroid follicular cell carcinomas

Thyroid carcinomas that originates from TFC are mainly divided into well-differentiated, poorly differentiated and undifferentiated carcinomas on the basis of clinico-pathological parameters such as prevalence, sex ratio, age, local/distant metastasis and survival rate (1). Papillary (PTC) and follicular (FTC) thyroid carcinomas are considered well-differentiated carcinomas (Kondo et al. 2006). Although initially defined by architectural criteria, the histological diagnosis of PTC rests on a number of nuclear features that predict the propensity for metastasis to local lymph nodes (9). The diagnosis of this most frequent type of thyroid malignancy (85–90% of thyroid malignancies) has been increasing, possibly owing to the changing recognition of nuclear morphological criteria. On the other hand, FTC is characterized by haematogenous spread, and the frequency of its diagnosis has been decreasing (10). Most well-differentiated thyroid cancers behave in an indolent manner and have an good prognosis. By contrast, anaplastic (ATC) thyroid carcinoma (undifferentiated) is highly aggressive and always fatal (9, 11). For this type of cancer there is not an effective treatment and, monstly, the prognosis is poor and death occurs within one year from the diagnosis. Finally, poorly differentiated (PDC) thyroid carcinomas are intermediate forms between

well-differentiated and undifferentiated, morphologically and behaviourally (9,12, 13).

Thyroid neoplasms, therefore, represent a good model for studying the events involved in epithelial cell multistep carcinogenesis (Figure 2), because they comprise a broad spectrum of lesions with different degrees of malignancy from benign adenomas (FA) (not invasive and very well differentiated) to the ATC (very aggressive and always fatal), with the intermediate forms represented by PTC and FTC (differentiated and with a good prognosis) (1, 14, 15).

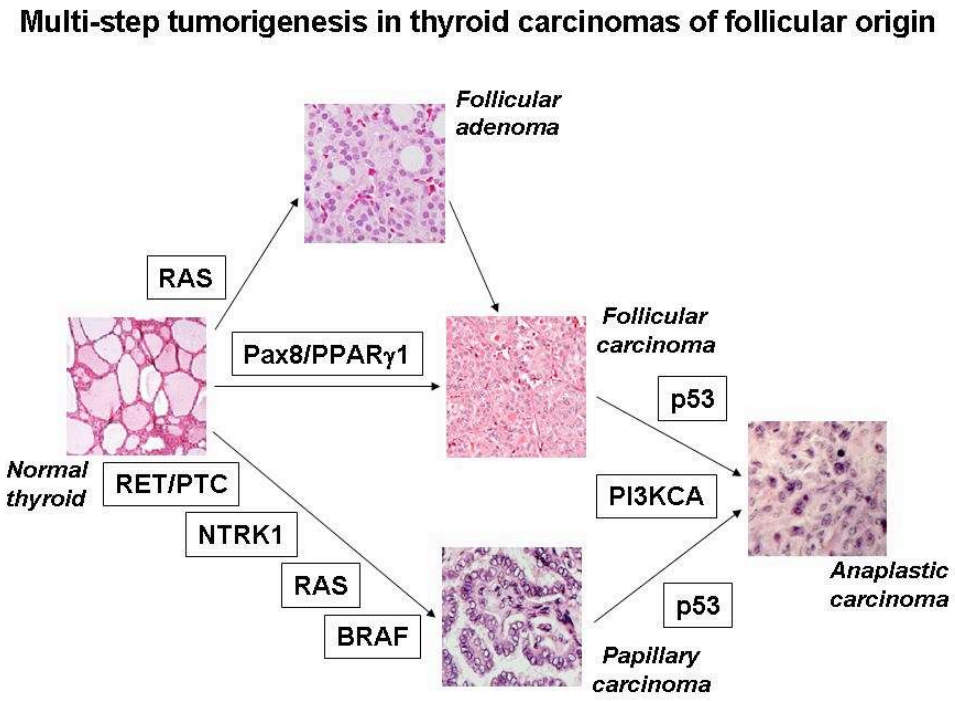


Figure 2 Model of multi-step carcinogenesis of thyroid neoplasms
On the basis of clinical, histological and molecular observations, three distinct pathways are proposed for neoplastic proliferation of thyroid follicular cells, including hyper-functioning follicular thyroid adenoma (tumours that are almost always benign lesions without a propensity for progression), follicular thyroid carcinoma and papillary thyroid carcinoma. Genetic defects that result in activation of RET or BRAF represent early, frequent initiating events that can be associated with radiation exposure. Most poorly differentiated and undifferentiated thyroid carcinomas are considered to derive from pre-existing well-differentiated thyroid carcinomas through additional genetic events, including p53 inactivation, but de novo occurrence might also occur.

Point mutations, rearrangements that activate proto-oncogenes, and loss-of-function in tumor suppressor genes have been identified in thyroid tumors (1). The involvement of several oncogenes has been demonstrated in PTC: activation of the RET/PTC oncogene, caused by rearrangements of the RET proto-oncogene, occurs in about 30% of cases (1), whereas the B-RAF gene is mutated in about 40% of cases (16). These tumors have also been associated with TRK gene rearrangements (17) and MET gene overexpression (18). RAS gene mutations (19) and PAX8-PPAR-γ rearrangements (20) are frequent in FTC, whereas impaired function of the p53 tumor suppressor gene is typical of ATC (21-23). Other genes have been implicated in thyroid neoplasias. Even though critical molecular mechanisms of thyroid carcinogenesis have been clarified, other molecular steps of neoplastic progression need to be investigated.

1.3 Microarray analysis of human carcinomas

DNA microarray technology was introduced in the scientific community in the mid-1990s to check the expression levels of thousands of genes simultaneously (24, 25). This kind of analysis was quickly adopted by the research community for the study of a wide range of biologic processes. Most of the early studies were performed simply to compare two biologic classes in order to identify the difference in gene expression, particularly of genes with potential relevance to a wide range of biological processes, such as the progression of cancer (26-29), the causes of asthma (30-32), heart diseases (33-35), neuropsychiatric disorders (36-38), and the analysis of factors associated with infertility (38-39).

Soon after its introduction, microarray technology was used to find new subclasses in disease states (40-41) and to identify biomarkers associated with disease (42), establishing that gene expression patterns could be used to distinguish subclasses of disease (28, 43-45).

This concept resulted in a proliferation of searches for patterns of expression that could be used to classify types of tumors (46) and to predict their outcome (47-48) and the response to therapy (48-49). In fact, gene expression profiling studies based upon this new technology has been successfully used to discover consistent gene expression patterns associated with a histological or clinical phenotype (48, 50-56) and also to predict the clinical outcome and survival as well as to classify different types of cancer (47, 48, 54-57). The research focus has now shifted toward identifying genetic determinants that are components of the specific regulatory pathways altered in cancers, leading to the discovery of novel therapeutic targets. In this view, our group, recently, investigated the Ras-dependent modulation of gene expression in thyroid cells (58). However, it is not easy to select few candidate genes for further studies from the lengthy gene lists generated from gene expression profiling studies (59). This approach, however, promises to provide diagnostic and prognostic markers that can be clinically used in the near future (60).

Therefore, to identify candidate genes involved in the process of thyroid carcinogenesis, we analysed a microarray, that represents 12.625 sequences, with RNAs extracted from normal human thyroid primary cell culture (NTPC), and six human thyroid carcinoma cell lines of different histotype (1 from a follicular carcinoma, 3 derived from papillary carcinomas and 2 from anaplastic

carcinomas). We focused our interest on the genes whose expression was drastically (at least 10 fold) up- or down-regulated in all of the six thyroid carcinoma cell lines in comparison to the normal thyroid primary cell culture assuming that the genes altered in their expression in all of the thyroid carcinoma cell lines might represent candidate genes involved in thyroid cell transformation.

1.4 CBX7 and Polycomb group genes

Within the context of this microarray study, we found that the CBX7 (Chromobox homolog 7) gene was down-regulated in the thyroid carcinoma-derived cell lines and also in thyroid fresh tumors. CBX7 encodes a novel protein that is a part of the Polycomb Group protein (61).

Polycomb Group genes (PcG) are a set of genes linking histone modifications with transcriptional repression: they are structurally diverse but functionally epistatic (62, 63). PcG proteins are part of two complexes with high molecular weight called Polycomb repressive complexes 1 and 2, PRC1 and PRC2 (Figure 3). (63, 64).

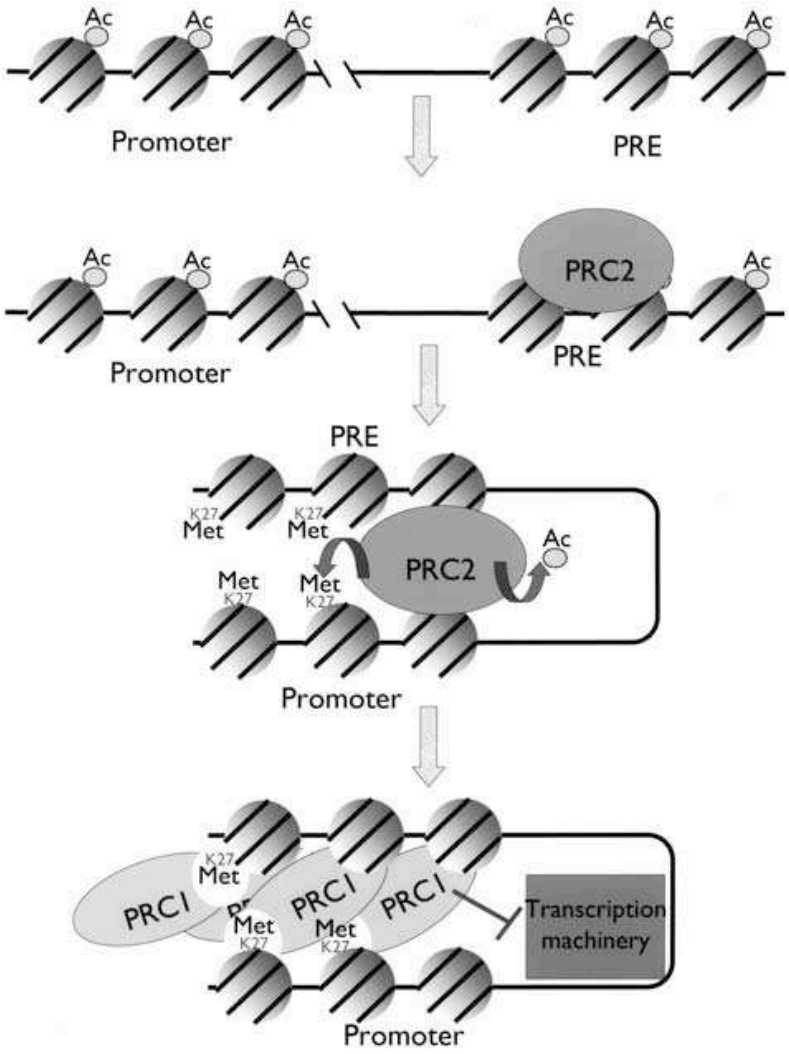


Figure 3 Mechanism of transcriptional repression by Polycomb Group proteins
The binding of Polycomb repressive complex 2 (PRC2) to Polycomb responsive elements (PRE) triggers histone deacetylation, and methylation of lysine 27 in Histone H3, both around the PRE and in the promoter region of the genes silenced. This modification constitutes one of the signals for the recruitment of the Polycomb repressive complex 1 (PRC1), or maintenance complex, which represses gene expression by inhibiting transcription initiation (Gil et al. 2005).

The PRC1 complex can recognize and bind to trimethylated K27 of Histone H3 through Polycomb proteins, and contributes to gene repression by inhibiting transcription initiation (65). The components of the core of this complex in mammals are the homologs of Drosophila Polycomb (Pc), Posterior sex combs (Psc), Sex combs extra (Sce) and Polyhomeiotic (Ph). The PRC2

complex, instead, has associated histone deacetylase and histone methyl transferase (specific for lysine of Histone H3, K27H3) activities; thus, it contributes to establish an histone repressive marks (66, 67). Among the known mammalian components of this second complex are Enhancer of Zeste (EZH2), Early embryonic deficient (EED), Suppressor of Zeste (SUZ12) and other associated proteins.

It is well known that PcG proteins deregulation is involved in tumorigenesis. In fact several studies have demonstrated that, for example, Bmi-1 is amplified in blood malignancies and medulloblastomas (68-71); EZH2 is involved in blood malignancies (68, 72), being its expression increased in prostate and breast tumors metastasis (73, 74); Mel18 has been suggested to act as a tumor suppressor gene (75); Rae28/Ph1 alterations are observed in acute lymphoblastic leukemia (76) and SUZ12 is up-regulated in colon tumors (77). Evidence is rapidly increasing that PcG genes are a novel class of oncogenes and anti-oncogenes, which may in future years become central to the development of novel cancer therapies based on epigenetic gene silencing (78, 79). Not only PcG genes such as BMI-1 and EZH2 are capable of cellular transformation, but also they are vital for cell survival. To date, abnormal PcG genes expression has been described in most human cancers. Also, the correlation between PcG expression and biological behavior of clinically defined cancer subtypes suggests that these genes play a central role in oncogenesis, and holds a promise for discovery of novel diagnostic markers (80).

CBX7 and the others PcG proteins contain a chromodomain in the N-terminus (CBX7 between amino acids 10 and 46). This domain was originally

identified in *Drosophila melanogaster* as a 37-amino-acid region of homology shared by heterochromatin protein 1 (HP1) and Polycomb (Pc) proteins (81). Phylogenetic and sequence analyses of the chromodomain (Figure 4A and B) revealed that CBX7 has a great similarity to other known or putative Pc proteins like CBX2 (Pc1), CBX4 (Pc2), CBX6 and CBX8 (Pc3). This similarity is less pronounced respect to the HP1 proteins, such as CBX1 (HP1β), CBX3 (HP1γ), CBX5 (HP1α) and, moreover, CBX7 does not contain the "chromo-shadow domain" that, instead, is a hallmark of HP1 proteins. CBX7 protein contains also other specific conserved residues that distinguish Pc and HP1 chromodomains: these residues specific of the Pc proteins are necessary both for Pc dimerization and recognition of tri-methylated Lys 27 on histone H3 (82).

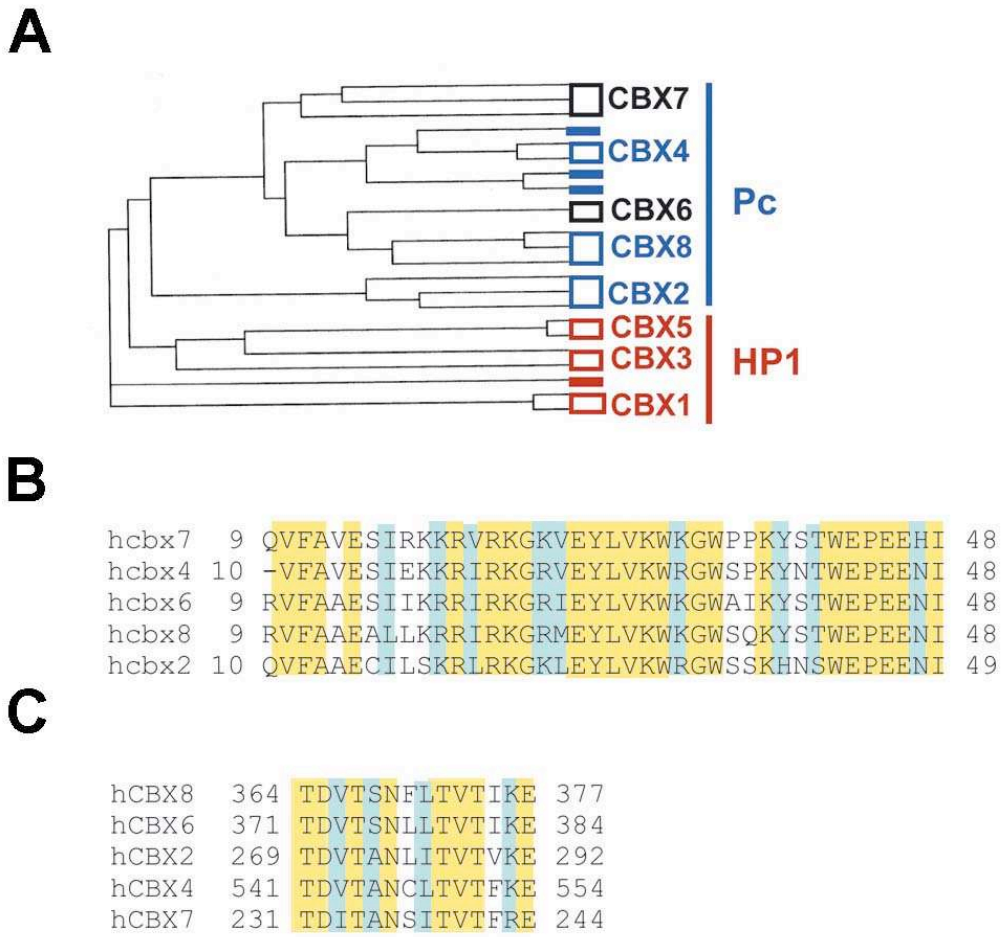


Figure 4 CBX7 and chromodomain
Identification of CBX7 as a Polycomb protein. (A) Phylogenetic analysis of Chromobox (CBX) proteins. Lines grouped by open rectangles denote orthologs of a same CBX protein (name in the right). Full rectangles refer to proteins that cannot be considered purely orthologs of any of the CBX proteins. Pc proteins are shown in blue, and HP1 proteins in red. CBX7 and CBX6 are shown in black. (B) Alignment of CBX7 chromodomain with that of other known and putative human Polycombs. Identical residues are shown in a yellow background, and similar residues are shown in a cyan background. (C) Alignment of CBX7 Polycomb box with that of other known and putative human Polycombs. Background colours are as described in (B) (from Gil et al. 2004).

In addition to the high homology within the chromodomain, CBX7 has also homology to Pc proteins in a carboxy-terminal region (Figure 4C) previously defined as the Pc box or C-box (83). It is known that the Pc box (amino acids 231–243) and the chromodomain (83) are both necessary for CBX7 function. It was seen that a fusion protein between CBX7 and the DNA-binding domain of

Gal4 is able to repress, in a dose dependent manner, the transcription over a 4xGal4-tk-luc reporter, suggesting that CBX7 functions by repressing transcription (61). So, CBX7 seems to act like the others Pc proteins exerting its effect through transcriptional repression.

Mouse CBX7 is, moreover, able to associate with facultative heterochromatin and with the inactive X chromosome, which indicates that CBX7 is really involved in the repression of gene transcription (84). CBX7 is also capable to interact with different Pc group members. In fact, similarly to other Pc proteins, CBX7 is able to inter-functions with itself and with the ring-finger protein Ring1 (61), as reported by Glutathione S-transferase (GST) pull-down experiments (85, 86). In addition, within the nucleus, CBX7 co-localizes with Ring1 to distinct foci-like structures termed Pc-bodies (87), in several cell lines, as demonstrated with immunofluorescence microscopy studies. However, there is not association between CBX7 and other PcG proteins, such as Bmi1, EED or EHZ2 (61) suggesting that CBX7 is part of different complexes than Bmi1 and Pc2.

Mouse CBX7 displays strong affinity for both H3K9me3 and H3K27me3 and is developmentally regulated in its association with chromatin(84). It has also been demostrated that chromodomain of CBX7 is essential for the binding to H3K9me3 and H3K27me3, in some cases associated with repressive state.

It has also been found that, *in vitro*, the chromodomain of CBX7 can bind RNA and that, *in vivo*, the interaction of CBX7 with chromatin, and the inactive X chromosome in particular, partly depends on its association with RNA (84). These data suggest that the capacity of this mouse Polycomb homolog to associate with

the inactive X chromosome, or any other region of chromatin, depends not only on its chromodomain but also on the combination of histone modifications and RNA molecules present at its target sites (84).

CBX7 is highly expressed in several tissues, including brain, kidney, heart and skeletal muscle. (61). A recent study has demonstrated that CBX7 expression is associated with extension of cellular life span in mouse embryonic fibroblasts and human prostate primary epithelial cells by downregulating expression of the Ink4a/Arf locus (61, 88). However, another study on ependymoma (89) reported a consistent down-regulation of CBX7 in the tumor samples, with the loss of at least one copy of the gene in 55% of the cases examined. This study also proposed other mechanisms than CBX7/p16 pathway, since in ependymoma, although gene mutations are rare, hypermethylation of p16Ink4a, Rb and p14ARF has been reported in a subgroup of tumors (4%–32%) (89).

1.5 Epithelial–mesenchymal transition

Epithelial–mesenchymal transition is an indispensable mechanism during morphogenesis which allows tissues and organs formation (90). Epithelial and mesenchymal cells differ in various functional and phenotypic characteristics. Epithelial cells grow in layers and communicate through specialized membrane structures, such as tight junctions, adherens junctions, desmosomes and gap junctions. Epithelial cells are motile and can move away from their nearest neighbours, while remaining within the epithelial layer (91). Mesenchymal cells, on the other hand, do not form an organized cell layer, nor do they have the same apical–basolateral organization and polarization of the cell-surface molecules and the actin cytoskeleton as epithelial cells.

Epithelial cells can convert into mesenchymal cells by a process known as the epithelial–mesenchymal transition. The term EMT describes a series of events during which epithelial cells loose many of their epithelial characteristics and take on properties that are typical of mesenchymal cells, which require complex changes in cell architecture and behaviour. The precise spectrum of changes that occur during EMT is probably determined by the integration of extracellular signals the cell receives, although this is still unclear.

The reverse process, known as mesenchymal–epithelial transition (MET), has also been reported, as shown in figure 5 (92-95).

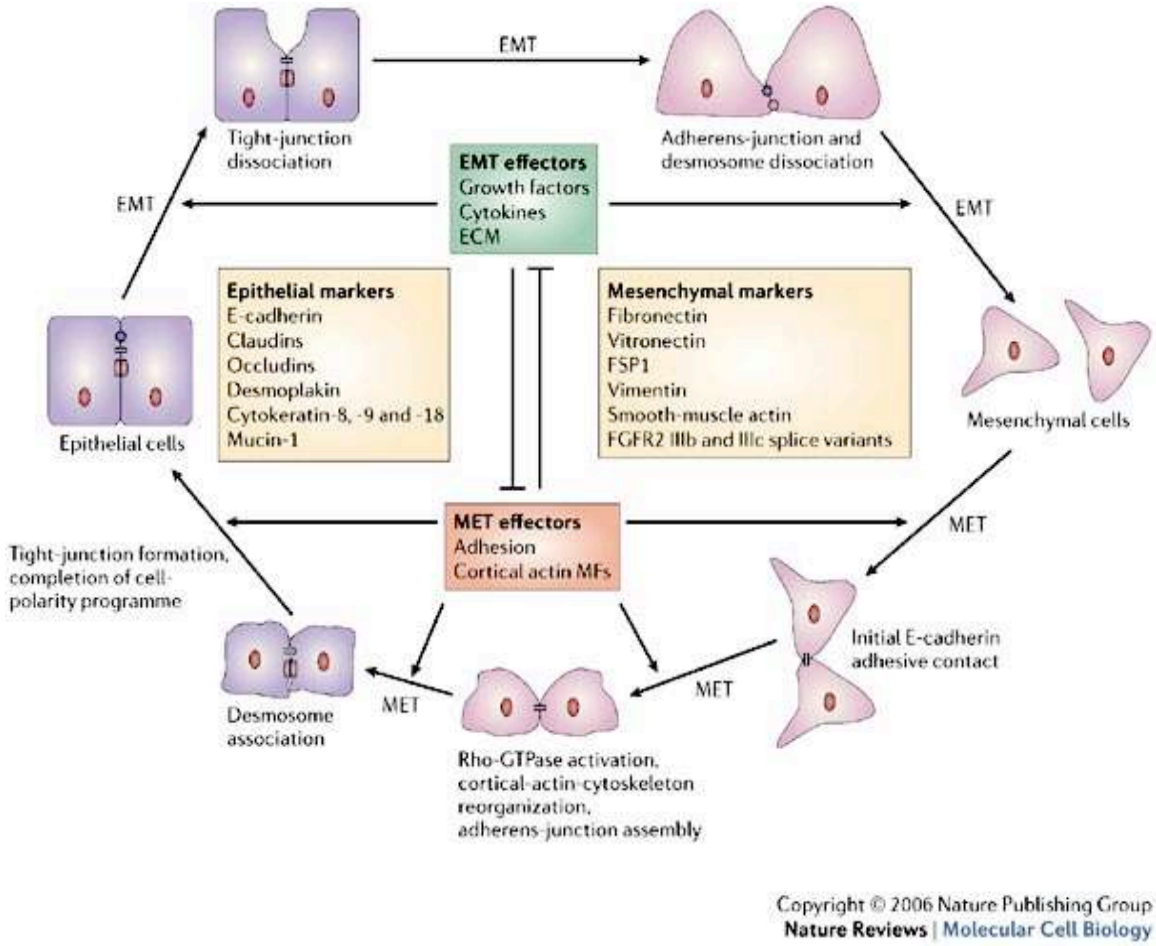


Figure 5 The cycle of epithelial-cell plasticity
The diagram shows the cycle of events during which epithelial cells are transformed into mesenchymal cells and vice versa. The different stages during EMT (epithelial–mesenchymal transition) and the reverse process MET (mesenchymal–epithelial transition) are regulated by effectors of EMT and MET, which influence each other. Important events during the progression of EMT and MET, including the regulation of the tight junctions and the adherens junctions, are indicated.

Without epithelial–mesenchymal transitions, in which polarized epithelial cells are converted into motile cells, multicellular organisms would be incapable of getting past the blastula stage of embryonic development (96). However, this important developmental programme has a more sinister role in tumour progression. Epithelial–mesenchymal transition provides a new basis for understanding cancer progression towards dedifferentiated and more malignant states (Figure 6).

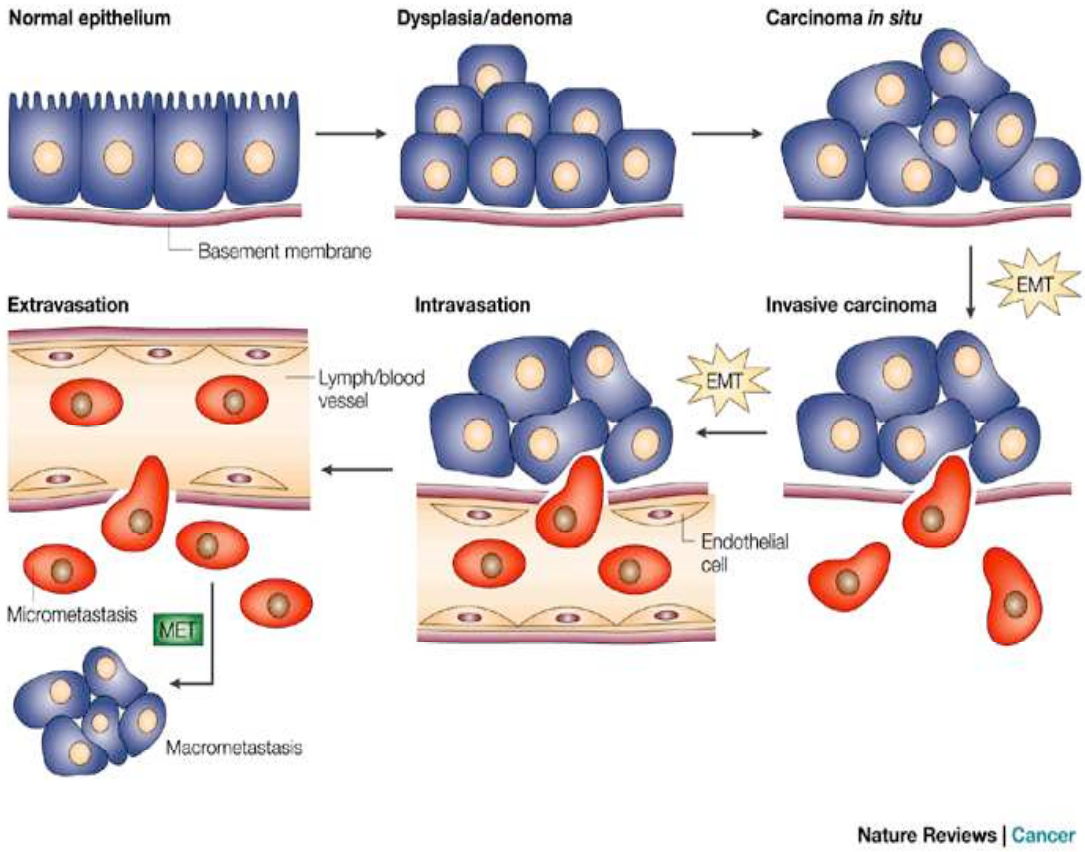


Figure 6 Sites of EMT and MET in the emergence and progression of carcinoma
Normal epithelia lined by a basement membrane can proliferate locally to give rise to an adenoma. Further transformation by epigenetic changes and genetic alterations leads to a carcinoma in situ, still outlined by an intact basement membrane. Further alterations can induce local dissemination of carcinoma cells, possibly through an epithelial—mesenchymal transition (EMT), and the basement membrane becomes fragmented. The cells can intravasate into lymph or blood vessels, allowing their passive transport to distant organs. At secondary sites, solitary carcinoma cells can extravasate and either remain solitary (micrometastasis) or they can form a new carcinoma through a mesenchymal—epithelial transition (MET).

Studies show that several extracellular activators can trigger EMT, that extensive crosstalk exists between the signalling pathways that activate and repress EMT, and that EMT-inducing signalling pathways have many common endpoints, including downregulation of E-cadherin expression and expression of EMT-associated genes (90).

Many effectors orchestrate the disassembly of junctional complexes and the changes in cytoskeletal organization that occur during EMT. The activation of signalling pathways also results in the activation of transcriptional regulators such as snail (now known as SNAI1)(97) and slug (now known as SNAI2), which regulate the changes in gene-expression patterns that underlie EMT.

A central target of these transcriptional regulators is the repression of the E-cadherin gene, an important caretaker of the epithelial phenotype. Downregulation of E-cadherin (or mutations that occur in cancer) has several important consequences that are of direct relevance to EMT. E-cadherin levels become limiting, which results in the loss of E-cadherin-dependent intercellular epithelial junctional complexes, and E-cadherin-mediated sequestering of β-catenin in the cytoplasm is abolished, as shown in figure 7 (90).

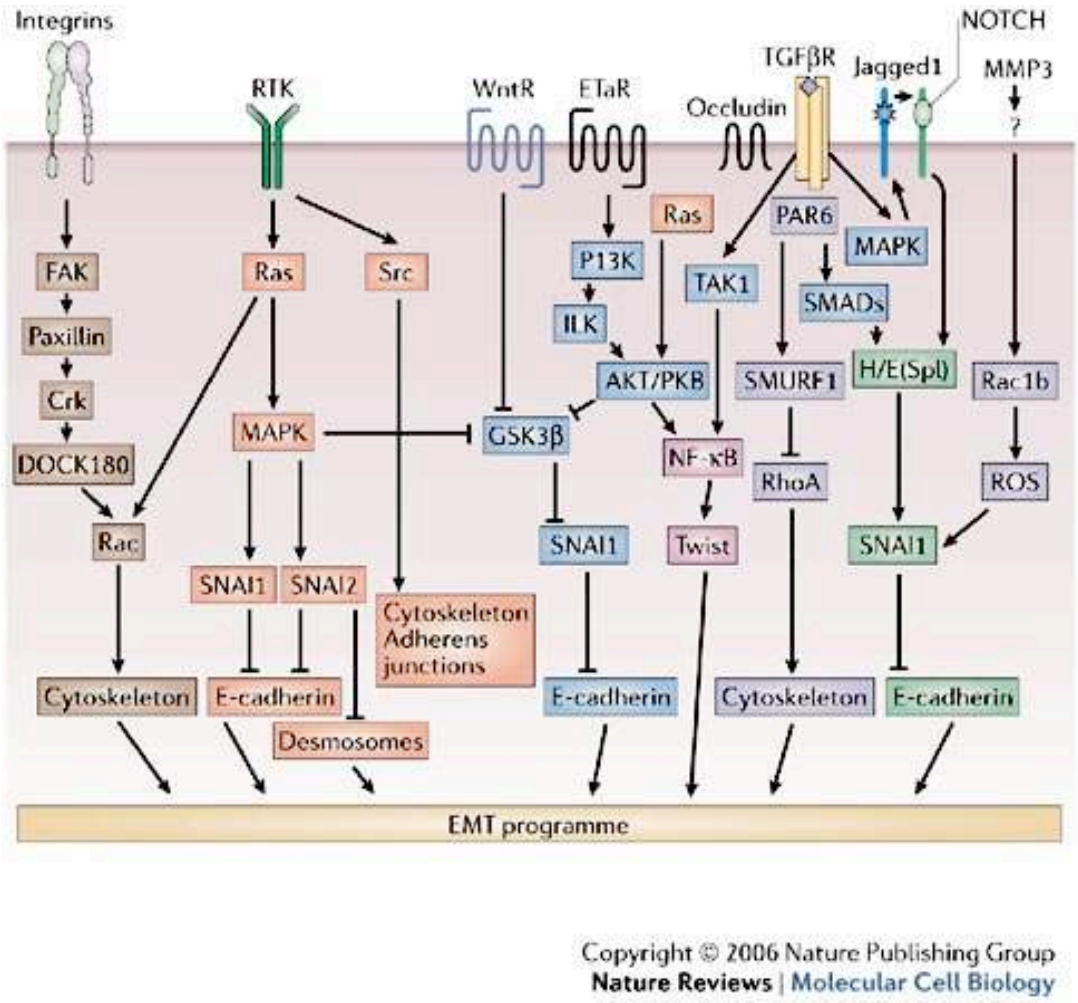


Figure 7 Overview of the molecular networks that regulate EMT
A selection of the signalling pathways that are activated by regulators of EMT and a limited representation of their crosstalk is illustrated. Activation of receptor tyrosine kinases (RTKs) is known to induce EMT in several epithelial cell types and *in vivo*, but it is now clear that the EMT process often requires co-activation of integrin receptors. The role of transforming growth factor-β (TGF) signalling in EMT is established for a limited number of normal and transformed cell lines, whereas in vivo data has indicated a mutual regulation of the TGF and NOTCH pathways during EMT. There is now increasing evidence that other signalling pathways could have an important role in EMT, including those elicited by G-protein-coupled receptors. Matrix metalloproteinases (MMPs) can also trigger EMT through as-yet-undefined receptors.

It took a long time for EMT to be recognized as a potential mechanism for carcinoma progression (96).

During progression to metastatic competence, carcinoma cells acquire mesenchymal gene-expression patterns and properties. This results in changed adhesive properties, and the activation of proteolysis and motility, which allows

the tumour cells to metastasize and establish secondary tumours at distant sites (98). In tissue culture this progression is accompanied by partial or complete EMT, and induction of EMT in many carcinoma cell lines results in the acquisition of metastatic properties *in vivo*. Infact, it is striking that the same signalling pathways that regulate developmental EMT are also activated during tumour progression.

Local invasion can be considered an initial and essential step in the malignancy of carcinomas, leading to the generation of usually fatal distant metastasis. Tumour invasion appears to be controlled by a coordinated series of cellular and molecular processes that enable tumour cells to dissociate and migrate from the primary tumor.

One of the hallmarks of EMT is the functional loss of E-cadherin (encoded by CDH1), which is currently thought to be a suppressor of invasion during carcinoma progression (96).

E-cadherin is therefore emerging as one of the caretakers of the epithelial phenotype. E-cadherin is the prototypic type I cadherin. Type I cadherins mediate homophilic interactions by forming adhesive bonds between one or several immunoglobulin (Ig) domains in their extracellular region and connecting to actin microfilaments indirectly via α-catenin and β-catenin in the cytoplasm.

So how is E-cadherin downregulated? In most cases, downregulation during carcinoma progression occurs by epigenetic mechanisms, including transcriptional repression and promoter hypermethylation. Occasionally, the E-cadherin gene is mutated, leading to the absence or to the expression of a non-functional protein(99-106).

1.6 E-cadherin

The cadherin superfamily of Ca^{2+}-dependent homophilic adhesion molecules plays a critical role in regulating cell-to-cell interactions. Alterations in cadherin expression or function occur frequently during carcinogenesis, such as the loss of the epithelial cadherin (E-cadherin) and/or the aberrant expression of other cadherins. Indeed, the aberrant expression of cadherins has been detected during carcinoma invasion, a process which is reminiscent of the epithelial-mesenchymal transition (EMT) so important in many critical developmental processes.

The different cadherins constitute a superfamily of transmembrane receptors that mediate Ca^{2+}-dependent cell-to-cell homophilic interactions. In this way, 5 different sub-families can be considered (Figure 8): 1) Classical cadherins (type I), mainly localised to adherens junctions; 2) Highly related type II cadherins; 3) Desmosomal cadherins (desmocollins and desmogleins) that form desmosomal junctions; 4) Protocadherins, mainly implicated in neural development; 5) Cadherin-related proteins, like the Flamingo and Fat-like cadherins (107).

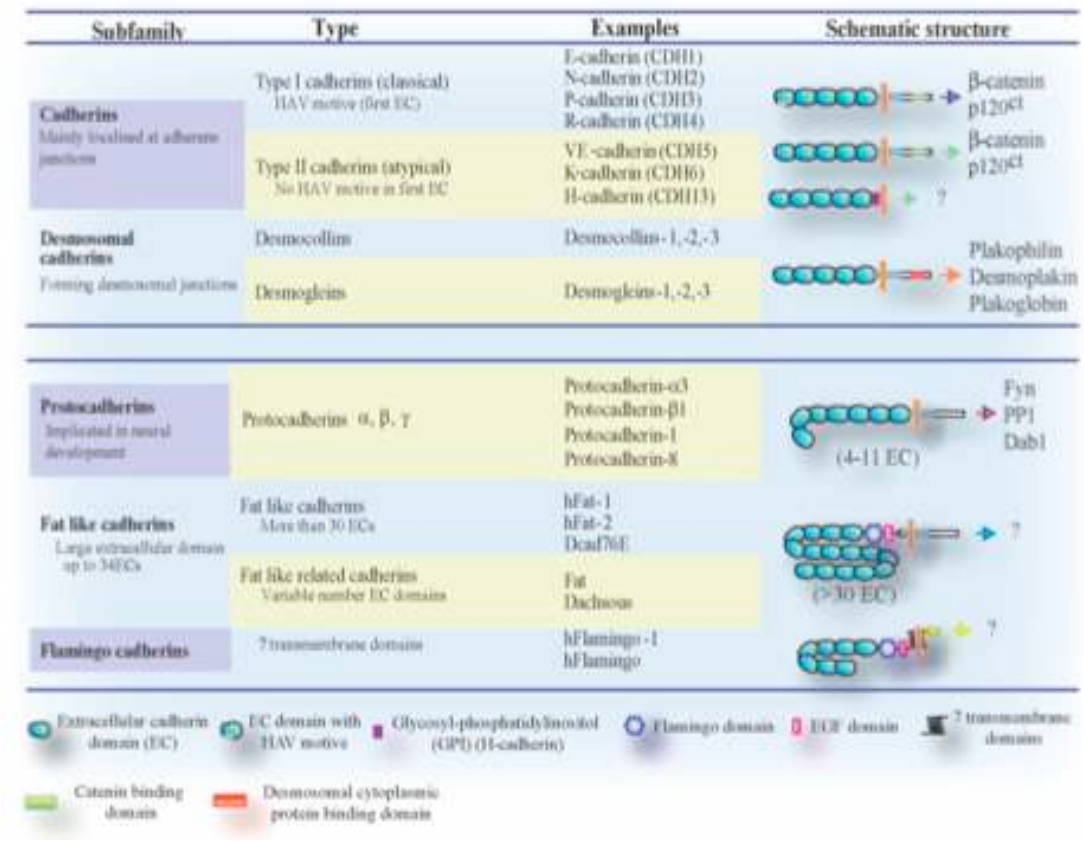


Figure 8 Organization of the cadherin superfamily
The main structural and functional characteristics of the cadherin subfamilies and types are indicated. Representative examples are shown, with gene denomination of some of them in brackets. The main structural domains and schematic structure of the different cadherin types are shown at the right.

Expression of E-cadherin is maintained in all epithelial tissues, but it is silenced during the process of EMT and in established mesenchymal cells (108, 109).

During the last decade, the expression of cadherins during tumour progression has been the subject of extensive studies, the majority of which have focused on the role of E-cadherin. The loss of E-cadherin expression and/or function has been observed during tumour progression of most carcinomas (110, 111), and this has been related to the induction of EMT which frequently occurs

during carcinoma invasion (96, 112). E-cadherin can be considered as an invasion suppressor gene (111).

Mechanisms of E-cadherin regulation during tumor progression

Most of our current knowledge regarding the way in which cadherin expression is controlled during tumour progression has derived from studies on E-cadherin. Indeed, the involvement of Ecadherin in epithelial morphogenesis and homeostasis, and its proven anti-invasive role in carcinoma progression has fomented interest in understanding the regulatory mechanisms that control E-cadherin expression under normal and pathological circumstances.

The human E-cadherin gene (CDH1) is located at the 16q22.1 locus and is comprised of 16 exons, spanning 99 Kb of genomic DNA. The mouse E-cadherin gene (cdh1) has a similar exon-intron organisation and is located on chromosome 8. Loss of E-cadherin function is thought to be one of the most important events during carcinoma progression (111), an assumption that is reinforced by the fact that germ line mutations in CDH1 induce a predisposition to diffuse-type gastric cancer (113). During tumour progression, E-cadherin can be functionally inactivated by different mechanisms, including somatic mutation and down-regulation of gene expression through promoter methylation and/or transcriptional repression. Indeed, down-regulation of the CDH1 gene expression is observed in many carcinomas (111), while inactivating mutations, frequent in diffuse gastric carcinomas and in lobular breast carcinomas, are rarely observed in other types of tumours (113, 114). Indeed, the most frequent genetic alterations of CDH1 in tumours involve exon-skipping and out-of-frame mutations, and in most cases

mutations occur in combination with a loss of heterozygosity of the wild type allele (114).

Epigenetic mechanisms of E-cadherin silencing

The hypermethylation of CpG islands is associated with the recruitment of methyl DNA binding proteins (MBDs) and of histone deacetylase activity (HDACs), which together contribute to the compaction of the DNA in the promoter region and hence, to gene inactivation. Of particular interest was the finding that in diffuse-type gastric cancer, methylation of the E-cadherin promoter serves as a second hit to somatic or germ-line genetic alterations in inactivating E-cadherin expression (115). Indeed, it has also been proposed that methylation of E cadherin CpG islands can act as a second hit for transcriptional expression to silence E-cadherin in ductal breast carcinoma (116).

On the other hand, methylation of the E-cadherin promoter is not always correlated with E-cadherin silencing (113), indicating that additional genetic or epigenetic modifications might be required for E-cadherin down-regulation. The molecular mechanisms involved in regulating E-cadherin methylation are starting to be unravelled. Recently, the association of several MBDs, such as MeCP2, MBD1 and MBD2, with the E-cadherin promoter have been reported in several cancer cell lines (117). Interestingly, methylation of E-cadherin CpG islands and MeCP2 expression have been shown to cooperatively and epigenetically regulate E-cadherin expression in colorectal cancer cells (118). It is also possible that other epigenetic mechanisms, like histone deacetylation, participate in regulating E-cadherin expression. Histone H3 deacetylation is associated with CpG

methylation-mediated gene silencing in other contexts and recently, the participation of both epigenetic modifications was reported for E-cadherin down-regulation in several cancer cell lines (117).

Transcriptional regulation of E-cadherin silencing

Besides regulation of E-cadherin by promoter hypermethylation and/or genetic alterations, direct transcriptional control of E-cadherin has emerged in the last years as an important regulatory mechanism of E-cadherin expression.

The mouse E-cadherin promoter was first isolated in 1991. Its initial characterization showed that it was a TATA-less promoter containing several potential proximal regulatory elements, including a CCAAT box (-65), a GC-rich region (-30 to -58) and a palindromic element (-70 to -90) composed of two adjacent E-boxes flanked by four inverted nucleotides called E-pal (120). The proximal CCAAT and GC-rich regions are required for basal E-cadherin expression and are recognized by constitutive AP2 and Sp1 transcription factors, and CAAT-binding proteins, respectively (119, 103, 120) (figure 9). The E-pal element was initially described as an epithelial-specific regulator (119), but subsequent studies showed it to be an active repressor in E-cadherin deficient cells (103, 120, 105).

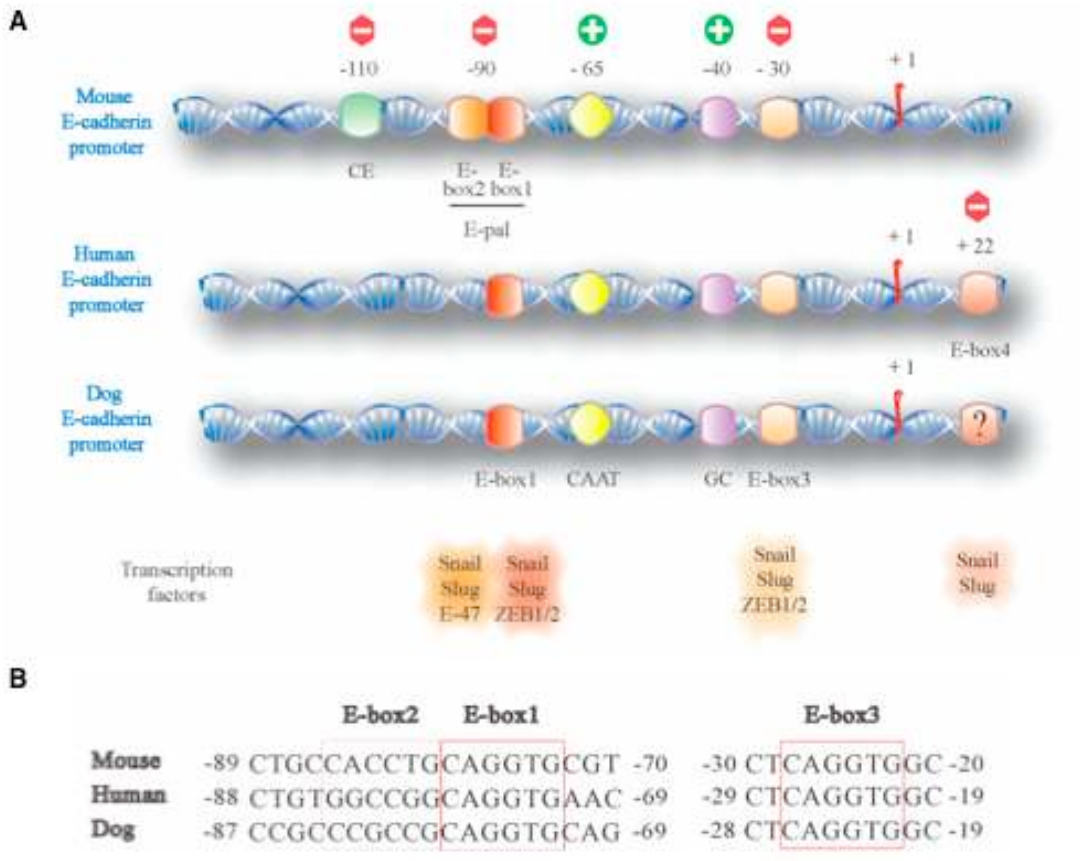


Figure 9 Schematic representation of the E-cadherin promoter region and E-box sequences in mouse, human and dog
(A) The E-cadherin promoter region is represented with its representative proximal control elements in the proximal region (exerting either a positive or negative effect on E-cadherin expression). The CAAT, GC and E-box1/E-box3 are conserved in all three promoters, while the E-box 2 is only present in the mouse promoter region. An additional E-box (E-box 4) downstream of the transcription initiation site is present in the human promoter, but is not conserved in the region encompassing the mouse promoter. Binding of identified factors to the different E-boxes is shown in the lower part. (B) Sequence of the E-boxes present in the proximal region of the mouse, human and dog E-cadherin promoter. Observe the high degree of conservation of E-box 1 and E-box 3.

Several studies have demonstrated that the E-boxes in the proximal E-cadherin promoter repress its expression. Point mutations abolishing the two E-boxes in the mouse E-pal, or E-box 1 and 3 in the human promoter, produced a strong induction of E-cadherin promoter activity in different carcinoma cell lines deficient in E-cadherin expression. Evidence has also been presented that the E-

box 4 also represses E-cadherin expression (121, 122), hence, factors able to bind to the E-boxes and repress E-cadherin expression must be present in these cells.

Characterization of E-cadherin repressors

The specificity of the factors identified was confirmed in a parallel screen with an E-pal element in which a point mutation abolished the repressor function (103, 120) and which should also preclude the binding of regulatory factors. More than 90% of the identified factors that specifically bound to the E-boxes corresponded to transcription factors, of which, two in particular were highly represented: 49% corresponded to the zinc finger factor Snail and 32% to the class I bHLH factors, E47 (123 , 124). The remaining factors corresponded mainly to an additional class I bHLH factor, mITF2 (also called E2-2), while a single clone corresponded to another member of the Snail family, Slug (125). Functional characterisation of Snail indicated that it does indeed act as a strong repressor of the mouse and human E-cadherin promoters (121, 123).

In all these instances, these repressors silence E-cadherin by binding to the proximal E-boxes, although some differences are observed between the mouse and human promoters. While the factors that act on the mouse promoter seem to preferentially interact with the E-pal element (including Snail, Slug and E47) (123, 124, 125), in the human promoter either all three E-boxes or E-box 1 and 3 appear to be required for the interaction of Snail/Slug or δEF1/SIP1, respectively (121, 122, 126, 127). These observations suggest that at least two of the three proximal E-boxes in the mammalian E-cadherin promoter are functionally similar in terms of recruiting repressors, regardless of their relative location. Interestingly,

all of the E-cadherin repressors identified exhibit expression patterns in developing embryos compatible with this function.

Although still preliminary, these data indicate that the different repressors may participate in silencing E-cadherin specifically in different types of tumour or, perhaps, at defined stages of tumour progression. A more detailed analysis of the expression of these repressors in large-scale tumour samples is clearly needed to study this hypothesis, an aim that is presently hindered by the lack of good specific antibodies to the different repressors.

Mechanism of E-cadherin repression: potential co-operation with epigenetic modifications

The information available regarding the different mechanisms involved in E-cadherin silencing make it difficult to define a simple model where E-cadherin expression is regulated by just a single genetic, epigenetic or transcriptional control mechanism. It seems more likely that a combination of different mechanisms is responsible for defining the status of E-cadherin expression during tumour progression. As such, a modification of the classical two hit hypothesis for E-cadherin silencing has recently been proposed by Cheng et al., in which a collaboration between genetic, epigenetic and transcriptional control of E-cadherin expression is at play in ductal breast cancer (116). In this system, Snail expression is correlated with the silencing of E-cadherin and hypermethylation of its promoter, rather than by methylation itself. Indeed, one interpretation of this model might invoke an interesting role for Snail in the co-ordination of both processes (116). The specific mechanisms by which Snail, and/or other repressors,

mediates E-cadherin silencing are still largely unknown. The modification of chromatin by the co-ordinated action of DNA and/ or histone methylation, and acetylation, has emerged as one of the major mechanisms for regulating the transcriptional activity of different regulatory genes. Indeed, several years ago the modification of chromatin was proposed as a mechanisms to regulate E-cadherin expression, and the participation of histone deacetylase activities (HDACs) has been used to explain the repressor activity of human Slug on artificial promoters (128). Very recently, details of the mechanism underlying Snail repression of E-cadherin were revealed, providing a link between transcriptional control and the epigenetic modifications of the E-cadherin promoter (107). Snail repression of the E-cadherin promoter involves the recruitment of a repressor complex formed, at least, by the co-repressor mSin3A, HDAC1 and HDAC2 (107) (Fig. 10). The recruitment of this complex is mediated by the N-terminal SNAG domain of Snail, previously thought to act as the repressor domain. The presence of this complex results in a net decrease in the amount of acetylated histones H3/H4 and an increase in methylated K9 of histone H3 in the endogenous E-cadherin promoter. In turn, this leads to a compact organisation of the chromatin (107). A preliminary analysis in the mouse skin carcinogenesis model indicates that Snail expression is also associated with the hypermethylation and silencing of the E-cadherin promoter, further supporting the notion of a connection between Snail and epigenetic modifications (Figure 10). The molecular mechanisms involved in this link remain to be elucidated, but the implication of MBDs and HDACs is an interesting aspect of E-cadherin silencing for further study, since treatment with

inhibitors of DNA methylation and HDACs promotes the reexpression of E-cadherin.(107)

In summary, current evidence indicates that silencing of E-cadherin transcription requires the participation of several repressor factors that interact with specific E-boxes in the proximal promoter. The repression of gene expression either involves the local modification of chromatin organisation through the recruitment of specific co-repressor complexes (Snail, Slug and potentially δEF1/SIP1), or heterodimerisation of repressors with specific partners (E47). These mechanisms of repression add an additional level of complexity to our understanding of the different factors that specifically repress E-cadherin expression. Thus, as well as the specific expression patterns and/or binding affinities for similar regulatory elements, repressor function might finally be regulated by the presence of, and/or affinity for, different co-repressors in specific cells or tissues. Thus, the regulation of expression and function of the different repressors needs also to be taken into account to obtain a comprehensive view of the transcriptional control of E-cadherin. In this sense, a recent study has elegantly shown that silencing of Snail expression can be triggered by the ligated estrogen receptor in breast cancer cell lines, involving the specific participation of the MTA3 protein in an additional repressor complex (129).

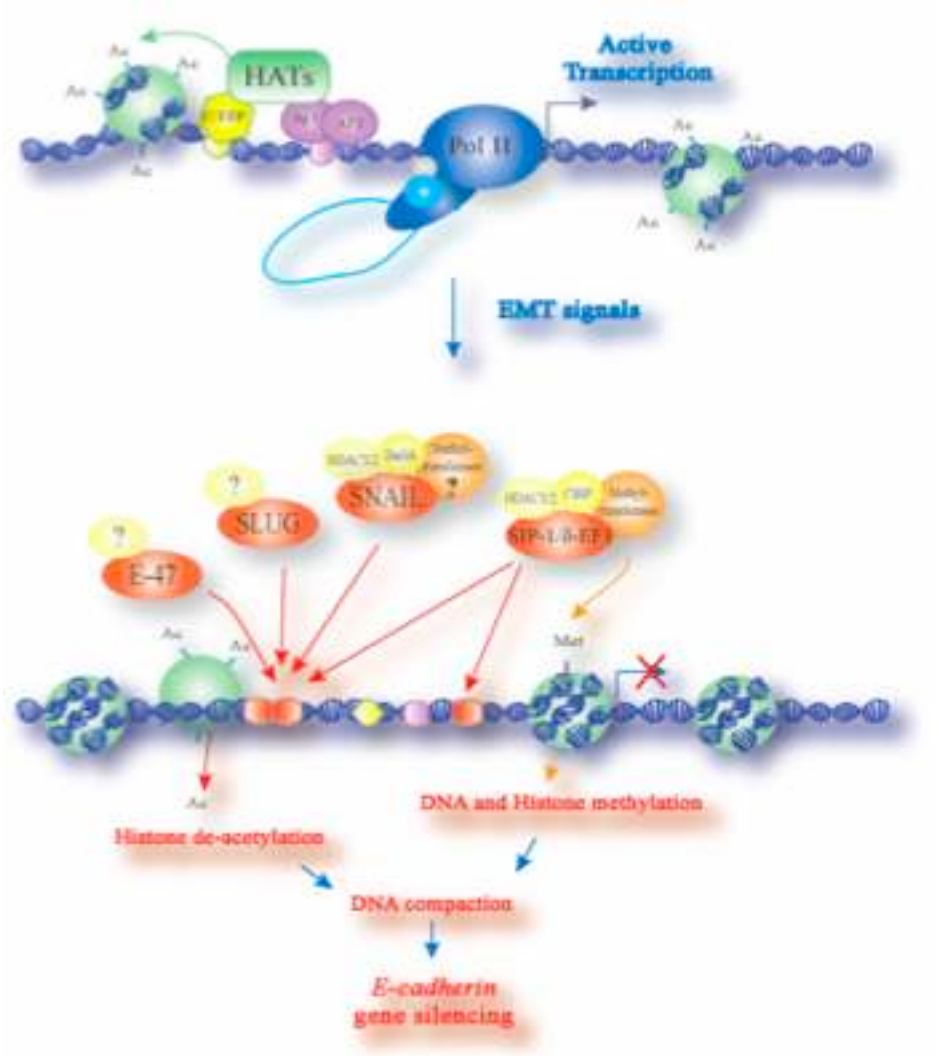


Figura 10 A model for the regulation of E-cadherin expression involving different repressors and epigenetic mechanisms

A schematic illustration of the putative epigenetic mechanisms involved in E-cadherin regulation is shown. In this model, active transcription is probably achieved through the interaction of several activators, and involves the acetylation of histone residues which promotes the chromatin to adopt an open conformation. Under the influence of the appropriate signals, several repressors and co-repressor molecules are recruited (depending on either the tissue or cellular context), attracting the repressor machinery which includes corepressor molecules, HDACs and methylases, and still unidentified molecules which will modify the acetylation and methylation status of chromatin, promoting E-cadherin silencing (see text for details).

2 AIMS OF THE STUDY

Thyroid tumors are the result of the accumulation of different modifications in critical genes involved in the control of cell proliferation. Although various therapeutic approaches are followed in clinical practice, most of them are not life-saving. Hence, the discovery of new approach to diagnose cancer at an early stage and to establish more effective therapies is a critical and urgent issue. To achieve this goal, identification and characterization of key molecules that participate in carcinogenesis are essential steps.

In this context, we studied the modulation of gene expression associated with the progression towards a malignant phenotype of thyroid cancer. To this aim, we used RNA samples, extracted from human thyroid carcinoma cell lines, to perform a screening with Affymetrix microarrays containing about 10.500 human genes.

Through this approach we have identified the gene: CBX7, a chromobox family protein which is down-regulated in thyroid carcinoma cells.

To gain insight into the role of CBX7 in malignant transformation we have decided to perform a functional proteomic assay, because the association of CBX7 with partners involved in a particular mechanism is strongly suggestive of the biological function of the protein (130).

Thus, the main aim of the study was the identification and characterization of proteins that interact with CBX7 and might contribute to the transformation process of the thyroid follicular cells

3 MATERIALS AND METHODS

3.1 Cell culture and transfections

In this study we used the following human thyroid carcinoma cell lines: TPC-1, WRO, NPA, ARO, FRO, NIM 1, B-CPAP, FB-1, FB-2, Kat-4 and Kat-18, which are described elsewhere (131). They were grown in DMEM (Gibco Laboratories, Carlsbad, CA) containing 10% fetal calf serum (Gibco Laboratories), glutamine 10mM (Gibco Laboratories) and ampicillin/streptomycin 100 μg/ml (Gibco Laboratories). Normal human thyroid primary culture cells have been established and grown as already described (132).

We have also used 293 Human embryonal kidney (HEK), maintained in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10% fetal calf serum, L-glutamine, and penicillin/streptomycin.

HeLa was maintained respectively in RPMI medium supplemented with 10% fetal calf serum, glutamine, ampicillin/streptomycin. All cell lines were maintained at 37°C under 5% CO_2 atmosphere.

Trichostatin A (TSA) (Sigma) was dissolved in ethanol and added to the culture medium at 300 nM. A corresponding volume of ethanol was added to control, untreated cells. TSA treatments were performed for 24 h after transfection.

Cells were transfected using Lipofectamine reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions.

3.2 Human thyroid tissue samples

Neoplastic human thyroid tissues and normal adjacent tissue or the controlateral normal thyroid lobe were obtained from surgical specimens and immediately frozen in liquid nitrogen. Thyroid tumors were collected at the Laboratoire d'Histologie et de Cytologie, Centre Hospitalier Lyon Sud, France, and at the Laboratoire d'Anatomie Pathologique, Hopital de L'Antiquaille, Lyon, France. The tumor samples were stored frozen until RNA or protein extractions were performed.

3.3 RNA isolation

Total RNA was extracted from tissues and cell cultures using the RNAeasy mini kit (Qiagen, Valencia, CA), according to the manufacturer's instructions. The integrity of the RNA was assessed by denaturing agarose gel electrophoresis.

3.4 Microarray analysis

In this study we used commercially available high density Affymetrix (Santa Clara, CA) oligonucleotide microarrays (HG_U95Av2) consisting of 12.625 probe set each representing a transcript. cRNA preparation, hybridization, scanning and analysis were performed as previously reported (133). Briefly, after cRNA hybridization to microarray, analysis of the readings from the quantitative scanning was performed using the Affymetrix Gene Expression Analysis Software (Affymetrix), according to Affymetrix protocols. Comparisons were made for each carcinoma cell line sample versus normal thyroid primary culture,

taking normal thyroid cells as baseline. The fold change values, indicating the relative change in the expression levels between the transformed cell line and the normal thyroid primary culture sample, were used to identify genes differentially expressed between these conditions.

3.5 Reverse transcription

1 μg of total RNA from each sample was reverse-transcribed with QuantiTect® Reverse Transcription Kit (Qiagen) using an optimized blend of oligo-dT and random primers according to the manufacturer's instructions.

3.6 PCR analysis

Semiquantitative RT-PCR

PCR was carried out on cDNA using the GeneAmp PCR System 9600 (Applied Biosystems, Foster City, CA). The RNA PCR Core Kit (Applied Biosystems) was used to perform amplifications. After a first denaturing step (94°C for 3 min), PCR amplification was performed for 25-30 (variable) cycles (94°C for 30 sec, 55-60°C (variable) for 30 sec, 72°C for 30 sec).

The sequences of primers used are shown in Table 1. To ensure that RNA samples were not contaminated with DNA, negative controls were obtained by carrying out the PCR on samples that were not reverse-transcribed, but otherwise identically processed. The PCR products were separated on a 2% agarose gel, stained with ethidium bromide and scanned with a Typhoon 9200 scanner (GE Healthcare, Piscataway, NJ).

Selection of primers and probes for qRT-PCR

To design a qRT-PCR assay we used the Human ProbeLibraryTM system (Exiqon, Vedbaek, Denmark). Briefly, using locked nucleic acid (LNA^{TM}) technology (134, 135), Exiqon provides 90 human prevalidated TaqMan probes of only 8-9 nucleotides that recognize 99% of human transcripts in the RefSeq database at NCBI (136, 137). Using the ProbeFinder assay design software (freely accessed on the web site www.probelibrary.com) we chose the best probe and primers pair. To amplify a fragment of CBX7 mRNA by Real-Time PCR, we entered its accession number (NM_175709.2) on the assay design page of the ProbeFinder software and we chose an amplicon of 71 nucleotides that spanned the 3rd and 4th exons. The probe number was "human 63" (according to the numbering of Exiqon's Human ProbeLibrary kit) and the primer sequences are reported in Table 1.

The same procedure was used to choose the probe and primers for the housekeeping gene G6PD (accession number X03674). We opted for an amplicon of 106 nucleotides that spanned the 3rd and 4th exons. The probe number was "human 05" (according to the numbering of Exiqon's Human ProbeLibrary kit) and the primer sequences are reported in Table 1. All fluorogenic probes were dual-labeled with FAM at 5'-end and with a black quencher at the 3'-end.

qRT-PCR

Real-Time Quantitative TaqMan PCR was carried out with the Chromo4 Detector (MJ Research, Waltham, MA) in 96-well plates using a final volume of 20 μl. For PCR we used 8 μl of 2.5x RealMasterMixTM Probe ROX (Eppendorf

AG, Hamburg, Germany), 200 nM of each primer, 100 nM probe and cDNA generated from 50 ng of total RNA. The conditions used for PCR were 2 min at 95°C, 45 cycles of 20 sec at 95°C and 1 min a 60°C. Each reaction was carried out in duplicate. Fold mRNA overexpression was calculated according to the formula $2^{(Rt-Et)}/2^{(Rn-En)}$ as described previously (138), where Rt is the threshold cycle number for the reference gene in the tumor, Et for the experimental gene in the tumor, Rn for the reference gene in the normal sample and En for the experimental gene in the normal sample. The primers used are reported in table 1.

Symbol	Acc.Num	Primer F	Primer R
KIF2C	U63743	CAGAACAAGGCTCACACC	AGCAGGCTTCCATCTCTT
CXCL1	X54489	CACCTGGATTGTGCCTAA	TGCAGGCTCCTCAGAAAT
PHLDA2	AF035444	GGCACGACATGAAATCC	GTGGTGACGATGGTGAAGT
CDC20	U05340	CCGTTACATTCCTTCCCT	AGTTGCCCTCTTTGATCC
PIB5PA	U45975	TGGAGGTGGCAGATGAGT	GGATGCTGTGGTTGTGAC
TFF3	L08044	ATGTCACCCCCAAGGA	TGGCAGCAATCACAGC
BTG2	U72649	ACCTCAACCTGGGGAACT	CCAAACGTCTCCCACTTC
RARRES3	AF060228	GCTATCGGGTCAACAACA	CCTAATCGCAAAAGAGCA
CBX7	NM_175709	GTCATGGCCTACGAGGAGAA	CTTGGGTTTCGGACCTCTCT
β-ACTIN	NM_001101	TCGTGCGTGACATTAAGGAG	GTCAGGCAGCCTCGTAGCTCT
CBX7 qRT-PCR	NM_175709	CGTCATGGCCTACGAGGA	TGGGTTTCGGACCTCTCTT
G6PD qRT-PCR	NM_000402	ACAGAGTGAGCCCTTCTTCAA	GGAGGCTGCATCATCGTACT

Table 1 List of primers

3.7 Immunostaining: technique, evaluation and statistical analysis

The cell distribution of the CBX7 protein was assessed by immunostaining formalin-fixed, paraffin-embedded thyroid tumor blocks retrieved from the files of the Dipartimento di Scienze Biomorfologiche e Funzionali at the University of Naples Federico II and selected to represent a wide range of thyroid neoplastic diseases. To ensure that we evaluated CBX7 expression only on technically adequate slides, we discarded slides that lacked a convincing internal control,

namely labeling of stromal, endothelial or lymphoid cells, shown to be positive in a preliminary normal tissue microarray analysis (data not shown). Based on these criteria, we scored paraffin-embedded stained slides from 20 cases of FA, 30 cases of classical PTC, 6 cases of TCV PTC, 32 cases of FTC, 12 cases of poorly differentiated carcinoma, and 12 cases of ATC. As controls, we selected areas of normal thyroid parenchyma from the lobe contralateral to the tumor in 20 surgical specimens of PTC.

Briefly, xylene-dewaxed and alcohol-rehydrated paraffin sections were placed in Coplin jars filled with a 0.01 M tri-sodium citrate solution, and heated for 3 min in a conventional pressure cooker (131). After heating, slides were thoroughly rinsed in cool running water for 5 min. They were washed in Tris-buffered saline pH 7.4 and then incubated overnight with the specific rabbit polyclonal primary antibody. Subsequently, tissue sections were stained with biotinylated anti-rabbit immunoglobulins, and then with peroxidase-labelled streptavidine (Dako, Carpinteria, CA). The signal was developed by using diaminobenzidine chromogen as substrate. Incubations both omitting and pre-adsorbing the specific antibody were used as negative controls.

3.8 Protein extraction

For total cell extract (TCE) preparation, cells were harvested 48 hours after the onset of transfection. Cells were then lysed in RIPA buffer (20 mM Tris-HCl pH 7.5, 5 mM EDTA, 150 mM NaCl, 1% Nonidet P40, and a mix of protease inhibitors), and clarified by ultracentrifugation at 13000 rpm at 4°C for 30 min.

For nuclear extract (NE), cells were washed twice in phosphate-buffered saline and resuspended in 3 volumes of a solution containing 10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl2, 0.1 mM EGTA, 0.5 mM DTT (homogenization solution). The cells were disrupted by passage through a 26-gauge needle. Nuclei were collected by centrifugation at 1500 rpm. and resuspended in a 1.2-volume of extraction solution containing 10 mM HEPES pH 7.9, 0.4 M NaCl, 1.5 mM MgCl2, 0.1 mM EGTA, 0.5 mM DTT, 5% glycerol, to allow elution of nuclear proteins by gentle shaking at 4°C. Nuclei were pelleted again by centrifugation at 12 000 rpm. and the supernatant was stored at -70°C until used. The protease inhibitors leupeptin (5 mM), aprotinin (1.5 mM), phenylethylsulfunilfluoride (2 mM), peptastatin A (3 mM), and benzamidine (1 mM) were added to both homogenization and extraction solutions. Protein concentration was determined by the Bradford protein assay (BioRad).

For Rnase A treatment protein estracts were incubated with 10μg/ml RNase A for 10 min at 37°C.

3.9 Western blotting and immunoprecipitation assay

Protein extracts and immunoprecipitated pellets were separated by SDS-PAGE, and then transferred onto Immobilon-P Transfer membranes (Millipore). Membranes were blocked with 5% non-fat milk proteins and incubated with antibody (Ab) at the appropriate dilutions. The filters were incubated with horseradish peroxidase-conjugated secondary antibody and the signals were detected by Western blotting detection system (ECL) (GE Healthcare). To ascertain that equal amounts of protein were loaded, the western blots were incubated with antibodies against the α-tubulin protein (Sigma). Membranes were then incubated with the horseradish peroxidase-conjugated secondary antibody (1:3000) for 60 min (at room temperature) and the reaction was detected by ECL.

For immunoprecipitation (IP) experiments, TCE were incubated with protein A-sepharose or G-sepharose beads (Amersham) for 1 h at 4 °C, then samples were centrifugated to 2000 rpm to eliminate beads and incubated overnight with antidoby. Then the IP were incubated with protein A-sepharose or G-sepharose beads for 1 h at 4 °C. The beads were collected and washed five times with lysis buffer, and boiled in Laemmli sample buffer for immunoblotting analysis. The Abs used for immunoprecipitation and Western blotting were: anti-CBX7 protein (Neosystem), anti-V5 (Invitrogen), anti-HDAC2 (Upstate), anti-HA (Roche), anti-MDB2 (Abcam), anti-RbAp46 (Abcam), anti-PRMT5 (BD transduction laboratories), anti-PRMT1 (Santa Cruz), anti-MTA3 (Novus Biologicals), anti-Tubulin (Santa Cruz Biotechnology Inc.) and anti-Vinculin (Santa Cruz Biotechnology Inc.).

3.10 Plasmid constructs

CBX7 expression plasmid was constructed by cloning the human cDNA sequence in a pcDNA3.1/V5-His-TOPO ® (Invitrogen). The primers used were: CBX7 forward 5'-ATGGAGCTGTCAGCCATC-3' and CBX7 reverse 5'-GAACTTCCCACTGCGGTC-3'. The expression of CBX7 was assessed by Western blotting. The same construct was also inserted in pCEFL-HA.

3.11 Samples preparation and immunoprecipitation for proteomic assay

For preparative gels, $8x10^7$ cells were transfected with 10 μg of empty vector or with CBX7-V5-His expression plasmid. For immunoprecipitation experiments, 1 mg of NE were incubated for 2 hours with agarose-beads saturated with aspecific antibody. Then the lisates were incubated with agarose anti-V5 antibodies over-night, previously saturated with milk (Abcam). The beads were collected and washed five times with lysis buffer, and the complex recovered by competition with peptide V5 (Abcam). Then the lisated were incubated again with agarose V5-beads and the complex again recovered by competition. For immunoblotting analysis, 50μg of TCE or immunoprecipitated proteins were separated by SDS-PAGE and transferred to Immobilon-P Transfer membranes (Millipore). Membranes were blocked with 5% nonfat milk proteins and incubated with the appropriate antibodies at the appropriate dilutions. The filters were incubated with horseradish peroxidase-conjugated secondary antibodies, and the revelations were performed by the ECL method (Amersham Biosciences).

3.12 Electrophoresis fractionation and in situ digestion

The immunoprecipitated protein mixture and the control were fractionated by SDS-PAGE on a 12% bis-acrylamide gel. The gel was stained by Colloidal Comassie (Pierce) and the corrisponding protein bands were excised from both lanes and digested *in situ* with trypsin(139). Peptide mixtures were extracted, freeze-dried and resuspended in 10 µl of 0.2% trifluoroacetic acid (TFA) for the Liquid Chromatography/Mass Spectrometry/Mass Spectrometr (LCMSMS) analysis

3.13 Liquid Chromatography/Mass Spectrometry/Mass Spectrometr (LCMSMS) analysis

Each peptide mixture was directly analyzed by LCMSMS using an Agilent ion trap mass spectrometer LC/MSD Trap XCT Ultra integrated with a new microfluidic chip-based technology for nanospray LC/MS (Agilent) and coupled on-line with a nano chromatography system Agilent 1100. After loading, the peptide mixture (8 µl) was first concentrated and washed at 4 µl/min onto a reverse-phase pre-column on the chip using 0.2% formic acid as eluent. The sample was then transferred and fractionated by reverse-phase chromatography onto a capillary column present on the chip at a flow rate of 300 nl/min using a linear gradient of eluent B (0.2% formic acid in 95% acetonitrile) in A (0.2% formic acid in 5% acetonitrile) from 7 to 60% in 50 min. The mass spectrometer was set up in a data-dependent MS/MS mode where a full scan spectrum (*m*/*z* acquisition range from 400 to 1600 Da/e) was followed by a tandem mass spectrum (*m*/*z* acquisition range until 1800 Da/e). The precursor ions were selected

as the most intense peaks of the previous scan. LC/MSD Trap software, provided by the manufacturers, was used to analyze raw MS and MS/MS spectra and to generate a peak list which was introduced in the MASCOT MS/MS ion search software for protein identification.

3.14 Transactivation assay

In the luciferase transactivation assay HEK 293 and HeLa cells were transiently transfected with the reporter construct (kindly provided by Dr ER Fearon, University of Michigan Medical School) in which the luciferase gene was driven by a fragment extending 1359 bp 5' and 125 bp 3'to the human E-cadherin transcription start site (104, 140) and normalized with the use of a co-transfected β-galactosidase construct. Luciferase activity was analysed by Dual-Light System (Applied Biosystems, Massachusetts, USA).

3.15 Expression and purification of recombinant proteins/GST pull-down experiments

GST fusion proteins were constructed by cloning the human cDNA sequence in a pGEX4T-1 (Promega). The primers used for GST-CBX7 were: CBX7 forward 5'-GAGCTGTCAGCCATCGGC-3' and CBX7 reverse 5'-GTCAGAACTTCCCACTGCGG-3'; for GST- CBX7 CHROMO: CBX7-CHROMO forward 5'-GAGCTGTCAGCCATCGGC-3' and CBX7 reverse 5'-TCACTTCTCCTTGCCCTTGGC-3';and for GST-CBX7 NOCHROMO: CBX7-NOCHROMO forward 5'-GCCAAGGGCAAGGAGAAGTGAC-3' and CBX7-NOCHROMO reverse 5'-GTCAGAACTTCCCACTGCGG-3'.

GST fusion proteins were produced in Escherichia coli BL21 cells. Stationary phase cultures of E. coli cells transformed with the plasmid of interest were diluted 5–400 ml in LB with ampicillin (100 mg/ml), grown at 30°C to an OD 600 of 0.6 and induced with 0.1mM IPTG. After an additional 2 h at 30°C, the cultures were harvested and resuspended in 10 ml of cold PBS (140 mM NaCl, 20mM sodium phosphate (pH 7.4)), 1mM phenylmethylsulfonyl fluoride (PMSF) and protease inhibitors (Boehringer). The cells were broken by French Press. For the GST proteins, the supernatant was then incubated at 4°C for 1 h with 250 μl of glutathione-Sepharose beads (Amersham Pharmacia Biotech). The expression of recombinant protein of CBX7 was assessed by Coomassie.

The recombinant proteins were eluted with a buffer containing PBS, 10mM reduced glutathione, and 10% (v/v) glycerol. For GST pull-down the fusion proteins were incubated with total cellular extracts (TCEs) in NETN buffer (20mM Tris-HCl, pH 8.0, 100mM NaCl, 1mM EDTA, and 0.5% Nonidet P-40) for 2 h at 4°C. The resins were then extensively washed in the same buffer and then resuspended in Laemmli buffer. The bound proteins were separated by SDS-PAGE and analysed by Western blotting. PRMT1 recombinant protein was provided by Upstate.

3.16 Electrophoretic mobility shift assay (EMSA) and supershift assay

For gel shift analysis, NE were prepared following the method of Dignam, et al (141). E-cadherin fragment were ^{32}P labeled (TransCruz™ Gel Shift Oligonucleotides) with ^{32}[P]-ATP to 50,000 cpm/ng by using polynucleotide kinase. Binding reaction mixtures (20 μl) were incubated for 20 minutes at room

temperature in 10 mM Tris pH7.5, 50 mM NaCl 1 mM dithiothreitol (DTT), 1 mM EDTA, 5% glycerol. Binding reaction contained DNA probe (50000 cpm), 5 µg nuclear extract and 1 µg poly dI-dC to inhibit non specific binding of the labeled probe to nuclear extract protein. The protein-DNA complexes were resolved by native PAGE (6% gel) in 0.5xTris/borate/EDTA and visualized by autoradiography.

For supershift analysis, assay were performed as described above with the exception that antibody (2 ug) was normally added 4 h prior to addition of labelled oligonucleotide probe.

3.17 HDAC activity assay

HEK 293, HeLa and NPA cells were transfected with increasing amount of CBX7 expression vector and were used to assay the HDAC activity. The total protein extract (30 µg) were incubated with 5 µl of [^{3}H]acetate-labeled histone H4 (1.8 nCi/µg) in 200 µl of activity buffer (25 mM Tris-HCl [pH 7.5], 10% glycerol, 1 mM EDTA, 125 mM NaCl) overnight at 37°C. The reaction was stopped by the addition of 50 µl of 1 N HCl-0.4 M acetate and the released [^{3}H]acetate was extracted with 600 µl of ethyl acetate(Upstate). After centrifugation, a 100 µl aliquot of the supernatant was counted in 5 ml of scintillation cocktail. All experiments were carried out three times, and samples were assayed in duplicate.

3.18 Chromatin immunoprecipitation (ChIP) and Re-ChIP assays

ChIP and Re-ChIP experiments were performed as reported elsewhere (142). Briefly, $5x10^6$ HEK 293 and HeLa cells were cross-linked using 1% formaldehyde and lysed. The pellets were then resuspended in 200 μl of lysis buffer (10mM EDTA, 50mM Tris-HCl pH 8.0, 1% SDS and protease inhibitors), and sonicated five times for 30 sec at maximum settings, obtaining fragments between 0.3 and 1.0 kb. The samples were cleared by centrifugation at 14 000 rpm for 15 min. After centrifugation, 20 μl of the supernatants were used as inputs, and the other part of the samples diluted 2.5-fold in Ip buffer (100mM NaCl, 2mM EDTA pH 8.0, 20mM Tris-HCl pH 8.0, 0.5% Triton X-100 and protease inhibitors).

The samples were subjected to immunoprecipitation with specific antibody (anti-V5 described in the previous section, and anti-HDAC2) after 2 h preclearing at 4 °C with Protein A Sepharose or Protein G Sepharose/BSA/Salmon Sperm (Upstate). Precipitated chromatin complexes were removed from the beads through 15 min incubation with 250 ml of 1% SDS, 0.1M $NaHCO_3$.

In Re-Chip experiments, complexes were eluted by incubation for 30 min at 37 °C in 250 μl of Re-Ip buffer and subjected again to the ChIp procedure. Crosslink was reversed by an overnight incubation at 65°C with 20 ml of 5M NaCl. Then were added 10 μl 0.5mM EDTA, 20 μl 1M Tris-HCl pH 6.5 and 20 μg of Proteinase K, and incubated for 1 h at 45°C. DNA was purified by Phenol/CHCl3, and precipitated by two volumes of ethanol in the presence of tRNA.

PCR was performed using specific primers:

pE-cadherin(P2)-L 5'-GAACTCAGCCAAGTGTAAAAGCC-3'

pE-cadherin(P2)-R 5'-GAGTCTGAACTGACTTCCGC-3'

p16(P1)-L 5'-GCAGTCCGACTCTCCAAAAG-3'

p16(P1)-R 5'-GGGTGTTTGGTGTCATAGGG-3'

pE2F1(P1)-L 5'-CGTTGGCTGTTGGAGATTTT-3'

pE2F1-R(P1)-R 5'-TTGCCTCACCCATGACATTA-3'

p1S100A4(P2)-L 5'-TCATCCAGTCCCCTGCTAGT-3'

p1S100A4(P2)-R 5'-AGAGCGGATACTGCCTTCCT-3'

pGAPDH(P2)-L 5'-TCCTGTTTCATCCAAGCGTG-3'

pGAPDH(P2)-R5'-GACTGTCGAACAGGAGGAG-3'

3.19 Ultracentrifugation of CBX7 transfected cells

Cells were vortexed at 4°C in ice-chilled lysis buffer. Lysis buffer L was derived from buffer A (10 mM HEPES [pH 7.9], 1.5 mM $MgCl_2$, 10 mM KCl, 200 mM NaCl, 0.2 mM EDTA) supplemented with 1 mM dithiothreitol, 40 U of RNasin (Amersham), protease inhibitor cocktail (P-8340; Sigma), and 0.5% Nonidet P-40. Cell lysates clarified by centrifugation for 30 min at 9,000 *g* at 4°C were loaded on top of 5 to 45% glycerol gradients in buffer A supplemented with 1 mM dithiothreitol and 10 U of RNasin (Amersham). The gradients were spun at 4°C for 16 h at 39 000 rpm in a Beckman SW41TI rotor. Fractions were collected from the top of the gradients. Gradient distribution was checked by Western blotting (143). RNAse A was provided by Fermentas.

3.20 *In vitro* methyltransferase assays

10 μg of CBX7-V5 construct were transfected by Lipofectamine (Invitrogen) into 100 mm diameter plate of HEK 293 cells. 1 μg of whole cell extracts was immunoprecipitated with the anti-V5. The immunoprecipitates were rinsed twice with Methylation buffer (50 mM Tris–HCl, pH 7.5, 1 mM EDTA and 1 mM EGTA) at room temperature, resuspended in 0.04 ml of Methylation buffer, and 5 μg of pooled histones (Sigma) and 2 ml (2.75 mCi) of [H3]Ado-Met(GE Healthcare) were added. Reactions were carried out at 30° for 30 min. Then 5 ml of the reaction were transferred onto the center of 2cm X 2cm P81 paper (Upstate). The square has been washed twice with 10% tricloroacetic for 15 minute and then with 95% ethanol for 5 minute. The square were dried and then transfer in a scintillation vial, added with scintillation cocktail. PRMT1 recombinant protein was provided by Upstate.

.

4 RESULTS AND DISCUSSION

4.1 Gene expression profile analysis of six thyroid carcinoma cell lines compared with normal human thyroid primary cells

To search for candidate genes involved in the neoplastic transformation of the thyroid gland, RNAs extracted from normal human thyroid primary cells and six human thyroid carcinoma cell lines of different origin (WRO cell line from a follicular carcinoma, TPC-1 and FB-2 cell lines, both deriving from papillary thyroid carcinoma, NPA cell line, which derives from a poorly differentiated papillary carcinoma, ARO and FRO cell lines originating from anaplastic carcinomas) were hybridized to an U95Av2 Affymetrix oligonucleotide arrays containing 12.625 sequences.

Among the transcripts represented on the array, 510 had a fold change higher than 10 and 320 lower than -10 in most of the cell lines versus normal thyroid cells (data not shown). To validate the results obtained by the microarray analysis we evaluated the expression of 50 transcripts by semiquantitative RT-PCR in a large number of thyroid carcinoma cell lines. The results essentially overlapped those achieved by microarray analysis, thus validating them. Some representative RT-PCR analyses are shown in Figure 11: in the panel A it is reported the expression of some genes, such as KIF2C, CXCL1, PHLDA2 and CDC20, that were up-regulated in thyroid carcinoma cell lines, whereas in panel B we show the expression of same of the genes that were down-regulated in the neoplastic cell lines, such as PIB5PA, TFF3, BTG2 and RARRES3.

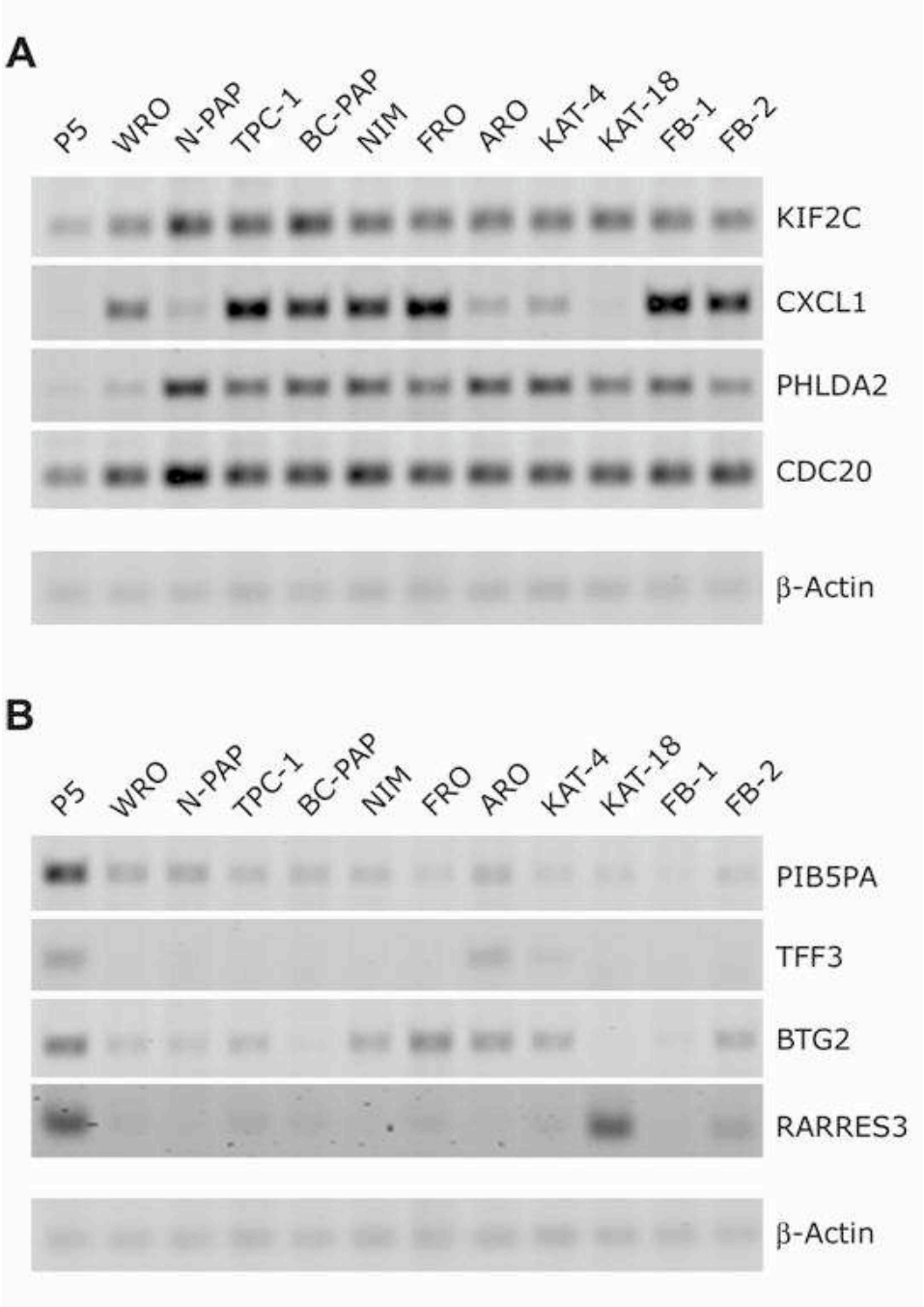


Figure 11 Validation of microarray data by semiquantitative RT-PCR
Semiquantitative RT-PCR of some up-regulated genes (A) in a panel of 11 thyroid carcinoma cell lines versus the normal thyroid cell line (P5) and some down-regulated genes (B) resulting by the microarray analysis. In both cases human β-Actin gene expression was evaluated as control to normalize the amount of RNA.

We have assumed that the up- or down-regulation of some genes in all of the thyroid carcinoma cell lines indicate the presence of a general event linked to thyroid cell transformation, independently from the genetic mechanisms underlying this process. Therefore we focused on the genes up- or down-regulated in all the carcinoma cell lines with a higher fold change. Some of them were further analysed by other techniques.

4.2 CBX7 gene expression is down-regulated in human thyroid carcinoma cell lines

Among the down-regulated genes, which represent potential tumor suppression, CBX7 was drastically down-regulated in all thyroid cancer cell lines tested with the cDNA microarray (Figure 12A). This result was confirmed by RT-PCR analysis in a large panel of thyroid carcinoma cell lines with normal thyroid primary culture cells as control (Figure 12B).

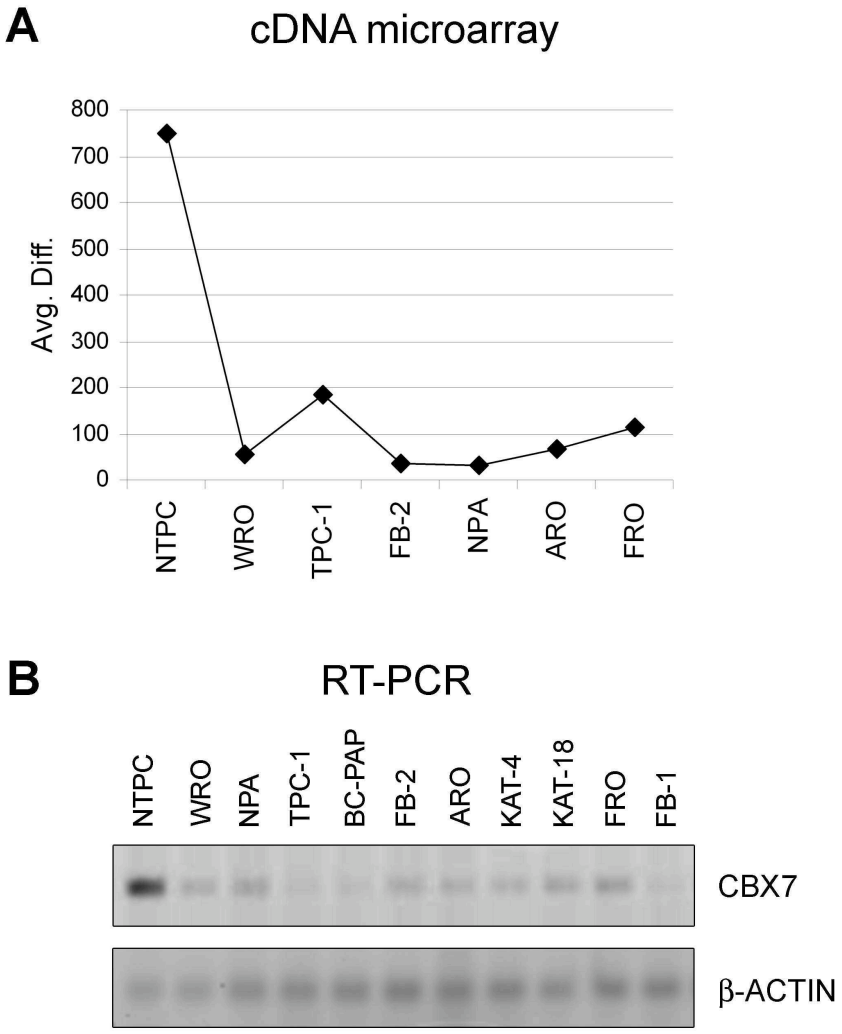


Figure 12 CBX7 expression in human thyroid carcinoma cell lines
(A) CBX7 gene expression by microarray analysis in human thyroid carcinoma cell lines vs normal human thyroid primary culture cells (NTPC). Avg. Diff, average difference: it is a quantitative relative indicator of a transcript expression level (∑(PM-MM)/pairs on average). (B) CBX7 gene expression analysis by RT-PCR in human thyroid carcinoma cell lines vs the normal human thyroid primary culture cells (NTPC). β-Actin gene expression served as loading control.

4.3 The loss of CBX7 expression correlates with a more aggressive phenotype in thyroid carcinomas

To determine whether the loss of CBX7 expression is a feature of thyroid tumors and not only of cultured thyroid carcinoma cell lines, we carried out an

immunohistochemical analysis (IHC) of paraffin-embedded tissues using polyclonal antibodies raised against the carboxy-terminal region of human CBX7 protein. As shown in Table 2, all 20 samples of normal thyroid parenchyma expressed CBX7 at a very high level, which coincided with the strong CBX7 staining in all follicles (Figure 13A).

		CBX7 expression [a]			
Histotype	*Cases*	*High*		*Low*	
Normal Thyroid Parenchyma	n=20	100%	(20/20)	0%	(0/20)
Follicular Adenomas	n=20	100%	(20/20)	0%	(0/20)
Follicular Carcinomas	n=32	34%	(11/32)	66%	(21/32)
Papillary Carcinomas - Classic Variant	n=30	43%	(13/30)	57%	(17/30)
Papillary Carcinomas - Tall Cell Variant	n=6	17%	(1/6)	83%	(5/6)
Poorly Differentiated Carcinomas	n=12	17%	(2/12)	83%	(10/12)
Anaplastic Carcinomas	n=12	0%	(0/12)	100%	(12/12)

Table 2 CBX7 expression in normal and neoplastic thyroid tissues by IHC
[a]High: cases with more than 50% of neoplastic cells showing CBX7 staining; Low: cases with less than 50% of neoplastic cells showing CBX7 staining.

The intensity of nuclear labeling of epithelial thyroid cells in follicular adenoma (Figure 13B) was similar to that of the internal control stromal and endothelial cells (Figure 13G), that express normally CBX7. Conversely, CBX7 expression was reduced in malignant lesions. As shown in Table 2, and in Figure 13C and D, the percent of low expressors was high in well-differentiated tumors, namely FTC (66%; 21/32 samples) and PTC (57%; 17/30 samples). It was even higher in the less differentiated tumors, namely PDC (83%; 10/12) (Table 2, and Figure 13F and G) and “tall cell variant”(TCV)-PTC (83%; 5/6) (Table 2, and Figure 13E). In the latter, neoplastic cells were almost devoid of CBX7 expression, which sharply contrasted with the intense staining of the infiltrating lymphocytes and stromal cells. Similarly, CBX7 expression was completely lost

in all cases of ATC (100%; 12/12) (Table 2, Figure 13H). No staining was observed when normal thyroid gland samples were stained with antibodies pre-incubated with CBX7 recombinant protein (Figure 13I) or in the absence of the primary antibodies (data not shown). Therefore, CBX7 expression was intense in normal thyroid and in benign neoplastic lesions, decreased in well-differentiated carcinomas and drastically reduced in aggressive thyroid tumors.

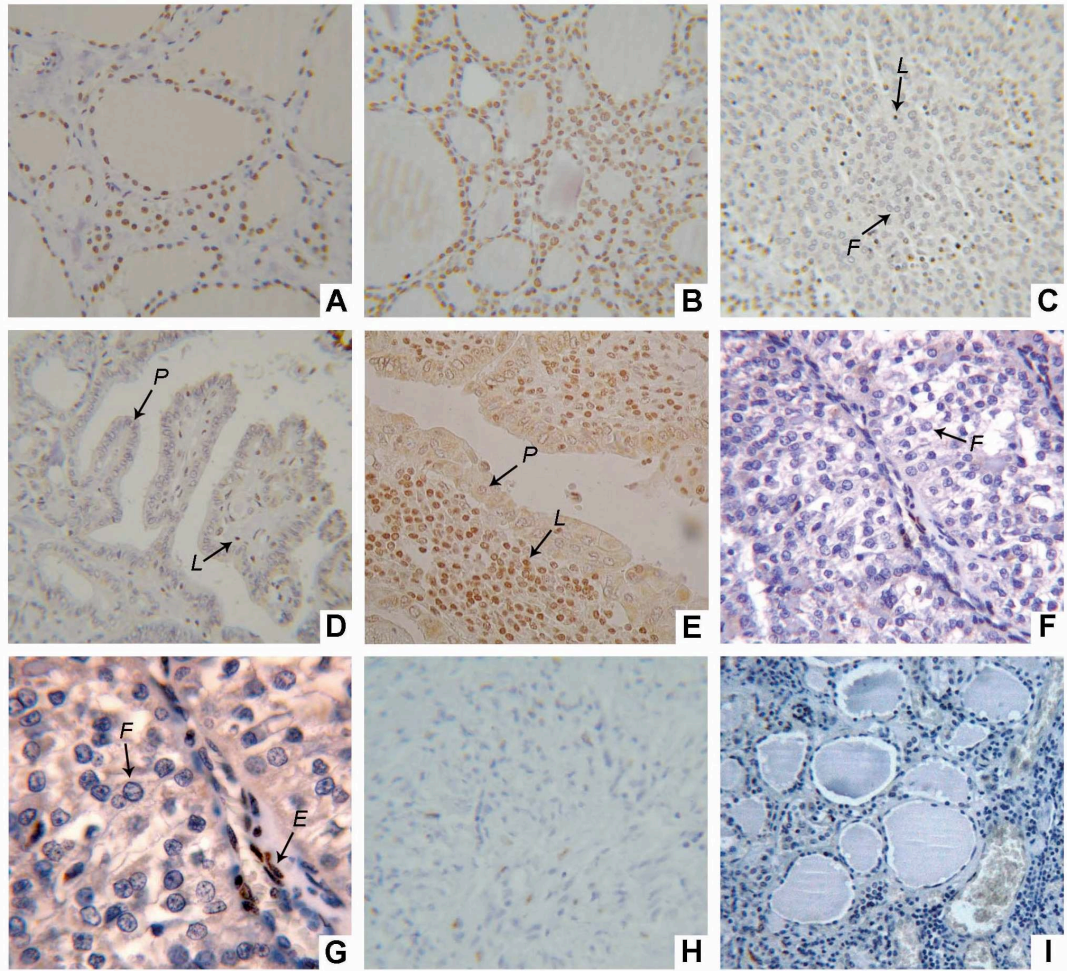


Figure 13 CBX7 staining in normal and neoplastic thyroid tissues
CBX7 nuclear staining was intense in benign follicular epithelial cells of normal thyroid (A) and of follicular adenoma (B), whereas it was weak in follicular carcinoma (C). CBX7 progressively decreased going from well-differentiated "classic variant" papillary (D) carcinomas to the "tall cell variant" of PTC (E), to poorly differentiated (F, G) and anaplastic (H) carcinomas. The signal disappeared after incubation of the sample with antigen (I). Sample showed in the G panel is the same of F, but magnified, to better show cell nuclei. Arrows with letters indicate the following sample features: *(P)* nuclei showing cytological features of PTC negative for CBX7 expression; *(F)* FTC nuclei negative for CBX7 expression; *(L)* lymphocyte showing CBX7 expression and providing positive internal control; *(E)* endothelial cells showing CBX7 expression and providing positive internal control.

4.4 Analysis of CBX7 expression in normal and neoplastic thyroid tissues by semiquantitative RT-PCR and quantitative Real-Time PCR

We also evaluated CBX7 expression by semiquantitative RT-PCR in a panel of matched normal/tumor tissues. The results confirmed the immunohistochemical data. In fact, there was an amplified band corresponding to CBX7 only in normal thyroid tissues (Figure 14A).

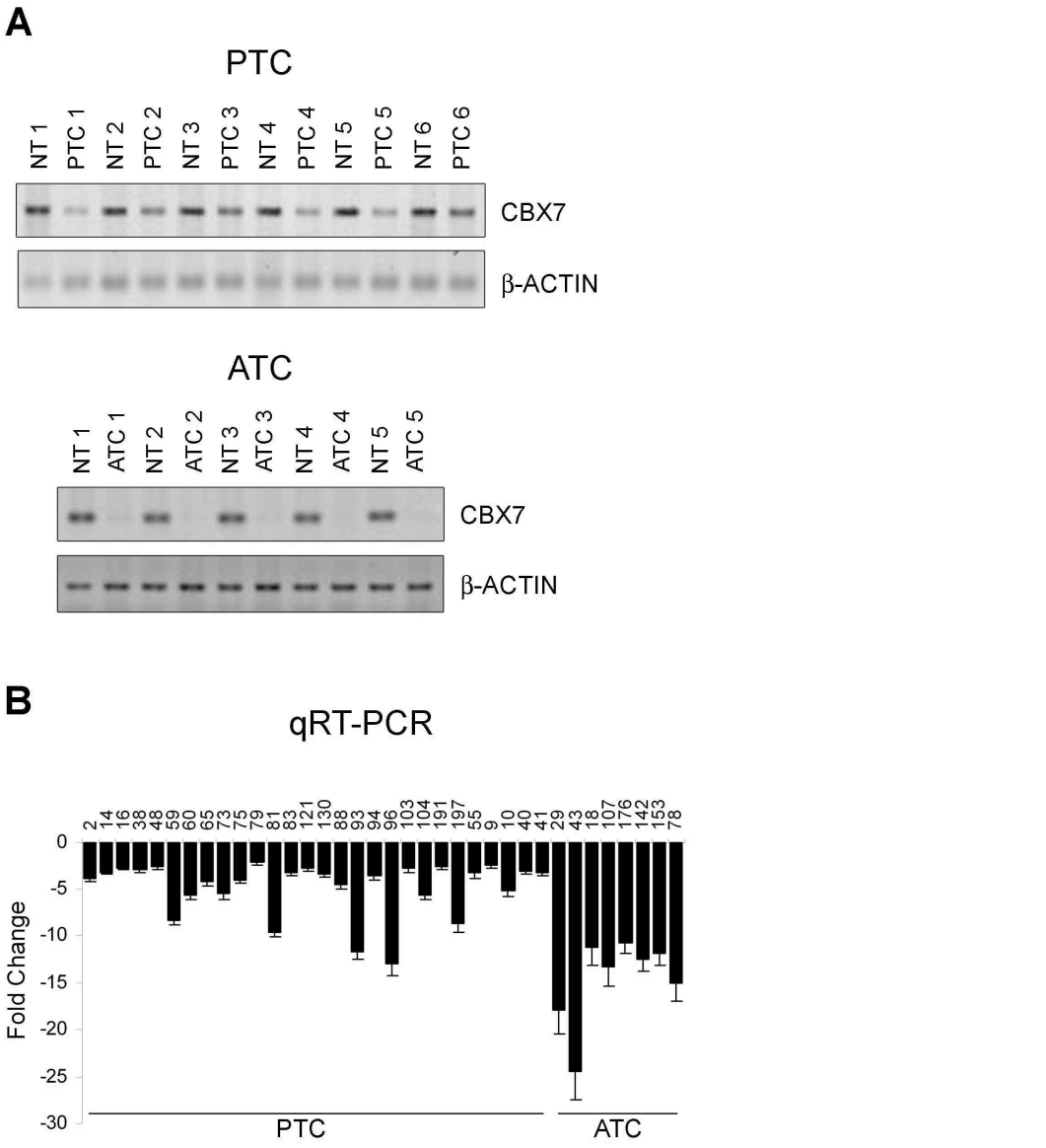


Figure 14 CBX7 expression in thyroid tumor samples
(A) RT–PCR analysis of CBX7 expression in human thyroid tumor samples vs their normal thyroid counterparts. β-Actin expression served as loading control. NT, normal thyroid tissue; PTC1 to PTC6, papillary thyroid carcinomas from different patients; ATC1 to ATC5, anaplastic

thyroid carcinomas from different patients. (B) qRT-PCR analysis of human thyroid tumor samples of different histotypes. The fold change indicates the relative change in expression levels between tumor samples and normal samples, assuming that the value of each normal sample is equal to 1.

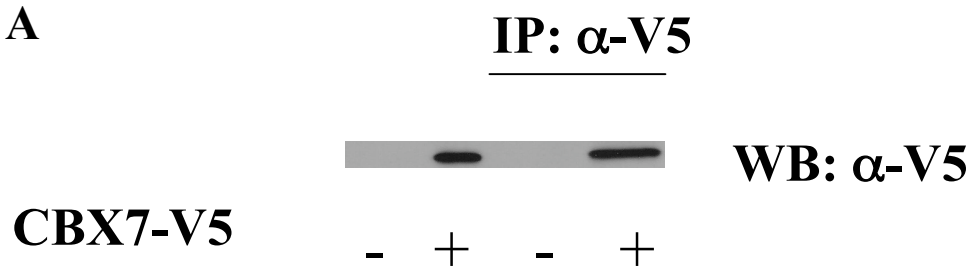
The amplified band decreased in PTC samples and almost disappeared in ATC. qRT-PCR analysis of a large number of human thyroid carcinoma samples of different histotypes confirmed a correlation between the reduction of CBX7 expression and a more malignant phenotype of thyroid neoplasias. In fact, as reported in Figure 14B, there was a negative fold change in CBX7 expression from –2.1 to –13 (average –4.8) in the PTC samples versus normal counterpart tissues. The reduction was even more pronounced in the anaplastic carcinoma samples, with a fold change ranging from –10.8 to –24.5 (average –14.6). These data are well correlated with the immunohistochemical ones and suggest that CBX7 expression is controlled at transcriptional level.

4.5 Proteomic analysis of CBX7 interacting proteins

In order to isolate and identify new CBX7 associated proteins, we performed proteomic analysis. To this aim, an expression vector coding for the full-length CBX7-V5, tagged with the V5 epitope, was transiently transfected in HEK 293 cells. The expression of CBX7 protein was verified by Western blotting analysis using anti V5 antibody.

About 1 mg of nucleas extract (NE) was then utilized for the proteomic analysis. In particular, proteins associated with CBX7 were purified by immunoprecipitation (IP) using anti-V5 agarose affiniy gel antibody.

To avoid having many aspecific proteins eluted in the loading buffer in the gel we have eluted the IP by competition with peptide V5. Preliminary SDS-PAGE experiments of the immunoprecipitated proteins showed a large number of protein bands. Therefore, before performing the preparative assay we have confirmed the expression of the CBX7-V5 construct in HEK 293 cells and the efficiency of the immunoprecipitation assay with anti-V5 agarose affinity gel antibody (Figura 15A). Then, the preparative immunoprecipitated material was fractionated on a 12% one-dimensional (1D) gel. Preparative gel was stained with Blue Coomassie as shown in the figure 15B.



B

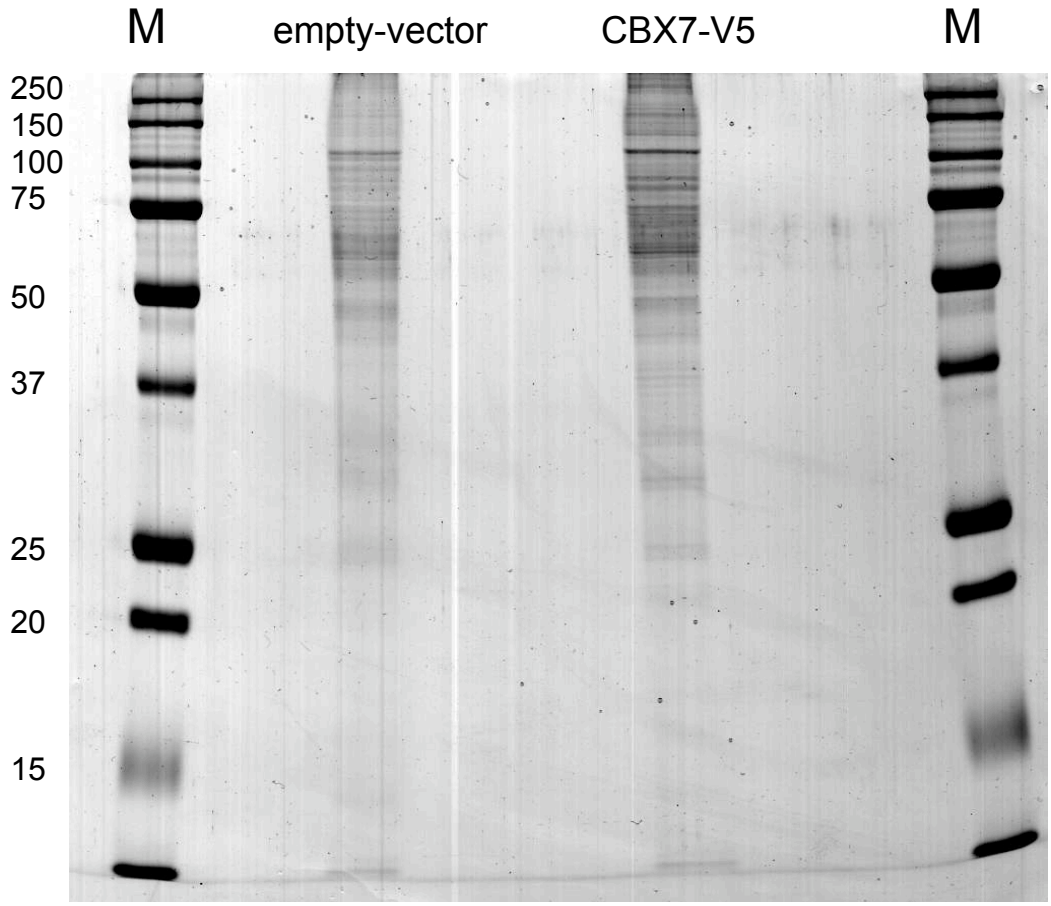


Figure 15 Proteomic analysis
A part of IP of empty vector and CBX7-V5 vector transfected HEK 293 cells were first fractionated on a 12% SDS-PAGE gel to verify the expression of CBX7 vector and the efficiency of the immunoprecipitation (A) and then on one-dimensional gel for proteomic assay. Next, the preparative gel was stained with Blue coomassie as shown in the figure (B) (M:marker)

The comparison between the "empty-vector" lane, corresponding to the immunoprecipitated lysate from empty-vector transfected cells, and the "CBX7-V5" lane, which corresponds to that of CBX7 overexpressing cells, allowed us to identify several bands that were present only in the "CBX7-V5" lane and absent in the control. Then, both gel lanes were sequentially cut in several fragments as shown in figure 16. Each protein band was digested in situ with trypsin (139) and

the peptide mixture obtained was directly analyzed by Liquid chromatography tandem mass spectrometry (LCMSMS).

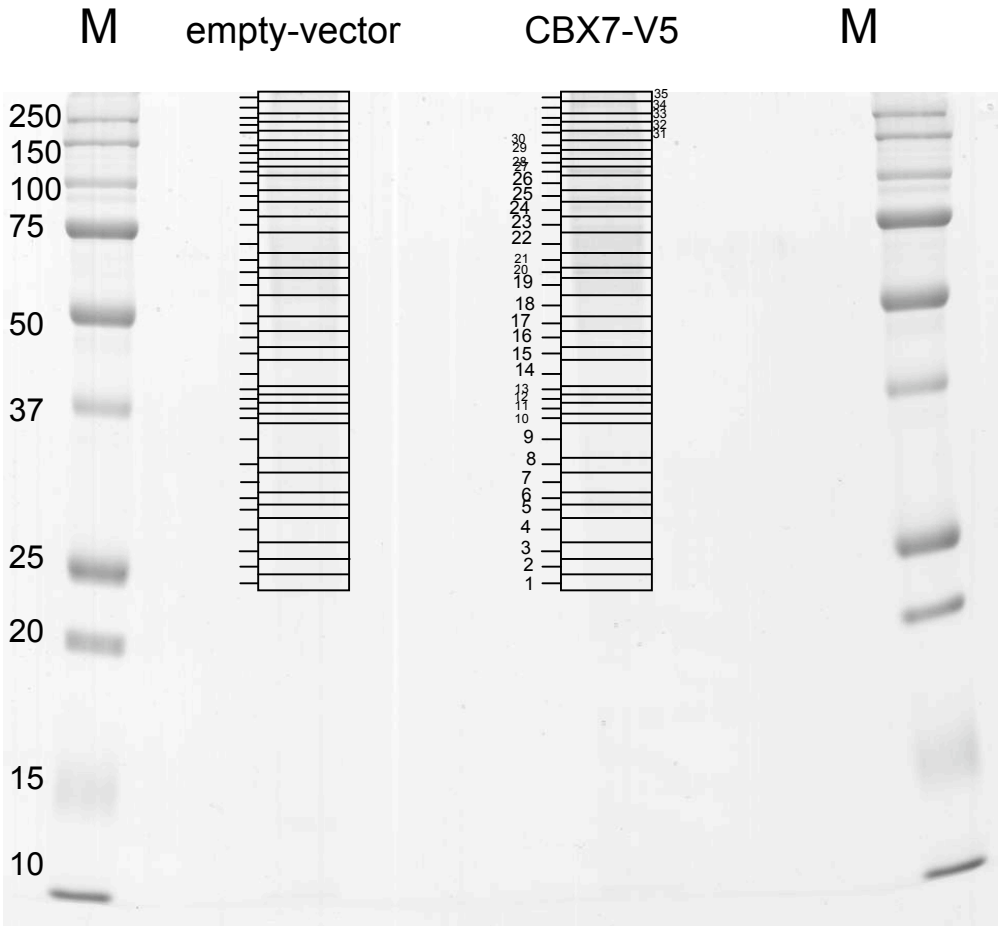


Figura 16 Proteomic analysis
IP of empty vector and CBX7-V5 were fractionated on one-dimensional gel. Preparative gel was stained with Blue coomassie and the gel was sequentially cut in several fragments covering all lane as shown in the figure (M:marker; number indicates a putative unique band)

The peak list of the peptides was introduced in the MASCOT MS/MS ion search software for protein identification. The final results are summarized in Table 3 (after the REFERENCES), where the names, accession numbers and function of the identified proteins are reported. It was not surprising that each excised band from 1-D gel separation contained more than one protein. We have

also clustered CBX7 interacting proteins observing that there are many proteins involved in DNA regulation, export and metabolism.

Generally, two peptide sequences are sufficient to identify a single protein, together with information about the score of peptides.

In our proteomic experiment we identified the bait CBX7 on the basis of a single peptide sequence, because the protein was not abundant; for this reason we decided to take in consideration proteins present exclusively in the CBX7 IP even if identified by just a single peptide recognition.

After MASCOT analysis we analized only the proteins that were present in the IP of CB7-V5.

4.6 *In vitro* characterization of CBX7 molecular complexes

All CBX7 putative partners identified by the proteomic approach still needed to be confirmed by independent experiments, such as GST pull-down or co-IP.

Among CBX7 interacting protein, we selected the HDAC2 because of its relevance in tumour biology.

To examine the specificity of this interactions and to map the regions of CBX7 proteins required for the binding to its molecular interactors, pull-down assays were performed incubating a total lysate deriving from HEK 293 cells with the CBX7 recombinant protein fused to GST and with two deletion mutants of CBX7, GST- CBX7 CHROMO and GST-CBX7 NOCHROMO, lacking from 55 to 251 aa and from 1 to 55 aa, respectively (Figure 17).

First, we analyzed the interaction between CBX7 and HDAC2. No associated

proteins were detectable in the complexes obtained incubating TCE with the GST protein alone or with the GST-CBX7 CHROMO fusion protein. Instead there was interaction between GST-CBX7 or the GST-CBX7 NOCHROMO mutant and HDCA2. These results clearly show that the interaction between CBX7 and HDCA2 involves the region of CBX7 outside the Chromodomain. Gels were stained with Blue coomassie to show that equal amount of GST-fusion proteins were used for the pull-down assay (data not shown).

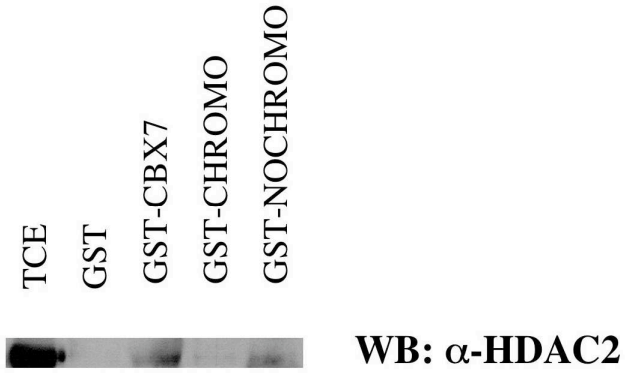


Figura 17 CBX7 interacts *in vitro* with HDAC2
5μg of each recombinant protein were incubated with 1 mg of TCE for 2 h. Then the beads were washed, the complexes resuspended in loading buffer and loaded on SDS-PAGE. The filter was incubated with antibody against HDAC2.

4.7 CBX7 physically interacts with HDAC2 protein

To verify the CBX7/HDAC2 interaction *in vivo*, HEK 293 cells were transiently transfected with the V5-tagged-CBX7 expression vector. Protein lysates were immunoprecipitated with anti-V5 antibody and immunoblotted with anti-V5 and anti-HDAC2 antibodies. As shown in Figure 8, we were able to detect the association between CBX7 and the endogenous HDAC2 protein, demonstrating that CBX7 and HDAC2 form complexes *in vivo*. The reciprocal experiment was performed immunoprecipitating with anti-HDAC2 and revealing with anti-V5 antibodies: it confirmed the interaction between the HDAC2 and CBX7 proteins (data not shown).

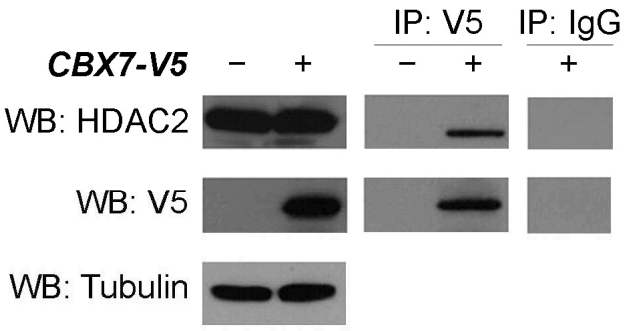


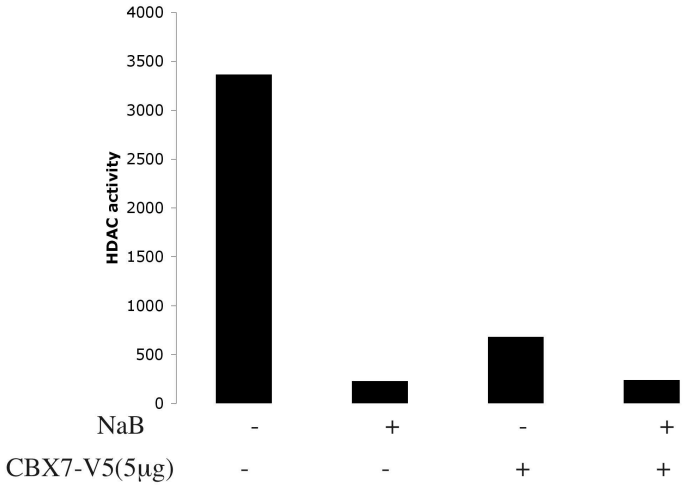
Figura 18 HDAC2/CBX7 *in vivo* interaction by co-immunoprecipitation
HEK 293 cells were transfected with V5-tagged-CBX7 vector and, after 48 h, total cell extracts were prepared and equal amounts of proteins were immunoprecipitated with anti-V5 antibodies. The immunocomplexes were analysed by Western blotting using anti-HDAC2 antibodies. The relative inputs are total cell extracts derived from HEK 293 untransfected cells. IgG indicates the negative control of immunoprecipitation using an unrelated antibody.

4.8 CBX7 inhibits HDAC activity

In order to evaluate the effects of the CBX7/HDAC2 interaction on the HDAC2 activity, we performed an HDAC activity assay on HEK 293 cells overexpressing CBX7.

As shown in figure 19A, CBX7 inhibited HDAC activity, almost as strongly as sodium butyrate (NaB), a robust HDAC inhibitory compound (144), used as positive control. In figure 19B the CBX7 expression levels in transfected cells are shown.

A



B

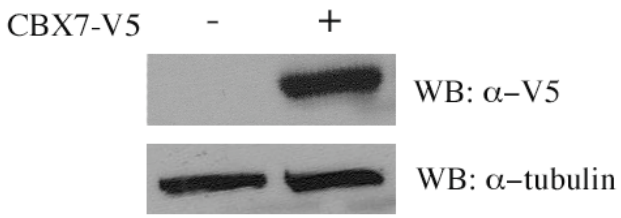


Figura 19 HDAC activity assay.
HEK 293 cells were transfected with empty vector(-) and CBX7 expression vector. (A) Nuclear extracts were assayed for HDAC activity after 48 h. (B)Samples treated with sodium butyrate (NaB) were used as positive control. Values represent the average of three experiments +/- SD(lower panel). Aliquots of the same lysates were resolved by SDS-PAGE, transferred to nitrocellulose membrane, and immunoblotted with the indicated antibodies.

To further confirm the inhibitory activity exerted by CBX7 on HDAC2 activity, we transfected HeLa and NPA cells with different amount of CBX7.

Remarkably, CBX7 inhibited the HDAC activity in the cell lines used in a dose-dependent manner (the percentage of inhibition was 65,15% and 42,55% in HeLa and NPA cells respectively, after the transfection of 10 μg of CBX7) (Figure 20). Again as positive control for HDAC inhibition we used sodium butyrate (NaB) (144).

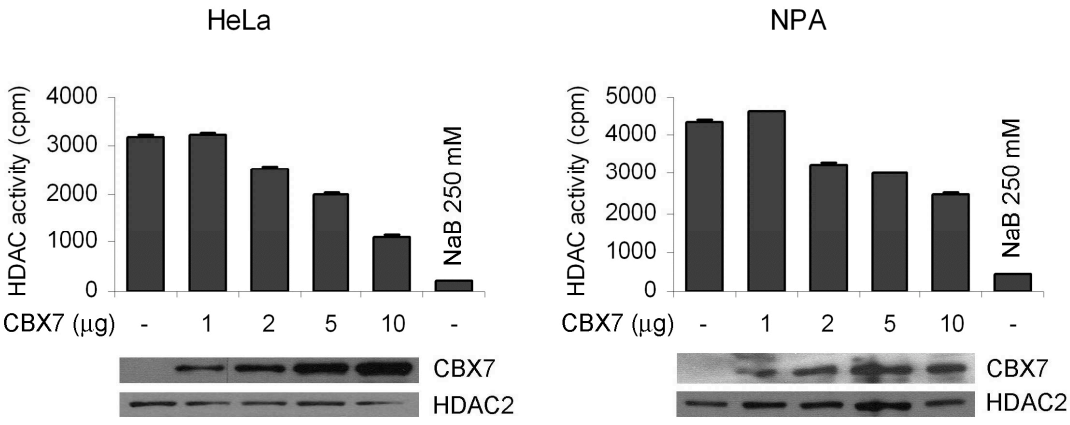


Figura 20 HDAC activity assay in Hela and NPA
HeLa and NPA cells were transfected with increasing amounts of CBX7 expression vector and nuclear extracts were assayed for HDAC activity after 48 h. Samples treated with sodium butyrate (NaB) were used as positive control. Values represent the average of three experiments +/- SD. Aliquots of the same lysates were resolved by SDS-PAGE, transferred to nitrocellulose membrane, and immunoblotted with the indicated antibodies (lower panel)

4.9 CBX7 and HDAC activity in tumors

As previously shown by RT-PCR and immunohistochemistry, CBX7 was strongly down-regulated in tumors. Therefore, we investigated HDAC activity in normal thyroid in comparison to anaplastic tumor (ATC) histotype. As expected, in absence of CBX7 there was an increase in HDAC activity (figure 21).

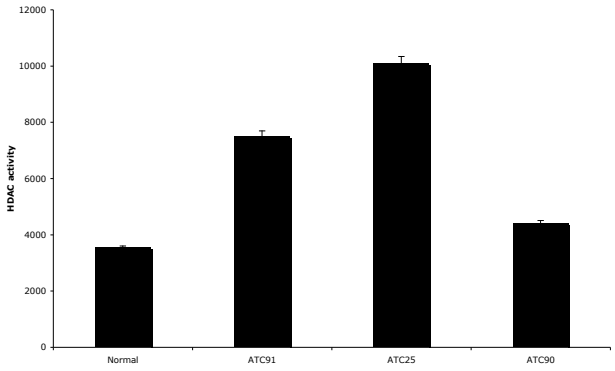


Figura 21 HDAC activity assay in anaplastic tumors (ATC)
Normal thyroid and sample of ATC were assayed for HDAC activity. Values represent the average of three experiments +/- SD.

4.10 CBX7 binds to E-cadherin gene promoter *in vivo*

HDACs catalyze the removal of acetyl groups from core histones (145). Because of their capacity to induce local condensation of chromatin, HDACs are generally considered repressors of transcription (145) and it is already known that HDAC2 is able to repress E-cadherin promoter (146). Given that CBX7 is a protein that probably binds chromatin, we asked whether the physical interaction between CBX7 and HDAC2 takes place on the human E-cadherin promoter, an important marker of EMT. Therefore, we first evaluated whether CBX7 protein binds the E-cadherin promoter *in vivo* by performing ChIP assays. HEK 293 and HeLa cells

were transfected with HA-tagged-CBX7 expression vector, crosslinked, and immunoprecipitated with anti-HA or anti-IgG antibodies. Immunoprecipitation of chromatin was subsequently analysed by semiquantitative PCR, using primers spanning the region of the E-cadherin promoter (300 bp upstream to 40 bp downstream to the transcription start site). Anti-HA antibodies precipitated this E-cadherin promoter region from HEK 293 and HeLa cells transfected with HA-tagged-CBX7 protein (Figure 22). No immunoprecipitation was observed with anti-IgG precipitates and with primers for the control promoter GAPDH indicating that the binding is specific for the E-cadherin promoter. We performed also PCR amplification of the immunoprecipitated DNA using primers for p16, E2F1 and S100A4 gene promoter to validate the presence of the CBX7 on the promoters of other genes.

Similar results were obtained when the NPA cells were used (data not shown). These results indicate that CBX7 protein binds the E-cadherin promoter region *in vivo*.

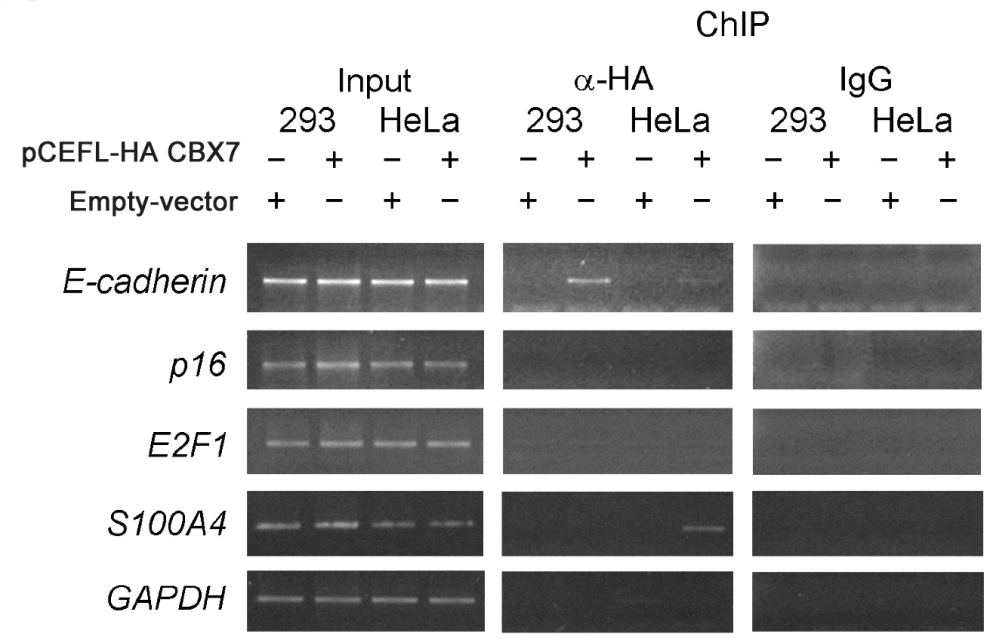


Figure 22 CBX7 binds to the E-cadherin promoter *in vivo*
Soluble chromatin from HEK 293 and HeLa cells transfected with HA-tagged-CBX7 was immunoprecipitated with anti-HA antibody. The DNAs were then amplified by PCR using primers that cover a region of human E-cadherin promoter (300 bp upstream to 40 bp downstream to the transcription start site). As an immunoprecipitation control, IgG was used. The panel shows also PCR amplification of the immunoprecipitated DNA using primers for p16, E2F1, S100A4 and GAPDH gene promoter

4.11 CBX7 occupies the E-cadherin promoter with HDAC2

To determine whether CBX7 occupies the E-cadherin promoter together with HDAC2, the anti-HA complexes were released, re-immunoprecipitated with anti-HDAC2 antibodies, and then analyzed by semiquantitative PCR (Re-ChIP). The results shown in the Figure 23 reveal that the antibodies against HDAC2 precipitate the E-cadherin promoter after their release from anti-HA, indicating that CBX7 occupies this region together with HDAC2. The reciprocal experiment provided comparable results (data not shown). Taken together, these results indicate that CBX7 binds to the human E-cadherin promoter *in vivo*, and participate in the same DNA-bound complex containing HDAC2.

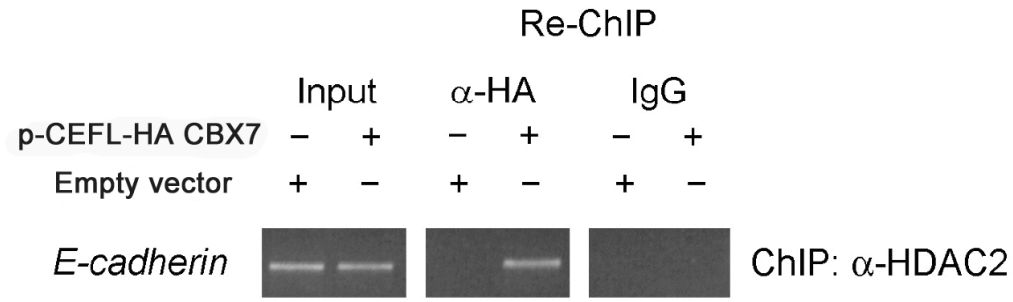


Figure 23 Re-ChIP experiment on E-cadherin promoter
In Re-ChIP experiments, soluble chromatin from transfected HEK 293 and HeLa cells was immunoprecipitated with anti-HDAC2 antibody, eluted, and reimmunoprecipitated with anti-HA. The purified DNA was used as template for PCR using the same primers described in figure 22.

4.12 CBX7 and E-cadherin expression are correlated in human thyroid carcinomas

In order to evaluate if the CBX7 binding to E-cadherin promoter could paly a role in thyroid carcinogenesis, the next step was to analyze E-cadherin expression in human thyroid carcinomas of different histotypes. As shown in the Figure 24 a positive correlation was found between CBX7 and E-cadherin expression. In fact, E-cadherin mRNA levels were drastically reduced in the ATC samples, whereas just a weak decrease was observed in the PTC ones. CBX7 expression showed the same trend when the same samples were analyzed. Immunohistochemical analysis confirmed that E-cadherin protein levels paralleled CBX7 protein levels (data not shown). These results are consistent with other reported data showing a down-regulation of the E-cadherin gene in human carcinomas of diverse anatomic origin characterized by a highly aggressive phenotype (107).

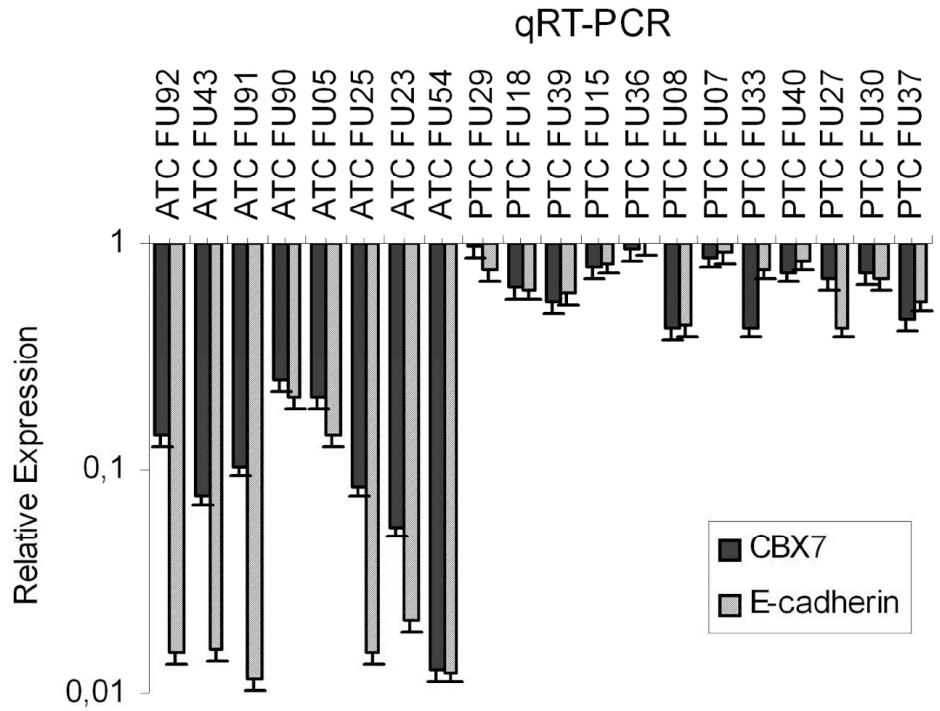


Figure 24 CBX7 and E-cadherin expression in fresh thyroid tumour samples analysed by qRT-PCR

The fold change indicates the relative change in expression levels between tumour samples and normal samples, assuming that the value of each normal sample is equal to 1. PTC, papillary thyroid carcinoma; ATC, anaplastic thyroid carcinoma. Numbers indicated the patients.

4.13 CBX7 modulates E-cadherin gene transcription

To evaluate the effect of CBX7 expression on E-cadherin transcription, HEK 293 cells were co-transfected with an expression vector encoding CBX7 and with a reporter vector carrying the luciferase gene under the control of the E-cadherin promoter. As shown in Figure 25, CBX7 was able to increase the transcriptional activity of the E-cadherin promoter in a dose-dependent manner. Conversely, no increase of luciferase activity was observed when CBX7 was co-transfected with the luciferase vector without E-cadherin promoter (data not shown).

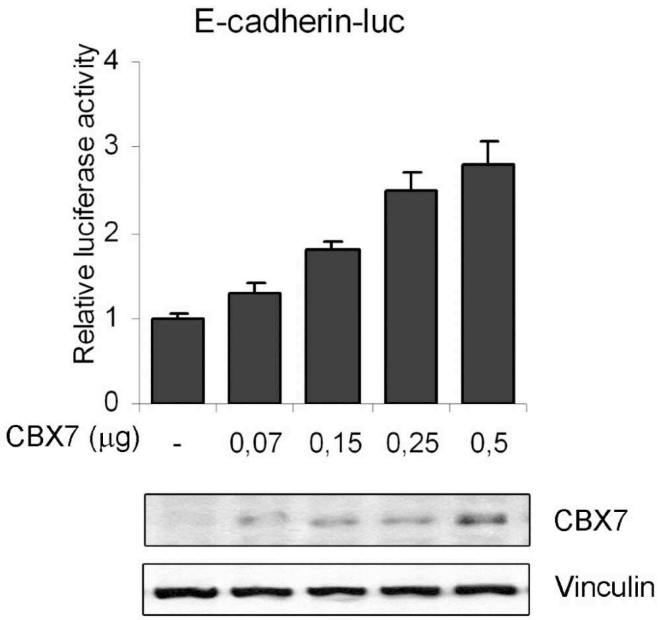


Figure 25 Activity of CBX7 on the E-cadherin promoter
Dose-response analysis of increasing amounts of CBX7 on the E-cadherin luciferase-reporter vector transiently transfected into HEK 293 cells. All transfections were performed in duplicate and the data are mean +/- SD of three independent experiments. Cells transfected with the backbone vector were used as control. Western blot for CBX7 from one representative experiment is shown in the lower panel. Vinculin was used to normalize protein loading.

The treatment of cells with trichostatin (TSA), a potent inhibitor of HDAC activity (147) was able to synergize with CBX7 in induce E-cadherin gene transcription (Figure 26). This results strongly suggest that CBX7 protein is involved in the E-cadherin gene transcription, likely counteracting the inhibitory effect of HDAC2 on this gene promoter.

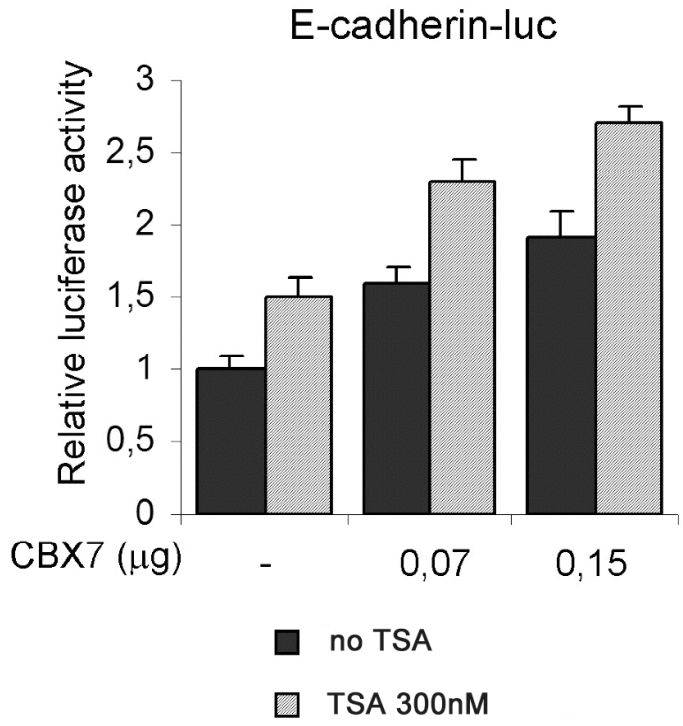
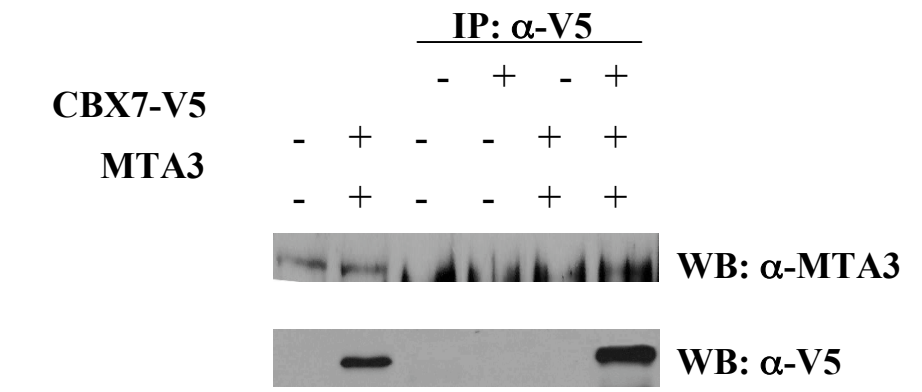


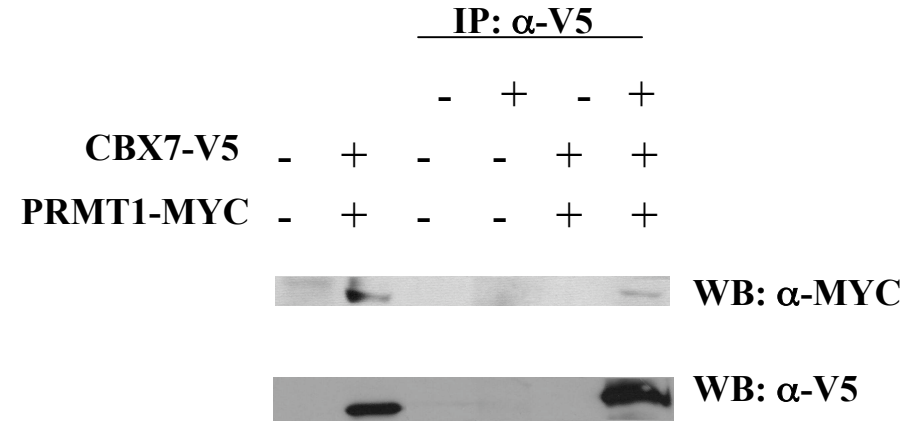
Figure 26 The CBX7 activity on the E-cadherin promoter is inhibited by HDAC inhibitor
Luciferase activities of extracts from HEK 293 cells cotransfected with the E-cadherin luciferase-reporter vector, increasing amounts of CBX7 and treated with the HDAC inhibitor trichostatin A **(**TSA) at 300 nM final concentration. All transfections were performed in duplicate and the data are mean+/- SD of three independent experiments.

4.14 *In vivo* characterization of CBX7 molecular complexes

To validate the interaction between CBX7 protein and some of the newly identified associated proteins *in vivo*, total cellular extracts from HEK 293 cells transfected with the empty vector or with the full-length CBX7 cDNA tagged with the V5 epitope were immunoprecipitated with the antibodies against V5 and transferred on filters that were subsequently immunoblotted with anti-V5 or antibodies against each interacting protein. As shown in Figure 27, a great amount of CBX7 protein was present in all of the immunocomplexes from HEK 293 cells overexpressing CBX7 protein, confirming that CBX7 and the indicated proteins really interact.

To examine the effective amount of the molecular interactors identified in HEK 293 cells, we evaluated their protein levels in empty vector or CBX7-overexpressing cells by Western blotting analysis. As shown in the Figure 27, the amount of each protein in these cells was comparable. As control, we confirmed the expression of CBX7 in transfected HEK 293 cells using anti-V5 antibody. These experiments demonstrate that the interactions, evaluated through proteomic analysis, are real, and that CBX7 protein belongs, at least in this cellular system, to new molecular complexes. We have performed the co-immunoprecipitation for other three proteins. In particular, MTA3 which might be a component of the same complex of CBX7 on the E-cadherin promoter and PRMT1 and PRMT5 could have a regulatory function in arginine methylation together with CBX7.





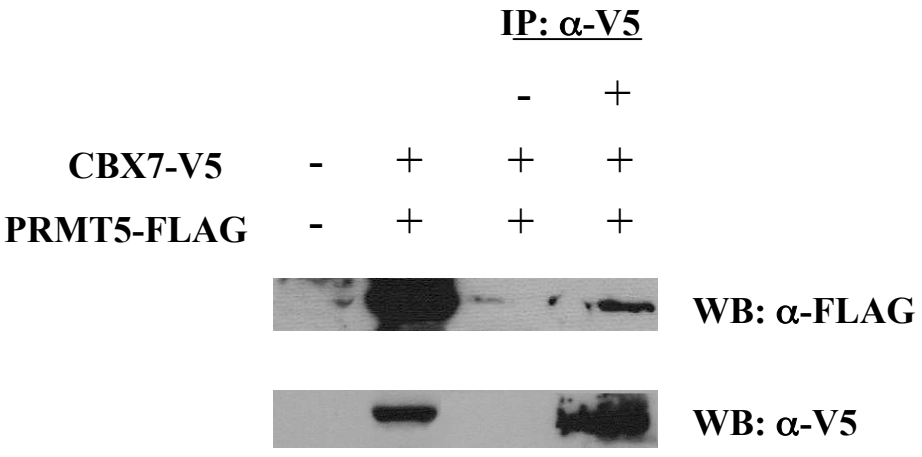


Figure 27 Confirmation of CBX7 interactions by co-immunoprecipitation
HEK 293 cells were transfected with V5-tagged-CBX7 vector and 48 h after transfection total cell extracts were prepared and equal amounts of proteins were immunoprecipitated with anti-V5 antibodies. The immunocomplexes were analysed by Western blotting using the indicated antibodies. The relative inputs are total cell extracts derived from HEK 293 untransfected and transfected cells.

4.15 Fractionation of CBX7-V5 transfected cells, with or without RNase A treatment

It has already been shown that CBX7 interacts also with RNA (84). To assay if the complexes of CBX7 are dependent on RNA binding we perfomed a fractionation on glycerol-gradient with and without RNase A, an endoribonuclease that specifically degrades single-stranded RNA.

We transfected HEK 293 cells with empty or CBX7-V5 expressing vector and performed NE. We treated 2 mg of NE of CBX7-V5 with RNase A and other 2 mg of NE with RNasin, an RNase inhibitor. Then we loaded the two samples of CBX7-V5 transfected cells on top of two 5 to 45% glycerol gradients. The gradients were spun at 4°C for 16 h at 39 000 rpm. Ten fractions were collected from the top of the gradients(1 to 10) plus a pellet(11).

To validate if RNAse A treatment changes the distribution of CBX7 complexes we analyzed all fractions by SDS-PAGE. To check the efficacy of the sedimentation we followed the sedimentation of CDK9, one of two component of Positive Transcription Elongation Factor b (P-TEFb). P-TEFb, a cyclin-dependent kinase, comprised of Cdk9 and in humans one of three cyclins, T1, T2, or K, is the key factor regulating eukaryotic mRNA transcription at the level of elongation. As well described in literature, P-TEFb sediments in two sets of fractions (143). Upper fractions contain active core P-TEFb, whereas the large kinase-inactive P-TEFb/7SK RNA complex is found in the lower fractions.

We found that the sedimentation of CBX7-V5 not change after treatment with RNAse, while P-TEFb sediments were in two sets of fractions (figure 28)

thus demonstrating that the complexes of CBX7 were not influenced by binding with RNA.

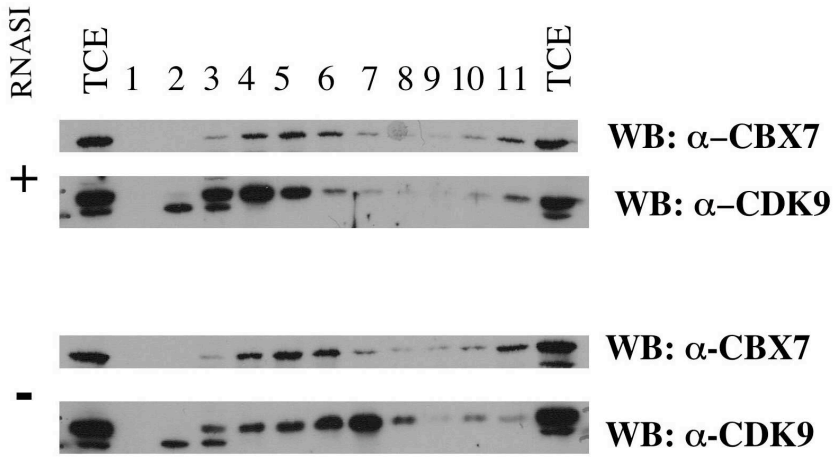


Figure 28 TCE separation by ultracentrifugation of CBX7-V5 transfected cells treated with RNase A
CBX7-V5 construct was transfected in HEK 293 cells and TCE was fractionated into 10 fractions (1-10) plus a pellet (11) by ultracentrifugation on a 5-45% glycerol-gradient. The fractions recovered were loaded on SDS-PAGE and analyzed for V5 expression. As control we followed the sedimentation of CDK9 (a component of p-TEFb) that sediment in two components.

4.16 Sedimentation of CBX7-V5 interacting proteins

In order to verify the segregation of CBX7-V5 interacting proteins we transfected HEK 293 cells with CBX7-V5 expressing vector. TCE were fractionated into 10 fractions (1-10) plus a pellet (11) by ultracentrifugation on a 5-45% glycerol-gradient. The fractions recovered were loaded on SDS-PAGE and analyzed for CBX7 interacting proteins, identified by the proteomic experiment, or proteins that have already been described in the same pathway (figure 29).

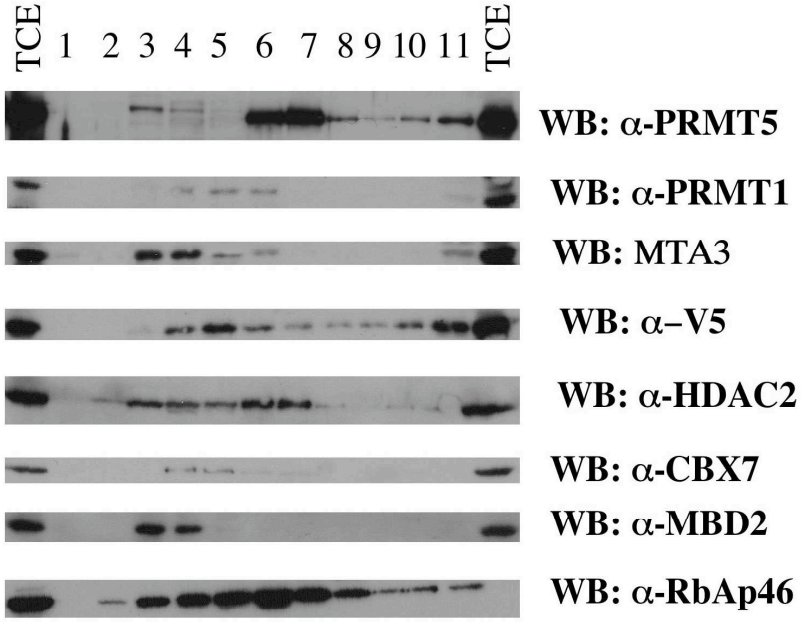


Figure 29 TCE separation of CBX7-V5 transfected cells
CBX7-V5 was transfected in HEK 293 cells and TCE was fractionated into 10 fractions (1-10) plus a pellet (11) by ultracentrifugation on a 5-45% glycerol-gradient. The fractions recovered were loaded on SDS-PAGE and analyzed for the expression of V5 and other proteins.

4.17 Characterization and function of CBX7 complexes

In an attempt to identify the CBX7 complexes that bind to the E-cadherin promoter we performed an EMSA assay. First, NE were extracted from HEK 293 cells transfected with an empty or a CBX7-V5 expressing vector. We performed ChIP experiments to verify the site of action of CBX7 on the E-cadherin promoter and to identify the minimal site to perform EMSA experiments (data not shown). We identified a fragment of E-cadherin promoter containing an E-box.

So we first performed an EMSA assay analyzing the capacity of nuclear extracts (5 μg) from HEK 293 cells transfected with an empty vector or a CBX7-V5 to bind a E-cadherin oligonucleotide, as we have seen in ChIP and Re-ChIP

experiments.

NE containing empty or CBX7-V5 expressing vector binds E-cadherin oligonucleotide specifically, forming complexes with various electrophoretic mobilities (complex A, B and D). When HEK 293 cells were transfected with CBX7-V5 we observed the formation of a fourth complex (C) (figure 30).

Binding specificity was demonstrated by competition experiments showing loss of binding on all the three complexes with the addition of a 100-fold molar excess of unlabelled E-cadherin oligonucleotide. In particular, in HEK 293 cells transfected with CBX7-V5 there was also the disappearence of the CBX7 specific complex (complex C).

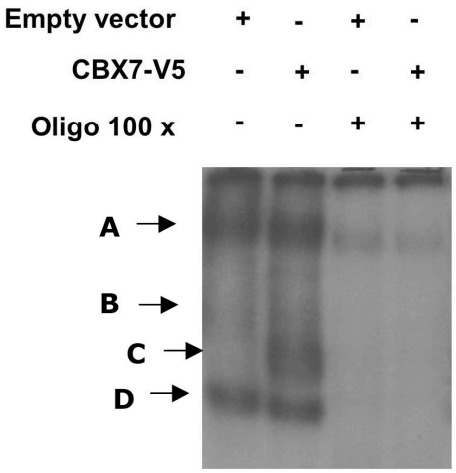


Figure 30 CBX7 binds to the E-cadherin promoter in vitro
EMSA performed with a radiolabelled oligonucleotide containing a fragment of the E-cadherin promoter (including E-Box) incubated with nuclear extracts (5 µg) from HEK 293 cells transfected with an empty vector or the CBX7-V5 expression vector. To assess the specificity of the binding, nuclear extracts protein was incubated in the presence of a 100-fold excess of unlabelled oligonucleotide used as competitor.

In an attempt to determine the composition of these complexes and the presence of proteins identified by proteomic assay, we used specific antibodies directed against the CBX7 interacting proteins in a supershift analysis.

We observed a significant supershift (complex E) with the antibody against Protein arginine N-methyltransferase 1 (PRMT1) in the lane with CBX7-V5 transfected cells extracts, indicating that PRMT1 partecipates in CBX7 complex (complexC) on E-cadherin promoter. (figure 31).

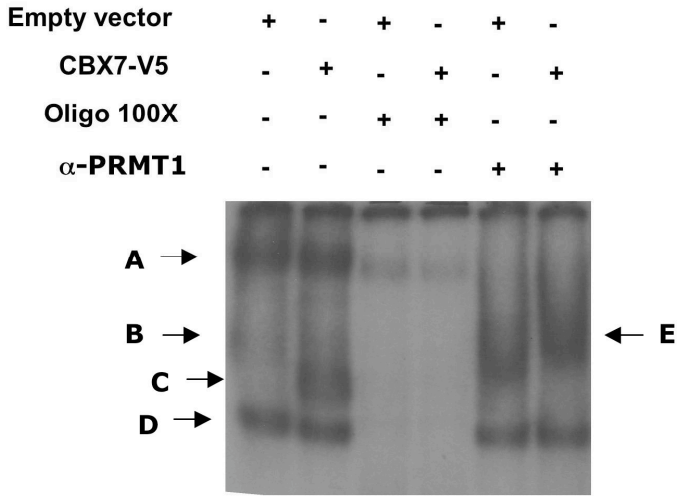


Figure 31 CBX7 and PRMT1 on E-cadherin promoter
EMSA performed with a radiolabelled oligonucleotide containing a fragment of the E-cadherin promoter (including E-Box) incubated with nuclear extracts (5 µg) from HEK 293 cells transfected with an empty vector or the CBX7-V5 expression vector. To assess the specificity of the binding, nuclear extracts protein was incubated in the presence of a 100-fold excess of unlabelled oligonucleotide used as competitor and with the antibody indicated. Arrows indicate specific complexes DNA/proteins.

Instead, in the case of the Metastasis associated 1 family, member 3, (MTA3) supershift the presence of antibody against MTA3 increased the binding of complex C to the E-cadherin oligonucleotide and decreased the amount of complex A (figure 32).

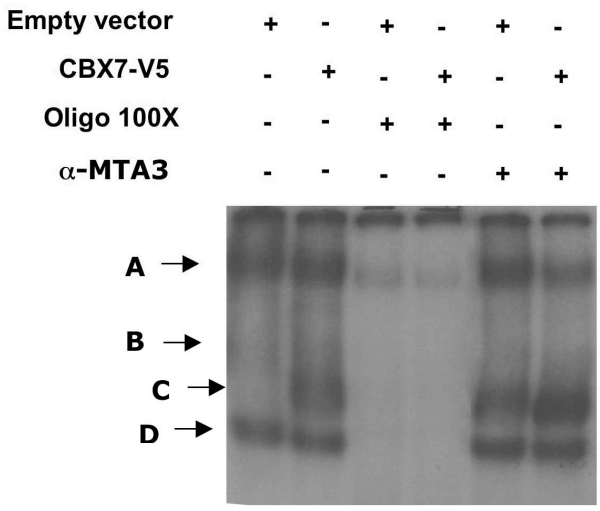


Figure 32 CBX7 and PRMT1 on E-cadherin promoter
EMSA performed with a radiolabelled oligonucleotide containing a fragment of the E-cadherin promoter (including E-Box) incubated with nuclear extracts (5 μg) from HEK 293 cells transfected with an empty vector or the CBX7-V5 expression vector. To assess the specificity of the binding, nuclear extracts protein was incubated in the presence of a 100-fold excess of unlabelled oligonucleotide used as competitor and with the antibody indicated. Arrows indicate specific complexes DNA/proteins.

4.18 CBX7 and Protein arginine N-methyltransferase 1 (PRMT1)

In order to better investigate the interaction between CBX7 and PRMT1 we performed a histone methyl transferase assay to validate:

a) If CBX7 has an associated arginine methylase activity, as suggested by the verified association with PRMT1;
b) if CBX7 is a substrate of PRMT1, as suggested by the evidence that the aminoacid sequence of CBX7 has two possible site for argine methylation.

To verify the first point we transfected HEK 293 cells with CBX7-V5 and immunoprecipitated with antidoby against V5.

Then, we assayed the ability of the immunoprecipitated proteins, i.e. the proteins interacting with CBX7, as PRMT1, to methylate arginine of the substrate.

In this assay we used as control a recombinant PRMT1 and as substrate core histones (figure 33).

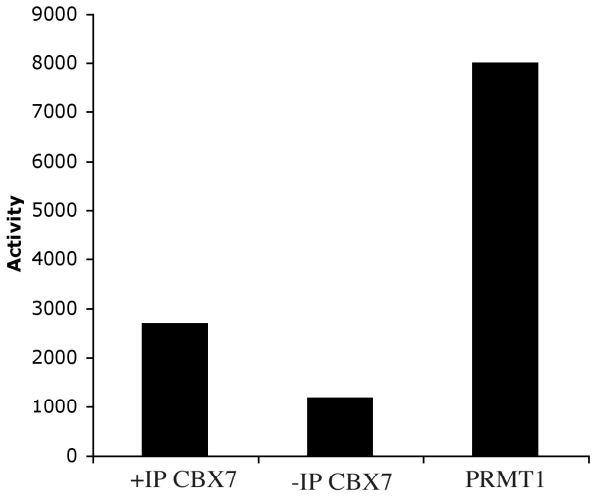


Figure 33 Histone methyltransferase assay
HEK 293 cells transfected with CBX7-V5 and immunoprecipitated were assayed for arginine methyltransferase activity (+IP: immunoprecipitation of CBX7-V5, -IP: negative control of immunoprecipitation, without antibody, PRMT1 is the positive control of the assay)

5 CONCLUSIONS

Thyroid tumors result from the accumulation of different modifications in critical genes involved in the control of cell proliferation. Although various therapeutic approaches are followed in clinical practice, most of them are not life-saving. Hence, the discovery of new approach to diagnose cancer at an early stage and to establish more effective therapies is a critical and urgent issue. To achieve this goal, identification and characterization of key molecules that participate in carcinogenesis are essential steps.

In this study we show that, by the use of cDNA microarray, it is possible to identify genes potentially involved in human cancers. In fact we identified CBX7, that could be actively involved in the process of thyroid carcinogenesis.

CBX7 is abundantly expressed in normal thyroid gland, but its expression decreased with malignancy and neoplasia stage. In fact, CBX7 expression was comparable to normal thyroid tissue in benign follicular adenomas, slightly reduced in papillary thyroid carcinomas displaying the classical histotype, and drastically reduced, and in most cases absent, in follicular thyroid carcinomas, in the tall cell variant of papillary thyroid carcinomas, in poorly differentiated and in anaplastic thyroid carcinomas. Thus, it is feasible that CBX7 levels could serve to differentiate between benign and malignant thyroid neoplasms.

In order to study CBX7 function in normal cells and understand why it is drastically down-regulated in thyroid cells, we have decided to investigate proteins interacting with CBX7, by a proteomic assay. Proteomic assays are very useful because the association of CBX7 with partners belonging to a specific

protein complex involved in a particular mechanism is strongly suggestive of the biological function of the protein (130).

In particular in our investigation we have found that CBX7 regulates an key mediator of EMT, E-cadherin, through HDAC2 activity.

EMT describes a series of events during which epithelial cells loose many of their epithelial characteristics and take on properties that are typical of mesenchymal cells, which require complex changes in cell architecture and behaviour. During progression of the tumor, carcinoma cells acquire mesenchymal gene-expression patterns and properties (90).

One of the hallmarks of EMT is the functional loss of E-cadherin (encoded by CDH1), which is currently thought to be a suppressor of invasion during carcinoma progression (96).

We demonstrated that CBX7 interacts with HDAC2, and is able to inhibit its activity in HeLa and in thyroid carcinoma cell lines. This result, by itself, may account for the increased aggressiveness of the neoplasias lacking CBX7 expression. In fact, HDACs are generally considered repressors of transcription, and HDAC activity is generally increased in human neoplasias, and, consistently, HDAC inhibitors represent a novel therapeutic strategy for cancer treatment (148).

Interestingly, we were able to demonstrate that both HDAC and CBX7 bind the E-cadherin promoter, and that CBX7 contrasts the inhibiting effect of HDAC on E-cadherin expression. This led us to hypothesize that the loss of CBX7 expression occurring in ATCs results in an increased HDAC activity on the E-cadherin promoter, then switching off its expression.

Consistently, E-cadherin expression essentially parallels CBX7 expression in thyroid carcinomas. Moreover, no E-cadherin promoter hypermethylation, frequently observed in other human neoplasias (107), has been detected in ATC (Sacchetti et al., unpublished data), suggesting that epigenetic mechanisms as those described here, may account for the silencing of the E-cadherin gene. Of course, it is reasonable to hypothesize that CBX7 may control through the same or other mechanisms the regulation of other genes involved in cell growth regulation and invasion.

In this scenario, we speculate that CBX7 could exert its effect on thyroid cancer progression also interacting with PRMT1 (Protein arginine N-methyltransferase 1) and MTA3, (metastasis associated 1 family, member 3).

PRMT1 is an arginine methyltransferase that is described to methylate H4R3, thus inducing transcription (149-152). Frequently PRMT1 methylates the same residues of PRMT5 as in the case of SPT5, that we have also found in our proteomic assay (153, 154). SPT5 and its binding partner SPT4 function both positively and negatively in regulating transcriptional elongation. SPT5 is specifically associated with PRMT1 and PRMT5. Also H4R3 is a site of methylation of both PRMT, with opposite consequences.

Sometimes deacetylation of the promoters is a prerequisite for PRMT5-mediated H3 and H4 arginine methylation (150). It is suggestive in case of CBX7 and HDAC2 on E-cadherin promoter. It has been also described the association of PRMT5 with Sin3A/HDAC2 complex (155).

Moreover, we have also to remember that CBX7 is described to interact with RNA and methylation of arginine is a common feature of RNA binding proteins (84).

Also the interaction of CBX7 with MTA3 could be very important. MTA3 has been yet described as a regulator of E-cadherin expression by transcription factor Snail in breast cells through a mechanism HDAC1/2 dependent (129, 156). For this reason it is very suggestive that CBX7 mediates activation of E-cadherin repressing HDAC2 activity.

It should be interesting to investigate the relationship of CBX7 and MTA3 in regulating E-cadherin. Infact, it has been also described the interaction of PRMT5 with HDAC2 and MTA3 (157).

In our proteomic experiment we have also identified many proteins involved in mRNA maturation, splicing and elongation. PRMT1 regulates also poly(A) binding protein in fission Yeast, another protein that interacts with CBX7 (158).

Protein that bind RNA have generally a common feature (RG) that is a site of methylation on arginine. So these data are in agreement with preliminary data suggesting a role for CBX7 in splicing and regulation of elongation.

As suggested by many important interacting protein of CBX7 and the pathway in which they are involved, CBX7 can be considerated a very important regulator of thyroid transcription and malignant transformation and a target of therapy.

6 REFERENCES

1. Kondo T, Ezzat S, Asa SL. Pathogenetic mechanisms in thyroid follicular-cell neoplasia. Nat Rev Cancer 2006;6:292–306.
2. De Felice M, Di Lauro R. Thyroid Development and Its Disorders: Genetics and Molecular Mechanisms. Endocrine Reviews 2004;25:722–46.
3. Mauchamp J, Mirrione A, Alquier C, Andre` F. Follicle-like structure and polarized monolayer: role of the extracellular matrix on thyroid cell organization in primary culture. Biol Cell 1998;90:369–80.
4. Le Douarin N, Fontaine J, LeLievre C. New studies on the neural crest origin of the avian ultimobranchial glandular cells. Interspecific combinations and cytochemical characterization of Ccells based on the uptake of biogenic amine precursors. Histochemie 1974;38:297–305.
5. Fontaine J. Multistep migration of calcitonine cell precursors during ontogeny of the mouse pharynx. Gen Comp Endocrinol 1979;37:81–92.
6. Gorbman A. Comparative anatomy and physiology. In: Ingbar SI, Braverman LE, eds. The thyroid. Philadelphia: JB Lippincott Company;1986:p. 43–52.
7. Lin HY, Harris TL, Flannery MS, Aruffo A, Kaji EH, Gorn A, Kolakowski LF Jr, Lodish HF, Goldring SR. Expression cloning of an adenylate cyclase-coupled calcitonin receptor. Science 1991;254:1022–4.
8. Nicholson GC, Moseley JM, Sexton PM, Mendelsohn FA, Martin TJ. Abundant calcitonin receptors in isolated rat osteoclasts. Biochemical and autoradiographic characterization. J Clin Invest 1986;78:355–60.
9. World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Endocrine Organs (eds DeLellis RA, Lloyd RV, Heitz PU, Eng C) (IARC Press, Lyon, 2004).
10. LiVolsi VA, Asa SL. The demise of follicular carcinoma of the thyroid gland. Thyroid 1994;4:233–6.
11. Kebebew E, Greenspan FS, Clark OH, Woeber KA, McMillan A. Anaplastic thyroid carcinoma. Treatment outcome and prognostic factors. Cancer 2005;103:1330–5.
12. Carcangiu ML, Zampi G, Rosai J. Poorly differentiated (‘insular’) thyroid carcinoma. A reinterpretation of Langhans’ ‘wuchernde Struma’. Am J Surg Pathol 1984;8:655–68.
13. Rodriguez JM, Parrilla P, Moreno A, Sola J, Pinero A, Ortiz S, Soria T. Insular carcinoma: an infrequent subtype of thyroid cancer. J Am Coll Surg 1998;187:503–8.
14. Hedinger C, Williams D, Sobin LH. The WHO Histological classification of thyroid tumours: a commentary on the second edition. Cancer 1989;63:908–11.
15. Wynford-Thomas D. Origin and progression of thyroid epithelial tumours. Cellular and molecular mechanisms. Horm Res 1997;47:145–57.
16. Xing M. BRAF mutation in thyroid cancer. Endocr Relat Cancer 2005;12:245–62.
17. Pierotti MA, Greco A. Oncogenic rearrangements of the NTRK1/NGF receptor. Cancer Lett 2006;232:90–8.

18. Di Renzo MF, Olivero M, Ferro S, Prat M, Bongarzone I, Pilotti S, Belfiore A, Costantino A, Vigneti R, Pierotti MA. Overexpression of the c-MET/HGF receptor gene in human thyroid carcinomas. Oncogene 1992;7:2549-53.
19. Suarez HG, Du Villard JA, Severino M, Caillou B, Schlumberger M, Tubiana M, Parmentier C, and Monier R. Presence of mutations in all three ras genes in human thyroid tumors. Oncogene 1990;5:565–70.
20. Kroll TG, Sarraf P, Pecciarini L, Chen CJ, Mueller E, Spiegelman BM, and Fletcher JA. PAX-8-PPARgamma1 fusion oncogene in human thyroid carcinoma. Science 2000;289:1357–60.
21. Ito T, Seyama T, Mizuno T, Tsuyama N, Hayashi T, Hayashi Y, Dohi K, Nakamura N, Akiyama M. Unique association of p53 mutations with undifferentiated but not differentiated carcinomas of the thyroid gland. Cancer Res 1992;52:1369–71.
22. Donghi R, Longoni A, Pilotti S, Michieli P, Della Porta G, and Pierotti MA. Gene p53 mutations are restricted to poorly differentiated and undifferentiated carcinomas of the thyroid gland. J Clin Invest 1993;91:1753-60.
23. Fagin JA, Matsuo K, Karmakar A, Chen DL, Tang SH, Koeffler HP. High prevalence of mutations of p53 gene in poorly differentiated human thyroid carcinomas. J Clin Invest 1993;91:179-84.
24. Lipshutz RJ, Morris D, Chee M, Hubbell E, Kozal MJ, Shah N, Shen N, Yang R, Fodor SP. Using oligonucleotide probe arrays to access genetic diversity. Biotechniques 1995;19:442–7.
25. Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science 1995;270:467–70.
26. DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, Trent JM. Use of a cDNA microarray to analyse gene expression patterns in human cancer. Nat Genet 1996;14:457-60.
27. Welford SM, Gregg J, Chen E, Garrison D, Sorensen PH, Denny CT, Nelson SF. Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization. Nucleic Acids Res 1998;26:3059–65.
28. Khan J, Simon R, Bittner M, Chen Y, Leighton SB, Pohida T, Smith PD, Jiang Y, Gooden GC, Trent JM, Meltzer PS. Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. Cancer Res 1998;58:5009–13.
29. Agrawal D, Chen T, Irby R, Quackenbush J, Chambers AF, Szabo M, Cantor A, Coppola D, Yeatman TJ. Osteopontin identified as lead marker of colon cancer progression, using pooled sample expression profiling. J Natl Cancer Inst 2002;94:513-21.
30. Rolph MS, Sisavanh M, Liu SM, Mackay CR. Clues to asthma pathogenesis from microarray expression studies. Pharmacol Ther 2006;109:284–94.
31. Erle DJ, Yang YH. Asthma investigators begin to reap the fruits of genomics. Genome Biol 2003;4:232.

32. Syed F, Panettieri RA Jr, Tliba O, Huang C, Li K, Bracht M, Amegadzie B, Griswold D, Li L, Amrani Y. The effect of IL-13 and IL-13R130Q, a naturally occurring IL-13 polymorphism, on the gene expression of human airway smooth muscle cells. Respir Res 2005;6:9.
33. Heymans S, Schroen B, Vermeersch P, Milting H, Gao F, Kassner A, Gillijns H, Herijgers P, Flameng W, Carmeliet P, Van de Werf F, Pinto YM, Janssens S. Increased cardiac expression of tissue inhibitor of metalloproteinase-1 and tissue inhibitor of metalloproteinase-2 is related to cardiac fibrosis and dysfunction in the chronic pressure-overloaded human heart. Circulation 2005;112:1136–44.
34. Ohki R, Yamamoto K, Ueno S, Mano H, Misawa Y, Fuse K, Ikeda U, Shimada K. Gene expression profiling of human atrial myocardium with atrial fibrillation by DNA microarray analysis. Int J Cardiol 2005;102:233–8.
35. Li T, Chen YH, Liu TJ, Jia J, Hampson S, Shan YX, Kibler D, Wang PH. Using DNA microarray to identify Sp1 as a transcriptional regulatory element of insulinlike growth factor 1 in cardiac muscle cells. Circ Res 2003;93:1202–9.
36. Evans SJ, Choudary PV, Vawter MP, Li J, Meador-Woodruff JH, Lopez JF, Burke SM, Thompson RC, Myers RM, Jones EG, Bunney WE, Watson SJ, Akil H. DNA microarray analysis of functionally discrete human brain regions reveals divergent transcriptional profiles. Neurobiol Dis 2003;14:240-50.
37. Pierce A, Small SA. Combining brain imaging with microarray: isolating molecules underlying the physiologic disorders of the brain. Neurochem Res 2004;29:1145–52.
38. Zvara A, Szekeres G, Janka Z, Kelemen JZ, Cimmer C, Santha M, Puskas LG. Over-expression of dopamine D2 receptor and inwardly rectifying potassium channel genes in drug-naive schizophrenic peripheral blood lymphocytes as potential diagnostic markers. Dis Markers 2005;21:61–9.
39. Chin KV, Seifer DB, Feng B, Lin Y, Shih WC. DNA microarray analysis of the expression profiles of luteinized granulosa cells as a function of ovarian reserve. Fertil Steril 2002;77:1214-8.
40. Alon U, Barkai N, Notterman DA, Gish K, Ybarra S, Mack D, Levine AJ. Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci U S A 1999;96:6745-50.
41. Perou CM, Jeffrey SS, van de Rijn M, Rees CA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Zhu SX, Lee JC, Lashkari D, Shalon D, Brown PO, Botstein D. Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. Proc Natl Acad Sci U S A 1999;96:9212–7.
42. Mick R, Ziober BL. c-Met expression in tall cell variant papillary carcinoma of the thyroid. Cancer 2003;98:1386–93.
43. Golub TR, Slonim DK, Tamayo P, Huard C, Gaasenbeek M, Mesirov JP, Coller H, Loh ML, Downing JR, Caligiuri MA, Bloomfield CD, Lander ES. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science 1999;286:531-7.

44. Bloom G, Yang IV, Boulware D, Kwong KY, Coppola D, Eschrich S, Quackenbush J, Yeatman TJ. Multi-platform, multi-site, microarraybased human tumor classification. Am J Pathol 2004;164:9-16.
45. Eschrich S, Yang I, Bloom G, Kwong KY, Boulware D, Cantor A, Coppola D, Kruhoffer M, Aaltonen L, Orntoft TF, Quackenbush J, Yeatman TJ. Molecular staging for survival prediction of colorectal cancer patients. J Clin Oncol 2005;23:3526-35.
46. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Eystein Lonning P, Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci U S A 2001;98:10869–74.
47. Beer DG, Kardia SL, Huang CC, Giordano TJ, Levin AM, Misek DE, Lin L, Chen G, Gharib TG, Thomas DG, Lizyness ML, Kuick R, Hayasaka S, Taylor JM, Iannettoni MD, Orringer MB, Hanash S. Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 2002;8:816-24.
48. Van De Vijver MJ, He YD, Van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R. A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med 2002;347:1999–2009.
49. Kihara C, Tsunoda T, Tanaka T, Yamana H, Furukawa Y, Ono K, Kitahara O, Zembutsu H, Yanagawa R, Hirata K, Takagi T, Nakamura Y. Prediction of sensitivity of esophageal tumors to adjuvant chemotherapy by cDNA microarray analysis of gene-expression profiles. Cancer Res 2001;61:6474–9.
50. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature 2000;403:503-11.
51. Bullinger L, Döhner K, Bair E, Fröhling S, Schlenk RF, Tibshirani R, Döhner H, Pollack JR. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. N Engl J Med. 2004 Apr 15;350:1605-16.
52. Dave SS, Wright G, Tan B, Rosenwald A, Gascoyne RD, Chan WC, Fisher RI, Braziel RM, Rimsza LM, Grogan TM, Miller TP, LeBlanc M, Greiner TC, Weisenburger DD, Lynch JC, Vose J, Armitage JO, Smeland EB, Kvaloy S, Holte H, Delabie J, Connors JM, Lansdorp PM, Ouyang Q, Lister TA, Davies AJ, Norton AJ, Muller-Hermelink HK, Ott G, Campo E, Montserrat E, Wilson WH, Jaffe ES, Simon R, Yang L, Powell J, Zhao H, Goldschmidt N, Chiorazzi M, Staudt LM. Prediction of survival in follicular lymphoma based on molecular features of tumor-infiltrating immune cells. N Engl J Med 2004;351:2159–69.

53. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K, Ferrari M, Egevad L, Rayford W, Bergerheim U, Ekman P, DeMarzo AM, Tibshirani R, Botstein D, Brown PO, Brooks JD, Pollack JR. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci U S A 2004;101:811–6.
54. Valk PJ, Verhaak RG, Beijen MA, Erpelinck CA, Barjesteh Van Waalwijk Van Doorn-Khosrovani S, Boer JM, Beverloo HB, Moorhouse MJ, Van Der Spek PJ, Lowenberg B, Delwel R. Prognostically useful gene-expression profiles in acute myeloid leukemia. N Engl J Med 2004;350:1617–28.
55. Lee JS, Chu IS, Heo J, Calvisi DF, Sun Z, Roskams T, Durnez A, Demetris AJ, Thorgeirsson SS. Classification and prediction of survival in hepatocellular carcinoma by gene expression profiling. Hepatology 2004;40:667–76.
56. Roepman P, Wessels LF, Kettelarij N, Kemmeren P, Miles AJ, Lijnzaad P, Tilanus MG, Koole R, Hordijk GJ, van der Vliet PC, Reinders MJ, Slootweg PJ, Holstege FC. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. Nat Genet 2005;37:182–6.
57. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D, Sondak V. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature 2000; 406: 536-40.
58. Visconti R, Federico A, Coppola V, Pentimalli F, Berlingieri MT, Pallante P, Kruhoffer M, Orntoft TF, Fusco A. Transcriptional profile of Ki-Ras-induced transformation of thyroid cells. Cancer Invest 2007; 25: 256-66.
59. Lee JS, Thorgeirsson SS. Comparative and integrative functional genomics of HCC. Oncogene 2006;25:3801–9.
60. Lossos IS, Czerwinski DK, Alizadeh AA, Wechser MA, Tibshirani R, Botstein D, Levy R. Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med 2004;350:1828–37.
61. Gil J, Bernard D, Martínez D, Beach D. Polycomb CBX7 has a unifying role in cellular lifespan. Nat Cell Biol 2004;6:67-72.
62. Otte AP, Kwaks TH. Gene repression by Polycomb group protein complexes: a distinct complex for every occasion? Curr Opin Genet Dev 2003;13:448–54.
63. Lund AH, van Lohuizen M. Polycomb complexes and silencing mechanisms. Curr Opin Cell Biol 2004;16:239–46.
64. Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. Hierarchical recruitment of polycomb group silencing complexes. Mol Cell 2004;14:637–46.
65. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 2002;298:1039–43.

66. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 2002; 16: 2893-905.
67. Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R, Farnham PJ. Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev 2004; 18: 1592-605.
68. Raaphorst FM, van Kemenade FJ, Blokzijl T, Fieret E, Hamer KM, Satijn DP, Otte AP, Meijer CJ. Coexpression of BMI-1 and EZH2 polycomb group genes in Reed-Sternberg cells of Hodgkin's disease. Am J Pathol 2000; 157: 709-15.
69. Beà S, Tort F, Pinyol M, Puig X, Hernández L, Hernández S, Fernandez PL, van Lohuizen M, Colomer D, Campo E. BMI-1 gene amplification and overexpression in hematological malignancies occur mainly in mantle cell lymphomas. Cancer Res 2001; 61: 2409-12.
70. Van Kemenade FJ, Raaphorst FM, Blokzijl T, Fieret E, Hamer KM, Satijn DP, Otte AP, Meijer CJ. Coexpression of BMI-1 and EZH2 polycomb-group proteins is associated with cycling cells and degree of malignancy in B-cell non-Hodgkin lymphoma. Blood 2001;97:3896–901.
71. Lessard J, Sauvageau G. Bmi-1 determines the proliferative capacity of normal and leukaemic stem cells. Nature 2003;423:255–60.
72. Visser HP, Gunster MJ, Kluin-Nelemans HC, Manders EM, Raaphorst FM, Meijer CJ, Willemze R, Otte AP. The Polycomb group protein EZH2 is upregulated in proliferating, cultured human mantle cell lymphoma. Br J Haematol 2001;112:950–8.
73. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG, Ghosh D, Pienta KJ, Sewalt RGAB, Otte AP, Rubin MA, Chinnaiyan AM. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 2002;419:624–9.
74. Kleer CG, Cao Q, Varambally S, Shen R, Ota I, Tomlins SA, Ghosh D, Sewalt RGAB, Otte AP, Hayes DF, Sabel MS, Livant D, Weiss SJ, Rubin MA, Chinnaiyan AM. EZH2 is a marker of aggressive breast cancer and promotes neoplastic transformation of breast epithelial cells. Proc Natl Acad Sci U S A 2003;100:11606–11.
75. Kanno M, Hasegawa M, Ishida A, Isono K, Taniguchi M. mel-18, a Polycomb group-related mammalian gene, encodes a transcriptional negative regulator with tumor suppressive activity. EMBO J 1995;14:5672–8.
76. Tokimasa S, Ohta H, Sawada A, Matsuda Y, Kim JY, Nishiguchi S, Hara J, Takihara Y. Lack of the Polycomb-group gene rae28 causes maturation arrest at the early B-cell developmental stage. Exp Hematol 2001;29:93–103.
77. Kirmizis A, Bartley SM, Farnham PJ. Identification of the polycomb group protein SU(Z)12 as a potential molecular target for human cancer therapy. Mol Cancer Ther 2003;2:113–21.
78. Nakao M, Minami T, Ueda Y, Sakamoto Y, Ichimura T. Epigenetic system: a pathway to malignancies and a therapeutic target. Int J Hematol 2004;80:103–107.
79. Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. Nature 2004;429:457–63.

80. Raaphorst FM, Vermeer M, Fieret E, Blokzijl T, Dukers D, Sewalt RG, Otte AP, Willemze R, Meijer CJ. Site-specific expression of Polycomb-group genes encoding the HPC-HPH/PRC1 complex in clinically defined primary nodal and cutaneous large B-cell lymphomas. Am J Pathol 2004;164:533–42.
81. Paro R, Hogness DS. The Polycomb protein shares a homologous domain with a heterochromatin-associated protein of Drosophila. Proc Natl Acad Sci U S A 1991;88:263–7.
82. Min J, Zhang Y, Xu R-M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev 2003;17:1823–8.
83. Jones DO, Cowell IG, Singh PB. Mammalian chromodomain proteins: their role in genome organisation and expression. Bioessays 2000;22:124–37.
84. Bernstein E, Duncan EM, Masui O, Gil J, Heard E, Allis CD. Mouse Polycomb Proteins Bind Differentially to Methylated Histone H3 and RNA and Are Enriched in Facultative Heterochromatin. Mol Cell Biol 2006;26:2560-9.
85. Satijn DP, Otte AP. Polycomb group protein complexes: do different complexes regulate distinct target genes? Biochim Biophys Acta 1999;1447:1–16.
86. Satijn DP, Gunster MJ, van der Vlag J, Hamer KM, Schul W, Alkema MJ, Saurin AJ, Freemont PS, van Driel R, Otte AP. RING1 is associated with the polycomb group protein complex and acts as a transcriptional repressor. Mol Cell Biol 1997;17:4105–13.
87. Saurin AJ, Shiels C, Williamson J, Satijn DP, Otte AP, Sheer D, Freemont PS. The human polycomb group complex associates with pericentromeric heterochromatin to form a novel nuclear domain. J Cell Biol 1998;142:887–98.
88. Bernard D, Martinez-Leal JF, Rizzo S, Martinez D, Hudson D, Visakorpi T, Peters G, Carnero A, Beach D, Gil J.CBX7 controls the growth of normal and tumor-derived prostate cells by repressing the Ink4a/Arf locus Oncogene 2005;24:5543–51.
89. Suarez-Merino B, Hubank M, Revesz T, Harkness W, Hayward R, Thompson D, Darling JL, Thomas DG, Warr TJ. Microarray analysis of pediatric ependymoma identifies a cluster of 112 candidate genes including four transcripts at 22q12.1-q13.3. Neuro-oncol 2005;7:20–31.
90. Thiery JP, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol 2006; 7: 131-42.
91. Schock F, Perrimon N. Molecular mechanisms of epithelial morphogenesis. Annu Rev Cell Dev Biol 2002; 18: 463–493.
92. Christ B, Ordahl CP. Early stages of chick somite development. Anat Embryol 1995; 191: 381–396.
93. Funayama N, Sato Y, Matsumoto K, Ogura T, Takahashi Y. Coelom formation: binary decision of the lateral plate mesoderm is controlled by the ectoderm. Development 1999; 126: 4129–4138.

94. Locascio A, Nieto MA. Cell movements during vertebrate development: integrated tissue behaviour versus individual cell migration. Curr Opin Genet Dev 2001;11: 464–469.
95. Nieto MA. The snail superfamily of zinc-finger transcription factors. Nature Rev Mol Cell Biol 2002; 3: 155–166.
96. Thiery JP. Epithelial-mesenchymal transitions in tumour progression. Nat Rev Cancer 2002; 2: 442-54.
97. Barallo-Gimeto A, Nieti MA. The Snail genes as inducers of cell movement and survival: implications in development and cancer. Development 2005; 132: 3151-3161.
98. Sleeman JP. The lymph node as a bridgehead in the metastatic dissemination of tumors. Recent Results Cancer Res 2000; 157: 55-81.
99. Berx G, Cleton-Jansen AM, Nollet F, de Leeuw WJ, van de Vijver M, Cornelisse C, van Roy F. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. EMBO J 1995; 14: 6107-15.
100. Risinger JI, Berchuck A, Kohler MF, Boyd J. Mutations of the E-cadherin gene in human gynecologic cancers. Nature Genet 1994; 7: 98–102.
101. Yoshiura K, Kanai Y, Ochiai A, Shimoyama Y, Sugimura T, Hirohashi S. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. Proc Natl Acad Sci U S A 1995; 92: 7416-9.
102. Hennig G, Behrens J, Truss M, Frisch S, Reichmann E, Birchmeier W. Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo. Oncogene 1995; 11: 475-484.
103. Hennig G, Lowrick O, Birchmeier W, Behrens J. Mechanisms identified in the transcriptional control of epithelial gene expression. J Biol Chem 1996; 271: 595–602.
104. Hajra KM, Ji X, Fearon ER. Extinction of E-cadherin expression in breast cancer via a dominant repression pathway acting on proximal promoter elements. Oncogene 1999: 18: 7274–7279.
105. Rodrigo I, Cato AC, Cano A. Regulation of E-cadherin gene expression during tumor progression: the role of a new Ets-binding site and the E-pal element. Exp Cell Res 1999; 248: 358–371.
106. Tamura G, Yin J, Wang S, Fleisher AS, Zou T, Abraham JM, Kong D, Smolinski KN, Wilson KT, James SP, Silverberg SG, Nishizuka S, Terashima M, Motoyama T, Meltzer SJ. E-Cadherin gene promoter hypermethylation in primary human gastric carcinomas. J Natl Cancer Inst 2000; 92: 569-73.
107. Peinado H, Portillo F, Cano A. Transcriptional regulation of cadherins during development and carcinogenesis. Int J Dev Biol 2004; 48: 365-75.
108. Takeichi M. The cadherins: cell-cell adhesion molecules controlling animal morphogenesis. Development 1988; 102: 639-655.
109. Takeichi M. Morphogenetic roles of classic cadherins. Curr Opin Cell Biol 1995; 7: 619-627.

110. Takeichi M. Cadherins in cancer: implications for invasion and metastasis. Curr Opin Cell Biol 1993; 5: 806-811.
111. Birchmeier W, Behrens J. Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. Biochim Biophys Acta 1994; 1198: 11-26.
112. Christofori G, Semb H. The role of the cell-adhesion molecule Ecadherin as a tumour- suppressor gene. Trends Biochem Sci 1999;24: 73-76.
113. Strathdee, G Epigenetic versus genetic alterations in the inactivation of E-cadherin. Semin Cancer Biol 2002; 12:373-379.
114. Becker KF, Hofler H, Van Roy F. Mutations of the human E-cadherin (CDH1) gene. Hum Mutat 1998; 12: 226-237.
115. Grady WM, Willis J, Guilford PJ, Dunbier AK, Toro TT, Lynch H, Wiesner G, Ferguson K, Eng C, Park JG, Kim SJ, Markowitz S. Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. Nat Genet 2000; 26: 16-7.
116. Cheng CW, Wu PE, Yu JC, Huang CS, Yue CT, Wu CW, Shen CY. Mechanisms of inactivation of E-cadherin in breast carcinoma: modification of the two-hit hypothesis of tumor suppressor gene. Oncogene 2001;20: 3814-3823.
117. Koizume S, Tachibana K, Sekiya T, Hirohashi S, Shiraishi M. (2002) Heterogeneity in the modification and involvement of chromatin components of the CpG island of the silenced human CDH1 gene in cancer cells. Nucleic Acids Res 2002;30: 4770-4780.
118. Darwanto A, Kitazawa R, Maeda S, Kitazawa S. MeCp2 And Promoter Methylation Cooperatively Regulate E-Cadherin Gene Expression In Colorectal Carcinoma. Cancer Sci 2003; 94: 442-447.
119. Behrens J, Lowrick O, Klein-Hitpass L, Birchmeier W. The E-Cadherin Promoter: Functional Analysis Of A G.C-Rich Region And An Epithelial Cellspecific Palindromic Regulatory Element. Proc Natl Acad Sci Usa 1991; 88: 1149511499.
120. Faraldo ML, Rodrigo I, Behrens J, Birchmeier W, Cano A. Analysis Of The E-Cadherin And P-Cadherin Promoters In Murine Keratinocyte Cell Lines From Different Stages Of Mouse Skin Carcinogenesis. Mol Carcinog 1997; 20: 33-47.
121. Batlle E, Sancho E, Franci C, Dominguez D, Monfar M, Baulida J, Garcia De Herreros A. The Transcription Factor Snail Is A Repressor Of E-Cadherin Gene Expression In Epithelial Tumour Cells. Nat Cell Biol 2000; 2: 84-89.
122. Hajra KM, Chen DY, Fearon ER. The Slug Zinc-Finger Protein Represses E-Cadherin In Breast Cancer. Cancer Res 2002; 62; 1613-1618.
123. Cano A, Perez-Moreno MA, Rodrigo I, Locascio A, Blanco MJ, Del Barrio MG, Portillo F, Nieto MA. The Transcription Factor Snail Controls Epithelial-Mesenchymal Transitions By Repressing E-Cadherin Expression. Nat Cell Biol 2000;2: 76-83.
124. Pérez-Moreno MA, Locascio A, Rodrigo I, Dhondt G, Portillo F, Nieto MA, Cano A. A New Role For E12/E47 In The Repression Of E-

Cadherin Expression And Epithelial-Mesenchymal Transitions. J Biol Chem 2001; 276: 27424-27431.
125. Bolós V, Peinado H, Perez-Moreno MA, Fraga MF, Esteller M, Cano, A. The Transcription Factor Slug Represses E-Cadherin Expression And Induces Epithelial To Mesenchymal Transitions: A Comparison With Snail And E47 Repressors. J Cell Sci 2003; 116: 499-511.
126. Grooteclaes ML, Frisch SM. Evidence For A Function Of Ctbp In Epithelial Gene Regulation And Anoikis. Oncogene 2000;19: 3823-3828.
127. Comijn J, Berx G, Vermassen P, Verschueren K, Van Grunsven L, Bruyneel E, Mareel M, Huylebroeck D, Van Roy F. The two-handed e box binding zinc finger protein sip1 downregulates ecadherin and induces invasion. Mol Cell 2001; 7: 1267-1278.
128. Hemavathy K, Ashraf SI, Ip YT. Snail/Slug Family Of Repressors: Slowly Going Into The Fast Lane Of Development And Cancer. Gene 2000; 257: 1-12.
129. Fujita N, Jaye DL, Kajita M, Geigerman C, Moreno CS, Wade PA. Mta3, A Mi-2/Nurd Complex Subunit, Regulates An Invasive Growth Pathway In Breast Cancer Cell 2003; 113: 207-219.
130. Monti M, Orrù S, Pagnozzi D, Pucci P. Interaction proteomics. Biosci Rep. 2005; 25: 45-56.
131. Pallante P, Berlingieri MT, Troncone G, Kruhoffer M, Orntoft TF, Viglietto G, Caleo A, Migliaccio I, Decaussin-Petrucci M, Santoro M, Palombini L, Fusco A. 2005. UbcH10 overexpression may represent a marker of anaplastic thyroid carcinomas. Br J Cancer. 93:464-71.
132. Curcio F, Ambesi-Impiombato FS, Perrella G, Coon HG. Long-term culture and functional characterization of follicular cells from adult normal human thyroids. Proc Natl Acad Sci U S A 1994;91:9004-8.
133. Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, Wolf H, Orntoft TF. Identifying distinct classes of bladder carcinoma using microarrays. Nature Genetics 2003;33:90-6.
134. Obika S, Nanbu D, Hari Y, Morio K, In Y, Ishii JK, Imanishi T. Synthesis of 2_-O,4_-C methyleneuridine and cytidine. Novel bicyclic Nucleosides Having a Fixed C3-endo Sugar Puckering. Tetrahedron Lett 1997;38:8735–8.
135. Koshkin AA, Singh SK, Nielsen P, Rajwanshi VK, Kumar R, Meldgaard M, Olsen CE, Wengel J. LNA (Locked nucleic acids): synthesis of the adenine, cytosine, guanine, 5-methylcytosine, thymine and uracil biocyclonucleoside monomers, oligomerisation and unprecedented nucleic acid recognition. Tetrahedron 1998;54:3607–30.
136. Mouritzen P, Nielsen PS, Jacobsen N, Noerholm M, Lomholt C, Pfundheller HM, Ramsing NB, Kauppinen S, Tolstrup N. The ProbeLibrary™ - Expression profiling 99% of all human genes using only 90 dual-labeled real-time PCR Probes. Biotechniques 2004;37:492–5.
137. Mouritzen P, Noerholm M, Nielsen PS, Jacobsen N, Lomholt C, Pfundheller HM, Tolstrup N. ProbeLibrary™ - Expression profiling 99% of all human genes using only 90 dual-labeled 2005. ProbeLibrary:A new method for faster design and execution of quantitative real-time PCR. Nat Methods 2005;2:313–7.

138. El-Rifai W, Frierson HF Jr, Moskaluk CA, Harper JC, Petroni GR, Bissonette EA, Jones DR, Knuutila S, Powell SM. Genetic differences between adenocarcinomas arising in Barrett's esophagus and gastric mucosa. Gastroenterology 2001;121:592-8.
139. Shevchenko A, Keller P, Scheiffele P, Mann M, Simons K Identification of components of trans-Golgi network-derived transport vesicles and detergent-insoluble complexes by nanoelectrospray tandem mass spectrometry. Electrophoresis 1997;18: 2591-600.
140. Ji X, Woodard AS, Rimm DL, Fearon ER. Transcriptional defects underlie loss of E-cadherin expression in breast cancer. Cell Growth Differ 1997; 8:773-8.
141. Dignam JD, Lebovitz RM, Roeder RG. Nucleic Acids Res 1983; 11: 1475-1489.
142. Pierantoni GM, Rinaldo C, Esposito F, Mottolese M, Soddu S, Fusco A. High Mobility Group A1 (HMGA1) proteins interact with p53 and inhibit its apoptotic activity. Cell Death Differ 2006; 13:1554-63.
143. Michels AA, Nguyen VT, Fraldi A, Labas V, Edwards M, Bonnet F, Lania L, Bensaude O. MAQ1 and 7SK RNA interact with CDK9/cyclin T complexes in a transcription-dependent manner. Mol Cell Biol. 2003; 23: 4859-69.
144. Kruh J. Effects of sodium butyrate, a new pharmacological agent, on cells in culture. Mol Cell Biochem 1982; 42:65–82.
145. Marks P, Rifkind RA, Richon VM, Breslow R, Miller T, Kelly WK. Histone deacetylases and cancer: causes and therapies. Nat Rev Cancer 2001; 1:194-202.
146. Ou JN, Torrisani J, Unterberger A, Provencal N, Shikimi K, Karimi M, Ekstrom TJ, Szyf M. Histone deacetylase inhibitor Trichostatin A induces global and gene-specific DNA demethylation in human cancer cell lines. Biochem Pharmacol 2007; 73: 1297-307.
147. Yoshida M, Kijima M, Akita M and Beppu T. 1990. Potent and specific inhibition of mammalian histone deacetylase both in vivo and in vitro by trichostatin A. J Biol Chem 1990; 265: 17174–17179.
148. Minucci S and Pelicci PG. Histone deacetylase inhibitors and the promise of epigenetic (and more) treatments for cancer. Nat Rev Cancer 2006; 6:38-51.
149. Bedford MT, Richard S. Arginine methylation an emerging regulator of protein function. Mol Cell. 2005; 18: 263-72.
150. Pal S, Sif S. Interplay between chromatin remodelers and protein arginine methyltransferases. J Cell Physiol. 2007; 213: 306-15.
151. Huang S, Litt M, Felsenfeld G. Methylation of histone H4 by arginine methyltransferase PRMT1 is essential in vivo for many subsequent histone modifications. Genes Dev 2005; 19: 1885-93.
152. Tan CP, Nakielny S. Control of the DNA methylation system component MBD2 by protein arginine methylation. Mol Cell Biol. 2006; 26: 7224-35.
153. Kwak YT, Guo J, Prajapati S, Park KJ, Surabhi RM, Miller B, Gehrig P, Gaynor RB.Methylation of SPT5 regulates its interaction with

RNA polymerase II and transcriptional elongation properties. Mol Cell.2003;11: 1055-66.
154. Ivanov D, Kwak YT, Guo J, Gaynor RB.Domains in the SPT5 protein that modulate its transcriptional regulatory properties. Mol Cell Biol 2000; 20: 2970-83.
155. Pal S, Yun R, Datta A, Lacomis L, Erdjument-Bromage H, Kumar J, Tempst P, Sif S. mSin3A/histone deacetylase 2- and PRMT5-containing Brg1 complex is involved in transcriptional repression of the Myc target gene cad. Mol Cell Biol. 2003; 23: 7475-87.
156. Peinado H, Ballestar E, Esteller M, Cano A. Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex. Mol Cell Biol. 2004; 24: 306-19.
157. Le Guezennec X, Vermeulen M, Brinkman AB, Hoeijmakers WA, Cohen A, Lasonder E, Stunnenberg HG.MBD2/NuRD and MBD3/NuRD, two distinct complexes with different biochemical and functional properties. Mol Cell Biol 2006; 26: 843-51.
158. Perreault A, Lemieux C, Bachand F.Regulation of the nuclear poly(A)-binding protein by arginine methylation in fission yeast. J Biol Chem. 2007; 282: 7552-62.

TABLE 2 CBX7 interacting proteins

BAND	PROTEIN ID	score	N° peptidi	MW	Function	Process
T1	(P62826) GTP-binding nuclear protein Ran (GTPase Ran) (Ras-like protein TC4)	27	2	25	GTP-binding protein involved in nucleocytoplasmic transport. Required for the import of protein into the nucleus and also for RNA export. Involved in chromatin condensation and control of cell cycle.Nuclear; becomes dispersed throughout the cytoplasm during mitosis. Interact with RANGAP1	GTPase activity/ Signal transduction ; Cell communication
T4	(P61981) 14-3-3 protein gamma (Protein kinase C inhibitor protein 1)	88	1	28	Adapter protein implicated in the regulation of a large spectrum of both general and specialized signaling pathway. Binds to a large number of partners, usually by recognition of a phosphoserine or phosphothreonine motif. Binding generally results in the modulation of the activity of the binding partner.Interacts with RAF1 and SSH1.Cytoplasm.	Receptor signaling complex scaffold activity/ Signal transduction; cell communication
T5	(O95931) Chromobox protein homolog 7	40	1	28	**BAIT**	
T7	(P04083) Annexin A1 (Annexin I) (Lipocortin I) (Calpactin II) (Chromobindin-9)	785	19	39	Calcium/phospholipid-binding protein which promotes membrane fusion and is involved in exocytosis. This protein regulates phospholipase A2 activity. It seems to bind from two to four calcium ions with high affinity.Phosphorylated by protein kinase C, epidermal growth factor receptor/kinase and TRPM7. Phosphorylation results in loss of the inhibitory activity.	Calcium ion binding/Signal transduction ; Cell communication
	(P07355) Annexin A2 (Annexin II) (Lipocortin II) (Calpactin I heavy chain) (Chromobindin-8)	284	5	39	Calcium-regulated membrane-binding protein whose affinity for calcium is greatly enhanced by anionic phospholipids. It binds two calcium ions with high affinity	Calcium ion binding/Signal transduction ; Cell communication
	(Q08188) Protein-glutamine gamma-glutamyltransferase E precursor	231	6	77	Catalyzes the cross-linking of proteins and the conjugation of polyamines to proteins. It is responsible for the later stages of cell envelope formation in the epidermis and the hair follicle	Transaminase activity/ Metabolism; Cell growth and/or maintenance

gil37183160 HRPE773 [Homo sapiens]	215	7	20	Q96DA0 HUMAN Similar to common salivary protein 1	Unknown
gil662841 heat shock protein 27	147	5	22	Involved in stress resistance and actin organization. Associates with alpha- and beta-tubulin and microtubules. Interacts with HSPB8.Cytoplasm. Nucleus. Cytoplasmic in interphase cells. Colocalizes with mitotic spindles in mitotic cells. Translocates to the nucleus during heat shock.	Chaperone activity/Cell proliferation
gil20334354 S100 calcium binding protein A14	144	2	12	Cytoplasm.	Calcium ion binding/Biological Process Unknown
gil4885607 small proline-rich protein 3 [Homo sapiens]	132	3	19	Cross-linked envelope protein of keratinocytes. Cytoplasm.	Structural molecule activity/ Cell growth and/or maintenance
gil31645 glyceraldehyde-3-phosphate dehydrogenase [Homo sapiens]	110	4	36	It has been suggested that glyceraldehyde-3-phosphate dehydrogenase (GAPDH) play a role in nuclear tRNA export	Catalytic activity/ Nuclear export
gil187302 epithelial cell marker protein 1 (14-3-3 sigma)	90	2	28	Adapter protein implicated in the regulation of a large spectrum of both general and specialized signaling pathway. Binds to a large number of partners, usually by recognition of a phosphoserine or phosphothreonine motif. Binding generally results in the modulation of the activity of the binding partner. p53-regulated inhibitor of G2/M progression. Cytoplasm. Or: Secreted protein. May be secreted by a non-classical secretory pathway. Belongs to the 14-3-3 family.	Receptor signaling complex scaffold activity /Signal transduction ; Cell communication
gil45592961 chromosome 20 open reading frame 70 [Homo sapiens]	76	2	27	containing a lipid binding domain - similar to short palate, lung and nasal epithelium carcinoma-associated protein 2	Unknown/Biological Process Unknown
gil18204442 Ribosomal protein L18	64	1	22	Ribosomal protein	Ribosomal subunit/ Protein metabolism

	gil895845 p64 CLCP	62	1	24	Seems to act as a chloride ion channel.. Mostly in the nucleus including in the nuclear membrane. Cytoplasm. Cell membrane. Small amount in the cytoplasm and the plasma membrane.SIMILARITY: Belongs to the chloride channel CLIC family. SUBCELLULAR LOCATION: Nucleus; nuclear membrane	Transmembrane receptor activity/ receptor
	(P09651) Heterogeneous nuclear ribonucleoprotein A1 (Helix-destabilizing protein)	58	1	39	Involved in the packaging of pre-mRNA into hnRNP particles, transport of poly(A) mRNA from the nucleus to the cytoplasm and may modulate splice site selection. Nucleus. Cytoplasm. Shuttles continuously between the nucleus and the cytoplasm along with mRNA. Component of ribonucleosomes.	Ribonucleoprotein/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism/ Nucleus export
T8	gil627650 transcription factor BTF2 chain p34 - human	56	1	35	FUNCTION: Component of the core-TFIIH basal transcription factor involved in nucleotide excision repair (NER) of DNA and, when complexed to CAK, in RNA transcription by RNA polymerase II. Anchors XPB.SUBUNIT: One of the six subunits forming the core-TFIIH basal transcription factor.SUBCELLULAR LOCATION: Nucleus.SIMILARITY: Belongs to the TFB4 family.	nucleotide-excision repair/ regulation of transcription, DNA-dependent
	(P29692) Elongation factor 1-delta (EF-1-delta)	171	3	31	EF-1-beta and EF-1-delta stimulate the exchange of GDP bound to EF-1-alpha to GTP	Guanyl-nucleotide exchange factor activity/ Translation regulator activity
T9	(Q15024) Exosome complex exonuclease RRP42 (Ribosomal RNA-processing protein 42)	73	1	32	Component of the exosome 3'->5' exoribonuclease complex. Required for the 3' processing of the 7S pre-RNA to the mature 5.8S rRNA. associated with the GTPase Ran. Interacts directly with CSL4. Nucleus; nucleolus	DNA binding/ Cell growth and/or maintenance
	gil34335134 SEC13-like 1 isoform b	45	1	36	Molecular function unknown (Putative nucleoporin interacting protein)	unknown
T10	gil12654583 Ribosomal protein P0	236	5	34	Localization: Citoplasm, nucleus, nucleolus	Ribosomal subunit/ Protein metabolism

T14	(O43684) Mitotic checkpoint protein BUB3	137	4	38	Required for kinetochore localization of BUB1. Nucleus.	protein serine/threonine kinae activity/Signal transduction; Cell communication
T15	(P68400) Casein kinase II subunit alpha	124	5	45	Casein kinases are operationally defined by their preferential utilization of acidic proteins such as caseins as substrates. The alpha and alpha' chains contain the catalytic site. Participates in Wnt signaling.Subunit alpha 1 e 2 interact with HISTONE DEACETILASE 2 (HDAC2)	Protein serine/threonine kinase activity/ Signal transduction ; Cell communication
	(Q01085) Nucleolysin TIAR	118	3	42	RNA-binding protein. Possesses nucleolytic activity against cytotoxic lymphocyte target cells. May be involved in apoptosis. Nuclear localization	RNA binding/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q15366) Poly(rC)-binding protein 2	84	2	39	Major cellular poly(rC)-binding protein. Binds also poly(rU). Nucleus. Cytoplasm. Loosely bound in the nucleus. May shuttle between the nucleus and the cytoplasm	RNA binding/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q14103) Heterogeneous nuclear ribonucleoprotein D0 (hnRNP D0)	76	3	38	Binds with high affinity to RNA molecules that contain AU-rich elements (AREs) found within the 3-prime untranslated regions of many protooncogenes and cytokine mRNAs. Also binds to double- and single-stranded DNA sequences in a specific manner and functions a transcription factor. Binding of RRM1 to DNA inhibits the formation of DNA quadruplex structure which may play a role in telomere elongation. May be involved in translationally coupled mRNA turnover. Implicated with other RNA-binding proteins in the cytoplasmic deadenylation/translational and decay interplay of the FOS mRNA mediated by the major coding-region determinant of instability (mCRD) domain. Nuclear; component of ribonucleosomes.	Ribonucleoprotein/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	gil1808644 arginine methyltransferase	64	2	40	Protein arginine methyltransferase that functions as a histone methyltransferase specific for H4	Methyltransferase activity/ Histone modifing protein

	gil4758158 Septin 2	50	1	42	Septins may form a mitotic scaffold for CENP-E and other effectors to coordinate cytokinesis with chromosome congression and segregation Interact with S100 calcium-binding protein A4 . Nucleus.	Receptor signaling complex scaffold activity/ Signal transduction ; Cell communication
	(O00303) Eukaryotic translation initiation factor 3 subunit 5 (eIF-3 epsilon)	108	2	38	Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA. Associates with the complex p170-eIF3.	Translation regulator activity/ Protein metabolism
T16	(Q12905) Interleukin enhancer-binding factor 2	80	3	43	Appears to function predominantly as a heterodimeric complex with ILF3. This complex may regulate transcription of the IL2 gene during T-cell activation. It can also promote the formation of stable DNA-dependent protein kinase holoenzyme complexes on DNA. Nucleus; nucleolus.	Transcription factor activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
T17	(Q86U42) Polyadenylate-binding protein 2 (Poly(A)-binding protein 2)	57	1	33	Involved in the 3'end formation of mRNA precursors (pre-mRNA) by the addition of a poly(A) tail of 200-250 nt to the upstream cleavage product. Stimulates poly(A) polymerase (PAPOLA) conferring processivity on the poly(A) tail elongation reaction and controls also the poly(A) tail length. Increases the affinity of poly(A) polymerase for RNA. Is also present at various stages of mRNA metabolism including nucleocytoplasmic trafficking and nonsense-mediated decay (NMD) of mRNA. Cooperates with SKIP to synergistically activate E-box-mediated transcription through MYOD1 and may regulate the expression of muscle-specific genes. Binds to poly(A) and to poly(G) with high affinity. May protect the poly(A) tail from degradation.Shuttles between the nucleus and the cytoplasm but predominantly found in the nucleus. Its nuclear import may involve the nucleocytoplasmic transport receptor transportin and a RAN-GTP-sensitive import mechanism. Is exported to the cytoplasm by a carrier-mediated pathway that is independent of mRNA traffic. Nucleus; nucleoplasm; nuclear speckle.	RNA binding/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism/ Nuclear export/import

	(Q8NFW8) N-acylneuraminate cytidylyltransferase	53	2	49	Catalyzes the activation of N-acetylneuraminic acid (NeuNAc) to cytidine 5'-monophosphate N-acetylneuraminic acid (CMP-NeuNAc), a substrate required for the addition of sialic acid. Has some activity toward NeuNAc, N-glycolylneuraminic acid (Neu5Gc) or 2-keto-3-deoxy-D-glycero-D-galacto-nononic acid (KDN). Nucleus. The BC2 (basic cluster 2) motif is necessary and sufficient for the nuclear localization and contains the catalytic active site. The localization in the nucleus is however not required for the enzyme activity	Ligase activity /Metabolism ; Energy pathways
	(P60842) Eukaryotic initiation factor 4A-I	47	1	46	ATP-dependent RNA helicase which is a subunit of the eIF4F complex involved in cap recognition and is required for mRNA binding to ribosome. Belongs to the DEAD box helicase family. eIF4A subfamily.	Translation regulator activity/ Protein metabolism
T18	(P68104) Elongation factor 1-alpha 1 (EF-1-alpha-1)	245	7	50	This protein promotes the GTP-dependent binding of aminoacyl-tRNA to the A-site of ribosomes during protein biosynthesis. Found in a nuclear export complex with XPO5, EEF1A1, Ran and aminoacylated tRNA. Interacts with XPO5.	Translation regulator activity/ Protein metabolism
T20	gil33327055 DERPC [Homo sapiens]	71	2	51	nuclear, strong basic protein; decreased expression in renal and prostate tumors; has cell growth inhibiting function	DNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism/ Cell proliferation
	gil8922679 smu-1 suppressor of mec-8 and unc-52 homolog	47	1	58	unknown	unknown

	gil48146467 HSPC117 [Homo sapiens]	108	3	56	unknown	unknown
T21	gil5901990 katanin p60 subunit A 1 [Homo sapiens]	51	2	56	Microtubules, polymers of alpha and beta tubulin subunits, form the mitotic spindle of a dividing cell and help to organize membranous organelles during interphase. Katanin is a heterodimer that consists of a 60 kDa ATPase (p60 subunit A 1) and an 80 kDa accessory protein (p80 subunit B 1). The p60 subunit acts to sever and disassemble microtubules, while the p80 subunit targets the enzyme to the centrosome. This gene encodes the p80 subunit. This protein is a member of the AAA family of ATPases.	ATPase activity/Metabolism; Cell maintenance
	gil47940653 Chaperonin containing TCP1, subunit 8 (theta)	68	2	60		Chaperone activity/ Cell maintenance
T22	gil56789228 Chaperonin containing TCP1, subunit 7 (eta) [Homo sapiens]	53	1	60		Chaperone activity/ Cell maintenance
	gil71051977 HDAC2 protein [Homo sapiens]	65	2	45	Histone deacetylase family. Deacetylase activity. Interact with casein kinase 2, alpha 1 polypeptide	Histone modifing protein
T23	(P14866) Heterogeneous nuclear ribonucleoprotein L (hnRNP L)	267	11	61	This protein is a component of the heterogenous nuclear ribonucleoprotein (hnRNP) complexes which provide the substrate for the processing events that pre-mRNAs undergo before becoming functional, translatable mRNAs in the cytoplasm. L is associated with most nascent transcripts including those of the landmark giant loops of amphibian lampbrush chromosomes. Nucleus; nucleoplasm	Ribonucleoprotein/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism

	(Q13283) Ras-GTPase-activating protein-binding protein 1 (ATP-dependent DNA helicase V)	180	4	52	May be a regulated effector of stress granule assembly. Phosphorylation-dependent sequence-specific endoribonuclease in vitro. Binds to the SH3 domain of Ras-GTPase-activating protein (RASA1) in proliferating cells. No interaction in quiescent cells.Interacts with USP10, and may regulate it. Forms homodimers and oligomers. Cytoplasm. Cytoplasm; cytosol. Cell membrane. Nucleus. Cytoplasmic in proliferating cells, can be recruited to the plasma membrane in exponentially growing cells. Cytosolic and partially nuclear in resting cells. Recruited to stress granules (SGs) upon either arsenite or high temperature treatment. Recruitment to SGs is influenced by HRAS.	Ribonuclease activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q12800) Alpha-globin transcription factor CP2 (Transcription factor LSF)	166	3	58	Binds a variety of cellular and viral promoters including fibrinogen, alpha-globin, SV40 and HIV-1 promoters. Activation of the alpha-globin promoter in erythroid cells is via synergistic interaction with UBP1. Functions as part of the SSP (stage selector protein) complex. Facilitates the interaction of the gamma-globin genes with enhancer elements contained in the locus control region in fetal erythroid cells. Interacts by binding to the stage selector element (SSE) in the proximal gamma-globin promoter.	Transcription factor activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	gil1809248 siah binding protein 1 [Homo sapiens]	107	2	58	Hypotetical Ribonucleoprotei/ RNA binding protein	RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	gil14124984 Chaperonin containing TCP1, subunit 3 (gamma) [Homo sapiens]	61	1	60		Chaperone activity/ Cell maintenance
T24	(P33240) Cleavage stimulation factor, 64 kDa subunit (CSTF 64 kDa subunit)	324	8	61	One of the multiple factors required for polyadenylation and 3'-end cleavage of mammalian pre-mRNAs. This subunit is directly involved in the binding to pre-mRNAs.Nucleus.	RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism

	gil57209129 paraspeckle component 1	199	5	59	Putative RNA binding protein	RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	gil23398584 Interferon regulatory factor 2 binding protein 1	80	1	62		Transcription regulator activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
T25	gil8923040 collaborates/cooperates with ARF (alternate reading frame) protein	74	1	61		RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q8WXI9) Transcriptional repressor p66 beta (p66/p68) (GATA zinc finger domain-containing protein 2)	47	1	66	Has transcriptional repressor activity. Targets MBD3 to discrete loci in the nucleus. Nucleus. Nuclear, in discrete foci.	Transcription regulator activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
T26c	gil178027 alpha-actin	258	9	42		Cell maintenance
	(P11940) Polyadenylate-binding protein 1 (Poly(A)-binding protein 1) (PABP 1)	48	1	71	Binds the poly(A) tail of mRNA. May be involved in cytoplasmic regulatory processes of mRNA metabolism. Its function in translational initiation regulation can either be enhanced by PAIP1 or repressed by PAIP2. Can probably bind to cytoplasmic RNA sequences other than poly(A) in vivo. May be involved in translationally coupled mRNA turnover. Implicated with other RNA-binding proteins in the cytoplasmic deadenylation/translational and decay interplay of the FOS mRNA mediated by the major coding-region determinant of instability (mCRD) domain. Cytoplasm. Nucleus. Shuttles between the cytoplasm and the nucleus.	RNA binding/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism

	(Q9BTC8) Metastasis-associated protein MTA3	58	2	68	Plays a role in maintenance of the normal epithelial architecture through the repression of SNAI1 transcription in a histone deacetylase-dependent manner, and thus the regulation of E-cadherin levels.Nucleus. Interact with histone deacetilase 2 non directly	Transcription regulator activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q9UBU9) Nuclear RNA export factor 1 (Tip-associating protein)	56	1	71	Involved in the nuclear export of mRNA species bearing retroviral constitutive transport elements (CTE) and in the export of mRNA from the nucleus to the cytoplasm. Nuclear; localized predominantly in the nucleoplasm and at both the nucleoplasmic and cytoplasmic faces of the nuclear pore complex. Shuttles between the nucleus and the cytoplasm.	RNA binding/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	gil57165052 Thyroid autoantigen 70kDa	53	1	70	Single stranded DNA-dependent ATP-dependent helicase. Has a role in chromosome translocation. The DNA helicase II complex binds preferentially to fork-like ends of double-stranded DNA in a cell cycle-dependent manner. It works in the 3'-5' direction. Binding to DNA may be mediated by p70. Involved in DNA nonhomologous end joining (NHEJ) required for double-strand break repair and V(D)J recombination. The Ku p70/p86 dimer acts as regulatory subunit of the DNA-dependent protein kinase complex DNA-PK by increasing the affinity of the catalytic subunit PRKDC to DNA by 100-fold. The Ku p70/p86 dimer is probably involved in stabilizing broken DNA ends and bringing them together. The assembly of the DNA-PK complex to DNA ends is required for the NHEJ ligation step. Required for osteocalcin gene expression	DNA repair protein/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
T27	gil15530265 Eukaryotic translation elongation factor 1 gamma	320	8	50		Translation regulator activity/ Protein metabolism
T28	gil34783356 RANGAP1 protein	99	2	64	GTPase activator for the nuclear Ras-related regulatory protein Ran, converting it to the putatively inactive GDP-bound state.	GTPase activator activity/ Signal transduction ; Cell communication

	gil14250668 Cortactin, isoform b	106	3	58		Cytoskeletal protein binding/Cell growth and/or maintenance
	gil82546879 GTP binding protein 1	106	2	73	FUNCTION: Not known, the C-terminus can act as an allosteric activator of the cholera toxin catalytic subunit.	GTPase activity/ Signal transduction ; Cell communication
T30	(Q14157) Ubiquitin-associated protein 2-like	97	3	104	Unclassified	Unknown/Biological_process unknown
	(P19447) TFIIH basal transcription factor complex helicase XPB subunit	97	2	90	ATP-dependent 3'-5' DNA helicase, component of the core-TFIIH basal transcription factor, involved in nucleotide excision repair (NER) of DNA and, when complexed to CAK, in RNA transcription by RNA polymerase II. Acts by opening DNA either around the RNA transcription start site or the DNA damage.	DNA repair protein/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q9BSJ2) Gamma-tubulin complex component 2	69	1	104	Gamma-tubulin complex is necessary for microtublule nucleation at the centrosome. Centrosome.	Cytoskeletal protein binding/ Cell growth and/or maintenance
	(Q12906) Interleukin enhancer-binding factor 3 (Nuclear factor of activated T-cells 90 kDa)	64	1	96	May facilitate double-stranded RNA-regulated gene expression at the level of post-transcription. Can act as a translation inhibitory protein which binds to coding sequences of acid beta-glucosidase (GCase) and other mRNAs and functions at the initiation phase of GCase mRNA translation, probably by inhibiting its binding to polysomes. Can regulate protein arginine N-methyltransferase 1 activity. May regulate transcription of the IL2 gene during T-cell activation. Can promote the formation of stable DNA-dependent protein kinase holoenzyme complexes on DNA. Nucleus; nucleolus.	Transcription factor activity/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism

	(Q01826) DNA-binding protein SATB1 (Special AT-rich sequence-binding protein 1)	60	1	86	Binds to DNA at special AT-rich sequences at nuclear matrix- or scaffold-associated regions. Thought to recognize the sugar-phosphate structure of double-stranded DNA. Nucleus.	Transcription factor activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(P33992) DNA replication licensing factor MCM5 (CDC46 homolog) (P1-CDC46)	46	2	83		DNA binding /Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	gil2494898 Periodic tryptophan protein 2 homolog	68	2	103	Nucleus; nucleolus. PWP2 is essential for viability and may play a role in the early G1 phase of the cell cycle.(La prot. Possiede diversi domini WD40)	Regulation of cell cycle/ Signal transduction ; Cell communication
T31						
	(Q02779) Mitogen-activated protein kinase kinase kinase 10	40	2	104	Activates the JUN N-terminal pathway	Protein serine-threonine kinase activity/ Signal transduction ; Cell communication
T32	(P43243) Matrin-3	160	6	95	May play a role in transcription or may interact with other nuclear matrix proteins to form the internal fibrogranular network. In association with the SFPQ-NONO heteromer may play a role in nuclear retention of defective RNAs. Nucleus; nucleoplasm; nuclear matrix.	RNA binding/Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(P11586) C-1-tetrahydrofolate synthase, cytoplasmic (C1-THF synthase)	110	2	102	Cytoplasm	Catalytic activity/ Metabolism; Energy pathways

	(P55884) Eukaryotic translation initiation factor 3 subunit 9	62	2	93	Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA	Translation regulator activity/ Protein metabolism
	(Q9H0D6) 5'-3' exoribonuclease 2 (EC 3.1.11.-) (DHM1-like protein)	61	1	109	Possesses 5'->3' exoribonuclease activity and may be involved in homologous recombination and RNA metabolism, such as RNA synthesis and RNA trafficking .Nucleus; nucleolus.	Ribonuclease activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(Q99613) Eukaryotic translation initiation factor 3 subunit 8 (eIF3 p110)	57	1	106	Binds to the 40S ribosome and promotes the binding of methionyl-tRNAi and mRNA	Translation regulator activity/ Protein metabolism
T33	(Q9ULK4) CRSP complex subunit 3 (Cofactor required for Sp1 transcriptional activation subunit 3)	102	2	158	Plays a role in transcriptional coactivation.	Transcription regulator activity/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
	(P26640) Valyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase)	88	2	142		Translation regulator activity/ Protein metabolism

(Q92900) Regulator of nonsense transcripts 1 (EC 3.6.1.-) (ATP-dependent helicase RENT1)	70	2	126	Part of a post-splicing multiprotein complex. Involved in nonsense-mediated decay (NMD) of mRNAs containing premature stop codons. Essential for embryonic viability.	RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
(O00267) Transcription elongation factor SPT5 (hSPT5)	56	2	121	Component of the DRB sensitivity-inducing factor complex (DSIF complex), which regulates mRNA processing and transcription elongation by RNA polymerase II. DSIF positively regulates mRNA capping by stimulating the mRNA guanylyltransferase activity of RNGTT/CAP1A. DSIF also acts cooperatively with the negative elongation factor complex (NELF complex) to enhance transcriptional pausing at sites proximal to the promoter. NUCLEO	transcription
(Q13435) Splicing factor 3B subunit 2 (Spliceosome-associated protein 145)	56	1	98	Subunit of the splicing factor SF3B required for 'A' complex assembly formed by the stable binding of U2 snRNP to the branchpoint sequence (BPS) in pre-mRNA. Sequence independent binding of SF3A/SF3B complex upstream of the branch site is essential, it may anchor U2 snRNP to the pre-mRNA. May also be involved in the assembly of the 'E' complex. Belongs also to the minor U12-dependent spliceosome, which is involved in the splicing of rare class of nuclear pre-mRNA intron. Nucleus.	RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
(O75533) Splicing factor 3B subunit 1 (Spliceosome-associated protein 155	52	2	146	Subunit of the splicing factor SF3B required for 'A' complex assembly formed by the stable binding of U2 snRNP to the branchpoint sequence (BPS) in pre-mRNA. Sequence independent binding of SF3A/SF3B complex upstream of the branch site is essential, it may anchor U2 snRNP to the pre-mRNA. May also be involved in the assembly of the 'E' complex. Belongs also to the minor U12-dependent spliceosome, which is involved in the splicing of rare class of nuclear pre-mRNA intron.	RNA binding/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism

	gil89047296 similar to SMC hinge domain containing 1	136	3	221		Unknown
T34						
	gil33946327 nucleoporin 214kDa [Homo sapiens]	54	3	214	integral to membrane.Interact with con Nuclear RNA export factor 1	porin activity/ protein import/ export
	gil71044479 death inducer-obliterator 1 isoform c	358	12	245	Apoptosis, a major form of cell death, is an efficient mechanism for eliminating unwanted cells and is of central importance for development and homeostasis in metazoan animals. In mice, the death inducer-obliterator-1 gene is upregulated by apoptotic signals and encodes a cytoplasmic protein that translocates to the nucleus upon apoptotic signal activation. When overexpressed, the mouse protein induced apoptosis in cell lines growing in vitro. This gene is similar to the mouse gene and therefore is thought to be involved in apoptosis. Alternatively spliced transcripts have been found for this gene, encoding multiple isoforms.	Putative DNA binding protein/ Cell proliferation/ Apoptosis/ Regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolism
T35						
	gil13606056 DNA dependent protein kinase catalytic subunit [Homo sapiens]	106	4	470	Serine/threonine-protein kinase that acts as a molecular sensor for DNA damage. Involved in DNA nonhomologous end joining (NHEJ) required for double-strand break (DSB) repair and V(D)J recombination. Must be bound to DNA to express its catalytic properties. Promotes processing of hairpin DNA structures in V(D)J recombination by activation of the hairpin endonuclease artemis (DCLRE1C). The assembly of the DNA-PK complex at DNA ends is also required for the NHEJ ligation step. Required to protect and align broken ends of DNA. May also act as a scaffold protein to aid the localization of DNA repair proteins to the site of damage. Found at the ends of chromosomes, suggesting a further role in the maintenance of telomeric stability and the prevention of chromosomal end fusion. Also involved in modulation of transcription. Recognizes the substrate consensus sequence [S/T-Q]. Phosphorylates Ser-139 of histone variant H2AX/H2AFX, thereby regulating DNA damage response mechanism.	protein kinase activity/DNA REPAIR