PERSPECTIVE

The CpG island methylator phenotype (CIMP): What’s in a name?

Laura A.E. Hughes a, b, Veerle Melotte b, Joachim de Schrijver c, Michiel de Maat d, Vincent T.H.B.M Smit e, Judith V.M.G. Bovée e, Pim J. French f, Piet A. van den Brandt a, Leo J. Schouten a, Tim de Meyer c, Wim van Criekinge c, Nita Ahuja g, James G. Herman g, Matty P. Weijenberg a and Manon van Engeland b *

Depts. of a Epidemiology and b Pathology, GROW-School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, the Netherlands;
c Dept. of Mathematical Modeling, Statistics and Bioinformatics, Ghent University, Ghent, Belgium;
d Dept. of Surgery, Orbis Medical Center, Sittard-Geleen, the Netherlands;
e Dept. of Pathology, Leiden University Medical Center, Leiden, the Netherlands;
f Dept. of Neurology, Erasmus University Medical Center, Erasmus University, Rotterdam, The Netherlands;
g The Johns Hopkins University School of Medicine, Baltimore, USA

* Corresponding author:
Manon van Engeland, dept. of Pathology, GROW-School for Oncology and Developmental Oncology, Maastricht University Medical Center, PO. Box 616, 6200 MD, Maastricht, The Netherlands. T: +31 433874622, F: +31 433876613; E: manon.van.engeland@mumc.nl
Abstract

Although the CpG island methylator phenotype (CIMP) was first identified and has been most extensively studied in colorectal cancer, the term ‘CIMP’ has been repeatedly used over the past decade to describe CpG island promoter methylation in other tumor types, including bladder, breast, endometrial, gastric, glioblastoma (gliomas), hepatocellular, lung, ovarian, pancreatic, renal cell and prostate cancers, as well as for leukemia, melanoma, duodenal adenocarcinomas, adrenocortical carcinomas, and neuroblastomas. CIMP has been reported to be useful for predicting prognosis and response to treatment in a variety of tumor types, but it remains unclear whether or not CIMP is a universal phenomenon across human neoplasia or if there should be cancer specific definitions of the phenotype. Recently, it was demonstrated that somatic isocitrate dehydrogenase-1 (IDH1) mutations, frequently observed in gliomas, establish CIMP in primary human astrocytes by remodelling the methylome. Interestingly, somatic IDH1 and IDH2 mutations, and loss-of-function mutations in ten-eleven translocation (TET) methylcytosine dioxygenase-2 (TET2) associated with a hypermethylation phenotype, are also found in multiple enchondromas of patients with Ollier disease and Mafucci syndrome, and leukemia, respectively. These data provide the first clues for the elucidation of a molecular basis for CIMP. Although CIMP appears a phenomenon that occurs in various cancer types, the definition is poorly defined and differs for each tumor. The current perspective discusses the use of the term CIMP in cancer, its significance in clinical practice, and future directions that may aid in identifying the true cause and definition of CIMP in different forms of human neoplasia.
Introduction

Unraveling the complexities of the epigenetic code has been instrumental in advancing our understanding of cancer etiology. It is now clear that epigenetic modifications including aberrant DNA methylation, histone modifications, chromatin remodeling, and non-coding RNAs play a significant role in cancer development [1]. Because such processes do not induce changes in the DNA sequence, but rather are self-propagating molecular signatures that are potentially reversible [2, 3], they provide novel targets for diagnosis and treatment strategies [1, 4, 5].

DNA hypermethylation of promoter-associated CpG islands of tumor suppressor and DNA repair genes, which leads to transcriptional silencing of these genes, has been the most studied epigenetic alteration in human neoplasia [1]. Widespread CpG island promoter methylation, also referred to as the CpG island methylator phenotype (CIMP), was first identified [6] and has been extensively studied in colorectal cancer (CRC). Recently, we systematically reviewed the body of CRC CIMP research and concluded that because there is no universal standard or consensus with respect to defining CIMP, establishing the true prevalence of CIMP in CRC will be challenging until its biological cause is determined [7]. Despite these limitations identified in CRC research, the term ‘CIMP’ has been repeatedly used over the past decade to describe the increased prevalence of CpG island promoter methylation in other tumor types, including bladder [8], breast [9-11], endometrial [12, 13], gastric [14-19], glioblastoma (gliomas) [20-22], hepatocellular [23-26], lung [27, 28], ovarian [29], pancreatic [30], prostate [31] and renal cell [32] cancers, as well as in leukemia [33-36], melanoma [37], duodenal adenocarinomas [38] adrenocortical carcinomas [39], and neuroblastomas [40, 41]. The primary purpose of these studies appears to have been to determine if CIMP is also present in these cancers, and if it can be used to distinguish between known phenotypes of the respective cancer type. However, in many cases, the
observation of CIMP for a tumor results from a self-fulfilling definition, where a subgroup of tumors with a greater degree of DNA methylation than the remaining tumors constitutes CIMP.

Although CIMP has been associated with environmental- and lifestyle factors [3, 42-48], the molecular basis for CIMP is only beginning to be explored. The first clues came from two studies showing that glioblastomas with a hypermethylator phenotype are associated with somatic mutations in isocitrate dehydrogenase-1 (IDH1) [20, 21], and that somatic mutations in IDH1, IDH2 as well as loss of function mutations in ten-eleven translocation (TET)-methylcytosine dioxygenase-2 (TET2) establish a hypermethylation phenotype in leukemia [49]. These are the first indications for a molecular basis of CIMP, and provide an explanation for a very distinct set of tumors with increased levels of hypermethylated DNA. Consequently, these studies have provided a framework for understanding the interplay between genetic and epigenetic changes, and also raise questions about the causes and importance of CIMP in other tumors types. Is ‘CIMP’ a universal phenomenon across human neoplasia caused by similar defects and characterized by similar hypermethylomes, or are there tumor type-specific causes and tumor type-specific definitions of the phenotype?

Addressing these questions is essential for directing research at exploiting CIMP. Here, we discuss the evolution in our understanding of CIMP in various tumors types and how the recent characterization of the human cancer genome and epigenome may influence future research.

CIMP: Roots in colorectal cancer

Molecular characteristics of CIMP tumors

Prior to any discussion on CIMP, it is important to briefly describe CIMP in CRC, as much of the research surrounding CIMP in other cancer types is based on this body of
evidence. It has been more than a decade since Toyota et al. first identified CIMP in CRC [6]. CRC tumors characterized by CIMP have distinctly different histology when compared to tumors derived from traditional adenoma-carcinoma pathway [50-53]. An early event in CIMP tumors appears to be the V600E BRAF mutation [53]. A tight association between the V600E BRAF mutation and CIMP, and mice data showing that the V600E BRAF mutation in the mouse gut induces increased DNMT3B expression, de novo methylation, and downregulation of specific CpG dinucleotides in p16INK4A exon 1, have been reported [54]. However, there is no functional evidence supporting that the V600E BRAF mutation is causal for CIMP. Therefore, it remains possible that BRAF mutation is a surrogate marker for another causal gene. Furthermore, most CIMP CRCs are characterized by promoter CpG island hypermethylation of the mismatch repair gene, MLH1, resulting in its transcriptional inactivation. Loss of MLH1 is thought to cause microsatellite instability (MSI), a form of genetic instability characterized by length alterations within simple repeated microsatellite sequences of DNA [51, 55]. Once MLH1 is inactivated, the rate of progression to malignant transformation is rapid [53].

In 2006, a major advance was made in CIMP research by using unsupervised hierarchical cluster analysis of methylation data; Weisenberger et al. identified a robust five-gene panel that recognized a distinct, heavily methylated subset of colorectal tumors that were also characterized by the V600E BRAF mutation and MSI [56]. This panel proved the validity of the phenotype in CRC, which has been further substantiated and validated in a large, population based sample [57]. Since then, combinations of genes in addition to those proposed in the Weisenberger panel have been suggested as the ‘best’ panel [58-61], but the idea that CIMP is tightly linked with the V600E BRAF mutation remains consistent in all studies. However, a cause or molecular mechanism for CIMP in CRC has not yet been identified and thus the sensitivity and specificity of this panel for defining CIMP remains to be established.
Another aspect that needs to be resolved is the question whether CRC CIMP cases should be further subgrouped in CIMP-high and CIMP-low CRCs [58-60, 62-64]. Although CIMP-low CRCs have been associated with KRAS mutations, this group has many clinical and pathological features in common with non-CIMP and consensus on how to define CIMP-low is currently lacking.

**CIMP translated to other cancer types**

From the literature, it is evident that many studies have investigated CIMP on the premise that the phenotype and genes that quantify the phenotype are not cancer type specific, but rather universal. For example, studies involving breast cancer and endometrial cancer have defined CIMP as ‘methylated multigenes in tumors’ [11], and ‘when multiple genes are concurrently methylated’ [13], respectively. The definition of ‘multiple’ is defined by each investigator to provide separations into subgroups of patients. Furthermore, it is not uncommon for researchers investigating tumors types other than CRC to reference the study by Weisenberger et al. [56] as a rationale for studying CIMP as a marker of cancer, even though the results of that study were very specific for CRC, especially for tumors characterized by the V600E BRAF mutation.

In our recent review, we detailed the use of various techniques and multiple gene panels and cut-off thresholds used to classify a CRC tumor as CIMP-positive [7]. Selection of gene panels and cut-off thresholds for defining CIMP and small sample sizes in other tumors types appears to be even more arbitrary than for CRC (table 1). Studies in gastric cancer [14-19] have often been based on the ‘classic’ gene panel first identified in CRC by Toyota et al. [6], prior to the Weisenberger panel [56]. Studies in ovarian cancer [29], breast cancer [11], hepatocellular carcinoma [23, 26], and melanoma [37] have in part chosen gene panels based on observations from CRC or gastric cancer research. It is not our intention to imply that such
studies are inherently flawed, but again, this type of selection assumes that CIMP is a universal process and not cancer specific.

Extensive studies of genetic and epigenetic changes in human cancers demonstrate that the transformation process differs greatly among tumors arising in different organs. Thus, if CIMP is ultimately organ or tissue specific, much of the true picture surrounding prevalence and prognostic value may not be recognized with the use of CIMP markers developed in another tumor type. For example, in a study of CIMP in endometrial cancer, genes were selected based on their high degree of methylation in other malignancies, including CRC [13]. However, a recent molecular characterization of endometrial tumors identified no V600E BRAF mutations in any of the 87 specimens considered [65]. Therefore, selecting a CIMP panel tightly associated with BRAF mutation may not be entirely relevant to quantifying or identifying CIMP in endometrial tumors. Similarly, results from a recent study on duodenal adenocarcinomas suggest that BRAF mutations are not involved in duodenal tumorigenesis, MSI, or CIMP development [38]. If one hypothesizes that CIMP is a general phenomenon, then the cause of CIMP should also be general and similar across different cancer types.

To assess just how universal CIMP is across tumor types requires genome wide characterization of the methylome. This is a relatively new direction in epigenetic research, and to our knowledge, has only been reported for gliomas [20], leukemia [49], breast cancer [10], benign non-hereditary skeletal tumors such as enchondroma [66], as well as most recently, renal cell carcinoma [32], melanoma [67], gastric cancer [68], and oral squamous cell carcinoma [69].

CIMP: genome wide characterization of the methylome

Glioma
Promoter-associated hypermethylation has been commonly reported in gliomas [70-76], but it wasn’t until 2010, when Noushmehr et al. utilized Ilumina array platform technology, that a CpG island methylator phenotype specific for a group of gliomas with distinct molecular and clinical characteristics was established [20]. They referred to this cluster of tumors as ‘G-CIMP’ to imply its specificity for this tumor type. G-CIMP loci were then validated with MethyLight technology, and perfect concordance with G-CIMP calls on the array platforms versus with the MethyLight markers was observed. Consequently, similar prevalence of the phenotype was shown, providing validation of the technical performance of the platforms and of the diagnostic marker panel. Furthermore, Noushmehr et al. showed that G-CIMP was very tightly associated with the somatic isocitrate dehydrogenase-1 (IDH1) mutation, and validated this in an independent subset of tumors [20].

In 2012, additional evidence for a causal role of IDH1 in generating CIMP was presented. Using immortalized human astrocytes, Turcan et al. demonstrated that the mechanistic process behind this involves the IDH1 mutation subtly remodeling the epigenome by modulating patterns of methylation on a genome-wide scale thereby changing transcriptional programs, and altering the differentiation state [21]. The authors suggest that the activity of IDH may form the basis of an ‘epigenomic rheostat’, which links alterations in cellular metabolism to the epigenetic state [21].

Mutations in IDH1 and IDH2 result in a reduced enzymatic activity toward the native substrate isocitrate. Mutant IDH1 catalyzes the reduction of α-ketoglutarate to 2-hydroxyglutarate (2-HG), a potential oncometabolite [77-80] affecting gene expression via various mechanisms. This is first accomplished via competitive inhibition of α-ketoglutarate dependent dioxygenases including Jumonji-C domain-containing histone demethylases (JHDMs), thereby altering histone methylation levels. In addition, 2-HG inhibits the TET family of 5-methylcytosine (5mC) hydroxylases that convert 5mC to 5-
hydroxymethylcytosine (5hmC) via direct competition with α-ketoglutarate resulting in an accumulation of 5-mC and thereby potentially altering the expression levels of large numbers of genes [49, 80]. Finally, a mechanism altering HIF expression is involved [81].

In their recent study, Turcan et al. showed that expression of wild-type IDH1 caused hypomethylation at specific loci, suggesting that both the production of 2-HG and the levels of α-ketoglutarate can affect the methylome [21]. Furthermore, unsupervised hierarchical clustering of methylome data showed that the hypermethylated genes included both genes that underwent de novo methylation as well as genes that originally possessed low levels of methylation but subsequently acquired high levels of methylation. Control astrocytes did not undergo these methylome changes. Mutant IDH1-induced remodeling of the methylome was reproducible and resulted in significant changes in gene expression [21].

Leukemia

For leukemia, the same story can be told. CIMP, defined by methylation of candidate genes, was reported in 2001 and 2002 [33, 36]. However, the mutational and epigenetic profiling data of Figueroa et al. in acute myeloid leukemia (AML) for the first time identified a causal relationship between IDH1, IDH2 and TET2 mutations and (overlapping) hypermethylation profiles and global hypermethylation [49]. Functional support for this relationship was provided in vitro in hematopoietic cells in that expression of mutant IDH1 and IDH2 leads to an increase in DNA methylation indicating that IDH1/2 and TET2 mutations contribute to leukemogenesis through a shared mechanism that disrupts DNA methylation. In vivo evidence comes from a conditional IDH1(R132H) knock-in mouse model which develops increased numbers of early hematopoietic progenitors, splenomegaly and anemia with extramedullary hematopoiesis. These alterations are accompanied by changes in DNA- and histone methylation profiles [82].
Echondroma and spindle cell hemangioma

Supporting the hypothesis that IDH1 mutation leads to DNA methylation, evidence shows that somatic mosaic mutations in IDH1 and, to a lesser extent IDH2, cause enchondroma and spindle cell hemangioma in patients with Ollier disease and Maffucci syndrome [66, 83]. These are rare skeletal disorders in which there is also an increased incidence of glioma [66]. Using Illumina HumanMethylation27 BeadChips, Pansuriya et al. examined possible differences in methylation between enchondromas with and without IDH1 mutations. Unsupervised clustering of the 2,000 most variable CpG methylation sites gave two subgroups, one of which showed an overall higher methylation at the examined CpG sites, and all but one enchondromas with an IDH1 mutation were positive for this ‘CIMP’ [83].

IDH mutations in other cancer types

In addition to glioma (>70%), leukemia (AML:15-30%) and echondroma (87%) and spindle cell hemangioma (70%), somatic IDH1 mutations are also found in sporadic chondrosarcoma (~50%) [49, 84] and at lower frequencies in anaplastic thyroid carcinoma (11%) [85], (intrahepatic) cholangiocarcinomas (10-23%) [86, 87] and melanoma (10%) [88], while in other solid tumors IDH1 mutations are infrequent (<5%) or absent [89, 90]. Interestingly, the IDH1/2 mutations in melanoma are also accompanied by a loss of 5-hmC in melanoma progression [67]. Therefore, it is interesting to speculate whether or not future research to establish the cause of CIMP in other cancer types should focus on genes that are functionally similar to the IDH family, such as TET2, or on totally different genes. More specifically, it remains uncertain whether CIMP in other cancer types is also caused by inhibition of the conversion of 5mC to 5hmC and subsequent demethylation or that other factors are responsible for the accumulation of 5-mC. In addition to CRC, another tumor type lacking IDH1/2 mutations, but with a putative CIMP phenotype, is breast cancer.
Breast cancer

To date, research that has investigated CIMP in breast cancer has not been conclusive [9, 91-94], with some studies going so far as saying that CIMP does not exist in breast cancer as a truly defined phenotype [9]. Recently, Fang et al. used unsupervised hierarchical clustering from data collected with the Infinium Human Methylation27 platform in an attempt to clarify this dispute [10]. Two DNA methylation clusters in a sample of breast cancer with diverse metastatic behavior were identified. One cluster encompassed a portion of hormone receptor (HR)$^+$ tumors (defined as estrogen receptor (ESR1)$^+$/progesterone receptor (PGR)$^+$, cluster 2) and one encompassed tumors that were ESR1$^+$/PGR$^+$ or ESR1$^-$/PGR$^-$ (cluster 1). Cluster 2 tumors had a highly characteristic DNA methylation profile with high coordinate cancer-specific hypermethylation at a subset of loci, similar to the CIMP phenotype seen in CRC. They referred to this as ‘B-CIMP’, and confirmed the composition of the phenotype through two independent clustering algorithms [10]. Although intriguing, these results should be interpreted with caution. Only 39 tumors were examined in the genome wide study, and three genes were chosen to validate the importance for outcome only. Furthermore, the definition for CIMP using these three genes could be interpreted as arbitrary, and the findings have yet to be validated in a separate cohort.

Nevertheless, this study provides interesting data for future studies to consider. For the first time, the question of whether CIMP targeted the same genes in different human tumors types was examined by repeating the hierarchical clustering to assess colon cancer (C-CIMP) and gliomas (G-CIMP) in additional tumors samples. With this analysis, Fang et al. showed that there was large-scale consensus between CIMP genes from the three cancer-types. CIMP in these different malignancies appeared to target many of the same genes, suggesting a common mechanistic foundation. However, despite the observed similarities, there was not 100% overlap between the Polycomb group PcG targets that comprise the B-, C-, and G-
CIMP, which may reflect a degree of tissue or organ specificity [10]. Although this supports the idea that \textit{IDH1} mutation has been determined as the cause of G-CIMP, this is not true for other cancers. The findings must be validated in additional cohorts before firm conclusions can be made.

**CIMP as a prognostic marker**

Through their methodology, the studies of Fang et al. [10] and Noushmehr et al. [20] were able to clearly show distinct clinical characteristics of tumors characterized by B-CIMP and G-CIMP. For instance, B-CIMP tumors were associated with ESR1/PGR status, a lower risk of metastasis, and an improved clinical outcome [10]. G-CIMP has been associated with improved survival, younger age at diagnosis and histological characteristics [20, 22]. Furthermore, using the Infinium array, a recent methylome analysis in a study of patients with primary clear cell renal carcinoma showed that CIMP characterized a specific cluster of tumors associated with aggressiveness and patient outcome [32]. Such findings reiterate that a major motivation for establishing whether CIMP is universal or cancer specific is because of its potential use as a prognostic marker.

Table 2 shows that CIMP is associated with both favorable and unfavorable prognosis, as well as different clinical characteristics, depending on the type of tumors. There are several possible explanations for these discrepancies. First, while CIMP has been identified in different types of cancer, it may simply not be a universal marker of good or bad prognosis. Second, as previously noted, it could be possible that for some cancers, the gene panels and cut-off thresholds used to define CIMP aren’t accurate for defining the ‘true’ phenotype. It is interesting to observe that CIMP is associated with a favorable prognosis for CRC and gliomas, two cancer types for which extensive research has been conducted with respect to
identifying genes that are associated with clinical and molecular features of the tumors, and in studies that included a relatively large number of cases [20, 57].

Moreover, it has been noted that the association of methylation at CIMP genes with good clinical outcome is not universally applicable to methylation at all genes. Methylation of specific candidate genes or groups of genes has been associated with poorer prognosis, and these genes may have an effect on tumors aggressiveness independent of CIMP [10].

Conclusions and future perspectives

Much like what has been observed in the field of CRC research [7], the study of CIMP in other tumor types has been quite heterogeneous in terms of how the phenotype has been defined. Recent studies considering genome wide characterization of the methylome in gliomas and leukemia have shown that CIMP is likely more than just a generic name to be used to describe aberrant methylation.

Although there is some overlap with respect to genes targeted by CIMP in colon cancer, breast cancer and gliomas, and although IDH1 and genes that affect the same (metabolic) pathway, such as IDH2 and TET2, have been demonstrated to be causally involved in of the generation of CIMP in gliomas and leukemia, cancer specific differences still exist and the cause of CIMP in the majority of cancer types remains to be identified. The causal relationship between somatic mutations in genes such as IDH1, IDH2 and TET2 and altered genome-wide DNA methylation profiles generated by next generation sequencing techniques is a promising clue on the cause of CIMP. The fact that these mutations impair histone demethylation and induce repressive histone methylation marks thereby blocking cell differentiation [95] provide clues on the complex relations between specific genetic alterations, CIMP and clinical characteristics such as histological features and prognosis.
In addition, analyzing the relationship between somatic mutations in chromatin remodeling genes and CIMP could yield interesting insights. For example, *AT-rich interactive domain-containing protein 1a (ARID1a)*, a member of the switch/sucrose non-fermentable (SWI-SNF) complex, has been reported to be mutated and inactivated in a subset of gastrointestinal cancers, the majority of which also exhibit another characteristic of C-CIMP namely microsatellite instability (MSI) [96-98].

To unify the field and to establish a standard definition for CIMP we present the following recommendations:

1. CIMP is not a single phenotype in all types of cancer. A simple variation from the standard nomenclature of ‘CIMP’ to make this distinction, such as ‘C-CIMP’ for colorectal cancer CIMP, ‘G-CIMP’ for glioma CIMP, ‘L-CIMP’ for leukemia CIMP and ‘B-CIMP’ for breast cancer CIMP should be adopted for clarity.

2. Multiple reports suggest a third category of CIMP in colorectal cancer by dividing CIMP into CIMP-high and CIMP-low. Although CIMP-low has repeatedly been associated with KRAS mutations, this group has many clinical and pathological features in common with non-CIMP, and thus without evidence that this is a distinct phenotype and without consensus on how to define CIMP-low, the use of CIMP-low should be discouraged.

3. A consensus meeting should be organized to:
   a) Obtain recommended guidelines on the optimal CIMP marker panel for each tumor type. This includes the number of markers in the panel, the specific loci (genes) included, and the defined region examined for methylation in each gene.
b) Obtain recommended guidelines on the method to measure CIMP. If quantitative methods are needed for CIMP classification, defined cutoffs must be established for each marker for subsequent validation.

4. Once CIMP markers have been identified, they should be validated in large, independent, well characterized patient series with clinical follow-up data (molecular pathological epidemiology approach) [99, 100].

5. A research effort for identifying the biological cause of CIMP among tumor types should be implemented once standard criteria for CIMP are established and validated. Focus should be on establishing causal relationships to find the driver(s) of CIMP.

6. Dissemination of the recommended guidelines to the field, as was done for Bethesda MSI markers [101], is crucial in standardizing research in the field of CIMP.

Hopefully, these recommendations will help to establish the true causes, manifestation, and proper definitions of CIMP.

Acknowledgements

This study was financially supported by a Cancer Research Foundation Limburg grant to MvE, MPW and PAvdB. JVMGB is supported by the Netherlands Organization for Scientific Research (917-67-315). PJF is supported by ZonMW project numbers 92003560 and 40-41200-98-9051, 95110051 and stophersentumoren.nl
References


<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>N</th>
<th>Gene panel a</th>
<th>Method</th>
<th>Marker threshold to assign CIMP-H</th>
<th>% CIMP-H b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrenocortical carcinomas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Barreau et al. 2012 [39]</td>
<td>France</td>
<td>51</td>
<td>Genome wide characterization of methylome and MS-MLPA of 33 genes identified in the genome-wide Infinium analysis</td>
<td>Infinium HumanMethylation27 arrays (Infinium HumanMethylation27 arrays)</td>
<td>Clustering analysis</td>
<td>16%</td>
</tr>
<tr>
<td>Bladder cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maruyama et al. 2001 [8]</td>
<td>USA</td>
<td>98</td>
<td>CHD1, RASSF1A, APC, CDH13, FHIT, RARB (RARβ), GSTP1, CDKN2A (p16), DAPK1 (DAPK), MGMT</td>
<td>MSP</td>
<td>≥4/10 genes methylated</td>
<td>16%</td>
</tr>
<tr>
<td>Breast cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bae et al. 2004 [9]</td>
<td>Korea/USA</td>
<td>109</td>
<td>RASSF1A, SCGB3A1(HIN1), TWIST1 (Twist), CCND2 (cyclin D2), RARB (RARβ), THRB (THRB), CDH1 (E-cadherin), ESR1 (ER), BRCA1, GSP1, BAX, RB1 (RB)</td>
<td>MSP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jing et al. 2010 [11]</td>
<td>China</td>
<td>50</td>
<td>RASSF1A, BRCA1, CDKN2A (p16), CDH1, ESR1 (ER), RARB (RARβ), PTGS2 (COX-2), APC, DAPK1 (DAPK), FHIT</td>
<td>MSP</td>
<td>≥3/10 genes methylated</td>
<td>78%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>50</td>
<td>non-tumor serum</td>
<td></td>
<td></td>
<td>9%</td>
</tr>
<tr>
<td>Fang et al. 2011 [10]</td>
<td>USA</td>
<td>39</td>
<td>Genome wide characterization of methylome</td>
<td>EpiTYPER system (Sequenom)</td>
<td>characterized by the presence or absence of coordinate hypermethylation at a large number of genes</td>
<td>44%</td>
</tr>
<tr>
<td>Endometrial cancer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whitcomb et al. 2003 [12]</td>
<td>USA</td>
<td>24</td>
<td>HOXA11, THBS1, THBS2, CTNNB1, VDR, MLH1, CDKN2A</td>
<td>COBRA</td>
<td>≥5/7 genes methylated</td>
<td>“it exists”</td>
</tr>
<tr>
<td>Zhang et al. 2011 [13]</td>
<td>China</td>
<td>35</td>
<td>CDKN2A (P14), CDKN2A (P16), ESR1 (ER), PTGS2 (COX2), RASSF1A</td>
<td>MSP</td>
<td>≥3/5 genes methylated</td>
<td>49%</td>
</tr>
<tr>
<td>Tumor Type</td>
<td>Authors</td>
<td>Country</td>
<td>Sample Size</td>
<td>Genes Assessed</td>
<td>Methylation Assay</td>
<td>Methylation Threshold</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------</td>
<td>---------</td>
<td>-------------</td>
<td>-------------------------</td>
<td>------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td>Duodenal adenocarcinoma</td>
<td>Fu et al. 2012</td>
<td>USA</td>
<td>98</td>
<td>CACNA1G, IGF2, NEUROG1, RUNX3, SOCS1</td>
<td>MethyLight</td>
<td>≥3/5 genes methylated</td>
</tr>
<tr>
<td>Gastric cancer</td>
<td>Toyota et al. 1999</td>
<td>USA</td>
<td>56</td>
<td>MINT1, MINT2, MINT12, MINT25, MINT31</td>
<td>MSP</td>
<td>≥3/5 genes methylated</td>
</tr>
<tr>
<td></td>
<td>Oue et al. 2003</td>
<td>Japan</td>
<td>103</td>
<td>MINT1, MINT2, MINT12, MINT25, MINT31</td>
<td>MSP</td>
<td>≥3/5 genes methylated</td>
</tr>
<tr>
<td></td>
<td>Kim et al. 2003</td>
<td>South Korea</td>
<td>79</td>
<td>MINT1, MINT2, MINT12, MINT25, MINT31</td>
<td>COBRA</td>
<td>≥3/5 genes methylated</td>
</tr>
<tr>
<td></td>
<td>Etoh et al. 2004</td>
<td>Japan</td>
<td>105</td>
<td>CDKN2A (P16), MLH1 (hMLH1), THBS1 (THBS-1), MINT1, MINT2, MINT12, MINT31</td>
<td>MSP</td>
<td>≥3/7 genes methylated</td>
</tr>
<tr>
<td></td>
<td>An et al. 2005</td>
<td>USA</td>
<td>83</td>
<td>MINT1, MINT2, MINT25, MINT31</td>
<td>MSP</td>
<td>≥2/4 genes methylated</td>
</tr>
<tr>
<td></td>
<td>Kusamo et al. 2006</td>
<td>Japan</td>
<td>78</td>
<td>MINT1, MINT2, MINT12, MINT25, MINT31</td>
<td>COBRA</td>
<td>≥4/5 genes methylated</td>
</tr>
<tr>
<td>Gliomas</td>
<td>Noushmehr et al. 2010</td>
<td>272</td>
<td>Genome wide characterization of the methylome</td>
<td>Infinium+ Golden Gate methylation assays</td>
<td>Clustering analysis</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>van den Bent et al. 2011</td>
<td>68</td>
<td>Genome wide characterisation of the methylome</td>
<td>Infinium Human Methylation27 arrays</td>
<td>Clustering analysis + Noushmehr definition</td>
<td>46%</td>
</tr>
<tr>
<td>Hepatocellular cancer</td>
<td>Shen et al. 2002</td>
<td>China, England, USA</td>
<td>85</td>
<td>CDKN2A (p16), CACNA1G, PTGS2 (cyclooxygenase-2), ESR1 (ER), MINT1, MINT2, MINT27, MINT31</td>
<td>MSP</td>
<td>≥2/8 genes methylated</td>
</tr>
<tr>
<td>Study</td>
<td>Country</td>
<td>Sample Size</td>
<td>Genes methylated</td>
<td>Methodology</td>
<td>Percentage</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------</td>
<td>-------------</td>
<td>---------------------------</td>
<td>-------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Zhang et al. 2007 [26]</td>
<td>China</td>
<td>50</td>
<td>CDKN2A (P14), CDKN2B (P15), CDKN2A (P16), TP53 (P53), RB1, ESR1 (ER), WT1 (WT1), RASSF1A, MYC (c-Myc)</td>
<td>MSP</td>
<td>≥5/8 genes methylated</td>
<td>70%</td>
</tr>
<tr>
<td>Cheng et al. 2010 [23]</td>
<td>China</td>
<td>60</td>
<td>CDKN2A (P14), CDKN2B (P15), CDKN2A (P16), CDKN1A (P21), SYK, TIMP3 (TIMP-3), WT1, CDH1 (E-cadherin), RASSF1A, RB1</td>
<td>MSP</td>
<td>≥4/10 genes methylated</td>
<td>32%</td>
</tr>
<tr>
<td><strong>Leukemia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toyota et al. 2001 [36]</td>
<td>USA</td>
<td>36</td>
<td>ESR1 (ER), CACNA1G, MINT2, MINT2, CDKN3A (p16INK4A), THBS1, CDKN2B (p15INK4B), PTC1H (PTC1A, PTC1B), ABCB1 (MDR1), MYOD1 (MYOD), SDCA, GRP37, PITX2, MLH1</td>
<td>Bisulfite-PCR</td>
<td>≥8/14 genes methylated</td>
<td>19%</td>
</tr>
<tr>
<td>Garcia-Manero et al. 2002 [33]</td>
<td>USA</td>
<td>80</td>
<td>ESRI (ER), CDKN2B (p15), CDKN2A (p16), ABCB1 (MDR1), THBS1, THBS2, ABL1 (C-ABL), PT73 (p73), MYOD1 (MYF3), MME (CD10)</td>
<td>Bisulfite-PCR</td>
<td>≥3/10 genes methylated</td>
<td>43%</td>
</tr>
<tr>
<td>Roman-Gomez et al. 2005 [34]</td>
<td>Spain</td>
<td>50</td>
<td>ADAMTS1 (ADAMTS-1), ADAMTS5 (ADAMTS-5), APAF1 (APAF-1), PPP1R1BB (ASPP-1), CDH1, CDH13, DAPK1 (DAPK), DIABLO, DKK3 (DKK-3), LAT51 (LAT-1), LATS2 (LATS-2), KLK10 (NES-1), CDKN2A (p14), CDKN2B (p15), CDKN2A (p16), CDKN1C (p57), PT73 (p73), PARK2 (PARK-2), PTEN, SFRP1/2/4/5 (sSFRP1/2/4/5), PTPN6 (SHP-1), SYK, PYCARD (TMS-1), WI1F1 (WIF-1)</td>
<td>MSP</td>
<td>≥3 methylated genes</td>
<td>76%</td>
</tr>
<tr>
<td>Roman-Gomez et al. 2006 [35]</td>
<td>Spain</td>
<td>54</td>
<td>38 genes involved in cell immortalization and transformation</td>
<td>MSP</td>
<td>≥3 methylated genes</td>
<td>63%</td>
</tr>
<tr>
<td>Figueroa et al., 2010 [49]</td>
<td>USA</td>
<td>385</td>
<td>Genome wide characterization of the methylome</td>
<td>Clustering analyses</td>
<td>--</td>
<td></td>
</tr>
<tr>
<td>Tumour Type</td>
<td>Study</td>
<td>Country</td>
<td>Tumour Count</td>
<td>Candidate Genes</td>
<td>Method</td>
<td>Cut-off/Result</td>
</tr>
<tr>
<td>---------------------</td>
<td>------------------</td>
<td>---------</td>
<td>--------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>Suzuki et al. 2006 [28]</td>
<td>Japan</td>
<td>150</td>
<td>TMEFF2 (HPP1), SPARC, RPRM (Reprimo), RBP1 (CRBP1), RARB (RARβ), RASSF1A, APC, CDH13, CDKN2A (p16)</td>
<td>MSP</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Liu et al., 2006 [27]</td>
<td>China</td>
<td>60</td>
<td>OGG1 (hOGG1), VHL, RARB (RAR-B), MLH1 (hMLH1), SEMA3B, RASSF1A, ZMYND10 (BLU), FHIT</td>
<td>MSP</td>
<td>≥4/8 genes methylated</td>
</tr>
<tr>
<td>Melanoma</td>
<td>Tanemura et al., 2009 [37]</td>
<td>USA</td>
<td>122</td>
<td>WIF1, TFPI2, RASSF1A, RARB (RARβ2), SOCS1, GATA4, MINT1, MINT2, MINT3, MINT12, MINT17, MINT25, MINT31</td>
<td>MSP</td>
<td>--</td>
</tr>
<tr>
<td>Neuroblastoma</td>
<td>Abe et al. 2005 [40]</td>
<td>Japan</td>
<td>140</td>
<td>17 members of PCDHB family, 13 members of PCDHA family, MST1 (HLP), DKFZp451I127, CYP26C1</td>
<td>qMSP</td>
<td>cut-off &gt;40% methylation of PCDHB family members</td>
</tr>
<tr>
<td></td>
<td>Abe et al. 2007 [41]</td>
<td>Germany</td>
<td>152</td>
<td>17 members of PCDHB family, MST1 (HLP), CYP26C1</td>
<td>qMSP</td>
<td>&gt;60% methylation of PCDHB family members and for samples with 40-60% PCDHB methylation, &gt;10% MST1 (HLP) methylation and/or &gt;70% CYP26C1 methylation</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>Strathdee et al., 2001 [29]</td>
<td>Scotland</td>
<td>93</td>
<td>BRCA1, HIC1, MLH1, CDKN2A (p16), TERC (hTR), CASP8, MINT25, MINT31, CDKN2B (p15), TP73 (p73)</td>
<td>MSP</td>
<td>Unclear, although they do make a conclusion about CIMP</td>
</tr>
<tr>
<td>Pancreatic cancer</td>
<td>Ueki et al., 2000 [30]</td>
<td>USA</td>
<td>45</td>
<td>RARB (RARβ), THBS1, CACNA1G, MLH1, MINT1, MINT2, MINT31, MINT32</td>
<td>MSP</td>
<td>≥4/8 genes methylated</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>USA</td>
<td>101</td>
<td>RARB (RARβ), RASSF1A, GSTP1, CDH13, APC, CDH1, FHIT, CDKN2A (p16INK4A), DAPK1 (DAPK), MGMT</td>
<td>MSP</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

* Genes names are reported as HUGO approved gene symbols, between brackets the gene symbols used in the original study

* CIMP-H refers to either CIMP or in the instance that a study reported three CIMP categories, CIMP-high

* Data not reported
Table 2: CIMP and clinicopathological features of different cancers

<table>
<thead>
<tr>
<th>Cancer type</th>
<th>Significant clinical associations</th>
<th>Prognosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adrenocortical Carcinomas [39]</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Bladder cancer [8]</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Breast cancer [10]</td>
<td>subset of hormone positive tumours (ESR1+/PGR+)</td>
<td>+</td>
</tr>
<tr>
<td>Colorectal cancer [56]</td>
<td>female, older age, proximal location, MSI, <em>BRAF</em> mutation</td>
<td>+</td>
</tr>
<tr>
<td>Duodenal adenocarcinomas [38]</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Endometrial cancer [12, 13]</td>
<td>early stage, COX-2 hypermethylation</td>
<td></td>
</tr>
<tr>
<td>Gastric cancer [14-19]</td>
<td>MSI, lymph node metastasis</td>
<td>+/-</td>
</tr>
<tr>
<td>Hepatocellular carcinoma [23-26]</td>
<td>serum α-fetoprotein (AFP), metastasis, TMN staging, CIMP in serum</td>
<td>-</td>
</tr>
<tr>
<td>Leukemia (adult acute lymphocytic) [33]</td>
<td>younger age</td>
<td></td>
</tr>
<tr>
<td>Leukemia (acute myeloid) [36]</td>
<td>younger age</td>
<td></td>
</tr>
<tr>
<td>Leukemia (T-cell acute lymphoblastic) [35]</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>Leukemia (childhood acute lymphoblastic) [34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lung cancer [27, 28]</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Melanoma [37]</td>
<td>advanced stage</td>
<td>-</td>
</tr>
<tr>
<td>Neuroblastoma [40, 41]</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Prostate cancer [31]</td>
<td>high pre-operative serum (PSA) levels, advanced stage</td>
<td>-</td>
</tr>
<tr>
<td>Renal cell carcinoma [32]</td>
<td>Tumor aggressiveness</td>
<td>-</td>
</tr>
</tbody>
</table>
The CpG island methylator phenotype: what's in a name?
Laura A.E. Hughes, Veerle Melotte, Joachim de Schrijver, et al.

Cancer Res  Published OnlineFirst June 25, 2013.