Abstract. The aim of the present study was to examine the whole-genome DNA methylation status of thymomas and identify differences in thymoma DNA methylation profiles. DNA methylation profiles of tissues (n=12) were studied using the Infinium MethylationEPIC BeadChip microarray (850K) and analyzed in relation to gene expression data. Functional annotation analysis of DNA methylation between the different groups was performed using the online tool GeneCodis3. In order to assess the diagnostic value of candidate DNA methylation markers, receiver operation characteristic (ROC) analysis was performed using the pROC package. A total of 10,014 CpGs were found to be differentially methylated (Δβ>0.2) between two thymoma types (type A and B). Combination analysis showed that 36 genes had differentially methylated CpG sites in their promoter region. ‘Pathways in cancer’, ‘focal adhesion’ and ‘regulation of actin cytoskeleton’ were the most enriched KEGG pathways of differentially methylated genes between tumor and controls. Among the 29 genes that were hypomethylated with a high expression, zinc finger protein 396 and Fraser extracellular matrix complex subunit 1 had the largest area under the curve. The present results may provide useful insights into the tumorigenesis of thymomas and a strong basis for future research on the molecular subtyping of epigenetic regulation in thymomas.

Introduction

Thymic epithelial neoplasms are uncommon thymic neoplasms that arise from epithelial cells of the thymus; they are the most frequent type of anterior mediastinal tumors in adults (1). Thymic epithelial neoplasms are divided into thymomas, thymic carcinomas and thymic neuroendocrine tumors (2). In clinical practice, different classifications have been proposed and used. The latest histological classification released in 2015 by the World Health Organization (WHO) suggested two main thymoma types; types A and B. More specifically, thymomas can be classified into five histological types (A, AB, B1, B2 and B3) based on the morphology of epithelial cells and the lymphocyte to epithelial cell ratio (3). Traditionally, the most commonly used classifications are the Masaoka and Masaoka-Koga staging systems (4). Thymomas are known to be associated with a variety of immunological diseases (5). Myasthenia gravis (MG) is the most frequent syndrome accompanying thymomas and occurs in 15‑20% thymoma patients (6,7).

The etiology and molecular pathogenesis of thymoma has not yet been elucidated. There are various mechanisms by which the pathogenesis of thymoma occurs, including epigenetic alterations, which are a hallmark of cancer due to their role in carcinogenesis initiation (8,9). Recent evidence has indicated that miR-145-5p is an important epigenetic regulation factor that may be involved in tumor progression and treatment response in thymic epithelial tumors (10). A set of prognostic and subtype-specific potential miRNAs have been identified in thymoma (11). A large miRNA cluster on chr19q13.42 was revealed as a transcriptional hallmark of type A and AB thymomas (12). Previous research has also provided evidence that DNA hypermethylation in promoter regions and global DNA hypomethylation serve an important role in the tumorigenesis of thymic epithelial tumors (13,14). Specific DNA methylation aberrations, which are associated with different thymic epithelial tumor histotypes or thymomas accompanied by MG, have previously been identified (15,16).
However, the function of aberrant DNA methylation in thymomas is less clear. The specific DNA methylation aberrations in thymoma vs. control, type A vs. type B thymomas and MG- vs. non-MG-thymomas remains largely unknown. Therefore, an array-based approach was used to uncover genome-wide DNA methylation profiles in fresh frozen thymoma and adjacent normal tissues in the present study. Following differential methylation analysis, a set of differentially methylated CpGs (DMCs) was identified. Furthermore, functional annotation analysis was performed on the corresponding differentially methylated genes. The present study may provide valuable insights into the epigenetic regulation of DNA methylation in thymoma and different thymoma subtypes.

Materials and methods

Study participants. For the genome-wide methylation analysis, eight patients with thymoma or atypical thymic carcinoid undergoing sternotomy were recruited at Peking Union Medical College Hospital (Beijing, China) between October 2014 and July 2015. The patients’ age range was 26-80 years (mean age, 49 years; 1:1 male:female). WHO histological subtypes were recorded as follows: Atypical type A (n=1), type A (n=1), type AB (n=1), type B1 (n=1), type B2 (n=1), type B3 (n=2) and atypical thymic carcinoid (n=1). Atypical thymic carcinoid is an extremely rare thymic neuroendocrine tumor derived from the neuroendocrine system (17). In total, 16 paired surgically resected tumor and adjacent normal tissue samples were collected and stored at -80°C until DNA extraction. Written informed consent was provided by all participants. The present study was approved by the Ethics Committee of Peking Union Medical College Hospital (Beijing, China) and was performed in compliance with the Declaration of Helsinki. Patient demographics and clinical characteristics are presented in Table I.

DNA isolation and bisulfite treatment. Genomic DNA was obtained from both the surgically resected tumor and adjacent normal tissue using the TIANamp Genomic DNA kit (Tiangen Biotech, Co., Ltd., Beijing, China) according to the manufacturer’s instructions. The concentration of extracted DNA was measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific, Inc., Waltham, MA, USA). The quality of extracted DNA was checked in 0.8% agarose gel electrophoresis. Only samples with a purity of 1.8-2.05 were selected for the microarray study. In total, four DNA samples were excluded from further analysis due to poor quality. Finally, DNA samples from six tumors (1T, 2T, 4T, 6T, 7T and 8T) and six normal tissues (2N, 3N, 4N, 5N, 6N and 7N) were maintained for downstream analysis. Genomic DNA (200-500 ng) from each sample was chemically modified and bisulfite-converted using the EZ DNA Methylation kit (Zymo Research Corp., Irvine, CA, USA) according to the manufacturer’s instructions, which converts unmethylated cytosines into uracil. Methylated cytosines remained unchanged during treatment.

Illumina 850K methylation microarray. Following bisulfite treatment, the DNA methylation status of case and control subjects was assayed using the recently developed Infinium MethylationEPIC BeadChip microarray from Illumina, Inc. (San Diego, CA, USA) according to the manufacturer’s instructions, which measured the methylation status of 853,307 CpG sites distributed over the whole genome. The image intensities were extracted using the Illumina iScan system (Illumina, Inc.) and quality-controlled using RnBeads (version 3.5) (18) in R (www.r-project.org/).

Microarray data preprocessing. The Illumina iScan system was used for image and data analysis of the BeadChips. The raw (.idat) files obtained from the methylation microarray were then transferred to the RnBeads software and a quality check of the raw data of each probe analysis was performed, including background correction, adjustment of probe type differences and probe exclusion. The single nucleotide polymorphism-associated probes were filtered, while those corresponding to the sex chromosomes were not. To avoid batch effect, all samples were processed together. Following these intra-sample normalization procedures, DNA methylation was scored as a β value, ranging from 0 (no methylation) to 1 (100% methylation). Unsupervised hierarchical clustering was then performed with Euclidian distance and complete linkage.

Searching for DMCs. To identify DMCs, the average β value was compared between the groups of interest (thymoma tumor tissues vs. adjacent normal tissue; type A vs. type B thymoma; MG- vs. non-MG-thymoma). Briefly, the CpGs were considered DMCs at an average DNA methylation differences (Δβ) between two groups of >0.2. The DMCs for each comparison were determined, following which DMCs were annotated with respect to defined CpG islands (CGIs), shores, shelves and relative to RefSeq genes 3’ untranslated region (UTR), gene body, exon 1, 5’UTR, transcription start site (TSS)1500, TSS200 and intergenic, according to the Infinium MethylationEPIC Microarray annotation file (www.infinium.com).

Expression data of thymomas. Gene expression data of thymomas (GSE29695) (19) were downloaded from the public data repository of Gene Expression Omnibus (GEO; www.ncbi.nlm.nih.gov/gds) (20). These expression data were obtained from 36 patients with thymomas, divided into two main groups: Type A (1 type A and 9 type AB) and B thymomas (20 type B1-B2 and 6 type B3). Genes with a poor signal quality across a maximal number of arrays were removed. As a result, 6,486 genes were found to have signals significantly above background. The raw data were quantile-normalized and log2-transformed prior to statistical analysis. Differential expression analysis was performed between type A and B thymomas. Statistical analysis was performed using Wilcoxon rank-sum tests. P-values were further corrected for false discovery by applying the Benjamini-Hochberg procedure. The genes were defined as differentially expressed genes (DEGs) at P<0.05.

Functional annotation analysis of differentially methylated genes. Once DMCs were annotated to the genes, those that showed significant differences in DNA methylation between different groups underwent functional annotation analysis, using the online GeneCodis3 tool.
Gene Ontology (GO) enrichment analysis was performed to classify the differentially methylated genes into categories of cellular component, biological process and molecular function (22). In addition, the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis was performed to detect the potential pathways of the differentially methylated genes (23).

Receiver operating characteristic (ROC) analysis. In order to assess the diagnostic value of candidate DNA methylation markers, ROC analysis was performed using the pROC package (24) in R. The area under the curve (AUC) was then calculated to assess the performance of each DNA methylation marker.

Results

Illumina 850K methylation microarray of subjects. In total, eight paired tumor samples and corresponding adjacent normal tissues were evaluated. The qualified DNA of six tumor (1T, 2T, 4T, 6T, 7T and 8T) and six normal (2N, 3N, 4N, 5N, 6N and 7N) tissues was used for genome-wide DNA methylation profiling using the Illumina 850K methylation microarray. A Manhattan plot was produced to display P-values that were generated by the -log10 (P-value) function for each CpG site (Fig. 1A). In addition, a volcano plot of CpG sites was constructed using $\Delta\beta$ and P-value, representing the methylation difference between tumors and controls by magnitude of change and statistical significance (Fig. 1B).

Identification of DMCs between tumor and control samples. In order to analyze DNA methylation differences between tumor and control samples, the average $\beta$ values between the groups were examined. A total of 19,118 probes were found to be significantly differentially methylated ($\Delta\beta>0.2$ and adjusted P<0.05), including 119 hypermethylated and 18,999 hypomethylated DMCs. Overall, there was a general decrease in tumor methylation, compared with the control. An unsupervised hierarchical clustering was performed and a heat map of the top 1,000 hypomethylated DMCs and all 119 hypermethylated DMCs was produced (Fig. 2). The heat map showed two robust DNA methylation clusters: One encompassing all tumors and another containing all controls. This indicated that tumors and controls had different DNA methylation characters and patterns.

Genomic features of DMCs between tumors and controls. The methylation categories of DMCs were analyzed in relation to genomic locations. Significant differences were observed between the hypo- and hypermethylated DMCs according to the functional genomic distribution, as well as the CpG content and neighborhood context (Fig. 3).
Of the 119 significantly hypermethylated DMCs, 9% were located in the 3’UTR region, 56% in the gene body, 8% in exon 1, 14% in the 5’UTR, 3% in TSS1500, 9% in TSS200 and none in intergenic region (Fig. 3A). Of the annotated significantly hypermethylated DMCs, 16% belonged to the CGI, 24% to the shore area, 3% to the shelf area, and the remaining 57% to the open sea area (Fig. 3A).

In contrast to hypermethylation, 3% of the 18,999 significantly hypomethylated DMCs were located in the 3’UTR region, 67% in the gene body, 1% in exon 1, 15% in the 5’UTR, 9% in the TSS1500, 3% in the TSS200, and the remaining 1% in intergenic region (Fig. 3B). With regard to CpG content and neighborhood context, 84% belonged to the open sea area of the genome, 9% to the shore area, 6% to the shelf area and 1% to the CGI (Fig. 3B).
.functional annotation of differentially methylated genes between tumor and controls. Of the 119 significantly hypermethylated DMCs, 81 DMCs represented 72 genes. Functional annotation of the 72 genes indicated that the most significantly enriched pathway was ‘natural killer cell mediated cytotoxicity’, as shown in Table II.

In addition, 10,953 of the 18,999 significantly hypomethylated DMCs represented 6,202 genes. To obtain further insight into pathways targeted by the hypomethylated DMCs, further functional annotation was performed. ‘Pathways in cancer’, ‘focal adhesion’, ‘regulation of actin cytoskeleton’, ‘axon guidance’, ‘calcium signaling pathway’ and ‘cell adhesion molecules (CAMs)’ were the most enriched KEGG pathways (Table II).

Identification of DMCs between type A and B thymomas. The DNA methylation profiling of two type A (1T and 2T) and three type B (4T, 6T and 7T) thymomas was obtained using the
Illumina 850K methylation microarray. In total, 10,014 CpGs were differentially methylated at $\Delta\beta>0.2$ and $P<0.001$ between type A and B thymoma subjects, which consisted of 3,998 hypermethylated and 6,016 hypomethylated DMCs. The top 1,000 DMCs were used for unsupervised hierarchical clustering. The results indicated that DNA methylation profiling of type A was significantly distinguished from that of type B thymoma (Fig. 4).

Functional annotation of differentially methylated genes between type A and B thymomas. Based on the combination analysis result of the 36 genes, the GO and KEGG pathway enrichment analysis was performed with a threshold FDR value of $<0.05$. The most enriched biological processes for the seven genes that were hypermethylated with low expression were ‘negative regulation of neutrophil differentiation’, ‘blood vessel endothelial cell differentiation’ and ‘negative regulation of monocyte differentiation’. The most enriched cellular components were ‘focal adhesion’, ‘cytoplasm’ and ‘cortical cytoskeleton’. The most enriched molecular functions were ‘phosphatidylinositol trisphosphate phosphatase activity’, ‘inositol-4, 5-bisphosphate 5-phosphatase activity’ and ‘PTB domain binding’. Furthermore, functional annotation showed that the seven genes that were hypermethylated with a lower expression were highly involved in five KEGG pathways: ‘Insulin signaling pathway’, ‘Fc gamma R-mediated phagocytosis’, ‘Fc epsilon RI singling pathway’, ‘CAMs’ and ‘focal adhesion’ (Table IV).
### Table III. Fifty-five common genes between the methylation and expression data of type A and B thymomas.

| NCBI gene ID | Symbol  | Mean.a  | Mean.b  | log2FC   | P-value.wilox |
|--------------|---------|---------|---------|----------|--------------|
| 51149        | C5orf45 | -2.1029 | -2.5033 | -0.2514  | 0.02117772   |
| 9637         | FEZ2    | 1.3034  | 1.0984  | 0.2468   | 0.00103725   |
| 1363         | CPE     | 1.5048  | 1.0218  | 0.5585   | 0.00042548   |
| 549          | AUH     | 0.3619  | -0.0936 | NA       | 0.00978003   |
| 89797        | NAV2    | 0.3573  | -0.3747 | NA       | 0.00156813   |
| 11069        | RAPGEF4 | -0.5569 | -1.1271 | -1.0171  | 0.00119310   |
| 7919         | BAT1    | 0.3833  | 1.0171  | -1.4078  | 0.03967444   |
| 3834         | KIF25   | -1.5249 | -1.7677 | -0.2131  | 0.04690174   |
| 26269        | FBXO8   | -0.3596 | -0.8502 | -1.2414  | 0.0105367    |
| 11167        | FSTL1   | 0.2920  | -0.3754 | NA       | 0.000672044  |
| 2581         | GALC    | 1.0096  | 0.9132  | 0.1448   | 0.00117772   |
| 23051        | ZHX3    | -1.4316 | -1.0666 | 0.4248   | 0.00430993   |
| 6146         | RPL22   | 1.9214  | 1.7237  | 0.1566   | 0.01434086   |
| 57176        | VARS2   | 0.5042  | 0.8529  | -0.7583  | 0.03164899   |
| 5791         | PTPre   | 0.6483  | 0.3587  | 0.8540   | 0.00544511   |
| 57404        | CYP20A1 | -1.0750 | -0.8289 | 0.3750   | 0.02117772   |
| 2060         | EPS15   | 0.6010  | 0.3958  | 0.7400   | 0.00610416   |
| 80144        | FRAS1   | 1.2244  | 0.5419  | 1.1758   | 5.26x10^5    |
| 3853         | KRT6A   | -0.9689 | -0.0782 | 3.6303   | 0.00156813   |
| 868          | CBLB    | 0.6216  | 0.2545  | 1.2886   | 1.51x10^5    |
| 5573         | PRKAR1A | 1.6307  | 1.3818  | 0.2390   | 0.01925341   |
| 488          | ATP2A2  | 0.6132  | 0.2930  | 1.0654   | 0.00015868   |
| 50640        | PNPLA8  | -0.1182 | -0.0782 | 3.6303   | 0.00156813   |
| 84283        | TMEM79  | -1.3515 | -0.7904 | 0.7737   | 0.00851184   |
| 9521         | EEF1E1  | -0.6646 | -1.1552 | -0.7976  | 0.01748649   |
| 167691       | LCA5    | -1.1525 | -1.9820 | -0.7822  | 7.71x10^5    |
| 10109        | ARPC2   | 1.1761  | 1.3098  | -0.1553  | 0.01925341   |
| 51389        | RWD1    | 0.2295  | 0.7348  | -1.6783  | 0.04690174   |
| 90627        | STARD13 | -1.1529 | -1.6594 | -0.5253  | 0.00484756   |
| 8462         | KLF11   | 0.3155  | -0.0838 | -0.3285  | 0.00264103    |
| 252884       | ZNF396  | -0.0627 | -0.7391 | -3.5582  | 9.29x10^5    |
| 65062        | ALS2CR4 | 1.1346  | 0.9718  | 0.2234   | 0.03338596   |
| 55721        | IQCC    | -1.6380 | -1.8191 | -0.1512  | 0.04690174   |
| 10978        | CLP1    | -1.3344 | -0.9553 | 0.4821   | 0.03967448   |
| 128387       | TATDN3  | -1.6457 | -1.4252 | 0.2075   | 0.03260885   |
| 11252        | PACSIN2 | 1.4083  | 1.1652  | 0.2733   | 0.00264103   |
| 23678        | SGK3    | -0.1551 | -0.5029 | -1.7471  | 0.00089667   |
| 64098        | PARVG   | 0.1042  | 0.5877  | -2.4945  | 0.01925341   |
| 160728       | SLC5A8  | -0.8111 | -0.3890 | 1.0597   | 0.03385967   |
| 81031        | SLC2A10 | 1.0323  | 0.6220  | 0.7306   | 0.00497803   |
| 2            | A2M     | 0.5004  | -0.0924 | NA       | 0.00013317   |
| 3635         | INPP5D  | -1.0257 | -0.7989 | 0.3604   | 0.04316489   |
| 8452         | CUL3    | -0.4168 | -0.9346 | -1.1650  | 0.00299359   |
| 3431         | SPI10   | 0.6534  | 0.8139  | -0.3168  | 0.02117772   |
| 3428         | IFI16   | -0.5547 | -0.0015 | 8.4988   | 0.01925349   |
| 152330       | CNTN4   | 0.0551  | -0.7010 | NA       | 0.00544511   |
| 2690         | GHR     | 0.8282  | 0.2883  | 1.5202   | 0.01296056   |
| 3385         | ICAM3   | -0.8282 | -0.0041 | 7.6322   | 0.04316489   |
| 83605        | CCM2    | 0.9092  | 1.2250  | -0.4301  | 0.02117772   |
| 54518        | APBB1P  | -0.2452 | 0.4220  | NA       | 0.01747864   |
| 81846        | SBF2    | 0.5247  | 0.1140  | 2.2063   | 4.32x10^5    |
In addition, the 29 genes that were hypomethylated with high expression were predominantly involved in the following biological processes: ‘Morphogenesis of an epithelium’, ‘regulation of protein phosphorylation’ and ‘cell communication’. ‘Cytoplasm’, ‘cytosol’ and ‘cAMP-dependent protein kinase complex’ were the most enriched cellular components. ‘cAMP binding’, ‘cAMP-dependent protein kinase regulator activity’ and ‘galactosylceramidase activity’ were the most enriched molecular functions (Table V). No significantly enriched KEGG pathways were identified for the 29 genes that were hypomethylated with high expression at a threshold FDR value of <0.05.

Evaluation of DNA methylation markers for type A and B thymomas. To assess the clinical functionality of DNA methylation markers as diagnostic biomarkers for type A and B thymomas, their sensitivity and specificity were determined using ROC curve analysis. For the seven genes that were hypermethylated with low expression, the AUC of ICAM3 (0.717), APBB1IP (0.748), IFI16 (0.745), PARVG (0.762), CCM2 (0.703), INPP5D (0.741) and SP110 (0.738) was >0.7, as shown in Fig. 6.

For the 29 genes that were hypomethylated with high expression, the AUC of FEZ2 (0.818), PTPRE (0.811), ATP2A2 (0.829), CBLB (0.853), C5orf45 (0.820), CPE (0.853), FSTL1 (0.815), ZNF396 (0.871), FRAS1 (0.885), NAV2 (0.801) and LCAS (0.846) was >0.8, as shown in Fig. 7A-K. Among these 11 genes, ZNF396 and FRAS1 had the largest AUC. For the diagnosis of type A and B thymomas, the sensitivity and specificity of ZNF396 was 84.6 and 90.9% (Fig. 7H), while

Table III. Continued.

| NCBI gene ID | Symbol  | Mean.a | Mean.b | log2FC | P-value.wilox |
|-------------|---------|--------|--------|--------|---------------|
| 27252       | KLHL20  | 0.692117181 | 0.541111423 | 0.355090623 | 0.036418783  |
| 10068       | IL18BP  | 0.367427467 | 0.117710084 | 1.642221566 | 0.043164899  |
| 185         | AGTR1   | -1.260285672 | -1.835884547 | -0.542724545 | 0.002993599  |

Figure 5. Genomic features of DMCs between type A and type B thymomas. (A) Graph showing percentages of hypermethylated DMCs according to their functional genomic distribution and CpG content/neighborhood context. (B) Graph showing percentages of hypomethylated DMCs according to their functional genomic distribution and CpG content/neighborhood context. DMCs, differentially methylated CpGs; uTR, untranslated region; TSS, transcription start site.
that of FRAS1 was 100 and 72.7% (Fig. 7I), respectively. The combination analysis of the above 11 genes increased sensitivity to 96.2% (Fig. 7I). The AUC of all 29 genes was shown in Fig. 8.

**Table IV. Enrichment analysis of the 7 genes hypermethylated with lower expression between type A and type B thymoma.**

| Term                      | ID              | Items                                                                 | FDR        |
|---------------------------|-----------------|----------------------------------------------------------------------|------------|
| Biological processes      | GO:0045659      | Negative regulation of neutrophil differentiation                      | 0.004604   |
|                           | GO:0060837      | Blood vessel endothelial cell differentiation                          | 0.004604   |
|                           | GO:0045656      | Negative regulation of monocyte differentiation                        | 0.005524   |
|                           | GO:0061154      | Endothelial tube morphogenesis                                         | 0.005524   |
|                           | GO:0001885      | Endothelial cell development                                           | 0.006138   |
|                           | GO:0045409      | Negative regulation of interleukin-6 biosynthetic process             | 0.006575   |
| Cellular components       | GO:0005925      | Focal adhesion                                                         | 0.003788   |
|                           | GO:0005737      | Cytoplasm                                                             | 0.011388   |
|                           | GO:0030863      | Cortical cytoskeleton                                                  | 0.015032   |
|                           | GO:0030054      | Cell junction                                                          | 0.015433   |
|                           | GO:0005886      | Plasma membrane                                                        | 0.017133   |
|                           | GO:0005884      | Actin filament                                                         | 0.017957   |
| Molecular functions       | GO:0034594      | Phosphatidylinositol trisphosphate phosphatase activity                | 0.003274   |
|                           | GO:0030487      | Inositol-4,5-bisphosphate 5-phosphatase activity                       | 0.003274   |
|                           | GO:0051425      | PTB domain binding                                                     | 0.004364   |
|                           | GO:0004445      | Inositol-polyphosphate 5-phosphatase activity                          | 0.005727   |
|                           | GO:0005515      | Protein binding                                                        | 0.023332   |
|                           | GO:0005178      | Integrin binding                                                       | 0.0412     |
| KEGG pathways             | hsa04910        | Insulin signaling pathway                                              | 0.031387   |
|                           | hsa04666        | Fc gamma R-mediated phagocytosis                                       | 0.032684   |
|                           | hsa04664        | Fc epsilon RI signaling pathway                                        | 0.036521   |
|                           | hsa04514        | Cell adhesion molecules (CAMs)                                          | 0.035423   |
|                           | hsa04510        | Focal adhesion                                                         | 0.039626   |

FDR, false discovery rate.

**Table V. Enrichment analysis of the 29 genes hypomethylated with higher expression between type A and type B thymoma.**

| Term                      | ID              | Items                                                                 | Hyp_c      |
|---------------------------|-----------------|----------------------------------------------------------------------|------------|
| Biological processes      | GO:0002009      | Morphogenesis of an epithelium                                       | 0.008491   |
|                           | GO:0001932      | Regulation of protein phosphorylation                                | 0.009307   |
|                           | GO:0007154      | Cell communication                                                   | 0.021245   |
|                           | GO:0060512      | Prostate gland morphogenesis                                          | 0.022083   |
|                           | GO:0033003      | Regulation of mast cell activation                                   | 0.022083   |
|                           | GO:0023051      | Regulation of signaling                                              | 0.022083   |
| Cellular components       | GO:0005737      | Cytoplasm                                                            | 0.000212   |
|                           | GO:0005829      | Cytosol                                                              | 0.028305   |
|                           | GO:0005952      | cAMP-dependent protein kinase complex                                 | 0.000883   |
| Molecular functions       | GO:0030552      | cAMP binding                                                         | 0.004318   |
|                           | GO:0008603      | cAMP-dependent protein kinase regulator activity                     | 0.006962   |
|                           | GO:0004336      | Galactosylceramidase activity                                         | 0.011623   |
|                           | GO:0004490      | Methylglutaconyl-CoA hydratase activity                               | 0.011623   |
|                           | GO:0031775      | Lutropin-choriogonadotrophic hormone receptor binding                 | 0.011623   |
|                           | GO:0005509      | Calcium ion binding                                                   | 0.023277   |
Identification of DMCs between MG- and non-MG-thymomas. Methylation microarray analysis was also used to obtain the DNA methylation pattern of the 2 MG- and 4 non-MG-thymomas. Using Δβ>0.2 and P<0.001, 121 DMCs were identified between the MG- and non-MG-thymoma subjects, including 22 hypermethylated and 99 hypomethylated DMCs. The result of unsupervised hierarchical clustering showed that the DNA methylation profiling of MG-thymoma was significantly distinguished from that of non-MG-thymoma (Fig. 9).

The 22 hypermethylated DMCs represented 20 genes, which were enriched in 4 KEGG pathways: ‘Mucin type O-Glycan biosynthesis’, ‘lysine degradation’, ‘p53 signaling pathway’ and ‘phototransduction pathway’. The 99 hypomethylated DMCs represented 73 genes, and no significantly enriched KEGG pathway was found.
Discussion

Epigenetic changes, particularly changes in DNA methylation, are important markers and widely studied in a variety of cancer types (26). The role of DNA methylation in oncogenesis is a topic of interest in the study of cancer biology. A previous report showed that epigenetic events have been implicated in thymomas (27). However, to the best of our knowledge, exhaustive analysis of genome-wide aberrant DNA methylation in the development and progression of thymomas has not been performed. In the current pilot study, the methylation profile of patients with thymoma was described. The landscape of methylation in thymomas was obtained using an illumina 850K methylation microarray.

A total of 19,118 DMCs were identified, 119 of which were hypermethylated and 18,999 hypomethylated. In general, a global decrease in methylation was observed in thymoma tissue, compared with the control. It is known that DNA methylation may be more dynamic outside the CGI (28); the results of the present study also indicated that the majority of DNA methylation changes do not occur at the CGI.

Figure 8. ROC curves for discrimination of type A and type B thymomas. ROC curves and AUC values were generated for 29 genes to compare gene expression in 10 type A and 26 type B thymoma cases from GSE29695. ROC, receiver operating characteristic; AUC, area under the curve.

Figure 9. Unsupervised hierarchical clustering dendrogram of the 121 DMCs between MG-thymoma and non-MG-thymoma. DNA methylation values were represented as colors, with red representing hypermethylated DMCs and green representing hypomethylated DMCs. MG, myasthenia gravis; DMCs, differentially methylated CpGs; T, tumor; N, normal control.
of DMCs were found within the gene body or open sea area. The DNA methylation of FHIT, MLH1, E-cadherin, MGMT, CDKN2A, HPP1 and DAP-K in thymomas has been previously described (13,14,29,30). The present study provided a more extensive list of candidate differentially methylated genes, which may be associated with thymomas. Further studies are required to evaluate the gene expression alterations in thymomas regulated by aberrant DNA methylation.

The latest histological classification recognizes two main thymoma types: A and B (1). The histological subtypes of thymomas seem to be of independent prognostic significance (31,32). Type A thymoma frequently follows a benign clinical course, whereas type B thymoma is considered a low to moderate malignant neoplasm (33). In the present study, the global methylation patterns of types A and B thymomas across the genome were studied. To avoid gender bias, all CpG probe and gene expression data were removed from chromosomes X and Y for the analysis. Differential methylation analysis identified 3,998 hypermethylated and 6,016 hypomethylated DMCs between type A and B thymoma subjects. Genomic features of DMCs also suggested that most DMCs were found within the gene body region or open sea area.

Furthermore, the methylation array data of type A and B thymomas were analyzed in relation to the gene expression array data from the GEO. According to the results, a set of 36 genes showed an inverse correlation between DNA methylation and expression alterations, which may have potential functional consequences, owing to aberrant promoter DNA methylation (TSS1500 and TSS200). Pathway enrichment analysis suggested that the seven genes that were hypermethylated with low expression (ICAM3, APBB1IP, IFI16, PARVG, CCM2, INPP5D and SP110) covered major pathways associated with Fc gamma R-mediated phagocytosis, Fc epsilon RI signaling pathway, cell adhesion molecules and focal adhesion, which serve an important role in tumor development and host-defense mechanisms. Therefore, the present results underlined the importance of aberrant DNA methylation in different subtypes of thymoma.

The methylation status of the 29 genes that were hypermethylated with a low expression were also evaluated using ROC curve analysis to distinguish type A from B thymomas. The results indicated that 11/29 genes (FEZ2, PTPRE, ATP2A2, CBLB, C5orf45, CPE, FSTL1, ZNF396, FRAS1, NAV2 and LCA5) may be potential biomarkers for the diagnosis of type A and B thymomas, with AUC>0.8. It has been reported that diagnostic information may be increased if the methylation of multiple genes is analyzed in combination (34). Herein, it was observed that combination analysis of the 11 genes increased sensitivity to 96.2%. The present results suggested that there are different epigenetic regulation mechanisms for type A and B thymomas. These 11 genes had potential functional consequences in type A and B thymomas, owing to aberrant promoter DNA methylation. Their roles in thymoma subtypes, as well as the utility of these biomarkers in a clinical setting, requires further study in a larger cohort of thymoma subjects.

In conclusion, the present study reported the dysregulated DNA methylation involved in thymoma using the Illumina 850K methylation microarray. Significant changes were observed in the DNA methylomes of thymoma tumor and normal samples, and between type A and B thymomas. To the best of our knowledge, the present study was the first global DNA methylation analysis of thymoma, which may set the foundation for understanding the mechanisms of tumorigenesis in thymoma, as well as for future investigation of epigenetic regulation in type A and B thymomas.

Acknowledgements

We thank Beijing Medintell Bioinformatic Technology Co., Ltd., for assistance in high-throughput sequencing and data analysis.

Funding

This study was funded by the Research Special Fund for the Public Welfare Industry of Health (grant no. 201402001), the National Natural Science Foundation of China (Youth Science Fund) (grant no. 81400068) and the CAMS Innovation Fund for Medical Sciences (grant no. 2016-I2M-1-002).

Availability of data and materials

The datasets used and analyzed during the present study are available from the corresponding author on reasonable request.

Authors’ contributions

YB, YM, and YN designed the study. YB, YM, SL, HL and JH contributed to the materials and performed experiments. YN, YZ, NL and LL performed the data analyses. YZ, NL, XM and JY interpreted the data. LL, XM, JY, BL, ZL and ZW performed the microarray analysis of DNA methylation and contributed significantly to writing the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Written informed consent was provided by all participants. The present study was approved by the Ethics Committee of Peking Union Medical College Hospital (Beijing, China) and was performed in compliance with the Declaration of Helsinki.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Guerrera F, Rendina EA, Venuta F, Margaritara S, Ciccone AM, Novellis P, Novero D, Anile M, Bora G, Rena O, et al: Does the World Health Organization histological classification predict outcomes after thymomectomy? Results of a multicentre study on 750 patients. Eur J Cardiothorac Surg 48: 48-54, 2015.
2. Carter BW, Benveniste MF, Madan R, Godoy MC, Groot PM, Truong MT, Rosado-de-Christenson ML and Marom EM: IASLC/ITMIG staging system and lymph node Map for thymic epithelial neoplasms. Radiographics 37: 758-776, 2017.
3. Marx A, Strobel P, Badve SS, Chalabreysse L, Chan JK, Chen G, de Leval L, Detterbeck F, Girard N, Huang J, et al.: ITMIG consensus statement on the use of the WHO histological classification of thymoma and thymic carcinoma: defined histological criteria, and reporting. J Thorac Oncol 9: 596-611, 2014.

4. Ried M, Marx A, Gotz A, Hamer O, Schalke B and Hofmann HS: State of the art: Diagnostic tools and innovative therapies for treatment of advanced thymoma and thymic carcinoma. Eur J Cardiothorac Surg 49: 1545-1552, 2016.

5. Okumura M, Fujii Y, Shiono H, Inoue M, Minami M, Usumi T, Kadota Y and Sawa Y: Immunological function of thymoma and pathogenesis of paraneoplastic myasthenia gravis. Gen Thorac Cardiovasc Surg 56: 143-150, 2008.

6. Aydemir B: The effect of myasthenia gravis as a prognostic factor in thymoma treatment. North Clin Istamb 3: 194-200, 2016.

7. Jamilloux Y, Frih H, Bernard C, Broussolle C, Petiot P, Girard N and Sève P: Thymoma and autoimmune diseases. Rev Med Interne 39: 17-26, 2018 (In French).

8. Venza M, Visalli M, Biondo C, Oteri R, Agliano F, Morabito S, Teti D and Venza I: Epigenetic marks responsible for cadmium-induced melanoma cell overgrowth. Toxicol In Vitro 29: 242-250, 2015.

9. Venza M, Visalli M, Beninati C, Biondo C, Teti D and Venza I: Role of genetics and epigenetics in mucosal, uveal, and cutaneous melanomagenesis. Anticancer Agents Med Chem 16: 528-538, 2016.

10. Bellissimo T, Ganci F, Gallo E, Assenov Y, Muller F, lutsik P, Walter J, lengauer T and Bock C: Comprehensive analysis of DNA methylation data with RnBeads. Anticancer Agents Med Chem 16: 528‑538, 2016.

11. Wei J, liu Z, Wu K, de Leval L, Detterbeck F, Girard N, Huang J, et al.: ITMIG consensus statement on the use of the WHO histological classification of thymoma and thymic carcinoma: defined histological criteria, and reporting. J Thorac Oncol 9: 596-611, 2014.

12. Radovich M, Solzak JP, Hancock BA, Concès ML, Atale R, Porter RF, Zhu J, Glasscock J, Kessler KA, Badve SS, et al.: A large microRNA cluster on chromosome 19 is a transcriptional hallmark of WHO type A and AB thymomas. Br J Cancer 114: 477-484, 2016.

13. Chen C, Yin N, Yin B and Lu Q: DNA methylation in thoracic neoplasms. Cancer Lett 301: 7-16, 2011.

14. Mokhtar M, Kondo K, Namura T, Ali AH, Fujita Y, Takai C, Cherry JM, Davis AP, Dolinski K, Dwight S, Eppig JT, et al.: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25-29, 2000.

15. Hirose Y, Kondo K, Takizawa H, Nakagawa Y, Toba H, Kajiura K, Suzuki M, Chen H, Shigematsu H, Ando S, Iida T, Nakajima T, Fujisawa T and Kimura H: Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts. Carcinogenesis 36: 1453-1463, 2015.

16. lopomo A, Ricciardi R, Maestri M, De Rosa A, Melfi F, Lucchi M, Mussi A, Coppede F and Migliore L: Gene‑specific methylation analysis in thymomas of patients with myasthenia gravis. Int J Mol Sci 17: 2121‑2131, 2016.

17. Zhu S, Wang ZT, Liu WZ, Zong SX and Li BS: Invasive atypical thymic carcinoma: Three case reports and literature review. Oncol Targets Ther 9: 6171-6176, 2016.

18. Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T and Bock C: Comprehensive analysis of DNA methylation data with RnBeads. Nat Meth 11: 1138-1140, 2014.

19. Badve S, Goswami C, Gokmen-Polar Y, Nelson RP Jr, Henley J, Miller N, Zaheer NA, Sledge GW Jr, Li L, Kessler KA, et al.: Molecular analysis of thymoma. PLoS One 7: e42669, 2012.

20. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al.: NCBI GEO: Archive for functional genomics data sets--Update. Nucleic Acids Res 41: D991-D995, 2013.

21. Tabas-Madrid D, Nogales-Cadenas R and Pascual-Montano A: GeneCodis3: A non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res 40: W478-W483, 2012.

22. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight S, Eppig JT, et al.: Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 25: 25-29, 2000.

23. Kanehisa M and Goto S: KEGG: Kyoto encyclopedia of genes and genomes. Nucl Acids Res 28: 27-30, 2000.

24. Carey V and Redestig H: ROC: Utilities for ROC, with uarray tool for functional genomics. Nucleic Acids Res 40: W478‑W483, 2012.

25. xie Y, liu J, Benbrahim-Tallaa L, Ward JM, Logsdon D, Diwan BA and Waalke MP: Aberrant DNA methylation and gene expression in layers of newborn mice transplacentally exposed to a hepatocarcinogenic dose of inorganic arsenic. Toxicology 236: 7-15, 2007.

26. Vizoso M, Puig M, Carmona FJ, Maqueda M, Velázquez A, Gómez A, Labernadie A, Lugo R, Gabasa M, Rigat-Bragarolas LG, et al.: Aberrant DNA methylation in non-small cell lung cancer-associated fibroblasts. Carcinogenesis 36: 1453-1463, 2015.

27. Hirobayashi H, Fujii Y, Sakaguchi M, Tanaka H, Yoon HE, Komoto Y, Inoue M, Miyoshi S and Matsuda H: GeneCodis3: A non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res 41: D991-D995, 2013.