The CpG Island Methylator Phenotype in Colorectal Cancer

Studies on Risk and Prognosis

Anna Dahlin
To my family
## THESIS SUMMARY

<table>
<thead>
<tr>
<th>Paper</th>
<th>Aim</th>
<th>Main findings</th>
</tr>
</thead>
</table>
| I Dahlin AM *et al.*  
*Int J Cancer*  
2008;122(9):2057-61 | To study the relation between prediagnostic plasma vitamin B12 concentrations and CRC risk. | Study subjects with higher plasma concentrations of vitamin B12 had a reduced risk of rectal, but not colon, cancer. |
| II Van Guelpen B, Dahlin AM *et al.*  
*Cancer Causes Control*  
2010;21(4):557-66 | To investigate components of one-carbon metabolism and CRC risk by CIMP subtypes. | Study subjects with the lowest levels of plasma folate had a reduced risk of CIMP-low or CIMP-high CRC. |
| III Dahlin AM *et al.*  
*Clin Cancer Res*  
2010;16(6):1845-55 | To study the effect of CIMP status on CRC patient prognosis. | CIMP-low CRC and MSS CIMP-high CRC were associated with a poorer patient prognosis compared to CIMP-negative CRC. |
| IV Dahlin AM *et al.*  
*Mod Pathol*  
2011 (Epub ahead of print) | To investigate the relation between tumor-infiltrating T cells and CRC patient prognosis. | Patients with tumors highly infiltrated by T cells had a better prognosis, regardless of MSI status. Patients with CIMP-low tumors poorly infiltrated by T cells had a very poor prognosis. |

CIMP, The CpG Island Methylator Phenotype  
CRC, colorectal cancer  
MSI, microsatellite instability  
MSS, microsatellite stable
# TABLE OF CONTENTS

THESIS SUMMARY i
TABLE OF CONTENTS ii
ABBREVIATIONS iii
ABSTRACT v
POPULÄRVETENSKAPLIG SAMMANFATTNING vi
ORIGINAL ARTICLES viii
INTRODUCTION 1
  Colorectal Cancer Incidence and Etiology 1
  Two Pathways of Tumorigenesis 6
  DNA Methylation and the CpG Island Methylator Phenotype 9
  Treatment and Prognosis in Colorectal Cancer 11
  Aims of This Thesis 15
MATERIAL AND METHODS 16
  Study Subjects 16
  Plasma Analyses 19
  DNA Methylation Analyses 20
  Genetic analyses 23
  Immunohistochemistry 24
  Statistics 26
  Ethical Approval 26
RESULTS AND DISCUSSION 27
  One-Carbon Metabolism in CRC Tumorigenesis 27
  Characterization of CIMP in a Northern Swedish Population 31
  CIMP and CRC patient survival 33
  Tumor-Infiltrating T Cells and Patient Survival 35
CONCLUSIONS 37
ACKNOWLEDGEMENTS 38
REFERENCES 40
ABBREVIATIONS

CD3+  CD3 positive
CI     Confidence interval
CIMP  The CpG island methylator phenotype
CpG   Cytosine preceding guanine (dinucleotide)
CRUMS The Colorectal Cancer in Umeå Study
DAB   3,3-diaminobenzidine chromogen
EGFR  Epidermal growth factor receptor
FAP   Familial adenomatous polyposis
HNPCC Hereditary nonpolyposis colorectal cancer
HR    Hazard ratio
IHC   Immunohistochemistry
LOH   Loss of heterozygosity
MONICA Multinational Monitoring of Trends and Determinants in Cardiovascular Disease
MSI   Microsatellite instability
MSS   Microsatellite stable
MSP   Mammography Screening Project
MTHFR Methylene tetrahydrofolate reductase
NSAID Non-steroidal anti-inflammatory drugs
NSHDS Northern Sweden Health and Disease Study
OR    Odds ratio
PCR   Polymerase chain reaction
PMR   Percent of methylated reference
VIP   The Västerbotten Intervention Project
ABSTRACT

Background Colorectal cancer (CRC) is the second most common malignancy in developed countries. The mortality is high, with nearly half of patients dying from the disease. The primary treatment of CRC is surgery, and decisions about additional treatment with chemotherapy are based mainly on tumor stage. Novel prognostic markers that identify patients at high risk of recurrence and cancer-related death are needed.

The development of CRC has been described in terms of two different pathways; the microsatellite instability (MSI) and chromosomal instability (microsatellite stable, MSS) pathway. More recently, the CpG island methylator phenotype (CIMP), characterized by frequent DNA hypermethylation, has been described as an alternative pathway of tumorigenesis. The event of DNA methylation is dependent on one-carbon metabolism, in which folate and vitamin B12 have essential functions.

The purpose of this thesis was to study CIMP in CRC. The specific aims were to investigate the potential role of components of one-carbon metabolism as risk factors for this subgroup of tumors, and the prognostic importance of CIMP status, taking into consideration important confounding factors, such as MSI and tumor-infiltrating T cells.

Methods CRC cases and referents included in the Northern Sweden Health and Disease Study (NSHDS, 226 cases and 437 referents) and CRC cases in the Colorectal Cancer in Umeå Study (CRUMS, n=490) were studied. Prediagnostic plasma concentrations of folate and vitamin B12 were analyzed in NSHDS. In both study groups, CIMP status was determined in archival tumor tissue by real-time quantitative PCR using an eight-gene panel (CDKN2A, MLH1, CACNA1G, NEUROG1, RUNX3, SOCS1, IGF2 and CRABP1). MSI screening status and the density of tumor-infiltrating T cells were determined by immunohistochemistry.

Results An inverse association was found between plasma concentrations of vitamin B12 and rectal, but not colon, cancer risk. We also found a reduced risk of CIMP-high and CIMP-low CRC in study subjects with the lowest levels of plasma folate.

We found that patients with CIMP-low tumors in both NSHDS and CRUMS had a poorer prognosis compared with CIMP-negative, regardless of MSI screening status. We also found that MSS CIMP-high patients had a poorer prognosis compared with MSS CIMP-negative. The density of tumor-infiltrating T cells and CIMP status were both found to be independent predictors of CRC patient prognosis. A particularly poor prognosis was found in patients with CIMP-low tumors poorly infiltrated by T cells. In addition, the density of T cells appeared to be more important than MSI screening status for predicting CRC patient prognosis.

Conclusion Rather than being one disease, CRC is a heterogeneous set of diseases with respect to clinico-pathological and molecular characteristics. We found that the association between risk and plasma concentration of vitamin B12 and folate depends on tumor site and CIMP status, respectively. Patient prognosis was found to be different depending on CIMP and MSI screening status, and the density of tumor-infiltrating T cells.
**POPULÄRVETENSKAPLIG SAMMANFATTNING**

Kolorektal cancer (KRC, cancer i tjock- och ändtarm) är en av de vanligaste cancersjukdomarna i västvärlden, och även i Sverige, där 5-6000 nya fall per år upptäcks. Förutom ålder och en viss ärftlighet, är riskfaktorerna för KRC till stor del kopplade till livsstil och kosthållning. Fysisk inaktivitet, rökning, och alkoholintag ökar risken, liksom en diet bestående av mycket kött och lite frukt och grönsaker.

Cancer uppstår på grund av mutationer i gener som kontrollerar celltillväxt och celldelning. Under senare år har man upptäckt att dessa gener även kan regleras genom felaktig metylering av DNA, och att detta är ett vanligt fenomen vid uppkomsten av KRC. De cellulära processerna som leder till DNA metylering är beroende av metylgrupper, vilket tillhandahålls via kosten och näringsämnet folat. Ett annat vitamin nödvändigt för att DNA metylering ska kunna ske är vitamin B12. Folat finns framför allt i frukt och grönsaker, medan vitamin B12 nästan uteslutande finns i animaliska livsmedel, såsom kött och mjölkprodukter.

I den första av de fyra studier som ingår i den här avhandlingen ville vi undersöka förhållandet mellan vitamin B12 och risken för att utveckla KRC. Genom att studera 226 fall av KRC och 436 friska kontrollindivider, fann vi att risken för rektalcancer var minskad bland personer med högre koncentration av vitamin B12 i plasma, vilket mättes i ett prov som lämnats upp till 12 år före cancerdiagnos.

Med tanke på den viktiga roll som folat och vitamin B12 har i omsättningen av metylgrupper, var vi speciellt intresserade av en grupp av kolorektala tumörer som kallas CIMP (the CpG island methylator phenotype) och som uppvisar en hög grad av DNA metylering. Inför den andra studien klassificerade vi därför tumörerna utifrån antalet metylerade gener. Om metylering inte kunde detekteras i någon av de 8 undersökta generna klassificerades tumörerna som CIMP-negativa. Tumörer med metylering i 1-5 gener klassificerades som CIMP-low, och metylering i 6-8 av generna som CIMP-high. Vi fann i denna studie att risken för uppkomst av en metylerad tumör (CIMP-low/high) var ökad bland personer med högre koncentration av folat i plasma.

Bedömning av patientens prognos, samt beslut om huruvida cytostatikabehandling ska ges efter att tumören avlägsnats med kirurgi, avgörs i dagsläget främst av tumörens stadium. Stadium tar hänsyn till
tumörens utbredning samt spridning till lymfkörtlar eller andra organ. Patienter med tumörstadium II anses ha liten risk för återfall, och ges generellt ingen ytterligare behandling utöver kirurgi. Ca 20% av dessa patienter drabbas trots tumörens låga stadium ändå av metastaserande sjukdom, vilket belyser behovet av ett sätt att identifiera denna grupp av patienter som potentiellt kan dra nytta av onkologisk behandling.

I den tredje studien ville vi därför undersöka hur tumörens CIMP-status påverkar patientens överlevnadsprognos. Utöver CIMP, undersökte vi även mikrosatellit instabilitet (MSI). Tumörer som uppvisar MSI uppkommer på grund av defekter i systemet som repararar de fel i DNA-sekvensen som uppstår vid celldelning, och har unika kliniska och molekylära egenskaper jämfört med de tumörer som inte är av MSI-typ och kallas mikrosatellit stabila (MSS).

Utöver de fall av KRC som ingick i de två ovan beskrivna studierna, kunde vi undersöka dessa faktorer i ytterligare 414 fall. Vi fann att patienter med tumörer av typen CIMP-low hade sämre prognos än CIMP-negativa. En sämre prognos sågs även hos patienter med tumörer av typen MSS CIMP-high.

Ett fenomen som ses ofta i tumörer av typen CIMP-high och MSI, är förekomsten av t-lymfocyter i och runt tumören, vilket representerar immunsystemets reaktion på tumören. En kraftig infiltration av t-lymfocyter i och runt tumören har tidigare förknippats med en god patientprognos, vilket vi också kunde bekräfta i avhandlingens fjärde studie. Vi fann också att mängden t-lymfocyter var viktigare för patientens prognos än tumörens MSI status, vilket tidigare föreslagits vara en viktig prognostisk faktor. En speciellt ogynnsam prognos sågs hos patienter med tumörer av typen CIMP-low som var sparsamt infiltrerade av t-lymfocyter.

Fynden från ovanstående beskrivna studier belyser hur KRC bör studeras med hänsyn till tumörens kliniska och molekylära egenskaper. Dessa egenskaper representerar sannolikt subgrupper av tumörer med olika mekanismer för uppkomst, vilket dessutom kan vara av vikt vid planering av behandling.
This thesis is based on the following original articles

**Paper I**

**Paper II**

**Paper III**

**Paper IV**

The articles in this thesis were reproduced with the kind permission of Wiley Periodicals, Inc. (paper I), Springer Science and Business Media (paper II), The American Association for Cancer Research, Inc. (paper III), and United States and Canadian Academy of Pathology (paper IV).
INTRODUCTION

Colorectal Cancer Incidence and Etiology

Incidence
Colorectal cancer (CRC) is the third most common malignancy in the world, and more than one million patients are diagnosed with the disease each year. The mortality is high, with nearly half of all patients eventually dying.\(^1\) The highest incidences are found in Japan, Australia, New Zealand, North America, and Europe,\(^1\) which reflects the association between a western life style and CRC risk.\(^2\) In contrast to many other malignancies, men and women are equally affected by CRC (ratio 1.2:1).\(^1\) In 2007, the incidence in Sweden was 5 873, including 3 885 colon and 1 988 rectal cancers, which makes the colorectum the third most common site for cancer after breast and prostate in this country.\(^3\)

Etiology
The development of CRC is believed to be influenced by both genetic (inherited) and environmental (lifestyle) factors. CRC is diagnosed in patients of all ages, but the risk is increased, and continues to increase each year, after the age of 50,\(^3\) making age one of the most important risk factors. Heredity is another important risk factor, and a substantially increased risk is found in relatives to CRC patients.\(^4\) However, only a few percent of hereditary cases are caused by known germline alterations.\(^5\)

The majority of CRCs are believed to arise from premalignant adenomatous polyps in the colon and rectum, and the finding of such polyps increases the risk of a CRC diagnosis later in life.\(^6\) The association between CRC risk and inflammatory bowel disease, including ulcerative colitis and Crohn’s disease, is also well established.\(^7\) Some, but not all, studies have observed an increased risk of CRC following cholezystectomy.\(^8\)\(^9\)
Patients with certain other medical conditions also experience an increased risk of CRC. Diabetes mellitus is one such condition, due not only to the typical western lifestyle risk factors associated with both diseases as described below, but possibly also to increased circulating levels of insulin and insulin-like growth factors. Acromegaly is a rare condition in which an increased risk of CRC has been explained by higher circulating levels of insulin-like growth factor-1.

Physical inactivity and obesity have both been associated with a significant increase in CRC risk, although a weaker association has been reported for rectal cancer. Cigarette smoking and a high intake of alcohol are also among the established risk factors for CRC. A protective effect from the use of aspirin and other non-steroidal anti-inflammatory drugs (NSAID) have been shown in large population-based studies and in some intervention studies. Hormone replacement therapy in women has also been shown to reduce the risk of CRC.

**CRC Risk Factors Related to a Western Diet**

The incidence of CRC is higher in countries with a typical “western lifestyle”, characterized by an unhealthy diet, physical inactivity, and obesity. The incidence is also rising in developing countries as they adopt a more western-like lifestyle. CRC incidence rises dramatically within a few generations in immigrants from low-risk to high-risk countries.

The relationship between CRC risk and the western diet, typically high in meats and animal fat, and low in fiber, fruits, and vegetables, has been extensively studied. An increased risk for CRC has been shown in individuals with a high intake of red and processed meat, whereas a protective effect has been seen for diets high in fiber, fruits and vegetables. However, randomized-control trials have not provided support for the latter finding. The role of dietary fat as a CRC risk factor is not established.
The exact mechanisms behind dietary influence on CRC risk are not known, but several micronutrients have been suggested to contribute. For example, higher intake of calcium and vitamin D has been related to a decreased risk of CRC,\textsuperscript{32-33} whereas high body iron stores have been associated with an increased risk.\textsuperscript{34-35} Among the dietary risk factors for CRC, folate has been one of the most extensively investigated because of its important function in one-carbon metabolism, as described in more detail below.

**Dietary Risk Factors Related To One-Carbon Metabolism**

Nucleotide synthesis and DNA methylation are two cellular processes that are both dependent on one-carbon metabolism (figure 1). An imbalance in the pathways of one-carbon metabolism may result in hypo- or hypermethylation of DNA and a reduced level of nucleotide synthesis. This in turn can cause aberrant gene expression and genomic instability, and hence have an influence on carcinogenesis.\textsuperscript{36-40} Methyl groups that are essential for these reactions are provided by folate from dietary sources such as green leafy vegetables, fruits, and liver, or by supplements and fortified foods. Many of the reactions in one-carbon metabolism are enzyme-driven and require cofactors, such as vitamin B12. Vitamin B12 is obtained almost exclusively from animal sources, such as meat, fish, and dairy products.

As described in figure 1, a reduced form of tetrahydrofolate (THF), 5,10-methyleneTHF, is required for both nucleotide synthesis and DNA methylation reactions. Before DNA methylation can occur, 5,10-methyleneTHF must be further reduced to 5-methylTHF. This reaction is driven by the enzyme 5,10-methyleneTHF reductase (MTHFR) and is irreversible. In the next step, homocysteine is remethylated to methionine by the donation of a methyl group from 5-methylTHF. This reaction is mediated by methionine syntase, which is dependent on vitamin B12 as a cofactor. Methionine is further converted to S-adenosylmethionine, which is used for the methylation of DNA and other structures in the cell. Because of their essential functions in one-
carbon metabolism, folate and vitamin B12 are two of the factors that have been studied in the context of cancer etiology in general, and CRC in particular.\textsuperscript{39, 41}

Based on large population-based epidemiological studies, higher plasma levels and dietary intake of folate are generally considered to protect against CRC.\textsuperscript{39, 42-43} However, this was not been found to be the case in the Northern Sweden Health and Disease Study (NSDHS) cohort.\textsuperscript{44} Also, chemopreventive studies have not supported a protective effect of folate supplementation on CRC and related outcomes.\textsuperscript{45-48} In addition, folate has recently been described to have a dual role in carcinogenesis. Initially, folate is believed to have a protective effect against changes in the normal colon mucosa, but in later stages of tumorigenesis, when neoplastic changes have already occurred, folate appears to promote tumor growth.\textsuperscript{18, 49} This hypothesis has been exemplified by the rise in CRC incidence after the introduction of mandatory food fortification in USA and Canada, which was implemented to prevent neural tube defects in developing fetuses.\textsuperscript{50}

Despite the important role of vitamin B12 as a methionine synthase cofactor (figure 1), and hence its potential role in carcinogenesis, this micronutrient has not been extensively investigated in epidemiological studies, and the role of vitamin B12 in CRC tumorigenesis is not clear.\textsuperscript{51-65}

Vitamin B6 is another coenzyme in one-carbon metabolism, which has been shown in several large studies to have a protective effect against CRC,\textsuperscript{63, 66} though evidence to the contrary also exists.\textsuperscript{67} Vitamin B6 will not be considered in this thesis.

Polymorphisms in genes encoding for enzymes in one-carbon metabolism have been shown to reduce enzyme activity and hence the pool of folate available for DNA methylation reactions.\textsuperscript{68} C677T and A1298C are two polymorphisms in the \textit{MTHFR} gene that have been extensively studied. In general, carriers of the C677T and/or A1298C alleles have a reduced risk of CRC.\textsuperscript{69-72}
Figure 1. DNA synthesis and DNA methylation reactions within a cell require one-carbon units, which is provided by folate. THF is reduced to 5, 10-methylTHF, which can be used either for DNA synthesis or be further reduced to 5-methylTHF, a reaction mediated by MTHFR. The enzyme methionine synthase and its cofactor vitamin B12 then mediate the remethylation of homocysteine to methionine. Methionine is further converted to SAM, which is used for the methylation of DNA. THF, tetrahydofolate; MTHFR, Methylene tetrahydrofolate reductase; SAM, S-adenosylmethionine; SAH, S-adenosylhomocysteine.
Two Pathways of Tumorigenesis

CRCs are believed to arise from premalignant lesions in the colon mucosa. During the development from barely visible aberrant changes in the colon mucosa to a benign tumor and later a malignant phenotype, genetic and epigenetic changes accumulate in the tumor cells (figure 2). APC is a tumor suppressor gene that is frequently mutated in both CRC and premalignant tumors and is therefore considered a key event in CRC development. Presence of a germline mutation in this gene is the underlying cause of Familial adenomatous polyposis (FAP), a condition characterized by the formation of numerous polyps in the colon that later develop into cancer (accounting for less than 1% of CRCs).

Alterations in several key genes are required for a cell to acquire malignant properties. However, due to rigid systems for primary prevention of DNA changes, as well as detection and removal of aberrant DNA, spontaneous mutation events are rare. The accumulation of mutations and subsequent malignant transformation of a cell are therefore believed to have underlying causes. Two major pathways of CRC tumorigenesis have been described; the microsatellite instability (MSI), and microsatellite stable (MSS) pathways (figure 2). Recently, a third pathway of tumorigenesis, the CpG Island Methylator Phenotype (CIMP), was described.

MSS

The MSS pathway is the underlying cause of the majority of CRCs and is often described in terms of chromosomal instability, which is an early event in MSS CRC tumorigenesis. The precise cause of chromosomal instability is not known, but it has been suggested to be a consequence of abnormalities in the mitotic checkpoint, centrosome number and function, telomere function, DNA damage response, or loss of heterozygosity (LOH, commonly found in chromosome 1, 5, 8, 17, and 18). A number of key events associated with the development of MSS CRC have been identified (figure 2), including mutations in tumor suppressor genes and oncogenes such as APC, TP53, KRAS, CTNNB1,
and \textit{PIK3CA}, and LOH in chromosome 18q (containing the tumor suppressor genes \textit{SMAD2}, \textit{SMAD4}, and \textit{DCC}).

It is generally believed that the majority of MSS tumors follow the pathway of tumorigenesis as explained above. However, it should be noted that MSS is not equivalent to chromosomal instability, since some tumors harbor only one of these two traits.

\textbf{Figure 2.} During the development from subtle changes in the colon mucosa to cancer, genetic and epigenetic changes accumulate in the tumor cells. MSS and MSI tumorigenesis are driven by chromosomal instability and microsatellite instability, respectively, and are prone to mutations in different genes. An alternative pathway of tumorigenesis, CIMP, is characterized by gene-specific hypermethylation, and is driven by epigenetic, rather than genetic events. CIMP, the CpG island methylator phenotype; CRC, colorectal cancer; MSI, microsatellite instability; MSS, microsatellite stable.
MSI
MSI is another pathway of tumorigenesis, accounting for approximately 15% of all CRC cases.\textsuperscript{79-80} MSI tumorigenesis is driven by the inactivation of mismatch repair genes, which are essential for the correction of occasional errors during DNA replication. Due to slippage of the replication machinery, repetitive DNA sequences (called microsatellites) are particularly prone to these errors, which cause the accumulation of frameshift mutations in mismatch repair deficient cells. Some of the identified key genes in MSI CRC that contain such sequences include $TGF\beta R II$, $BAX$, and $IGFIIR$.\textsuperscript{80} Sporadic MSI tumors are commonly caused by promoter hypermethylation of the mismatch repair gene $MLH1$ resulting in the inactivation of this gene.\textsuperscript{81} The familial form of MSI CRC is hereditary nonpolyposis CRC (HNPCC, or Lynch syndrome),\textsuperscript{82} which is caused by germline mutations in the mismatch repair genes $MLH1$, $PMS2$, $MSH6$, or $MSH2$, and accounts for about 3% of all CRC cases.\textsuperscript{5,83}

Compared to MSS, MSI tumors are more often located in the proximal colon, poorly differentiated, and of a mucinous, or signet ring, histological type.\textsuperscript{79, 84-85} Another common finding in MSI CRC is the presence of tumor-infiltrating T cells.\textsuperscript{86-87} MSI tumors have often been associated with a better patient prognosis compared with MSS.\textsuperscript{88-89} In 1999 an alternative pathway of tumorigenesis, the CpG Island Methylator Phenotype (CIMP), was described by Toyota et al.\textsuperscript{78} This phenotype is discussed below.
DNA Methylation and the CpG Island Methylator Phenotype – An Alternative Pathway of Tumorigenesis

DNA Methylation

In normal cells, DNA methylation has important functions in the maintenance of genomic stability, tissue-specific gene regulation, and imprinting (inactivation of one parental allele). Methylation occurs on cytosines preceding guanines (CpGs) which are found either scattered as single dinucleotides throughout the genome, or clustered in promoter regions (so-called CpG islands, figure 3), where they silence gene expression when methylated.\textsuperscript{90-91} More than half of human genes have been estimated to be regulated in this manner, by promoters containing CpG islands.\textsuperscript{92-93} In cancer, CpG islands are often found aberrantly hypermethylated, causing inappropriate silencing of gene expression.\textsuperscript{40, 94} In contrast, the scattered CpG dinucleotides, which are methylated in normal cells, have been found unmethylated in cancer. Genomic hypomethylation is believed to contribute to tumorigenesis by creating prerequisites for genomic instability.\textsuperscript{38, 40}

The CpG Island Methylator Phenotype - CIMP

In 1999, Toyota \textit{et al.} described a subgroup of CRCs that showed frequent promoter hypermethylation, which they called CIMP (short for the CpG Island Methylator Phenotype).\textsuperscript{78} The aberrantly hypermethylated, and thus silenced, genes were found to be unmethylated in normal colorectal mucosa and in CRC without this phenotype. The discovery of CIMP led to the proposal of a tumorigenic pathway of CRC driven by promoter hypermethylation and hence epigenetic, rather than genetic, inactivation of tumor suppressor genes.\textsuperscript{95} Several genes that are hypermethylated in CIMP have important functions in the cell, (e.g. \textit{CDKN2A}, the gene coding for the tumor suppressor p16), whereas others have unknown functions.
Figure 3. In normal cells, single CpG sites scattered throughout the genome are generally methylated, whereas CpG islands within promoter regions are unmethylated, which allows for active gene transcription (top). In cancer cells, hypermethylation of CpG islands within promoter regions cause silencing of gene expression, and hypomethylation of single CpG sites is believed to contribute to genomic instability (bottom). Adapted from Herman et al. 2003 NEJM.

The exact definition of CIMP has not been uniform among studies. To be able to discuss CIMP status across individual study definitions in this thesis, extensively hypermethylated tumors are referred to as CIMP+, and non-CIMP+ tumors are referred to as CIMP-. In studies where an intermediately methylated subgroup has been defined, this is referred to as CIMP-intermediate. The terms CIMP-negative, CIMP-low, and CIMP-high are used only for tumors defined according to the specific CIMP definition used in papers II-IV, which is based on a panel of eight genes, and according to which CIMP-negative, 0 genes methylated; CIMP-low, 1-5 genes methylated; and CIMP-high, 6-8 genes methylated.

Several large studies have reported a strong association between CIMP+ and proximal tumor location, MSI, and BRAF mutation.\textsuperscript{78, 96-103} Not as frequently reported, and not as strong, are observed associations between CIMP+ and female sex, higher age, higher tumor stage and grade, a mucinous histological type, wild type TP53, KRAS mutations, and the presence of tumor-infiltrating lymphocytes.\textsuperscript{97, 99-106} Although not as widely studied, the CIMP-intermediate subgroup has been associated with male sex and KRAS mutation.\textsuperscript{107-109}
Figure 4. Although colorectal adenomas can be large in size, they are limited to the mucosa (a, the innermost layer of the colon). Stage I tumors are limited to the submucosa (b), or have invaded, but not grown completely through, the circular and longitudinal muscular layers of the bowel wall (c). Stage II tumors are characterized by growth into the subserosa (d). Stage III are tumors with spread to lymph nodes, regardless of depth of bowel wall infiltration, and stage IV (advanced) tumors have metastasized to distant organs.

Treatment and Prognosis in Colorectal Cancer

The CRC Staging System
CRC is staged I-IV according to the depth of bowel wall infiltration and spreads to lymph nodes and distant organs (figure 4). In the earliest stage, stage I, the tumor has not invaded through the thick muscular layer of the bowel wall. 5-year survival in this subgroup is 90-100%. Risk of recurrence and cancer-related death increases as the tumor extends through muscularis propria (stage II; 5-year survival 75-85%) and spread to lymph nodes (stage III; 5-year survival 45-60%) and distant organs (stage IV). Until recently, patients with metastatic disease rarely survived more than 2 years. However, surgical resection of liver metastases has become a successful treatment option, contributing greatly to the improved survival in this patient group.
**CRC Treatment**
The primary treatment of CRC is surgery. Additional treatment with adjuvant chemotherapy is sometimes administered, depending mainly on tumor stage. Patients with stage I tumors have an excellent prognosis and are therefore generally treated with surgery alone. Stage II patients are considered for adjuvant chemotherapy only if at an increased risk of recurrence and cancer-related death, based on criteria described below. Stage III tumor patients, having a high rate of recurrence and cancer-related death, are generally considered for adjuvant chemotherapy. In addition, rectal cancers are often considered for preoperative radiotherapy. Advanced (stage IV) CRC patients are treated with surgical resection of the primary tumor and, when possible, liver metastases, as well as palliative chemotherapy. Many stage IV patients eventually succumb to their disease, which reflects the limited effect of chemotherapy in advanced cases.

Whereas a survival benefit of adjuvant chemotherapy has been shown for stage III CRC patients, results for patients with stage II disease have not been conclusive. However, a recent large meta-study including almost 3 000 stage II CRC patients presented evidence for an effect of adjuvant chemotherapy in this subgroup. A great interest exists in finding markers that identify stage II tumor patients with a high risk of recurrence and cancer-related death, who would potentially benefit from adjuvant chemotherapy. Another challenge in the development of CRC treatment is the identification of stage III patients with a good prognosis, who would benefit little from, and therefore do better without, adjuvant chemotherapy.

**Prognostic Markers**
Stage is the most important factor when predicting CRC patient prognosis. It is also the main factor considered in treatment decision protocols. In addition to stage, positive resection margins, lymphovascular invasion, and poor cellular differentiation are among the established risk factors for a poor prognosis. Obstruction and
perforation at presentation,\textsuperscript{117} performance status,\textsuperscript{118} and preoperative carcinoembryonic antigen levels\textsuperscript{119} reflect tumor burden, and are also indicators of a poor prognosis.

Although tremendous effort has been put into identifying novel prognostic markers, few have been shown to provide additional information beyond that of tumor stage, and not even the most promising have been implemented in clinical practice.\textsuperscript{120-121} Some of the more robust and thoroughly investigated prognostic markers include LOH in 18q (including the gene \(DCC\)),\textsuperscript{122-123} chromosomal instability,\textsuperscript{124} and MSI.\textsuperscript{88-89} Other extensively investigated markers include the expression of thymidylate synthase,\textsuperscript{125} proliferation markers, such as Ki-67,\textsuperscript{126} and mutations in genes with tumor suppressor or oncogene function, such as TP53,\textsuperscript{127-129} KRAS,\textsuperscript{130-135} and BRAF.\textsuperscript{135-137} The presence of tumor-infiltrating inflammatory cells has also been shown to be important for a favorable patient prognosis.\textsuperscript{138-147} Modern technology has allowed for new approaches, such as multigene expression signatures\textsuperscript{148-149} which have recently generated promising results but await further validation.

**Treatment-Specific Predictive Markers**

The standard adjuvant treatment for non-metastatic CRC is based on 5-fluorouracil and oxaliplatin. Increased expression of thymidylate synthase, one of the main targets for 5-fluorouracil, as well as certain polymorphisms in the gene coding for this enzyme, have been shown to modulate the effect of treatment.\textsuperscript{125, 150-154} \(MTHFR\) polymorphisms confer a reduced folate pool and have been implicated in 5-fluorouracil function and as modulators of its effect.\textsuperscript{155-158} Polymorphisms in the gene for dihydropyrimidinid dehydrogenase, an enzyme involved in the metabolism of 5-fluorouracil, have also been related to therapeutic response.\textsuperscript{159-160}
Although studies have not been altogether conclusive, CRC patients with MSI tumors are believed to have only a minor effect of 5-fluorouracil-based therapies.\textsuperscript{88, 161-165}

Treatment targeting the epidermal growth factor receptor (EGFR) is sometimes administered to patients with metastatic disease. KRAS and BRAF are two of the downstream mediators of EGFR signaling, and mutations in these genes confer a lack of response to anti-EGFR treatment.\textsuperscript{166-172} Testing for \textit{KRAS} mutations before the initiation of such treatments has been implemented in clinical practice.\textsuperscript{173}

\textbf{DNA Methylation, CIMP, and CRC Patient Prognosis}

In recent years much attention has been focused on DNA hypermethylation of specific genes and its ability to predict CRC patient outcome. One of the most studied genes is \textit{CDKN2A}.\textsuperscript{174-175} However, a recent meta-analysis concluded that the hypermethylation of \textit{CDKN2A} is not an independent prognostic factor in CRC.\textsuperscript{175} Results have been reported for an abundance of other, less frequently studied, genes, but for none has a prognostic role been established.

The prognostic importance of concurrent hypermethylation of several genes, as in CIMP, has also been studied.\textsuperscript{97, 102-103, 136-137, 176-188} Although the results are not conclusive, many studies have found a poor prognosis in MSS CIMP+ CRC patients.\textsuperscript{97, 102-103, 136, 178, 180, 182} The importance of a CIMP-intermediate subgroup in CRC patient survival is unclear.\textsuperscript{102, 137, 178, 183, 187, 189-190} The interaction between CIMP status and response to chemotherapy have been investigated, but the results have been contradictory.\textsuperscript{178-179, 186}
Aims of This Thesis

General Aims
CRC is increasingly being recognized as a heterogeneous set of diseases that varies in clinicopathological and molecular characteristics, and that has different precursor lesions and initiating events. Risk factors should be investigated in terms of CRC subtypes based on clinicopathological and molecular features, which has seldom been done in the past. There is also a need for additional prognostic markers, in order to improve the prediction of patient prognosis.

The aim of this thesis was to investigate CIMP in CRC in a northern Swedish population, with a focus on the role of components of one-carbon metabolism as risk factors for CRC and subgroups of CRC based on CIMP status, and the role of CIMP and related factors in determining patient prognosis.

Specific Aims

- To investigate how prediagnostic plasma vitamin B12 concentrations relate to the risk of developing CRC.

- To relate factors of one-carbon metabolism to the risk of developing CIMP-negative, CIMP-low and CIMP-high CRC.

- To investigate the association between CIMP status and CRC patient prognosis, taking into consideration MSI screening status.

- To investigate the relation between tumor-infiltrating T cells and CRC patient prognosis, taking into consideration CIMP and MSI screening status.
MATERIAL AND METHODS

Study Subjects

The studies described in this thesis are based on two study groups, the Northern Sweden Health and Disease Study (NSHDS) and the Colorectal Cancer in Umeå Study (CRUMS). The clinico-pathological characteristics of the CRCs cases included in NSHDS and CRUMS are described in table 1.

The Northern Sweden Health and Disease Study (NSHDS)

The first study group includes CRC cases and matched referents identified among subjects recruited to the NSHDS cohort, which is a prospective study comprising three cohorts, namely the Northern Sweden Multinational Monitoring of Trends and Determinants in Cardiovascular Disease (MONICA) Project, the Västerbotten Intervention Project (VIP), and the local Mammography Screening Project (MSP). The recruited subjects were men and women aged 25-74 years in MONICA, aged 30 (only 1985-1996), 40, 50, and 60 years in VIP, and women aged approximately 50-70 years in MSP. The subjects of MONICA and VIP were invited to a free health examination and were asked to complete an extensive lifestyle questionnaire and to donate a blood sample for future research. In MSP, women were asked, at the time of their mammography screening, to complete a questionnaire concerning reproductive history and to donate a blood sample.

CRC cases were identified from the National Cancer registry (ICD-10 18.0, 18.2-9, 19.9, and 20.9) in 2002. For each case in NSHDS, two referents were matched by sex, age (±1 year), subcohort, date of health survey (± 6 months, generally within one to two months), and fasting status at blood sample donation. Exclusion criteria included a previous CRC diagnosis, insufficient plasma sample available, no matching referent, and prioritization to other studies.
The nested case-referent study approach is useful for the investigation of risk factors for a disease, as in papers I and II. Cases and references are recruited before diagnosis, and hence any problems with recall bias and reverse causation are minimized. Also, by matching referents from the same cohort to each case (hence the term nested), impractical and expensive investigation of all referent subjects within a large cohort is avoided.

**The Colorectal Cancer in Umeå Study (CRUMS)**

The second study group, CRUMS, comprises patients who underwent primary CRC resection at Umeå University Hospital (Umeå, Sweden), during the period 1995 to 2003, and for whom archival tumor tissue and clinico-pathological data were available.

Patients were recruited to CRUMS after their CRC diagnosis, and the data were thus collected retrospectively. This study design is useful in the investigation of patient and tumor characteristics in relation to disease outcome, as in papers III and IV.

**Patient and Tumor Characteristics**

A pathologist verified all diagnoses in NSHDS and CRUMS. Pathology reports and patients records were used to extract tumor characteristics, patient data, and cancer-specific survival. Vital status was obtained from the Swedish population registry, and patients in NSHDS and CRUMS were followed up until 2008 and 2005, respectively.
Table 1. Clinico-pathological characteristics of colorectal cancer cases in the Northern Sweden Health and Disease Study (NSHDS) and the Colorectal Cancer in Umeå Study (CRUMS).

<table>
<thead>
<tr>
<th></th>
<th>NSHDS</th>
<th>CRUMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases, n</td>
<td>226</td>
<td>490</td>
</tr>
<tr>
<td>Referents, n</td>
<td>437</td>
<td>-</td>
</tr>
<tr>
<td>Study design</td>
<td>Nested case-referent (prospective)</td>
<td>Retrospective</td>
</tr>
<tr>
<td>Studied in papers</td>
<td>I, II, III</td>
<td>III, IV</td>
</tr>
<tr>
<td>Age at diagnosis, median years (range)</td>
<td>63 (35-74)</td>
<td>71 (26-96)</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>94 (41.6)</td>
<td>270 (55.1)</td>
</tr>
<tr>
<td>Women</td>
<td>132 (58.4)</td>
<td>220 (44.9)</td>
</tr>
<tr>
<td>Tumor site, n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right-sided colon</td>
<td>72 (31.9)</td>
<td>154 (31.8)</td>
</tr>
<tr>
<td>Left-sided colon</td>
<td>70 (31.0)</td>
<td>148 (30.5)</td>
</tr>
<tr>
<td>Rectum</td>
<td>84 (37.2)</td>
<td>183 (37.7)</td>
</tr>
<tr>
<td>Tumor stage, n* (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>36 (18.4)</td>
<td>75 (15.7)</td>
</tr>
<tr>
<td>II</td>
<td>69 (35.2)</td>
<td>188 (39.2)</td>
</tr>
<tr>
<td>III</td>
<td>46 (23.5)</td>
<td>102 (21.3)</td>
</tr>
<tr>
<td>IV</td>
<td>45 (23.0)</td>
<td>114 (23.8)</td>
</tr>
<tr>
<td>MSI screening status, n* (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSS</td>
<td>175 (87.5)</td>
<td>402 (84.5)</td>
</tr>
<tr>
<td>MSI</td>
<td>25 (12.5)</td>
<td>74 (15.5)</td>
</tr>
<tr>
<td>CIMP status, n* (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CIMP-negative</td>
<td>95 (50.0)</td>
<td>247 (50.4)</td>
</tr>
<tr>
<td>CIMP-low</td>
<td>68 (35.8)</td>
<td>183 (37.3)</td>
</tr>
<tr>
<td>CIMP-high</td>
<td>27 (14.2)</td>
<td>60 (12.2)</td>
</tr>
<tr>
<td>BRAF V600E, n* (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wild type</td>
<td>161 (82.1)</td>
<td>418 (86.0)</td>
</tr>
<tr>
<td>Mutated</td>
<td>35 (17.9)</td>
<td>68 (14.0)</td>
</tr>
<tr>
<td>Degree of infiltration of tumor by T cells, n* (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>-</td>
<td>147 (32.3)</td>
</tr>
<tr>
<td>Intermediate</td>
<td>-</td>
<td>163 (35.8)</td>
</tr>
<tr>
<td>High</td>
<td>-</td>
<td>145 (31.9)</td>
</tr>
</tbody>
</table>

*The following number of missing cases were present in NSHDS/CRUMS: tumor site, 0/5; tumor stage, 30/11; MSI, 26/14; CIMP, 36/0; BRAF mutation, 30/4; total CD3 score (degree of infiltration by T cells), - /35.
Plasma Analyses

As a part of the study protocol, a blood sample was collected from NSHDS subjects at recruitment as previously described in detail. To account for any differences in analyses dependent on fasting status or handling procedures in MONICA, VIP, and MSP, all referent samples were matched to their index case for fasting status and subcohort.

The concentrations of folate, vitamin B12, and total homocysteine in plasma were measured at the Department of Medical Biosciences, Clinical Chemistry, Umeå University (Umeå, Sweden). For folate and vitamin B12, the Quantaphase II B12/Folate Radioassay (Bio-Rad, Hercules, CA, USA) was used according to the manufacturer’s instructions. In brief, the plasma sample was mixed with radio-labeled folate and vitamin B12 and added to immobilized folate and vitamin B12-binding protein. Due to competitive binding of labeled and unlabeled vitamins, the radioactivity of bound vitamins is inversely proportional to the amount of folate and vitamin B12 in the original plasma sample.

Total plasma homocysteine was measured using a fluorescence polarization immunoassay in an IMx unit (Abbott, Chicago, IL, USA) according to the manufacturer’s instructions. In brief, all forms of homocysteine in the plasma sample were reduced and converted to S-adenosyl-L-homocysteine. The sample was then mixed with a fluorescein tracer and added to immobilized antibodies that bind S-adenosyl-L-homocysteine as well as the fluorescein tracer. The amount of bound fluorescein tracer is measured in the IMx unit, and is inversely proportional to the amount of homocysteine in the plasma sample.
DNA Methylation Analyses

Since the nucleotide sequence is not affected by methylation of the CpGs within it, regular PCR based methods cannot be used to detect DNA methylation. However, if DNA is treated with bisulfite, unmethylated cytosines are converted to uracil, while methylated cytosines remain unchanged (figure 5). Hence, after bisulfite modification, PCR can be used to discriminate between methylated and unmethylated DNA. We used a quantitative real-time PCR-based method called MethyLight, which was first described by Peter W. Laird and colleagues.194

CIMP status is determined based on the promoter methylation status of several genes. We used a well-established and thoroughly evaluated eight-gene CIMP panel including the genes CDKN2A, MLH1, CACNA1G, NEUROG1, RUNX3, SOCS1, IGF2 and CRABP1. These genes are methylated in CIMP+ CRC, but not in normal colorectal tissue or CIMP- CRC.101, 195 In addition, this panel identifies CIMP-low, which is a group of tumors with an intermediate number of genes methylated and with characteristics different from both CIMP-negative and CIMP-high CRC.98, 108-109

**MethyLight**

Tumor DNA was extracted from formalin-fixed, paraffin-embedded CRC tissue. When necessary and possible, the proportion of tumor cells was maximized by macrodissection. The EZ DNA methylation kit (Zymo Research, Orange, California, USA) was used for bisulfite conversion of DNA.

A detailed MethyLight protocol has been published by Weisenberger et al.195-196 For each gene in the CIMP panel, primers covering CpG sites of the promoter are designed to anneal only to methylated and bisulfite-treated DNA. In addition to methylation-specific primers, a probe is designed to hybridize to a sequence within the amplicon that also contains methylated CpG sites. The use of both primers and probe
minimizes the detection of partly methylated DNA that may not have an impact on gene expression. All primers and probes were according to Weisenberger et al.\textsuperscript{195}

To normalize for the amount of bisulfite-treated DNA present in the reactions, a reference reaction, \textit{ALU}, was run for all tumor samples. \textit{Alu} sequences are short, interspersed, repeated sequences that are present in over a million copies in the human genome.\textsuperscript{198} Moreover, the \textit{ALU} reaction is designed from an \textit{Alu} sequence depleted of CpG dinucleotides by evolutionary deamination, and is hence methylation independent.\textsuperscript{199} To normalize results for the efficacy of PCR amplification, a reaction with presumably fully methylated reference DNA (human genomic DNA treated with CpG Methyltransferase M.SssI) was run for each gene. A standard curve was generated with \textit{ALU} primers and probes in reactions with serial dilutions of the fully methylated reference DNA.

For each sample and gene, the percent of methylated reference (PMR) value was calculated by the following equation: (quantity of the gene-specific reaction of the sample/quantity of the \textit{ALU} reaction of the sample)/(mean quantity of the gene-specific reaction for the methylated reference sample/mean quantity of the \textit{ALU} reaction for the methylated reference sample).\textsuperscript{195} For each sample, a gene was considered hypermethylated if an exponential amplification plot was obtained, and the PMR was $>10$ (figure 5). A sample was considered uninformative if the PCR threshold cycle for \textit{ALU} was $>22$.\textsuperscript{195,196} A tumor was defined as CIMP-negative if no genes were hypermethylated, CIMP-low if 1-5 genes were hypermethylated, and CIMP-high if 6-8 genes were hypermethylated.\textsuperscript{101}
Figure 5. MethyLight work flow and amplification plots generated from a CIMP-high sample (left, eight genes and control reaction positive) and CIMP-negative sample (right, only control reaction positive). CIMP, the CpG island methylator phenotype
Genetic analyses

Taqman allelic discrimination method was used to determine the MTHFR C677T and A1298C genotypes and BRAF V600E mutation status as previously described.\textsuperscript{44, 200} MTHFR genotypes for CRC cases and referents in NSDHS were determined in DNA extracted from buffy coat. BRAF V600E mutation analysis was performed on DNA extracted from formalin-fixed paraffin-embedded tumor tissue from CRC cases in both NSHDS and CRUMS. The genetic analyses were carried out at the Department of Medical Biosciences, Center for Genome Research, Umeå University (MTHFR genotyping), and the Department of Clinical Genetics, Umeå University Hospital (BRAF mutation).

<table>
<thead>
<tr>
<th>Table 2. Antibodies used for immunohistochemical stainings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protein</strong></td>
</tr>
<tr>
<td>MLH1</td>
</tr>
<tr>
<td>PMS2</td>
</tr>
<tr>
<td>MSH6</td>
</tr>
<tr>
<td>MSH2</td>
</tr>
<tr>
<td>CD3</td>
</tr>
<tr>
<td>p53 (NSHDS)</td>
</tr>
<tr>
<td>p53 (CRUMS)</td>
</tr>
</tbody>
</table>

\* in microwave oven
\dagger in 2100 retriever pressure cooker (Biocare Medical)
Immunohistochemistry

Immunohistochemistry (IHC) is a method commonly used in both clinical and research settings for the detection of proteins in tissue sections. This method is based on the use of antibodies that bind specifically to the protein of interest and are visualized by enzymatic conversion of a chromogen. After counterstaining with hematoxylin for visualization of cell nuclei that are not stained with the protein-specific antibodies, the sections can be examined under light microscope.

For IHC stainings, we used formalin-fixed, paraffin-embedded CRC tissue specimens obtained in accordance with routine clinical protocols after primary tumor surgery. Deparaffinized, rehydrated 4 µm sections were pretreated and stained with primary antibodies according to table 2. The iVIEW DAB Detection kit (Ventana Inc, Tuscon, AZ, USA) or MACH3 detection kit (Biocare Medical, Concord, CA, USA) was used for visualization. Both systems are based on the enzymatic conversion of 3,3-diaminobenzidine (DAB) by horseradish peroxidase, resulting in a brown color that indicates the presence and location of the protein of interest. Non-malignant lymph nodes were used as a positive control for CD3 staining, and non-malignant stromal cells were used as internal positive controls for MLH1, MSH2, MSH6, PMS2, and p53. Sections lacking positivity in control cells were considered uninformative. Overexpression of p53 was defined as present when $\geq 25\%$ of tumor cells were positive for nuclear staining of p53. Overexpression of p53 detected by IHC has been closely associated with the presence of mutations in the corresponding gene.\(^{201}\)

MSI Screening Status

MSI CRC is caused by inherited or acquired inactivation of mismatch repair genes. In order to determine MSI screening status, we used IHC to detect mismatch repair protein loss. A positive MSI screening status was defined a lack of nuclear staining for MLH1, MSH2, MSH6, or PMS2, and is referred to in this thesis as MSI. Cases with a negative MSI screening status are referred to as MSS. The gold standard to determine
MSI status of a tumor is by PCR amplification of a panel of five microsatellite markers (defined by the Bethesda guidelines in 1997), according to which tumors with MSI in ≥2 of the markers are considered MSI-high. The use of IHC has been shown to accurately identify MSI-high CRC, whereas it does not detect the presence of MSI-low. The sensitivity and specificity for detecting MSI by IHC, as described above, compared to gold standard MSI testing, is 92% and 99.8%, respectively.

Quantification of Tumor-Infiltrating T Cells
To determine the density of tumor-infiltrating T cells, we used IHC staining of tumor tissue from patients in CRUMS. The antibody used, anti-CD3, recognizes the epsilon chain of the CD3 molecule, which is a part of the T cell receptor complex present on all T cells.

When evaluating the infiltration of T cells in the IHC-stained tissue slides, the following semi-quantitative scale was used: 1, no, or sporadic CD3 positive (CD3+) cells; 2, moderate numbers of CD3+ cells; 3, abundant occurrence of CD3+ cells; and 4, highly abundant occurrence of CD3+ cells. A separate score was assigned for the tumor front (CD3+ cells in stroma, adjacent to the invasive tumor margin), the tumor center (CD3+ cells in stroma, within the tumor mass), and the intraepithelial compartment (CD3+ cells within tumor cell nests). The total CD3 score was calculated as the sum of the CD3 scores. Patients were then divided into three equally sized groups with low (3-4), intermediate (5-6), or high (7-12) total CD3 scores. Other research groups have used a similar approach.
Statistics

For comparison of clinico-pathological characteristics in subgroups, Kruskal-Wallis tests were used for continuous variables and $\chi^2$-tests or Fisher’s exact tests (when observed or expected frequencies were less than five) were used for categorical variables. Correlations between distributions of CD3 scores were determined by non-parametric Spearman’s Rho.

Odds ratios (OR) for disease and 95% confidence intervals (CI) were calculated using conditional logistic regression. Kaplan-Meier plots were constructed for survival analyses, and survival differences between groups were tested by log rank tests. Multivariate Cox proportional hazard models were employed in order to take multiple clinico-pathological factors into consideration in the survival analyses. Cancer-specific events were defined as death with known disseminated or recurrent disease. Cases were censored at the end of follow-up or at time of death by other causes. CRUMS patients who died with postoperative complications within one month after surgery were excluded from survival analyses. Data on deaths due to postoperative complications were not available for NSHDS cases, but only a few cases died within one month after surgery.

For all statistical analyses, SPSS (version 14.0, 15.0, or 17.0) or PASW Statistics 18.0 (SPSS Inc., Chicago, Illinois, USA) were used.

Ethical Approval

The use and handling of tissue samples and patient data in papers I-IV have been approved by the ethical committee, Umeå University, and the Swedish National Computer Data Inspection Board.
RESULTS AND DISCUSSION

One-Carbon Metabolism in CRC Tumorigenesis

An imbalance in the pathways of one-carbon metabolism may result in hypo- or hypermethylation of DNA and/or a reduced level of nucleotide synthesis, which in turn can cause aberrant gene expression and genomic instability.36-40 These events are identified as early and crucial in CRC tumorigenesis.40, 209-211 Folate, vitamin B12, homocysteine, and the MTHFR enzyme are essential components of one-carbon metabolism, and may, therefore, be implicated in CRC tumorigenesis.

To investigate these factors in the context of CRC risk, we used the 226 CRC cases and 436 matched referents included in NSHDS. Blood samples from all study subjects were collected prospectively. A number of risk modulating factors, including body mass index, physical activity, smoking, and alcohol intake, could be considered in the multivariate analysis. However, information on family history of CRC and the intake of NSAID, were lacking and could not be taken into account.

Vitamin B12

A 70% reduction in rectal cancer risk was found in individuals with the highest compared to lowest plasma concentrations of vitamin B12 (adjusted OR, 0.30 [95% CI, 0.08-0.99]; paper I). For colon cancer, and in particular for cancer located in the left colon, the trend was in the opposite direction.

Few studies have examined the role of vitamin B12 in CRC risk, and even fewer have presented data for rectal cancer separately.52, 54-63, 65 Although most studies have found no association between dietary intake or circulating levels of vitamin B12 and rectal cancer52, 57, 59, 63 or CRC55, 60-62, 65 risk, one other study have reported an inverse association.56 Notably, some studies have reported opposing results that indicate an increased risk for rectal cancer58 or CRC54 with increasing dietary intake of vitamin
B12. A possible explanation for the contradictive findings may be the confounding effect by intake of red meat, which is a major source of vitamin B12. Intake of meat in general, and red meat in particular, is an established risk factor for CRC, which may depend on its high content of animal fat, or on carcinogens that arise during some cooking procedures. Red meat is also rich in iron, which is another suggested risk factor for CRC.

The findings in paper I support a protective effect of higher levels of circulating vitamin B12 against the development of rectal, but not colon, cancer. Differences in the etiology of CRC depending on localization, have previously been suggested, with a possible explanation in the different distribution of molecular subtypes in the colon and rectum.

Few rectal cancers are of MSI and/or CIMP+ type. In sporadic CRC, both MSI and CIMP+ arise in the background of DNA promoter hypermethylation (of MLH1 and multiple genes, respectively). Speculatively, a protective effect of vitamin B12 apparent in the rectum might be diminished in the colon due to the contributory role of vitamin B12 to DNA methylation and, thereby, to the development of MSI and CIMP+ cancer. Paper II was thus initiated, in order to investigate whether the discrepancy in results for vitamin B12 in colon and rectal cancer might depend on the different distributions of CIMP and MSI screening status at these localizations. A related aim of paper II was to determine whether previous findings from the same cohort of a reduced risk of CRC in individuals with very low levels of folate, a one-carbon donor, might be due to a lower risk of DNA hypermethylation, and hence MSI and CIMP+ CRC, in these subjects.

Components of One-Carbon Metabolism and CIMP

In paper II, CIMP and MSI screening status were assessed in the CRC cases in NSHDS. The risk of subtypes of CRC with different CIMP and MSI screening status were then determined for folate, vitamin B12, homocysteine, and the MTHFR C677T and A1298C polymorphisms.
The risk of developing CRC with hypermethylation in one or more genes (i.e. CIMP-low or CIMP-high) was increased in subjects with higher plasma concentrations of folate. ORs for quintiles two to five were all above one, and the adjusted OR (taking into account body mass index, current smoking, recreational physical activity, and alcohol intake) for quintile two to five versus one, was 2.96 (95% CI, 1.24–7.08). Although the risk for CIMP-negative CRC also tended to increase with increasing levels of plasma folate (quintile 2-4), the results were not statistically significant, and an increased risk was not seen for the highest folate levels (quintile 5). The findings for an adequate “methyl availability” status (defined as plasma concentrations ≥6.8 nmol/l for folate, and ≥148 pmol/l for vitamin B12, and <15 µmol/l for total homocysteine) suggested a possible reduced risk of CIMP-negative CRC (adjusted OR, 0.61 [95% CI, 0.32-1.15]), whereas the OR was in the opposite direction for CIMP-low/high CRC. Plasma concentrations of neither vitamin B12 nor homocysteine were associated with the risk of CIMP-low/high or CIMP-negative CRC.

Several studies have presented results in line with our findings, including associations between higher dietary intake and circulating levels of folate and vitamin B12 and gene-specific hypermethylation in CRC as well as in normal colon mucosa. In addition, animal studies have shown an increase in DNA methylation events in the colon mucosa of mice supplemented with folate. However, null findings, and results that contradict our findings have also been reported. Studies of “methyl availability” based on dietary components and the MTHFR polymorphisms have found no associations, or an increased risk of CIMP+ CRC in subjects with low “methyl availability”.

When considering the results from papers I and II, it should be noted that due to the dietary pattern in northern Sweden, and the absence of mandatory folic acid fortification of foods in this country, the plasma concentrations of folate in the study population is low, whereas vitamin B12 deficiency is uncommon. Hence, the studies in this thesis
represent the investigation of circulating levels of folate and vitamin B12 at the lower and higher ends of the physiological range, respectively.

The MTHFR C677T and A1298C polymorphisms, which result in a lower enzyme activity\(^6\) and hence a lower pool of folate available for DNA methylation, have been associated with lower CRC risk.\(^59\)\(^-\)\(^72\) The inclusion of these polymorphisms in paper II allowed the investigation of whether the reduced risk is confined to CRCs with or without hypermethylation.

The MTHFR C677T polymorphism was associated with a reduced risk of both CIMP-negative and CIMP-low/CIMP-high CRC. The MTHFR A1298C polymorphism was associated with an increased risk, which was confined to CIMP-negative CRC. In line with these findings, some studies have reported lower levels of methylation in tumors and normal mucosa of subjects with the MTHFR C677T polymorphism.\(^230\)-\(^232\) Others have shown the inverse association and an increased risk of tumors with hypermethylation in MTHFR C677T and A1298C carriers,\(^177\),\(^218\),\(^233\)-\(^235\) or no association.\(^236\)-\(^238\) The combination of a less active variant of the MTHFR enzyme and a low folate intake have also been associated with hypermethylation.\(^227\)

**Components of One-Carbon Metabolism and MSI**

In paper II, CRC cases with MSI tumors had higher prediagnostic levels of plasma folate compared with MSS tumor patients. MSI in sporadic CRC is caused by MLH1 hypermethylation, which may be influenced by higher methyl availability. However, hypermethylation of MLH1 was not associated with higher plasma folate concentrations in paper II. Other studies have reported opposite findings for MSI versus MSS.\(^239\)-\(^240\) Schernhammer et al. found an inverse association between dietary folate intake and the risk of both MSI and MSS CRC in a prospective setting.\(^241\) Associations between the MTHFR C677T and A1298C polymorphisms and MSI and/or MLH1 hypermethylation have been reported in several studies.\(^236\),\(^239\),\(^242\)-\(^243\)
Characterization of CIMP in a Northern Swedish Population

To study clinico-pathological characteristics and patient outcome in CIMP-negative, CIMP-low, and CIMP-high CRC, we included a second set of study subjects, namely 414 patients from CRUMS. The frequencies of CIMP-negative, CIMP-low, and CIMP-high tumors were similar in NSHDS (50%, 36%, and 14%, respectively) and CRUMS (51%, 37%, and 12%, respectively). Similar frequencies were also found in an American population with the same methodology and CIMP gene-panel (15% CIMP-high and 38% CIMP-low), whereas lower frequencies of CIMP-high (7.5%), and higher frequencies of CIMP-low (74%), were found in a Korean population.

Clinico-Pathological Characteristics of CIMP Subtypes

The previously reported associations between CIMP+ and a proximal tumor site, mucinous histological type, MSI, BR-4F-mutation and normal p53 screening status were confirmed in CIMP-high cases in both NSHDS and CRUMS (paper III). Other features previously associated with CIMP+, such as higher age at diagnosis, female sex, poor differentiation, and higher tumor stage were consistent, but not statistically significant in NSHDS and CRUMS. Several of the clinico-pathological characteristics associated with CIMP-high are also related to MSI. However, the associations between CIMP-high and a right-sided location, mucinous histological type, and BR-4F mutation were confirmed in analyses of MSI and MSS tumors separately, indicating that these traits are associated to CIMP-high independently of MSI screening status.

The recruitment protocols were different in the two studies, resulting in a younger age at diagnosis and a higher percentage of female cases in NSHDS compared with CRUMS. Also noteworthy, 12 of 13 MSI CIMP-high tumor patients in NSHDS were women, whereas a skewed sex distribution in this subgroup was not evident in CRUMS.
The Definition of CIMP

The definition of CIMP-negative, CIMP-low, and CIMP-high in this thesis is based on a well-established, eight-gene panel proposed by Ogino et al.\textsuperscript{101} Other commonly used gene panels include the MINT (methylated in tumor) panel used by Toyota et al. in the original paper in which they first described CIMP+,\textsuperscript{78} and the five-gene panel developed by Weisenberger et al. for the precise detection of CIMP+.\textsuperscript{195}

The many different methods to detect hypermethylation and panels to define CIMP+, as well as the inconsistent use of a CIMP-intermediate subgroup, make comparisons between studies difficult. Although much research has been done to find the most appropriate genes to define CIMP+ and CIMP-intermediate (which may be an altogether different set of genes and not simply a lesser number of methylated genes), a consensus panel for CIMP+ has not yet been established.
CIMP and CRC patient survival

Although the proportion of CRC patients who survive their illness has increased with advances in the fields of surgery and oncologic therapy, CRC mortality is still high.\textsuperscript{1,3,244} Patient prognosis is estimated in clinical practice based mainly on tumor stage. Although much effort has been put into identifying novel prognostic biomarkers, few have been implemented in clinical practice. In paper III, the impact of CIMP status on patient outcome was assessed in CRC cases in NSHDS and CRUMS.

CIMP-high and Patient Survival

Patients with CIMP-high CRC in both NSHDS and CRUMS showed a trend of reduced cancer-specific survival compared with CIMP-negative, which was not statistically significant. However, in the MSS tumor subgroup, CIMP-high was associated with a very poor prognosis, which was statistically significant in univariate and multivariate analysis in NSHDS and in univariate analysis in CRUMS. The multivariate hazard ratio (HR, adjusted for sex, age at diagnosis, tumor location, tumor stage, and adjuvant chemotherapy) in NSHDS was 3.05 (95\% CI, 1.40-6.63), and in CRUMS 1.38 (95\% CI, 0.62-3.07).

Other studies have reported a shorter survival in CIMP+ patients, although the results have not always been statistically significant.\textsuperscript{103,179,181,185,187} The evidence for a poor prognosis in MSS CIMP+ has been more convincing.\textsuperscript{97,102-103,136,178,180,182} However, some studies have reported null,\textsuperscript{176-177,183-184} or even opposite results.\textsuperscript{137,188}

CIMP-low and Patient Survival

Patients with CIMP-low CRC had a worse prognosis compared with CIMP-negative. This finding was statistically significant in multivariate analyses in NSHDS (multivariate HR, 2.01 [95\% CI, 1.20-3.37]) and of borderline significance in CRUMS (multivariate HR, 1.48 [95\% CI, 1.00-2.22]). The poor survival in CIMP-low patients was seen in both MSI and MSS subgroups. Few other studies have presented results for
survival of CIMP-intermediate tumor patients separately. In line with the results presented in here, Barault et al. found a worse prognosis in patients with CIMP-intermediate CRC. Most other studies have reported non-significant trends of a worse prognosis in CIMP-intermediate compared to CIMP-, or null findings.

Adjusting for *BRAF*

The V600E mutation in *BRAF* is strongly associated with CIMP+, and has been related to a poor patient prognosis. Several studies have suggested that the poor outcome of CIMP+ patients is caused by the high frequency of *BRAF* mutations in this subgroup. In paper III, the finding of a poorer prognosis in CIMP-low and MSS CIMP-high tumor patients remained after adjusting the results for *BRAF* mutation, which was in line with another large study. However, statistical significance was reduced, which may reflect the limited power of these analyses (*BRAF* mutation was found in 70.4% and 85.4% of CIMP-high tumors in NSHDS and CRUMS, respectively).

The relationship between CIMP+ and *KRAS* mutations is less established. Some studies have found a high frequency of *KRAS* mutations in CIMP-intermediate tumors. In addition, *KRAS* mutation has been associated with a poor patient prognosis. Yagi et al. found a poorer survival in patients with CIMP-intermediate and *KRAS* mutated, but not CIMP-intermediate and *KRAS* wild-type, tumors.

In paper III, information on *KRAS* mutations was lacking and could therefore not be considered in the survival analyses, which was a limitation of the study. Also, although a large number of well-characterized cases was included, analyses stratified by CIMP and MSI screening status often suffered from low power. The MSI CIMP+ subgroup of patients comprise only 10-12% of all CRCs, and are therefore particularly difficult to study.
Tumor-Infiltrating T Cells and Patient Survival

One of the features associated with CIMP+ and MSI in CRC is the finding of lymphocytes within the tumor. The presence of tumor-infiltrating lymphocytes is believed to represent the host response against the tumor, and has been associated with a favorable patient prognosis in several large studies. However, few large studies have considered the potential confounding effect of MSI and CIMP in this context.

To investigate the impact of tumor-infiltrating T cells on CRC patient survival, T cells were visualized and quantified in tumor sections immunohistochemically stained for anti-CD3 in 484 CRUMS cases. Patients whose tumors were highly infiltrated by T cells had a better prognosis compared with patients whose tumors were poorly infiltrated. The multivariate HR (adjusted for sex, age, tumor site, tumor stage, and adjuvant chemotherapy) for high versus low T-cell infiltration was 0.57 (95% CI, 0.31-1.00; paper IV). This finding was also statistically significant in separate analyses of stage II tumors (for high versus low infiltration by T cells, multivariate HR, 0.25 [95% CI, 0.08–0.75]. Based on these findings and other investigations of stage II CRC, a randomized clinical trial exploring the potential benefit of adjuvant chemotherapy in patients with stage II tumors poorly infiltrated by T cells is warranted.140, 143, 207, 253-254, 257-258

Tumor-Infiltrating T cells, MSI, CIMP, and Patient Survival

In analyses combining MSI screening status and the degree of infiltrating T cells, patients with MSI tumors highly infiltrated by T cells had the best prognosis (5-year cancer-specific survival, 87%), whereas patients with poorly infiltrated MSS tumors had the worst prognosis (5-year cancer-specific survival, 39%). However, patients with highly infiltrated MSS tumors had better prognosis compared with patients with MSI tumors what were less infiltrated (log rank p, 0.013). This finding indicates that the density of tumor-infiltrating T cells is more important for patient prognosis than MSI screening status. These findings are in
in line with other studies,\textsuperscript{208, 248, 252-255, 259-260} although contradictory observations have also been reported.\textsuperscript{146, 257}

CIMP-low tumors with low T-cell infiltration were associated with a particularly poor patient prognosis (5-year cancer-specific survival, 29\%). In analyses restricted to poorly infiltrated tumors, the multivariate HR for CIMP-low \textit{versus} CIMP-negative was 3.07 (95\% CI, 1.53-6.15). When including both CIMP status and the degree of tumor-infiltrating T cells in the multivariate model, both factors were identified as independent prognostic markers.

\textbf{Radiotherapy, Tumor-Infiltrating T cells, and Patient Survival in Rectal Cancer}

As previously shown by others,\textsuperscript{261-262} a reduction in the density of tumor-infiltrating T cells was found in rectal cancers that had been treated with preoperative radiotherapy. In line with previous findings,\textsuperscript{261} a better prognosis was found in rectal cancer patients with highly infiltrated tumors, regardless of whether radiation therapy had been administered.
CONCLUSIONS

In paper I, a reduced risk of rectal, but not colon, cancer with increasing plasma concentrations of vitamin B12 was observed. In paper II, a reduced risk of CIMP-high and CIMP-low CRC was found in study subjects with the lowest levels of plasma folate.

In paper III, patients with MSS CIMP-high tumors were identified as a subgroup with poor prognosis. Patients with CIMP-low CRC also had a poorer prognosis, particularly in combination with a low T-cell infiltration of the tumor (paper IV). In paper IV, a high density of tumor-infiltrating T cells was a prognostic marker for a good patient prognosis, independently of MSI screening status.

Based on the evidence to date, including the results presented in this thesis, the relation between one-carbon metabolism and DNA hypermethylation is not clear. Given the important role of DNA methylation in CRC tumorigenesis, and the ongoing debate concerning mandatory folic acid fortification of foods, further investigation is warranted.

Rather than being one disease, CRC includes several subtypes of tumors with different clinico-pathological and molecular characteristics, which should be considered in future investigations of risk factors and patient outcome.
ACKNOWLEDGEMENTS


Tack Bethany van Guelpen för att du är en förebild på flera plan. Du har hjälpt mig att växa in i rollen som forskare och till på köpet gjort mig till en bättre och starkare atlet.

Tack Maria Henriksson för att du hjälpt mig att utvecklas i labbet. Jag är imponerad över att du hållit masken när jag har ställt mina mest osmarta (för det finns ju inga dumma) frågor.

Och tack Ake Öberg för att du hållit i spakarna när det gäller de mest kliniska delarna av projekten. Att ha en kirurg i sitt handledarteam har varit ovärderligt.

Det har känts oerhört tryggt att ha ett team av handledare med expertis som täcker alla områden inom translationell forskning - from bench to bedside, vid min sida.

Colorektalgruppens nuvarande och tidigare medlemmar, och speciellt

Roger Stenling    Sofia Edin
Kerstin Näslund   Maria Jacobsson
Jörgen Rutegård   Martin Isaksson-Mettävainio
Kent Persson      Anna Löfgren-Burström
Vincy Eklöf


Speciellt tack till Inger Cullman för all hjälp med MethyLight och mycket annat under de första åren. Även denna vänlighet bidrog till att jag utan att tveka ville bli doktorand. Johan Hultdin såg till att jag fick erfarenhet av hur plasmaanalyser utförs på lab, och såg dessutom till att jag förstått saken rätt. Även tack till Ingegerd Johansson och
Göran Hallmans som är drivande krafter inom biobanken och dess projekt. Göran är dessutom den som styrde mig rätt här i livet genom att vidarebefordra min förfrågan om sommarjobb till colorektalgruppen, vilket jag är mycket tacksam för.


Tack Karin Bodén, Åsa Lundsten, Terry Persson och Helena Fousinette för att ni gör vårt arbete lättare och vår arbetsplats så trevlig. Tack även för ert överseende med en som inte begriper sig riktigt på vilket papper som ska vart inom administrationen.

Stort tack till följande fonder som bidragit till min forskning under åren:

Cancerforskningsfonden i Norrland
Stiftelsen J C Kempes Minnes Stipendiefond
Anna Cederbergs stiftelse för medicinsk forskning

Kollegor och finansiärer i all åra, men utan stödet hemmifrån hade det inte gått. Den här boken skulle inte finnas utan min mamma och pappa som alltid uppmuntrat mig till studier. Och med en storebror vid sin sida är man inte rädd för nåt! Tack för att ni stöttat mig vad jag än tagit mig för (med undantag för motorcykelkortet).

Tack också till Familjen Eriksson för att ni alltid får mig att känna mig så välkommen till Dalarna och Leksand.

Eva och lillasyster Elina, ni gör det extra roligt att komma hem.

Och vad är livet utan vänner? Jag vill tacka alla er som förgyller mitt liv.

Utan Vincy vid min sida de senaste 7 åren hade allt varit mycket tråkigare. Klanmötena tillsammans med Susanne Gidlund och Magdalena Hägg är alldeles för sällan.

Jag är Kristina Persson evigt tacksam för att du lärt mig åka skidor, och jag är så glad att du min vän. Värmland är för långt bort!

Till medlemmarna i bokklubben, våra möten är något jag alltid ser fram emot. Och tack Emma Andersson för att du startade den!

Familjen Hjelte/Björkén, Umeå är sig inte likt sen ni flyttade, och jag saknar er!


Jag önskar vi sågs lite oftare.

Johan, du är det bästa jag vet. Tack för att du gör mig lycklig!
REFERENCES


145. Nagtegaal ID, Marijnen CA, Kranenbarg EK, et al. Local and distant recurrences in rectal cancer patients are predicted by the nonspecific immune response; specific immune response has only a systemic effect--a histopathological and immunohistochemical study. BMC Cancer 2001;1:7.


