Chapter 5

Phenotype-specific DMR (pDMR) screen

5.1 Introduction

The MHC is associated with many complex diseases including infectious, autoimmune and inflammatory diseases as well as cancer. In many cases, their aetiologies are polygenic and involve genetic, epigenetic and environmental factors (de Bakker et al., 2006; Garcia-Lora et al., 2003; Rioux and Abbas, 2005; Vyse and Todd, 1996). Although past studies have generated extensive data for the genetics of the MHC (Horton et al., 2008; Nejentsev et al., 2007; Stewart et al., 2004; Traherne et al., 2006) resulting in important contributions to medicine, further studies are necessary to better understand the causes of such diseases.

My main aim was to elucidate the role of differentially methylated regions (DMRs) (see chapter 1) within the MHC in the context of a phenotype that is associated with defects in the MHC class I processing and presentation pathway (Chang et al., 2003; Groothuis et al., 2005; Parham, 2005).

MHC class I molecules have two critical functions: 1. to bind small (8-10mer) peptides derived from protein antigens, and 2. to present bound peptides to T-cell receptors (Flutter and Gao, 2004; Held and Mariuzza, 2008). They are cell surface glycoproteins consisting of two subunits (figure 5.1): a highly polymorphic heavy chain (α-chain) encoded by one of the three classical MHC loci *HLA-A*, *HLA-B* and *HLA-C* and a non-polymorphic light chain (β–chain) called β2-microglobulin (B2M) encoded outside the MHC. Proper folding of MHC class I molecules requires the formation of three disulfide bonds, one in the α3 immunoglobulin domain, one in the α2 immunoglobulin domain and one in the β chain (figure 5.1).

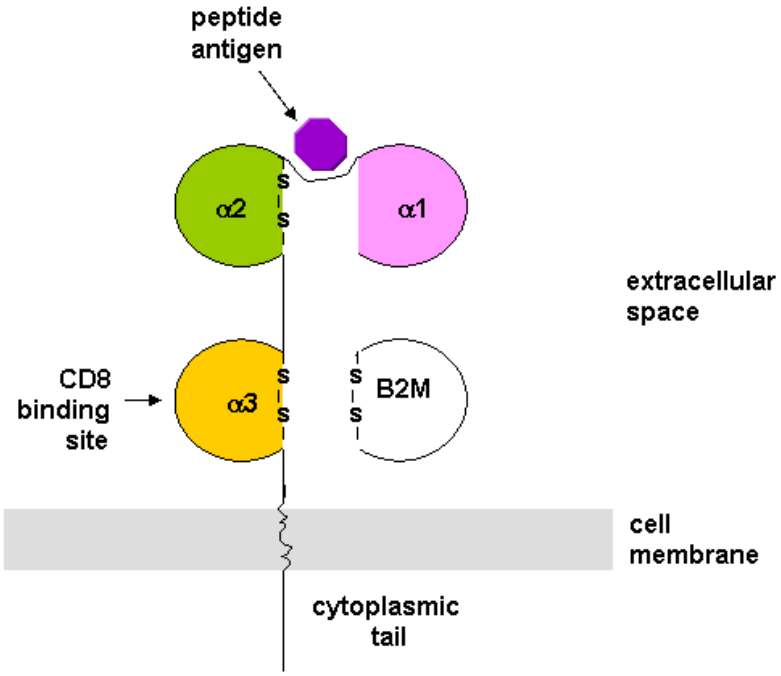

Figure 5.1. **MHC class I molecule.** Diagram of the MHC class I heavy chain associated with B2M and a peptide antigen presented on the cell surface. The heavy chain, also referred as α chain, is a 43kDa transmembrane glycoprotein consisting of three extracellular domains α1, α2 and α3. α1 and α2 are polymorphic and form a deep grove where the peptide antigen can bind. B2M which represents the light chain, has a molecular weight of 12 kDa and is non-polymorphic. The disulfide bonds (-S-S-) are also shown.

The peptide-MHC class I complexes are assembled in the endoplasmatic reticulum (ER) while going through a multifactorial pathway called the MHC class I pathway (Hewitt, 2003; Pamer and Cresswell, 1998). The latter involves at least eight components in addition to HLA-A, HLA-B, HLA-C and B2M: PSMB8, PSMB9, TAP1, TAP2 and TAPBP which are encoded within the MHC as well as the non-MHC encoded proteins ERp57, calnexin (CANX) and calreticulin (CALR) (figure 5.2). The components of the MHC class I pathway are discussed in more detail in sections 5.3 and 6.2.

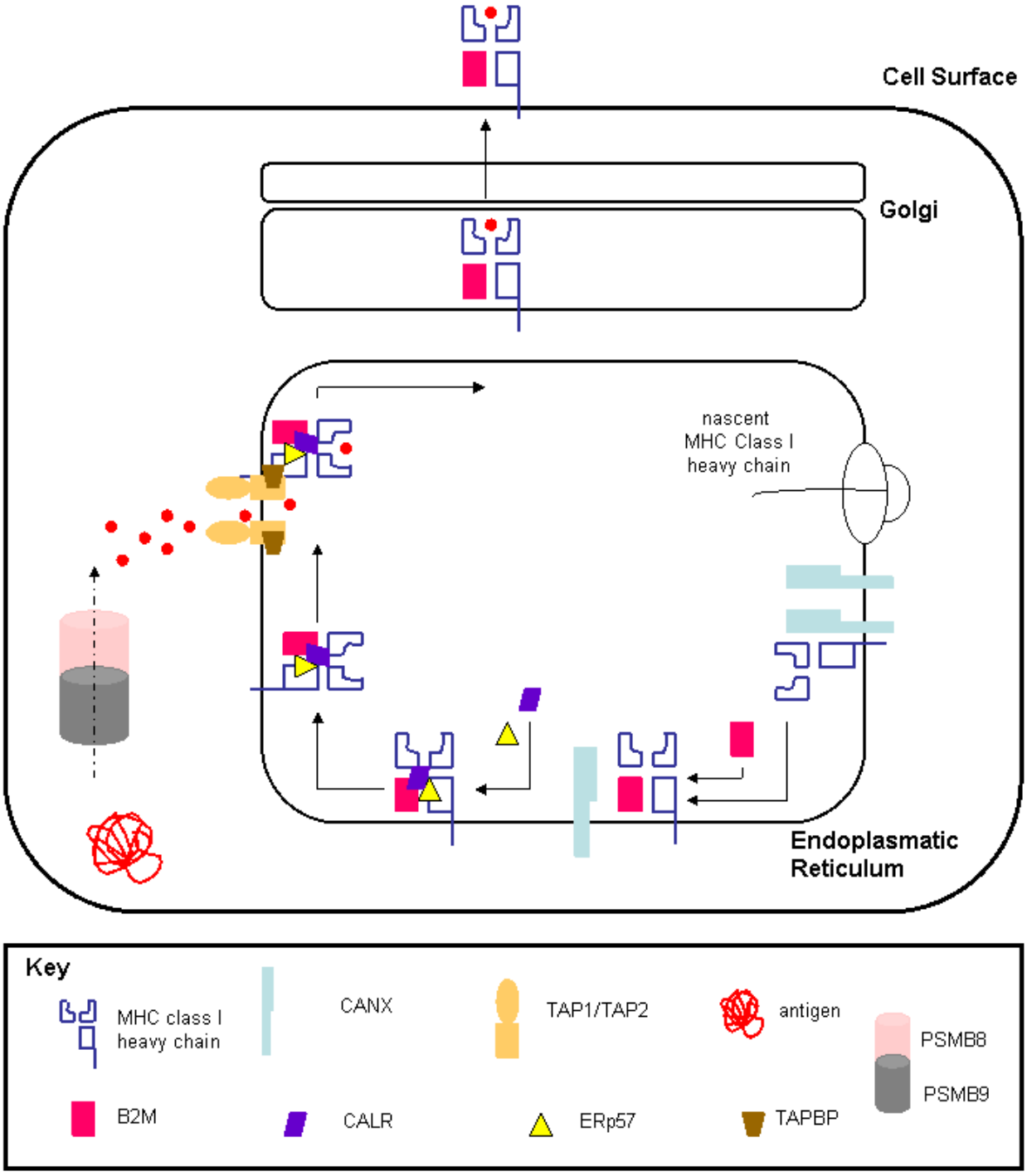

Figure 5.2 **MHC class I antigen presentation pathway.** Antigens are proteolytically processed in the cytosol by the proteasome (PSMB8/PSMB9). Peptides generated by the proteasome are translocated into the ER lumen by TAP. MHC class I molecules (heavy chain and associated B2M) fold and assemble in the ER lumen with the aid of the ER chaperones CANX, CALR and ERp57. The complex of MHC class I

molecules, chaperones, TAP and TAPBP facilitates peptide binding. Peptide-loaded MHC class I molecules dissociate from TAP and are transported through the secretory pathway to the cell surface.

In cases where any of the MHC class I pathway components are defective, intracellular MHC class I molecules (figure 5.1) can be subjected to unfolding and degradation resulting into MHC class I^- phenotypes (Hughes et al., 1997). Such phenotypes have been reported to be associated with many MHC related diseases including cancer where immune evasion can occur through repression of antigen presentation (Chang et al., 2003; de Visser et al., 2006). Although the MHC class I^- phenotype has been studied extensively, the underlying mechanism(s) remain unclear. Recent studies have implicated DNA methylation in the expression of MHC class I genes in cancer but their analysis was restricted to promoter regions only (Fonsatti et al., 2007; Fonsatti et al., 2003; Manning et al., 2008; Nie et al., 2001).

I aimed to pursue a more comprehensive study of the role of DNA methylation in the MHC class I pathway. For this, I used the samples described in section 5.2 and performed an analysis that can be divided into three parts:

Part 1. Expression analysis of the genes *HLA-A, HLA-B, HLA-C, TAP1, TAP2, PSMB8, PSMB9* and *TAPBP* which I call candidate genes (section 5.3). These genes are encoded within the MHC and are involved in the MHC class I antigen presentation pathway. The effect of inhibition of DNA methylation in the expression of these eight genes was also tested (section 5.4)

Part 2. Identification of differentially methylated regions (DMRs) between cell lines displaying the MHC class I^- phenotype (nine cancer cell lines) and two control cell lines (sections 5.5.1 and 5.5.2). To this end I used the MeDIP-MHC tiling array approach (chapter 3).

Part 3. Data generated under parts 1 and 2 were combined to identify:

(i). DMRs that can be associated with the MHC class I^- phenotype (section 5.5.3). I call these DMRs phenotype-specific DMRs (pDMRs).

(ii). DMRs overlapping with the coding regions of the eight candidate genes that may be hypermentylated as the result of low transcriptional levels (section 5.6). Low transcriptional activity has been reported to drive DNA hypermethylation in some instances (Bird, 2002; Meissner et al., 2008).

(iii) Prominent DMRs within the MHC region which, although not associated with the MHC class I^- phenotype, may be important for the regulation of MHC genes in general (section 5.7).

This analysis is described in more detail in the following sections.

5.2 Samples used for pDMR screen

For the pDMR screen I used samples with abnormal expression levels of one or more of the genes involved in the MHC class I antigen presentation pathway. I used cancer cell lines for two reasons: i. MHC class I^- phenotype is known to be associated with many cancer types; ii cancer cell lines were easy to obtain and work with. Based on published data (Blanchet et al., 1992) and availability, nine cancer cell lines were chosen for this screen. As positive controls for MHC class I pathway gene expression, two peripheral blood EBV-transformed cell lines (GM15510 and GM10851) were selected. Table 5.1 shows the characteristics of each of the cell lines chosen.

	Ethnicity	Age	Gender	Tissue	Disease	Levels of MHC class I expression
T47D		54	female	duct	carcinoma	low
578T	Caucasian	74	female	breast	carcinoma	
H69	Caucasian	55	male	lung	carcinoma	low
Colo-205	Caucasian	70	male	colon	adenocarcinoma	high
CCRF-CEM	Caucasian	4	female	blood	leukemia	low
MCF7	Caucasian	69	female	breast	adenocarcinoma	
MDA-MB-231	Caucasian	51	female	breast	adenocarcinoma	high
MDA-MB-361	Caucasian	40	female	breast	adenocarcinoma	intermediate
K562	Caucasian	53	female	bone marrow	leukemia	low
GM15510				blood	-	normal
GM10851	Caucasian	52	male	blood	-	normal

Table 5.1 **Characteristics of cell lines used in the pDMR screen.** The indicated expression status is based on published protein analysis (Blanchet et al., 1992). Empty boxes indicate that the corresponding information was not available.

MHC class I molecules are expressed in a wide range of tissues and cell types, including the tissues of origin of the cell lines selected here (Lechler, 2000). Although cancer cell lines are known to frequently display karyotypic variability (Roschke et al., 2003) this should not affect the methylation analysis as my experimental design (MeDIP-enriched versus total DNA) normalizes for the given genotype (chapter 3).

Finally, it should be noted that the two EBV transformed cell lines (GM15510 and GM10851), although normal with respect to the MHC class I^- phenotype, are expected to have modified methylation patterns compared to the primary cells from which they have been established. EBV virus has been shown to activate DNA methyltransferase activity by increasing the expression of DNMTs (Flanagan, 2007; Tsai et al., 2002).

5.3 Relative expression of MHC class I pathway genes encoded within the MHC

There are eight genes involved in the MHC class I pathway that are encoded within the MHC region: *HLA-A, HLA-B, HLA-C, TAP1, TAP2, PSMB8, PSMB9* and *TAPBP*. All eight genes are encoded within the MHC. Previous studies have shown that genes involved in the same pathway (e.g. members of the TGF-beta signalling pathway) as well as genes encoded within the same chromosomal band (e.g. 4Mb band of chromosome 2q.14.2) can be epigenetically silenced (Frigola et al., 2006; Hinshelwood et al., 2007). These studies support the notion that the eight genes investigated here maybe regulated concordantly by epigenetic mechanisms. These genes are discussed in more detail below:

HLA-A, HLA-B and *HLA-C* genes

Each of these genes encodes an α-chain of a class I molecule (figure 5.1). They are the most polymorphic human loci known to date (de Bakker et al., 2006; Horton et al., 2008; Traherne et al., 2006). It has been proposed previously that promoter methylation of these three loci may be involved in their down-regulation in cancer cells (Nie et al., 2001; Serrano et al., 2001).

TAP and *PSMB* genes

These genes are encoded within a tight cluster in the MHC class II region. The products of the two *TAP* genes, *TAP1* and *TAP2*, are members of the ATP-binding cassette (ABC) transporter superfamily (Townsend and Trowsdale, 1993). They form a complex in the endoplasmatic reticulum (ER) membrane that translocates peptide antigens from the cytoplasm into the lumen of the ER (figure 5.2) (Androlewicz and Cresswell, 1994; Kelly et al., 1992).

The *PSMB* genes, *PSMB8* and *PSMB9*, encode components of the proteasome (figure 5.2) (Gaczynska et al., 1993; Glynne et al., 1991). *PSMB9* has been implicated in proteolytic digestion of cytoplasmic proteins. Production of *PSMB8* and *PSMB9* components has been shown to alter the proteolytic activity of the proteasome to favour antigen peptides capable of binding to the peptide groove of MHC class I molecules (figures 5.1 and 5.2).

TAP1 and *PSMB9* share a bidirectional promoter (Wright et al., 1995). This is a minimal 593bp region which is sufficient for concurrent expression in both directions which implicates that these two genes are controlled simultaneously by common elements.

In cancer cells and human papilloma virus 16-associated tumours, epigenetic induction of MHC class I surface expression has been shown to be associated with the up-regulation of the following genes: *TAP1*, *TAP2*, *PSMB8* and *PSMB9* (Manning et al., 2008; Setiadi et al., 2007).

TAPBP

TAPBP is encoded within the class II region of the MHC. Its product acts as a chaperone that facilitates the association of MHC class I molecules with peptide antigens (figure 5.2) (Lauvau et

al., 1999). *TAPBP* has been shown to be epigenetically regulated in melanoma cells (Khan et al., 2008)

5.3.1 Expression analysis

According to published data MHC class I molecules show differential levels of expression and abundance at the surface of the cells I chose to use (Blanchet et al., 1992) (table 5.1). I analysed further the expression levels of all MHC encoded genes (eight genes) involved in the MHC class I pathway. To this end I performed quantitative real time PCR and calculated the difference in total mRNA levels of each of the eight genes between the cancer and the two normal EBV-transformed cell lines (shared controls). At this point it should be noted that all these genes have multiple isoforms. In order to simplify the analysis, primers were designed to capture all possible isoforms of each of the eight genes. Analysis of the data was done as described in section 2.2.4.3.

According to my data, *HLA-A, HLA-B, TAP1* and *PSMB8* were down-regulated in all cancer cell lines, in both biological replicates (figure 5.3), relative to the two controls (fold difference >1.5). *HLA-C* mRNA levels were close to normal in three cancer cell lines (CCRF-CEM, Colo-205 and MDA-MB-361; fold difference <1.5) and reduced in the rest six. *TAP2*, *PSMB9* and *TAPBP* levels were normal only in the T47D cell line and down regulated in the rest, apart from *PSMB9* which was the only gene that showed about 2-fold up-regulation in one cancer cell line (CCRF-CEM). Expression data are shown in figure 5.3 and are summarised in table 5.2.

With few exceptions, the eight MHC genes tested in this section seem to be co-ordinately down-regulated in almost all cancer cell lines tested. This finding agrees with previous publications (Johnsen et al., 1998; Meissner et al., 2005; Romero et al., 2005) and indicates that the eight MHC genes may share a common regulatory pathway. This is further supported by studies

showing that treatment of antigen presenting cells with the cytokines INF-γ or TNF-α induces coordinated changes at different steps of the MHC class I processing and presentation pathway (Dejardin et al., 1998). In addition, it has been shown previously that treatment of cancer cell lines with DNA methylation inhibitors resulted in up-regulation of MHC class I molecules, suggesting that DNA methylation plays a role in the regulation of MHC class I pathway genes (Fonsatti et al., 2007; Fonsatti et al., 2003; Nie et al., 2001).

Epigenetic modifications have the potential to target expression of multiple genes simultaneously. This can be done either by the epigenetic silencing or up-regulation of a common regulatory factor involved in expression of all genes that are part of the same pathway, or by simultaneous aberrant alterations of epigenetic marks at loci of multiple genes involved in the same pathway or encoded in the same chromosomal region (Frigola et al., 2006; Hinshelwood et al., 2007). I attempted to investigate the role of DNA methylation in the concordant silencing of genes involved in the MHC class I pathway as described in the following sections.

	HLA-A	***HLA-B***	***HLA-C***	***TAP1***	***TAP2***	***PSMB8***	***PSMB9***	***TAPBP***
T47D	**-**	**-**	**-**	**-**	**+/-**	**-**	**+/-**	**+/-**
MDA-MB-231	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
MDA-MB-361	**-**	**-**	**+/-**	**-**	**-**	**-**	**-**	**-**
CCRF-CEM	**-**	**-**	**+/-**	**-**	**-**	**-**	**+**	**-**
Colo-205	**-**	**-**	**+/-**	**-**	**-**	**-**	**-**	**-**
H69	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
578T	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
K562	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
MCF7	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**

Table 5.2. **Summary of MHC encoded MHC class I pathway gene expression.** Relative expression levels were calculated for all cancer cell lines and each of the eight genes involved in the class I pathway compared to two normal control cell lines. Symbol – indicates reduced expression levels; + indicates upregulation and +/- denotes that the expression of the corresponding gene is similar in both cancer and

control cell lines. Only differences greater than 1.5 fold (in both biological replicates, figure 5.3) were considered to be significant.

Figure 5.3. **Relative expression of MHC encoded MHC class I pathway genes.** mRNA levels of *HLA-A, HLA-B, HLA-C, TAP1, TAP2, PSMB8, PSMB9* and *TAPBP* were determined by quantitative RT-PCR. After normalizing expression to *UBC* (section 2.2.4.3), the fold change in expression levels was calculated relative to the two normal EBV-transformed cell lines (shared controls). Figure shows data corresponding to two biological replicates. The mean of three technical replicates (for each biological replicate) is shown.

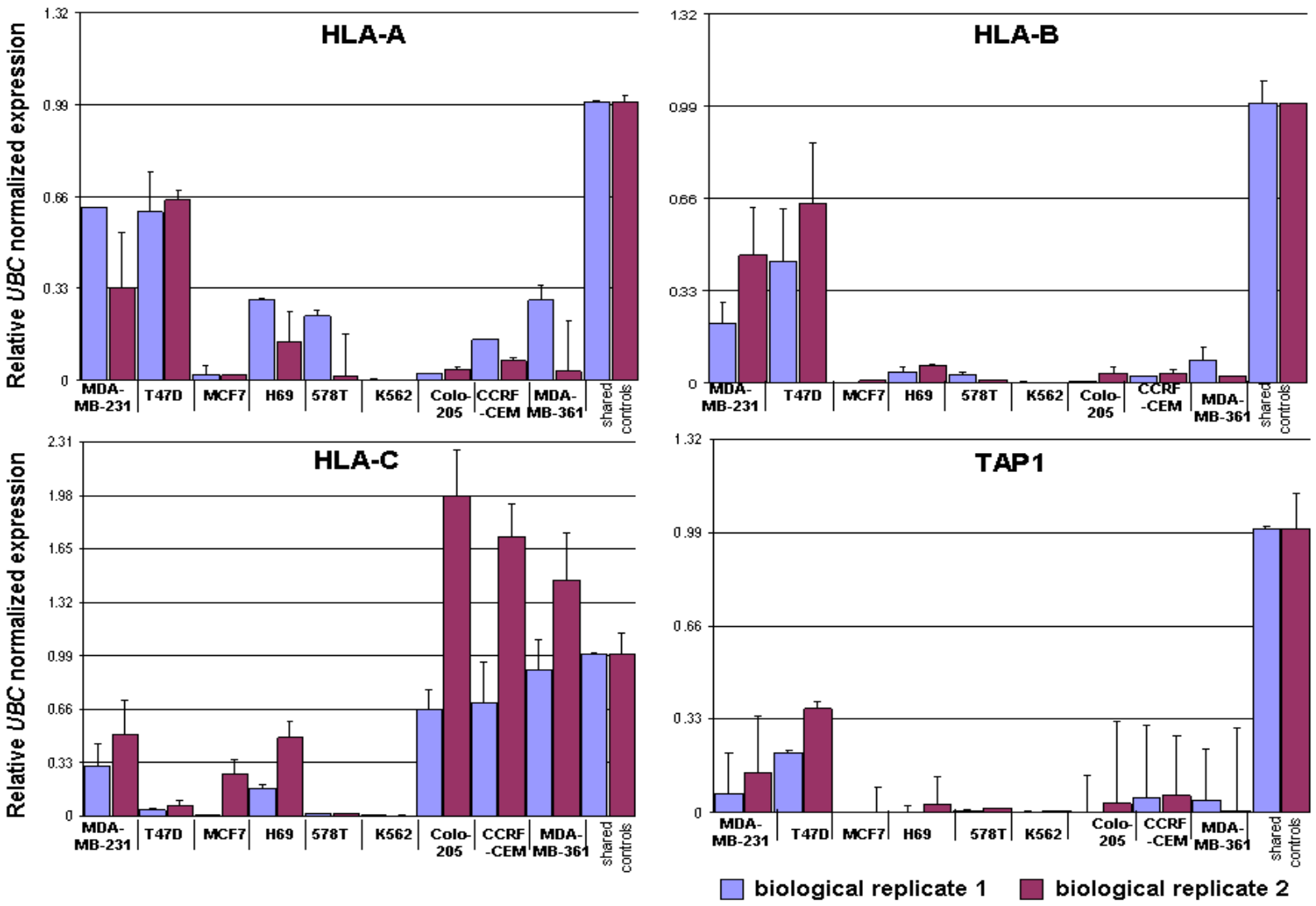
HLA-A
HLA-B
HLA-C
TAP1
Relative UBC normalized expression
1.32
0.99
0.66
0.33
0
2.31
1.98
1.65
MDA-MB-231
T47D
MCF7
H69
578T
K562
Colo-205
CCRF-CEM
MDA-MB-361
shared controls
biological replicate 1
biological replicate 2

Figure 5.3

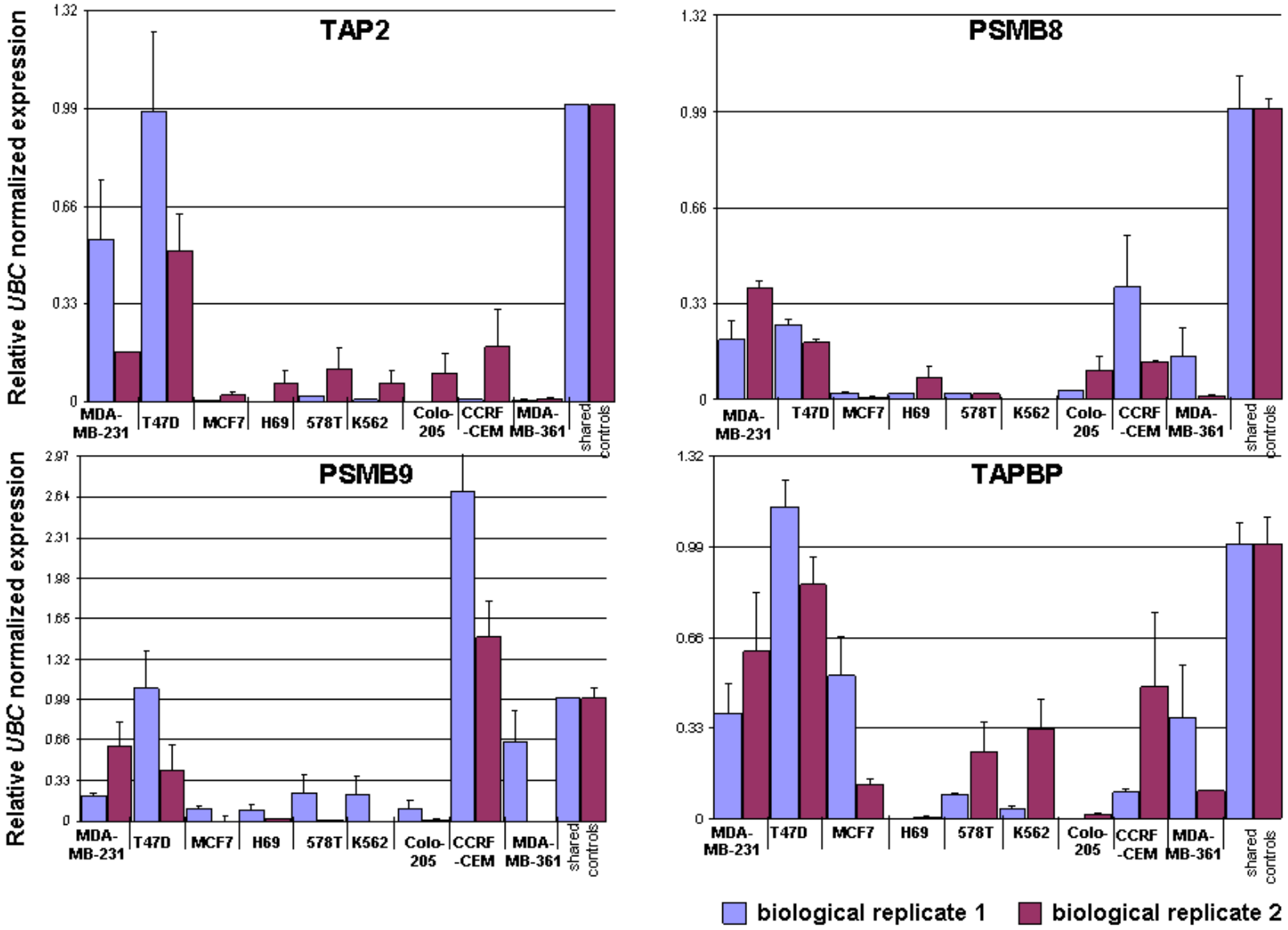
TAP2
PSMB8
PSMB9
TAPBP
Relative UBC normalized expression
1.32
0.99
0.66
0.33
0
2.97
2.64
2.31
1.98
1.65
MDA-MB-231
T47D
MCF7
H69
578T
K562
Colo-205
CCRF-CEM
MDA-MB-361
shared controls
biological replicate 1
biological replicate 2

Figure 5.3

5.4 Effect of DNA methylation inhibition on MHC encoded class I pathway genes

DNA methylation has been shown to co-ordinately change the expression of multiple genes in a chromosomal segment as well as genes involved in the same pathway (Frigola et al., 2006; Hinshelwood et al., 2007). The eight genes under investigation in this chapter are encoded within the same chromosomal band and are all involved in the same pathway, the MHC class I presentation pathway. According to my expression analysis (section 5.3) MHC-encoded MHC class I pathway genes are co-ordinately expressed.

I sought to investigate the effect of DNA methylation on the down-regulation of MHC class I pathway genes. For this I selected two cancer cell lines with the lowest expression of MHC class I pathway genes (MCF7 and 578T; figure 5.3) and treated them with increasing doses of the DNA methyltranferase inhibitor 5-aza-2'-deoxycytidine (5-aza-CdR) (sections 1.3.5 and 2.2.1.3.5). Real time qPCR analysis of MHC class I pathway gene expression revealed that 5-aza-CdR treatment can induce a marked increase of expression for most of the genes, in a dose-dependent manner (figure 5.4). In MCF7 cells, expression of all eight genes was up-regulated compared to untreated cells. The most dramatic effect was on the expression of the *PSMB8* gene (~6000-fold increase) whereas *TAPBP* and *TAP2* gene expression levels were the least affected (4-fold and 9-fold increase respectively) (figure5.4a). In 578T cells, methylation inhibition affected mainly the expression of *HLA-A* and *HLA-C* (20-fold increase) followed by *PSMB8* (6-fold increase). Treatment had no effect on the expression of the *TAP2* and *TAPBP* genes whereas for the rest of the genes, a small increase (4-fold) in mRNA levels was observed (figure 5.4b).

Although the degree of expression restoration of the MHC class I pathway genes varies between MCF7 and 578T cells, it is clear that 5-aza-CdR can co-ordinately up-regulate

genes involved in the pathway. Hence, this implicates a role for DNA methylation in the regulation of MHC class I genes.

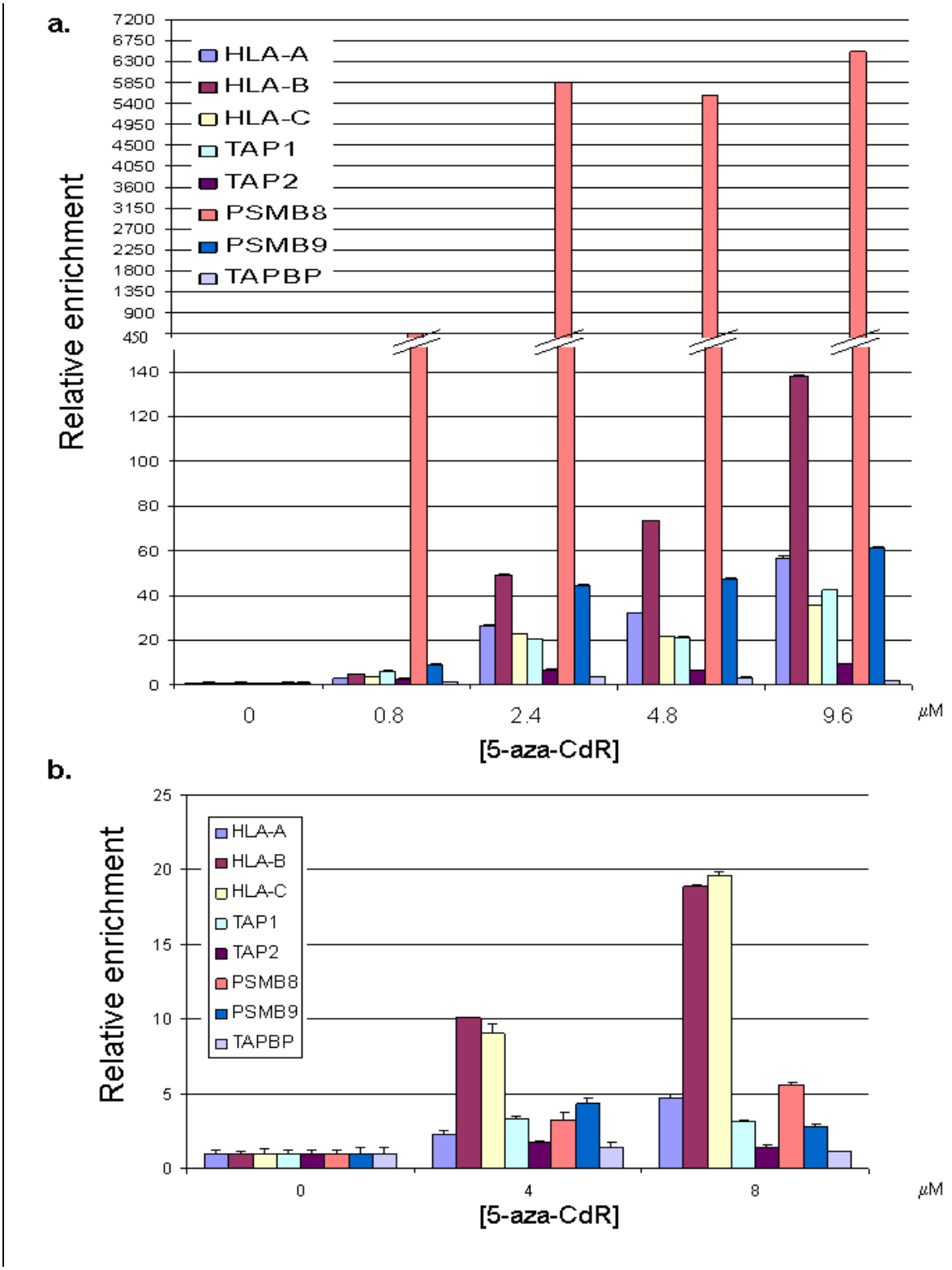

Figure 5.4. **Restoration of gene expression after DNA methylation inhibition in two cancer cell lines.** mRNA levels of *HLA-A, HLA-B, HLA-C, TAP1, TAP2, PSMB8, PSMB9* and *TAPBP* were determined in MCF7 (a) and 578T (b) in both untreated and 5-aza-CdR treated cells. After normalizing expression to *UBC*, the fold change in expression was calculated relative to untreated cells (0 μM 5-aza-CdR). The mean of six measurements (two biological replicates and three technical replicates for each) is shown in part a. The mean of three technical replicates is shown in part b.

At this point it is worth mentioning that, although the DNA demethylation capabilities of 5-aza-CdR are well characterised, this drug may exhibit alternative mechanisms of transcription reactivation. It has been shown for instance that 5-aza-CdR can induce expression of genes lacking CpG methylation (Scott et al., 2006; Soengas et al., 2001) and that it can be associated with H3K9 demethylation (Coombes et al., 2003; Fahrner et al., 2002; Nguyen et al., 2002). Also, as 5-aza-CdR is a global demethylation agent, 5-aza-CdR assays are not efficient to confirm that methylation of a specific genomic region affects expression of a gene. Therefore, the possibility that other epigenetic marks, as well as DNA methylation, and that demethylation of potential control regions outside the MHC may have an effect on MHC class I regulation should not be ignored.

Following verification of the role of DNA methylation in the concordant down-regulation of MHC class I pathway genes I sought to investigate further if aberrant DNA methylation patterns within the MHC region are involved in this phenotype.

5.5 MHC DMRs associated with the MHC class I$^-$ phenotype

The main aim of this chapter was to investigate if there are DMRs within the MHC that could be associated with the expression of MHC class I genes encoded within the MHC. To this end:

5.5.1 I generated methylation profiles for all eleven cell lines used here (table 5.1) by using the MeDIP-MHC tiling array approach (chapter 3).

5.5.2 I identified methylation differences between each of the nine cancer cell lines (table 5.1) and the shared controls. Methylation differences between two samples are referred to as DMRs.

5.5.3 I correlated the DMRs (5.5.2) with expression data based on the analysis described in section 5.3. DMRs that showed an association with expression data were called phenotype specific DMRs (pDMRs).

5.5.1 Generation of DNA methylation profiles within the MHC region

I generated methylation profiles for all eleven cell lines tested using the MHC tiling array in combination with MeDIP as described in chapter 3. In accordance to what was observed in chapter 4, the overall methylation profiles along the MHC are very similar in all 11 cell lines tested (figure 5.5). The profiles show similar patterns with the C+G content across the MHC region. Although MeDIP enrichment profiles do not allow absolute DNA methylation values to be called, they enable detection of relative methylation differences between samples (Keshet et al., 2006; Weber et al., 2005). These profiles were used to identify methylation differences between the shared control cell lines and the cancer cell lines tested here and subsequently for pDMR identification as described below.

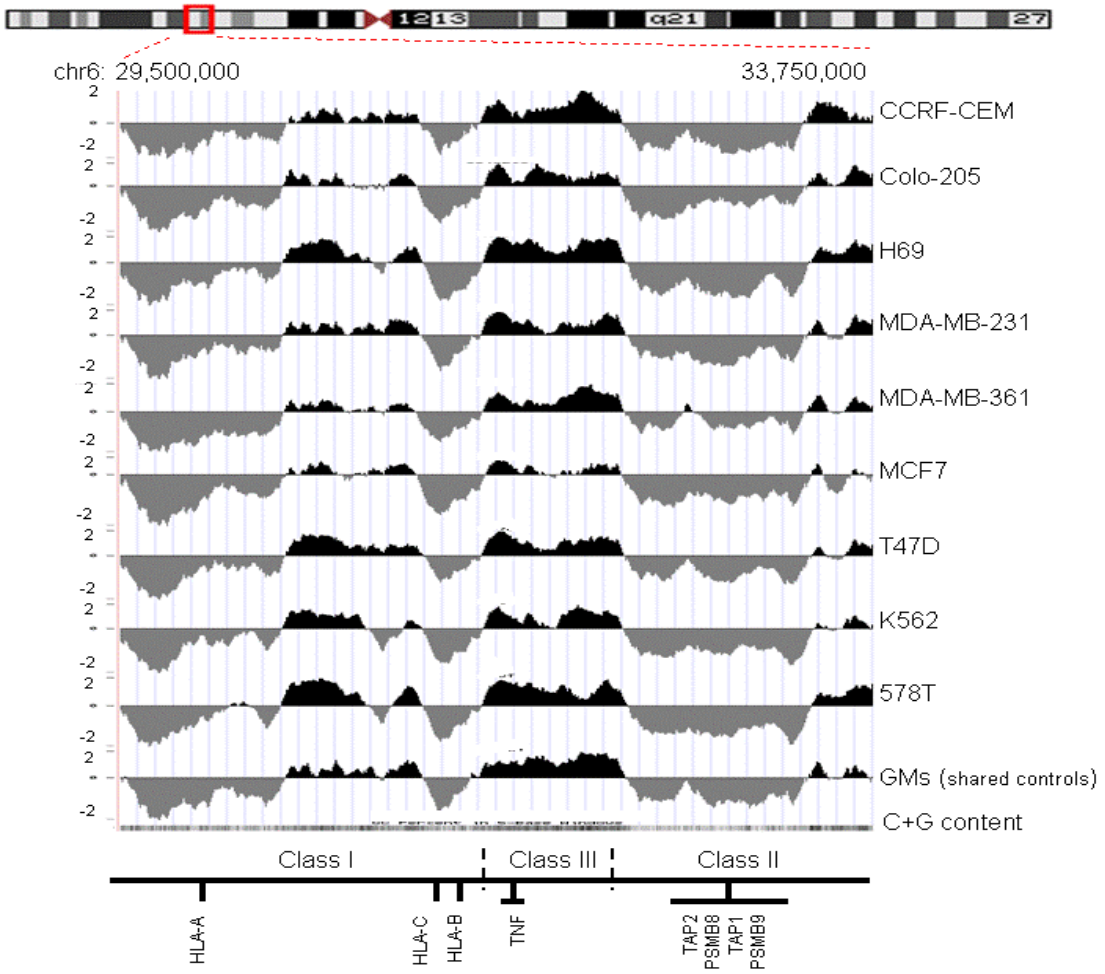

Figure 5.5 **DNA methylation profiles of the MHC.** For each of the 11 cell lines tested the $\log_2$ signal ratios (MeDIP/input) were uploaded as individual tracks to the UCSC genome browser using the 'smooth' function. For each sample the mean of two biological replicates was calculated. GMs-normal refers to the average values corresponding to the two control cell lines GM10851 and GM15510. Regions enriched or depleted in DNA methylation are shaded in black and grey, respectively. Also shown is a track of the C+G content (the darker the shading, the higher the C+G content). The approximate positions of the MHC class I, II and II subregions and some landmark genes are indicated.

5.5.2. DMRs between the cancer cell lines and shared controls

I identified differentially methylated regions (DMRs) between each of the nine cancer and the two normal EBV-transformed cell lines. The analysis for DMR identification was done in a similar manner as for tDMR identification (chapter 4).

At this point it is worth mentioning that the control cell lines originate from peripheral blood whereas the cancer cell lines originate from a variety of tissues (table 5.1). Hence, it should be expected that some of the DMRs are due to tissue-specific differences between the control and the corresponding cancer cell line. I reasoned that I should remove from the DMR list those DMRs that have already been characterised as tDMRs by the tDMR screen (chapter 4). Of the DMRs identified, 18 overlapped with tDMRs. These 18 DMRs were removed from further analysis. Their coordinates are given in appendix table 5.1. The genomic location of the remaining DMRs is shown in figure 5.6 (coordinates are in appendix table 5.2). A total of 552 putative DMRs were identified, of which 139 were present in multiple comparisons whereas 144 were identified only once. Hence 283 loci (average size 2kb) within the MHC region show methylation differences between the cancer cell lines and the two shared controls. I defined these 283 loci as non-redundant putative DMRs (to reflect the non-redundancy at the sequence level) in a similar manner as described in section 4.7 (non-redundant tDMRs). Figure 5.6 shows the location of the non-redundant DMRs as well as the 55 non-redundant tDMRs identified in chapter 4. The five fold difference between the number of the non-redundant DMRs and non-redundant tDMRs may be the outcome of more pair-wise comparisons (nine in the pDMR sceen and six in the tDMR) in the pDMR sceen.

The cell line with the most DMRs is MCF7 (141 DMRs) and the cell line with the least DMRs (43 DMRs) is T47D. The CCRF-CEM cell line, which is the only cancer cell line of the same tissue source as the shared controls (table 5.1), has 46 DMRs.

All of the 283 non-redundant DMRs are candidate methylation regulatory elements for the MHC class I^- phenotype. In the following section I filtered this list of DMRs further to identify those that show the highest co-occurrence with the phenotype under investigation.

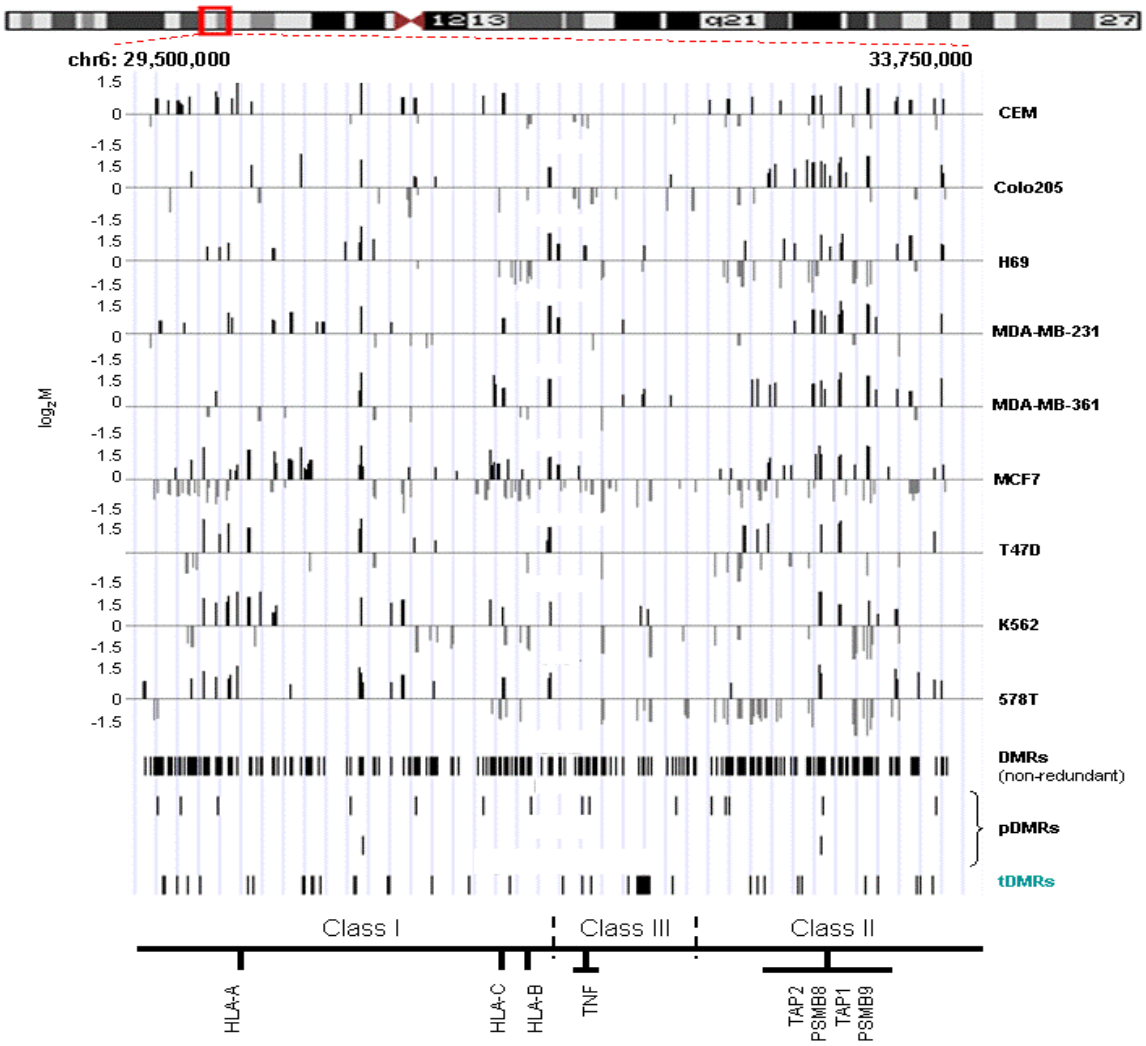

Figure 5.6 **DMRs identified between the cancer cell lines and shared controls**. Pair-wise comparisons (nine in total) of the MHC array-derived DNA methylation profiles were performed using t-statistics. A significance threshold p-value < 0.001 was used. Genomic location of DMRs identified for each pair-wise comparison (cancer cell line versus the mean values of the shared controls GM15510 and GM10851) are shown. The vertical axis shows the $\log_2$ ratio of the two corresponding methylation profiles (ie cancer cell line versus shared controls). Each line represents a DMR (average size 2kb). Black lines represent DMRs that are more methylated in cancer compared to the control cell lines (the identities of cancer cell lines within each comparison are given on the right) whereas grey lines represent less methylated DMRs. A track of the C+G content and the approximate positions of the MHC class I, II and II sub-regions and some landmark genes are also indicated. The genomic location of the 283 non-redundant DMRs (section 5.5.2) and 55 non-redundant tDMRs (section 4.7) are shown as black lines (average size 2kb). The pDMRs identified in section 5.5.3 are also shown. Upper track shows the 14 pDMRs

associated with *PSMB9* up-regulation whereas the lower track shows the two pDMRs associated with *HLA-A, HLA-B, PSMB8* and *TAP1* down-regulation.

5.5.3 pDMR identification

In the previous section I presented the DMRs between each of the nine cancer cell lines that display the MHC class I^- phenotype and shared normal control cell lines. In this section I analysed which of these DMRs can be highly linked with the MHC class I^- phenotype and hence called pDMRs.

This analysis was based on the expression analysis described in section 5.3. I aimed to identify the DMRs that could be connected with the expression of each of the eight genes involved in the MHC class I pathway and encoded within the MHC region (table 5.2; figure 5.3). The genes *HLA-A, HLA-B, TAP1* and *PSMB8* are down-regulated in all nine cancer cell lines compared to shared controls (>1.5 fold difference). I reasoned that MHC-DMRs linked with their expression should be present in all nine comparisons (figure 5.6). In a similar manner, *HLA-C* associated DMRs should be present in all but the CCRF-CEM, Colo-205 and MDA-MB-361 cell lines. *TAP2* and *TAPBP* DMRs should be present in all cell lines apart from T47D. Finally DMRs associated with *PSMB9* down-regulation should be present in MDA-MB-231, MDA-MB-361, Colo-205, H69, 578T, MCF7 and K562 comparisons. DMRs present only in CCRF-CEM comparison could be associated with *PSMB9* up-regulation (table 5.2).

5.5.3.1 pDMRs associated with *HLA-A, HLA-B, TAP1* and *PSMB8* expression

There are two putative pDMRs that can be associated with the expression of *HLA-A*, *HLA-B, TAP1* and *PSMB8* genes (figure 5.6). Interestingly, one of these DMRs overlaps with the bidirectional promoter of *TAP1* and *PSMB9* (figure 5.7a) (Wright et al., 1995). This pDMR is hypermethylated in all cancer cell lines compared to the shared controls and, because of its genomic location it can have a putative role in the regulation of

TAP1. Although this pDMR could also regulate the *PSMB9* gene, it seems that this is not the case as this pDMR is present in all cell lines, including those expressing *PSMB9* at normal levels (figure 5.3). The methylation status of the genomic region corresponding to this pDMR was verified further in all cell lines tested by bisulphite sequencing (figure 5.7b). These data show that about 10 CpG sites within this DMR are hypermethylated in all cancer cell lines. In the K562 cell line there are 20 additional CpG sites that are hypermethylated but these cannot be correlated with expression patterns (*TAP1* is downregulated in all cell lines, not only in K562 cells). About 36% of CpG sites within this pDMR were refractory for bisulphite sequencing analysis. Bisulphite sequencing analysis normally fails due to: i. failure in designing primers specific for bisulphite converted DNA (see section 1.3.6.1) and ii: poor quality of sequence traces corresponding to bisulphite converted DNA.

I attempted to verify further the role of the hypermethylation of the 10 CpG sites reported above. To this end I performed bisulphite sequencing analysis of DNA extracted from 5-aza-CdR treated MCF7 cells (section 5.4). This analysis revealed a 5-fold reduction in methylation levels of the 10 CpG sites after inhibition of DNA methylation (figure 5.7c). This result in combination with the expression analysis in 5-aza-CdR treated cells (section 5.4), where the expression of the *HLA-A, HLA-B, TAP1* and *PSMB8* genes is up-regulated (figure 5.4), indicates that methylation levels of these 10 CpG sites may be involved in the regulation of *HLA-A, HLA-B, PSMB8* and *TAP1* genes.

Interestingly, although these 10 CpG sites are located within the *TAP1*/*PSMB9* promoter, they do not overlap with any known control element within this region (figure 5.8). Instead they are located downstream of the start codon and TSS of *PSMB9*. It is possible that I have identified a new regulatory element within the promoter region. Functional analysis experiments, which are discussed in section 7.4.1, are required to verify further this possibility.

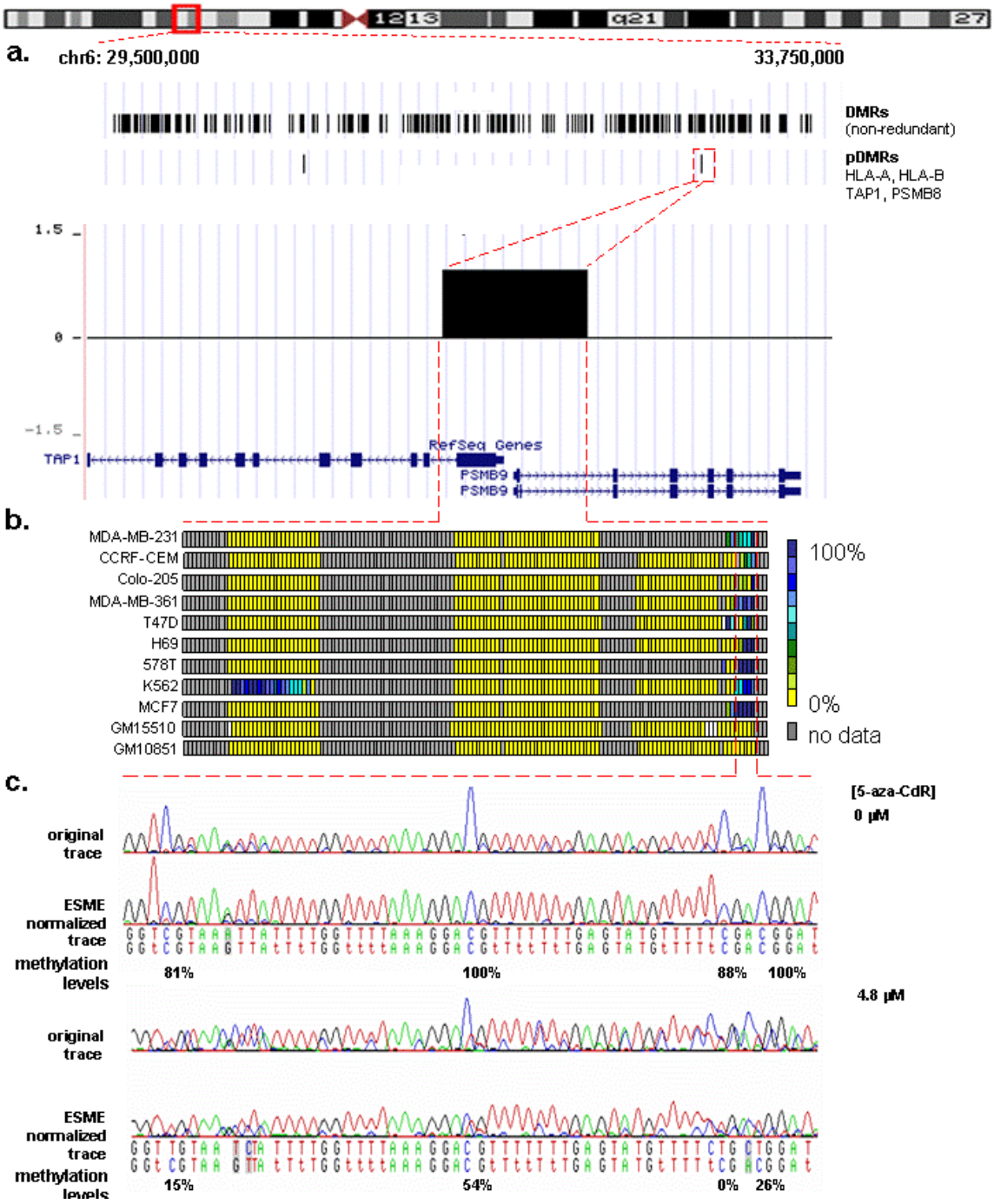

Figure 5.7. **A pDMR within the *TAP1*/*PSMB9* bidirectional promoter.** A pDMR within the *TAP1*/*PSMB9* promoter was associated with the down-regulation of *HLA-A, HLA-B, TAP1* and *PSMB8* but not *PSMB9*. According to MeDIP-MHC array data this pDMRs is hypermethylated in all cancer cell lines compared to the shared controls. a. Genomic location of the pDMR. b. The methylation patterns within this pDMR were verified by bisulphite sequencing. Each square represents a CpG site. The colour code indicates methylation percentage as calculated by ESME analysis (section 2.2.3). c. Effect of 5-aza-CdR. DNA extracted from MCF7 5-aza-CdR treated

and untreated cells, was subjected to bisulphite sequencing analysis. Methylation of the 10 CpG sites that were found to be hypermethylated (b) in cancer cell lines was tested. Traces corresponding to 5-aza-CdR treated and untreated samples as well as traces before and after ESME normalization (section 2.2.3.5) are shown. Concentration of 5-aza-CdR used is indicated. Upon 5-aza-treatment methylation values drop around 5-fold (on average). A representative part of the sequencing data generated is shown.

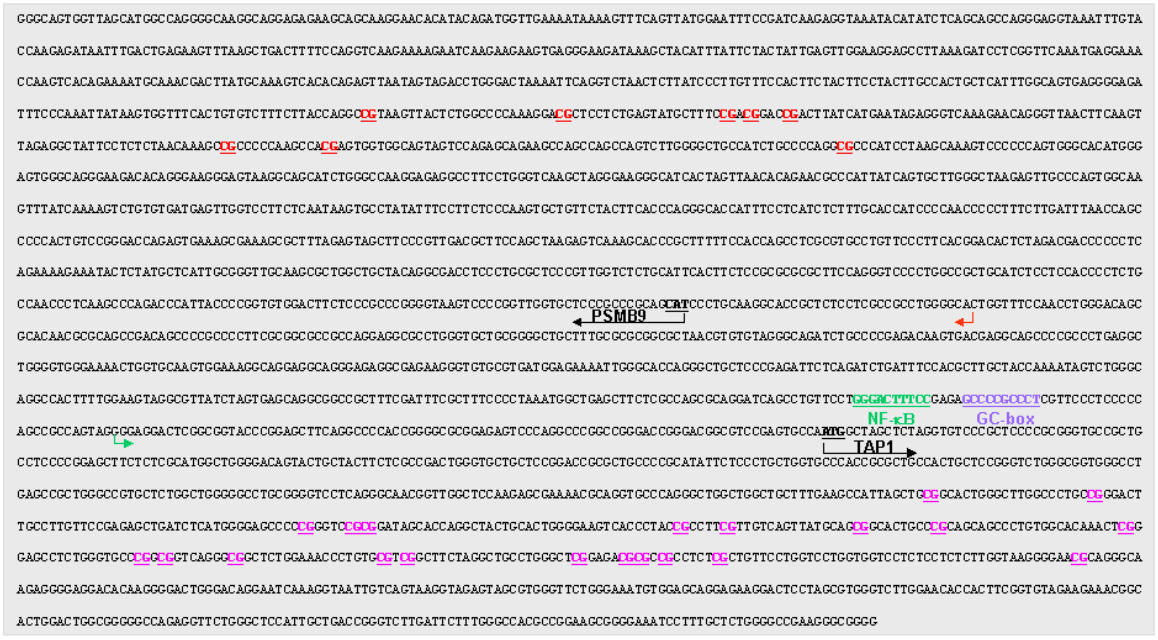

Figure 5.8. **Sequence of the DMR overlapping with the *TAP1*/*PSMB9* bidirectional promoter** The ATG translation start codons are indicated by a black arrow and the gene name. The most prominent transcription start sites are depicted by red and green arrows for PSMB9 and TAP1 respectively. Two functional transcription factor binding sites defined in Wright et al (Wright et al., 1995) are indicated. Green letters correspond to NF-kB and purple to GC-box binding sites. The ten CpG sites hypermethylated in all cancer cell lines are coloured red. The CpG sites hypermethylated only in the K562 cell line are coloured pink.

The second pDMR overlaps with the coding and 3'UTR regions of Nurim (NRM) which is an inner nuclear membrane (INM) protein (figure 5.9) (Holmer and Worman, 2001; Rolls et al., 1999). To date, there are no data implicating NRM in the MHC class I pathway indicating that I might have identified a novel control element for the expression of

members of the MHC class I pathway. However, further experiments are required to prove or refute this possibility.

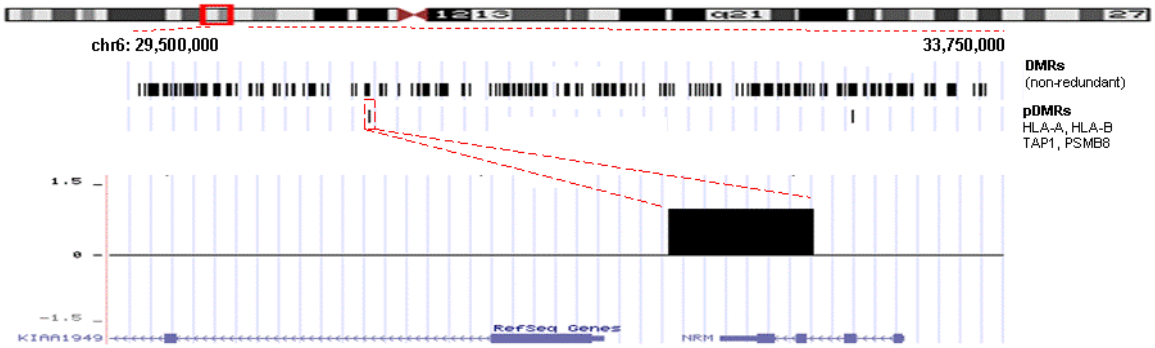

Figure 5.9. **pDMR associated with *HLA-A, HLA-B, TAP1* and *PSMB8* expression**. Pair-wise comparisons between each of the cancer cell lines and the two shared controls identified a putative pDMRs at the NMR loci in the MHC class I region. This pDMR is hypermethylated in all cancer cell lines compared to the shared controls. Tracks, also show the 283 non-redundant DMRs and the 2 pDMRs associated with HLA-A, HLA-B, PSMB8 and TAP1, were uploaded to the UCSC browser.

It is worth mentioning that none of these two pDMRs can be connected directly with the expression of *HLA-A, HLA-B* and *PSMB8* genes but this does not exclude the possibility that they regulate the expression of these genes. It is known that transcriptional control elements can regulate expression of genes within a distance on the same chromosomal band or even on different chromosomes (figure 4.1) (Maston et al., 2006).

5.5.3.2. pDMRs associated with *TAP2, TAPBP, HLA-C* and *PSMB9* expression

According to this analysis there are no pDMRs within the MHC associated with the expression of *TAP2*, *TAPBP* and *HLA-C*. Based on data of 5-aza-CdR treatment (section 5.4), expression levels of *TAP2* and *TAPBP* were the least affected. Hence it is possible that DNA methylation is not involved in the regulation of these two genes and that is why no pDMR was found to be associated with them. On the other hand *HLA-C* has been shown to be up-regulated after 5-aza-CdR treatment. It is possible that due to array

resolution I was not able to detect all pDMRs, particularly small pDMRs. Alternatively, the expression of *HLA-C* may be controlled by other epigenetic marks, for example histone marks. 5-aza-CdR has already been associated with H3 lysine 9 methylation (Coombes et al., 2003; Fahrner et al., 2002; Nguyen et al., 2002). Mapping histone marks across the MHC by performing Chromatin Immunoprecipitation (ChIP) in combination with the MHC tiling array will be informative on this matter.

Finally, 14 DMRs were present only in the CCRF-CEM cell line (figure 5.10). CCRF-CEM is the only cancer cell line in this study that shows increased levels of *PSMB9* mRNA (figure 5.3). Hence these 14 DMRs can be connected with *PSMB9* up-regulation and hence can be characterised as pDMRs. Although all of the 14 pDMRs can be equally important, the one overlapping with the coding and 3'UTR region of the *PSMB9* gene is the most likely to control *PSMB9* up-regulation. A previous study has shown that 3'UTR hypomethylation is associated with gene silencing (Shann et al., 2008). However, it is possible that the combination of 3'UTR hypomethylation together with 5'UTR hypermethylation, which is the case for the *PSMB9* gene in the CCRF-CEM cell line, leads to transcription activation (figures 5.10).

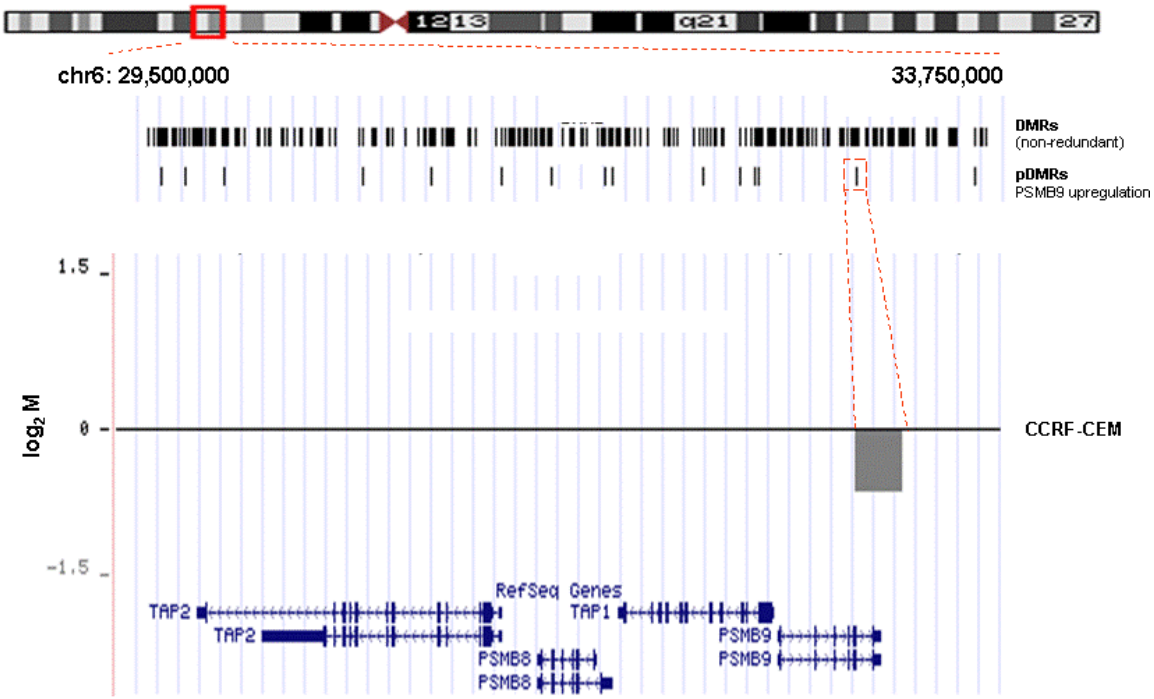

Figure 5.10. **pDMRs associated with *PSMB9* up-regulation.** 14 DMRs present only in the CCRF-CEM comparison were associated with *PSMB9* up-regulation. Of these the one overlapping the 3'UTR of *PSMB9* is the most likely to control expression of the gene. It seems to be less methylated in CCRF-CEM as shown in the lower part of the figure. RefSeq genes are also shown. Tracks showing the genomic location of DMRs were uploaded to the UCSC genome browser.

In summary, in this section I describe the identification of two pDMRs associated with the down-regulation of *HLA-A, HLA-B, TAP1* and *PSMB8* as well as of 14 pDMRs associated with the up-regulation of *PSMB9*. One interesting question arising from pDMR identification is the mechanism that drives their formation during the development of a phenotype, in this case the MHC class I^- phenotype. As it was discussed in the general introduction of this thesis (chapter 1), there is supporting evidence suggesting that DNA methylation is the result of low transcriptional activity. This implies that DNA methylation is a secondary event during the process of gene silencing (Bird, 2002; Clark and Melki, 2002; Turker, 2002).

In the following section I attempt to correlate increasing methylation levels with reduction of transcriptional activity.

5.6 DNA methylation and levels of transcriptional activity

It has been proposed that gene silencing is a gradual and evolving process that induces promoter DNA methylation at a very late stage of the transcriptional silencing process (Bird, 2002; Clark and Melki, 2002; Turker, 2002). This is consistent with recent data suggesting a “Use it or Lose it” model (Meissner et al., 2008). Based on this model, genes with low levels of transcriptional activity in a given cell type are susceptible to be locked into this state by DNA methylation. In an attempt to investigate this possibility further I identified all the DMRs (section 5.5.2) that overlap with the eight MHC encoded genes involved in the MHC class I pathway and I correlated them with the corresponding expression levels (section 5.3; figure 5.3). Although all candidate genes are down-regulated in almost all cases (figure 5.3), the level of down-regulation differs between cell lines.

Based on this analysis there is: (i). one DMR overlapping with the bidirectional promoter of *TAP1* and *PSMB9* which is present in all cell lines and hence cannot be correlated with differential expression levels. (ii). one DMR present within the *HLA-B* locus which is not correlated with expression levels and hence it is not discussed further within this thesis and (iii). a DMR overlapping with the promoter of *PSMB8* gene which is present in three cell lines: 578T, MCF7 and K562 (figure 5.11a). Interestingly these three cell lines are those with the lowest *PSMB8* expression levels compared to the rest (figures 5.3 and 5.11c). I tested this DMR further as described below.

5.6.1 DMRs overlapping with the PSMB8 promoter region

In three cancer cell lines (578T, K562, MCF7), a DMR was identified within the 5'UTR of the *PSMB8* gene (figure 5.11a). Bisulphite sequencing analysis confirmed that this region is heavily methylated in these three cell lines (figure 5.11b). As the *PSMB8* gene is down-regulated in all cancer cell lines (figures 5.3 and 5.11c) this DMR does not qualify as pDMR based on the analysis described in section 5.5.3. However it is important to note that 578T, K562 and MCF7 show the lowest expression levels of *PSMB8* compared to the rest of cancer cell lines (figure 5.11c) and this may be linked to the proposal that gene silencing is a gradual and evolving process that induces promoter DNA methylation (Clark and Melki, 2002; Meissner et al., 2008; Turker, 2002). Hence, I speculate that this can be the case for the silencing event of the *PSMB8* gene. Initially *PSMB8* silencing is the result of factor(s) other than DNA methylation. When transcription levels fall below a threshold (relative expression < 0.05) (figure 5.11c) induces an increase in local methylation and this could explain the high levels of methylation in the 578T, K562 and MCF7 cell lines. Treatment of MCF7 cells with 5-aza-CdR reduces methylation of the *PSMB8* DMR 5-fold (on average) (figure 5.11d) and this is in inverse correlation with the expression levels of the *PSMB8* gene in 5-aza-CdR treated cells (figure 5.4). Treatment of additional cell lines, other than MCF7, 578T and K562, with methylation inhibitors would be informative. The expectation is that treatment with 5-aza-CdR should not have an effect on the expression levels of the PSMB8 gene in cell lines lacking the DMR within the *PSMB8* 5'UTR.

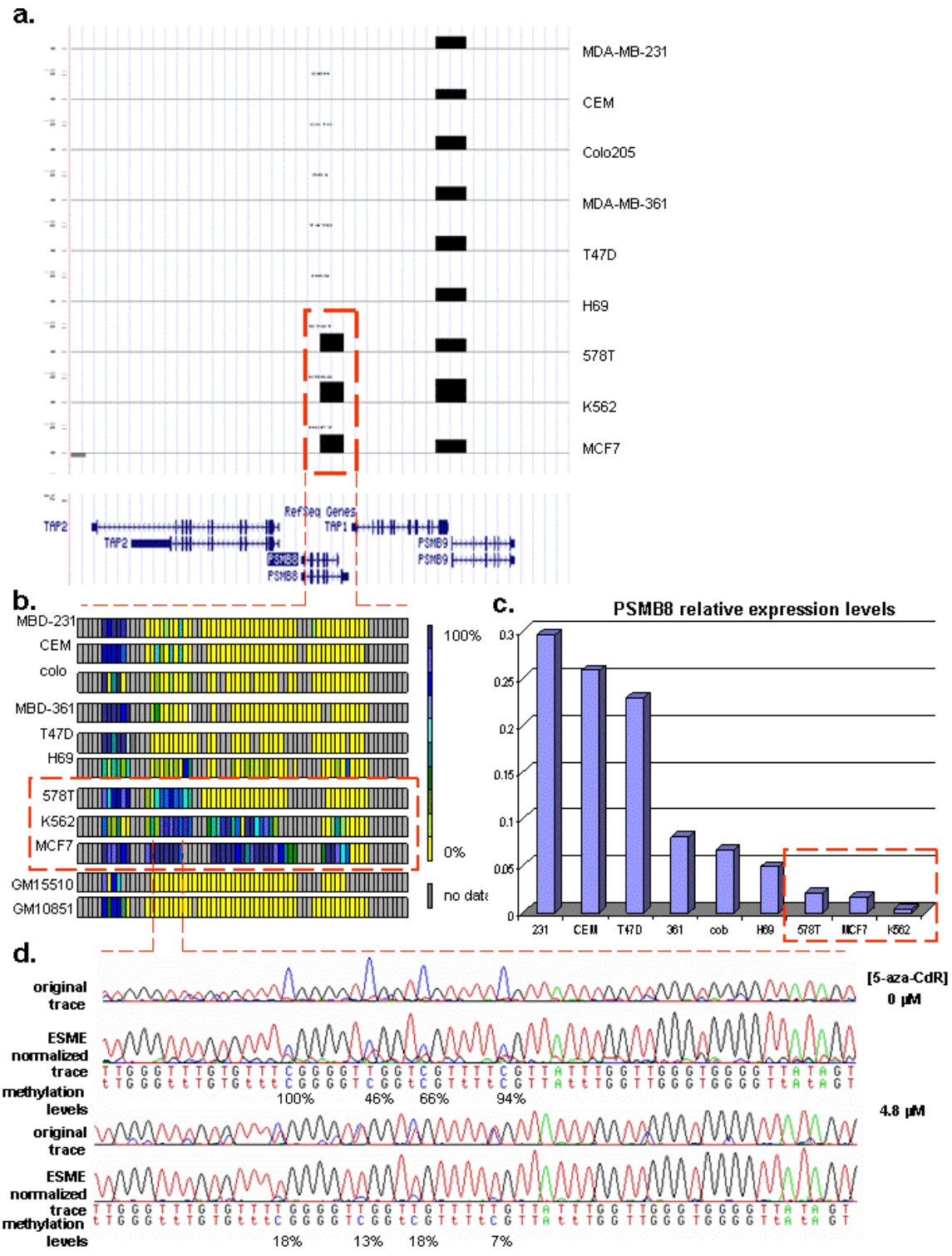

Figure 5.11 **DMRs within the PSMB8 promoter.** Pair-wise comparisons between the cancer and the shared control cells revealed a DMR within the *PSMB8* promoter in the cell lines 578T, K562 and MCF7.(a). DMRs identified by the MeDIP-MHC tiling array approach. The red box indicates the DMR within the *PSMB8* promoter. The pDMR within the *TAP1/PSMB9* promoter (discussed in

section 5.5.3.1; figure 5.7) is also shown. Tracks were uploaded to the UCSC browser. Vertical axis shows the $\log_2$ ratio of the cancer versus normal cell line profile. Identity of the corresponding cancer cell line is shown on the right. (b). DMR overlapping with *PSMB8* was validated further by bisulphite sequencing analysis. Each square represents a CpG site. The colour code indicates methylation percentage as calculated by ESME analysis. (c). Relative expression levels for *PSMB8* in the nine cancer cell lines (taken from figure 5.3). Red box indicates the three cell lines (K562, 578T, MCF7) with the lowest *PSMB8* levels (fold enrichment <0.05). Graph was reproduced using data shown in figure 5.3. (d).DNA extracted from MCF7 5-aza-CdR treated and untreated cells, was subjected to bisulphite sequencing analysis. Traces corresponding to 5-aza-CdR treated and untreated samples as well as traces before and after ESME normalization are shown. Concentration of 5-aza-CdR used is indicated. Upon 5-aza-CdR treatment methylation values dropped 5-fold. A representative part of the sequencing data generated is shown.

5.7 Prominent DMRs within the MHC region

One of the great advantages of the MeDIP-MHC tiling array approach compared to the bisulphite sequencing strategy, which followed by the HEP pilot study for the MHC region and covered only 2% of the MHC, is that it allows the unbiased methylation analysis of the complete MHC region, albeit at 2kb resolution. I wished to look for prominent DMRs within the MHC which, although they were not found to be associated with the MHC class I^- phenotype, may be important in the regulation of MHC genes in general.

To this end I used the list of DMRs that resulted by the analysis described in section 5.5.2 (pair-wise comparisons of DNA methylation profiles of cancer cell lines versus the shared controls) (appendix table 5.2) and I looked for DMRs present in the majority of comparisons. The most prominent DMRs were present in the tumour necrosis factor (TNF) cluster in the MHC class III region. These DMRs are discussed in the following sections.

5.7.1 The tumour necrosis factor cluster

The tumour necrosis factor (TNF) cluster contains genes for three cytokines (LTA, TNF-α and LTB) and is located in the MHC class III region (figure 1.1). These cytokines play multiple roles in the development and function of the immune system (Aggarwal et al., 2002; McDevitt et al., 2002). The regulation of expression of these genes is complex: transcription is controlled in a tissue- and stimulus-specific manner (Falvo et al., 2000; Falvo et al., 2000; Tsai et al., 1996). Recent evidence has shown that the *TNF-α* gene is epigenetically regulated. More specifically, it has been shown that in K562 cell (which is one of the cancer cell lines tested here), the promoter region of *TNF-α* is hyper-methylated (Sullivan et al., 2007). According to the same study K562 cells do not produce TNF-α upon lipopolysaccharide (LPS) treatment. LPS is one of the major inducers of TNF production both *in vivo* and *in vitro* (Dumitru et al., 2000).

The analysis conducted under section 5.5.2 revealed three DMRs within the TNF cluster, overlapping with the *LTA, TNF-α* and *LTB* genes. In the following sections I studied these DMRs further and, more specifically, I explored how they can be associated with the expression of the corresponding genes.

5.7.2 DMRs within the TNF cluster

According to the MHC array analysis (5.5.2) the gene body of *TNF-α* is hypermethylated in all cancer cell lines except CCRF-CEM. The *TNF-α* promoter region is hypermethylated in all but two cancer cell lines (K562 and CCRF-CEM). DMRs were also identified in some cell lines within the gene bodies of *LTA* and *LTB* (figure 5.12a). The DMRs identified within the TNF cluster were validated further by bisulphite sequencing (figure 5.12b), which agrees with the array data. The two EBV-transformed cell lines GM10851 and GM15510 (shared controls) appear to be unmethylated. CCRF-

CEM cells are either unmethylated or less methylated compared to the rest of the cancer cell lines within the regions tested. In addition, according to bisulphite sequencing data and previous studies the *TNF-α* promoter in K562 cells is hypermethylated. However the MeDIP-MHC approach failed to detect the corresponding DMR in these cells. This can be due to lower methylation levels in K562 cells compared to the other cancer cell lines (except CCRF-CEM). The lower methylation levels observed for K562 cells may not be sufficient to be detected by my experimental approach (MeDIP-MHC tiling array). Methylation values for the CpG sites within these DMRs, which were found to be refractory to bisulphite sequencing analysis, may be informative in this context

It was interesting to observe that the region overlapping with the *TNF-α* promoter-DMR contains the binding site for the transcription factor termed lipopolysaccharide-induced TNF-a factor (LITAF). LITAF is a regulatory element that has been shown to mediate LPS-induced *TNF-α* gene expression in THP-1 cells (Tang et al., 2003). Hence it is possible that the DMR identified within the promoter of TNF-α plays a role in the regulation of this gene. In a similar manner the additional DMRs identified within the TNF cluster, although they do not overlap with promoters of the *LTA* and *LTB* genes, may regulate their transcription.

In the following section I attempt to investigate how the three DMRs discussed herein may affect transcriptional levels of the corresponding genes.

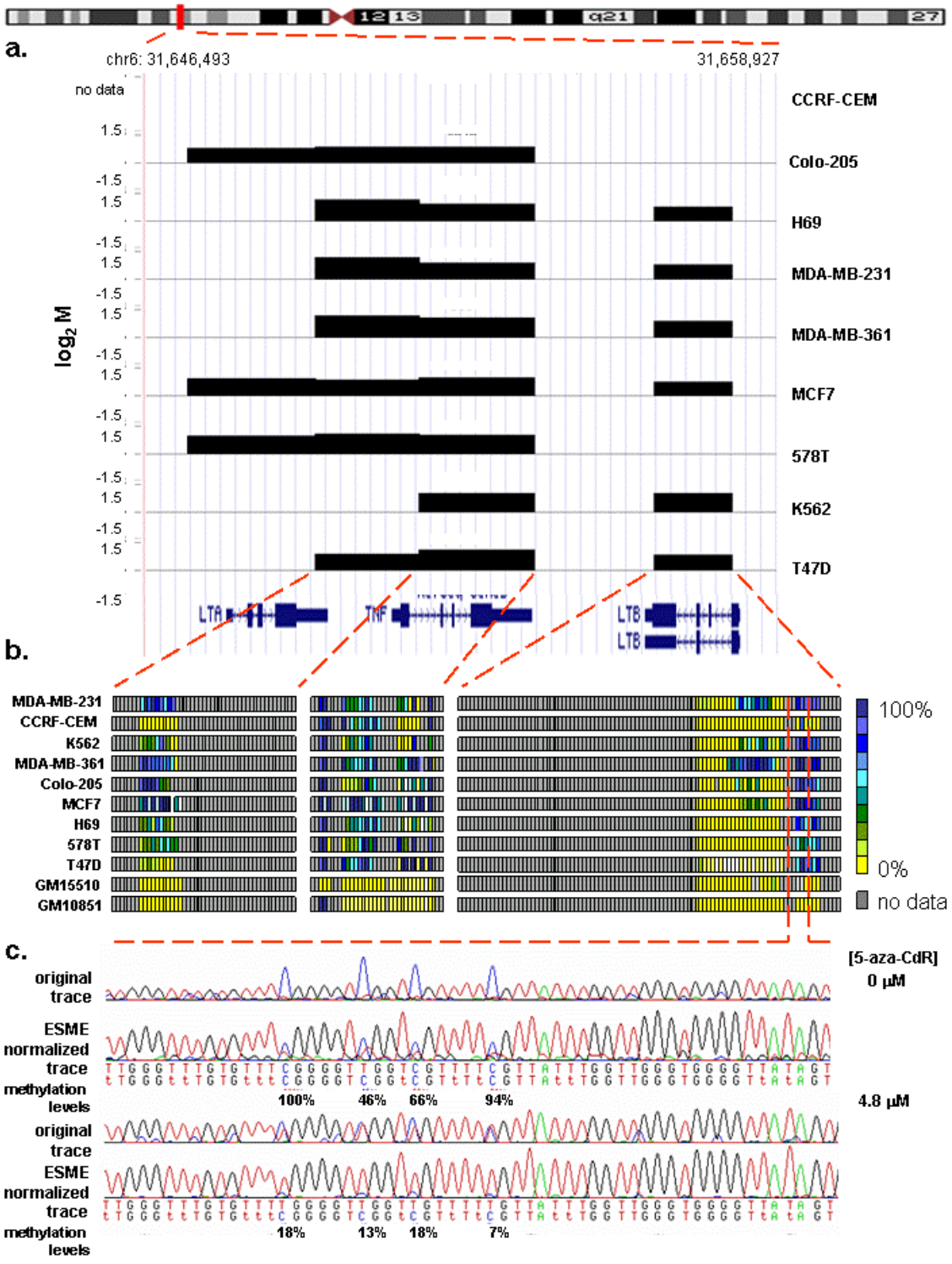

Figure 5.12. **DMRs within the TNF cluster.** Pair-wise comparisons between the cancer and control cell lines revealed four DMRs within the TNF cluster. a. Tracks were uploaded to the UCSC genome browser. Vertical axis shows the $\log_2$ ratio of the cancer versus normal cell line profile. b. DMRs overlapping with the *TNF-α* promoter, *TNF-α* gene body and *LTB* were validated

further by bisulphite sequencing analysis. Each square represents a CpG site. The colour code indicates methylation percentage as calculated by ESME analysis. c. Effect of 5-aza-CdR. DNA extracted from MCF7 5-aza-CdR treated and untreated cells, was subjected to bisulphite sequencing analysis. Traces corresponding to 5-aza-CdR treated and untreated samples as well as traces before and after ESME normalization are shown. Concentration of 5-aza-CdR used is indicated. Upon 5-aza-CdR-treatment, a 5-fold reduction (on average) in methylation levels was observed. A representative part of the sequencing data generated is shown.

5.8.3 Expression of TNF cluster genes and correlation with DMRs

To correlate the DMRs identified in the previous section with mRNA levels of the *LTA*, *LTB* and *TNF-α,* the expression of these genes in the cancer cell lines was compared to the shared controls. To this end, I performed real time qPCR using primer sets corresponding to *LTA, LTB* and *TNF-α* cDNAs respectively. Differences in expression between each of the cancer cell lines and the two shared controls were calculated as described in section 2.2.4. In all cancer cell lines all three genes were down-regulated (fold >1.5) compared to the two controls (figure 5.13)

Interestingly, the expression of *TNF-α* and *LTB* was about 4.5 fold higher in the CCRF-CEM line compared to the other cancer cell lines and this may explain the lower methylation levels in CCRF-CEM (figure 5.12a,b).

The correlation of low expression levels of genes encoded within the TNF cluster with the DMRs presented in the previous section suggests that DNA methylation may be a transcription regulator for these genes. In order to test this further, expression of *LTA*, *TNF-α* and *LTB* was determined in 5-aza-CdR treated MCF7 and 578T cells as it was done above for the MHC class I pathway genes (section 5.4). 5-aza-CdR was capable of inducing expression of *TNF-α, LTA* and *LTB* in both cell lines although the effect was

.

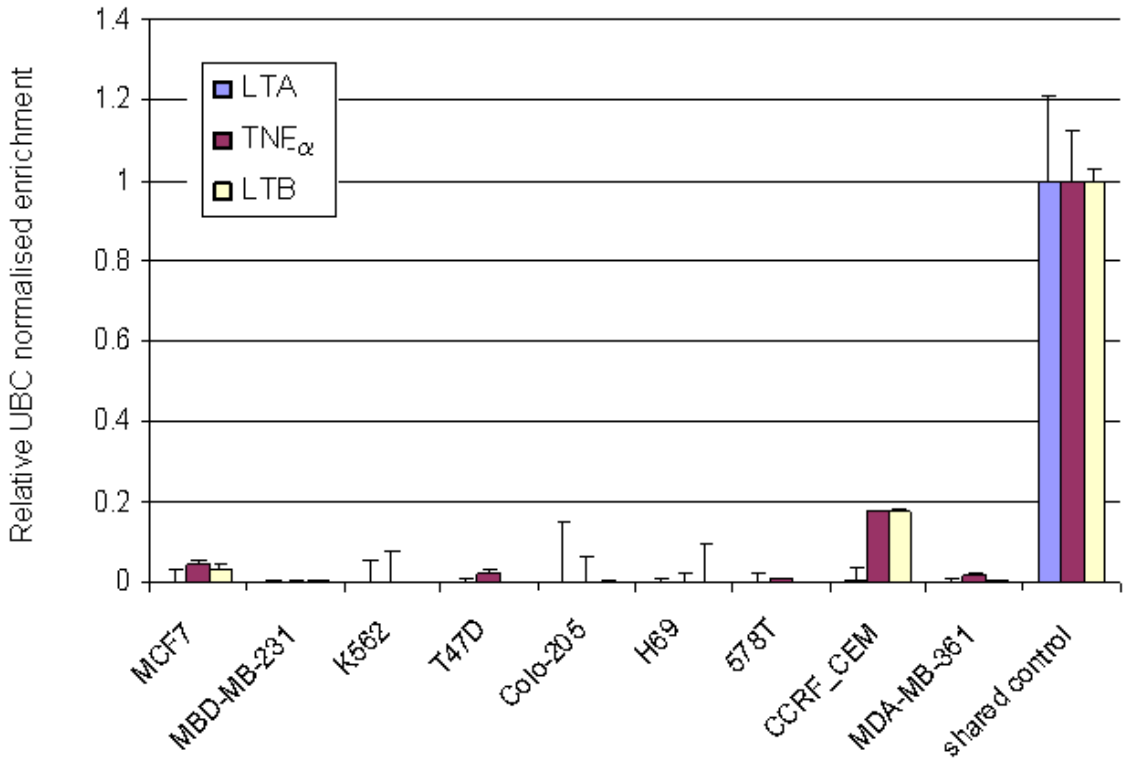

Figure 5.13. **Relative expression of TNF cluster genes.** mRNA levels of *LTA, TNF-α* and *LTB* were determined by quantitative RT-PCR. After normalizing expression to *UBC* (section 2.2.4.3) the fold change in expression levels was calculated relative to the two EBV transformed control cell lines (shared controls). Figure shows the mean of data corresponding to two biological replicates and three technical replicates of each (six measurements in total).

more significant in the MCF7 cell line (figure 5.14). In addition, bisulphite sequencing data corresponding to DNA extracted from 5-aza-CdR treated cells confirmed that methylation levels were decreased in the TNF cluster DMRs (figure 5.12c) follwing inhibition of DNA methylation.

At this point I would like to mention that TNF-α together with interferon-γ (IFN-γ) are two cytokines known to be involved in regulation of MHC class I molecules (Johnson, 2003; Johnson and Pober, 1994; Ohmori and Hamilton, 1995). In the next section, I discuss how methylation of the *TNF-α* promoter can be correlated with the expression of genes involved in the MHC class I pathway.

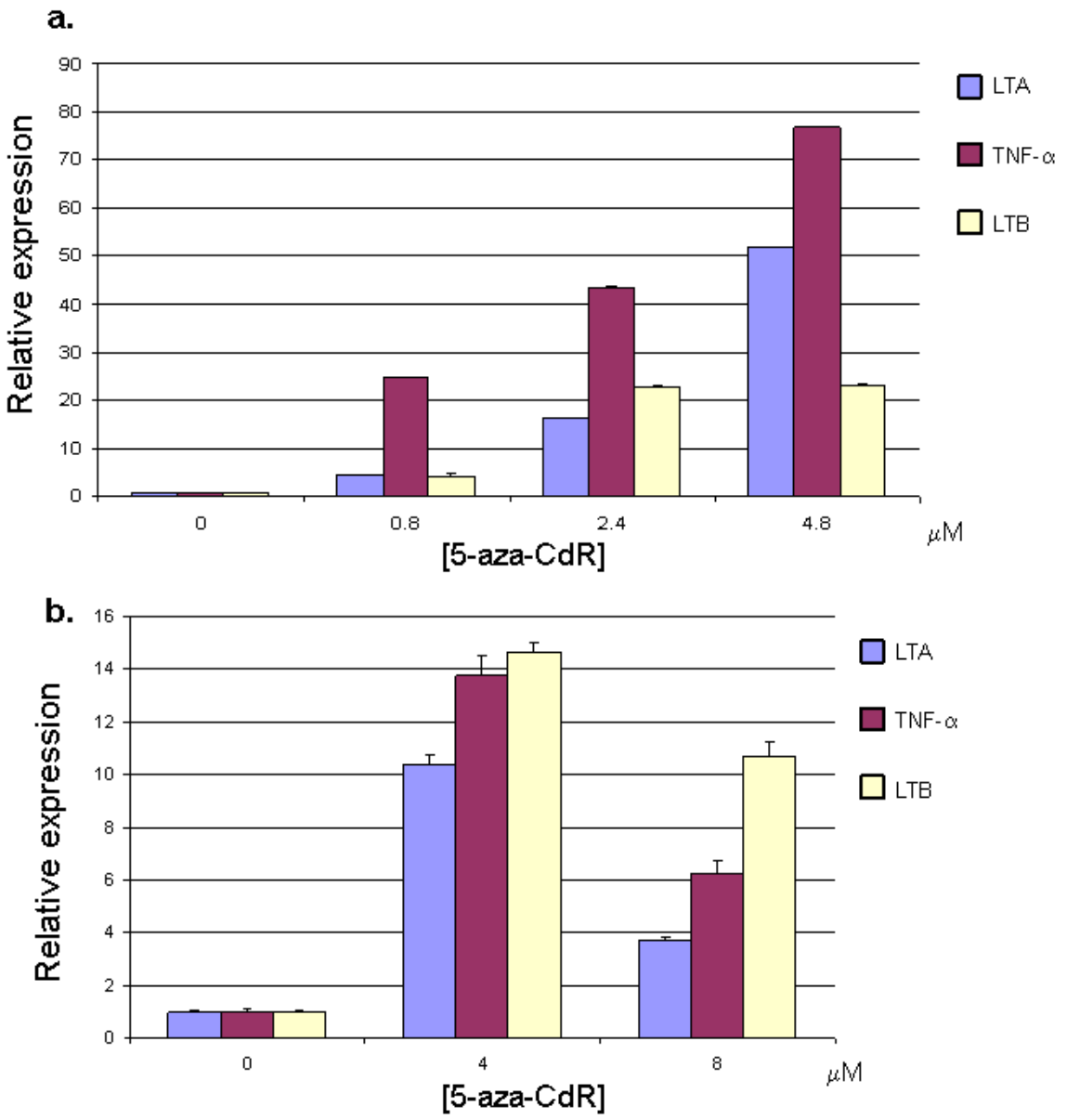

Figure 5.14 **TNF-cluster gene expression after DNA methylation inhibition in two cancer cell lines.** mRNA levels of *LTA, TNF-α* and *LTB* were determined in MCF7 (a) and 578T (b) in both untreated and 5-aza-CdR treated cells. After normalizing expression to *UBC*, the fold change in expression was calculated relative to untreated cells (0 μM 5-aza-CdR). The mean of six measurements (two biological replicates and three technical replicates for each) is shown in part a. The mean of three technical replicates is shown in part b.

5.8 Discussion

Previous studies provided supportive evidence for a role of DNA methylation in the regulation of genes involved in the MHC class I pathway (Fonsatti et al., 2007; Fonsatti et al., 2003; Nie et al., 2001). In this chapter I aimed to elucidate this further. Specifically, I attempted to test if DNA methylation within the MHC region is involved in the MHC class I^- phenotype. This is associated with many diseases. Hence elucidating the mechanism that leads to this phenotype can contribute to diagnosis and treatment.

My data confirmed the concordant down-regulation of the eight MHC class I pathway genes encoded within the MHC region in the nine cancer cell lines (displaying the MHC class I^- phenotype) used, compared to two shared controls (EBV-transformed cell lines). Treatment of two cancer cell lines (MCF7 and K562) with a DNA methyltransferase inhibitor (5-aza-CdR) revealed that changes in DNA methylation patterns in the cancer cell lines may lead to the concordant aberrant expression of the eight genes.

I used the MeDIP-MHC tiling path assay, which I developed, to investigate if there are any DNA methylation regulatory elements within the MHC region that may control the expression levels of the *HLA-A, HLA-B, HLA-C, TAP1, TAP2, TAPBP, PSMB8* and *PSMB9* genes. The advantage of this approach is that it allows for unbiased detection of DMRs within the complete MHC region. The MHC was one of the first genomic regions in which higher order chromatin architecture was shown to affect gene expression (Christova et al., 2007; Volpi et al., 2000). Consistent with this, a recent study using the tiling array I developed, has shown the existence of chromatin loops within the MHC which act as insulators for the transcription of MHC genes (Ottaviani, 2008). This indicates that the expression of MHC genes is not controlled solely by local control elements, but also by more distant regulatory elements (figure 4.1). Hence, it was important to expand this study beyond the methylation analysis of the eight loci implicated in the MHC class I pathway.

My analysis revealed 283 candidate DMRs that could be involved in the control of the expression of these genes. Correlation of these DMRs with relative expression levels of the eight candidate genes identified two DMRs that can be linked with *HLA-A, HLA-B, PSMB8* and *TAP1* down-regulation and 14 DMRs that can be linked with the up-regulation of *PSMB9*. I defined these DMRs phenotype-specific, pDMRs, to reflect their concordance with the phenotype under investigation.

Of the two pDMRs associated with *HLA-B, HLA-B, PSMB8* and *TAP1*, one is located within the *TAP1/PSMB9* bidirectional promoter and can be linked to the expression of the *TAP1* gene. Interestingly, it could not be associated with the expression of *PSMB9* because this pDMR was present in cell lines expressing this gene at normal levels (figures 5.3 and 5.7). It is possible that different elements within the promoter region control the two genes separately. More work has to be done to further elucidate this matter. Further bisulphite sequencing analysis, to reveal all the CpG sites with differential methylation patterns, and functional analysis (e.g. deleting the region with the promoter that contains the key CpG sites) will be informative. The second pDMR (NRM locus) cannot be linked directly to any of the genes of the MHC class I pathway (figure 5.9) but it may be part of a genomic loop that controls MHC genes. Chromatin loops have been associated with the MHC region (Ottaviani, 2008).

In a similar manner, although the 14 pDMRs associated with the PSMB9 up-regulation could not be linked directly with *PSMB9* expression (only one was located within its coding region) may be equally important for the regulation of the *PSMB9* gene.

My analysis did not reveal any pDMRs associated with *HLA-C, TAP2, TAPBP* and *PSMB9* down-regulation despite the fact the 5-aza-CdR treatment increases *HLA-C* and *PSMB9* expression in both cell lines tested (578T and MCF7) (figure 5.4). This can be explained by: i) the MHC tiling array resolution (2kb) (I have not performed any systematic analysis to define the number of CpG sites required to be differentially

methylated in order for a DMR to be detected), and ii). the possibility that other epigenetic marks (e.g. histone modification) may control the expression of these two genes. It has been shown that histone deacetylase inhibitors induce *TAP* and *TAPBP* genes (Khan et al., 2008).

The identification of a validated DMR within the *PSMB8* promoter in three cell lines (578T, K562 and MCF7) which show the lowest expression for *PSMB8* (enrichment < 0.05) allowed me to speculate that the hypermethylation observed may follow the silencing of the gene by other factors as it has been proposed by previous studies (Clark and Melki, 2002; Meissner et al., 2008; Turker, 2002). Hence, it is possible that the *PSMB8* promoter DMR is not the cause but rather the consequence of gene silencing. Treatment of cell lines other than MCF7, K562 and 578T with methylation inhibitors will be informative.

Finally three DMRs that could be correlated with expression of genes within the TNF cluster were identified, demonstrating the advantage of the unbiased MeDIP-MHC tiling array approach. This agrees with previous data showing the *TNF-*α expression to be controlled epigenetically (Sullivan et al., 2007). My data will be informative for future studies aiming to further elucidate the regulation of the *LTA, TNF-*α and *LTB* genes.

It is worth mentioning that TNF-α, together with IFN-γ, is a cytokine known to be an immune modifier acting on the MHC class I processing and presentation pathway by inducing expression of the *PSMB8, PSMB9, TAP1, TAP2* and MHC class I genes. A kB-like element within the promoter of these genes is responsible for the response upon TNF-α stimuli (Johnson and Pober, 1994). The connection of the DMRs within the TNF cluster and the MHC class I^- phenotype is discussed further in the final discussion (chapter 7) of this thesis.

5.9 Conclusion

A pDMR screen aiming to identify DMRs within the MHC region associated with the MHC class I^- phenotype was conducted. For this purpose the MeDIP-MHC tiling array approach was employed. In total 16 pDMRs were identified and validated further. The presence of a DMR within the *PSMB8* promoter in cells expressing this gene at the lowest level indicated that, in some cases, DNA methylation may be a secondary event during the process of gene silencing. Finally, three DMRs were identified within the TNF cluster, providing a broad basis for better understanding of how genes within this cluster are controlled. The data generated within this chapter will be of great value for future studies regarding MHC associated phenotypes.