Molecular understanding of KRAS- and BRAF-mutated colorectal cancers

Ida Lundberg
To my family
   - I love you to the moon and back
# Table of contents

Table of contents i
Abstract ii
Populärvetenskaplig sammanfattning iii
List of abbreviations vi
List of original papers vii
Introduction 1
  Colorectal cancer 1
    The biology of CRC 2
    KRAS/BRAF signalling 3
    KRAS- and BRAF-mutated CRCs 5
  Cancer stem cells 8
    The stem cell factor SOX2 10
    The intestinal epithelial factor CDX2 10
    Other markers of cellular stemness 11
  Tumour immunity 12
    Immune factors involved in tumour rejection and promotion 12
    Immune cells in CRC progression 14
  miRNA in gene regulation and cancer 14
    miRNAs: regulators of gene expression 14
    miRNAs in CRC 15
Aims of the thesis 16
Materials and methods 17
  CRC cell lines 17
  Patient cohort 18
  Gene expression analysis 18
  Protein expression analysis/IHC 19
Results and discussion 21
  Paper I 21
  Paper II 22
  Paper III 25
  Paper IV 27
Conclusions 30
Acknowledgements 32
References 34
Abstract

Colorectal cancer (CRC) is the third most commonly diagnosed malignancy in both men and women, and one of the leading causes of cancer-related deaths worldwide. One frequently mutated pathway involved in oncogenesis in CRC is the RAS/RAF/MAP kinase pathway. Oncogenic activation of KRAS and BRAF occur in 30–40% and 5–15% of all CRCs, respectively, and the mutations are mutually exclusive. Even though KRAS and BRAF are known to act in the same pathway, KRAS- and BRAF-mutated CRCs have different clinical and histopathological features. For example, BRAF mutation in CRC is tightly linked to microsatellite instability (MSI) and a CpG island methylator phenotype (CIMP), which is not seen in KRAS-mutated tumours. BRAF-mutated CRCs are also more often found in right-sided tumours. However, the underlying molecular reasons for these differences have not yet been defined.

The overall aim of this thesis was to investigate molecular differences between KRAS- and BRAF-mutated CRCs to understand how KRAS and BRAF mutations differentially affect tumour progression. We used an in vitro cell culture system to explore molecular differences between KRAS- and BRAF-mutated CRCs and verified our findings using CRC tissue specimens from the Colorectal Cancer in Umeå Study (CRUMS).

We found that BRAF mutation, but not KRAS mutation, was associated with expression of the stem cell factor SOX2. Furthermore, SOX2 was found to be correlated to a poor patient prognosis, especially in BRAF-mutated cancers. We further investigated the role of BRAF in regulation of SOX2 expression and found that SOX2 is at least partly regulated by BRAF in vitro. We continued by investigating the functional role of SOX2 in CRC and found that SOX2-expressing cells shared several characteristics with cancer stem cells, and also had down-regulated expression of the intestinal epithelial marker CDX2. There was a strong correlation between loss of CDX2 expression and poor patient prognosis, and patients with SOX2 expression were found to have a particularly poor prognosis when CDX2 levels were down-regulated. In conclusion, in these studies we identified a subgroup of BRAF-mutated CRCs with a particularly poor prognosis, and having a cancer stem cell-like appearance with increased expression of SOX2 and decreased expression of CDX2.

Tumour progression is regulated by interactions with cells of the immune system. We found that BRAF-mutated CRCs were more highly infiltrated by Th1 lymphocytes than BRAF wild-type tumours, while the opposite was true for KRAS-mutated CRCs. Interestingly, we found that part of this difference is probably caused by differences in secreted chemokines and cytokines between KRAS- and BRAF-mutated CRCs, stimulating different arms of the immune response.

Altered levels of expression of miRNAs have been seen in several malignancies, including CRC. We found that BRAF- and KRAS-mutated CRCs showed miRNA signatures different from those of wild-type CRCs, but the expression of miRNAs did not distinguish KRAS-mutated tumours from BRAF-mutated tumours.

In summary, our findings have revealed possible molecular differences between KRAS- and BRAF-mutated CRCs that may explain some of the differences in their clinical and histopathological behaviour.
Populärvetenskaplig sammanfattning

I Sverige får varje år ca 6000 människor diagnosen kolorektalcancer (cancer i tjock- och ändtarm), vilket gör den till den tredje vanligaste formen av cancer efter prosstata- och bröstcancer. Även om dödligheten i kolorektalcancer har minskat de senaste årtiondena till följd av framsteg i behandling, så är det fortfarande runt en tredjedel som inte överlever. Detta beror på att cancern ofta redan före diagnos har hunnit sprida sig (metastasera) till andra delar av kroppen. För att kunna förbättra behandling och därmed också prognosen hos patienter med kolorektalcancer är det viktigt att studera de molekylära mekanismerna som ligger bakom sjukdomens uppkomst och spridning. Denna kunskap kan hjälpa till att identifiera de subgrupper av kolorektalcancer som har sämre prognos och därmed motivera intensivare behandling i syfte att förbättra deras överlevnad.

Endast 10% av alla kolorektalcancerfall har en tydlig ärtlig förklaring, resterande 90% anses som sporadiska (icke ärtliga). Cancer är resultatet av en stegvis ansamling av förändringar (mutationer) i gener, vilket bland annat leder till förlust av tumörspressorgener och ökat uttryck av onkogener. KRAS och BRAF är två onkogener som ofta är muterade i kolorektalcancer, ca 40% är muterade i KRAS och 10% är muterade i BRAF. KRAS och BRAF är proteiner båda involverade i samma signaleringsväg, MAPK-signaleringsvägen, som reglerar uttrycket av många andra gener, tex gener som styr celldelning. Mutationer i KRAS och BRAF resulterar i proteiner som är ständigt aktiva och som leder till en överaktiv signalering oberoende av de tillväxtfaktorer som normalt aktiverar signaleringsvägen. Även om både KRAS och BRAF är involverade i samma signaleringsväg så har tumörer med dessa mutationer flera olika kliniska och histopatologiska skillnader. BRAF-mutation i kolorektalcancer är till exempel kopplat till höger-sidiga tumörer och mikrosatellit instabilitet (MSI), dvs de uppkommer pga inaktivering av DNA-reparationsproteiner. Till skillnad från BRAF-muterade kolorektalcancrar är de KRAS-muterade tumörerna oftare kromosomalt instabila, även kallade mikrosatellit-stabila (MSS). En annan skillnad kan ses i tumörernas tillväxtmönster, där KRAS-muterade tumörer oftare har ett infiltrativt växtsätt.

Syftet med studierna i denna avhandling var att undersöka molekylära skillnader mellan KRAS- och BRAF-muterade kolorektalcancrar för att öka förståelsen om hur KRAS- och BRAF-mutationer påverkar tumörens progression. För att studera detta använde vi oss av både in vitro-studier med cancercell-linjer muterade i KRAS eller BRAF, och tumörvävnad från patienter med kolorektalcancer.

Tumörens progression har visat sig påverkas av interaktioner mellan tumörceller och cells som ingår i immunsystemet. Infiltration av T-lymfocyter, en typ av vita blodkroppar, i och runt tumören har förknippats med en god prognos hos patienter med kolorektalcancer. Vi fann att BRAF-muterade tumörer var mer infilterade av en subgrupp av T-lymfocyter, så kallade Th1-lymfocyter, medan KRAS muterade tumörer istället var mindre infilterade. Ett immunsvår involverande Th1-lymfocyter anses motverka tumörens progression, medan en infiltration av Th2-lymfocyter istället anses främja tumörens progression. Vi fann att BRAF- och KRAS-muterade kolorektalcancerar utsöndrar olika signalmolekyler (cytokiner och kemokiner), som kan stimulera olika delar av immunsvaret. De BRAF-muterade tumörerna uttryckte signalmolekyler som kan attrahera och aktivera Th1-lymfocyter, medan KRAS-muterade tumörer istället visade sig uttrycka signalmolekyler kopplat till ett immunsvår involverande Th2-lymfocyter. Vi fann även att BRAF-muterade kolorektalcancerar med låg infiltration av Th1-lymfocyter var kopplade till en särskilt dålig överlevnad, vilket indikerar att detta är en subgrupp som kan vara i behov av en utökad behandling.

MikroRNA (miRNA) är korta, icke-kodande RNA-sekvenser (dvs de översätts inte till proteiner). miRNA kontrollerar istället uttrycket av andra protein-kodande gener. Nivåer av olika miRNA har visat sig vara förändrade i cancer. Vi ville därför undersöka om KRAS- och BRAF-muterade
kolorektalcancrar har olika miRNA-signaturer, som kan förklara en del av de kliniska och molekylära skillnaderna mellan KRAS- och BRAF-muterade tumörer. Genom in vitro-cellstudier fann vi att genuttrycket av olika miRNA verkade skilja sig mellan cancerceller muterade i KRAS och cancerceller muterade i BRAF, men dessa resultat kunde inte verifieras i tumörvävnad. Däremot fann vi att uttrycket av de miRNA undersökta i den här studien, skiljde sig i KRAS- och BRAF-muterade tumörer jämfört mot tumörer som inte var muterade i KRAS eller BRAF. Detta tyder på att mutationer i KRAS och BRAF, som båda resulterar i en överaktiv MAPK-signalering, på lika sätt reglerar uttrycket av de miRNA inkluderade i den här studien.

Sammanfattningsvis påvisar studierna i den här avhandlingen flera molekylära skillnader och likheter mellan KRAS- och BRAF-muterade kolorektalcancrar, vilka möjligen kan förklara en del av de kliniska och histopatologiska skillnaderna.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APC</td>
<td>Adenomatous polyposis coli</td>
</tr>
<tr>
<td>CDX2</td>
<td>Caudal type homeobox 2</td>
</tr>
<tr>
<td>CIN</td>
<td>Chromosomal instability</td>
</tr>
<tr>
<td>CIMP</td>
<td>The CpG island methylator phenotype</td>
</tr>
<tr>
<td>CRC</td>
<td>Colorectal cancer</td>
</tr>
<tr>
<td>CRUMS</td>
<td>The Colorectal Cancer in Umeå Study</td>
</tr>
<tr>
<td>CSC</td>
<td>Cancer stem cell</td>
</tr>
<tr>
<td>CTL</td>
<td>Cytotoxic lymphocyte</td>
</tr>
<tr>
<td>ECM</td>
<td>Extracellular matrix</td>
</tr>
<tr>
<td>EMT</td>
<td>Epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>FAP</td>
<td>Familial adenomatous polyposis</td>
</tr>
<tr>
<td>FFPE</td>
<td>Formalin-fixed, paraffin-embedded</td>
</tr>
<tr>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
</tr>
<tr>
<td>GAP</td>
<td>GTPase-activating protein</td>
</tr>
<tr>
<td>GEF</td>
<td>Guanine nucleotide exchange factor</td>
</tr>
<tr>
<td>HNPCC</td>
<td>Hereditary non-polyposis colorectal cancer</td>
</tr>
<tr>
<td>IHC</td>
<td>Immunohistochemistry</td>
</tr>
<tr>
<td>MAPK</td>
<td>Mitogen-activated protein kinases</td>
</tr>
<tr>
<td>miRNA</td>
<td>MicroRNA</td>
</tr>
<tr>
<td>MMP</td>
<td>Matrix metalloproteinase</td>
</tr>
<tr>
<td>MSI</td>
<td>Microsatellite instability</td>
</tr>
<tr>
<td>MSS</td>
<td>Microsatellite stable</td>
</tr>
<tr>
<td>RTK</td>
<td>Receptor tyrosine kinase</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>SOX2</td>
<td>SRY (sex-determining region Y)-box 2</td>
</tr>
<tr>
<td>TGFβ</td>
<td>Transforming growth factor beta</td>
</tr>
<tr>
<td>Th1/2</td>
<td>T helper 1/2</td>
</tr>
<tr>
<td>Treg</td>
<td>T regulatory</td>
</tr>
</tbody>
</table>
List of original papers

Paper I

Paper II

Paper III

Paper IV

The original articles were reprinted with permission from the publishers.
Introduction

Colorectal cancer

Colorectal cancer (CRC) is one of the most commonly diagnosed malignancies worldwide, the third most common cancer in men, and the second most common cancer in women. In 2011, around 6,000 new cases of CRC were diagnosed in Sweden, representing 10.7% of all malignancies. Only breast cancer and prostate cancer have a higher incidence.

The pathogenesis of CRC is complex, and includes both hereditary and environmental factors. In approximately 10% of CRC patients there is a clear familial association. The two most known CRC syndromes are familial adenomatous polyposis (FAP) and hereditary non-polyposis colorectal cancer (HNPCC), also called Lynch syndrome. FAP accounts for less than 0.1% of all CRC cases, while HNPCC accounts for 1–3% of them. Apart from these genetic CRC syndromes, the risk of developing CRC is increased in families with a history of CRC, especially if a close relative is diagnosed with CRC before the age of 55 or more than one relative has been diagnosed with CRC. Other risk factors for developing CRC are high age, previous premalignant polyps in the large bowel, smoking, and inflammatory bowel disease. The incidence of CRC is significantly higher in countries with a typical “western lifestyle”, characterised by an unhealthy diet, low physical activity, and obesity.

CRC is also one of the leading causes of death from cancer, with approximately one-third of CRC patients dying from the disease, mostly involving metastasis. The stage of the disease at diagnosis is the most important prognostic factor in CRC. CRC is staged from I to IV according to the depth of bowel wall infiltration, spread to lymph nodes, and/or distant metastases. Surgery is the most important curative treatment. In more advanced stages, adjuvant or palliative chemotherapy is given. In addition, rectal cancers are often considered for preoperative radiotherapy. In the past, patients with distal metastases have had a very poor prognosis, but the introduction of surgical resection of liver metastases has contributed to improved survival in this patient group. Despite similar clinicopathological characteristics, such as tumour stage and tumour histology, the clinical course and the response to treatment vary. Even though our knowledge of molecular subgroups of CRC is improving, our understanding of the molecular basis of why some tumours are more aggressive than others is still poor.
The biology of CRC

CRC is a biologically heterogeneous disease with different molecular characteristics, and about 90% of all cases are diagnosed as being sporadic. Sporadic tumours arise from pre-neoplastic lesions derived from normal colonic epithelium, which progress to malignant tissue through a multi-step process. This process is known as the adenoma-to-carcinoma sequence and depicts the genetic events required for a normal colon epithelium to be transformed into a malignant phenotype in most of the sporadic CRC cases. Today, it is known that CRCs arise through three major pathways, involving combinations of genetic and epigenetic changes. These are the chromosomal instability (CIN) pathway, the microsatellite instability (MSI) pathway, and the CpG island methylator phenotype (CIMP) pathway.

The CIN pathway is responsible for approximately 70–85% of all sporadic CRCs, and is characterised by an increased rate of chromosomal change, resulting in aneuploidy. Tumours that arise through this pathway are classified as microsatellite stable (MSS). Genomic changes identified in the MSS adenoma-carcinoma sequence include activation of oncogenes such as KRAS, and inactivation of tumour suppressor genes such as adenomatous polyposis coli (APC) and tumour suppressor p53 (TP53), as well as loss of heterozygosity in chromosome 18q (containing the genes DCC, SMAD2, and SMAD4).

The MSI pathway accounts for approximately 10–15% of all sporadic CRC cases. BRAF is an oncogene that is commonly mutated in this pathway, and CTNNB1 (encoding β-catenin) is also mutated in some cases. MSI tumorigenesis results from inactivation of the DNA mismatch-repair (MMR) system. MMR proteins normally prevent mutations by correcting errors in DNA that spontaneously occur during DNA replication. In MSI tumours, inactivation of one or several MMR genes leads to accumulation of mutations (i.e. insertions or deletions). These mutations are more prone to occur in repetitive DNA sequences, called microsatellites, since the DNA polymerase is more susceptible to making errors when replicating these short repeated sequences. The most common aberration in MSI tumours is caused by promoter hypermethylation of the MMR gene MLH1.

The CIMP pathway is found in approximately 15–20% of all sporadic CRC cases, and involves epigenetic changes that were first described by Toyota et al. DNA methylation commonly occurs at CpG sites, where a cytosine nucleotide is located next to a guanine nucleotide. The CIMP pathway is driven by promoter hypermethylation of CpG islands, regions with a high frequency of CpG sites, thereby silencing genes such as tumour suppressors. Several genes involved in CRC progression have been shown to be
epigenetically silenced by hypermethylation, such as \( APC \), \( MLH1 \), and \( CDKN2A \) (p16). As the majority of MSI tumours also display a CIMP phenotype, there is substantial overlap between these two pathways.

Figure 1 summarises the most common genetic and epigenetic events associated with progression of CRC.

**Figure 1.** Common genetic and epigenetic events in the progression from normal colonic mucosa to invasive cancer.

**KRAS/BRAF signalling**

The progression from normal colon to CRC involves activation of oncogenes and/or repression of tumour suppressor genes. One frequently mutated pathway involved in oncogenesis in CRC is the mitogen-activated protein kinase (MAPK) pathway, which includes the two major oncogenic drivers of CRC, KRAS and BRAF. The MAPK pathway consists of a chain of proteins that transduces signals from the cell surface to the cell nucleus, and activates genes implicated in many important cellular functions such as cell growth, division, and differentiation.
KRAS acts as a central component of signalling pathways originating from cell surface receptor tyrosine kinases (RTKs) such as the epidermal growth factor receptor (EGFR). The KRAS gene codes for the protein KRAS, a cell membrane-bound small GTPase functioning as a molecular on/off switch that is activated by cell surface receptors 27. Activation of KRAS is stimulated through guanine nucleotide exchange factors (GEFs), which catalyse the exchange of GDP to GTP. Active GTP-bound KRAS is in turn negatively regulated by GTPase-activating proteins (GAPs), which accelerate the hydrolysis of GTP into GDP (Figure 2) 28.

The BRAF gene, encoding the protein BRAF, is a downstream effector of KRAS in the MAPK signalling transduction pathway 29. Following activation, KRAS associates with the RAS-binding domain of BRAF and recruits the complex to the cell membrane 30. The binding of KRAS to BRAF, in combination with conformational changes and regulatory phosphorylations/dephosphorylations, activates BRAF 31, 32. BRAF is a serine/threonine kinase that—when activated—stimulates a downstream signalling cascade by phosphorylating and activating MEK, which in turn activates ERK 33. Activated ERK then phosphorylates and activates different targets, either in the cytoplasm or in the nucleus. The targets of ERK include for example transcription factors, protein kinases/phosphatases, and cytoskeletal proteins, thereby regulating multiple cellular processes 34, 35. ERK is also involved in several negative feedback loops in the MAPK pathway, for example by directly inhibiting BRAF activity or by inhibiting the binding of adaptor proteins to activated RTKs, which is needed for downstream signalling events 36.

BRAF is not the only downstream target of KRAS (Figure 2). Active KRAS can also stimulate the PI3K pathway, which regulates cell survival and cell proliferation 37. The RalGEF/Ral pathway is another downstream effector of KRAS, involving the Rho-family GTPases CDC42 and Rac 38, 39. Apart from KRAS and BRAF, there are also other RAS and RAF family members. HRAS, KRAS, and NRAS (amongst others) belong to the RAS family 40, and the RAF family consists of ARAF, BRAF and CRAF 41. However, in the work in this thesis we only studied KRAS and BRAF, since they are the most commonly mutated gene members of the RAS and RAF families in CRC.
Figure 2. Signalling events downstream of KRAS.

BRAF is the most well-known downstream target of KRAS and is involved in the MAPK kinase cascade. The next-best characterised downstream effector of KRAS is PI3K. KRAS is also known to activate the RalGDS pathway.

**KRAS- and BRAF-mutated CRCs**

Approximately 40% of all CRCs have activating mutations in the *KRAS* gene, while activating mutations in the *BRAF* gene occur in about 10% of the tumours. Targeted therapy against EGFR, an activator of the MAPK signalling pathway, is used today in the clinic and has improved the survival of metastatic CRC patients. However, only 10–20% of the patients benefit from anti-EGFR treatment. Resistance to the therapy is largely ascribed to *KRAS* and *BRAF* mutations, since these mutations result in a constitutively active signalling downstream of EGFR; upstream inhibition, such as anti-EGFR treatment, is therefore ineffective.
Ninety per cent of all KRAS mutations in CRC occur at codons 12 and 13, where the most common mutations are G12D (G→A transition) and G12V (G→T transversion) in codon 12 and G13D (G→A transition) in codon 13 (Figure 3). It has been shown that the interaction of KRAS with GAPs is impaired in these mutations, indicating that the glycine residues at positions 12 and 13 are important sites for inactivation of KRAS. Thus, mutations at these positions result in a constitutively active KRAS.

BRAF mutations in CRC are mostly seen at codon 600, and almost all are single-nucleotide mutations resulting in replacement of valine with glutamic acid (V600E) (Figure 3). The valine at position 600 has been shown to be an important site for holding the BRAF protein in an inactive conformation. The V600E mutation results in a constitutively active BRAF that phosphorylates MEK independently of KRAS binding.

**Figure 3.** Frequencies of KRAS and BRAF mutations in CRC.

Both KRAS and BRAF mutations are classified as driver mutations, since they are both mutated prior to malignant conversion and are therefore primary genetic events in CRC tumorigenesis. KRAS and BRAF mutations are also associated with a poor patient prognosis in CRC, but BRAF has been shown to be a more potent oncogene than KRAS. The two mutations are believed to be mutually exclusive in CRC.

One explanation for the
mutual exclusivity is that the two mutations may be incompatible, since an excess of signalling through the MAPK cascade could result in cell cycle arrest.56

KRAS mutation can occur together with loss of APC, while BRAF and APC aberrations are rarely found together in CRC. It has also been found that KRAS mutation synergizes with APC inactivation in CRC progression, resulting more often in invasive adenocarcinomas.57,58 BRAF mutation is strongly associated with a serrated histology, both in the earliest precursors of CRC and in adenomas, but this is not seen in KRAS-mutated CRCs with the same frequency.59,60 BRAF-mutated CRCs also often display a CIMP-high phenotype, with promoter hypermethylation of numerous genes.61-63. BRAF mutation is more often found in MSI tumours with dysfunctions in the DNA mismatch-repair system.64,65. MSI is associated with a good prognosis in CRC, and BRAF-mutated MSI tumours have a better patient prognosis than BRAF-mutated MSS tumours.50,66.

Figure 4. The main differences that can be distinguished between KRAS- and BRAF-mutated CRCs.
Clinically, \textit{BRAF}-mutated CRCs are more often located in the right colon and are seen more frequently in female patients. A high infiltration rate of lymphocytes, a poorly differentiated tumour, and a mucinous histology are other characteristics of \textit{BRAF}-mutated CRCs \textsuperscript{67}. Unlike \textit{BRAF} mutation, \textit{KRAS}-mutated CRCs are associated with MSS tumours and a CIMP-low or CIMP-negative phenotype \textsuperscript{65, 68}. Another difference between \textit{KRAS}- and \textit{BRAF}-mutated CRCs can be found in their growth pattern. We have observed that \textit{KRAS}-mutated tumours more often show an infiltrating growth pattern. Figure 4 summarises the main differences between \textit{KRAS}- and \textit{BRAF}-mutated CRCs.

In summary, \textit{KRAS} and \textit{BRAF} are both major oncogenic drivers of CRC. They are both involved in the same signalling pathway, but have different genetic/epigenetic, pathological, and clinical characteristics. A better understanding of the molecular mechanisms that lead to differences between \textit{KRAS}- and \textit{BRAF}-mutated CRCs is of utmost importance, since it may provide information necessary for advances in diagnosis, prediction and new or improved therapeutic strategies.

\section*{Cancer stem cells}

Recent research has revealed that a small subgroup of tumour cells have characteristics associated with stem cells, and they have therefore been called cancer stem cells (CSCs). CSCs have the capability of self-renewal and multi-lineage differentiation, features that can drive both tumour growth and the emergence of new tumours \textsuperscript{69, 70}.

In tumorigenesis, the traditional stochastic model proposes that a small population of tumour cells acquire the appropriate mutations necessary for tumour initiation and propagation. Within this model, all cells have the same probability to become tumorigenic (Figure 5A). However, the CSC model proposes that tumours are organized hierarchically, with CSCs lying at the apex and driving and sustaining tumour growth \textsuperscript{71, 72} (Figure 5B). Just like normal stem cells, CSCs divide asymmetrically, with one daughter cell retaining stem cell properties and the other daughter cell differentiating and generating the bulk of the tumour \textsuperscript{73}. However, it has recently been proposed that differentiated cells can become de-differentiated and generate CSCs \textsuperscript{74}. This model is called the plastic CSC theory, and is a dynamic bidirectional conversion between stem cell and non-stem cell states and implies that CSCs can arise \textit{de novo} from already differentiated cells \textsuperscript{75} (Figure 5C). This model reconciles the stochastic and hierarchical models into one model and could be one explanation for the tumour heterogeneity seen in CRC.
CSCs were first identified in leukemia 76, but were later on found in many solid malignancies, including CRC 77. Several studies have shown that CSCs are linked to patient prognosis in CRC 78-80, indicating that CSCs are a subgroup of tumour cells of clinical significance, and putative therapeutic targets.

The capabilities required for tumour cells to metastasise can probably only be seen in a small number of tumour cells. Amongst other features, they have to be able to migrate, to evade the immune system, and to colonize at a new tumour site. CSCs have several of these characteristics, which suggests that they would be likely candidates for the subset of cells that gives rise to metastases. If the plastic CSC theory is true, CSCs are the tumour cells most capable of forming metastases, and their suggested plasticity provides an advantage in adapting to the foreign environment. According to the “seed and soil” hypothesis 81, which proposes that metastasis depends on interaction between tumour cells (the “seeds”) and the microenvironment (the “soil”), CSCs have been suggested to be the “lethal seeds” that can re-initiate tumour growth at the distant site 82. CSCs also have the ability to create spheres in culture. They grow slowly and form clusters that do not require adherence to survive, which is not the case for differentiated cells. Compared to adherent cells, sphere-forming CSCs have been shown to be more aggressive, and thereby metastatic 83. The quiescent slow-cycling phenotype seen in CSCs (compared to differentiated cells) would at least partly explain their resistance to therapeutic agents and their possible role in tumour recurrence 84, 85.

Figure 5. Different models of tumorigenesis: the stochastic model (A), the CSC model (B), and the plastic CSC model (C).
**The stem cell factor SOX2**

The large SRY (sex-determining region Y)-box (SOX) gene family consists of transcription factors that are characterised by containing a high-mobility group (HMG). The HMG of the protein is the DNA-binding domain, which allows the SOX protein to regulate gene transcription by binding to specific DNA sequences. SOX proteins are so-called architectural transcription factors since they induce a bend in the DNA and thereby alter its conformation. The SOX proteins are also known to interact with other transcription factors and cofactors in order to efficiently regulate gene transcription \(^{86}\). The SOX family is known to be important in transcription of genes involved in the regulation of developmental processes, such as determination of cell fate \(^{87}\).

The key member SOX2 belongs to the SOXB1 subgroup together with the SOX1 and SOX3 genes, which are closely related to SOX2. SOX2 is important in the maintenance of cell pluripotency and self-renewal of stem cells \(^{88}\), as well as CSCs \(^{89}\), suggesting a potential oncogenic role for SOX2. SOX2 is one of the factors that are necessary to induce pluripotency in stem cells by reprogramming of differentiated cells \(^{90}\). Recent cancer research has suggested that SOX2 has an important role in tumorigenesis, with aberrant expression of SOX2 being reported in CRC \(^{91-93}\) as well as in several other malignancies including breast and pancreatic cancer \(^{94, 95}\). An association between SOX2 expression and cellular stemness has been shown in several malignancies, such as human ovarian, cervical, pancreatic, head and neck squamous-cell, and breast carcinoma \(^{96-98}\), but to date no such association has been proven for CRC.

**The intestinal epithelial factor CDX2**

Caudal type homeobox 2 (CDX2) is a transcription factor that regulates the expression of intestine-specific genes involved in cell differentiation \(^{99}\). Since the expression of CDX2 is restricted to the gastrointestinal tract, it is used today as a diagnostic marker to identify normal intestinal epithelium and tumours of intestinal origin \(^{100}\).

Homeobox genes are known to be aberrantly expressed in cancers, and both loss and gain of expression have been associated with tumour development \(^{101}\). In the normal epithelium of the colorectum, CDX2 is expressed at high levels, but these levels are reduced or lost in a subset of CRCs. CDX2 has been proposed to have a tumour suppressor role in CRC \(^{102}\), and loss of CDX2 have been reported to be associated with poor prognosis \(^{103-105}\). Previous studies have linked loss of CDX2 to different tumour
features, such as CIMP-high, BRAF mutation, and MSI \textsuperscript{106-108}. A reduced expression of CDX2 has also been linked to tumours with poor differentiation \textsuperscript{103, 108, 109}. As described above, the presence of CSCs is also associated with poorly differentiated tumours, indicating a possible link between CSCs and CDX2.

**Other markers of cellular stemness**

CD44 has been proposed as a CSC marker in several malignancies including CRC \textsuperscript{110, 111}. CD44 is a transmembrane glycoprotein and a major cell surface receptor for hyaluronic acid, and is thereby important for cell adhesion to the extracellular matrix (ECM) and to other cells \textsuperscript{112}. RTKs such as EGFR can be activated by CD44 binding to hyaluronic acid, indicating that CD44 is also involved in activating signalling cascades such as the MAPK and PI3K pathways, and thereby possibly also tumour initiation \textsuperscript{113}. When tumour cells metastasise, they need to adhere to the blood vessels, both for intravasation and extravasation \textsuperscript{114}. It has been shown that there is a correlation between expression of CD44 on tumour cells and both adhesion of cancer cells to endothelial cells and metastasis \textsuperscript{115}. In CRC, increased expression of CD44 by CSCs has been suggested to initiate the metastatic process \textsuperscript{78}. Reduced expression of CD44 has been shown to reduce clonal formation \textit{in vitro} and also to reduce tumorigenicity \textit{in vivo} in CRC, indicating that CD44 may be a putative CSC marker in CRC \textsuperscript{116}. Expression of a cluster of stem-like markers, including the two markers SOX2 and CD44, has recently been investigated in CRC and has been shown to be associated with a significantly worse prognosis \textsuperscript{117}.

Another suggested CSC marker in CRC is CD133, but CD133 is not restricted to stem cells—and both CD133-positive and CD133-negative metastatic CRC cells have the capability of tumour initiation, indicating that CD133 may not be a very good CSC marker in CRC \textsuperscript{118}.

CD24 is a cell adhesion molecule that has also been suggested as a CSC marker in CRC. An increased expression of CD24 has been found in CRC and is associated with lymph node metastasis, indicating that it can be used as a CSC marker with prognostic value \textsuperscript{119}. CD24 is also a ligand for P-selectin, which is expressed on endothelial cells and platelets, and CD24 can thereby mediate adhesion and rolling of tumour cells on the endothelium, facilitating the metastatic process \textsuperscript{120}. CD24 has also been shown to be regulated by SOX2 \textsuperscript{92}. 
Tumour immunity

It is well known that the cells in the tumour microenvironment, such as stromal cells and immune cells, affect tumour growth. These surrounding non-malignant cells can support the tumour by modulating the differentiation and proliferation of tumour cells, for example by secreting factors such as proteases and cytokines. This will create an environment that allows the tumour to gain properties necessary for growth, invasion, and metastasis.

As early as 1909, it was suggested by Ehrlich that the immune system might control tumour growth and thereby repress tumour development. Fifty years later, the cancer immunosurveillance concept was postulated by Burnet and Thomas, as the capacity of the immune system to recognize and destroy nascent transformed cells. For several decades, the concept of cancer immunosurveillance was questioned, until it was validated by animal experiments. This concept was then refined into the cancer immune-editing hypothesis, which describes the immune system-tumour interactions from immunosurveillance to tumour escape. Cancer immuno-editing describes three different phases, i.e. elimination, equilibrium, and escape, where the fate of the tumour is guided by the immune system. In the elimination phase, the tumour cells are destroyed by the immune system, as described by the immunosurveillance concept. Tumour cells that survive the elimination phase enter the equilibrium phase, where the tumour cells co-exist with the immune system, and where growth of the tumour is restrained by the immune system. In the escape phase, the immunologically educated tumour cells escape the immune system and tumour progression proceeds.

Immune factors involved in tumour rejection and promotion

Most human tumours, including CRC tumours, are infiltrated by immune cells. The immune response has been shown to play opposing roles in cancer progression, displaying both tumour-promoting and tumour-suppressive activities. The different roles of the immune response in tumour progression probably involve variations in the balance between different subsets of immune cells, as summarised in Figure 6. Briefly, on the one hand, the pro-inflammatory M1 macrophages promote and enhance the cytotoxic T lymphocyte (CTL) response as well as T helper 1 (Th1) polarisation, favouring tumour rejection. On the other hand, the anti-inflammatory M2 macrophages support the T helper 2 (Th2) and T regulatory (Treg) subsets, leading to immune suppression and tumour promotion.
The complex interplay between tumour cells and immune cells is influenced by cytokines and chemokines expressed by immune cells, and also by tumour cells and other non-cancerous cells found in the tumour microenvironment \cite{129}. M1 macrophages present antigen, produce pro-inflammatory cytokines (e.g. TNF, IL-12, and IL-23), and produce chemokines (e.g. CXCL9 and CXCL10) that support the recruitment and activity of Th1 cells and CTLs \cite{130}. Factors produced by Th1 are required for CTL activation, and when CTLs are activated, they have the ability to lyse tumour cells. M2 macrophages, on the other hand, produce immunosuppressive cytokines (e.g. IL-10 and transforming growth factor beta (TGFβ)) and chemokines (e.g. CCL17 and CCL22), supporting the immunosuppressive Th2/Treg axis \cite{130}. Th2 cells and Tregs promote tumour growth by suppressing the functionality of CTLs \cite{131}. Factors produced by M2 macrophages and Tregs (e.g. TGFβ, VEGF, and PDGF) also stimulate angiogenesis and tissue remodelling, promoting tumour growth and metastasis \cite{132}. Tumour cells also release factors that can influence the balance between tumour-promoting and tumour-suppressive immunity, such as TGFβ, TNF, and VEGF, adding another level of complexity to the anti-tumour immune response \cite{133}.
**Immune cells in CRC progression**

Immune cell infiltration is a strong prognostic factor in CRC \(^{134, 135}\), and is suggested to be even more predictive of tumour recurrence than tumour stage \(^{136}\). Increased immune cell infiltration, particularly of the Th1/CTL axis, is associated with better prognosis in CRC \(^{135, 137}\) and also in several other tumour types \(^{129}\). The prognostic role of infiltration of the Th2 and Treg subsets, on the other hand, vary between different cancers but are more often correlated to a poor patient prognosis \(^{129}\).

The complex interactions between tumour cells and the immune response are far from being well understood, and few studies have looked at subgroups of CRC in relation to T lymphocyte infiltration. One subgroup of CRC that has been studied in relation to immune infiltration is the MSI tumours, which have been shown to be associated with increased infiltration of lymphocytes \(^{138, 139}\). This can be explained by the increased number of frame-shift mutations seen in MSI tumours, leading to increased production of abnormal peptides that may be immunogenic \(^{140}\). It has also been suggested that \(KRAS\) mutation is associated with reduced expression of genes related to the Th1 axis \(^{141}\). A better characterisation of how the immune response is regulated in different molecular subgroups of CRC is very important, considering the many immunotherapies that are being introduced into clinical practice today.

**miRNA in gene regulation and cancer**

**miRNAs: regulators of gene expression**

MicroRNAs (miRNAs) are short, non-coding, single-stranded RNAs 18 to 24 nucleotides in length that post-transcriptionally modulate gene expression by binding to specific mRNA target sites \(^{142}\). The maturation of miRNA occurs through a multi-step process. First, RNA polymerase II or III transcribes miRNA genes as long primary transcripts that form a hairpin-shaped precursor miRNA. The precursor miRNA is then transported to the cytoplasm where it is cleaved to its mature length, removing the hairpin loop and creating a double-stranded miRNA duplex molecule. The functional strand of the miRNA duplex assembles into the RNA-induced silencing complex, while the passenger strand is degraded. Depending on how well the miRNA base pairs with the mRNA target, gene expression will be negatively regulated by either mRNA cleavage or translational repression. mRNA cleavage is dependent on perfect complementarity between the miRNA and the mRNA sequence, which is often seen in plant cells \(^{143}\) but not in animal
cells—where the nucleotides do not pair up as well. In animal cells, the 5´-end of the miRNA generally targets the 3´-untranslated region (UTR) of the mRNA, interfering with mRNA translation 144. Animal miRNAs also accelerate target mRNA de-adenylation (shortening of the poly(A) tail), which leads to mRNA decay 145.

One miRNA may have hundreds of different mRNA targets, and one mRNA may in turn be regulated by several different miRNAs 146. Today, thousands of miRNAs have been described in animal and plant cells, and about one-third of the human genome is thought to be regulated by miRNAs 147.

**miRNAs in CRC**

By targeting mRNAs, miRNAs are involved in several cellular processes such as tumour initiation and growth 148, 149. miRNAs may directly regulate tumour progression by acting as oncogenes or tumour suppressor genes. Different miRNAs have been found to show both reduced and increased levels in cancer, compared to normal tissue. Different mechanisms have been suggested to explain the changes in miRNA expression seen in cancer. These include transcriptional events regulated by transcription factors and DNA methylation, and also post-transcriptional events such as dysregulation of miRNA processing 150. The first study to reveal altered expression levels of miRNAs in CRC was published in 2003, where miR-143 and miR-145 were shown to have reduced levels in tumour tissue 151. Since then, altered expression levels of various miRNAs have been reported in CRC as well as in other malignancies 152-154.

Several different miRNAs have been suggested to regulate CRC development and progression. One of these is miR-21, with increased expression levels in CRC tumour tissue 155, 156. Another miRNA that is aberrantly expressed in CRC is miR-31. The expression of miR-31 has been associated with advanced tumours 157, 158. Both *KRAS* and *BRAF* mutations have been associated with miR-31 in CRC 159, 160, and it has been reported that miR-31 is the most up-regulated miRNA in *BRAF*-mutated cases 159. In a recent study in which miRNA expression was compared in tumours of different molecular subtypes, the majority of differences in miRNA expression were found between MSI and MSS tumours 161. However, our understanding of miRNA alterations in different subgroups of CRC is limited.
Aims of the thesis

General aim

The aim of this thesis was to investigate KRAS- and BRAF-mutated CRCs at the molecular level, in search of possible differences that may explain their different clinical and histopathological behaviour. A better molecular understanding of KRAS- and BRAF-mutated CRCs might lead to important advances in diagnosis, prediction and treatment.

Specific aims

Paper I

To investigate the expression of the stem cell marker SOX2 in CRC in relation to KRAS and BRAF mutation; the prognostic role of SOX2, and the regulation of SOX2 expression by KRAS and BRAF mutation.

Paper II

To investigate the functional role of SOX2 in CRC, and associations with epithelial to mesenchymal transition (EMT) or CSCs.

Paper III

To investigate the prognostic importance of Th1 lymphocyte infiltration in CRC, and a possible relation between Th1 lymphocyte infiltration and KRAS and BRAF mutation.

Paper IV

To investigate if KRAS- and BRAF-mutated CRCs have different miRNA signatures.
Materials and methods

CRC cell lines

In this thesis, we mostly used the colon cancer cell line Caco2. The Caco2 cell line has been classified as CIMP-negative and MSS, and is wild-type in both KRAS and BRAF, representing one of the largest subgroups of sporadic CRCs.

To be able to study molecular differences between KRAS- and BRAF-mutated CRCs, we introduced a mutation in KRAS or BRAF into the Caco2 cell line. We could thereby compare effects caused by the specific mutations in relation to each other, and in relation to the control cell line, which was wild-type in both genes. By introducing the different mutations into cells of the same genetic background, we could limit the influence of genotypic differences between cell lines.

We chose to create stable transfectants instead of using transient transfections. In this way, we were able to establish a more stable gene expression system that could be used for long-term experiments. The stable transfectants expressing mutant BRAF (Caco2-BRAFV600E) or mutant KRAS (Caco2-KRASG12V) were created by transfecting Caco2 cells with the plasmids pMCEF-BRAFV600E (a kind gift from Prof R. Marais) or pcDNA3-KRASG12V (a kind gift from Dr N. Ignatenko), containing inserts with the mutated genes. The BRAFV600E mutation was chosen since the vast majority of BRAF-mutated CRCs have this V600E substitution. The KRASG12V mutation was chosen since it is one of the most common KRAS mutations in CRC, and because it has a higher oncogenic potential than other KRAS mutations. The plasmids used contained the neomycin-resistance gene, allowing the use of G418 for selection of transfected cells. Only the cells that had successfully incorporated the plasmid DNA into their genome would survive in long-term culture with G418 treatment, resulting in selection and expansion of the transfected cells. To verify that the stable cell lines actually expressed mutant KRAS or mutant BRAF, we used the digital droplet PCR method to detect both wild-type and mutant sequences in the same reaction. Generation of the Caco2 cell line stably expressing SOX2 (Caco2-SOX2) was performed in the same way as the KRAS- and BRAF-mutated cell lines, but instead using the plasmid pcDNA3.3-SOX2, containing the SOX2 insert.
Patient cohort

To study associations between expression of mutated KRAS and BRAF and tumour characteristics, and also patient prognosis, we used the Colorectal Cancer in Umeå Study (CRUMS) \textsuperscript{164}. CRUMS involved primary CRCs that underwent tumour-resective surgery at Umeå University Hospital during the years 1995 and 2003. Both formalin-fixed paraffin-embedded (FFPE) tissue and fresh frozen tissue had been collected. One pathologist did all histopathological classifications, such as growth pattern, tumour type, and tumour grade (differentiation), by reviewing routinely stained tumour sections. Clinical data, such as tumour site and patient age and sex, were obtained from the patient records and a last update of survival data was done during the autumn of 2012.

KRAS mutations in the CRUMS cohort were detected by Sanger sequencing of codon 12 and 13, as described in a previous study \textsuperscript{65}. For detection of the BRAF\textsuperscript{V600E} mutation in the CRUMS cohort, DNA samples were subjected to allelic discrimination of single-nucleotide polymorphisms by the use of real-time PCR (Taqman allelic discrimination assay) \textsuperscript{165}.

In CRUMS, data were collected retrospectively since the patients were recruited after their diagnosis. This is a good study design when investigating the prognostic importance of different tumour features. A great advantage of CRUMS is that numerous clinicopathological and molecular tumour characteristics have been defined for all the patients included. Another advantage is that most of the tumours were not treated, or were treated only with basic chemotherapy or radiotherapy before surgery, while the treatment plans today are often more complex. Preoperative treatment can disrupt the tumour histology, thereby making it difficult to study, for example, the location of different cellular and molecular components. Possible selection of subclones of tumour cells that can survive the treatment is another problem when studying preoperatively treated tumours.

Gene expression analysis

For gene expression analysis in cell lines and tissue samples, semi-quantitative real-time PCR was used. The polymerase chain reaction (PCR) amplifies a particular region of DNA, generating thousands or even millions of copies, making it possible to detect and quantify gene expression even with small amounts of input RNA. By using reverse transcription, RNA is reverse-transcribed into complementary DNA (cDNA) before the PCR reaction. The PCR reactions consist of cycles of temperature changes, where the double-stranded DNA is first denatured, followed by primer annealing to
the single-stranded DNA and extension. Compared to conventional PCR, 
real-time PCR measures the amplified DNA at each PCR cycle, instead of 
relying on an endpoint measurement. This allows measurements during the 
exponential phase of the amplification, where doubling of product occurs 
every cycle, giving a more precise measurement than with conventional PCR. 
To monitor the PCR reaction in real time, fluorescence has to be generated 
throughout the reaction and measured at the end of each cycle. We used 
SYBR green, which is one of the most commonly used fluorescence 
technologies in real-time PCR. SYBR green is a dye which fluorescence 
dramatically increases when it binds to double-stranded DNA. So when more 
DNA is amplified, more SYBR green is bound and more fluorescence is 
generated. Real-time PCR is a very simple and effective method for gene 
expression analysis.

Total RNA was isolated from cell lines, FFPE samples, or fresh frozen 
tissue samples, and then reverse-transcribed into cDNA. Primers 
complementary to the DNA target sequences of interest were used. The level 
of gene expression was quantified in relation to reference genes, which have 
been shown to have a relatively constant expression in the cell lines or 
tissues studied. For this, the $2^{-\Delta\Delta CT}$ method was used. When analysing 
RNA expression, GAPDH was used for normalisation in cultured cells and 
RPL13A was used for fresh frozen tumour specimens. When analysing 
miRNA expression in cultured cells and FFPE specimens, RNU6-2 was used 
for normalisation.

**Protein expression analysis/IHC**

Immunohistochemistry (IHC) is a widely used method for detecting antigens 
in tissue samples, both in the clinic and in research. This technique uses 
antibodies that specifically bind to the antigen of interest. The antibody-
antigen interaction is visualised by use of a secondary antibody conjugated to 
an enzyme that converts a chromogen into a coloured precipitate. The 
advantages of IHC are that it allows visualisation of the location of the 
antigen of interest in the tissue and that it is a fast, inexpensive, and simple 
method. One disadvantage of IHC is the risk of non-specific staining. This 
can be prevented by using different controls, such as positive and negative 
tissue controls that are known to express or not to express the antigen of 
interest. The specificity of the secondary antibody can also be tested by 
leaving out the primary antibody.

FFPE CRC specimens were used for IHC staining in the work described in 
this thesis. Before IHC staining, the tumour tissue samples were cut in 4-μm 
sections and then dried before deparaffinization and rehydration. All IHC
staining was performed with a Ventana Benchmark Ultra staining machine and the iVIEW DAB Detection Kit was used for visualisation. In order to visualise the tissue structures, all tissue sections used were also counterstained with haematoxylin.

Evaluation of IHC staining was performed under a light microscope, where the same observer evaluated each tumour twice under the supervision of an experienced pathologist. In cases of discord, a third, final evaluation was done. The expression of SOX2 was evaluated as positive or negative staining in the cell nuclei. Cytoplasmic or stromal staining, seen only in a few cases, was disregarded since SOX2 is a transcription factor with nuclear functions. The expression of the transcription factor CDX2 was also evaluated in the cell nuclei and scored as: 5% positive tumour cells, 5–50% positive tumour cells, or > 50% positive tumour cells. The normal mucosa, included in most of the tissue sections, was used as positive control since CDX2 is expressed in intestinal epithelial cells. T-bet expression was evaluated at different subsites within the tumour: the invasive tumour front, the centre of the tumour, and within the tumour epithelium (intraepithelial expression), as described in one of our previous papers. The T-bet staining was scored numerically as follows: no/sporadic (1), moderate (2), abundant (3), or highly abundant (4), and then a total score was calculated by adding together the expression at the different tumour subsites. A total score of 3–4 represented low expression; 5–6 represented moderate expression; and 7–12 represented abundant expression.
Results and discussion

Paper I

**SOX2 is regulated by BRAF and contributes to poor patient prognosis in colorectal cancer**

In this study, we investigated the expression of the stem cell marker SOX2 in the CRUMS cohort of CRC patients. The expression of SOX2 was evaluated by IHC in 441 CRC specimens, and assessed as positive or negative nuclear staining. Forty-seven (10.7%) of the CRC specimens were scored as positive for SOX2. SOX2 was expressed in tumour tissue and not in adjacent normal mucosa, which was included in most of the tumour samples. These findings are in line with previous studies on CRC 92, 169, 170. Aberrant expression of SOX2 has been shown in other cancer types as well, such as breast 96, pancreas 95, and gastric cancer 171, indicating that it is a common phenomenon in malignant tissue.

In the CRUMS cohort, expression of SOX2 was associated with high tumour grade, high tumour stage, and *BRAF* mutation. Interestingly, even though *KRAS* is an upstream activator of *BRAF*, no correlation between SOX2 expression and *KRAS* mutations was found. Since SOX2 is a stem cell marker, the correlation with high-grade (poorly differentiated) tumours was not surprising; it has been reported previously in several cancers 172-176. Furthermore, *BRAF*-mutated CRCs are more often poorly differentiated 67. The association between SOX2 and *BRAF* mutation may be one explanation for this.

We also found that SOX2 expression was associated with a poor patient prognosis, which is in accordance with studies of several other cancer types 177-179. In CRC, an association between SOX2 expression and increased metastasis has been reported, which could explain part of the poor patient prognosis seen in patients with SOX2-expressing tumours 91, 169. Since SOX2 is a stem cell factor, the correlation with metastases in CRC is interesting. Due to their potential for unlimited division, and the fact that they are more likely to survive at a distant site than differentiated cells, stem cells tend to have metastatic capacity 180. In this study, we also explored SOX2 expression in a small patient cohort (n = 13) with tissue specimens from both primary tumours and corresponding metastases. We did indeed find that patients with SOX2-positive primary tumours often had SOX2-positive metastases, indicating that SOX2-positive cells are more likely to migrate. Even though so few cases were studied, it strengthens previous findings. Interestingly, we
found that the poor prognosis seen in patients with SOX2-expressing tumours was restricted to BRAF-mutated CRCs. No association between SOX2 and patient survival was found in BRAF wild-type CRC cases, indicating that BRAF-mutated SOX2-expressing tumours represent a subgroup with an especially poor patient prognosis.

Using in vitro cell culture studies, we were able to explore the association between SOX2 and BRAF mutation further. BRAF-mutated CRC cells (Caco2-BRAF\(^{V600E}\)) showed increased expression of SOX2 compared to KRAS-mutated (Caco2-KRAS\(^{G12V}\)) and KRAS/BRAF wild-type (Caco2) CRC cells. By blocking BRAF downstream signalling with a MEK inhibitor, the expression of SOX2 was reduced, indicating that BRAF regulates the expression of SOX2. This finding also demonstrated that the regulation of SOX2 by BRAF is at least partly mediated through the MEK-activating activity of BRAF. KRAS and BRAF are both involved in the MAPK signalling pathway, leading to activation of MEK. Since no association between KRAS mutation and SOX2 expression was found, either in the tumour tissue samples or in the CRC cell lines, this suggests that additional BRAF downstream signalling events may also regulate SOX2. To our knowledge, this is the first study to reveal a correlation between BRAF mutation and SOX2 expression in CRC, demonstrating that SOX2 is at least partly regulated by BRAF.

**Paper II**

**SOX2 expression is associated with a cancer stem cell state and down-regulation of CDX2 in colorectal cancer**

In this study, we investigated the functional role of SOX2 in CRC. Some studies have suggested that SOX2 expression is associated with epithelial to mesenchymal transition (EMT) and increased migration of tumour cells \(^{91,92}\). EMT is a process whereby epithelial cells acquire a mesenchymal phenotype, a process that facilitates cell motility and invasion. However, when analysing our cancer cell lines (Caco2 and Caco2-SOX2), increased SOX2 expression was not found to be associated with an EMT profile. Even though the expression of epithelial marker E-cadherin was reduced by SOX2, expression of mesenchymal markers such as fibronectin and vimentin was not elevated—but instead significantly reduced—in Caco2-SOX2 cells. The expression of different matrix metalloproteinases (MMPs) was also found to be down-regulated in Caco2-SOX2 cells. Since MMPs facilitate tumour cell migration \(^{181}\), this would suggest that SOX2-expressing cells would have a
decreased migratory activity, which we were also able to demonstrate using a cell migration assay. These findings suggest that SOX2 expression does not regulate EMT or migration, at least not in our cell model. Further studies are required to explain the contradictory effects of SOX2 on EMT and cellular migration in CRC.

We found that Caco2-SOX2 cells had a lower proliferative rate, were less adherent, and showed a spheroid morphology compared to Caco2 cells. Cells express several different adhesion molecules required for cell-to-cell contact, and contact between cells and the ECM. Cells expressing SOX2 were found to show reduced expression of adhesion molecules, such as several different integrins, which would explain their less adherent growth pattern.

CSCs are known to share characteristics usually associated with normal stem cells, such as a quiescent slow-cycling phenotype and anchorage-independent growth. In a recent study, it was shown that a subpopulation of cultured colon cancer cells with low adherence displayed stemness properties, such as higher self-renewal ability. Interestingly, this subpopulation of cells also had a lower proliferative rate and a high tumorigenic and metastatic ability. These findings, in combination with the fact that SOX2 is a stem cell factor, indicate that cancer cells expressing SOX2 might represent a subpopulation of CSCs. To further verify a possible relation between SOX2 and CSCs, we analysed the expression of other CSC markers in our cell model. The expression of CD44 and CD24, which have been linked to colon CSCs, was found to be significantly elevated in Caco2-SOX2 cells. In a previous study on CRC, CD24 has been suggested to be regulated by SOX2. The prognostic value of cancer cells with a stem-cell-like phenotype has been evaluated to some extent in CRC. Elevated expression of stem-cell-like factors, including SOX2 and CD44, has been associated with a significantly worse prognosis. In models of skin squamous-cell carcinoma, deletion of SOX2 was shown to result in reduced tumour initiating potential and tumour regression, and SOX2 has been suggested as a marker and also a driver of cancer stemness. Together, these results indicate that the poor prognosis seen in SOX2-expressing CRCs might in part be explained by acquired cancer stemness properties, favouring tumour progression.

In paper I, we found that SOX2 was associated with poorly differentiated tumours. Decreased expression of the intestinal epithelial marker CDX2 has previously been linked to poorly differentiated CRCs. We therefore investigated the relation between SOX2 and CDX2 using our in vitro cell culture models. CRC cell lines expressing SOX2 were indeed found to have significantly reduced levels of expression of CDX2. The reduced CDX2
expression could be verified using fresh frozen tumour tissue from 25 CRC patients.

Next, we evaluated CDX2 expression in the CRUMS patient cohort. CDX2 expression was investigated by IHC in tumour tissue sections from 445 patients; 43.4% of the patients were found to have CDX2 expression in less than 50% of the tumour cells, and 14.4% showed CDX2 expression in less than 5% of the tumour cells. Decreased expression of CDX2 was more often found in poorly differentiated tumours and in tumours expressing SOX2. Reduced expression of CDX2 was also associated with a higher tumour stage and right-sided tumours, findings that have been reported from other studies as well 107. BRAF mutation, CIMP-high, and MSI were molecular characteristics associated with down-regulated expression of CDX2, which is in line with results from previous studies 107. These associations of CDX2 with SOX2 expression and BRAF mutation are also in line with the results from paper I, where BRAF mutation was found to regulate SOX2 expression. Furthermore, down-regulated expression of CDX2 was significantly associated with poor patient survival, as also seen in other reports on CRC 104, 105, 108.

SOX2-expressing tumours with reduced expression of CDX2 were correlated to poorer patient survival compared to SOX2-expressing tumours with retained CDX2 expression. This suggests that part of the poor prognosis seen in SOX2-positive tumours might be through SOX2-mediated down-regulation of CDX2. However, several of the SOX2-negative tumours were also found to have reduced expression of CDX2, and also here CDX2 down-regulation was associated with poor patient survival. These findings suggest that CDX2 can also be down-regulated by SOX2-independent mechanisms. SOX2 is only one potential CSC factor in CRC. It may be that other factors linked to cellular stemness can also regulate CDX2 expression. Further studies are required to elucidate the link between cellular stemness and CDX2.

SOX2 and CDX2 have both been associated with poorly differentiated tumours, a feature shared by CSCs. By combining the evaluation of expression of SOX2 and CDX2, it has been found to be possible to segregate patients with gastric cancer into different prognostic groups 187. Patients with tumours expressing SOX2, but not CDX2, had the worst outcome. In colon, it has recently been suggested that CDX2 expression is down-regulated by SOX2, but that SOX2 requires induction of SOX21 to repress CDX2 188. In CRC, it has also been suggested that CDX2 is a target of the Hippo pathway 189. The Hippo pathway has important roles in tumour suppression by regulating cell proliferation, growth, and apoptosis. When deregulated, the pathway is instead involved in tumour initiation and progression 190, 191.
Hippo pathway has been shown to regulate the inactivation of the oncogenic YAP protein. It has also been reported that SOX2, through increased expression of YAP1, disrupts the Hippo pathway in order to maintain cellular stemness\textsuperscript{192, 193}. Therefore, the link between SOX2 expression and reduced CDX2 expression can either be through increased expression of SOX2\textsuperscript{1} or through deregulation of the Hippo pathway, by increased expression of YAP1.

**Paper III**

**The infiltration of, and prognostic importance of Th1 lymphocytes vary in molecular subgroups of colorectal cancer**

In this study, we investigated the infiltration of Th1 lymphocytes in molecular subgroups of CRC. Immune cell infiltration is a strong prognostic factor in CRC\textsuperscript{134, 135}. The complex interplay between cancer cells and immunity can lead to both promotion and rejection of tumour growth, and it is likely that the balance between different subsets of immune cells determines the outcome (Figure 6).

The infiltration of Th1 lymphocytes was investigated in the CRUMS cohort of CRC patients. The Th1 lymphocyte marker T-bet was semi-quantitatively evaluated by IHC in 418 CRC tumour specimens. Sixty-five percent of the CRC patients showed modest to massive infiltration of Th1 cells, while the other 35% showed weak or no infiltration. Infiltration of Th1 cells was more often found in right-sided colon tumours. Furthermore, a strong inverse association was seen with tumour stage. In the CRUMS cohort, an increased infiltration of Th1 cells was significantly associated with a better patient prognosis, which is in line with findings from other studies\textsuperscript{135, 137}.

We also investigated the Th1 lymphocyte infiltration in relation to molecular subgroups, *KRAS* and *BRAF* mutation, MSI and CIMP status. Increased infiltration of Th1 cells was found in *BRAF*-mutated tumours, while *KRAS*-mutated tumours were less infiltrated. In addition, Th1 infiltration was more frequent in MSI and CIMP-high tumours. An increased density of Th1 cells in MSI tumours has previously been shown by Boissiere-Michot et al.\textsuperscript{139}. We found that patients with MSI tumours that were highly infiltrated by Th1 lymphocytes had a favourable prognosis. Furthermore, MSI tumours with increased infiltration of Th1 cells were also *BRAF*-mutated to a significantly higher extent. In contrast, patients with *BRAF*-mutated tumours poorly infiltrated by Th1 lymphocytes were found to have an especially poor prognosis. A close association between MSI and *BRAF*
mutation has been established, and MSI CRCs have been shown to be associated with a better patient prognosis. However, we found in a previous study that BRAF mutation could be of prognostic importance within the MSI subgroup, suggesting that the improved prognosis found with Th1 infiltration in this study was not only dependent on MSI.

To study KRAS and BRAF mutations further in relation to Th1 lymphocyte infiltration, we used our in vitro cell culture system of CRC cells stably expressing KRAS or BRAF mutations. Cytokines and chemokines, secreted by immune cells as well as by tumour cells, can recruit and stimulate different subsets of immune cells. We therefore investigated whether KRAS- or BRAF-mutated CRC cell lines had different patterns of cytokine and chemokine expression, stimulating the recruitment and polarisation of different subsets of T lymphocytes. We found that expression of the Th1-attracting chemokine CXCL10 was increased in Caco2-BRAFV600E cells compared to both Caco2 and Caco2-KRASG12V cells. Caco2-BRAFV600E cells also expressed reduced levels of the chemokine CCL22 and of the cytokine TGFβ, compared to both Caco2 and Caco2-KRASG12V cells. CCL22 and TGFβ are both known to stimulate the Th2/Treg axis. CXCL10 expression was also found to be significantly up-regulated in BRAF-mutated tumour tissue specimens compared to KRAS-mutated tumours.

Gene expression levels do not necessarily correlate with protein expression levels. We therefore also analysed proteins secreted by the in vitro cultured cells, using a multiplex immunoassay (Proseek, Olink). Conditioned media from Caco2, Caco2-BRAFV600E, and Caco2-KRASG12V cells were analysed with the Oncology I panel, which included both CXCL10 and TGFβ. Caco2-BRAFV600E secreted higher levels of CXCL10 than Caco2 and Caco2-KRASG12V cells, while Caco2-KRASG12V cells secreted higher levels of TGFβ (unpublished observation). These results both verify and strengthen previous findings, since they confirm that increased gene expression also results in increased protein translation as well as secretion. The cell lines used in this study were MMR-proficient, indicating that the results from the in vitro studies were independent of MSI.

In contrast to our study, a previous study using different CRC cell lines has shown that wild-type cells express higher levels of CXCL10 than KRAS- and BRAF-mutated cells. However, the authors investigated KRAS and BRAF mutation in cell lines of different genetic backgrounds, which could have influenced the results. In another study, CRCs with an MSI phenotype were found to express higher levels of CXCL10, findings that are in line with our results. High expression of genes coding for the chemokines CX3CL1, CXCL9, and CXCL10 have previously been found to be associated with infiltration of Th1 lymphocytes and with a good patient prognosis.
contrast, *KRAS* mutation has previously been shown to up-regulate the expression of VEGF and TGFβ, stimulating angiogenesis and promoting tumour growth 196. *KRAS*-mutated CRCs have also been found to express reduced levels of genes related to the function of Th1 lymphocytes, including CXCL10 141. These results together support the idea that *KRAS* and *BRAF* mutation may stimulate different axes of the immune response in CRC, which might in part explain the different clinical and histopathological behaviour of *KRAS*-mutated tumours and *BRAF*-mutated tumours.

**Paper IV**

**microRNA expression in *KRAS*- and *BRAF*-mutated colorectal cancers**

In the last study, we investigated whether *KRAS*- and *BRAF*-mutated CRCs express different miRNA signatures. miRNAs can post-transcriptionally regulate the expression of many important genes, including those of transcription factors, and are highly dysregulated in CRC 197 as well as in numerous other cancer types 153. Differences in miRNA expression might therefore partly explain the different clinical and histopathological behaviour of *KRAS*- and *BRAF*-mutated CRCs.

miRNA expression profiles have previously been studied in tissue specimens of *KRAS*- and *BRAF*-mutated CRCs 160, 198, 199. An miRNA expression signature differentiating *KRAS*-mutated tumours from *KRAS* wild-type tumours has been reported 198, and also an miRNA signature associated with *BRAF*-mutated tumours but not *KRAS*-mutated tumours 199. In this study, we focused on investigating differences between *KRAS*- and *BRAF*-mutated tumours and *KRAS/BRAF* wild-type tumours, using mutated cell lines of the same genetic background and also tumour tissue specimens.

We investigated miRNAs differentially expressed in *KRAS*- and *BRAF*-mutated CRC cell lines (relative to *KRAS/BRAF* wild-type cells) using an miRNA PCR array that included 84 different miRNAs known to be aberrantly expressed in cancer. Nine miRNAs that were found to be expressed at significantly different levels in the different CRC cell lines, were chosen for further investigations (let-7a, let-7i, miR-10a, miR-10b, miR-100, miR-181a, miR-181b, miR-372, and miR-373). We also investigated the *in vitro* expression of miR-31 in our cell lines, since it has been shown by others to be up-regulated in *BRAF*-mutated CRCs 159, 160, 199.

We found that expression of miR-31 was significantly up-regulated in the *BRAF*-mutated cell line and also in *BRAF*-mutated tumour tissue samples.
The tumour tissue samples with KRAS mutations also showed a tendency to have increased expression of miR-31 relative to the wild-type tumours, but this difference was smaller than for the BRAF-mutated tumours. The KRAS-mutated cell line expressed lower levels of miR-31 than the BRAF-mutated cells, indicating that miR-31 is more highly associated with BRAF mutation, in agreement with the results of other studies \(^{159, 160, 199}\). In CRC, elevated expression of miR-31 has been found to be associated with an aggressive phenotype and also with poor patient survival \(^{159, 200}\). Studies have also shown that miR-31 promotes the proliferation of colon cancer cells \(^{200, 201}\). Furthermore, a correlation has been found between an increased invasion and metastasis rate and up-regulated expression of miR-31 in CRC \(^{200}\), explaining the poor prognosis seen in patients with tumours that express high levels of miR-31.

miR-100 was found to be differentially expressed in KRAS- and BRAF-mutated cells in the PCR array, but this could not be verified in tumour tissues. In our study, KRAS- and BRAF-mutated specimens showed a slightly up-regulated expression of miR-100 compared to wild-type tumours, a difference with borderline significance for KRAS-mutated CRCs. In CRC, expression of miR-100 has previously been found to be lower in tumours than in normal tissues \(^{202}\), and also lower in tumours with lymph node metastases than in tumours with no metastases \(^{203}\). In addition, since miR-100 has been shown to inhibit the proliferation and migration of CRC cells, miR-100 has been suggested to be a tumour suppressor in CRC \(^{202}\).

miR-372 expression and miR-373 expression were both down-regulated in KRAS- and BRAF-mutated cell lines relative to the wild-type cell line. In tumour tissues, the same tendencies were found, with miR-373 being significantly less expressed in BRAF-mutated tumours than in wild-type tumours. But once again, no clear difference was found between KRAS- and BRAF-mutated tumours.

The remaining six miRNAs, let-7a, let-7i, miR-10a, miR-10b, miR-181a, and miR-181b, did not show clear differences between the groups when analysed in tumour tissue.

Our results indicate that KRAS- and BRAF-mutated tumours may have different miRNA signatures than KRAS/BRAF wild-type tumours. However, we were unable to find miRNA signatures that could distinguish between KRAS- and BRAF-mutated CRCs. KRAS and BRAF are both involved in the MAPK signalling pathway, and mutations of the genes result in constitutive activation of the signalling pathway. An increased MAPK signalling, irrespective of how it is activated, may therefore be associated with a specific miRNA signature in CRC. Thus, KRAS- and BRAF-mutated CRCs might be considered as one subgroup, with increased MAPK signalling, when
investigating miRNA profiles. However, since we only analysed a proportion of all the miRNAs present in human cells, further investigation of additional miRNAs is required to support this conclusion.
Conclusions

The studies in this thesis have revealed both differences and similarities between KRAS- and BRAF-mutated CRCs at the molecular level.

BRAF-mutated CRCs were found to express the stem cell factor SOX2, which was correlated with an especially poor patient prognosis. SOX2 was also shown to be at least partly regulated by BRAF in vitro.

CRC cells expressing SOX2 were found to have properties of CSCs, which may explain the poor patient prognosis seen in SOX2-expressing tumours. SOX2-expressing cells also had reduced expression of the intestinal epithelial marker CDX2. Decreased expression of CDX2 was associated with poor patient survival in CRC. Furthermore, SOX2-expressing tumours had a poorer prognosis when CDX2 was down-regulated, suggesting that CDX2 down-regulation might explain part of the negative effect of SOX2 on prognosis.

Taken together, these studies identified a subgroup of CRC patients with a particularly poor prognosis, carrying tumours with BRAF mutation and a CSC-like appearance. Identification of poor prognostic subgroups of CRC is of considerable importance, since these patients would probably need a more robust treatment plan. The findings from our studies, and from others, suggest that therapeutic approaches aimed at targeting CSCs may be of value for these patients. Further studies are needed to explain the role of cancer stemness in tumour progression and in prognosis in patients with CRC.

We found a correlation between increased infiltration of Th1 lymphocytes and better patient prognosis in CRC. In addition, BRAF-mutated tumours were found to have increased infiltration of Th1 cells, while KRAS-mutated tumours had a decreased infiltration. Interestingly, BRAF-mutated CRCs were found to express increased levels of the Th1-attracting chemokine CXCL10 and decreased levels of CCL22 and TGFβ, stimulating the Th2/Treg axis, compared to KRAS-mutated CRCs. These results suggest that the difference in Th1 infiltration rate between KRAS- and BRAF-mutated CRCs is probably partly caused by differences in chemokine and cytokine secretion, stimulating different arms of the immune response. This study identified different molecular subgroups of CRC with different patient prognosis, depending on Th1 lymphocyte infiltration. Poorly infiltrated BRAF-mutated tumours were found to have an especially poor prognosis, and an extended treatment plan should be considered for this subgroup of patients.
Using CRC cell lines, KRAS and BRAF mutations appeared to differentially affect the expression of miRNAs, but this could not be verified in tumour tissues. Even though expression of miRNAs could not help to distinguish between KRAS- and BRAF-mutated tumours, KRAS- and BRAF-mutated CRCs had different miRNA profiles from those of KRAS/BRAF wild-type CRCs. These results suggest that KRAS and BRAF mutations, both of which result in increased MAPK signalling, do not differentially regulate miRNA expression in CRC.

In summary, these studies together reveal possible molecular differences and similarities between KRAS- and BRAF-mutated CRCs that may explain part of their different clinical and histopathological behaviour. Our findings may contribute to future advancements in diagnosis, prediction and therapy of patients with KRAS- and BRAF-mutated CRCs.
Acknowledgements

Till sist vill jag tacka alla personer som har hjälpt och stöttat mig under dessa år. Utan er skulle inte den här boken finnas!

Min huvudhandledare Richard Palmqvist. Tack för den här tiden och tack för att du har trott på mig. Du har en fantastisk förmåga att få alla att känna sig välkomna och delaktiga, vilket har gjort att jag har trivts i gruppen från dag ett. Trots din tidsoptimism har ju det här gått galant! 😊

Min biträdande handledare (och före detta huvudhandledare) Maria Wikberg. Tack för allt du har lärt mig, framförallt all laborativ kunskap du har gett mig! Jag uppskattar att du har tagit dig tid att förklara allt och lärt mig från grunden, man ska inte alltid använda massa färdiga kit, eller hur?! Tack också för dina peppande ord och din positivitet, det har behövts mer än du tror 😊.

Min biträdande handledare Sofia Edin, tack för allt du har lärt mig under dessa år! Det är alltid lika roligt att spåna idéer tillsammans med dig. Du är grymt duktig på att skriva och du har lärt mig mycket på den fronten, utan alla dina ”petiga” kommentarer (som du själv ibland tycker har varit för mycket) skulle jag inte ha lärt mig så mycket som jag har gjort 😊. Tack även för allt stöd du har gett mig, du har alltid funnits där för mig och vi har kunnat prata om allt möjligt, stort som smått, vilket har betytt mycket! Anna Löfgren-Burström, du har också varit ett fantastiskt stöd och har alltid sett efter mig, lite som en extra-mamma (på ett bra sätt!). Till erat rum har dörren alltid stått öppen, tack! 😊

Mina rumskamrater Mona Shirdel, Björn Gylling och Agnes Ling. Tack för en väldigt rolig tid, och trots att det stundtals har varit väldigt trångt i vårat rum har humöret ändå hållits uppe. Har nog aldrig ätit så mycket kakor och druckit så mycket kaffe som under dessa år 😊.

Resten av CRC-gruppen (nuvarande och tidigare medlemmar), och framförallt Carl Zingmark, Xingru Li, Robin Myte, Roger Stenling, Vincy Eklöf, Bethany van Guelpen, Kerstin Näslund och Åsa Stenberg. Tack för allt ni har hjälpt mig med och tack för alla roliga stunder vi har haft, ni är ett fantastiskt gäng!

Karin Bodén, Terry Persson, Carina Ahlgren, Clas Wikström, Lennart Österman och Åsa Lundsten, tack för all administrativ hjälp och för allt jobb ni gör!
Elin Forsgren och Erik Bovinder Ylitalo. Det är snart 10 år sedan vi började på universitetet tillsammans, och tänk att vi alla tre snart har disputerat, helt galet!! Tack för er vänskap och för allt stöd ni har gett mig 😊.

Till ”fika-gänget” på 2 trappor och ”doktorand/lunch-gänget”, ibland behöver man tänka på annat än jobb, tack för alla roliga diskussioner om allt mellan himmel och jord ☺.

Till alla er utanför jobbet som alltid är där och stöttar mig och förgyller mitt liv:

Mina allra bästa vänner Ida, Bea och Emelie. Tack för att ni alltid finns där, i vätt och torrt. Att det är över 20 år sedan vi lärde känna varandra tror jag förklarar hur mycket vi har varit med om tillsammans 😊, finns få personer som känner mig så bra. Jag älskar er♡. Dessutom har jag fått tre andra fina vänner på köpet, Joel, Andreas och Mattis, och även två nya tillskott i gänget, Benjamin och Nellie, bästaste barnen ☺. Tack till er alla!

Till min underbara mamma och pappa, tack för att ni alltid tror på mig och för allt ni har lärt mig genom livet. Tack för att ni har gjort mig till den jag är, jag älskar er♡. Tack också till resten av min fantastiska familj och släkt (storebror, Mormor, farmor och alla ni andra).

Till min andra familj, framförallt till Stina, Staffan och Iris. Tack för att ni alltid ställer upp, ni är bland de mest omtänksamma människor jag känner ☺.


/ Ida
References


64. Li, W.Q. et al. BRAF mutations are associated with distinctive clinical, pathological and molecular features of colorectal cancer independently of microsatellite instability status. *Mol Cancer* 5, 2 (2006).


38


140. Tougeron, D. et al. Tumor-infiltrating lymphocytes in colorectal cancers with microsatellite instability are correlated with the number and spectrum of frameshift mutations. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc* 22, 1186-1195 (2009).


158. Slaby, O. et al. Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer. *Oncology* 72, 397-402 (2007).


