Biomarker mRNAs for staging and prognosis of colorectal cancer

Lina Ohlsson

Department of Clinical Microbiology, Immunology
Umeå
2011

Responsible publisher under swedish law: the Dean of the Medical Faculty
This work is protected by the Swedish Copyright Legislation (Act 1960:729)
ISBN: 978-91-7459-318-1
ISSN: 0346-6612-1458
Front cover: Arkitektkopia. Cover design: Connie Eliasson. Photo: Peter Andersson.
Elektronisk version tillgänglig på http://umu.diva-portal.org/
Printed by: Arkitektkopia
Umeå, Sverige 2011

To my parents!

TABLE OF CONTENTS

ABSTRACT

Mesenteric lymph node (ln) metastasis is the single most important prognostic characteristic in colorectal cancer (CRC). The ln status is used for staging and is a decisive selection criterion for postoperative adjuvant therapy. However, it is difficult to accurately determine ln status by routine histopathology (H&E). Thus, ~25% of CRC patients, who by H&E are considered to lack tumor cells in their lns, i.e. stage I+II, die from CRC.

To explore the utility of biomarker mRNA analysis for staging and prognosis of CRC, lns were collected at surgery and mRNA levels for fourteen biomarkers, including carcinoembryonic antigen (CEA), kallikrein 6 (KLK6), cytokeratin 20 (CK20), guanylyl cyclase C (GCC), CEACAM1-S, CEACAM6 and mucin 2 (MUC2), were determined by quantitative RT-PCR with RNA copy standards. Results were compared to routine H&E analysis.

The biomarkers were analyzed for capacity to detect disseminated tumor cells in lns. mRNA levels were determined in CRC- and control lns, primary tumor, normal colon, immune cells and fibroblasts. Lack of expression in immune cells and fibroblasts and high and homogenous expression in primary tumors showed to be the determining factors. CEA fulfilled these criteria best, followed by KLK6, CK20, GCC, and MUC2.

Utility of the biomarker mRNAs for staging and prognosis was examined in 174 CRC patients. CEA was the best predictor of disease-free survival time after surgery with a 71 months difference between CEA(+) and CEA(-) patients and a hazard ratio of 5.1 for risk of recurrence for CEA(+) patients. CEA, CK20 and MUC2 were more sensitive than H&E in that these biomarkers identified patients who succumbed from recurrent CRC although H&E analysis had failed to detect the disseminated tumor cells. Combined analysis of CEA and MUC2 mRNAs improved prediction of outcome. Patients with high risk for recurrence had low MUC2/CEA ratios.

KLK6 mRNA was identified as a potential progression marker by genome-wide microarray analysis of gene expression. It was found to be ectopically expressed in CRC tumor cells. KLK6(+) lns was an indicator of poor prognosis (hazard ratio 3.7). Notably, the actual level was of importance for outcome. The higher the KLK6 mRNA levels the greater the risk of recurrence. At the 90^{th} percentile the hazard risk ratio for KLK6(+) patients was 5.6. KLK6 positivity in lns with low numbers of tumor cells, as indicated by low CEA mRNA levels, indicated poor prognosis (hazard ratio 2.8). Thus, KLK6 adds prognostic information to CEA analysis.

Increased levels of mRNA for the proinflammatory cytokine interferon-γ and the down-regulatory cytokine interleukin-10 in lns of CRC patients suggested ongoing immune reactions against the infiltrating tumor cells. Elevated TGF-β1 levels correlated weakly with survival, suggesting protection by the antiproliferative effect of TGF-β1 in sporadic cases.

CEA mRNA was the best single biomarker for staging and prediction of disease-free survival time and risk of recurrence after surgery. In addition to CEA, KLK6 positivity and low MUC2/CEA ratio correlate with poor prognosis. Thus, CEA, MUC2 and KLK6 mRNAs form a strong “trio” for staging and prediction of outcome for CRC patients.

POPULÄRVETENSKAPLIG SAMMANFATTNING PÅ SVENSKA

Användbarhet av biomarkör-mRNA for stadieindelning och prognos vid kolorektalcancer

Cancer i tjock- och ändtarmen (även kallad kolorektalcancer) är ett stort folkhälsoproblem, särskilt i västvärlden. Kolorektalcancer är den tredje vanligaste cancerformen efter bröst- och prostatacancer. Årligen insjuknar mer än 1 miljon människor i världen, varav ungefär hälften avlider i sjukdomen. I Sverige insjuknar ungefär 6000 personer varje år, varav 2600 dör av sjukdomen.

Kirurgi är i dagsläget den viktigaste behandlingsformen. Vid ett kirurgiskt ingrepp avlägsnas all synlig tumörvävnad i tarmen samt dränerande lymfkörtlar. I samband med att tumören tas bort görs även en mikroskopisk undersökning där tumörens utbredning i tarmväggen, spridningen till lymfkörtlar och andra organ undersöks. Patienter med tumörspridning till lymfkörtlar är en högriskgrupp och ungefär hälften av patienterna i denna grupp får återfall i sjukdomen. Högriskpatienter erbjuds därför tilläggsbehandling i form av strålterapi eller cytostatikabehandling, vilka har visat sig minska återfall med en tredjedel. Att hitta tumörceller i lymfkörtlar är därför mycket viktigt och avgörande för bedömning av patientens återfallsrisk och därmed val av efterbehandling och uppföljningsstrategi. Rutinmässigt används Hematoxylin och Eosin (H&E) infärgning av snitt från lymfkörtlar. Bedömningen av vävnadssnitt kräver en tränad patolog. Metoden är svår, arbetskrävande och inte tillräckligt känslig.

Målet med detta avhandlingarbete var att försöka hitta en bättre metod att detektera spridda tumörceller i lymfkörtlar från patienter med kolorektalcancer.

Kvantitativ realtids RT-PCR är en molekylärbiologisk teknik som kan detektera mycket små mängder av budbärar RNA för specifika gener (biomarkör-mRNA). Genom att mäta mängden av budbärar RNA för gener, som uttrycks av tumörceller, erhålls ett kvantitativt mått på mängden av, och med vissa budbärar RNA typen av, tumörceller i lymfkörtlar (biomarköranalys). Vid operationer samlades lymfkörtlar som ligger i tumörområdet. Dessa delades på mitten, varav ena halvan undersöktes enligt rutinmässiga principer med H&E-infärgning och den andra halvan undersöks med biomarköranalys.

Vi har undersökt 14 olika biomarkörer med avseende på dessas förmåga att upptäcka spridda tumörceller hos kolorektalcancer-patienter. Ett stort kliniskt material bestående av 160 kolorektalcancer-patienter och 20 kontrollpatienter (totalt över 600 lymfkörtlar) har analyserats. Tre av

biomarkörerna gav mycket lovande resultat: carcinoembryonalt antigen (CEA), kallikrein 6 (KLK6) och mucin2 (MUC2). CEA visade sig detektera tumörceller i lymfkörtlar med hög känslighet och vara viktig för stadieindelning och prediktion av prognos. Medelöverlevnad för patienter med CEA positiva lymfkörtlar [CEA(+)] var 44 månader jämfört med 115 månader för CEA(-) patienter, vilket gav en skillnad i överlevnad med nästan 6 år (71 månader). Risk för återfall i sjukdom var 5 gånger större för CEA(+) patienter än för CEA(-) patienter. KLK6 visade sig vara ett mått på tumörcellernas aggressivitet. Patienter med KLK6(+) lymfkörtlar hade sämre prognos jämfört med KLK6(-) patienter. Även nivån KLK6 visade sig vara av stor betydelse för patienternas prognos. Höga nivåer MUC2 per tumörcell visade sig istället ha en skyddande effekt och associerades med god prognos. MUC2 i relation till antalet tumörceller (CEA-värde) och nivån av aggressivitet (KLK6) gav därmed ytterligare viktig prognostisk information.

Biomarkörerna CEA, KLK6 och MUC2 kan tillsammans detektera tumörceller som spridit sig till lymfkörteln med hög känslighet, och prediktera patienternas prognos. Om patienter får korrekt prognos har de större chans att få rätt behandling och chansen för överlevnad ökar. Kvantitativ RT-PCR med biomarkörerna CEA, MUC2 och KLK6 ser mycket lovande ut för att i framtiden kunna användas till stadieindelning och prediktion av prognos hos patienter med kolorektalcancer.

ABBREVIATIONS

APC	Adenomatous polyposis coli
CDCP1	CUB domain-containing protein 1
CD	Crohn's disease
CEA	Carcinoembryonic Antigen
CEACAM	CEA Cellular Adhesion Molecule
CK20	Cytokeratin 20
CRC	Colorectal cancer
EC	Epithelial cell
ECM	Extracellular matrix
FAP	Familial adenopolyposis coli
GC	Goblet cell
GCC	Guanylyl cyclase C
IEC	Intestinal epithelial cell
IEL	Intraepithelial lymphocyte
IFN-γ	Interferon-γ
IL	Interleukin
H&E	Hematoxylin and Eosin
HNPCC	Heriditary non polyposis colorectal cancer
IBD	Inflammatory Bowel Disease
KLK	Human tissue kallikrein
KRAS	v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog
LI	Large intestine
MHC	Major histocompatibility complex
MMP7	Metalloproteinase 7
MSI	Microsatellite instability
MUC	Mucin
SI	Small intestine
TP53	Tumor protein 53
TGF-β	Transforming growth factor- β
TNF-α	Tumor necrosis factor-α
UC	Ulcerative colitis

PAPERS IN THE THESIS

This thesis is based on the following papers, referred as Paper I-IV:

PAPER I

Biomarker selection for detection of occult tumour cells in lymph nodes of colorectal cancer patients using real-time quantitative RT-PCR
L Ohlsson, M-L Hammarström, A Israelsson, L Näslund, Å Öberg, G Lindmark and S Hammarström
Br J Cancer 2006;95:218-25

PAPER II

Lymph node CEA and MUC2 mRNA as useful predictors of outcome in colorectal cancer
L Ohlsson, A Israelsson, Å Öberg, R Palmqvist, H Stenlund, M-L Hammarström, S Hammarström and G Lindmark
Int J Cancer 2011 May 26 doi: 10.1002/ijc.26182

PAPER III

Lymph node kallikrein 6 mRNA – a new progression marker for colorectal cancer
L Ohlsson, G Lindmark, A Israelsson, R Palmqvist, Å Öberg, M-L Hammarström and S Hammarström
(Submitted)

PAPER IV

Pro-inflammatory and down-regulatory cytokine expression in tumor infiltrated mesenteric lymph nodes of patients with colorectal cancer
L Ohlsson, S Hammarström, Å Öberg, G Lindmark and M-L Hammarström
(Manuscript)

The published papers are reproduced with the permission of the relevant publishers.

1. INTRODUCTION

1.1 CANCER

Cancer is by definition an uncontrolled proliferation of cells that show varying degrees of fidelity to the cell from which it originated. Tumors can be divided into benign and malignant. Benign tumors do not invade adjacent tissue or metastasize, and are often more differentiated than malignant ones. Malignant tumors invade surrounding tissues and metastasize to distant sites, where they grow and invade further. Malignant tumors are classified according to the tissue of origin from which the tumor is derived. Carcinomas, the most common tumor type, arise from endodermal or ectodermal tissue such as skin and epithelium of internal organs. Leukemias and lymphomas originate from hematopoietic cells of the bone marrow. Sarcomas arise from mesodermal connective tissue such as bone and fat. Metastasis is defined as transfer of tumor cells from one site to another, where the site of transfer is not physically directly connected with the site of origin (1).

1.1.1 Tumor microenvironment

Tumors are much more complex tissues than just masses of proliferating cancer cells. They are composed of multiple specific cell types that interact with each other: cancer stem cells, cancer cells, endothelial cells, pericytes, immune cells, cancer-associated fibroblasts and stromal cells. All these different cell types, except of course the cancer cells, constitute the “tumor microenvironment” and can be seen in Figure 1 (2). Tumors are suggested to originate from tissue stem cells or progenitor cells that, through misregulation of self-renewal, form cancer stem cells (CSCs). CSCs possess stem cell properties such as self-renewal and differentiation. They have long telomeres and infinitive lifespan and replication potential. CSCs are either in a dormant or a proliferative phase. Dormant cancer stem cells have been shown to be more resistant to chemotherapy than ordinary cancer cells, if not targeted and, have the ability to regenerate a new tumor after therapy is completed (3). The cancer cells proliferate from CSCs and promote the tumor progression. These cells carry oncogenes, mutated or overexpressed genes that induce and/or inhibit cellular proliferation and/or regulate programmed cell death. Cancer cells within the same tumor are initially considered to be a homogenous cell population until the cancer advances further. Endothelial cells that construct the tumor-associated vasculature can be activated by vascular endothelial growth factor (VEGF). Pericytes are

a specialized mesenchymal cell type that wrap around the endothelial lining of blood vessels and interact with endothelial cells to synthesize the vascular membrane. Infiltrating immune cells, mainly macrophages, monocytes and T cells are common constituents of tumors. Collectively, they have both tumor-antagonizing and tumor-promoting properties. Cancer-associated fibroblasts are present in the tumor stroma and are divided into either myofibroblasts or cells that are similar to fibroblasts that support most normal tissues. Other constituents of the tumor are various stromal cells, which may be recruited from the adjacent normal tissue (2).

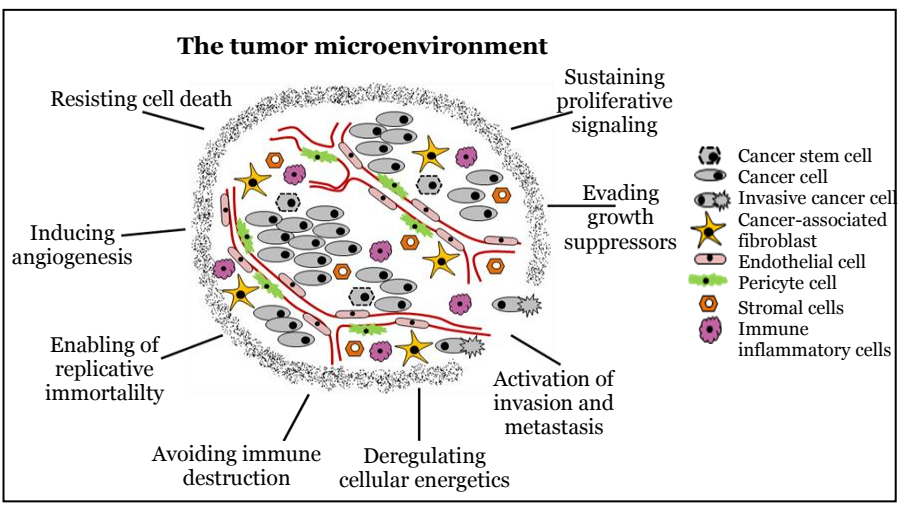


Figure 1. Illustration of the tumor microenvironment and eight functional capabilities required for its maintenance of malignancy, adapted from Hanahan et al. 2011 (2).

1.1.2 Genome instability and tumor promoting-inflammation

Functional capabilities (Figure 1) that allow cancer cells to survive, proliferate and disseminate are possible due to two important characteristics: genome instability and tumor-promoting inflammation (2).
Normal cells have complex systems that carefully detect and repair defects in the DNA to maintain genomic stability. DNA-damage response pathways stimulate the DNA repair system when the DNA is damaged in a cell. DNA damage is recognized by a variety of sensor and adaptor proteins, which activate upstream and effector kinases that target the major cell cycle control machinery. Depending on the extent of DNA damage, either all lesions in the cells are repaired and they can re-enter the cell cycle, or cells with irreparable DNA damage are eliminated by apoptosis or remain permanently

arrested (known as senescence). Spontaneous mutations are thereby normally very low during each cell generation (4). Defects in the maintenance of DNA and/or the repair system of the genome are critical for tumor progression. Cancer cells usually have an increased rate of mutations depending on increased sensitivity to mutagenic agents and/or through breakdown of components of the genomic maintenance machinery. Defects are found in genes that encode for products involved in detecting DNA damage and activating the repair system, directly repairing damaged DNA, inactivating or intercepting mutagenic molecules before they damage DNA, and/or telomerase activity. These types of functions are usually lost during tumor progression and the extent of genome alterations varies between different types of tumors (2).

Tumors are usually infiltrated by immune cells from both the innate and the adaptive immune systems, which induce inflammatory conditions. The tumor-associated inflammatory response may enhance tumor progression by supplying the tumor with bioactive molecules such as growth factors needed for proliferative signaling, survival factors that limit cell death, proangiogenic factors, extracellular matrix (ECM)-modifying enzymes that promote angiogenesis, and invasion and metastasis. Inflammatory cells can also release chemicals such as reactive oxygen species, which are mutagenic for nearby cancer cells, thus accelerating the mutations of the cancer cell genome (2).

1.1.3 Molecular mechanisms

There are at least eight functional capabilities required for cancer cells to survive, proliferate, disseminate and metastasize successfully: sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, deregulating cellular energy metabolism and avoiding immune destruction (Figure 1).

Normal tissues carefully control the cell growth and division cycle to maintain normal tissue architecture and function. Cancer cells manage to sustain and induce proliferation by deregulating production of growth factors and their receptors. The growth factors stimulate normal cells within the tumor-associated stroma and/or elevate levels of receptor proteins at their cell surface. Altered expression of growth factors or their receptors can result in uncontrolled proliferation (2). Cancer cells can also affect signal transduction pathways and transcription factors. Ras proteins are a family of GTPases involved in cellular signal transduction for control of cellular proliferation, differentiation and survival. About 20-25% of all tumors have mutations in ras oncogenes. Mutations in the ras oncogenes prevent GTP hydrolysis and lead to constant activation of Ras, i.e., uncontrolled cellular

growth, differentiation and survival (5). The Myc protein is a transcription factor responsible for activation of numerous genes. Myc promotes proliferation and cell survival and is frequently found to be upregulated in many types of cancer (6).

Cancer cells have to avoid the action of growth suppressor genes, which inhibit excessive cell proliferation. The most common suppressors of proliferation are the retinoblastoma-associated (RB) gene and the tumor protein 53 gene (TP53, also known as p53). RB functions as a critical and important gatekeeper of cell cycle progression, where it determines if cells will proliferate or become senescent and/or wheather apoptosis should be activated (2). p53 is a transcription factor that induces cell cycle arrest if it receives input such as stress and abnormalities from the intracellular signaling system. p53 is inactivated in almost all types of cancer leading to uncontrolled cell proliferation (7).

Programmed cell death by apoptosis is triggered in response to physiologic stress, DNA damage and/or imbalance in signaling. Apoptotic activation is regulated by both upstream regulators and downstream effector components (2). Caspases are important proteolytic initiators and executioners of apoptosis and are triggered in cells requiring termination. The regulators of apoptosis receive signals both from the extracellular and intracellular space, which results in initiation of proteolysis and degradation of DNA. Phagocytic cells consume damaged cells. Insufficient caspase activation results in a defect in apoptosis (8). The tumor suppressor p53 also functions to induce apoptosis in cells with damaged DNA or other chromosomal abnormalities. Cancer cells manage to avoid apoptosis by mutations in or loss of the p53 gene (7). Necrosis is another variant of cell death, where cells swell and explode, and release their entire cell contents into the local tissue environment. Necrosis triggers recruitment of immune cells that may bring growth-stimulating factors to the surviving cells, and thus is tumor promoting as well as tumor inhibiting (2).

Telomeres (guanine-rich repeated sequences) are located at the ends of chromosomes, protecting them, as well as being necessary for cells to pass through the cell growth and division cycle. Telomere repeats become shorter and shorter after each cell division cycle. Telomerase is an enzyme that adds telomere repeat segments to the ends of telomeric DNA and therefore in that way allows the cell to go on dividing cycle after cycle. Cancer cells manage to enable replicative immortality by upregulation of telomerase, which avoids triggering of senescence and apoptosis.

Induced angiogenesis is of great importance for cancer cells to get nutrients and oxygen, and for elimination of metabolic waste and carbon dioxide. Cancer cells induce expression of endothelial growth factor-A, which is involved in the production of new blood vessels, allowing the vasculature to continually generate new vessels (2).

The invasion and metastatic cascade represents a multi-step process starting with invasion of a local tumor cell followed by intravasation of the cancer cells into blood and/or lymphatic vessels. Finally, the tumor cells escape from the circulation and colonize at distal sites where micrometastases form macroscopic tumors. Less than 0.1% of the disseminated tumor cells succeed in developing a distant metastasis. Tumor cells can invade either by moving collectively or as single cells. Invasion of clusters of tumor cells occurs in epithelial cancers such as breast and colorectal cancer (CRC). The collectively moving tumor cells have intact cell-cell junctions, traction force for movement by multicellular coordination of polarity and cytoskeletal activity, and remodeling of the ECM. During cancer progression, the shape, mode of attachment to surrounding cells and the ECM of tumor cells becomes altered. A well-characterized type of alteration in epithelial cancer is the epithelial to mesenchymal transition (EMT). EMT causes loss of the cell-to-cell adhesion molecule E-cadherin between epithelial cells, and upregulation of the adhesion molecule N-cadherin associated with cell migration. The hematogenous route has been suggested as the main route for metastatic spread, but the lymphatic system is also a key player in tumor cell dissemination. Cancer cells enter the lymphatic system via the thin-walled lymphatic capillaries consisting of single layers of endothelial cells, endothelial transmigration (9).

Cancer cells need fuel for cell growth and division. They obtain this fuel through alterations in energy metabolism. Under aerobic conditions, glucose is processed by normal cells to pyruvate by glycolysis in the cytosol and then to carbon dioxide in the mitochondria. Cancer cells reprogram and limit the energy metabolism to glycolysis (2).

Cells and tissues are constantly monitored by the immune system. The role of the immune system in tumor development is summarized in the concept of immunoediting. Immunoediting includes three phases: elimination, equilibrium and escape. In the elimination phase, tumor cells are successfully recognized and eliminated by the immune system. The elimination of tumor cells includes both innate and adaptive immunity. The innate immune response includes effector cells such as natural killer (NK) cells, natural killer T (NKT) cells and γδ T cells, which are activated by inflammatory cytokines released from the tumor cells, surrounding macrophages and stromal cells. The adaptive immune response against tumor cells involves T lymphocytes (T cells) activated by tumor-specific antigens presented by dendritic cells (DCs). If the tumor cells are not completely eliminated, they proceed into the equilibrium phase, where the immune system controls tumor cell growth but is not able to eliminate them. Lymphocytes and interferon-γ (IFN-γ) play a critical role in this phase. Clinical evidence suggests that tumors can remain dormant for many years, and that they may cause relapse after a long period of time. Tumor cells can

be recognized by tumor-specific antigens and tumor-associated antigens. Unfortunately, most tumor antigens are not unique to tumor cells; tumor antigens may be expressed at low levels by normal adult cells and in much higher levels by tumor cells. However, tumors manage to arise in the presence of a functional immune system. The escape phase includes tumor cells that are no longer susceptible to immune attack. Cancer cells may avoid destruction by the immune system by lacking of tumor-specific antigens mediated by alterations of effector molecules, which complicates recognition and activation of the immune system. Tumors can also escape the immune system by production of soluble factors that are immunosuppressive such as VEGF, interleukin-10 (IL-10) and transforming growth factor-β (TGF-β). These immunosuppressive factors paralyze cytotoxic T cells and natural killer cells. Reduction or loss of major histocompatibility complex (MHC) class I on the surface of tumor cells reduces the action of cytotoxic T cells (10).

1.1.4 Causes

There are five known causes of cancer: oncogenic viruses; mutagenic chemicals; UV-irradiation in genetically susceptible individuals; radioactive irradiation and prolonged physical irritation by certain substances such as asbestos (1,11-12).

In humans, human T cell leukemia virus 1 (HTLV1) causes T cell leukemia/lymphoma. HHV8 causes Kaposi's sarcoma in untreated AIDS patients, Human papilloma virus 16 and 18 cause cervix cancer. Epstein-Barr virus causes Hodgkin's lymphoma and Burkitt's lymphoma, and Hepatitis B causes liver cancer. The viruses contain viral oncogenes corresponding to mostly mutated, cellular genes that regulate cellular growth and proliferation and/or death (1,11).

Lung cancer caused by smoking tobacco is perhaps the most important example of cancer caused by mutagenic chemicals. Chemical carcinogens can either induce mutations directly or be converted to metabolites that can cause mutations. Polycyclic aromatic hydrocarbons are an important group of chemical carcinogens.

Ultraviolet radiation and asbestos can also cause cancer and are classified as physical carcinogens (12).

1.2 THE IMMUNE SYSTEM

The immune system protects the human body from bacteria, viruses, fungi and parasites by generating a variety of cells and molecules that together recognize and eliminate these foreign invaders. It is composed of two main collaborating systems called the innate and adaptive immune systems.
The innate immune system is the first line of defense. It is strong and fast, but less specific than the adaptive immune system and includes components present before the onset of infection. Components of the innate immune system are physical barriers, phagocytes and pattern-recognition molecules. Skin and mucosal epithelium are examples of physical barriers and the acidity of the stomach is an example of a chemical barrier. The ingestion of extracellular particles is called phagocytosis and is conducted by specialized cells such as blood monocytes, neutrophils and tissue macrophages. Soluble factors such as lysozyme, antimicrobial peptides (AMPs), interferon proteins and complement system factors are other important components of the innate immune response. Cells and molecules involved in the innate immune system recognize molecules that are unique to microbes, therefore it can distinguish between self and pathogens. If the pathogen breaks through the primary barriers it will be detected and taken up by phagocytic cells causing an acute inflammatory response. Usually most infections in healthy humans are cleared within a few days by the innate immune system (11).
The adaptive immune system is the highly specific second line of defense that is activated in response to an infection. It takes five to six days after exposure to foreign molecules (antigens) to produce a primary response. Immunological memory is one of the hallmarks of adaptive immunity. Thus, upon reinfection by the same pathogen the vertebrate organism mounts a faster and stronger immune response. The major cellular components of adaptive immunity are lymphocytes and antigen-presenting cells. Lymphocytes have antigen-specific receptors on their surfaces. Immune T lymphocytes and antibodies produced by plasma cells are the main effectors of the adaptive immune system. Antigenic specificity, diversity, immunologic memory and self-nonself discrimination are four important characteristics of adaptive immunity. Lymphocytes are produced in the bone marrow by hematopoiesis. Lymphocytes can be divided into two major groups: B lymphocytes (B cells) and T lymphocytes (T cells). B cells mature in the bone marrow and when they are released each B cell expresses a unique antigen-binding receptor on its surface, the B-cell receptor. The B cell receptor can be secreted and is then called antibody or immunoglobulin. Two identical heavy chains and two identical light chains kept together by disulfide bonds and non-covalent forces form the basic “Y-like” structure of antibodies. Antibodies recognize specific epitopes on the antigen. When a naive B cell

encounters an antigen with an epitope that matches its specific B cell receptor, the B cell starts to proliferate and these B cells differentiate into either memory B cells or effector B cells (also known as plasma cells). The plasma cells produce and secrete enormous amounts of antibodies, which are the major effector molecules of humoral immunity. T cells are produced in the bone marrow, but migrate to mature in the thymus. Mature T cells express a unique T-cell receptor on their surface, which can bind to a specific epitope on the antigen. T cells are divided into two major subpopulations: T helper (Th) and T cytotoxic (Tc) cells. Th cells express CD4 molecules and Tc cells express CD8 molecules on their surfaces. CD4+ T cells consist of a large number of cell subsets. The major subsets are Th1 cells that secrete IFN-γ and participate in cellmediated immunity, Th2 cells that secrete IL-4 and IL-13 and conduct humoral immunity, Th17 cells that produce IL-17A and indirectly conduct antibacterial responses, and T regulatory (Treg) cells that produce IL-10 and/or TGF-β1 and downregulate immune responses. The T-cell receptor can only recognize its specific antigen if the epitope of the antigen is bound to a cell membrane protein called major histocompatibility complex (MHC). MHCs are glycoproteins with a peptide-binding pocket. There are two main classes: class I and class II. The latter class occurs on cell membranes of antigen-presenting cells (APCs) such as B cells, macrophages and dendritic cells whereas almost all nucleated cells express MHC class I molecules. The naive T cell is activated and starts to proliferate and differentiate into effector T cells upon interaction with an antigen presented on MHC class II of an APC cell. The activated T cell secretes cytokines that activate B cells, Tc cells, macrophages and several other cell types. Tc cells can be induced to form cytotoxic T cells by stimulation of certain cytokines. Cytotoxic T cells are capable of eliminating foreign invaders that display the specific antigen recognized by the T cell (11).

Antibodies present in plasma, lymph and tissue fluids constitute the humoral arm of adative immunity and protect the host from extracellular bacteria, free viruses, etc. Cell-mediated immunity is mediated by antigen-specific T cells in collaboration with other nonantigen-specific cells of the immune system. It protects the host from intracellular bacteria, viruses and cancer. As the effector molecules of the humoral response, antibodies bind to antigens constituting, for example, surface components of an invading bacterium. This antibody binding allows phagocytic cells to efficiently recognize and eliminate the bacterium. Antibody binding will also activate the complement system via the classical pathway resulting in lysis of the invader. Antibodies can also neutralize toxins produced by bacteria and neutralize virus particles by preventing them from binding to host cells. Activated Th cells and cytotoxic T cells are the effector cells of the cell-mediated response Th cells produce cytokines that can activate phagocytic

cells enabling them to phagocytose more efficiently, and cytotoxic T cells kill altered autologous cells, including virus-infected cells and tumor cells (11).
The innate and adaptive immune systems are highly interactive and cooperative to generate a stronger and more effective total immune response. Cytokines and chemokines are important signaling molecules in the cross-talk between the two systems.

1.2.1. Anatomy of the immune system

The lymphatic system consists of a network of capillaries, vessels and lymph nodes that returns filtered plasma-derived interstitial fluids (lymph) to the bloodstream. The blood circulates under pressure, which leads to the seeping of blood plasma through the thin walls of the blood capillaries into the surrounding tissue. This fluid is called lymph and delivers nutrients, oxygen and hormones to the cells in the tissues. Afterwards, the lymph leaves the tissue and brings cellular waste products by entering thin-walled tubes. These thin-walled tubes are also called lymphatic capillaries and are composed of a single layer of endothelial cells. Lymphocytes, dendritic cells, macrophages and other cells can enter these capillaries and join the lymphatic flow. From these tiny tubes the lymph flows to larger lymphatic vessels. The final and largest lymphatic vessel is called the thoracic duct, which empties the lymph into the subclavian vein, returning it to the bloodstream. The lymphatic system is designed so the lymph flows upward through the body from the feet and hands to the neck. The slow and low-pressure flow of the lymphatic system is accomplished by movement of the body's muscles that squeeze the vessels. The lymphatic vessels contain a series of one-way valves that ensure one-directional flow of the lymph. The lymphatic system drains all the tissues in the body of foreign antigens and transports them to organized lymphoid tissues such as lymph nodes (Figure 2), where the antigens are trapped and lymphocytes are activated.
Several different kinds of organized lymphoid tissues are situated along the lymphatic vessels: diffuse collections of lymphocytes and macrophages, networks of dendritic cells and resting B cells surrounded by a network of draining lymphatic capillaries (lymphoid follicles), lymphoid follicles within the lamina propria and highly organized lymphoid organs such as lymph nodes with distinct regions of T-cell and B-cell activity surrounded by a fibrous capsule (13).
Lymph nodes are encapsulated bean-shaped structures containing a network of packed lymphocytes, macrophages and dendritic cells that serve as a site for concentration of foreign antigens, allowing intimate contact with the immune cells and leading to lymphocyte activation. Lymph nodes are divided into three regions: the cortex, the paracortex and the medulla. The cortex contains lymphocytes (mostly B cells), macrophages and follicular

dendritic cells, which are arranged in primary follicles. Adjacent to the cortex is the paracortex, which contains mostly T cells but also dendritic cells that have migrated from tissues. These dendritic cells express MHC class II on their surfaces and are thereby able to present antigens to Th cells. The medulla is less heavily populated by lymphoid cells compared to the cortex and paracortex. It contains the antibody-producing plasma cells. When the antigen is carried into the lymph node by the lymph, it is initially trapped, processed and presented by the dendritic cells in the paracortex. This results

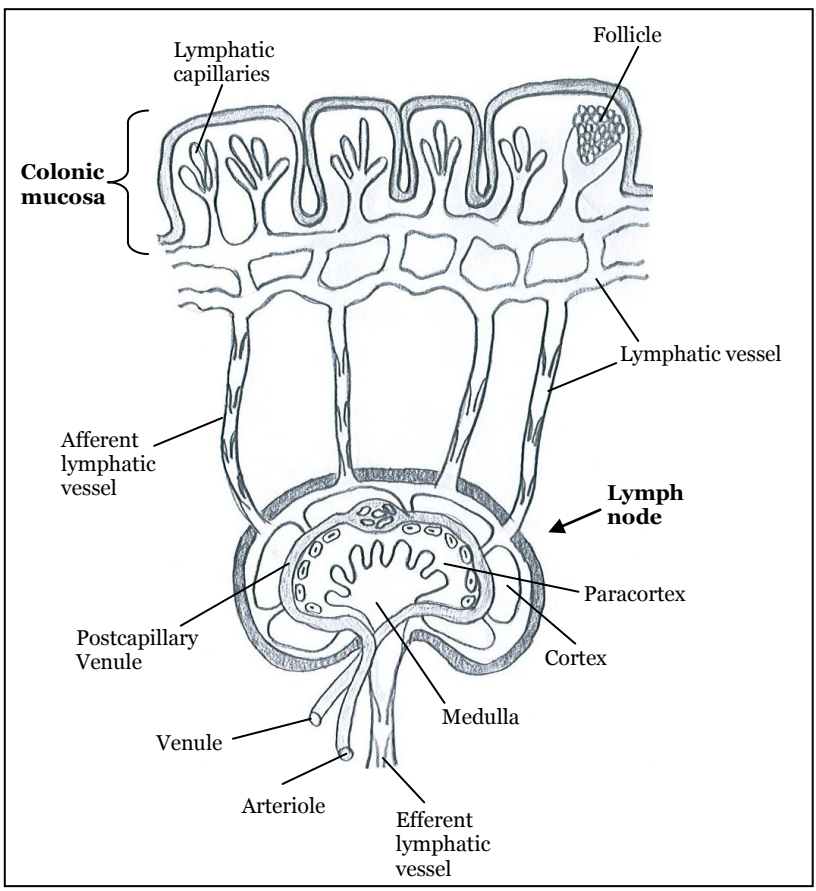


Figure 2. Schematic drawing of a lymph node, lymphatic vessels and lymphatic capillaries.

in activation of Th cells that later activate B cells and form small foci consisting mainly of proliferating B cells. Some of the B cells differentiate into plasma cells at the edges of the paracortex. After a few days, some of the

Th cells and B cells migrate to primary follicles in the cortex and form secondary follicles with a germinal center, which produce plasma cells that migrate to the medulla and the bone marrow. A lymph node has numerous incoming lymphatic vessels, but only one outgoing vessel called the subcapsular sinus. The lymph enters the cortex and then flows through the paracortex and finally the medulla before it leaves the lymph node. During this process, antigens become trapped by phagocytotic cells in the cortex, the immune response is activated by Th cells and B cells, and finally the lymph leaves the lymph node containing high levels of antibodies and lymphocytes. The increased concentration of lymphocytes in the outgoing lymph is due to lymphocyte proliferation within the node, but also the accumulation of lymphocytes that have migrated from the blood into the lymph node through the postcapillary venules (Figure 2; 13).

1.3 THE HUMAN INTESTINE

The human intestine is part of the alimentary canal extending from the stomach to the anus. It consists of two segments: the small intestine (SI) and the large intestine (LI). The SI is subdivided into the duodenum, jejunum and ileum, and its primary function is to digest and absorb nutrients and minerals from the digested food passing through. The LI is subdivided into the cecum, colon ascendens, colon transversum, colon descendens, colon sigmoideum and rectum (Figure 3A; 14). The main function of the LI is to absorb water from the remaining indigestible food and finally to pass useless material out of the body. In addition, anerobic bacteria fermentate carbohydrates leading to production of short fatty acids, which are absorbed by the colonic mucosa (15). The anatomy of the SI and LI are similar. The mucosa is the first layer closest to the gut lumen. It consists of a glandular epithelium, lamina propria and a thin smooth muscle layer (muscalaris mucosa). The mucosa is folded to increase the surface area of the intestine. These folds are called invaginations (crypts of Lieberkühn) and villi (finger-like projections) that project into the lumen. The mucosa of the SI consists of both villis and crypts, whereas the LI has an essentially flat surface with multiple crypts. Underneath the mucosa is the submucosa, consisting of loose connective tissue with collagen and elastic fibers. The muscalaris externa, two muscle layers made up of an inner circular and outer longitudinal muscle sheath, and finally the serosa complete the structure of the intestine (Figure 3B; 14).

The intestine hosts an enormous number of different microorganisms, mostly bacteria, in the luminal space. The composition of the intestinal microflora varies from individual to individual, although recent studies indicate that the very complex fecal flora can be divided into three distinct enterotypes based on the occurrence of a few key bacteria (16). Diet, age and genetic factors probably influence the composition of the microflora. The concentration of bacteria in the SI is approximately 10^{4-6} bacteria per milliliter of fluid. The colon has a less acidic pH, a larger volume, a lower concentration of bile salts and a slower luminal flow than the SI. The SI also contains Paneth cells that continuously produce anti-microbial defensins and lysozymes. Thereby, a higher concentration of bacteria, 10^{11-12} bacteria per gram of content, is present in the LI. The intestinal microflora has influence over the digestion and absorption of food, interacts with the intestinal immune system, produces vitamins and secretes hormones (17-18). Dominating bacterial genera in the SI are *Streptococcus* and *Neisseria* (19), and in the LI *bifidobacteria* and *Bacteroides* (17).

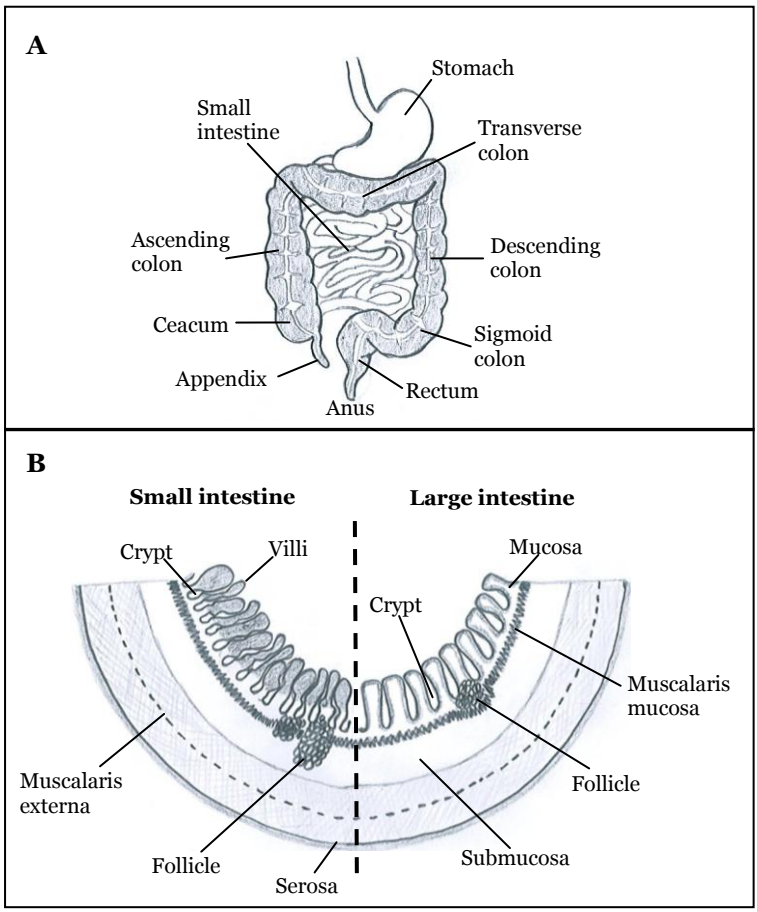


Figure 3. A) Schematic drawing of the gastrointestinal tract focusing on the large intestine and **B)** the layers of the intestinal wall of the small and large intestine.

1.3.1 Intestinal epithelium

The colonic epithelium consists of columnar epithelial cells (ECs), goblet cells (GCs), endocrine cells and intraepithelial lymphocytes (IELs). ECs and GCs are joined together by tight junctions. Single or small clusters of intraepithelial lymphocytes (IELs) are situated between the ECs close to the basal lamina (Figure 4A; 20). The colon mucosa also harbours follicles belonging to the immune system. A particular type of thin ECs, called microfold cells (M cells) cover the follicle. They transport macromolecules and particles, such as fragments of microbes, through the epithelium to the immune cells in the follicle. M cells express carcinoemryonic antigen (CEA)-family members at their apical surface (21). The hormone-producing

endocrine cells are found throughout the colonic epithelium. The most differentiated ECs are found at the free luminal surface, whereas the zone at the base of the crypt harbours the stem cells. Paneth cell-like cells are also present at the base of the crypts and have the function of nursing the stem cells (22).

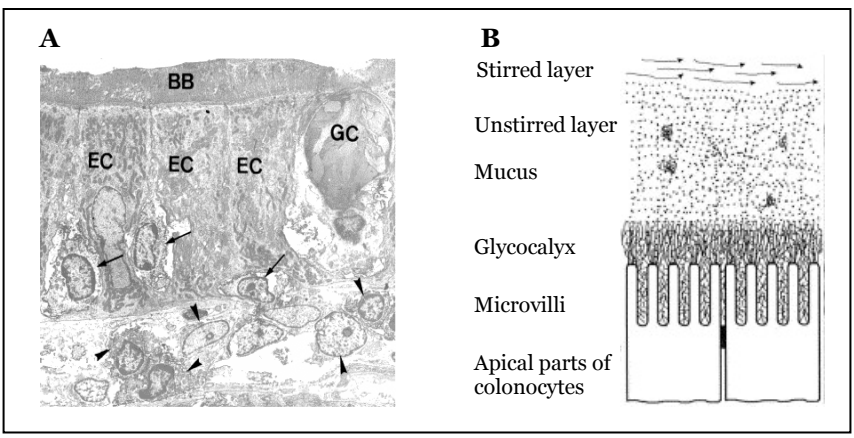


Figure 4. A) Low-power electron micrograph of human colonic mucosa, showing the epithelial layer and the underlying lamina propria. The colonic epithelium consists of columnar epithelial cells (EC), goblet cells (GC) and intraepithelial lymphocytes (arrows). The brush border (BB) comprises epithelial villi covered by the glycocalyx. **B)** Schematic drawing of the components in the barrier that protects the human colon epithelium (20).

The colonic epithelial cells are covered by a protective layer of hydrated viscous mucus, which consists of mucins synthesized and secreted by goblet cells and epithelial cell anchored mucins, water, electrolytes and fragments of the shredded host cells. Mucins (MUCs) are high molecular weight glycoproteins with a high content of clustered oligosaccharides. MUCs can be either secreted or membrane bound. MUC1, MUC2, MUC3, MUC4, MUC5B and MUC6 are expressed in human colon, where MUC2 is the main secreted mucin and a major component of the mucus layer (23). Below the mucus layer and on top of the apical surface of the colonic epithelial cells is the glycolax. It consists of a filamentous network of oligosaccharide chains and core molecules of plasma membrane-bound glycoproteins such as CEA-family members, glycolipids, proteoglycans and membrane-bound mucins (Figure 4B; 20).
Lamina propria is situated underneath the colonic epithelium. This is a loose areolar connective tissue occupied by components of the gut-associated lymphoid tissue, which includes immune cells such as T cells, B cells, plasma

cells, macrophages and dendritic cells, and also solitary lymphoid follicles and lymphatic vessels draining lymph from the tissue to the mesenteric lymph nodes (14).

1.3.2. Mucosal immunity

The intestinal mucosa is a critical, highly sensitive barrier against the surrounding world. Mucosal immunity must allow recognition of antigens of potentially harmful microorganisms and at the same time neglect to react against antigens of food constituents (oral tolerance). Oral tolerance is essentially a property of the small intestine and will not be dealt with here.
The intestinal epithelial cells (IECs) play important roles both in innate and adaptive immune responses. The colonic epithelium is protected from the turbulent luminal environment, firstly by the mucus layer. The carbohydrate moieties of the mucins provide potential binding sites for bacteria. Bound bacteria will move with the flow of mucus and be discharged into the gut lumen. Moreover, the mucus layer probably acts as an anchoring matrix for secreted agents such as secretory IgA, lysozyme and AMPs (e.g., defensins). A microorganism that is able to penetrate through the mucus layer, for example by possessing a mucinase, then meets the glycocalyx loaded with other types of receptor moieties. Binding to the glycocalyx, however, does not ensure access to the epithelial cells because components of this layer are easily released from the apical surface of the microvilli. The glycocalyx is also thought to serve as a size-selective diffusion barrier that excludes particles such as bacteria and viruses. Through tight junctions the epithelial cells maintain a physical barrier that is selective and permeable to prevent uncontrolled passage. In addition to these barrier functions, the IECs are responsible for interaction between the epithelial monolayer and the mucosal immune system by participating in the recruitment of immune cells to sites of infection (20,24). The IECs manage this through their ability to express and secrete chemokines and affect the local immune responses by expressing cytokines and producing antimicrobial effector molecules such as lysozyme and defensins (25).
The colonic epithelium also contains IELs in direct contact with the ECs. IELs probably have an important role in local immunosurveillance of the epithelial surface. IELs are T cells, which can recruit effector cells to the epithelium and lamina propria or act as costimuli for lymphocyte activation (26). Colonic IELs express interleukin (IL) IL-1β, IL-2, interferon-γ (IFN-γ), tumor necrosis factor-α (TNF-α), IL-8, and can be induced to express IL-10 and transforming growth factor-β1 (TGF-β1) (27). Epithelial cells release defensins like human beta defensin 1 and lysozyme and can be induced to release other defensins upon exposure to bacteria, e.g., human beta defensin 2. They also function to maintain a sterile environment in the crypt, thereby

having a protective function in defending the stem cells from microbial damage (22).

The lamina propria contains components of the gut-associated lymphoid tissue (GALT), which provides the colonic epithelium with adaptive immune protection. Involved in this adaptive immune response are CD8+ and CD4+ T cells, B cells, plasma cells, macrophages and dendritic cells. The plasma cells produce mainly IgA and the T cells are of T helper type 1 (Th1) and Treg cell type (28). Part of the GALT is also the lymphoid follicles present in the colonic mucosa. These are covered by M cells through which antigen can penetrate and reach the immune cells situated just below the M-cell. The follicles are inductive sites for an adaptive immune response. In the small intestine particularly the ileum, lymphoid follicles occur in relatively large groups called Peyer's patches.

Approximately 100-150 lymph nodes in the human body lie in the mesentery of the abdomen and are called mesenteric lymph nodes. These lymph nodes drain the intestine and play important role in activation of immune responses to antigens passing through the lumen (13).

1.4 COLORECTAL CANCER

Colorectal cancer (CRC) is the third most common form of cancer after breast and prostate cancer and represents one of the major public health problems worldwide, accounting for approximately 1 million new cases and half a million deaths every year (29). In Sweden more than 6000 new cases and 2600 deaths can be related to this disease every year (30-31).
Most cases of CRC are sporadic, but there is also a small fraction of patients in which hereditary factors play an important role (see below). Most likely, there is more than one cause of CRC. A major risk factor for CRC is long-standing inflammatory bowel disease. Obesity, a high- fat and low-fiber diet, alcohol consumption and cigarette smoking are other suggested environmental risk factors (32-33). The incidence of CRC increases with age and regular physical activity is associated with lower risk of CRC (33).
Common symptoms associated with CRC are blood in the feces and feces with mucus. However, it is also common for the patients to be asymptomatic (1).
CRC is defined by cancerous growths in the colon, rectum and/or appendix. More than 95% of colorectal cancers are adenocarcinomas, which evolve from the glandular epithelium. CRC progresses in a series of well-defined morphologic steps, which usually develops slowly over a long period of time. The origin of CRC is a small, benign tumor or polyp in the colorectal epithelium, called an adenoma, which gradually grows and becomes more and more disorganized in its intercellular organization until it develops into a malignant phenotype (1).

1.4.1 Pathogenesis

The pathogenesis of CRC is complex and is dependent on a wide variety of different factors. Both genetic alterations and epigenetic pathways play important roles in the development of malignant tissue. Four main molecular pathways have been identified to be involved in malignant transformation of CRC: the chromosomal instability (CIN) pathway, the CpG island methylator phenotype (CIMP) pathway, the microsatellite instability (MSI) pathway and the serrated pathway. In addition to these pathways other genetic alterations and chronic inflammation are also of importance.
The CIN-pathway is the most common genetic aberration in CRC. It is present in almost 85% of all CRC cases and is associated with poor prognosis. It is recognized by aneuploidy and structural alterations of the chromosomes and is associated with mutations in the adenomatous polyposis coli (APC) and v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS) genes. Mutation of the APC may co-exist with a deletion of the gene deleted in colorectal cancer (DCC), Mothers against

decapentaplegic homolog 2 (SMAD2) and/or SMAD4. SMAD2 and SMAD4 are involved in controlling the proliferation pathways. APC is a tumor suppressor gene and mutations of this gene are present in most CRCs as well as early in the development of adenomas. The APC protein has a variety of functions such as a role in microtubule formation and interaction with several proteins such as β-catenin. Inherited mutations in the APC gene are known as familial adenopolyposis coli (FAP). KRAS is a proto-oncogene important for transduction and intracellular signaling that controls cell proliferation. KRAS protein is located on the inner surface of the cellular membrane and is bound to GTP in its active conformation, where hydrolysis of guanosine triphosphate (GTP) to guanosine diphosphate (GDP) makes it inactive. Mutations of downstream mediators of KRAS lead to decreased GTPase activity, resulting in a constantly active KRAS. More than 50% of all CRC cases have KRAS mutations (34).
The second most common pathway promoting malignant transformation of colonic tissue is the CIMP-pathway, accounting for 15% of all CRC cases. This pathway is characterized by epigenetic instability, which is caused by hypermethylation of tumor suppressor gene promoters. The promoters contain CpG islands that are susceptible to hypermethylation. A defined CIMP CRC tumor must have methylation in at least three loci from a panel of CpG island-associated genes.

MSI arises when mutations occur in nucleotide repeat sequences throughout the genome. MSI is closely related to the mismatch repair (MMR) system, which functions to correct errors made by DNA polymerase during replication of DNA. MSI is usually subdivided into MSI-high and MSI-low depending on the degree of instability. CRC patients with MSI-positive tumors usually have better prognosis than patients with CIN tumors. Hereditary non polyposis colorectal cancer (HNPCC) patients have germline mutations in the MMR genes.

Previously the hyperplastic polyps were considered to be benign, but they are now classified as traditional serrated adenomas (TSAs), sessile serrated adenomas (SSAs) or true hyperplastic polyps. TSAs contain a uniform population of abnormal cells, whereas SSAs do not. Most serrated adenocarcinomas are MSI-low and arise from TSA. MSI tends to play an overlapping role even in this pathway. KRAS mutations are found in 80% of the TSA lesions and BRAF mutations in SSA lesions. Both KRAS and BRAF genes play roles in cell proliferation (34).

In addition to the four main pathogenic pathways there are other pathways affected by mutations that contribute to development of malignant colonic tissue. The TGF-β pathway is involved in cellular proliferation, differentiation and apoptosis, and alterations of this pathway are common in malignant colon tissue. Loss of the tumor suppressor gene DCC is also common in CRC and is associated with worse prognosis. Mutations of p53

are found in 50% of all CRCs. p53 is a negative regulator of cell division, which when DNA damage occurs increases and prevents cells from entering the S-phase, thereby allowing time for DNA repair. The EGFR-pathway plays a role in cell proliferation, angiogenesis and apoptosis. Increased expression of EGFR is associated with a more advanced stage of CRC and correlates with the invasiveness of the tumor (34).

1.4.2 Heredity

About 5% of all CRC cases are hereditary. The two main forms of hereditary CRC are familial adenopolyposis coli (FAP) and hereditary nonpolyposis colorectal cancer (HNPCC; Lynch syndrome) that account for approximately 1% and <5%, respectively, of CRC cases. FAP patients have mutations in the well-known tumor suppressor gene APC and this disease is autosomally dominant inherited. FAP is characterized by progressive development of innumerable adenomatous polyps of the colon. Young patients may have few polyps, but after some years the polyps increase and thousands of adenomas occur in the colon. At the age of 40 all these patients get CRC no matter what. HNPCC, also known as Lynch syndrome, is an autosomally dominant inherited disease. Patients with Lynch syndrome have an 80% risk of developing CRC. HNPCC patients have mutations in genes associated with the mismatch repair system (MMR), which corrects errors made by DNA polymerase during replication of DNA. These mutations lead to microsatellite instability (35).

1.4.3 Screening

Overall survival would improve if more patients were diagnosed at an early stage of CRC, because early-stage disease and precancerous lesions can be treated more successfully than advanced stage disease. Important reasons for CRC-screening are: that CRC causes significant morbidity and mortality, that treatment of advanced disease is a financial burden to society, and that effective treatments are present for early stages which often are symptom free. Available methods for screening for CRC are colonoscopy and sigmoidoscopy, CT- and MR-colonography, capsule endoscopy, and DNA and occult blood test in feces. Colonoscopy is considered the gold standard for detection of neoplastic colonic lesions. It allows examination of the entire colon and one has the ability to take biopsies and remove potential adenomas. Colonoscopy is widely used for elective screening of CRC in the USA, but it is an expensive method and therefore not optimal for general population screening. The faecal occult blood test (FOBT) has shown to reduce mortality caused by CRC and is cheap, simple and noninvasive. FOBT detects hemoglobin in the stool either enzymatically or immunologically. Using FOBT, the detection rate of early-stage (stage I and II) CRC has

significantly increased. In countries with general population screening, FOBT-positive patients are offered subsequent colonoscopy and other complementary diagnostic evaluation procedures. The fecal DNA test, which analyzes cellular elements containing genetic information shed with the stools, is also an option (36). KRAS seems to be a good candidate gene for these kinds of analyses, but unfortunately this method is more expensive and involves more complicated stool-sampling procedures than FOBT (34). A challenge for future screening would be development of a simple blood analysis for the detection of proteins and genes related to CRC. CEA was the first soluble biomarker accepted for clinical use for CRC, but is only used for monitoring of disease progression after surgery. Plasma tissue inhibitors of both metalloproteinase-1 and septin 9 have been suggested for early detection of CRC, but still no single test appears to be sufficiently specific for CRC (36).

1.4.4 Staging, diagnosis and prognosis

Staging CRC can be defined as combined measuring of the penetration of the cancer into the surrounding tissue and of distant metastases.

The only recently abandoned Dukes' classification system focused on the local extent of the tumor, lymphatic spread, venous spread and histologic grading, and included four different stages (Dukes' A-D). Dukes' A included tumors that are limited to the intestinal mucosa; Dukes' B tumors are those that had invaded into the intestinal wall; Dukes' C tumors could be of A or B origin, but with lymph node involvement; and Dukes' D tumors could be of A or B origin, with or without lymph node involvement, but with distant metastasis (37).

The current staging system, theTNM classification, is the latest and today most frequently used staging system of CRC assigns patients to stages I-IV depending on the extent and size of the primary tumor (T), regional lymph node status (N) and distant metastasis (M) status (Table 1; 38). T describes the degree of invasion of the intestinal tissue, N describes the degree of lymph node involvement and M describes the presence or absence of distant metastases. It is recommended, that as many regional lymph nodes as possibly but at least 12 lymph nodes should be harvested facilitating a correct lymph node status (38-39).

In clinical practice histopathology is used for tumor staging. In histopathologic examinations tissue sections are stained with hematoxylin and eosin (H&E) to investigate the local tumor invasion and the presence of CRC metastases in lymph nodes.

In addition to TNM classification, there are other well-established important prognostic factors such as tumor grade, venous and lymphatic invasion, budding and border configuration. Tumor grade is based on the extent of gland formation. Tumor budding is described as the transition from

Table 1. TNM classification system categorys and definitions

Stage category	Definitions		
TX	Primary tumor cannot be assessed		
T0	No evidence of primary tumor		
Tis	Carcinoma *in situ*; intraepithelial or invasion of lamina propria		
T1	Tumor has invaded the submucosa		
T2	Tumor has invaded into muscularis propria		
T3	Tumor has invaded through the muscularis propria and into the pericolorectal tissues		
T4a	Tumor perforates visceral peritoneum		
T4b	Tumor directly invade other organs or structures		
NX	Regional lymph nodes cannot be assessed		
N0	No lymph node metastasis		
N1	Metastasis in one to three regional lymph nodes		
N1a	Metastasis in one lymph node		
N1b	Metastasis in two to three lymph nodes		
N1c	Tumor deposits in the subserosa, mesentery or non-peritonealized pericolic or perirectal tissues without lymph node metastasis		
N2	Metastasis in four or more regional lymph nodes		
N2a	Metastasis in four to six lymph nodes		
N2b	Metastasis in seven or more lymph nodes		
M0	M0 distant metastasis		
M1	Distant metastasis present		
M1a	Distant metastasis in one organ or site		
M1b	Distant metastasis in more than one organ or site		
Stage I	T1-2N0M0		
Stage II	T3-4N0M0	Stage IIA	T3N0M0
		Stage IIB	T4aN0M0
		Stage IIC	T4bN0M0
Stage III	AnyTN1-2M0	Stage IIIA	T1-2N1M0
			T1N2AM0
		Stage IIIB	T3-4aN1bM0
			T2-3N2aM0
		Stage IIIC	T2-3N2bM0
			T4aN2aM0
			T3-4aN2bM0
			T4bN1-2M0
Stage IV	AnyTanyNM1	Stage IVA	AnyTanyNM1a
		Stage IVB	AnyTanyNM1b

glandular structures to single cells or clusters of up to four cells at the invasive margin of primary CRC tumors, which have malignant stem cell properties, and high potential of re-differentiation both locally and at distant sites (40).
Biomarkers may be useful as prognostic factors. Different biomarkers, both at the mRNA and the protein level, have been examined for this purpose. CEA is a well-known and established serum marker in CRC, but has so far not been found useful as a prognostic factor and guide for selection of treatment. CEA has proven useful for monitoring of disease after surgery (41). CA 19-9, CA 242, CA 72-4, CA 50 and tissue inhibitor of metalloproteinase 1 (TIMP-1) are other suggested potential serum markers for CRC (42). During the transformation from normal mucosa to malignant tissue, multiple molecular alterations occur, which provide many candidate biomarkers. Oncogenes such as KRAS, tumor suppressor genes (p53 and APC) and MSI are all usually altered in CRC (34). Preliminary data using cell- and tissue-based biomarkers appear promising for aiding diagnosis of early CRC, but are still not recommended for clinical use (42).

1.4.5 Mucinous Colorectal Carcinoma

Ten to twenty percent of all CRCs are mucinous. In mucinous CRC large amounts of the colon goblet cell mucin MUC2 is produced. The definition of mucinous CRC is that more than 50% of the tumor area must be covered by mucinous elements on histologic examination (43). MSI has been associated with mucinous carcinoma and MSI is at least twice as high as in non-mucinous carcinoma. Patients with mucinous CRC usually have better prognosis than patients with nonmucinous cancers (23).

1.4.6 Treatment

Surgery is the primary treatment for colon cancer and continues to have the greatest impact on CRC survival. In early disease surgery usually is a successful treatment, but residual micrometastases can cause relapse (44). To improve survival, surgery can be complemented with adjuvant chemotherapy or radiation therapy depending on the TNM stage. Surgery can be combined with adjuvant radiotherapy in stage II and III rectal cancer patients. Radiation therapy damages the DNA of cancer cells, which causes them to die or reproduce more slowly. Cancer cells have turned off the DNA-repair machinery during development from normal cells into cancerous cells, and are thereby more susceptible to radiation than normal cells (45). Fluorouracil- and levamisole-based chemotherapy is usually recommended to stage III colon cancer patients (46). Fluorouracil (5-FU) functions as an inhibitor of the enzyme thymidylate synthase. This enzyme is necessary for synthesis of thymidine, which is a nucleotide required for DNA replication.

Levamisole is an immunomodulator with immune stimulatory effects (44,47). Adjuvant chemotherapy is not regularly administered to stage I and II patients, but is reserved only for specific high-risk stage II patients (46). Follow-up procedures after curative resection are very important. In Sweden, follow-up regimens include computed tomography of the chest and abdomen as well as serum levels of CEA. Colonoscopy at five year intervals is normally performed to examine for metachronous colorectal neoplasia (46). CRC patients with metastatic disease who receive palliative chemotherapy may have improved survival and quality of life.

1.4.7 Survival

Overall five-year survival of CRC patients was shown by O'Connel et al to be approximately 65% (48). Dividing the patients into stage I-IV, the overall five-year survival rates were 93%, 83%, 60% and 8%, respectively. Most of the recurrences and deaths from CRC occur within two years after surgery (48). In Sweden approximately the same percentages of overall five-year survival rates apply (49). It has been shown that there is a significant difference in overall five-year survival rates between Dukes' C patients with a lymph node yield of less than nine nodes compared to Dukes' C patients with more than 9 retrieved lymph nodes (50). Adjuvant chemotherapy has been shown to reduce the number of deaths by one third in stage III patients (51-52).

1.4.8 Limitations and future challenges

Detection of mesenteric lymph node metastasis represents evidence of tumor cell dissemination beyond the primary location and is the single most important prognostic characteristic in CRC (53). Accurate staging of CRC tumors is important for selection of patients who may benefit from complementary adjuvant therapy. In lymph node analysis of tumor spread, H&E can detect one cancer cell among 200 normal cells, but only a fraction (less than 0.1%) of the available volume of the lymph node is analyzed, making it difficult to detect small numbers of tumor cells and micrometastases. Failure to identify micrometastases will result in undertreatment of patients. Understaging may also be explained by insufficient number of retrieved lymph nodes for examination and inadequate sensitivity of the detection method. Due to the difficulty in detecting disseminated tumor cells in lymph nodes, there is a need for more sensitive and better methods than the routinely used method of H&E staining (54-55). The prediction of outcome for CRC patients especially patients who are considered free of lymph node metastases, remains challenging and important. On one hand, it is important to improve the methods for identifying patients who appear to be free of disease after

surgery but will later develop recurrent CRC. On the other hand, it is equally important to more accurately identify patients who can be spared unnecessary adjuvant therapy treatment because their risk of recurrence is low (56).

One important group of CRC patients is those with mucinous CRC. Because of the small fraction of mucinous CRC patients, not much effort is put into identifying these patients. Instead they may receive unnecessary treatment, which can have critical life consequences. Therefore it is of vital importance to be able to detect mucinous CRC in a more sensitive and specific manner.

Alternative techniques have been suggested for staging the tumors of patients with CRC, such as immunohistochemical staining of tumor associated proteins, qualitative RT-PCR of molecular markers in regional lymph nodes, and detection of tumor cells in blood. A number of studies have demonstrated that qRT-PCR with specific biomarkers is more sensitive than traditional microscopy and allows the simultaneous testing of the entire available volume of each lymph node (57-60). One method is to use biomarkers for detection of disseminated tumor cells and for prediction of outcome. In this thesis, we show that screening for a combination of three biomarker mRNAs has great potential for both detecting tumor dissemination and predicting disease outcome.

1.5 INFLAMMATORY BOWEL DISEASE

Chronic inflammatory disease in the intestine or human inflammatory bowel disease (IBD) includes two major forms: ulcerative colitis (UC) and Crohn's disease (CD). Ulcerative colitis is a chronic inflammatory disease of the colon that mainly affects the mucosa and is characterized by chronic diarrhea and rectal bleeding. UC can occur at any age, but more commonly affects people between 20-40 years of age. Other symptoms such as lower abdominal cramps and weight loss for UC patients are also common. Crohn's disease is a chronic, granulomatous and inflammatory disease extending to the bowel wall. The inflammation mainly occurs in the small intestine and more rarely in the colon, where it is called Crohn's colitis. Abdominal pain, diarrhea and fever are common symptoms of CD. The inflammation usually affects all layers of the intestine. In contrast to UC, the inflamed tissue of CD patients is separated by segments of apparently normal tissue, so-called skip lesions (1). UC and CD patients have periods of disease remission, not only active disease.

There are diverse mechanisms that underlie UC and CD and thereby the cause is still unknown. There is extensive individual variability between patients with UC and CD in terms of clinical manifestations and responses, and patients usually need different medications. UC and CD are suggested to result from a combination four basic components: immunological factors, genetics, the intestinal flora and epidemiology. None of these four components is thought to individually manage triggering or maintaining the intestinal inflammation, thus a combination of various factors or most likely all four factors are needed (61).

No curative treatment of UC and CD is available, but treatment based on anti-inflammatory medications and/or other drugs to block the ongoing inflammation are effective in reducing the severity and symptoms of the disease. Ongoing research is based on antibiotics and probiotics to manipulate the intestinal flora (61).

UC and Crohn's colitis patients have increased risk for developing CRC, which accounts for approximately 1-2% of all cases of CRC. IBD with colon involvement is among the high-risk factors for CRC. About 10-15% of all IBD patient deaths are caused by CRC (32). The chronic inflammation is believed to promote carcinogenesis, where the long duration and severe anatomic extent of the inflammation is believed to be an increased risk for CRC (62).

1.6 BIOMARKERS

1.6.1 The Carcinoembryonic antigen-family (CEACAMs)

Carcinoembryonic antigen (CEA) is one of the oldest and best characterized tumor markers of all. Gold and Freedman discovered CEA in 1965 (63). Initially, CEA was believed to be expressed during fetal life, absent in adult life and re-expressed in cancer cells. Today we know that CEA is expressed in adult tissue as well, but with a restricted expression pattern found mainly in the epithelial cells in the colon. Shortly after CEA was discovered also the CEA-related cellular adhesion molecules (CEACAMs) were identified. These molecules were discovered mainly because of their cross-reactivity with CEA-antiserum and through sequence homology studies. Finally, CEA (also known as CEACAM5) and the other CEACAMs were collectively named the CEA-family. The CEA-family is a member of the immunoglobulin superfamily. CEACAMs are heavily glycosylated glycoproteins attached to the plasma membrane of cells. CEACAM1, 3 and 7 each contain a hydrophobic transmembrane domain followed by a cytoplasmic domain, while CEACAM2, 4, CEA and CEACAM6 are attached to the plasma membrane via a glycosylphosphatidyl inositol anchor. CEACAM1 occurs in two different forms with respect its cytoplasmic tail, which can be either long (CEACAM1-L) or short (CEACAM1-S) cytoplasmic tail (Figure 5; 64-65). The former contains ITAM/ITIM motifs.

CEA and CEACAM7 have selective epithelial expression. CEACAM3 and CEACAM8 are mainly expressed in granulocytes while CEACAM1 and CEACAM6 are more broadly expressed in human tissue. CEACAM1 is expressed in variety of epithelia in tissues such as stomach, kidney, colon, pancreas, and bile canaliculi, as well as on granulocytes, B-lymphocytes, T-lymphocytes and monocytes. CEACAM6 also has a broad expression pattern, with expression in the epithelia of different organs, in granulocytes and monocytes.

CEA, CEACAM1, CEACAM6 and CEACAM7 are expressed in healthy adult colon. The molecules are localized on the apical surface of the mature columnar epithelial cells. CEA and CEACAM6 are also found in the mucous droplets of the goblet cells. These CEACAMs are released into the gut lumen from the apical surface of mature columnar cells along with mucous, and then are expelled with the feces. In colon cancer the tumor cells have lost much of their polarity and express apical components like CEA together with baso-laterally expressed components. The tumor grows inward and tumor cells and their fragments reach the blood either via lymph vessels or directly through blood vessels.

The biological functions of all the members of the CEA-family are still poorly understood. CEACAMs act as intercellular adhesion molecules, probably as

receptors for signaling and have been suggested to play a role in the innate immune defense against microbes (64-65). CEA has been shown to mediate Ca2+-independent, homotypic aggregation of human colon adenocarcinomas cells (66). Bacterial binding to CEACAMs prevents infection-induced epithelial-cell detachment from the extracellular matrix by an integrin-dependent process, where CEA stimulates integrin-dependent binding of epithelial cells to fibronectin (67).

In recent years, CEACAM1 has become one of the most studied CEA-family members. The N-domain of CEACAM1 is responsible for the important cell-cell adhesion function and phosphorylation of the cytoplasmic domain for the signaling events (64-65). CEACAM1 has been shown to regulate several physiological and pathological processes such as tumor biology and leukocyte activation (64-73). Normal and tumor vessels of CEACAM1 deficient mice have been shown to have hyperactivation of the endothelial nitric oxide synthetase causing increased vascular permeability and reduced vessel maturation. Thus, CEACAM1 has been identified as a regulator of basal as well as acute vascular permeability (68). Moreover, peritumoral endothelial CEACAM1 expression has been shown to be critical for the formation of organized tumor matrix and intratumoral vessel maturation in mammary carcinomas (69). Moreover, CEACAM1 has been shown to play an important role in cell polarization and lumen formation in a model of breast morphogenesis (70). It was also shown that mammary carcinoma cells that lack expression of CEACAM1 also lack the ability of form intercellular lumina. When transfected with CEACAM1-4S morphogenesis is restored (71). CEACAMs are involved in the modulating of immune responses associated with infection, inflammation and cancer functioning as co-receptors for both lymphoid and myeloid cells. CEACAM1, particularly CEACAM1-L, has been shown to influence T cell function when expressed at the surface of T cells, predominately as inhibitor of T cell cytokine production, proliferation and/or cytotoxic activity. Additionally, CEACAM1-L on the surface of B cells and NK cells acts as an inhibitor of cellular functions (72-73). Neutrophils are the only immune cells that express CEACAM3, CEACAM4, CEACAM6 and CEACAM8 in addition to CEACAM1. CEACAMs are mobilized to the cell surface upon various activating signals. At the cell surface they may contribute to intercellular CEACAM-CEACAM binding interactions. It has been reported that CEACAM1-dependent intercellular binding is involved in survival signals that prevent neutrophils apoptosis (73).

CEACAM1 has also been shown to be upregulated in gastric adenocarcinomas in contrast to downregulated expression in CRC. CEACAM1 is thereby suggested to function as a tumor inhibitor in CRC (74-75).

CEA is expressed in cancers of epithelial origin particularly in CRC and pancreatic cancer but also in lung (adenocarcinoma), breast and some other cancers. The great interest in CEA emerged from the finding that CEA could be detected in the blood of patients with CRC, where patients with advanced CRC had high CEA levels in the blood compared to healthy individuals (64-65). Today, CEA is the most frequently applied serum tumor marker for follow-up of CRC patients (42). Due to restricted and high expression in normal colon epithelial cells with retained expression in CRC, CEA has also been used as a target in several immunotherapy trials. MEDI-565 is a recombinant construct composed of a human single-chain antibody recognizing CEA and a single-chain antibody specific for CD3 (76). MEDI-565 in combination with T cells has been shown to inhibit cell growth and promote apoptosis of tumor cells in CRC patients by antigen-specific T cell-mediated killing (77). On the other hand, treatment of patients with metastatic CRC using autologous T cells with a high-avidity T cell receptor that recognizes CEA demonstrated a decrease in CEA serum levels and objective regression of cancer metastatic to lung and liver. Unfortunately, severe transient inflammatory colitis was also induced (78).

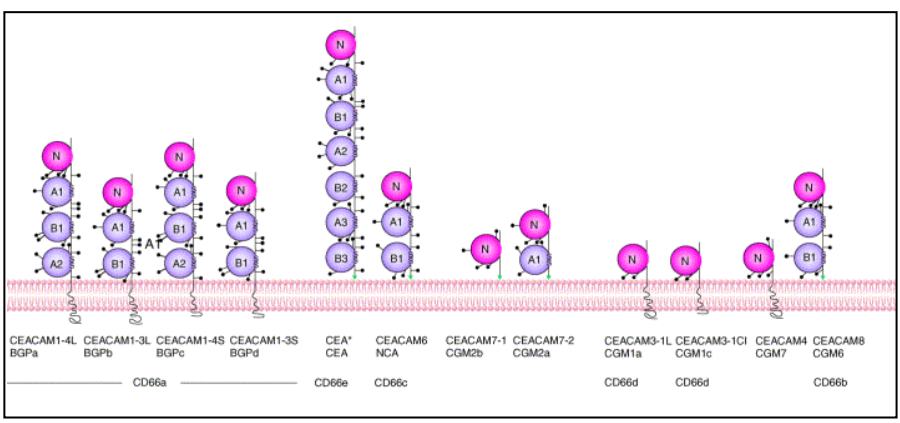


Figure 5. In this figure, models of the CEA-subgroup molecule are illustrated as models. The dark grey domains illustrate the immunoglobulin V-like amino-terminal ends and the light grey domains illustrate the IgC-like amino-terminal ends. The GPI (glycosylphosphatidylinositol) linkage to the cell membrane is shown as an arrowhead. Four forms of CAECAM1 can be seen, which are formed through alternative mRNA splicing (long and short cytoplasmic domain and three or four extra-cellular domains respectively). Glycosylation sites are illustrated as lollipops at the different CEA-members. The names are given in three different nomenclatures. The upper one is the newest nomenclature that is used today (64).

CEACAM6 is expressed in several cancer types and has also been suggested to be an important marker for CRC (79-80). CEACAM6 was shown to be

expressed at higher levels in colon metastases compared to the primary tumor (80). Antibodies targeting CEA and CEACAM6 affect cell migration, cell invasion and cell adhesion resulted in improved survival of mice with metastases (81).

1.6.2 Human tissue kallikrein-family (KLKs)

The human tissue kallikrein (KLK)-family consists of 15 secreted serine proteases with diverse expression patterns and physiological functions. KLKs are often expressed in, or localized at, the granular epithelium of different tissues. KLKs are secreted by the secretory epithelial cells and can therefore be found in body fluids such as sweat, milk of lactating women, seminal plasma and/or cerebrospinal fluid. Transcription of some KLKs can be regulated by sex-steroid hormones. KLKs are translated as inactive pro-enzymes, where the pro-peptide needs to be cleaved off in order for the serine protease domain to become catalytically active. The pro-enzymes of KLK can be substrates for already activated KLKs, thereby starting potential proteolytic cascades. The activation of pro-KLK proteins is irreversible. Activated KLKs are controlled by protease inhibitors, such as α2-macroglobulin and serpins. The kallikrein-family members are known for their dysregulated expression in a wide variety of carcinomas. KLK3 (also known as prostate-specific antigen, PSA) is used clinically as a serum biomarker for prostate cancer. KLKs are suggested to be involved in tumor progression and to play important roles in many cancer-related processes, such as cell growth and differentiation, tissue remodeling, angiogenesis, invasion and metastasis. Thus, members of the KLK-family are potential biomarkers for diagnosis, prognosis and monitoring of cancer (82-83).
Epression of human kallikrein 6 (KLK6, also known as neurosin) is restricted to the brain and central nervous system (CNS; 84). The physiologic function of this protein is not well established but recent work has demonstrated that KLK6 has a protective function in the CNS (85). A decreased level of KLK6 has been observed in CNS of Alzheimer's and Parkinson's patients (86-87).
KLK6 is suggested to be involved in tumor invasion and metastasis of cancer (86). KLK6 degrades basic constituents of the extracellular matrix (ECM) and the basement membrane, such as fibrinogen and collagens, which suggests that KLK6 plays a role in tissue remodelling and tumor invasion (88). Moreover, KLK6 downregulates the activity of E-cadherin. This may facilitate dissemination of tumor cells by reducing adhesion between cells (89-90). Elevated serum levels of KLK6 correlates with poor prognosis of ovarian cancer patients (83,91). In CRC, KLK6 mRNA was shown to occur at elevated levels in primary tumors compared to normal colon tissue and to be associated with a more advanced Dukes' stage, liver metastasis and poor

prognosis (89,92). KLK6 functions as a mediator of KRAS-dependent migration (93).
The human kallikrein 7 (KLK7, also known as human stratum corneum chymotryptic enzyme, HSCCE) is normally expressed in skin, esophagus, heart and liver (84). KLK7 catalyzes the degradation of intracellular cohesive structures in the outmost layer of the skin and contributes to the cell shedding process (94). The extracellular matrix protein fibronectin is a substrate for KLK7, which suggests that excretion of KLK7 leads to degradation of ECM (95). KLK7 cleaves the extracellular domain of E-cadherin, which produces a soluble fragment that decreases cell aggregation in pancreatic cancer (96). Both of these studies suggest that KLK7 is implicated in enhancing tumor cell invasion by directly degrading components of the ECM. KLK7 is suggested to be a potential biomarker in ovarian, breast and colon cancer, where high levels of KLK7 predict poor outcome (83,97-99). In colon cancer high KLK7 mRNA level is especially associated with liver metastasis (99).
Human kallikrein 11 (KLK11, also known as trypsinlike serine protease, TLSP) has a broad tissue expression pattern with most abundant expression in prostate and testis. It is secreted at high concentrations into seminal plasma. The physiological function(s) of KLK11 is still unknown (84). Similar to other members of the KLK-family, KLK11 shows dysregulated expression in cancer. KLK11 has been suggested as a potential prognostic biomarker in ovarian cancer, prostate cancer and low rectal carcinoma (LCR; 100-102). Interestingly, ovarian cancer patients with KLK11-positive protein expression in their ovarian tumors had a favorable prognosis (83,100). Moreover, upregulation of KLK11 mRNA in prostate cancer tissue was associated with less advanced and less aggressive tumors. Also, a lower concentration of PSA in serum was found in patients with overexpression of KLK11 (101). In contrast, LCR patients with high levels of KLK11 in tumor tissue had poor survival rates (102).

1.6.3 Mucin 2 (MUC2)

Mucins are a family of high molecular weight epithelial glycoproteins with a high content of clustered oligosaccharides. Mucins have a protective function in normal colon as they are the major glycoprotein in mucus, which covers and protect the epithelial surface. Structurally and functionally mucins can be divided into two distinct classes: 1) secreted, gel-forming mucins and 2) transmembrane mucins (23). The MUC-family contains 17 mucins (MUC1-2; MUC3A, B; MUC4; MUC5A-C; MUC6-17) with different tissue distribution. Six of them (MUC1, MUC2, MUC3, MUC4, MUC5AC and MUC12) are expressed in normal colon and in colon cancer cells. MUC2 is the main structural constituent of the mucus layer in the intestine, where it is

synthesized and secreted by the goblet cells (23,103). Colon cancer develops into distinct histologic subtypes, where one major type is mucinous colon cancer (43).

1.6.4 Cytokeratin 20 (CK20)

The mammalian cytoskeleton is responsible for the structural integrity of the cell, for protection against mechanical trauma and it has been suggested to have a role in communication between adjacent cells (104). The cytoskeleton is a highly dynamic and reorganizing structure. It consists of three different filament systems: microtubules (25 nm), intermediate filaments (10-12 nm) and actin-containing filaments (7-10 nm). Cytokeratins are intermediate keratin filaments present in the cytoskeleton of epithelial tissue. There are 20 known cytokeratins, which are divided into two subgroups: the low molecular weight, acidic (cytokeratins 9-20), and the high molecular weight, basic to neutral cytokeratins (cytokeratins 1-8). Cytokeratins are always expressed in specific pairs for every type of tissue, composed of one type I unit and one type II unit (104). CK20 is expressed in human intestinal epithelium, urothelium and Merkel cells in the epidermis. Because of its restricted expression in human tissue and its lack of or low level of immunological cross-reactivity between members of the cytokeratin-family, CK20 has become an important tool for detecting metastatic epithelium-derived cancer cells by immunohistochemistry and by PCR analysis (58,105).

1.6.5 Guanylyl Cyclase C (GCC)

Guanylyl cyclase C (GCC, also known as GUCY2C) is a member of the guanylyl cyclase family, which is a family of enzymes that catalyzes the conversion of GTP to cGMP. The guanylyl cyclase family comprises both membrane-bound and soluble isoforms. Guanylyl cyclases are regulated by diverse extracellular agonists, such as peptide hormones, bacterial toxins and free radicals. Guanylyl cyclase stimulation results in accumulation of cGMP, which regulates complex signaling cascades and plays a central role in the regulation of diverse (patho) physiological processes including vascular smooth muscle motility, intestinal fluid and electrolyte homeostasis, and retinal phototransduction (106). GCC is a transmembrane receptor for the bacterial heat-stable enterotoxins (STs) and for the paracrine hormones guanylin and uroguanylin. GCC has limited expression outside of the intestinal epithelium compared to the rest of the members of guanylyl cyclase family, which are more widely expressed (106-107). GCC plays a central role in maintaining mucosal homeostasis in the intestine. During colon carcinogenesis, the hormone ligands of GCC, i.e., paracrine hormones, are frequently lost. Dysregulation of GCC together with loss of paracrine hormone expression initiates neoplastic formation and tumor

progression (108). The expression of GCC is conserved during CRC and metastatic disease. GCC has been reported to be a colon-specific biomarker to detect disseminated tumor cells in lymph nodes of patients with CRC (56,59,109). Detection of GCC mRNA in lymph nodes has been associated with risk of recurrence for node-negative (pN0) CRC patients (56,59,60,109). Recently, a new molecular test called Previstage™ GCC Colorectal Cancer Staging Test has been made available. This test identifies disseminated tumor cells in lymph nodes and provides more accurate information about the staging of CRC patients (110).

1.6.6 Matrixmetalloproteinase 7 (MMP7)

Tissue remodeling is essential for maintaining normal tissue architecture, as well as for tumor progression, invasion, and metastasis, where matrix metalloproteinases (MMPs) play an important role. MMPs are zinc-dependent endopeptidases that belong to a large family of proteases. MMPs are involved in many processes such as degrading extracellular matrix proteins, cell proliferation, migration, apoptosis and differentiation (111). MMPs are secreted as proenzymes that must be cleaved to become active enzymes. Matrix metalloproteinase 7 (MMP7, also known as matrilysin) is expressed in the invasive front of colorectal tumors. Only MMP7 and membrane type-1 MMP can be secreted by the tumor itself (112). Expression of MMP7 has been related to decreased survival, and elevated level of MMP7 has been suggested as a predictor of disease recurrence and liver metastasis of in patients with CRC (113).

1.6.7 CUB domain-containing protein 1 (CDCP1)

CDCP1 is a type I transmembrane protein. Its extracellular domain contains three CUB domains and several possible glycosylation sites. CUB domains are structurally related to immunoglobulin domains and have important functions in cell adhesion (114). CDCP1 is associated with stem cells. It is found in hematopoietic stem cells and progenitor cells, and in tissues, including skeletal muscle, kidney, small intestine, colon, placenta, lung, stomach, esophagus and rectum (115-116). Dysregulated expression of CDCP1 is associated with cancer of the lung, kidney, colon and breast. An increased level of CDCP1 has been shown to correlate with poor prognosis, a higher relapse rate and occurrence of metastases. Also CDCP1 has been shown to function as an antiapoptotic molecule that during metastasis promotes tumor cell survival (116). Thus, CDCP1 emerges as a potential tumor marker and target to disrupt cancer progression. In addition to solid tumors, CDCP1 has also been suggested to be an independent marker of leukemia, bone marrow and mesenchymal stem/progenitor cells, and neural progenitor cells (116-117).

1.7 CYTOKINES

Many different cell types are involved in the development of an effective immune response. Cytokines play a central role in the communication between these cells. They are low molecular weight regulatory proteins or glycoproteins that bind with high affinity to specific receptors on the membrane of target cells. Binding triggers signal transduction in the target cell and results in altered gene expression. Due to high affinity between the cytokine and the receptor very low concentrations of cytokine are needed to mediate biological effects. Cytokines can bind to the same cell that secreted it (autocrine action), bind to a receptor on targets cells close to the producing cells (paracrine action) or bind to target cells at distant sites of the body (endocrine action). Cytokines are short-lived proteins and their most common mode of action is paracrine action. Pleiotropy, redundancy, synergy, antagonism and cascade function are typical characteristics for cytokines that enable them to regulate the cellular activity in a coordinated and interactive way. Pleiotropic action means that the same cytokine can have different biological effects on different target cells. Redundant action means that two or more cytokines mediate similar functions. Synergism of cytokines occurs when the effect of two cytokines is greater than the effect of individual cytokines on cellular activity, or an additional activity can be induced. Cytokines can act antagonistically, which means that they inhibit or offset the effects of other cytokines. Cytokines are structurally divided into four groups: the hematopoietin family, the interferon family, the chemokine family and the tumor necrosis factor family. The most common producers of cytokines are Th cells, dendritic cells and macrophages. Cytokine action is not antigen-specific. Cytokines regulate the development, intensity and duration of cellular and humoral immune responses. Moreover, cytokines regulate hematopoiesis, induction of inflammatory responses, and the healing of wounds (11).

Interleukin-10 (IL-10) is an anti-inflammatory cytokine that acts as a modulator in both innate and adaptive responses. The primary sources of IL-10 are monocytes/macrophages/dendritic cells and lymphocytes particularly Treg cells. IL-10 suppresses Th2 and Th1 cell growth. In contrast, IL-10 enhances B-cell survival, proliferation and antibody production. Lymphocyte-produced IL-10 downregulates the expression of MHC class II molecules and co-stimulatory molecules on macrophages and dendritic cells. Thereby, IL-10 inhibits antigen presentation. IL-10 has important effects on the intestinal immune response, where it is a key mediator for maintenance of gut immune homeostasis. Impaired IL-10 production is involved in the pathogenesis of Crohn's disease, while IL-10 levels are increased in active UC (118-119).

Interferon-γ (IFN-γ) is a proinflammatory cytokine critical for innate and adaptive immunity against viral and intracellular bacterial infections, and tumor control. CD4 Th1 cells and CD8 cells produce IFN-γ. IFN-γ is also produced by natural killer cells and natural killer T cells, which constitutively express IFN-γ mRNA that allow rapid induction upon infection. IFN-γ increases antigen presentation by upregulating the surface expression of MHC class I and II molecules. IFN-γ is also important for activation of macrophages in response to intracellular bacteria and viruses. Decreased IFN-γ induction or signaling is associated with increased susceptibility to intracellular bacteria, certain viruses and tumor induction (119).

Interleukin-17A (IL-17A, initially known as IL-17) is a proinflammatory cytokine that plays an important role in the defense against bacterial and fungal infections, but also in development of autoimmunity and inflammatory diseases. IL-17A is mainly produced by a specific subset of CD4+ Th cells called Th17 cells, but can also be produced by other T cell subsets such as CD8+ cells and gdT cells. IL-17A affects a variety of cells, which respond by upregulating expression of proinflammatory cytokines, chemokines, and metalloproteases. By upregulating expression of chemokines IL-17A recruits neutrophils for protection against extracellular pathogens. Moreover, IL-17A has been shown to be involved in several inflammatory disorders such as multiple sclerosis and IBD (119-120).

Tumor necrosis factor-α (TNF-α) is a proinflammatory cytokine that has important functions in the immune response to bacterial and certain fungal, viral and parasitic infections. Several cell types produce TNF-α. Major producers during infection are macrophages. TNF-α induces production of acute phase proteins during an inflammatory response. It induces and initiates a cascade of cytokines and increases vascular permeability, thereby recruiting both macrophages and neutrophils to the site of infection. TNF-α causes necrosis in some cancers, promotes growth of other types of tumor cells, participates in several inflammatory disorders and appears to be a major cytokine in Crohn's disease (121). In CRC, TNF-α can also be produced by the CRC tumor cells, which indirectly may lead to downregulation of the humoral immune response promoting tumor progression (122).

Transforming growth factor-β1 (TGF-β1) is a potent immunosuppressive cytokine that regulates multiple types of immune cells and can be secreted by most leukocytes. TGF-β1 suppresses effector Th cell differentiation, macrophages, dendritic cells and natural killer cells, and inhibits the conversion of naive T cells to Treg cells, proliferation of T and B cells and effector cytokine production. It is believed that TGF-β1 produced by T cells provides the most important downregulatory effects, including inhibition of T cell development, induction and maintainance of immune tolerance and homeostasis (123). TGF-β1 plays important roles both as suppressor and promoter in tumor initiation and progression (124).

2. AIM

Explore the utility of biomarker mRNA analysis of mesenteric lymph nodes from colorectal cancer patients for staging and prediction of outcome.

3. MATERIAL & METHODS

3.1 Patients

All patients included in this thesis underwent surgery at the Department of Surgery and Perioperative Sciences, University Hospital, Umeå, and at the Department of Surgery, Helsingborg Hospital/Lund University, Helsingborg, Sweden. Informed consent was obtained from all patients and in one case from the patient's parents. The Research Ethics Committee of the Medical Faculty, Umeå University, Sweden approved the study. A locally radical resection of the tumors with wide lymph node dissection was carried out in all patients. To characterize the tumors, Dukes' and/or TNM classification systems were applied including also tumor differentiation grade and mucinous status. No patients were lost for follow-up.
Clinical data on the patients is presented in Papers I-IV.

3.2 Lymph nodes

Lymph nodes for biomarker mRNA analysis were immediately dissected from the surgically removed specimen and bisected under sterile conditions with separate scalpels to prevent RNA cross-contamination. One half of each node was fixed in 10% buffered formalin for routine H&E staining and the other half was snap frozen in liquid nitrogen and stored at -70°C until RNA extraction (Figure 6). A median of two lymph nodes (range 1-15) per patient was subjected to both biomarker mRNA analysis and H&E staining. H&E staining was performed on additional lymph nodes that were obtained after the resected specimen was fixed in formalin overnight, resulting in total of a median of 13 lymph nodes (range 1-51) per patients for H&E analysis.
For further details, see Papers I-IV.

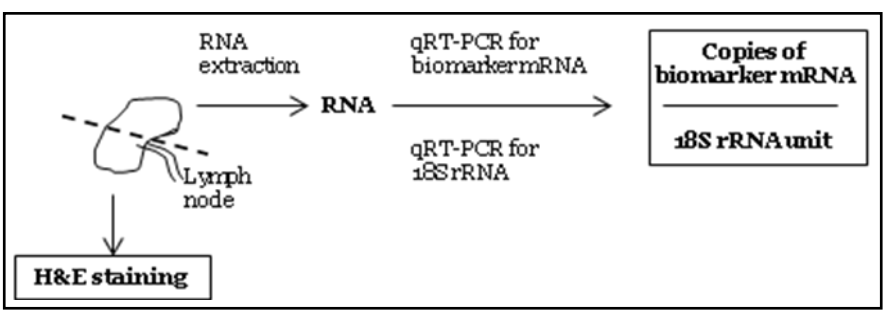


Figure 6. Schematic illustration of the procedure for collection of lymph nodes.

3.3 RNA extraction

Total RNA was extracted by using the acid guanidine phenol chloroform method (125) by adding 0.5 ml of a solution D (4 M guanidinium thiocyanate, 25 mM sodium citrate (pH 7.0), 0.5% sarcosyl and 0.1 M 2-mercaptoethanol) per 25 mg tissue or up to 2.5 x 10^6 cells in the first homogenization step. Extracted RNA was dissolved in RNAse-free water containing the RNAse inhibitor RNAsin. Further details of the protocol are described in Hayat 2009 (126).

3.4 Real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Comparative analyses of different biomarker mRNAs were based on mRNA level determinations by using real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) normalized to the expression of a housekeeping gene. In this method, the RNA is first reverse transcribed into cDNA and then amplified by PCR. RNA can be reverse transcribed into cDNA by using an oligo-dT primer, random hexamers or a sequence-specific primer. The oligo-dT primer binds to the poly-A tail of all mRNAs available in the sample and optimally produces full-length cDNA. Random hexamers bind to the RNA and produce short pieces of cDNA from all RNAs. The sequence-specific primer binds to a specific mRNA and produces its specific cDNA. We have used random hexamers for the RT step for the housekeeping gene, 18S rRNA, and sequence-specific primers for the biomarker mRNAs. Sequence-specific primers are particularly useful for determination of small amounts of RNA, as in our case with biomarker mRNAs in lymph nodes. This method utilizes a pair of primers (reverse and forward) that are complementary to the two cDNA strands of a defined RNA sequence of interest. The primers are then extended by a DNA polymerase. The DNA strands are amplified in an exponential manner. Real-time is defined as a continuous measurement of the amplified DNA during the PCR amplification. For the detection of products, a sequence-specific DNA probe labeled with a fluorescent reporter dye at the 5´end was used. The probe was also labeled with a quencher dye at the 3´end. The probe will not emit any fluorescence as long as it is intact: after cleavage by the DNA polymerase, the reporter dye is released and emits fluorescence (127). Figure 7 describes the details of the qRT-PCR steps.

All qRT-PCR assays for quantitative determination of biomarker mRNAs were constructed in the laboratory using Primer Express (Applied Biosystems) and performed using the Taqman EZ technology (Taqman EZ RT-PCR, Applied Biosystems). Specific reverse and forward primers were constructed and placed in different exons of the gene of interest. A sequence-specific dye-labeled probe was placed over the boundary between the two

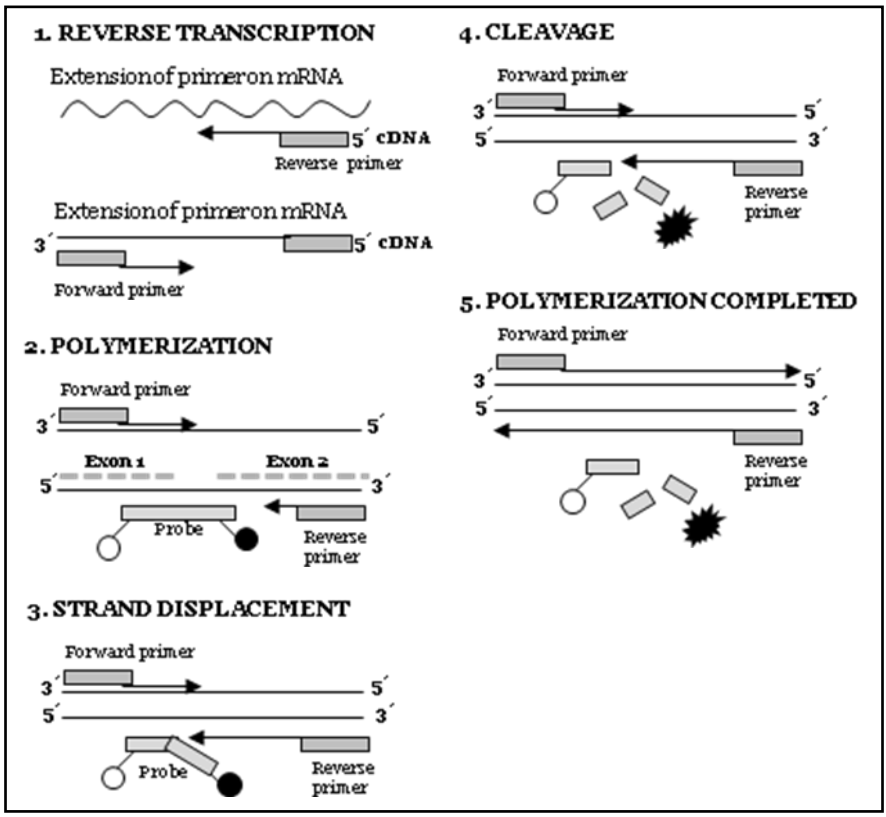


Figure 7. Determination of biomarker mRNA analysis: **1)** The mRNA is reverse transcribed into cDNA by extension of the reverse primer on mRNA followed by cDNA synthesis of the first cDNA strand. The second complementary strand of cDNA is amplified by extension of the forward primer followed by synthesis of the second complementary cDNA strand. **2)** The cDNA strands are separated by denaturation at 95°C, and then by decreasing the temperature, the primers and the probe can bind to the cDNA and begin polymerase chain reaction polymerization. The probe consists of a reporter dye at the 5´end and a quencher dye at the 3´end, which as longas it is intact, will not emit fluorescence. **3)** The DNA polymerase starts to build the complementary cDNA strand. **4)** The DNA polymerase cleaves the probe while building the new cDNA strand. The probe fragments are displaced from the cDNA and the reporter dye (fluorochrome) is released, emits fluorescence and can be measured. **5)** Finally, DNA polymerase completes the polymerization of the complementary cDNA strand. The first cycle is completed and double-stranded cDNA has been produced from the mRNA sequence of interest. Then the PCR cycle is repeated 40-45 times (127).

exons to make the assays specific for RNA and to avoid amplification of genomic DNA. The Taqman EZ technology uses r*Tth* DNA polymerase, which functions both as a thermoreactive reverse transcriptase and a thermostable DNA polymerase. This means that both the reverse

transcription and the DNA polymerization can take place without adding any subsequent buffers or enzymes (127). Primer sequences, probe sequences, qRT-PCR profiles, and reporter/quencher dyes for all examined mRNA assays are described in Papers I-IV and in Hayat 2009 (126). Emission from released reporter dye was measured by using ABI Prism 7700 and 7900 Sequence Detection System (Applied Biosystems).

An RNA copy standard was included in every qRT-PCR run to relate the fluorescence to absolute number of copies by using the linear relationships between log concentration and PCR cycles. Serial dilutions of the respective RNA copy standard over a six-log concentration range were included in each qRT-PCR run from which a standard curve was fitted and used as an external standard for determination of concentrations in unknown samples (Figure 8). All analyses of unknown samples were carried out in triplicate and expressed as copies of mRNA/µl. The reproducibility of most of these qRT-PCR assays are have been described in Papers II and III. Detailed descriptions of the preparation of RNA copy standards have been described in Fahlgren et al 2003 and Hayat 2009 (25,126).

The concentration of 18S rRNA was also determined in each sample by qRT-PCR according to the manufacturer's protocol (Applied Biosystems). 18S rRNA has been proven to be a stable and reliable housekeeping gene suitable for normalizing mRNA levels among immune cells. As no copy number standard is available for the 18S rRNA assay, 18S rRNA content was expressed as arbitrary units defined as the amount of 18S rRNA in 1 pg of total RNA extracted from polyclonally stimulated PBMCs (128). Results are expressed as mRNA copies per unit of 18S rRNA, thus yielding directly comparable levels between different biomarker mRNAs (Figure 6).

3.5 Illumina gene expression analysis

To search for CRC progression markers, we performed genome-wide screening using Illumina gene expression array technology with HumanRef-8 beadchips. Genome-wide screening or microarray analysis provides a view of gene activity in different biological samples as in this case the difference in expression levels of genes between CRC samples and controls. HumanRef-8 beadchips provide genome-wide transcriptional coverage of updated and well-characterized content from the National Center for Biotechnology Information (NCBI) reference sequences (RefSeq). Eight different samples are profiled in parallel on HumanRef-8 covering more than 18.000 unique curated genes.

The Illumina gene expression array technology uses silica microspheres (beads) as array elements. The beads are introduced randomly to the surface of a slide-sized silicon substrate with ordered microwells. Gene-specific

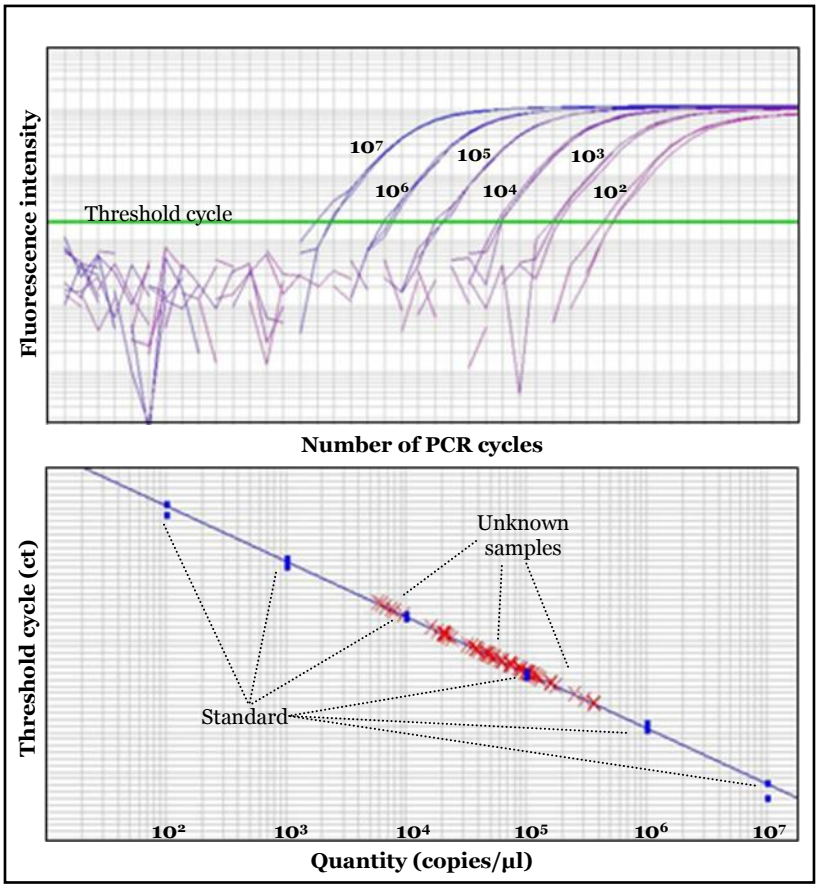


Figure 8. The appearance of a serial dilution of copy standard in a typical qRT-PCR run, which is fitted into a standard curve.

oligonucleotides (oligos) are covalently attached to beads, which consist of a 29-base address sequence and a 50-base probe sequence.
The address is used to map and decode the array, while the probe is used for quantification of transcript expression levels. Both sequences of each oligo have been carefully selected using bioinformatics. Each bead carries more than 100.000 identical oligos and is represented on average 30 times on the beadchip (129).
Total RNA was extracted using the acid guanidine phenol chloroform method (125) as described previously Paper II. The RNA concentration was measured in a NanoDrop ND-1000 Spectrophotometer (NanoDrop Technologies), and the integrity of the RNA was determined in a 2100 Bioanalyzer using an RNA nanoassay (Agilent Technologies). Total RNA was converted to biotinylated double-stranded cRNA according to the Illumina Totalprep RNA Amplification Kit (Ambion). First, the RNA is reverse

transcribed into cDNA followed by cDNA purification, in vitro transcription to synthesize and amplify cRNA, and finally cRNA purification. The in vitro transcription step incorporates biotin-labeled nucleotides. Thus, the output of this step is biotin-labeled cRNA. The labeled cRNA is then hybridized to the probe on the beadchips and stained with streptavidin-Cy3. Finally, fluoroscence emission by Cy3 is quantitatively detected by Illumina Beadstation GX (Illumina; Figure 9).

The data from the Illumina gene expression analysis was analyzed by using Illumina Beadstudio software (version 3.3) with the direct hybridization assay. Intensity data were normalized by Beadstudio's cubic spline algorithm with subtracted background. Significant differences in expression were calculated using the Beadstudio software Error Model Illumina Custom with multiple testing corrections using Benjamini and Hochberg False Discovery Rate (FDR). Differences in gene expression were calculated as fold change, dividing the signal in the CRC samples of interest over the average signal of control samples.

Real-time qRT-PCR was used for verification of microarray data. Taqman Gene Expression Assays were purchased from Applied Biosystems and performed by using the Taqman EZ technology (Applied Biosystems) as described in Paper II.

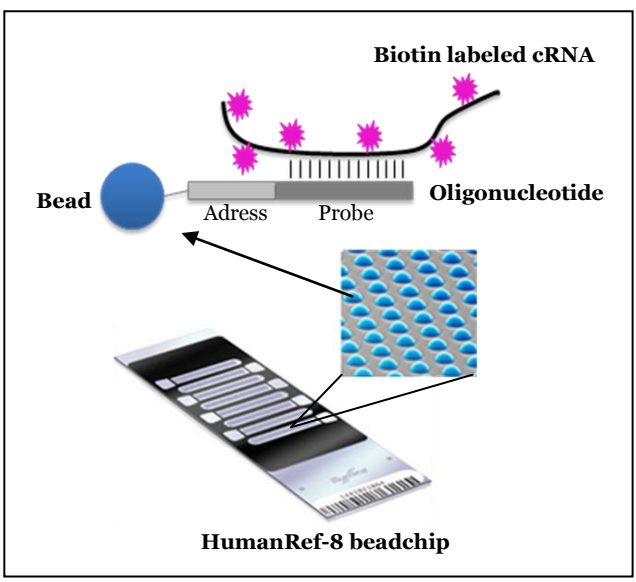


Figure 9. Schematic illustration of the Illumina Gene Expression Technology. Biotin-labeled cRNA is hybridized to the probe sequence of the oligonucleotides attached to the beads in the microwells of the HumanRef-8 beadchip.

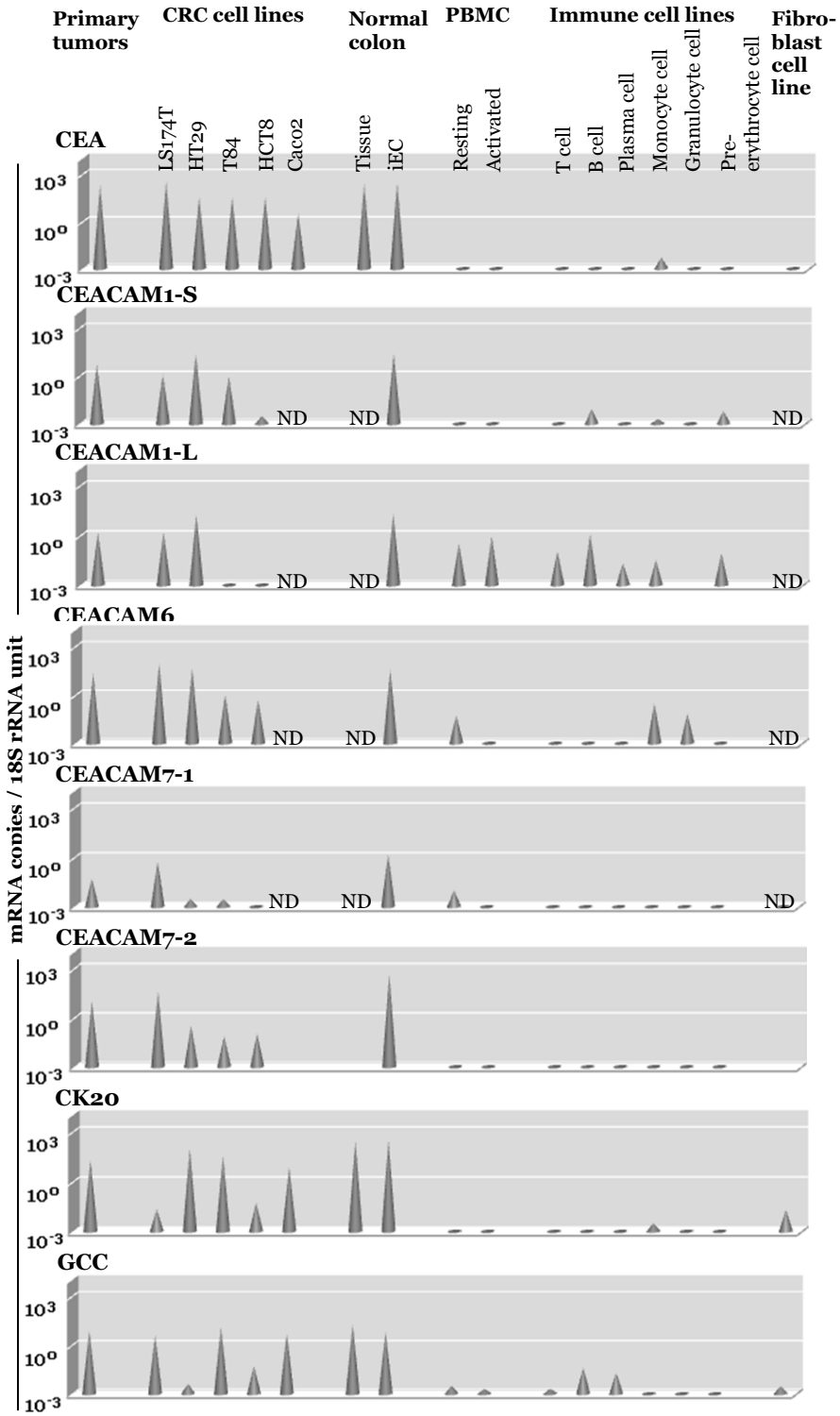
4. RESULTS

In this thesis, a total of 14 biomarker mRNAs, i.e., CEA, CEACAM1-S, CEACAM1-L, CEACAM6, CEACAM7-1, CEACAM7-2, CK20, MUC2, MMP7, GCC, CDCP1, KLK6, KLK7 and KLK11 were investigated for their ability to detect disseminated tumor cells of CRC patients. Accurate and sensitive detection of tumor cells is of fundamental importance for staging and prediction of prognosis. In Paper I, clinical material consisting of 113 lymph nodes from 51 CRC patients and 10 control patients were examined for biomarker mRNA levels. In Paper II, a significantly larger clinical material consisting of 635 lymph nodes from 174 CRC patients and 24 control patients, and in Paper III, 611 lymph nodes from 166 CRC patients and 23 control patients were examined. To investigate the ongoing immune response in lymph nodes containing disseminated tumor cells, levels of 5 cytokine mRNAs, i.e., IFN-γ, TNF-α, IL-17A, IL-10 and TGF-β1 were investigated. The cytokine mRNAs were also investigated for their utility as prognostic markers (Paper IV).

4.1 Expression levels of biomarker mRNAs in primary CRC tumor, CRC cell lines and normal colon

The expression levels of the 14 biomarker mRNAs varied considerably in the primary tumors. CEA mRNA had the highest expression level (median 176 mRNA copies/18S rRNA unit) and KLK7 had the lowest (median 0.009 mRNA copies/18S rRNA unit). Eight of the biomarkers were analyzed for homogeneity of expression within the same tumor. CEA, CK20, GCC and CDCP1 proved to be relatively homogeneously expressed in primary tumors, while MUC2, KLK6, KLK7 and KLK11 showed large intra-tumor variation. Five different CRC cell lines were analyzed for expression levels of all 14 biomarker mRNAs. CEA, KLK6 and CDCP1 had the highest expression levels with relatively small variation between CRC cell lines. The other biomarkers showed lower levels and larger variation in expression levels between the different CRC cell lines (Figure 10, Papers I-III).

mRNA expression levels of particularly CEA and CEACAM6, but also CEACAM1-S, GCC and CDCP1 were similar in primary CRC tumors and in normal colon epithelial cells and/or normal colon tissue, while CEACAM7-1/2, CEACAM1-L, MUC2, CK20 and KLK11 clearly displayed lower levels in primary CRC tumors. MMP7 was the only biomarker mRNA showing higher expression in tumor cells compared to normal colon epithelial cells. KLK6 and KLK7 were not expressed in normal colon at all, but were ectopically expressed in primary tumors (Figure 10, Papers I-III).


Primary tumors
CRC cell lines
Normal colon
PBMC
Immune cell lines
Fibro-blast cell line
LS174T
HT29
T84
HCT8
Caco2
Tissue
iEC
Resting
Activated
T cell
B cell
Plasma cell
Monocyte cell
Granulocyte cell
Pre-erythrocyte cell
CEA
CEACAM1-S
CEACAM1-L
CEACAM6
CEACAM7-1
CEACAM7-2
CK20
GCC
ND
mRNA copies / 18S rRNA unit
10³
10⁰
10⁻³

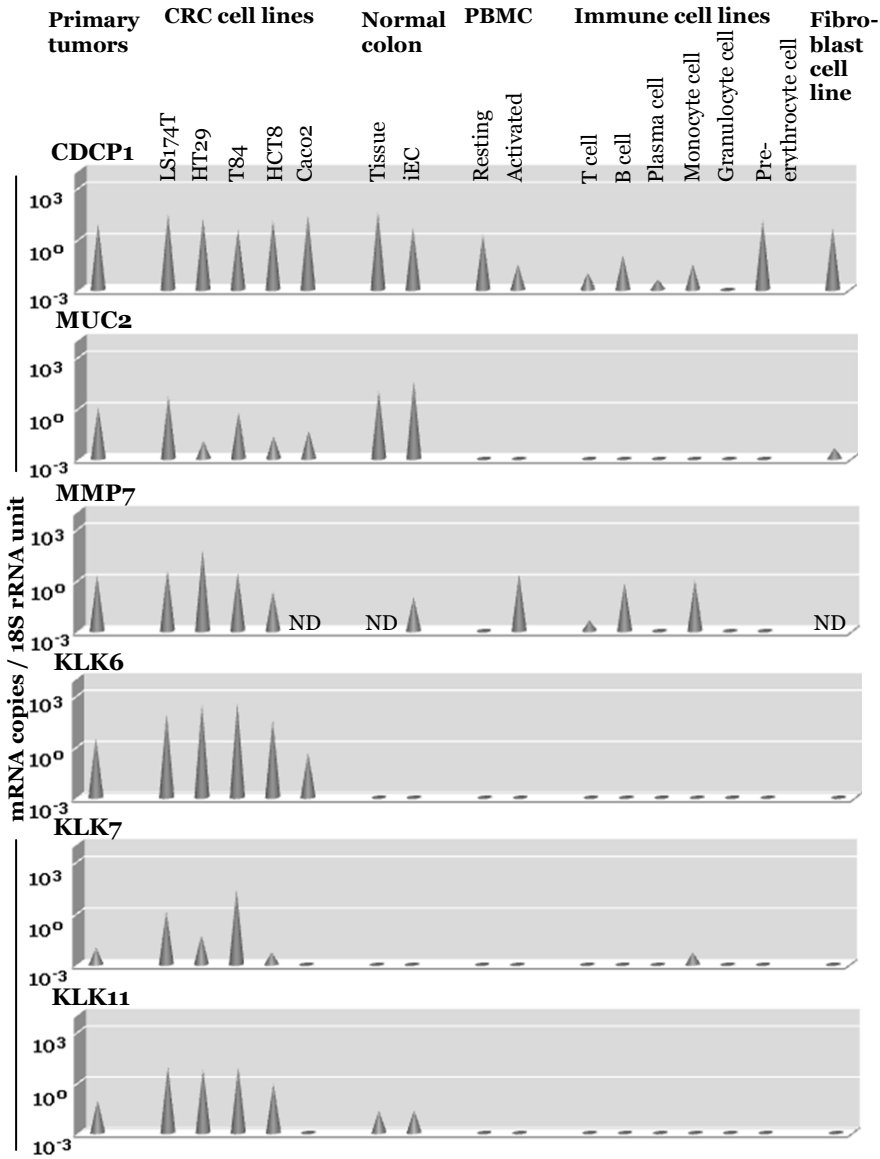


Figure 10. Expression levels of biomarker mRNAs in primary tumors, the CRC cell lines LS174T, HT29, T84, HCT8 and Caco2, normal colon tissue and isolated epithelial cells (iECs), freshly isolated (resting) PBMC and PBMC subjected to polyclonal T cell activation (activated), immune cell lines and a fibroblast cell line. The numbers of primary tumors analyzed were 85 for KLK6, 58 for KLK7 and KLK11, 57 for CEA, CK20, GCC, CDCP1 and MUC2, and 20 for CEACAM1-S/L, CEACAM6, CEACAM7-1/2 and MMP7. Expression of primary tumors is represented by the median value. ND = not determined.

4.2 Biomarker mRNA levels in immune cells and fibroblasts

To detect disseminated tumor cells in lymph nodes containing many immune cells and fibroblasts, it was important to examine the specificity of the biomarker for tumor cells. Therefore, it was essential to determine whether immune cells and fibroblasts also expressed the biomarkers. We were able to classify the investigated biomarkers into four different groups: 1) no expression in any of the immune cell types - KLK6 and KLK11; 2) only trace amount of expression in one type of immune cell - CEA, CEACAM7-2, MUC2, CK20 and KLK7; 3) low expression in more than one type of immune cell - CEACAM1-S, CEACAM7-1 and GCC; 4) high expression levels in several of the immune cell types - CEACAM1-L, MMP7 and CDCP1. Eight of the biomarkers were also analyzed for expression in a fibroblast cell line. CEA, MUC2, GCC, KLK6, KLK7 and KLK11 showed no or very low expression in fibroblasts, while CK20 showed some expression and CDCP1 showed high expression (Figure 10, Papers I-III).

4.3 Expression of biomarker mRNAs in lymph nodes of patients with benign intestinal disease

Expression levels of biomarker mRNAs were determined in lymph nodes from patients with benign intestinal disease (UC, Crohn's colitis, rectal prolapse and colon lipoma). KLK6 was not detected in lymph nodes of control patients, and KLK7 and KLK11 were not expressed or expressed only at very low levels in a few control nodes. CEA, CK20, MUC2 and GCC were expressed at low levels ($<10^{-2}$ mRNA copies/18S rRNA unit) in the control nodes, while CEACAM1-S, CEACAM6, CEACAM7-1 and CEACAM7/2 were expressed at higher levels (10^{-2}-10^{0} mRNA copies/18S rRNA unit). CEACAM1-L, MMP7 and CDCP1 had the highest expression levels ($>10^{0}$ copies/18S rRNA unit). The highest control node for its respective biomarker was used as a cut-off for discrimination between nodes from controls and CRC patients (Papers I-III).

4.4 Biomarker specificity of CRC

To rank the biomarkers with respect to specificity for CRC, two different specificity indexes were calculated: 1) the median value in primary CRC tumors/highest value in any type of immune cell and fibroblasts, and 2) the median value in primary CRC tumors/highest value of control lymph nodes (Table 2). Considering both indexes, the following ranking was obtained: KLK6>CEA>CK20>GCC>MUC2=KLK11>CEACAM7-2>CEACAM1-S>CEACAM6>KLK7>CEACAM7-1=MMP7>CDCP1>CEACAM1-L. The differences between the two indexes are mainly due to the type of immune cell in which the biomarker is expressed at the highest level and the frequency of this cell type in lymph nodes. KLK6 and CEA stand out as

highly specific for tumor cells of the colon. CDCP1, CEACAM1-L, MMP7, CEACAM7-1 and KLK7 showed essentially no specificity for CRC (Table 2).

Table 2. Specificity indexes for biomarker mRNAs

Biomarker	Median CRC tumor value/highest value of any immune cell type	Median CRC tumor value/highest value of control lymph nodes
CEA	35200	13538
KLK6	30000	30000
MUC2	10000	161
CEACAM7-2	9000	90
CK20	6000	1819
KLK11	800	167
CEACAM1-S	638	39
GCC	175	200
CEACAM6	11	2000
KLK7	2	19
CEACAM7-1	1	6
MMP7	1	1
CEACAM1-L	1	0.4
CDCP1	0.6	4

4.5 Biomarker mRNA levels in lymph nodes of CRC patients – relationship to stage

Table 3 summarizes the results of biomarker mRNA analysis of lymph nodes from CRC patients in the first and second study (Papers I-III). Results are shown as percentage of patients of each stage that show biomarker mRNA values over the levels of the control group. Note that this cut-off level is different from the clinical cut-off level (see below). Each patient was represented by the lymph node with the highest value. Paper I indicates that CEA, CK20, MUC2, and CEACAM6 and CEACAM1-S show promise as biomarkers for CRC because the frequency of positive values generally increased with stage. The same trend was seen when median values for the different groups were compared. However, since this study was small with only 6 and 9 patients in stage I and stage IV, respectively, the actual percentages are rather inprecise. CEACAM1-L, CEACAM7-1/2 and MMP7 had no value as CRC lymph node biomarkers.

Paper II, which included lymph nodes from three times the number of CRC patients, a much more precise estimate of the frequency of biomarker positive lymph nodes was obtained. CEA, CK20 and KLK6 detected between 54 and 94 % of all stage III and IV patients, while GCC, MUC2 and CDCP1

detected between 37 and 69% of the patients in these two groups. The frequency of stage I and II patients with biomarker mRNA positive lymph nodes varied between 9 and 28% for all markers except CEA. For CEA mRNA the corresponding frequencies were 52-57%. The finding that CEA mRNA analysis detects such a high frequency of patients with positive nodes is most likely due to its high expression level in CRC tumor cells allowing detection of very few tumor cells/ lymph node.

Table 3. Percentage of CRC patients with biomarker mRNA levels over cut-off of controls

Paper	Number of CRC patients	Biomarker	% of CRC patients with values over cut-off of controls			
			Dukes'A /stage I	Dukes'B /stage II	Dukes'C /stage III	Dukes'D /stage IV
I	51	CEA	100	88	100	78
		CEACAM6	17	27	80	44
		CEACAM1-S	17	21	70	25
		CEACAM1-L	17	13	20	0
		CEACAM7-1	0	8	20	13
		CEACM7-2	0	20	30	29
		CK20	50	42	90	56
		MUC2	17	35	80	50
		MMP7	0	12	50	25
II-III	166-174	CEA	52	57	72	94
		CK20	28	24	54	75
		MUC2	10	9	37	44
		GCC	14	16	37	69
		CDCP1	17	11	35	69
		KLK6	21	11	54	75

4.6 Correlations between biomarker mRNAs

Pair-wise comparisons of biomarker mRNA levels in primary tumors and lymph nodes were performed. In primary tumors there was no significant correlation between any of the investigated biomarkers. However, when lymph nodes were compared biomarker mRNA levels in many cases were correlated to each other. Comparing lymph nodes from all stages, we found that particular CEA and CK20 mRNA levels were significantly ($p<0.0001$) correlated with a high correlation coefficient ($r=0.76$). When the different stages were analyzed separately it was noted that particularly lymph nodes from Dukes' stage C/stage III patients showed significant correlations between CEA on one hand and CK20/GCC/MUC2 on the other. Moreover, CK20 and GCC, as well as CEA and KLK6 were correlated with each other. All these comparisons had r-values >0.70. Lymph nodes of Dukes' stage D/stage IV patients also showed significant correlations between CEA on one hand and CEACAM7-2/CK20/GCC/MUC2/KLK6 on the other and between CEACAM7-2 and CK20/MUC2. In lymph nodes from Dukes' A/stage I and Dukes' B/stage II patients only CEA amd CK20 mRNA values were strongly

correlated with each other. The findings that independent epithelial-specific biomarker mRNAs levels are correlated with each other showing higher correlation coefficients with increasing stage is a strong indication that they detect the same cells i.e. disseminated tumor cells (Papers I-III).

4.7 Poor correspondance between tumor cell detection by biomarker mRNA analysis and H&E staining

Figure 11 shows the result of comparison between biomarker mRNA analysis and H&E staining of lymph nodes of CRC patients. All lymph nodes (95 in Paper I and 600 in Paper II) are included in the figure. The dashed line indicates the cut-off level corresponding to the highest value of control nodes. Considering firstly, lymph nodes that are H&E positive (filled circles) it can be seen that in Paper I, CEA, CEACAM6 and CEACAM1-S mRNA detected all H&E positive nodes. In contrast, CEACAM1-L and CEACAM7-1/2 missed 50% or more of the H&E positive nodes disqualifying them as CRC tumor markers. In Paper II CEA missed two and KLK6 three H&E positive nodes and the other four markers between 5 and 10 H&E positive nodes. For CEA this constitutes 5 % of the H&E positive nodes. Interestingly, these two H&E positive nodes were also missed by all other biomarkers.
Considering H&E negative nodes (empthy circles), it is interesting to note that a rather large fraction of these nodes were biomarker mRNA positive and thus contained tumor cells missed by H&E staining. For CEA mRNA as much as 57% of the H&E negative nodes were marker positive; some with very high CEA values.

4.8. Biomarker mRNA analysis identified several stage I and II CRC patients who died from tumor recurrence

Stage I and II CRC patients are by definition considered to lack tumor cells in their lymph nodes. In Paper I one CEA mRNA positive stage I and three CEA mRNA positive stage II patients died from recurrent CRC within the follow-up time after surgery. Using CEA mRNA biomarker analysis, these patients would have been classified as stage III patients and been eligible for chemotherapy, which possibly could have saved them. These four patients were also identified by CK20 mRNA and three of them also by MUC2 mRNA (Paper I).

4.9. Determination of clinically relevant cut-off level for biomarker mRNA

As previously described a larger clinical material was analyzed in Paper II and III. This quantity of clinical material allowed determination of clinically

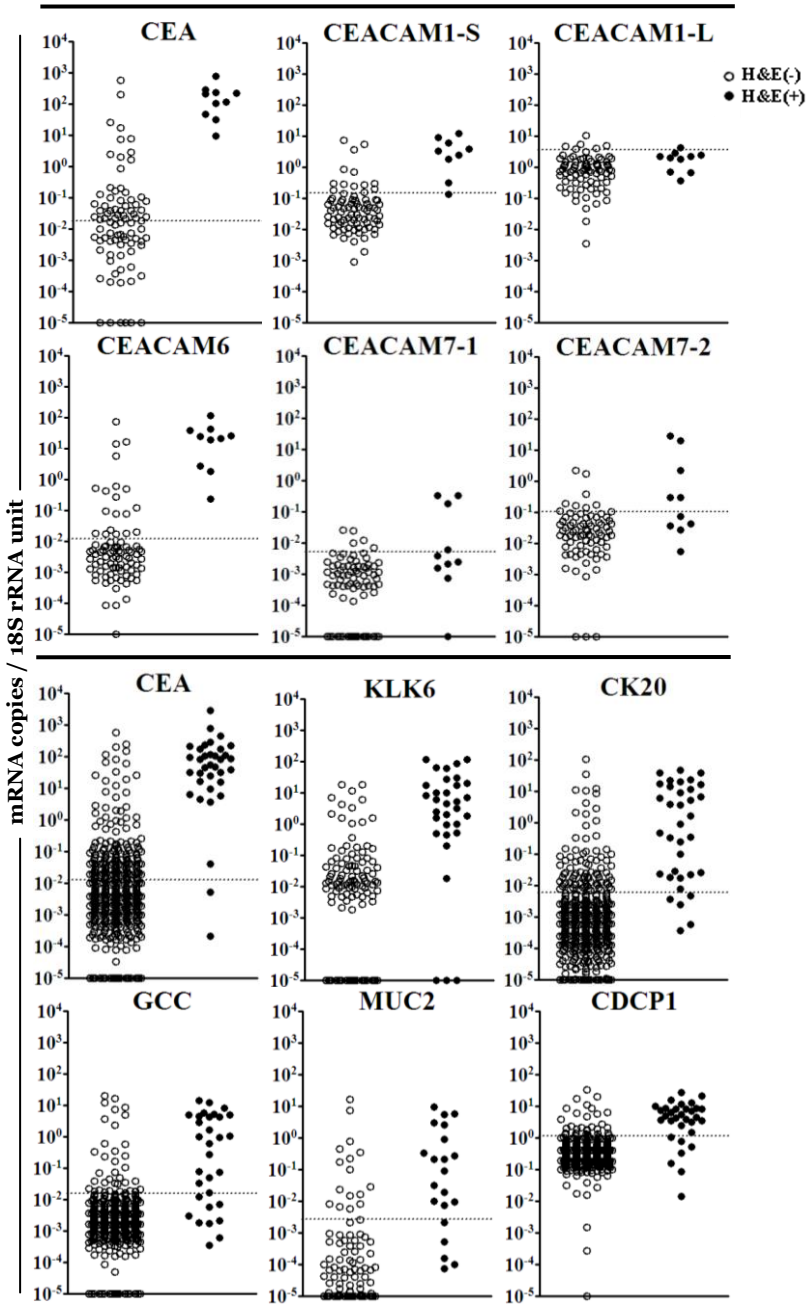


Figure 11. Biomarker mRNA levels in H&E-negative (unfilled circles) and H&E-positive (filled circles) lymph nodes of CRC patients. Dotted line indicates highest control level of respective biomarker. CEACAM1-S/L, CEACAM6 and CEACAM7-1/2 depict the results from the smaller clinical material (n=95 lymph nodes). For comparison results for CEA are given for both the smaller and the larger clinical material (n=95 and 517, respectively).

relevant cut-off levels. During the follow up time 56 patients had either died from CRC or were living with recurrent disease, wheras 118 patients were apparently cured or had died from causes unrelated to CRC. We defined the biomarker value at the 80th percentile as the cut-off level by applying Cox regression analysis on these patients' survival data.
Using this, so-called, clinically relevant cut-off level for CEA mRNA, we found that only 6 stage I+II patients were CEA(+). Two of these succumbed from CRC, another two patients died from unrelated diseases and could have had recurrent disease at the time of death. The remaining two patients were alive without recurrent disease. The same two patients who died from recurrent CRC were also identified by KLK6-, CK20-, GCC- and MUC2 mRNA, but not by CDCP1 mRNA (Papers II and III).

4.10 Difference in survival time after surgery between biomarker(+) and biomarker(-) CRC patients

The clinical cut-off levels for biomarker mRNA analysis of lymph nodes of CRC patients, as defined in the previous section, was also used to investigate differences in survival time between the biomarker(+) and the biomarker(-) group of patients (Paper II). Follow-up was performed on January 1, 2011 for all patients and no patients were lost for follow-up. Cumulative survival was calculated for CEA, CK20, GCC, CDCP1, MUC2 and KLK6 by using Kaplan-Meier. CEA mRNA showed the greatest difference in mean survival time between CEA(-) patients compared to CEA(+) patients. It was 71 months. CK20 and GCC were quite close with differences of 66 and 63 months, respectively (Paper II).
The mean survival time of KLK6(-) patients was 117 months compared to only 63 months for patients with at least one lymph node with KLK6 expression, i.e. KLK6(+) patients, giving a difference in mean survival time of 54 months (Table 4). Interestingly, also the actual KLK6 mRNA level was related to survival time. KLK6(+) patients with higher KLK6 mRNA levels had shorter survival times compared to KLK6(+) patients with lower KLK6 mRNA level. Actually, the group of CRC patients constituting the 90 -100 percentile of KLK6 values had a very short mean survival time of 40 months giving a difference of 76 months compared to the KLK6(-) patients population (Paper III). Survival time based on H&E classification gave a difference of 73 months (Table 4).
Cumulative survival time according to Kaplan–Meier had not been calculated for analysis based on the mRNA biomarkers CEACAM1-S/L, CEACAM6, CEACAM7-1/2 and MMP7. This has now been done using the relatively small clinical material of 51 CRC patients in Paper I (Table 5). For comparative purposes cumulative survival on the same limited clinical material was also calculated for CEA, CK20, MUC2 and KLK6. The patients

Table 4. Average survival time after surgery for biomarker(+) compared to biomarker(-) patients determined by cumulative survival according to Kaplan-Meier.

Biomarkers	Clinical cut-off	Average survival time after surgery (months)			
		Marker(-) patients	Marker(+) patients	Difference	p-value
CEA	3.67[1)]	115[2)]	44	71	<0.0001
CK20	0.0099	114	48	66	<0.0001
GCC	0.035	112	49	63	<0.0001
CDCP1	1.83	109	74	35	<0.0001
MUC2	0.0062	111	62	49	<0.0001
KLK6	-	117	54	63	<0.0001
H&E[3)]	-	115	42	73	<0.0001

1) mRNA copies/18S rRNA unit. 2) Average survival time after surgery calculated by cumulative survival analysis according to Kaplan-Meier for CRC patients (n=174 for CEA, CK20, GCC, CDCP1 and MUC2; n=166 for KLK6). Follow-up was performed January 1, 2011. 3) The same lymph nodes were used as those for biomarker mRNA analysis.

Table 5. Average survival time after surgery for biomarker(+) compared to biomarker(-) patients determined by cumulative survival according to Kaplan-Meier.

Biomarkers	Cut-off	Average survival time after surgery (months)			
		Marker(-) patients	Marker(+) patients	Difference	P-value
CEACAM1-S	0.16[1)]	110[2)]	53	57	0.005
CEACAM1-L	1.20	-	-	-	ns
CEACAM6	0.019	117	50	67	<0.0001
CEACAM7-1	0.0023	-	-	-	ns
CEACAM7-2	0.071	-	-	-	ns
MMP7	0.98	-	-	-	ns
CEA	*0.19*	*122*	*36*	*86*	*<0.0001*
CK20	*0.0064*	*122*	*36*	*86*	*<0.0001*
MUC2	*0.0012*	*113*	*59*	*54*	*0.004*
KLK6	-	*125*	*31*	*94*	*<0.0001*
H&E[3)]	-	110	29	81	<0.0001

1) mRNA copies/18S rRNA unit. 2) Average survival time after surgery calculated by cumulative survival analysis according to Kaplan-Meier for 51 CRC patients. Follow-up was performed January 1, 2011. 3) The same lymph nodes were used as those for biomarker mRNA analysis.

were divided into three groups and analyzed by univariate Cox regression analysis. The cut-off level was set to the 66.7th percentile. CEACAM1-S and CEACAM6 mRNA analysis showed significant difference of mean survival time between biomarker(-) and biomarker(+) patients. CEACAM1-L, CEACAM7-1/2 and MMP7 showed no significant difference of survival time after surgery (Table 5).
When using the cut-off calculated for this limited clinical material CEA together with CK20 showed the greatest difference of survival time after surgery between biomarker (-) patients and biomarker (+) patients, which resulted in a difference of 86 months that was highly significant for both markers. KLK6 showed the highest difference of 94 months in mean survival between KLK6(-) and KLK6(+) patients (Table 5).

4.11 Risk of recurrence for biomarker(+) patients compared to biomarker(-) patients

Univariate Cox regression analysis was used to calculate hazard risk ratios for CEA, CK20, GCC, CDCP1, MUC2 and KLK6, comparing biomarker(-) and biomarker(+) patients as defined by the clinical cut-off levels (Paper II-III). CEA mRNA analysis showed the highest hazard ratio. CEA(+) patients had a 5.1-fold greater risk for recurrence of CRC compared to CEA(-) patients. CK20, GCC, CDCP1, MUC2 and KLK6 gave lower hazard risk ratios. CEA mRNA analysis gave the same hazard ratio as H&E classification (Table 6).
Keeping in mind that the clinical material in the first study is relatively small we also calculated hazard ratios for CEACAM1-S/L, CEACAM6, CEACAM7-1/2 and MMP7. CEACAM1-S and CEACAM6 mRNA analysis showed significant hazard ratios of 3.6 and 4.9, respectively, whereas CEACAM1-L, CEACAM7-1/2 and MMP7 showed no significant hazard ratios (Table 7).
KLK6(+) patients had a 9.6–fold greater risk for recurrence than KLK6(-) patients if the risk was calculated using the same material as that included in Paper I. Also, both CEA and CK20 mRNA analysis showed high hazard ratios of 7.8 (Table 7).

4.12 Multimarker approach of biomarker mRNA analysis

As a single biomarker, none of the other 13 investigated biomarkers performed better than CEA alone for staging and prediction of prognosis of CRC patients (Papers I-III). According to CEA mRNA analysis, patients with CEA mRNA levels higher than the established clinical cut-off (>3.67 mRNA copies/18S rRNA unit) in lymph nodes have a high risk for recurrence of CRC and should therefore be offered adjuvant treatment and follow-up. However, there is a group of patients with intermediate CEA levels, CEA(int) patients, who have CEA levels higher than the highest control values but

Table 6. Risk of recurrence analyses in biomarker(+) compared to biomarker(-) surgically treated CRC patients determined by Cox proportional hazard model.

Biomarker	Clinical cut-off	Risk for recurrence	
		Hazard ratio	
CEA	3.67[1)]	5.1	**CEA**
CK20	0.0099	4.2	**CK20**
GCC	0.035	3.7	**GCC**
CDCP1	1.83	2.7	**CDCP1**
MUC2	0.0062	2.9	**MUC2**
KLK6	-	3.7	**KLK6**
H&E[2)]	-	5.1	**H&E**[2)]

1) mRNAcopies/18S rRNA unit. 2) The same lymph nodes as those used for biomarker mRNA analysis were analyzed by H&E.

Table 7. Risk of recurrence analyses in biomarker(+) compared to biomarker(-) surgically treated CRC patients determined by Cox proportional hazard model.

Biomarker	Cut-off	Risk for recurrence	
		Hazard ratio	p-value
CEACAM1-S	0.16[1)]	3.6	0.009
CEACAM1-L	1.20	-	ns
CEACAM6	0.019	4.9	<0.0001
CEACAM7-1	0.0023	-	ns
CEACAM7-2	0.071	-	ns
MMP7	0.98	-	ns
CEA	*0.19*	*7.8*	*<0.0001*
CK20	*0.0064*	*7.8*	*<0.0001*
MUC2	*0.0012*	*3.5*	*0.007*
KLK6	-	*9.6*	*<0.0001*
H&E[2)]	-	6.1	<0.0001

1) mRNA copies/18S rRNA unit. 2) The same lymph nodes as those used for biomarker mRNA analysis were analyzed by H&E.

lower levels than the clinical cut-off. It is difficult to know whether these patients, i.e. CEA(int), require further treatment or not. They definitely harbour tumor cells in their lymph nodes but probably in relatively small numbers. These patients should therefore be evaluated further using complementary biomarker mRNAs that may identify properties of tumor cells that helps to determine their aggressiveness (Papers II and III). We

believe that KLK6 may be used as a complemetary biomarker to CEA to identify particularly aggressive CRC tumor cells in lymph nodes. KLK6 mRNA was the only biomarker investigated in this thesis that had no detectable mRNA levels in normal colon and control lymph nodes. Moreover, CEA(int) patients with KLK6(+) lymph nodes had 2.8-fold greater risk for recurrence and a 53 months shorter mean survival time compared to CEA(int)KLK6(-) patients. These results suggest that presence of KLK6 mRNA in lymph nodes is associated to more aggressive tumor cells. Thus, KLK6 also adds important prognostic information to CEA mRNA analysis (Paper III).

Patients with mucinous CRC tumors have been suggested to have better prognosis than patients with non-mucinous tumors (23) and MUC2 is the major colon mucin. Therefore, we investigated whether a combination of MUC2 mRNA and CEA mRNA analysis would add prognostic information. The ratio between MUC2 and CEA was calculated. A low MUC2/CEA ratio was significantly associated with high risk for recurrence of CRC, particularly for the MUC2(+) patients. MUC2(+) patients with low MUC2/CEA ratio had shorter survival time after surgery and higher hazard risk ratio than MUC2(+) patients with high MUC2/CEA ratio. Thus, abundane of MUC2 in relation to tumor cell number is associated with better prognosis (Paper II).

To summarize, CEA, MUC2 and KLK6 mRNAs form a strong trio for detection of disseminated tumor cells in lymph nodes for staging and prediction of prognosis of CRC patients (Papers I-III).

4.13 Cytokine mRNA levels in lymph nodes from CRC patients and controls

To investigate whether we could find evidence for an ongoing immune response in lymph nodes containing disseminated tumor cells, mRNA levels of three pro-inflammatory cytokines, i.e. IFN-γ, TNF-α and IL-17A, and two downregulatory cytokines, i.e. IL-10 and TGF-β1 were analyzed in mesenteric lymph nodes from 45 CRC patients and 12 control patients (UC, Crohn's colitis and colon lipoma patients). The lymph nodes included in the study were mainly from stage III and IV CRC patients that had different levels of CEA and KLK6 mRNAs, thus representing lymph nodes with different numbers of tumor cells and of different degrees of aggressiveness. The mRNA levels of proinflammatory IFN-γ and down-regulatory IL-10 were found to be expressed at significantly higher levels in CRC nodes compared to control nodes (Paper IV). Possibly, IL-10 produced by Tregs counteracts the effect of T cell produced IFN-γ. Of note was that the expression levels of the cytokine mRNAs correlated with each other in lymph nodes from CRC patients, but not in controls.

4.14 Cytokine mRNA levels in relation to outcome of CRC patients

TGF-β1 mRNA was expressed at the highest levels of the investigated cytokines. Interestingly, the mRNA levels of this cytokine were correlated to patient's survival. Patients with high levels of TGF-β1 had a better prognosis than patients with low levels (Paper IV). However, since the significance was weak (P=0.04) and TGF-β1 was the only cytokine that correlated to outcome of CRC patients, we performed a genome wide screening with lymph nodes from CRC patients compared to nodes of controls to search for immune- or tumor cell-derived factors. IL-1 receptor antagonist (IL1RA) and macrophage migration inhibition factor (MIF) showed increased expression in CRC lymph nodes compared to controls (Paper IV).

5. DISCUSSION

In this thesis, we have explored the possibility of using real-time quantitative RT-PCR assays for carefully selected biomarker mRNAs as an alternative to routine H&E staining for detection of disseminated tumor cells in lymph nodes, staging and prediction of prognosis of CRC patients (Papers I-III).

There are several reasons to search for an alternative method to routine H&E. Importantly, only a few sections per node can in practice be subjected to H&E analysis, which means that more than 99% of lymph node volume remains unexamined. Furthermore there is an important element of subjectivity involved in traditional histopathology. It is based on the ability of the observer to differentiate between tumor cells and normal cells on basis of their morphology and organization in the tissue. This requires skills that usually take years of experience to develop and the quality of the diagnosis will vary depending on the skills of the observer. Thus, tumor cells within a lymph node may remain undetected by H&E analysis.

It is wellknown that about 25% of all stage I+II CRC patients will develop recurrent disease after surgery (130). In this work, we found that a similar proportion of stage I+II CRC patients, i.e. 20-30%, had disseminated tumor cells in their lymph nodes at the time of surgery, and that several of these patients succumbed from recurrent disease. One likely explanation for tumor recurrence in these patients is that they indeed had aggressive disseminated tumor cells in their lymph nodes already at operation that subsequently formed distant metastases that killed them. If the patients had not been under-staged by H&E they probably would have received additional teraphy and may have survived.

Reasons for chosing real-time quantitative RT-PCR assays for biomarker mRNA analysis are that: 1) the method is specific, sensitive and objective and has a large measuring range; 2) A large volume of the lymph nodes can be examined, ideally the entire node. 3) The method is amendable for automation; 4) It allows for a multimarker approach; 5) When automated it will probably be more cost-effective than conventional histopathology.

There are, however, still a number of problems, which need to be addressed before this method, could be recommended in clinical practice. The main problem is to identify the cocktail of biomarker mRNAs that would optimally identify lymph nodes with aggressive tumor cells. In this thesis we have identified CEA, KLK6 and MUC2 mRNAs as biomarkers with great potential. Hovever, one must keep an open mind, acknowledging that there may other markers that in combination may perform better. A second problem is that all our analysis is based on fresh tissues. To find mesenteric lymph nodes in fresh tissue is both difficult and time-consuming compared to finding them

in formalin fixed specimen. However, it may be possible to use formalin fixed lymph nodes for the mRNA analysis in particular for CEA mRNA because of its extreamely high expression levels. Work is under way to address this very important practical question. Moreover, we consider it to be of great importance to be able to perform mRNA analysis on formalin fixed lymph nodes since a large study where analysis of the selected biomarker mRNA cocktail is compared to conventional immunohistopathology back to back will be required in order for the alternative to become accepted.

In the search for suitable biomarkers for analysis of lymph nodes from CRC patients we have found that it was important to determine the specificity for tumor cells of the biomarker mRNA in question. Naturally, it was important to investigate whether the biomarker was expressed in immune cells. We chose to analyze a series of cell lines representing T cells, B cells, plasma cells monocytes, granulocytes, and pre-erythrocytes. In addition we analyzed peripheral blood mononuclear cells (PBMC) and activated PBMC. The latter was important since activated cells express certain molecules that are only marginaally expressed in resting cells. Compare for example MMP7 (Paper I). Another cell type that was analyzed was fibroblasts. Such cells are present almost everywhere not least in the tumor stoma. Finally we studied the expression levels in lymph nodes from patients with benign intestinal diseases. To be useful as a biomarker for detection of tumor cells in lymph nodes of CRC patients it proved to be important to show no or only marginal expression in immune cell lines, PBMC, activated PBMC and fibroblasts. High expression in lymph nodes from patients with benign intestinal diseases is not acceptable and several potential biomarkers belonging to the CEACAM family, i.e. CEACAM1-L, CEACAM7-1/2 as well as MMP7 disqualified on this basis. However, low expression levels in benigh lymph nodes are acceptable although it constitutes a problem. We found that all markers except KLK6 expressed low levels. The explanation for the low, but significant, expression of CEA and CK20 mRNA in lymph nodes from IBD patients may be that they actually contain epithelial cells released from the inflammed colon.

Another aspect is the actual expression level of the biomarker in tumor cells and the frequency of marker positive tumors. Ideally a biomarker should be expressed at very high levels and in all individual tumors. CEA was unusual in that it was expressed at very high levels and in all individual primary tumors investigated showing only moderate variation in expression levels between individual tumors. In contrast, the promising biomarker KLK7 was disqualified because of its very low expression levels in CRC tumors. Of the 14 investigated biomarker mRNAs, CEA and KLK6 mRNAs followed by CK20 and MUC2 fullfilled these criteria best. Interestingly, the tumor cell specificity of each biomarker mRNA reflected the ability of the biomarker to

discriminate between control nodes and CRC lymph nodes, and to predict prognosis of CRC patients. Thus, it is of vital importance to investigate the biomarker tumor cell specificity before starting to evaluate the expression levels of biomarker mRNAs in lymph nodes of CRC patients (Papers I-III).
First, we compared the biomarker expression levels in lymph nodes from CRC patients and controls. Second, a clinical cut-off level was determined for each biomarker using Cox regression analysis and patient survival data. Patients with biomarker mRNA level higher than clinical cut-off were considered to be biomarker(+) and patients with level lower than clinical cut-off as biomarker(-). These two groups of patients were compared in mean survival time after surgery and risk for recurrent disease. Cumulative survival was calculated by using Kaplan-Meier and hazard risk ratio by univariate Cox regression analysis. To determine which of the 14 biomarker mRNAs is the best one need to consider the size of the investigated clinical material and the lengths of the median observation time. The first material was relatively small (51 patients) but the observation time was longer (median 117 months), the second material was larger (166-174 patients) but the observation time shorter (median 75 months). The size of the clinical material determines how accurate the cut-off level can be determined. For the smaller material the cut-off level was established to the 66.7th percentile and for the larger material at the 80th percentile. It is reasonable to put more emphasis to the results obtained in the larger study. Weighing these factors together CEA and H&E would seem to have a slight edge over the others. However, a much larger study is needed to definitely rank the four top candidates i.e. CEA, KLK6, CK20 and H&E.
Another important factor that has bearing on evaluating the merits of biomarker mRNAs is the number of lymph nodes/patient that is avalible for analysis. The fact that relatively few lymph nodes, median 2, were available for <u>both</u> biomarker mRNA analysis and H&E staining puts the mRNA analysis at a disadvantage. Staging in this study was based on H&E staining of median 13 lymph nodes. Thus, only H&E negative nodes were available for mRNA analysis from 26 stage III and IV patients (Papers II and III). In further compative studies a larger number of lymph nodes should be analyzed by both types of assays.
When biomarker mRNA analysis and H&E staining was compared side by side for all the about 600 nodes examined (Figure 11), CEA mRNA analysis failed to detect two H&E(+) lymph nodes. These two nodes were not detected by any of the other biomarker mRNA either. Explanations for this might be that the the tumor cells were only located in the half of the lymph node that was subjected to H&E analysis or that there has been a mix up between different nodes at the lab. Both H&E(+) patients missed by CEA mRNA analysis had at least one H&E(-) node with CEA mRNA level clearly above clinical cut-off, indicating that the two patients could produce

"normal" epithelial CRC tumors. Thus, we believe that CRC tumor cells identified as such by morphological criteria that lack CEA and CK20 do not excist or at least are extreamely rare.

The high sensitivity of mRNA for CEA and CK20 in detecting disseminated tumor cells in lymph nodes in comparison to H&E staining stands out most clearly when the two assays are compared back to back for all about 600 nodes (Figure 11). About 57% of the H&E (-) nodes are CEA positive. The merits of in particular CEA mRNA assay over H&E staining is also demonstrated by the finding that four stage I and II patients, who succumbed from recurrent CRC were clearly biomarker mRNA positive. These patients were classified as H&E(-) and therefore not recommended for further treatment.

It can be concluded that CEA mRNA analysis has great potential for staging and prediction of outcome of CRC patients. CEA(+) patients with lymph nodes with CEA-values above clinical cut-off, have high risk of recurrent disease and should be offered adjuvant therapy and follow-up. CEA(-) patients with lymph node CEA values less than highest control cut-off are considered to be free of tumor cells in their lymph nodes and cured by surgery. However, there is the group of patients in the middle with CEA values higher than the control cut-off and clinical cut-off, i.e., CEA(int) patients. CEA(int) patients have clearly detectable tumor cells in their lymph nodes, but not high levels enough to reach the clinical cut-off. It is difficult to know wheather these patients need further treatment or not (Papers II-III). Especially for the CEA(int) group of patients there is a need for complementary biomarkers to CEA mRNA that may be able to determine the aggressiveness of the tumor cells (Papers II-II).

Patients with prominent expression of mucin in the primary tumor have been suggested to have better prognosis than patients with nonmucinous tumors (23). Our results showing that a high MUC2/CEA mRNA ratio in lymph nodes is associated with favorable prognosis particularly for MUC2(+) patients is in line with this notion (Paper II). Whether this indicates that these colon tumors generally are more differentiated and/or more of a goblet cell-type is presently not known.

We searched for a tumor progression marker for CRC that could be used as a complement to CEA mRNA analysis for identification of the tumor cell aggressiveness. This type of biomarker may hopefully identify patients with poor prognosis in need of adjuvant chemotherapy, but also patients with good prognosis, who could be spared unnecessary treatment. KLK6 mRNA was identified by genome wide screening as a potential progression marker for CRC. KLK6 was highly upregulated in CRC samples compared to controls. KLK6 mRNA was the first biomarker we have investigated so far that was not expressed in normal colon but highly expressed in primary tumors. Therefore, KLK6 mRNA is tumor cell specific in CRC. This made

KLK6 very interesting as a potential biomarker for CRC. The results showed that KLK6 mRNA expression in lymph nodes of CRC patients was associated to poor prognosis. Also the actual KLK6 level was directly related to outcome, where patients with high KLK6 levels in their nodes had the shortest survival time and highest risk of recurrence. Notably, all KLK6(+) lymph nodes also where CEA(+). CEA mRNA alone performed better in detecting tumor cells and for prediction of outcome of CRC, but by using CEA mRNA analysis for detection of tumor cell number in combination with KLK6 mRNA as a marker for tumor cell aggressiveness additional important prognostic information was obtained. Thus, KLK6 mRNA is a valuable progression marker for CRC (Paper III).

To summarize, we suggest that biomarker mRNA analysis of CEA, KLK6 and MUC2 should be used for staging of CRC patients. A similar number of lymph nodes as are routinely subjected to H&E staining should be analyzed. In evaluating the results the CEA mRNA level should be considered firstly, and secondly both the MUC2/CEA ratio and KLK6 levels as complementary factors. Table 7 summarizes the different possible biomarker mRNA combinations with staging and recommended further treatment and/or follow-up. CEA(+) patients should be offered further treatment and follow-up. CEA(-) patients are considered to have been treated by surgery. CEA(int) patients should be evaluated further in respect to KLK6 and MUC2/CEA ratio. If the CEA(int) patients are KLK6(+) and/or MUC2/CEA(low), i.e. presence of aggressive tumor cells they should be offered further treatment and follow-up. But if they are KLK6(-) and MUC2/CEA(high), indicating presence of non-aggressive tumor cells, they should be spared unnecessary adjuvant treatment and considered treated by surgery (Papers I-III). For this approach to be accepted and useful in clinical practice extended comparative studies are needed.

In this study we have focused on the expression of biomarker mRNA in regional lymph nodes of CRC patients. However, the biomarkers are also expressed in the primary tumor. To make a long story short, with the possible exception of KLK6, we did not find that the expression levels in the primary tumor of any of the biomarkers were useful for predicting outcome for CRC patients. For KLK6 mRNA there was a tendency that levels above median KLK6 expression in primary tumor was associated to poor outcome and short survival, but the difference was not significant. Thus, it seems advisable to continue analysis of the tumor cells disseminated to the lymph nodes.

Neither biomarker analysis nor routine histopatology of lymph nodes detected all patients that succumbed from CRC within the follow-up period. There could be a number of reasons for this. Too few lymph nodes were analysed by either method resulting in understaging and inadequate

supplementary teraphy and follow-up. Another possibility is that tumor cells spread via the hematogenous route.

Table 7. Biomarker expression level combinations with stage and recommended further treatment after surgery.

Biomarker levels			Stage	Explanation	Recommended treatment
CEA	**MUC2/CEA ratio**	**KLK6**			
+	high or low	+ or -	III- IV	High load of tumor cells	Adjuvant chemotherapy and follow-up
int	low	+	III-IV	Presence of aggressive tumor cells with low levels of favorable MUC2	
int	low	-	III-IV	Presence of non-aggressive tumor cells with low levels of favorable MUC2	
int	high	+	III-IV	Presence of aggressive tumor cells with high levels of favorable MUC2	
int	high	-	III-IV	Presence of non-aggressive tumor cells with high levels of favorable MUC2	No further treatment
-	-	-	I-II	No presence of tumor cells	

Finally, we searched for evidence for an ongoing immune response in tumor cell infiltrated lymph nodes by determining levels of the proinflammatory cytokine mRNAs IFN-γ, TNF-α and IL-17A and the downregulatory cytokines IL-10 and TGF-β1. The cytokine mRNA levels in CRC lymph nodes were also analyzed for their utility of predicting outcome. The lymph nodes included in the study had different number of tumor cells (CEA mRNA levels) with different degree of aggressiveness (KLK6 mRNA levels; Paper IV). The study showed that there indeed was an ongoing functional immune response probably against the infiltrating tumor cells in the lymph node that involved increased levels of IFN-γ and IL-10 mRNAs in CRC nodes compared to control nodes. IFN-γ producing T cells and Tregs producing IL-10 have probably been activated.

TGF-β1 mRNA was the only cytokine that were significantly decreased in lymph nodes from those CRC patients that have died from CRC compared to the patients that were alive without recurrent disease. This result suggests that high TGF-β1 mRNA levels are protective, which might be due to the

antiproliferative action of this cytokine. If true, it remains to be established which cell types provide the protective effect since TGF-b1 is produced by a number of different cell types (Paper IV).

6. CONCLUSIONS

- Biomarker mRNA analysis shows great promise as an alternative method for detection of disseminated tumor cells in regional lymph nodes of patients with colorectal carcinoma. Using real-time quantitative RT-PCR with a specific probe and a RNA copy standard, the method is objective, specific, highly sensitive, and allows comparison of expression levels between biomarkers.

- As a single biomarker carcinoembryonic antigen (CEA) mRNA performed best of 14 investigated biomarkers for detection of disseminated tumor cells in lymph nodes and for staging and prediction of prognosis of CRC patients. The hazard risk ratio for CEA mRNA(+) patients was 5.1-7.8 and the mean difference in mean survival time after surgery between CEA mRNA(+) and CEAmRNA(-) patients was 71-86 months. A number of understaged patients were detected.

- Mucin2 (MUC2) seems to have a protective function. A high MUC2 mRNA/CEA mRNA ratio in CRC lymph nodes was associated with a favorable prognosis.

- Tissue kallikrein 6 (KLK6) mRNA was found to be ectopically expressed in tumor cells from CRC patients. Expression of KLK6 mRNA in lymph nodes of CRC patients was associated to poor prognosis. Noteably, the actual level was of importance for outcome. The higher the KLK6 mRNA levels the greater the risk of recurrence. At the 90^{th} percentile the hazards risk ratio for KLK6(+) patients was 5.6.

- Lymph node CEA, MUC2 and KLK6 mRNAs form a strong multi-marker "trio" for staging and prediction of prognosis of CRC patients.

- Immune cells in the lymph nodes are responding to the presence of tumor cells by production of IFN-γ and IL-10, suggesting involvement of different T cell subsets. Elevated TGF-β1 levels were weakly correlated with survival, suggesting protection by the antiproliferative effect of TGF-β1 in sporadic cases.

7. ACKNOWLEDGEMENT

Det finns många personer jag vill tacka så hjärtligt på olika sätt efter åren som doktorand fram till min disputation!

Först och främst vill jag tacka mina handledare **Sten Hammarström** och **Marie-Louise Hammarström** att ni inspirerade mig att söka examensarbetet hos er. Ni gav mig chansen att få vara en del i CRC-projektet. Jag är oerhört tacksam för att ni alltid har trott på mig, anförtrott mig stort ansvar och med stort engagemang givit mig möjlighet till den stora utveckling och mål jag nått idag! Ni är fantastiska medmänniskor, positiva och entusiastiska. Tack för att ni delat med er av ovärderliga kunskaper och passion till forskningen. Ni är en stor inspirationskälla och kan inte få någon att tycka annat än att det är roligt med forskning! Sist men inte minst vill jag tacka er för den ödmjukhet och för att ni med stort hjärta alltid brytt er om mig även utanför arbetet.

Anne Israelsson som från första dag noggrant och med ett glatt humör instruerat mig på lab! Du har verkligen varit en positiv och frisk fläkt under dessa år och med dina tekniska och laborativa kunskaper hjälpt mig massor. Alltid snäll och omtänksam.

Gudrun Lindmark och **Åke Öberg** för ett mycket trevligt, lärorikt och roligt samarbete samt ert tappra insamlande av kliniskt material till projektet. Ert engagemang i projektet och mig.

Richard Palmqvist för givande samarbete och positiv inspirerande forskaranda!

Lena Hallström-Nyhlén som alltid hjälpt mig med diverse administrativa ärenden. Du har besparat mig mycket tid och besvär. Alltid så snäll, glad och omtänksam. Har uppskattat alla våra pratstunder om väder och vind, och inte minst vårt gemensamma intresse djur.

Mia Sundström för din otroliga energi, ditt glada humör och din förmåga att inte vara rädd för att riva tag i saker. Tack för att du funnits vid min sida som en kär vän och ställt upp i både roliga och tråkiga stunder. Jag är verkligen glad att jag hittade en vän som du!

Veronika Sjöberg för roligt och givande samarbete på olika sätt forskningsmässigt men även alla trevliga stunder som god vän utanför jobbet.

Ann-Christine Stenqvist för alla givande jobbrelaterade diskussioner men också som en fin vän.

Tack till alla som arbetar/arbetat på immunologen: **Marianne Sjöstedt, David Eriksson, Theres Lindgren, Aziz Bitar, Greg Pietz, Viqar Banday, Rahda Thyagarajan, Maria Hedberg, Kristina Lejon, Torgny Stigbrand, Vladimir Baranov, Rifat Ekici, Gangwei Ou** och **Arman Rahman** för alla givande och roliga diskussioner om både forskning och annat, och för att ni alla tillsammans skapat en sådan trivsam, rolig och stimulerande arbetsmiljö!

Jag vill tacka alla ovanstående för den positiva och roliga inblick jag har fått i forskning under dessa år. Jag känner verkligen att jag hittat rätt och är laddad för vad framtiden har att erbjuda inom forskning!

♥ **Mina föräldrar** ♥ som alltid stått vid min sida i vått och torrt. Hjälpt mig upp när jag varit låg och delat min glädje när livet varit på topp. Ni är verkligen helt fantastiska! Vem kunde tro att konstant ihopplåstrade och/eller ihopsydda vilda Lina 0-15 år faktiskt i vuxen ålder skulle lyckas skada sig lite mer sällan och istället använda energin till studier och cancerforskning. Tack för att ni alltid trott på mig, låtit mig vara den jag är och låtit mig prova mina vingar!

Min storasyster **Malin** med sambo **John** som alltid funnits vid min sida trots långt avstånd. Tack för all omtanke, kärlek, trygghet och glädje ni gett mig. Man kan inte önska sig en bättre storasyster än du! Och självklart stort tack till era trollungar **Malcom** och **Tuva**, som förgyller allas våra liv! ♥

Malin Cederlöf en fantastisk och god vän! Så glad att vi lärde känna varandra och tacksam för att du alltid finns vid min sida. Alltid lika ärlig, rak, omtänksam, rolig, snäll och glad! Jag behöver bara höra din röst så blir jag glad! Du ger mig massor energi varje dag! Klakt giver!

Jenny & Mattias Lundberg som fått mig att känna mig som en i familjen med allt vad det innebär. Ni har varit ett ovärderligt stöd på alla sätt och vis och min trygghet här i Umeå. Tack för alla givande pratstunder, prommisar, myskvällar och annat roligt! Jag och Tokiga Tiger ser fram emot att följa Jeppe i sitt sökande efter nöten!!

Jenny Österberg som alltid trott på mig och mer eller mindre övertalade mig att doktorera. Det gjorde du helt rätt i! Det började i en korridor på Carlshem för 11 år sedan och ännu är du en av mina allra bästa vänner som stöttar mig i ur och skur!

Min vän **Adrian Elliott** för åtskilliga års vänskap, roliga stunder och för att du alltid tror på mig!

Liselotte Ohlsson för din vänskap och alla våra roliga äventyr med och utan de fyrbenta i sällskap!

Karin & Peter Andersson för er vänskap och stöd, roligt umgänge både i skog, mark och vid utställningsringen med eller utan hundar!

Min lilla familj under studietiden: **Erica Nygren, Jessica Ilshagen** och **Shandy Wikman** för er vänskap och omtanke!

Robin Landström som senaste tiden med din värme, omtanke och ditt glada humör gett mig massor av energi och hopp som fått mig att orka med denna intensiva höst!

Sist men absolut inte minst min trogna och lojala bästa kompis ♥ **Tiger** ♥, som varit med från början till slut och sett till att jag blivit ordentligt motionerad, givit mig sällskap och massor av glädje varje dag!

TACK!
/Lina

This work was financially supported by grants from **Stig and Ragna Gorthon's Foundation** in Helsingborg (G Lindmark), **the Gunnar Nilsson's Cancer Foundation** (G Lindmark), **the Swedish Cancer Foundation** (S Hammarström), **the Norrland's Cancer Foundation** (M-L Hammarström), **the County of Västerbotten** (Å Öberg), and **a PhD-study support grant from the Medical Faculty of Umeå University** to L Ohlsson.

8. REFERENCES

1. Rubin E: Essential Pathology, 3rd Edition, 2001, Lippincott Williams & Wilkins, Baltimore & Philadelphia, USA.

2. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646-74.

3. Soltanian S, Matin M. Cancer stem cells and cancer therapy. Tumor Biol 2011;32:425-40.

4. Medema RH, Macurek L. Checkpoint control and cancer. Oncogene 2011 doi: 10.1038/onc.2011.451.

5. Pylayeva-Gupta Y, Grabocka E, Bar-Sagi D. RAS oncogenes: weaving a tumorigenic web. Nat Rev Cancer 2011;11:761-74.

6. Wolfer A, Ramaswamy S. MYC and metastasis. Cancer Res 2011;71:2034-37.

7. Khoury M, Bourdon JC. P53 isoforms: an intracellular microprocessor? Genes & Cancer 2011;2:453-65.

8. Olsson M, Zhivotovsky B. Caspases and cancer. Cell death and differentiation 2001;18:1441-49.

9. van Zijl F, Krupitza G, Mikulits W. Initial steps of metastasis: cell invasion and endothelial transmigration. Mutat Res 2011;728:23-34.

10. Kim R, Emi M, Tanabe K. Cancer immunoediting from immune surveillance to immune escape. Immunology 2007;121:1-14.

11. Kindt T, Goldsby R, B Osborne: Immunology, 6th Edition, 2007, W. H. Freeman and Company, New York, USA.

12. Yang M. A current global view of environmental and occupational cancers. J Environ Sci Health C Environ Carcinog Ecotoxicol Rev 2011;29:223-49.

13. Willard-Mack CL. Normal structure, function, and histology of lymph nodes. Toxicol Pathol 2006;34:409-24.

14. Kierszenbaum AL: Histology and cell biology-an introduction to pathology, 2002, Mosby, St. Louis, Missouri, USA.

15. Bingham S. Meat, starch, and nonstarch polysaccharides and large bowel cancer. Am J Clin Nutr 1988;48:762-67.

16. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, et al. Enterotypes of the human gut microbiome. Nature 2011;473:174-80.

17. Walter J, Ley RE. The Human Gut Microbiome: Ecology and Recent Evolutionary Changes. Annu Rev Microbiol 2010 Sep 28. Epub ahead of print.

18. Schiffrin EJ, Blum S. Interactions between the microbiota and the intestinal mucosa. Eur J Clin Nutr 2002;3:60-64.

19. Ou G, Hedberg M, Hörstedt P, Baranov V, Forsberg G, Drobni M, et al. Proximal small intestinal microbiota and identification of rod-shaped bacteria associated with childhood celiac disease. Am J Gastroenterol 2009;104:3058-67.

20. Hammarström S, Baranov V. Is there a role for CEA in innate immunity in the colon? Trends Microbiol 2001;9:119-25.

21. Baranov V, Hammarström S. Carcinoembryonic antigen (CEA) and CEA-related cell adhesion molecule 1 (CEACAM1), apically expressed on human colonic M cells, are potential receptors for microbial adhesion. Histochem Cell Biol 2004;121:83-89.

22. Simons BD, Clevers H. Stem cell self-renewal in intestinal crypt. Exp Cell Res 2011 Jul 20. Epub ahead of print.

23. Byrd JC, Bresalier RS. Mucins and mucin binding proteins in colorectal cancer. Cancer Metastasis Rev 2004;23:77-99.

24. Pitman RS, Blumberg RS. First line of defense: the role of the intestinal epithelium as an active component of the mucosal immune system. J Gastroenterol 2000;35:805-14.

25. Fahlgren A, Hammarström S, Danielsson A, Hammarström ML. Increased expression of antimicrobial peptides and lysozyme in colonic epithelial cells of patients with ulcerative colitis. Clin Exp Immunol 2003;131:90-101.

26. Lundqvist C, Baranov V, Hammarström S, Athlin L, Hammarström ML. Intra-epithelial lymphocytes. Evidence for regional specialization and extrathymic T cell maturation in the human gut epithelium. Int Immunol 1995;7:1473-87.

27. Lundqvist C, Melgar S, Yeung MM, Hammarström S, Hammarström ML. Intraepithelial lymphocytes in human gut have lytic potential and a cytokine profile that suggest T helper 1 and cytotoxic functions. J Immunol 1996;157:1926-34.

28. Brandtzaeg P. Mucosal immunity: induction, dissemination, and effector functions. Scand J Immunol 2009;70:505-15.

29. Garcia M, Jemal A, Ward EM, Center MM, Hao Y, Siegel RL, et al. American Cancer society: Global cancer facts & figures, Atlanta, USA 2007.

30. The National Board of Health and Welfare. Statistics-Health and Diseases. Cancer incidence in Sweden 2009.

31. The National Board of Health and Welfare. Statistics-Health and Diseases. Causes of death in Sweden 2007.

32. Mattar MC, Lough D, Pishvaian MJ, Charabaty A. Current management of inflammatory bowel disease and colorectal cancer. Gastrointest Cancer Res 2011;4:53-61.

33. Wei EK, Giovannucci E, Wu K, Rosner B, Fuchs CS, Willett WC,et al. Comparison of risk factors for colon and rectal cancer. Int J Cancer 2004;108:433-42.

34. Harrison S, Benziger H. The molecular biology of colorectal carcinoma and its implications: a review. Surgeon 2011;9:200-10.

35. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med 2003;348:919.

36. Nielsen HJ, Jakobsen KV, Christensen IJ, Brünner N. Screening for colorectal cancer: possible improvements by risk assessment evaluation? Scand J Gastroenterol 2011 Aug 19. Epub ahead of print.

37. Dukes CE, Bussey HJ. The spread of rectal cancer and its effect on prognosis. Br J Cancer 1958;12:309-20.

38. AJCC. Cancer staging manual, 7th edition, New York, Springer, 2010.

39. Obrocea FL, Sajin M, Marinescu EC, Stoica D. Colorectal cancer and the 7th revision of the TNM staging system: review of changes and suggestions for uniform pathologic reporting. Rom J Morphol Embryol 2011;52:537-44.

40. Zlobec I, Lugli A. Prognostic and predictive factors in colorectal cancer. J Clin Pathol 2008;61:561-69.

41. Sturgeon C. Practice guidelines for tumor marker use in the clinic. Clin Chem 2002;48:1151-59.

42. Duffy MJ, van Dalen A, Haglund C, Hansson L, Holinski-Feder E, Klapdor R, et al. Tumour markers in colorectal cancer: European Group on Tumour Markers (EGTM) guidelines for clinical use. Eur J Cancer 2007;43:1348-60.

43. Park ET, Oh HK, Gum JR, Crawley SC, Kakar S, Engel J, et al. HATHI expression in mucinous cancers of the colorectum and related lesions. Clin Cancer Res 2006;12:5403-10.

44. Meyerhardt JA, Mayer RJ. Systemic therapy for colorectal cancer. N Engl J Med 2005; 352: 476-87.

45. Bedford JS, Dewey WC. Radiation Research Society. 1952-2002. Historical and current highlights in radiation biology: has anything important been learned by irradiating cells? Radiat Res 2002;158:251-91.

46. Weitz J, Koch M, Debus J, Höhler T, Galle PR, Büchler MW. Colorectal cancer. Lancet 2005;365:153-65.

47. Chrisp P, McTavish D. Levamisole/fluorouracil. A review of their pharmacology and adjuvant therapeutic use in colorectal cancer. Drugs Aging 1991;1:317-37.

48. O'Connell J, Maggard M, Ko C. Colon cancer survival rates with the new american joint comittee on cancer sixth edition staging. J of the Nat Canc Inst 2004;96:1420-25.

49. Damber L, Nathanaelsson L, Andersson G. Nationella Kvalitetsregistret för Colon Cancer 2008. (National Registry for Colon Cancer, 2008). Oncologic Center, University Hospital of Norrland, Umeå, Sweden, 2010.

50. Fretwell VL, Ang CW, Tweedle EM, Rooney PS. The impact of lymph node yield on Duke's B and C colorectal cancer survival. Colorectal Dis 2010;12:995-1000.

51. Moertel GC, Fleming TR, MacDonald JS, Haller DG, Laurie JA, Tangen CM, et al. Fluorouracil plus levamisole as effective adjuvant therapy after resection of stage III colon carcinoma: a final report. Ann Intern Med 1995;122:321-26.

52. Flemming ID, Cooper JS, Henson DE, Hutter RVP, Kennedy BJ, Murphy GP, et al, eds. American Joint Committe on Cancer (1997) Cancer staging manual. 5th edition. Lippincott-Raven: Philidelphia.

53. Nicastri DG, Doucette JT, Godfrey TE, Hughes SJ. Is occult lymph node disease in colorectal cancer patients clinically significant? A review of the relevant literature. J Mol Diagn 2007;9:563-71.

54. Pantel K, Cote RJ, Fodstad O. Detection and clinical importance of micrometastatic disease. J Natl Cancer Inst 1999;91:1113-24.

55. Baxter NN, Virnig DJ, Rothenberger DA, Morris AM, Jessurun J, Virnig BA. Lymph node evaluation in colorectal cancer patients: a population-based study. J Natl Cancer Inst 2005; 97:219-25.

56. Haince J-F, Houde M, Beaudry G, L'espérance S, Garon G, Desaulniers M, et al. Comparison of histopathology and RT-qPCR amplification of guanylyl cyclase C for detection of colon cancer metastases in lymph nodes. J Clin Pathol 2010;63:530-37.

57. Öberg Å, Lindmark G, Israelsson A, Hammarström S, Hammarström M-L. Detection of occult tumour cells in lymph nodes of colorectal cancer patients using real-time quantitative RT-PCR for CEA and CK20 mRNAs. Int J Cancer 2004;111:101-110.

58. Nordgård O, Oltedal S, Kørner H, Aasprong OG, Tjensvoll K, Gilje B, et al. Quantitative RT-PCR detection of tumor cells in sentinel lymph nodes isolated from colon cancer patients with an ex vivo approach. Ann Surg 2009;249:602-07.

59. Schulz S, Hyslop T, Haaf J, Bonaccorso C, Nielsen K, Witek ME, et al. A validated quantitative assay to detect occult micrometastases by reverse transcriptase-polymerase chain reaction of guanylyl cyclase C in patients with colorectal cancer. Clin Cancer Res 2006;12:4545-52.

60. Sargent DJ, Resnick MB, Meyers MO, Goldar-Najafi A, Clancy T, Gill S, et al. Evaluation of Guanylyl Cyclase C Lymph Node Status for Colon Cancer Staging and Prognosis. Ann Surg Oncol 2011 May 1. Epub ahead of print.

61. Schirbel A, Fiocchi C. Inflammatory bowel disease: Established and evolving considerations on its etiopathogenesis and therapy. J Dig Dis 2010;11:266-76.

62. Ullman TA, Itzkowitz SH. Intestinal inflammation and cancer. Gastroenterology 2011;140:1807-16.

63. Gold P, Freedman SO. Specific carcinoembryonic antigens of the human digestive system. J Exp Med 1965;122:467-81.

64. Hammarström S. The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues. Cancer Biology 1999;9:67-81.

65. Horst A K, Wagener C (2004) CEA-related CAMs. In handbook of Experimental Pharmacology, Cell Adhesion Behrens J and Nelson WJ (eds), Vol. 165, pp 283-341. Heidelberg: Springer-Verlag.

66. Benchimol S, Fuks A, Jothy S, Beauchemin N, Shirota K, Stanners CP. Carcinoembryonic antigen, a human tumor marker, functions as an intercellular adhesion molecule. Cell 1989;57:327-34.

67. Muenzner P, Rohde M, Kneitz S, Hauck CR. CEACAM engagement by human pathogens enhances cell adhesion and counteracts bacteria-induced detachment of epithelial cells. J Cell Biol 2005;170:825-36.

68. Nouvion AL, Oubaha M, Leblanc S, Davis EC, Jastrow H, Kammerer R, et al. CEACAM1: a key regulator of vascular permeability. J Cell Sci 2010;123:4221-30.

69. Gerstel D, Wegwitz F, Jannasch K, Ludewig P, Scheike K, Alves F, et al. CEACAM1 creates a pro-angiogenic tumor microenvironment that supports tumor vessel maturation. Oncogene 2011;30:4275-88.

70. Huang J, Hardy JD, Sun Y, Shively JE. Essential role of biliary glycoprotein (CD66a) in morphogenesis of the human mammary epithelial cell line MCF10F. J Cell Sci 1999;112:4193-205.

71. Kirshner J, Chen CJ, Liu P, Huang J, Shively JE. CEACAM1-4S, a cell-cell adhesion molecule, mediates apoptosis and reverts mammary carcinoma cells to a normal morphogenic phenotype in a 3D culture. Proc Natl Acad Sci U S A 2003;100:521-26.

72. Nagaishi T, Chen Z, Chen L, Iijima H, Nakajima A, Blumberg RS. CEACAM1 and the regulation of mucosal inflammation. Mucosal Immunol 2008;1:S39-42.

73. Gray-Owen SD, Blumberg RS. CEACAM1: contact-dependent control of immunity. Nat Rev Immunol 2006;6:433-46.

74. Neumaier M, Paululat S, Chan A, Matthaes P, Wagener C. Biliary glycoprotein, a potential human cell adhesion molecule, is down-regulated in colorectal carcinomas. Proc Natl Acad Sci U S A 1993;90:10744-48.

75. Kinugasa T, Kuroki M, Takeo H, Matsuo Y, Ohshima K, Yamashita Y, et al. Expression of four CEA family antigens (CEA, NCA, BGP and CGM2) in normal and cancerous gastric epithelial cells: up-regulation of BGP and CGM2 in carcinomas. Int J Cancer 1998;76:148-53.

76. Lutterbuese R, Raum T, Kischel R, Lutterbuese P, Schlereth B, Schaller E, et al. Potent control of tumor growth by CEA/CD3-bispecific single-chain antibody constructs that are not competitively inhibited by soluble CEA. J Immunother 2009;32:341–52.

77. Osada T, Hsu D, Hammond S, Hobeika A, Devi G, Clay TM, et al. Metastatic colorectal cancer cells from patients previously treated with chemotherapy are sensitive to T-cell killing mediated by CEA/CD3-bispecific T-cell-engaging BiTE antibody. Br J Cancer 2010;102:124-33.

78. Parkhurst MR, Yang JC, Langan RC, Dudley ME, Nathan DA, Feldman SA, et al. T cells targeting carcinoembryonic antigen can mediate regression of metastatic colorectal cancer but induce severe transient colitis. Mol Ther 2011;19:620-26.

79. Jantscheff P, Terracciano L, Lowy A, Glatz-Krieger K, Grunert F, Micheel B, et al. Expression of CEACAM6 in resectable colorectal cancer: a factor of independent prognostic significance. J Clin Oncol 2003;21:3638-46.

80. Blumenthal RD, Leon E, Hansen HJ, Goldenberg DM.Expression patterns of CEACAM5 and CEACAM6 in primary and metastatic cancers. BMC Cancer 2007:7:2.

81. Blumenthal RD, Hansen HJ, Goldenberg DM. Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic Antigen). Cancer Res 2005;65:8809-17.

82. Borgono CA and EP Diamandis. The emerging roles of human tissue kallikreins in cancer. Nat Rev Cancer 2004;4:876-90.

83. Paliouras M, Borgono C and EP Diamandis. Human tissue kallikreins: The cancer biomarker family. Cancer Lett 2007;249:61-79.

84. Shaw JL, Diamandis EP. Distribution of 15 human kallikreins in tissues and biological fluids. Clin Chem 2007;53:1423-32.

85. Scarisbrick IA, Blaber SI, Tingling JT, Rodriguez M, Blaber M, Christophi GP. Potential scope of action of tissue kallikreins in CNS immune-mediated disease. J Neuroimmunol 2006;178:167-76.

86. Diamandis EP, Yousef GM, Petraki C, Soosaipillai AR. Human kallikrein 6 as a biomarker of alzheimer's disease. Clin Biochem 2000;33:663-67.

87. Ogawa K, Yamada T, Tsujioka Y, Taguchi J, Takahashi M, Tsuboi Y, et al. Localization of a novel type trypsin-like serine protease, neurosin, in brain tissues of Alzheimer's disease and Parkinson's disease. Psychiatry Clin Neurosci 2000;54:419-26.

88. Magklara A, Mellati AA, Wasney GA, Little SP, Sotiropoulou G, Becker GW, et al. Characterization of the enzymatic activity of human kallikrein 6: Autoactivation, substrate

specificity, and regulation by inhibitors. Biochem Biophys Res Commun 2003;307:948-55.

89. Kim JT, Song EY, Chung KS, Kang MA, Kim JW, Kim SJ, et al. Up-regulation and clinical significance of serine protease kallikrein 6 in colon cancer. Cancer 2011;117:2608-19.

90. Klucky B, Mueller R, Vogt I, Teurich S, Hartenstein B, Breuhahn K,et al. Kallikrein 6 induces E-cadherin shedding and promotes cell proliferation, migration, and invasion. Cancer Res 2007;67:8198-206.

91. Diamandis EP, Yousef GM, Soosaipillai AR, Bunting P. Human kallikrein 6 (zyme/protease M/neurosin): a new serum biomarker of ovarian carcinoma. Clin Biochem 2000;33:579-83.

92. Ogawa K, Utsunomiya T, Mimori K, Tanaka F, Inoue H, Nagahara H, et al. Clinical significance of human kallikrein gene 6 messenger RNA expression in colorectal cancer. Clin Cancer Res 2005;11:2889-93.

93. Henkhaus RS, Gerner EW, Ignatenko NA. Kallikrein 6 is a mediator of K-RAS-dependent migration of colon carcinoma cells. Biol Chem 2008;389:757-64.

94. Ekholm IE, Brattsand M, Egelrud T. Stratum corneum tryptic enzyme in normal epidermis: a missing link in the desquamation process? J Invest Dermatol 2000;114:56-63.

95. Ramani VC, Haun RS. The extracellular matrix protein fibronectin is a substrate for kallikrein 7. Biochem Biophys Res Commun 2008;369:1169-73.

96. Johnson SK, Ramani VC, Hennings L, Haun RS. Kallikrein 7 enhances pancreatic cancer cell invasion by shedding E-cadherin. Cancer 2007;109:1811-20.

97. Talieri M, Diamandis EP, Gourgiotis D, Mathioudaki K, Scorilas A. Expression analysis of the human kallikrein 7 (KLK7) in breast tumors: a new potential biomarker for prognosis of breast carcinoma. Thromb Haemost 2004;91:180-86.

98. Talieri M, Mathioudaki K, Prezas P, Alexopoulou DK, Diamandis EP, Xynopoulos D, et al. Clinical significance of kallikrein-related peptidase 7 (KLK7) in colorectal cancer. Thromb Haemost 2009;101:741-7.

99. Inoue Y, Yokobori T, Yokoe T, Toiyama Y, Miki C, Mimori K, et al. Clinical significance of human kallikrein7 gene expression in colorectal cancer. Ann Surg Oncol 2010;17:3037-42.

100. Diamandis EP, Borgoño CA, Scorilas A, Harbeck N, Dorn J, Schmitt M. Human kallikrein 11: an indicator of favorable prognosis in ovarian cancer patients. Clin Biochem 2004;37:823-29.

101. Stavropoulou P, Gregorakis AK, Plebani M, Scorilas A. Expression analysis and prognostic significance of human kallikrein 11 in prostate cancer. Clin Chim Acta 2005;357:190-95.

102. Yu X, Tang HY, Li XR, He XW, Xiang KM. Over-expression of human kallikrein 11 is associated with poor prognosis in patients with low rectal carcinoma. Med Oncol 2010;27:40-44.

103. Derrien M, van Passel MWJ, van de Bovenkamp JHB, Schipper RG, de Vos WM, Dekker J. Mucin-bacterial interactions in the human oral cavity and digestive tract. Gut Microbes 2010;1:254-68.

104. Jacques C, Moura de Aquino A and Ramos-e-Silva M. Cytokeratins and dermatology. Skinmed 2005;11:354-60.

105. Wildi S, Kleeff J, Maruyama H, Maurer CA, Friess H, Büchler MW, et al. Characterization of cytokeratin 20 expression in pancreatic and colorectal cancer. Clinical Cancer Research 1999;5:2840-47.

106. Lucas KA, Pitari GM, Kazerounian S, Ruiz-Stewart I, Park J, Schulz S, et al. Guanylyl cyclases and signaling by cyclic GMP. Pharmacol Rev 2000;52:375-414.

107. Frick G S, Pitari G M, Weinberg D, Hyslop T, Schulz S, Waldman SA. Guanylyl cyclase C: a molecular marker for staging and postoperative surveillance of patients with colorectal cancer. The British library, Expert Rev Mol Diagn 2005;5:701-13.

108. Pitari GM, Li P, Lin JE, Zuzga D, Gibbons AV, Snook AE, et al. The paracrine hormone hypothesis of colorectal cancer. Clin Pharmacol Ther 2007;82:441-47.

109. Waldman SA, Hyslop T, Schulz S, Barkun A, Nielsen K, Haaf J, et al. Association of GUCY2C expression in lymph nodes with time to recurrence and disease-free survival in pN0 colorectal cancer. J Am Med Assoc 2009;301:745-52.

110. Carlson M. Previstage™ GCC Colorectal Cancer Staging Test A new molecular test to identify lymph node metastases and provide more accurate information about the stage of patients with colorectal cancer. Mol Diag Ther 2009;13:11-14.

111. Parsons SL, Watson SA, Brown PD, Collins HM, Steele RJ. Matrix metalloproteinases. Br J Surg 1997;84:160-66.

112. Adachi Y, Yamamoto H, Itoh F, Arimura Y, Nishi M, Endo T, et al. Clinicopathologic and prognostic significance of matrilysin expression at the invasive front in human colorectal cancers. Int J Cancer 2001;95:290-94.

113. Fang YJ, Lu ZH, Wang GQ, Pan ZZ, Zhou ZW, Yun JP, et al. Elevated expressions of MMP7, TROP2, and survivin are associated with survival, disease recurrence, and liver metastasis of colon cancer. Int J Colorectal Dis 2009;24:875-84.

114. Büring HJ, Kuci S, Conze T, Rathke G, Bartolović K, Grünebach F, et al. CDCP1 identifies a broad spectrum of normal and malignant stem/progenitor cell subset of hematopoietic and nonhematopoietic origin. Stem cells 2004;22:334-43.

115. Scherl-Mostageer M, Sommergruber W, Abseher R, Hauptmann R, Ambros P, Schweifer N. Identification of a novel gene, CDCP1, overexpressed in human colorectal cancer. Oncogene 2001;20:4402-08.

116. Deryugina EI, Conn EM, Wortmann A, Partridge JJ, Kupriyanova TA, Ardi VC, et al. Functional role of cell surface CUB domain-containing protein 1 in tumor cell dissemination. Mol Cancer Res 2009;7:1197-211.

117. Wortman A, He Y, Deryugina EI, Quigley JP, Hooper JD. The cell surface glycoprotein CDCP1 in cancer – insights, opportunities, and challenges. IUBMB Life 2009;61:723-30.

118. Paul G, Khare V, Gasche C. Inflamed gut mucosa: downstream of interleukin-10. Eur J Clin Invest 2011 doi: 10.1111/j.1365-2362.2011.02552.x.

119. Akdis M, Burgler S, Crameri R, Eiwegger T, Fujita H, Gomez E, et al. Interleukins, from 1 to 37, and interferon-γ: receptors, functions, and roles in diseases. J Allergy Clin Immunol 2011;127:701-21.

120. Iwakura Y, Ishigame H, Saijo S, Nakae S. Functional specialization of interleukin-17 family members. Immunity 2011;34:149-62.

121. Aggarwal BB, Gupta SC, Kim JH. Historical perspectives on tumor necrosis factor and its superfamily: twenty-five years later, a golden journey. Blood. 2011 Nov 3. Epub ahead of print.

122. Grimm M, Lazariotou M, Kircher S, Höfelmayr A, Germer CT, von Rahden BH, et al. Tumor necrosis factor-α is associated with positive lymph node status in patients with recurrence of colorectal cancer-indications for anti-TNF-α agents in cancer treatment. Cell Oncol 2011;34:315-26.

123. Yoshimura A, Wakabayashi Y, Mori T. Cellular and molecular basis for the regulation of inflammation by TGF-beta. J Biochem 2010;147:781-92.

124. Yang L, Pang Y, Moses HL. TGF-beta and immune cells: an important regulatory axis in the tumor microenvironment and progression. Trends Immunol 2010;31:220-27.

125. Chomczynski P and Sacchi N. Single-step method of RNA isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform extraction. Analyt Biochem 1987;162:156-59.

126. Ohlsson L, Israelsson A, Öberg Å, Hammarström M-L, Lindmark G, Hammarström S. Detection of tumor cells in lymph nodes of colon cancer patients using real-time quantitative reverse transcription-polymerase chain reaction. Hayat MA: Methods of cancer diagnosis, therapy, and prognosis, 2009;4:257-270, Springer, New York, USA.

127. TaqMan EZ RT-PCR Kit manual, Applied Biosystems, Part Number 402877, Foster City, CA, USA.

128. Bas A, Forsberg G, Hammarström S, Hammarström M-L. Utility of the housekeeping genes 18S rRNA, β-actin and glyceraldehyde-3-phosphate-dehydrogenase for normalization in real-time quantitative reverse transcriptase-polymerase chain reaction analysis of gene expression in human T lymphocytes. Scand J Immunol 2004;59:566-73.

129. Kuhn K, Baker S, Chudin E, Lieu M-H, Oeser S, Bennet H, et al. A novel, high-performance random array platform for quantitative gene expression profiling. Genome Res 2004;14:2347-56.

130. Chang G, Rodriguez-Bigas M, Skibber J, Moyer V. Lymph node evaluation and survival after curative resection of colon cancer: systematic review. J Natl Cancer Inst 2007;99:433-41.