Chapter 4

Tissue-specific DMR (tDMR) screen

4.1 Introduction

Tissue specific gene expression in higher eukaryotes involves the activation or silencing of transcription at the appropriate time, and at the right genomic location during cell differentiation. Gene expression is controlled by promoter sequences located immediately upstream of the transcription start site (TSS) of a gene, and by additional regulatory elements located close to the gene that they control, or at a certain distance, or even on a different chromosome (figure 4.1) (Maston et al., 2006).

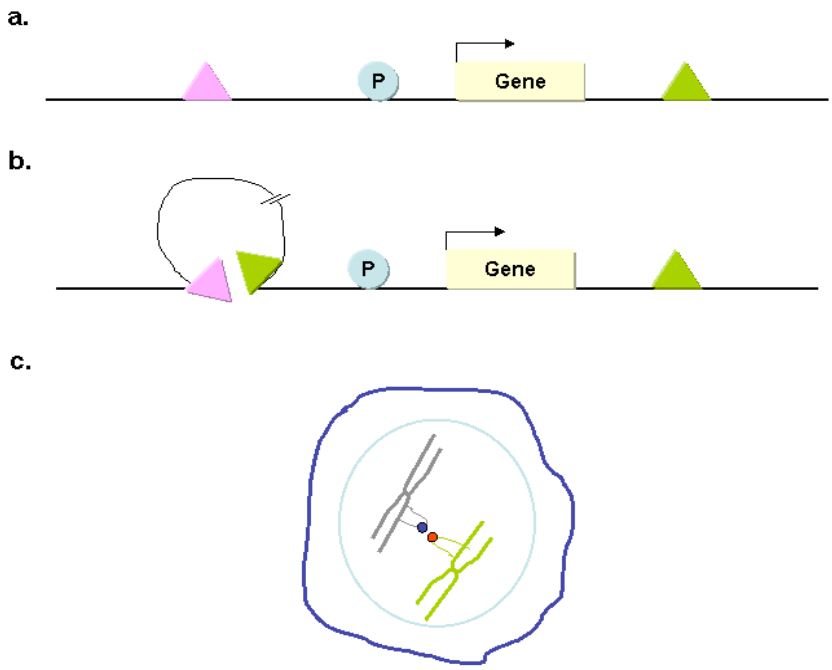

Figure 4.1 **Regulation of gene expression.** a. Linear view of gene regulation. The promoter (P) near the start of a gene provides the minimal information needed for gene expression. The function of the promoter is supplemented by enhancers or silencers farther away (triangles), where regulatory proteins bind to activate or repress transcription of the gene (arrow). b. The looping model of gene regulation. Proteins binding to control regions (triangles) scan through large portions of DNA, looping the intervening region out, until they find the relevant gene. c. Genes from different chromosomes might come into contact when the chromatin containing them loops out from their chromosome 'territory'. Figure was adapted (with some modifications) from Kioussis D., 2005 (Kioussis, 2005).

What determines the gene expression pattern that uniquely defines different tissues with otherwise identical genetic material remains a fundamental biological question.

Landmark publications in the mid-1970s speculated on the role of differential methylation of CpG sites in tissue specific gene expression (Holliday and Pugh, 1975; Riggs, 1975). However this idea remained controversial as a number of studies failed to correlate promoter methylation with expression of known tissue specific genes (Walsh and Bestor, 1999; Warnecke and Clark, 1999). It was almost 20 years later when a study by Futscher BW., et al (Futscher et al., 2002) showed that promoter methylation of the MASPIN gene controls its cell type specific expression.

It is of note that all studies referenced above looked only at the methylation patterns of the promoters of genes with known tissue specific gene expression. However, it should be kept in mind that expression is not controlled solely by promoter regions (figure 4.1). Hence, a different way to investigate if there is a contribution of DNA methylation in tissue specific gene expression is to first identify regions with tissue specific methylation patterns (tDMRs), and subsequently try to correlate them with tissue-specific gene expression.

Advances in methylation analysis technology (section 1.3.6) have eased the way for tDMR identification. Large-scale studies were reported during the past few years that aimed to identify tissue-specific methylation patterns (Eckhardt et al., 2006; Illingworth et al., 2008; Rakyan et al., 2004). One such study, the Human Epigenome Project (HEP), was launched in 1999 and aimed to systematically analyse DNA methylation in the regulatory regions of all known genes in most major tissues and cell types using bisulphite sequencing (Beck et al., 1999). About 25% of the amplicons investigated by the HEP were tDMRs. Interestingly, tDMRs present within CpG islands (CGIs) were located several kilobases away from the nearest annotated genes (Eckhardt et al., 2006). This possibly explains why previous studies reported few tissue specific methylation patterns.

However, when my study was designed (April 2005), only data from the pilot HEP study were available (figure 1.10) (Rakyan et al., 2004). The latter has generated DNA methylation data for about 2.5% of the human MHC region. A significant

proportion (10%) of the MHC loci analysed showed tissue-specific DNA methylation patterns.

In this chapter I will describe a more comprehensive study looking for tDMRs within the entire MHC region. This was part of an effort to identify epigenetic control elements that may be involved in the regulation of MHC genes. For this purpose I generated a MHC tiling array, which I used in combination with MeDIP to enable me to generate methylation data for about 97% of the MHC region (chapter 3). Four samples, also tested as part of the HEP study, were used. In the following sections I present the tDMRs I identified and their characteristics, including correlation with tissue-specific gene expression.

4.2 Samples used for the tDMR screen.

For the tDMR screen, DNA extracted from two tissues (liver and placenta), $CD8^+$ T lymphocyte cells and sperm was used. Two biological replicates of each sample were tested (table 2.1)

These samples were chosen because they have been used previously for methylation analysis across the MHC as part of the HEP study (Eckhardt et al., 2006; Rakyan et al., 2004). Based on HEP, MHC associated tDMRs are present within these tissues. Based on previous studies failing to identify significant sex-specific DNA methylation differences on autosomes (excluding imprinted regions) and the MHC (Rakyan et al., 2004, Weber et al., 2005, Eckhardt et al., 2006), the samples studied here were not controlled for sex.

Tissue-specific DNA methylation profiles of the MHC

Comprehensive methylation profiles of the MHC region were generated using the MHC tiling array in conjunction with Methylated DNA Immunoprecipitation (MeDIP). For this purpose, DNA extracted from 2 tissues (liver and placenta), $CD8^+$ T lymphocytes and sperm was used (section 4.2). I co-hybridised each immunoprecipitated sample with its corresponding untreated sheared control DNA on

the MHC array and analysed the data as described in chapters 2 and 3. Methylation profiles along the MHC region in the four samples tested are shown in figure 4.2.

At this (megabase) resolution, three main observations can be made: (i) The overall profiles correlate significantly ($0.83 < R^2 < 0.93$), suggesting few or no large-scale (>100 Kb) differences in DNA methylation, except perhaps in liver, where some regions appear to be lower in methylation than in other tissues. (ii) As expected from the MeDIP validation qRT-PCR data (figure 3.7) (although CpG density was analysed here), the profiles correlate very well with C+G content, clearly demarcating the boundaries of the MHC class I, II, III and extended class II regions. (iii) The profiles further show the vast improvement in coverage compared to the 253 amplicons, analysed as part of the Human Epigenome Project (Rakyan et al., 2004).

These profiles were used for the identification of tDMRs within these four samples as described in the following section (section 4.4).

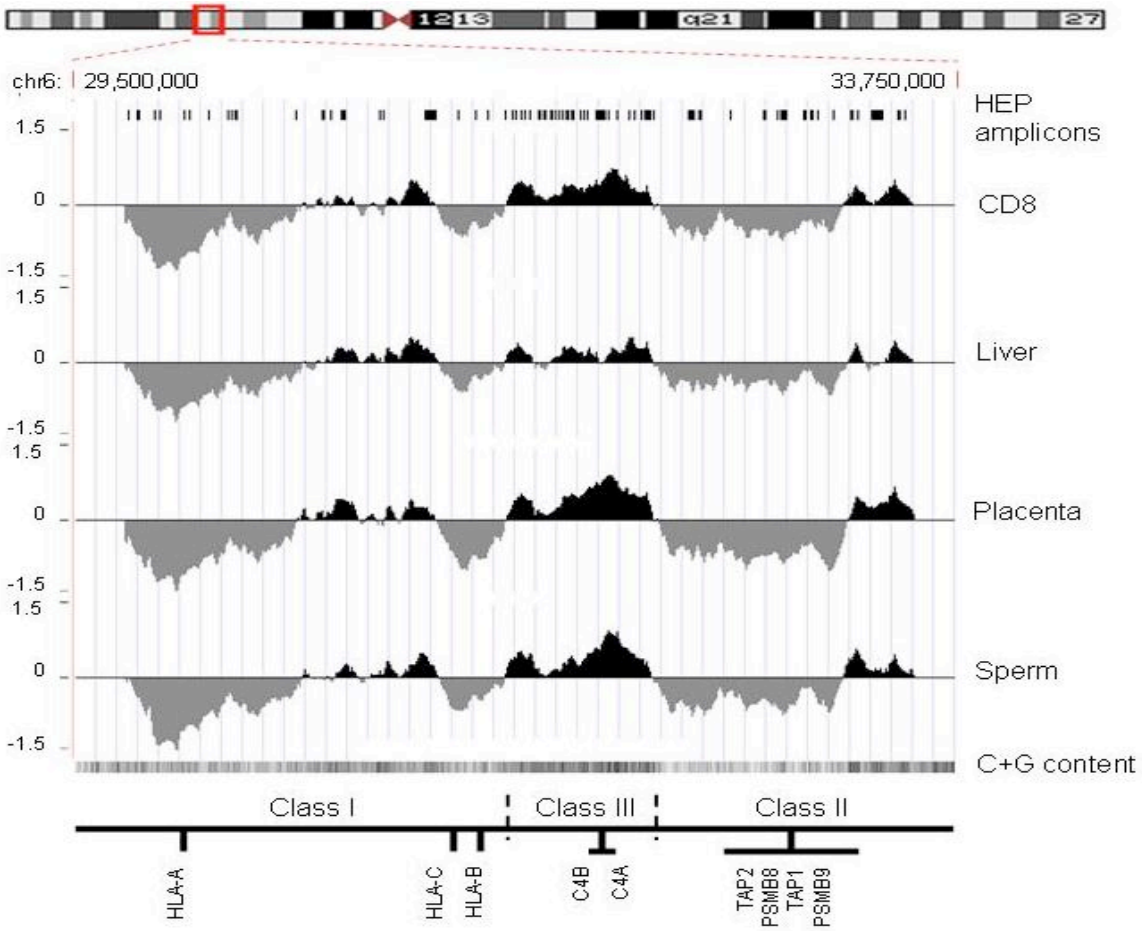

Figure 4.2 **DNA methylation profiles of the MHC.** For each of the four samples tested ($CD8^+$ lymphocytes, liver, placenta, sperm), the log_2 signal ratios (MeDIP/input) were uploaded as individual tracks to the UCSC genome browser using the 'smooth' function. Regions enriched or depleted in DNA methylation are shaded in black and grey, respectively. Also shown are the locations of HEP amplicons and a track of the C+G content (the darker the shading, the higher the C+G content). For orientation, the approximate positions of the MHC class I, II and II subregions and some landmark genes are indicated.

4.4 tDMR identification

For the identification of tDMRs, I performed pair-wise comparisons (six in total – $CD8^+$ T lymphocytes versus placenta, liver versus placenta, placenta versus sperm, $CD8^+$ T lymphocytes versus sperm, liver versus sperm, and liver versus $CD8^+$ T lymphocytes) of the array-derived DNA methylation profiles (section 4.3). At 2 kb, the probe resolution was not high enough to determine if more than one tDMR was contained within a probe or if positive, adjoining probes were part of the same tDMR. Therefore, each differentially methylated probe was considered to be a separate tDMR. According to this definition, I identified a total of 90 putative tDMRs of which 35 were present in more than one comparison (Figure 4.3; Appendix table 4.1).

According to the pair-wise analyses, sperm is most frequently differentially methylated which agrees with the findings of the Human Epigenome Project. The majority of tDMRs identified in sperm are hypomethylated compared to the other samples (65% of tDMRs in placenta-sperm comparison; 93% of tDMRs in CD8-sperm comparison; 32% of tDMRs in liver-sperm comparison). It is known that DNA extracted from sperm is hypomethylated compared to somatic cell DNA (Farthing et al., 2008; Reik, 2007; Schaefer et al., 2007). Notable exceptions are the tDMRs identified in the complement region which seem to be less methylated in liver than any of the other samples. DMRs within this region are discussed further in section 4.6.

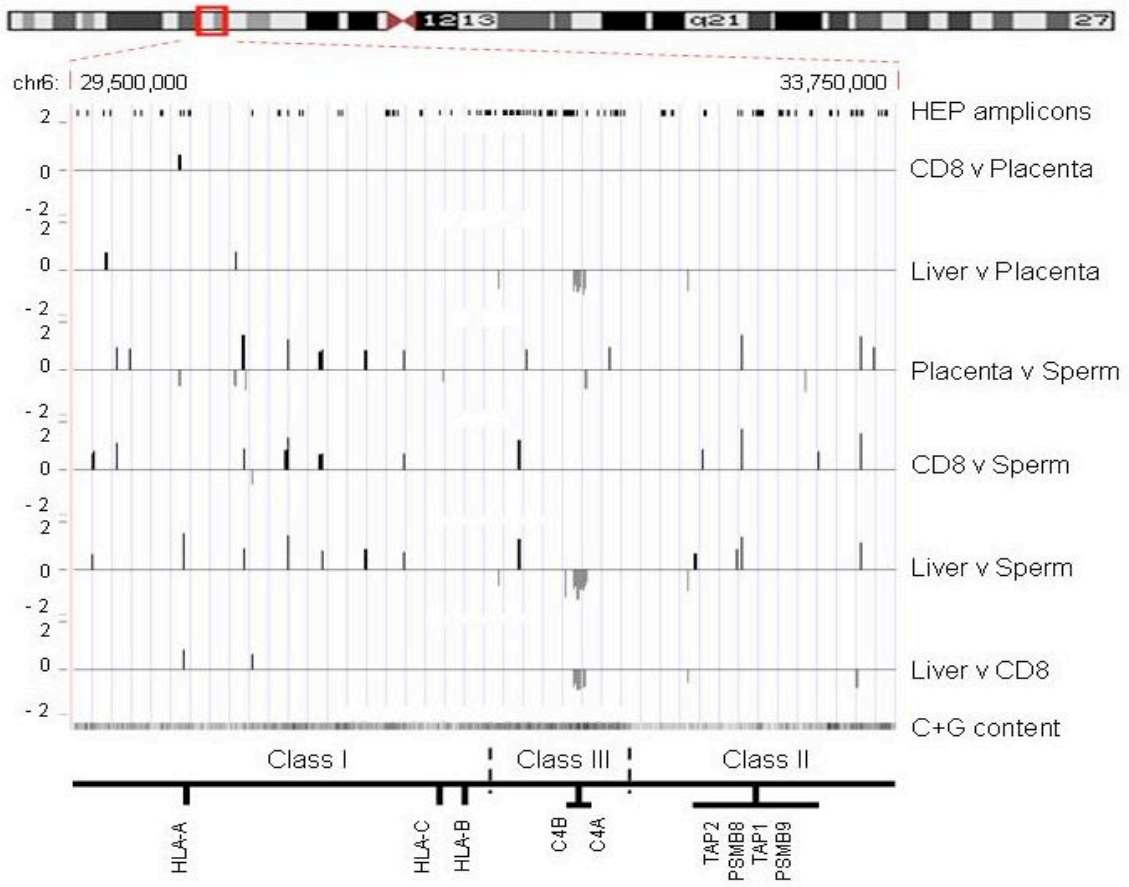

Figure 4.3 **tDMRs within the MHC region.** Pair-wise comparisons (six in total) of the MHC array-derived DNA methylation profiles were performed using t-statistics. A threshold of p-value < 0.001 was used. In total 90 putative tDMRs were identified. Vertical axis shows the $\log_2$ ratio of the two corresponding methylation profiles. Each line represents a tDMR (average size 2kb). Black lines represent tDMRs more methylated in one sample compared to the other (the identities of the pair-wise comparisons are given on the right) whereas grey lines represent less methylation. The majority of tDMRs are present in comparisons with sperm. The locations of HEP amplicons, a track of the C+G content and the approximate positions of the MHC class I, II and II subregions and some landmark genes are also indicated. The Class III region encoding for C4 genes seems to be less methylated in liver.

4.5 Validation of tDMRs

In this section I describe the validation of the tDMRs reported in the previous section (section 4.4). This was done in two steps:

4.5.1. I randomly selected six tDMRs and subjected them to independent methylation analysis using bisulphite DNA sequencing.

4.5.2. I correlated the tDMRs identified by my analysis with the corresponding HEP data (Eckhardt et al., 2006; Rakyan et al., 2004).

4.5.1 Validation of tDMRs by bisulphite sequencing.

I randomly selected six tDMRs, irrespective of their functional relevance, and I subjected them to bisulphite sequencing analysis. The latter is an independent method for methylation analysis (section 1.3.6.1). While with the MeDIP-MHC tiling approach I could only identify methylation differences between two samples (DMRs) (chapter 3), bisulphite sequencing analysis assigns absolute methylation values for each CpG site within a region of about 300- 400 bp. Hence, it is not appropriate to directly compare data generated by the two approaches. However, it is possible to compare relative methylation differences between two samples based on data generated independently by the two methods.

Figure 4.4 shows the genomic locations of the six tDMRs (a), their methylation status based on comparison of their respective MeDIP array profiles (b) and their absolute methylation values based on bisulphite sequencing (c). In all six cases, the bisulphite sequencing results were consistent with the array data, indicating that that the array is suitable for the identification of tDMRs.

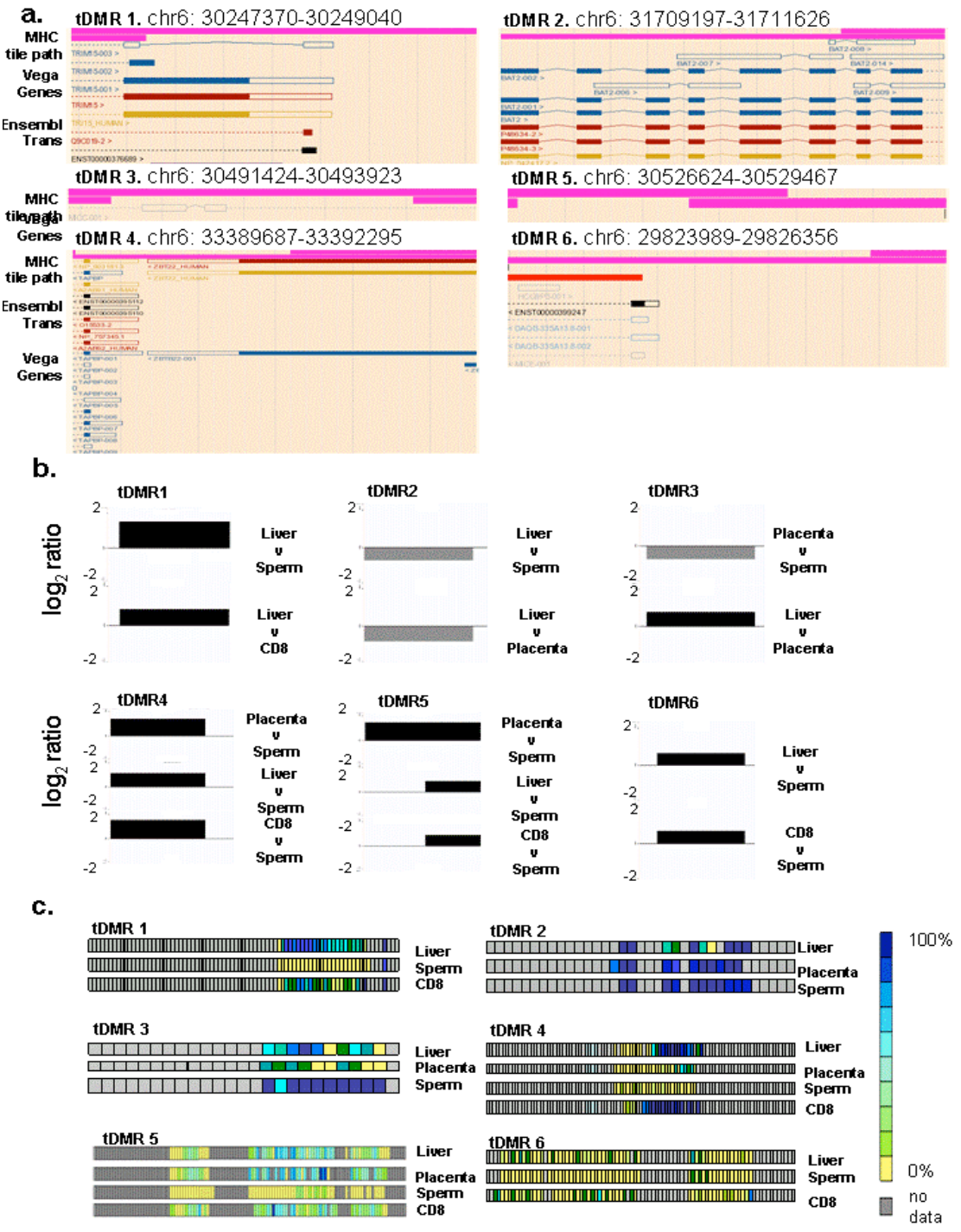

Figure 4.4 **tDMR validation.** Six tDMRs were randomly selected and subjected to bisulphite sequencing analysis. a. Genomic location and associated Ensembl annotation of the six tDMRs. The given chromosome 6 coordinates refer to the tiles (pink track) involved in the formation of these tDMRs. b. tDMR status based on pair-wise comparisons of the $\log_2$ MeDIP enrichment ratios of the indicated tissues/cell types. Black boxes represent tDMRs that are more methylated in sample 1 of the comparison and grey boxes represent tDMRs that are more methylated in sample 2 of the comparison. c. Absolute DNA methylation values of individual CpG sites in tDMRs based on bisulphite sequencing analysis. Because of assay and or technical limitations, bisulphite data could only be obtained for about 50% of the CpG

sites involved in the putative tDMRs. Each square represents a CpG site. The colour code indicates methylation values as calculated by ESME (section 2.2.3). Grey squares indicate CpG sites for which no data could be obtained. Based on this analysis, bisulphite data essentially agree with array data in all cases. It should be noted that I have not performed a systematic analysis defining the smallest size of a DMR that can be identified using the MeDIP-MHC array approach. Therefore, it is not certain if the three CpG sites that appear to be less methylated in liver-tDMR2 are sufficient for detection. It is possible that additional CpG sites are hypomethylated in liver (tDMR2) but bisulphite sequencing analysis was not successful for all CpG sites within this tDMR. Correlation of tDMRs identified by this study with tDMRs identified by MeDIP-chip (Nimblegen arrays - 50bp resolution) (Rakyan et al., 2008) will be informative in determining the minimum number of differentially methylated CpG sites to determine a tDMR. At the time this thesis was written there data were not available.

4.5.2 Correlation with HEP data

As part of the HEP study, a total of 253 unique amplicons corresponding to regulatory exonic and intronic regions associated with 90 MHC-genes were analysed for DNA methylation levels in multiple tissues and cell types (Eckhardt et al., 2006; Rakyan et al., 2004). Of these, 57 amplicons were from liver, sperm, placenta and $CD8^+$ T lymphocytes, the tissues/cell types used here. As it was noted in the previous section and in chapter 3, directly correlating methylation levels of these 57 values with the corresponding MeDIP-MHC tiling array data it is not possible. For this reason I only compared the 55 non-redundant tDMRs (see section 4.7) identified by my analysis with the tDMRs identified by the HEP. There was only one HEP tDMR overlapping with a tDMR identified by my analysis (figure 4.5). Based on MeDIP-MHC array data, this region (chr6:33,389,694-33,391,696) is less methylated in sperm compared to the other samples (figure 4.5a) and this agrees with the HEP values (figure 4.5b).

Eleven additional tDMRs were identified by the HEP within the four samples tested in this study. These tDMRs failed to be identified by the MeDIP-MHC tiling array approach probably due to the low resolution (2kb) of the MHC tiling array. Hence, it is possible that the MeDIP-MHC tiling array approach is not sufficient to detect DMRs smaller than 300 bp (300 bp is the average size of bisulphite sequencing amplicons;

section 2.2.3). On the other hand with the MeDIP-MHC tiling array approach I have managed to identify 54 additional tDMRs that were not reported by the HEP study. This may reflect the lower coverage of the MHC region by the HEP (2.5%) compared to my study (97%).

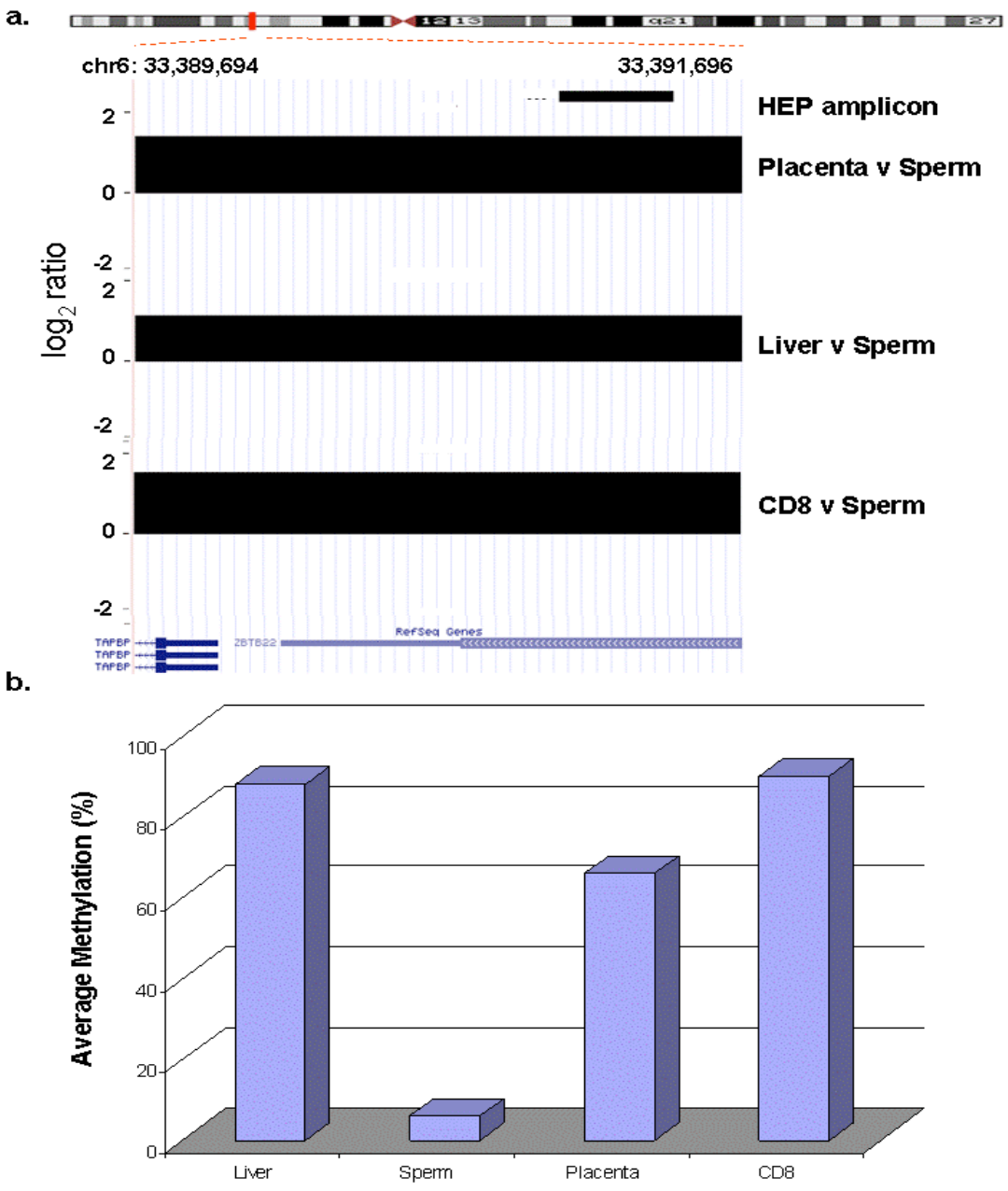

Figure 4.5 **Example of a tDMR identified by both HEP and MeDIP-MHC tiling array studies.** Comparisons of methylation profiles between samples (liver, placenta, sperm and CD8) identified a tDMR common to both studies. a. tDMR status based on pair-wise

comparisons of the $\log_2$ MeDIP enrichment ratios of the indicated tissues/cell types. Black boxes represent a tDMR that is more methylated in sample 1 of the comparison. b. Average methylation values based on HEP data. Absolute methylation levels of each CpG within the HEP amplicon (HEP amplicon ID: 536) were calculated using ESME (Lewin et al., 2004). Average methylation values for all CpGs in each of the four samples are indicated. Sperm is clearly less methylated compared to the other samples.

4.6 Correlation of tDMRs with expression data

I correlated the tDMRs with gene expression using data publicly available from the Genomics Institute of the Novartis Research Foundation Gene Expression Atlas database. This database contains whole-genome mRNA expression data obtained using human U95A Affymetrix microarray chips and mRNA extracted from a number of tissues, including liver, placenta and $CD8^+$ T lymphocytes (sperm was not included in this database) (Su et al., 2002). I identified the probes on the U95A Affymetrix which corresponding MHC loci overlapping with tDMRs according to the liver versus placenta, liver versus $CD8^+$ T lymphocytes and $CD8^+$ T lymphocytes versus placenta comparisons described above. Seven such probes were identified and the genomic features of the corresponding tDMRs are shown in Table 4.1 (see below). One of the probes (Affymetrix ID 40766_at that corresponds to C4A and C4B transcripts) shows a high inverse correlation between expression and methylation at these loci (Figure 4.6). Both loci are highly expressed and hypomethylated in the liver.

4.7 Non-redundant tDMRs

35 out of the 90 identified putative tDMRs were observed in more than one comparison (Figure 4.3; Appendix Table 4.1). Hence, there are 55 loci (average size 2kb), within the MHC region, that according to this study show tissue-specific methylation levels. I define these 55 loci as non-redundant tDMRs (to reflect the non-redundancy at the sequence level) and show their genomic locations in Figure 4.7a and Table 4.1. Based on this definition, about 3% of the MHC loci (average size 2kb)

show tissue specific methylation patterns. This is lower than what was found in the HEP study. According to HEP, 10% of the MHC loci analysed were characterised as tDMRs although only 2.5% of the MHC was tested. This difference can be due to: (i). array resolution and may indicate that there are additional tDMRs which the MeDIP-MHC tiling array approach failed to identify; or (ii). the eight additional samples tested by the HEP study (Eckhardt et al., 2006).

The C4 complement region is the region within the MHC with the highest density of non-redundant tDMRs (18) (Figure 4.7b). As discussed in the previous section, these tDMRs show inverse correlation with C4A and C4B expression patterns.

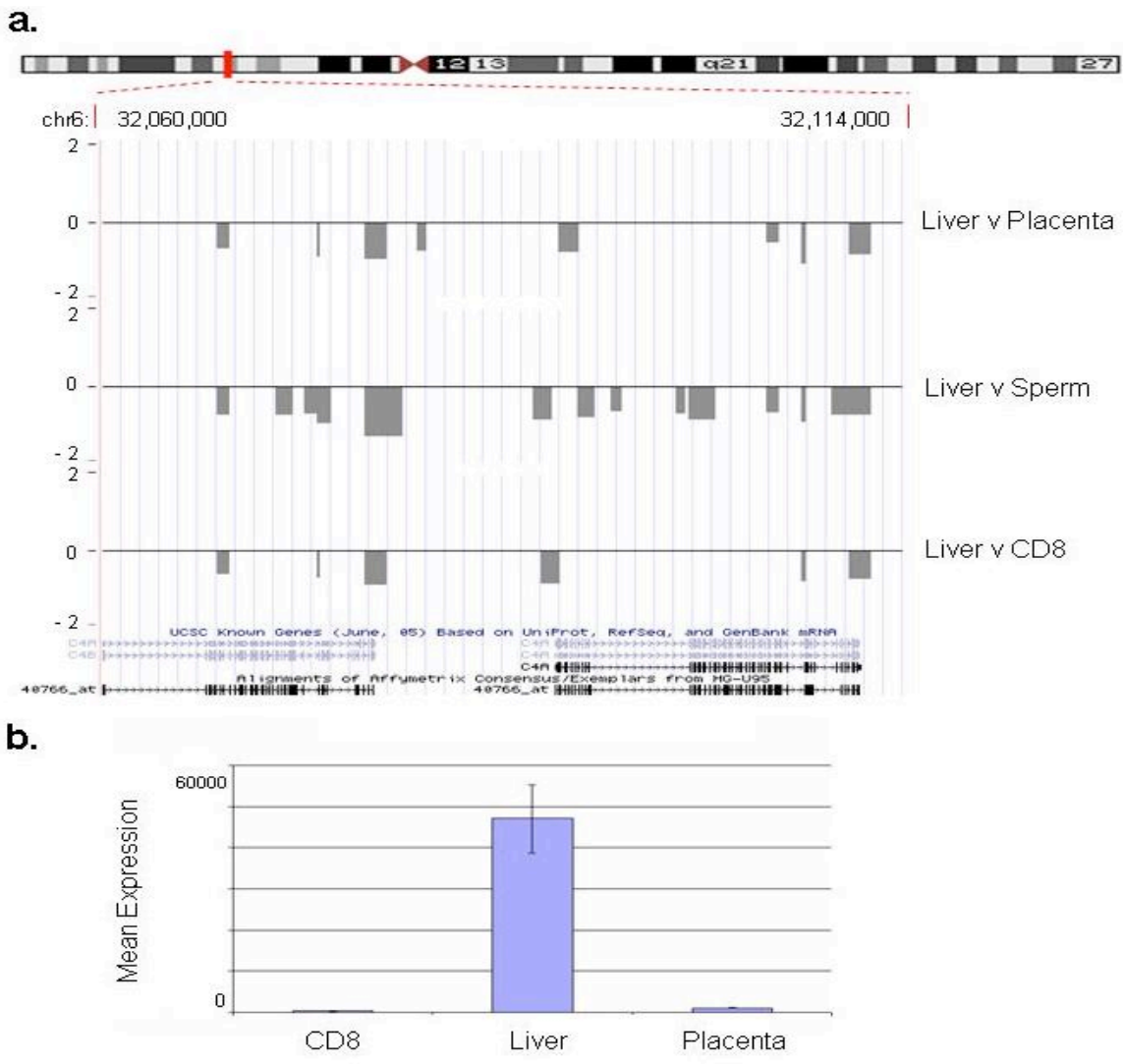

Figure 4.6. **Example of tDMRs correlating with tissue-specific gene expression.** a) tDMRs within the region encoding the C4A and C4B genes. Vertical axis shows the $\log_2$ ratio

of the two corresponding methylation profiles. Grey lines indicate regions (average size 2 kb) that are less methylated in liver compared to the other samples (placenta, sperm, CD8). Known genes and Affy_U95 expression array probes within this region are also shown. b). Expression of C4A and C4B. Graph shows the mean expression values of the probe corresponding to C4A and C4B (Affy_ID: 40776_at) transcripts in three samples tested: CD8, liver, placenta. C4A and C4B transcripts are highly expressed in liver tissue only. Data were taken from the GNF SymAtlas.

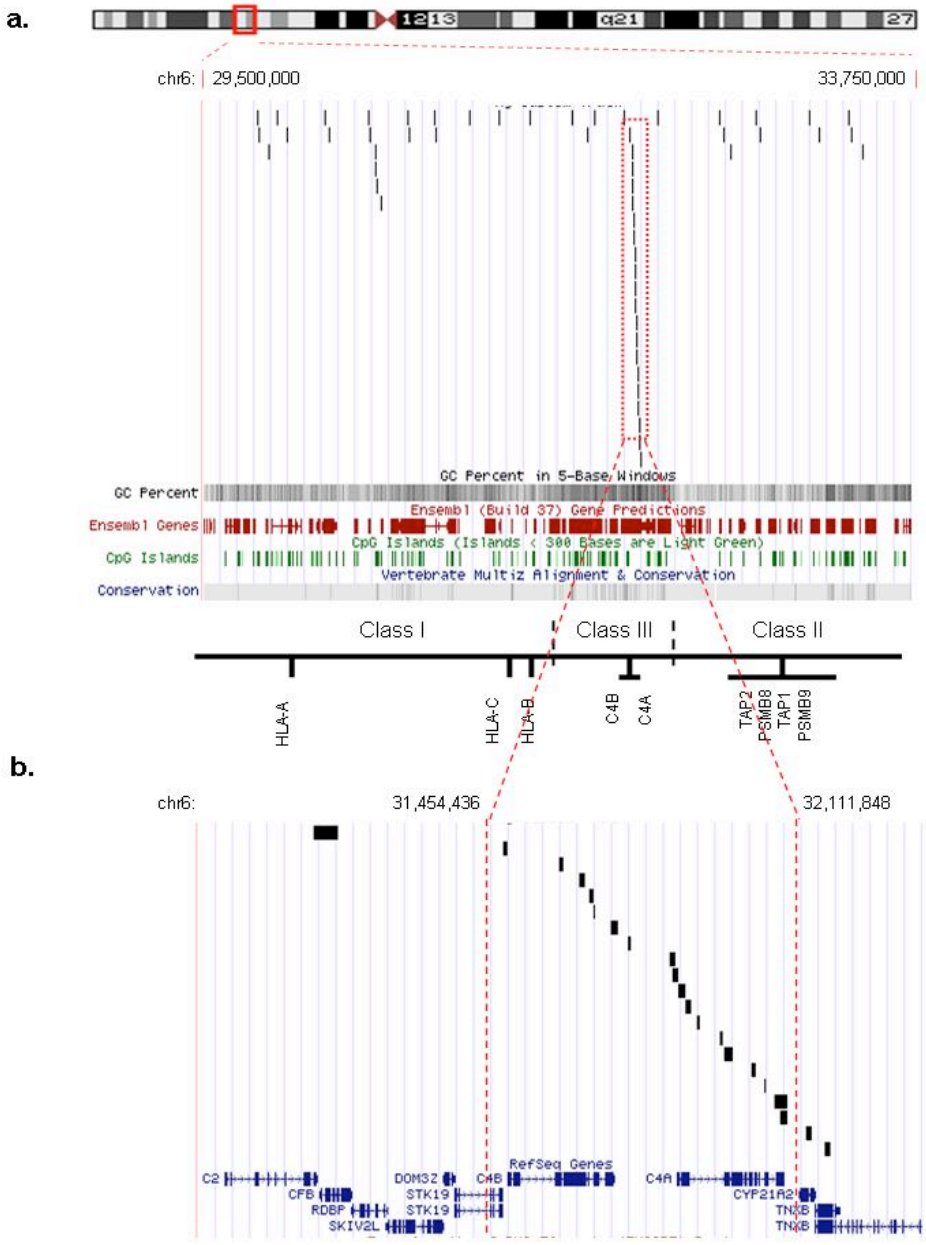

Figure 4.7 **Non-redundant tDMRs within the MHC region**. a. Screen-shot showing the locations of 55 non-redundant tDMRs identified in the MHC region after uploading of the data to the UCSC genome browser. Each vertical black line represents a putative tDMR. The high density of 18 tDMRs within the C4A and C4B complement region is clearly visible (boxed with

red doted line). Tracks showing C+G content, Ensembl genes, CpG islands and conservation are also shown. b. Enlargement of the C4A and C4B complement region showing the 18 overlapping or adjacent tDMRs (delimited by red dotted lines) which could be part of one large tDMR spanning the entire C4 complement region.

4.8 Genomic features of non-redundant tDMRs

To characterize the potential functionality of the 55 non-redundant tDMRs reported in the previous section, I analyzed them for a number of genomic features using the ENSEMBL functional build (Hubbard et al., 2007), as described in section 2.2.11. The result of this analysis is summarized in Table 4.1 and figure 4.8. I found the majority (39) of these tDMRs to map to intragenic regions and the minority (16) to map to intergenic regions. While repetitive elements were overrepresented within the intergenic tDMRs (44%), DNAse I sites and evolutionary conserved elements (ECRs) were overrepresented within the intragenic tDMRs (15%). Furthermore, only 2% of the tDMRs contained transcription start sites (TSS) and about 7% CpG islands and RNA polymerase II binding sites. In all, 21% of the tDMRs contained features significantly ($P < 0.05$) associated with regulation, such as CpG islands, DNase1 and RNA polII binding sites, TSSs and ECRs. Although only few other epigenetic data are yet publicly available for the MHC, I also analyzed the tDMRs for features associated with epigenetic function. Based on this analysis, 6 (11%) tDMRs have insulator protein (CTCF) binding sites, 13 correlated with the transcription-activating histone marks (H3K4me2, H3K36me3, H3K4me3 and H3K4me1) and two with the transcription-silencing mark H4K20me1. Interestingly, 54% of the H3K4me3 sites overlapping with both intragenic and intergenic tDMRs appeared to be close to DNaseI sites. Presence of both H3K4me3 and DNaseI sites indicates promoter regions. Finally, two tDMRs were associated with the histone variant H2AZ.

	chr6 coordinates (NCBI_35)	TSS	CTCF	H4K20me1	PolII	H3K4me2	H3K36me3	H3K4me3	DnaseI	H3K4me1	CpG island	ECR	H2AZ	repeats %	CpG%
1	29823989-29826356		x					x	x		x		x	28.63	2.94
2	30000805-30003606				x			x	x					12.6	9.28
*3	30228982-30231712													26.44	2.64
*4	30247370-30249040		x					x	x		x			4.73	9.1
*5	30565890-30568365			x	x		x	x						11.95	4.77
6	30721858-30724158							x					x	5.91	10.52
7	30731648-30734384													42.71	2.27
8	30891136-30893651													95.08	5.24
9	31709197-31711626											x		79.93	5.01
*10	31803609-31806450		x					x	x			x		0	2.88
11	31841070-31843352											x		0	11.05
12	32020686-32023216											x		18.4	5.61
13	32056738-32058031	x												6.57	3.935
*14	32067481-32068550													0	3.09
*15	32071709-32072864													0	5.61
16	32073608-32074514													0	5.02
17	32074474-32074660													13.67	3.97
*18	32077678-32079121													0	5.35
*19	32081199-32081780													19.53	5.4
20	32088659-32090434													0	3.685
*21	32088718-32090526													0	3.41
*22	32090749-32092076													1.96	4.22
23	32092057-32093147													0	2.38
24	32094350-32095101													100	1.33
25	32098656-32099323													100	3.59
26	32099573-32100214													0	3.875
*27	32104734-32105602													0	5.29
*28	32107212-32107398													0	5.35
29	32109195-32110435													9.995	5.6
*30	32110416-32111859													19.53	5.4
31	32115381-32116535											x		0	7.27
32	32119000-32120024											x		0	7.61
33	32223988-32226638							x			x			4.19	11.09
34	32659407-32660508							x	x		x			9	5
35	32817677-32820582													20.44	2.96
36	32836042-32838492													9.42	7.26
37	33192620-33193912		x											32.79	6.5
*38	33372651-33375048			x	x	x	x	x	x	x		x		10.93	8.76
39	33389687-33392295				x			x				x		2.3	7.51
40	29830203-29832660													6.08	7.77
41	29889483-29892066													21.2	4.15
42	29937894-29939594													38.78	2.48
43	30484481-30486798													96.59	1.04
44	30491424-30493923													1.4	2
45	30526624-30528439							x	x					3.54	7.33
46	30527803-30529467		x					x						3.54	6.73
47	30534798-30537070													42.42	1.26
48	30881555-30884300													98.55	4.27
49	31092038-31094660													74.72	7.51
50	31270669-31273172		x					x						4.59	2.75
51	31454436-31456982													8.59	5.84
*52	32590480-32591619													44.47	2.11
53	32622631-32625110													91.09	5.97
54	33132309-33134479								x					24.37	1.01
55	33450625-33452501													84.66	2.88

Table 4.1. **Genomic features of non-redundant tDMRs.** A total number of 55 non-redundant putative tDMRs (see text for definition – section 4.7) were identified. tDMRs are divided into 2 groups: intragenic and intergenic and their co-ordinates on chromosome 6 are provided. Enrichment of genomic features, including CpG islands, DNAseI sites, TSSs, ECRs, CTCF binding sites, RNA PolII binding sites, H4K20me1, H3K4me2, H3K4me3, H3K36me3 and

H2AZ was tested and marked by symbol 'x' if enrichment was statistically significant ($P < 0.05$). Percent CpG and repeat density were also determined and are shown for each tDMR. tDMRs 14 – 30 (intragenic) and 12 (intergenic) are mapping to the region encoding for C4A and C4B genes. Asterisks indicate the tDMRs that overlap with Affy_U95 expression array probes.

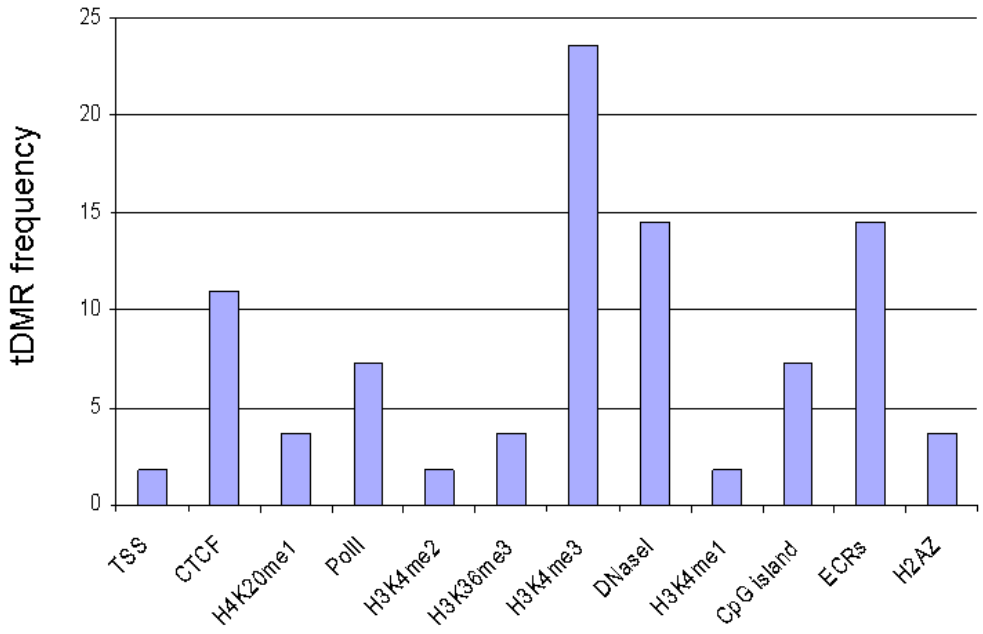

Figure 4.8. **Genomic features of putative tDMRs.** Proportion of putative tDMRs overlapping with genomic features. Histone mark H3K4me3 has the highest frequency whereas TSS, H3K4me2 and H3K4me1 have the lowest.

4.9 Discussion

I used the MHC-tiling array for DNA methylation profiling of four samples previously used for the HEP study: two tissues (liver and placenta), $CD8^+$ T lymphocytes and sperm. Comparison of these profiles allowed me to identify 55 putative, non-redundant tDMRs (90 in total). From these, I randomly selected 10% (6 tDMRs) for validation by an independent method. In all cases, tDMR status could be confirmed, indicating that the array is suitable for DNA methylation analysis and DMR identification. While the analysis carried out here is informative with respect to differential methylation between samples, it did not allow assigning absolute DNA methylation values to each tDMR. This is not a shortcoming of the array but a

limitation of the MeDIP assay which is highly dependent on CpG density as discussed in chapter 3 and illustrated in figure 3.7. Therefore, it was not possible to directly compare my data to the HEP data which, in any case, only cover about 2.5% of the MHC. Only one tDMR was identified by both studies. The on-going development of a novel algorithm employing a Bayesian de-convolution strategy to normalize MeDIP array data for CpG density is likely to overcome this current limitation in the near future (Down et al., 2008). For the same reason as mentioned above, the limited number of samples did not allow me to analyse the data for inter-individual variation which was observed in the HEP (Rakyan et al., 2004) and other studies (Flanagan et al., 2006).

I also correlated gene-associated tDMRs with expression data of the cognate genes available from the GNF SymAtlas. I found a strong correlation within the region encoding, for instance, the fourth component of the human complement (C4). C4 is an essential factor of the innate immunity and consists of two isoforms (C4A and C4B) that differ only in five nucleotides (Szilagyi et al., 2006). C4A and C4B are examples of copy number variants (CNVs) in the human genome. I show that regions within the 5'-UTR, 3'-UTR and the gene body of C4A and C4B are less methylated in liver than in sperm, placenta and $CD8^+$ T lymphocytes. As these two genes are expressed only in liver, it is possible that DNA methylation is the underlying mechanism controlling their expression. At this point, sensitivity and specificity should also be considered. While sensitivity is not an issue in this case (the experimental design normalizes for the genotype of the sample DNA), specificity is. As neither my array nor the Affymetrix U95 array can discriminate between C4A and C4B (which are more than 99% identical), it was not possible to ascertain whether or not these two loci are differentially methylated in this case. Selective hypermethylation is a known mechanism for silencing of duplicated genes (Rodin and Riggs, 2003).

Finally, genomic features associated with the 55 putative tDMRs were identified. Interestingly, only 21% of the tDMRs overlap with known genomic features. It is

possible that the rest of the tDMRs either do not have any genomic function or they are associated with novel genomic features and control elements that may be interesting to investigate further.

4.10 Conclusion

Using MeDIP, I have demonstrated the application of the MHC tiling array for DNA methylation profiling and the identification of tissue-specific differentially methylated regions (tDMRs). Based on the analysis of two tissues and two cell types, I identified 90 tDMRs within the MHC and described their characterisation. Its successful application for DNA methylation profiling indicates that this array represents a useful tool for molecular analyses of the MHC in the context of medical genomics. In the following chapter I describe its application for the investigation of a MHC class I^- phenotype which is commonly associated with cancer (Seliger et al., 2002; The International HapMap Project, 2003).