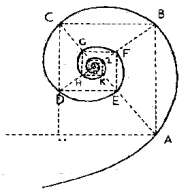
UNIVERSITÀ DEGLI STUDI DI MILANO

Facoltà di Medicina e Chirurgia
Dipartimento di Scienze e Tecnologie Biomediche
Settore scientifico disciplinare BIO/10



SCUOLA DI DOTTORATO DI RICERCA IN MEDICINA MOLECOLARE
Curriculum di Oncologia Molecolare
Ciclo XXIII

TESI DI DOTTORATO

Mesenchymal features mediated by *Twist1* in Colorectal cancer cells and microenvironment

Il dottorando
Dr. Giuseppe Celesti
Matr.R07705

Il Tutore
Prof. Massimo Roncalli
Il Correlatore
Dr. Luigi Laghi

Il Direttore
Prof. Mario Clerici

ANNO ACCADEMICO 2009/2010

ABSTRACT

Background. Colorectal cancer (CRC) is a major cause of death for cancer in western countries, ranking third in both sexes. Therapeutic developments in the past decades have extended life expectancy in patients with advanced disease (i.e., stage III), and even for those with distant metastasis (i.e., stage IV). Most treatments for advanced disease nowadays include a combination of chemotherapy with target therapy. Despite advances, the fact that metastatic colorectal cancer remains largely incurable, pushes to pursue a better understanding of the factors underlying cancer progression. Nowadays, a major field of investigation is the relationship between epithelial tumor cells and the surrounding compartment, namely tumor microenvironment, and in particular its contribution to cancer progression.

The tumor microenvironment essentially comprises tumor infiltrating cells, vasculature, extracellular matrix, plus other matrix associated molecules. Transformed cells can modulate the functions of stromal cells, likely to facilitate their own growth and survival. In this Darwinian perspective, outgrowth of cancer cells goes together with local changes. Such changes, like clonal ones, are likely progressive, from the stage of local invasion, up to regional lymph-node colonization and finally to the development of distant metastasis.

Infiltrating cells are a mix of populations having myeloid or mesenchymal origin, including tumor-associated macrophages, myeloid -derived suppressor cells, mast cells, monocytes, neutrophils, CD3+ T cells, natural killers, dendritic cells, endothelial cells, mesenchymal stem cells and cancer-associated fibroblasts. Taken together, all these players are involved in a double-faced game, in which an anti-tumor effect (such as that exerted by CD3+ cells in early CRC, [1]) is counteracted by a cancer promoting one (e.g., that exerted by macrophages [2]. Currently, this cancer-microenvironment match implies contradictory and controversial data, largely depending upon the investigated cell population and upon the experimental setting. This Ariadnes' thread seems unravelled, mainly because of the contemporaneous evolution of both tumor and microenvironment cells during multi-stage tumor progression. What is certainly perceived today, is that a switch from a genetic to a non-clonal prospective is required to understand tumor evolution. Clearly, the dynamic architecture of the stromal compartment and the interactions therein, need to be reconciled with the evidences concerning genetic irreversible changes in stromal cells. The latter include loss of heterozygosis, microsatellite instability (MSI) [3], trisomy of Chromosome 7 in connective tissues elements of CRC [4], and p53 mutations (in the stroma of breast carcinoma[5]). These surprising results rise the possibility that the stromal compartment contains not only an admixture of non-neoplastic cells, but also cancer cells with an aggressive and invasive phenotype which became able to invade the surrounding tissues by mimicking fibroblast morphology. This metamorphosis would be possible through the re-use of an embryonic program by cancer cells, that is the **epithelial to mesenchymal transition** (**EMT**). Originally, in a murine model of spontaneous metastatic breast cancer, Weinberg and Coll. demonstrated that the highest aggressive potential of cancer cell was reached after their transition from an epithelial to a mesenchymal morphology (i.e., EMT), driven by *Twist1* gene [6]. Twist1 expression and EMT are strictly associated with the acquisition of a spindle-like fibroblast morphology, the down-regulation of the

epithelial marker E-cadherin and the expression of mesenchymal ones (N-cadherin and vimentin), with metastases development, and with the inhibition of the key pathway of senescence p16 driven [7]. Currently, many studies demonstrate that this embryonic program is activated by a set of transcription factors which act pleiotropically. These include *Snail*, *Slug*, *Zeb1/2*, and obviously *Twist1*. These regulators are expressed in various combination in a number of malignant tumor types and have been shown in experimental models of carcinoma formation to be casually important for programming invasion [8], [9], [10], [11].

The issues. The multi-step process of metastasis development, recapitulated by local invasion, intravasation, extravasation, colonization and growth in distant organs, postulates that cells undergoing EMT should be able to complete each individual step. However, no experimental evidence of the EMT process has been provided for the first step of local invasion (nor for the following ones) in tissues specimens of human malignancies. Similarly, proof of EMT in human epithelial cancer cells remains partial, largely derived from murine models, and based on ectopic expression of EMT regulators, or on stimulation to achieve transient mesenchymal features. Although several studies demonstrate the expression of EMT transcription factors in both the tumoral and stromal compartments of different cancers [12], [13], [14], they do not directly link this expression to the presence of cancer infiltrating cells.

The experimental work. We moved from the unexplained evidence of genetic and chromosomal abnormalities in the stromal compartment of solid tumors, and from the role of Twist1 as EMT regulator in cancer cells.

Moving by microarray data of a pool of CRC cells, we first show that tumor cells with a permanent mesenchymal signature, in a stable EMT state, can arise from epithelial CRC cells, both in humans and mice. Then we clarify that, within the mesenchymal signature, Twist1 plays a crucial role in the migration and invasion of CRC tumor cells.

By a combination of immuno-based and cytogenetic methods we detected in human CRC a Twist1+subpopulation of stromal cells, with a mesenchymal phenotype. This subpopuation was more represented in MS-stable CRC than in less metastatic MS-unstable CRC, and was associated with advanced CRC stage and with worse survival. Additionally, we showed that Twist1 transcript is degraded in MSI CRC, due to a frameshifted 3'-UTR, and propose that this mechanism contributes to the low capability of MSI CRC to exploit EMT to undergo metastasis. Finally, we identified Twist1+, stromal cells which share genetic changes with epithelial cancer cells, establishing a genetic link between epithelial and mesenchymal components of CRC.

In summary, we provide data from an original cellular model to tissue studies to prove the occurrence of EMT in human CRC. We demonstrated the presence of tumor cells in the stroma of CRC tumors and we give a method to identify those cells undergone to EMT and able to invade the surrounding tissues. Thus, our results readdress the study approach to the stromal compartment from that of a recipient of non-neoplastic cells to an incubator of cancer EMT cells able to disrupt tissues by mimicking activated fibroblast.

The experimental demonstration of human CRC in stable EMT, coupled to the identification of previously postulated tumor cells with mesenchymal morphology in the stromal compartment, add substantial evidence to EMT, translating a model it into a real phenomenon contributing to the metastatic process of human CRC. In a

clinical perspective, our study highlights the importance to re-evaluate the target therapy of solid tumors from the epithelial compartment to include the mesenchymal one.

1. INTRODUCTION

It is now ascertained that human cancer can be characterized by patterns of changes in gene expression. Genes that mediate tumorigenesis can be broadly characterized as *oncogenes*, which are activated by alterations; and *tumour suppressor genes*, which are inactivated during tumorigenesis. Oncogenes can encode growth factors or their receptors, signalling molecules, regulators of the cell cycle, and other factors that regulate cell proliferation and survival. Their oncogenicity can be induced by mutations that lead to overactive gene products, amplifications that alter copy number, alterations or rearrangements that affect promoter function, or modified interactions with regulators of transcription or epigenetic modification. Tumour suppressors restrain growth and proliferation, passage through the cell cycle, motility, invasion, or other functions related to stable differentiation. Genes that encode tumour suppressors are commonly inactivated by deletion, mutations, promoter methylation, or other changes in regulation. Colorectal cancer (CRC),is a major public health problem in the United States and globally. In the United States, it is the third in cancer incidence, and the second leading cause of cancer mortality. In 2010, there will be an estimated 50,000 death associated with this disease. Worldwide, there will be nearly 1,000,000 of new cases diagnosed in 2008, resulting in about 500,000 deaths. It has now been well established that CRC develop gradually over a long period of time through the sequential accumulation of genetic alterations that overcome the redundant control mechanisms built into each cell. Only a few mutations are common to most colorectal tumours, but each tumour has a unique combination of genetic alterations. It is possible that no two CRCs are alike, a daunting consideration for the rational planning of treatment.

1.1 Models of CRC Pathogenesis

It was not immediately obvious that CRCs would be so diverse genetically. The initial attempt to characterize multistep carcinogenesis resulted in a novel conceptual model in which specific genetic alterations were associated with the sequential evolution of the neoplastic phenotype in the colon.

CRCs develop through an ordered series of events i.e. the so-called “adenoma–carcinoma sequence”, beginning with the transformation of normal colonic epithelium to an adenomatous intermediate and then ultimately adenocarcinoma [15]. Multiple genetic events are required for tumour progression. Genomic instability is now recognized as an essential cellular feature that accompanies the acquisition of these mutations. In CRC, at least three distinct pathways of genomic instability have been described, the chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) pathways.

In 1990, Fearon and Vogelstein [16] first proposed a multistep genetic model of colorectal carcinogenesis that has come to serve as a paradigm for solid tumour progression. Inactivation of the *adenomatous polyposis coli* (*APC*) tumour suppressor gene occurs first, followed by activating mutations of *KRAS* (Figure 1.1). Subsequent malignant transformation is driven by additional mutations in the transforming growth factor–β, PIK3CA, and TP53 pathways [15,16,17,18,19,20] This model predicts that at least seven distinct mutations are required. Recent genome-wide sequencing efforts have calculated as many as 80 mutated genes

per colorectal tumour, but a smaller group of mutations (<15) were considered to be the true “drivers” of tumorigenesis. [21,22] Although continue to be proposed refinements to the original model, three key principles have been today well established: a) multiple genetic hits are required, b) there are discrete intermediates in the progression to cancer, and c) the temporal acquisition of these genetic changes matters. For example, *APC* mutations serve as the initiating event in adenoma formation in human and mouse models. In contrast, mutational activation of *KRAS* cannot initiate cancer *in vivo*, and only when combined with a mutation in *APC* does mutant *KRAS* promote tumour progression [23]

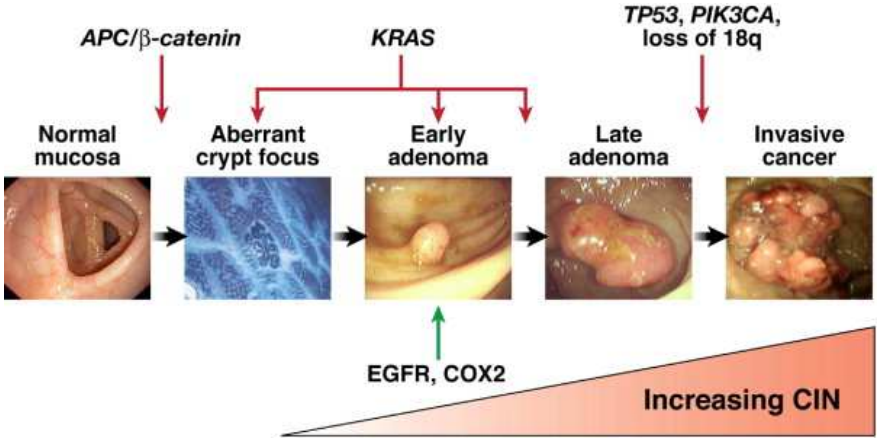


Figure 1.1 *Multistep genetic model of colorectal carcinogenesis (The Chromosomal instability in Colorectal cancer, Gastroenterology, 2010)*

1.1.1 Genomic Instability and CRC

Baseline mutation rates are insufficient to account for the multiple mutations that are required for cancer to develop. The rate of mutations per nucleotide base pair is estimated to be as low as 10^{-9} per cellular generation [24]. As proposed by Loeb *et al*. cancer cells must acquire some form of intrinsic genomic instability, a so-called “mutator phenotype,” that increases their rate of new mutations [25]. CIN is observed in 65%–70% of sporadic CRCs; the term refers to an accelerated rate of gains or losses of whole or large portions of chromosomes that results in karyotypic variability from cell to cell [26]. The consequence of CIN is an imbalance in chromosome number (aneuploidy), subchromosomal genomic amplifications, and a high frequency of loss of heterozygosity (LOH). A defect in the DNA *mismatch repair system* leads to a second pathway characterized by instability in stretches of DNA microsatellites. The third pathway is designated the CIMP, which exhibits gene silencing due to hypermethylation of CpG islands [30]. Because the definitions of these three pathways are not mutually exclusive, a tumour can occasionally exhibit features of multiple pathways. For example, up to 25% of MSI CRCs can exhibit chromosomal abnormalities [31] In addition, whereas CIMP can account for most of the MSI-positive/CIN-negative CRCs, up to 33% of CIMP-positive tumours can exhibit a high degree of chromosomal aberrations[29]

Conversely, as many as 12% of CIN-positive tumours exhibit high levels of MSI[32]. The significance and implications of these overlapping features are not yet fully defined.

1.1.2 Mechanisms Leading to Chromosomal Instability

It is now known that more than 100 genes can cause chromosomal instability in the yeast *Saccharomyces cerevisiae*. Although many of these genes have human homologues, only a limited number have been implicated in human tumours.

1.1.2.1Chromosome Segregation Defects and CIN

The CIN phenotype could result from defects in pathways that ensure accurate chromosome segregation. The mitotic checkpoint (also known as the spindle assembly checkpoint) is the major cell cycle control mechanism that ensures high fidelity of chromosome segregation by delaying the onset of anaphase until all pairs of duplicated chromatids are properly aligned on the metaphase plate. Defects in checkpoint signalling lead to chromosome mis-segregation and subsequent aneuploidy with abnormal numbers of chromosomes being distributed to daughter cells. Nearly 100 years ago, Theodor Boveri suggested that malignant tumours may arise from a single cell with an abnormal genetic constitution acquired as a result of defects in the mitotic spindle apparatus [33]

Products of the genes *mitotic arrest–deficient* (*MAD*) and *budding uninhibited by benzimidazoles* (*BUB*) operate as checkpoint sensors and signal transducers that control sister chromatid separation. Their activation causes inhibition of the anaphase-promoting complex/C (APC/C), a large ubiquitin-protein ligase, and cell cycle arrest. MAD3/BUBR1, MAD2, and BUB3 associate with the APC/C-activating molecule CDC20 to form the mitotic checkpoint complex and induces a conformational change in the APC/C, which prevents binding and ubiquitination of its substrates [34]. Activation of the APC/C leads to degradation of securin and activation of separases, a class of caspase-related proteases. Separase regulates a multiprotein complex termed cohesin, which creates physical links between sister chromatids that are maintained until late mitosis. Errors in this complex and ordered process result in unequal chromosomal segregation. Cahill *et al* demonstrated that cancer cell lines with mutations in *hBUB1* and *hBUBR1* failed to arrest when spindles were depolymerized[35]. Mutations in *hBUB1* were dominant because exogenous expression of the mutant gene conferred an abnormal spindle checkpoint and CIN to chromosomally stable cell lines [36]. To date, *hMAD2* and *hMAD1* mutations have been identified in breast cancer and leukemia, respectively, but not in CRC [37,38]. However, mutations in the *hZw10*, *hZwilch*/*FLJ10036*, and *hRod*/*KNTC* kinetochore proteins that function at the spindle checkpoint, and in *Ding*, which is essential for proper chromosome disjunction, have been reported in CRC [39]. Another protein involved in the mitotic checkpoint is the centromere-associated protein E (CENP-E), a kinetochore-associated kinesin-like protein. CENP-E stabilizes the interaction between microtubules and kinetochores and stimulates the recruitment and kinase activity of BubR1. Fibroblasts derived from mice heterozygous for *CENP-E* show significantly higher rates of aneuploidy than wild-type fibroblasts due to chromosomal mis-segregation. Another proposed cause of CIN is the abnormal centrosome number and function. Centrosomes coordinate mitosis by serving as an anchor for the reorganization of cytoplasmic microtubules into a mitotic spindle apparatus. It has been proposed that extra centrosomes lead to the formation of multiple spindle

poles during mitosis, resulting in the unequal distribution of chromosomes. Evidence of a causal relationship between centrosome amplification and cancer has recently been provided in flies[42]. Transplantation of larval brain cells with extra centrosomes induced the formation of metastatic tumours, suggesting that centrosome amplification can initiate tumorigenesis.
A pathogenic role for centrosome-associated Aurora and Polo-like kinases (Plk) has been also identified. There are three human Aurora kinases with distinct subcellular localizations and roles. Aurora A (*AURKA*) is involved in centrosome function and duplication, mitotic entry, and bipolar spindle assembly. Aurora A overexpression is associated with centrosome amplification, arrested mitosis with incomplete cytokinesis, and multinucleation [43] *AURKA* is amplified and positively associated with the degree of CIN in CRC [44] Aurora B is the catalytic component of the chromosomal passenger complex that regulates the accurate segregation of chromatids at mitosis, histone modification, and cytokinesis. Although its involvement in tumorigenesis is less clear, Aurora B overexpression correlates with advanced stages in CRC [45] Plk1 to Plk4 are serine/threonine kinases, which regulate entry into mitosis, centrosome duplication, transition from metaphase to anaphase, and cytokinesis. Two recent studies have provided evidence of a cross-talk between Plk1 and Aurora A. [46,47] Elevated expression of Plk1 has been observed in up to 73% of primary CRCs, and this correlated with tumour invasion, lymph node involvement, and Dukes' stage [48].

1.1.2.2 Telomere Dysfunction and CIN

Evidences indicate that CIN can be driven by telomere dysfunction. Telomeres are DNA-protein complexes that consist of stretches of hexameric repeats (TTAGGG in humans) that protect the ends of eukaryotic chromosomes from fusing and breaking during segregation. In somatic cells, a portion of telomeric DNA is lost after each round of cell division, due to the inability of DNA polymerase to completely synthesize the 3′ end of linear chromosomes, also known as the “end-replication problem.” When a critical short telomere length is reached, DNA damage checkpoints trigger cell senescence pathways and apoptosis. Cells that fail to undergo senescence enter a crisis-like state characterized by massive cell deaths. Cells that survive the crisis checkpoint activate the enzyme complex that elongates telomeres called telomerase, or alternatively the Alternative Lengthening of Telomeres pathway.
When the telomere end protection is compromised, chromosomal ends enter in breakage-fusion-bridge cycles that can continue for multiple cell generations leading to dramatic genome reorganization. Genomic amplification at fragile sites can commonly result from breakage-fusion-bridge cycles [49].The importance of telomere loss and breakage-fusion-bridge cycles has been demonstrated *in vivo*. In mice deficient in the RNA component of telomerase ($Terc^{-/-}$), telomere shortening led to increased rates of spontaneous tumour development of aberrant crypt foci and microadenomas in the gastrointestinal tract [50]. Additionally it has been also demonstrated that telomeres are shortened at the stage of high-grade dysplasia [51]. Several Authors have reported shorter telomeres rate ranging between 77% and 90% in CRC tissues, compared with adjacent normal counterparts[52,52,54,55].These observations led to hypothesize that whereas telomere shortening promotes the chromosomal instability that drives early carcinogenesis, telomerase activation during later stages confers immortality to the tumour cells. Consistent with this view, longer telomeres have been identified in

tumors staged as Dukes' C and D versus those scored as A and B, and increasing telomere length and telomerase expression has been correlated with an increased invasiveness.[56]

1.1.2.3 DNA Damage Response and CIN

It is well known that DNA damage response machinery protects cells from exogenous as well as endogenous genotoxic stress by initiating a cascade that culminates in cell cycle arrest to allow sufficient time to repair the damage or, in the case of irreparable damage, by inducing senescence or apoptosis. Some DNA repair proteins are involved in human cancers, such as the ataxia telangiectasia mutated (*ATM*) and Rad3-related (*ATR*) protein kinases.[59] Inactivating mutations in these genes predispose to specific syndromes, including ataxia-telangiectasia (*ATM* mutations), Seckel (*ATR* mutations), Li-Fraumeni (*TP53* mutations), and hereditary breast-ovarian cancer (*BRCA1* and *BRCA2* mutations) syndromes, all of them characterized by an increased susceptibility to cancer Among the checkpoint genes, *TP53* has been the only one directly implicated in human CRC, having at least a permissive role for development of CIN. Recently, Wang *et al* identified >1000 genes that could cause chromosomal instability based on their homology to genes of yeast and *Drosophila melanogaster*[39]. Today, 100 candidate genes were sequenced in a panel of CRCs and,19 mutations in five genes were identified, while one of these genes, *MRE11*, is involved in DNA double-strand break repair.

New insights suggesting a direct relationship between the DNA damage response and CIN have been provided by *in vivo* studies. Haploinsufficiency of histone *H2AX*, an ATM and ATR substrate, can compromise genomic integrity, and in a p53-deficient background, enhance tumour susceptibility [60,61] *ATM* and *H2AX* are closely linked, map to the same cytogenetic region (11q23), and often deleted in various human cancers. Mouse embryonic fibroblasts derived from *ATM* and *H2AX* double-deficient mice show severe genomic instability [62]. In addition, Peddibhotla *et al* reported that deficiency of *Chk1*, another DNA damage checkpoint effector, causes *in vivo* multiple mitotic defects and displaces *Aurora B* from late mitosis structures, resulting in failure of cytokinesis and multinucleation [63]

1.1.2.4 Cytogenetic of CIN CRC: 'the monosomic and trisomic type'

CIN CRC show several forms of genomic instability, characterized mainly by chromosomal rearrangements and numerical abnormalities at a greatly increased rate compared with normal cells [64].

Originally, the CIN phenotype was used to describe tumors with a high degree of intercellular heterogeneity in chromosome number, ascertained by counts for a restricted set of chromosome-specific centromeres [26]. CIN was further employed to describe cancers with either aneuploid or polyploid DNA content as measured by cytometry or cytogenetics, or multiple gains or deletions of chromosomes or chromosome arms, or frequent losses of heterozygosity (LOH) [65]. The first cytogenetic observations on cancer cells from colorectal adenocarcinomas were described by Dutrillaux [65]. With classical cytogenetic approach he observed two distinct patterns of chromosomal anomalies. The first one, called “monosomic type”, was characterized by many chromosomes losses, including the losses or deletions of chromosomes 18, 17 (short arm = p), 1p, 4, 14, 5 (long arm = q) and 21. He noticed that this condition frequently evolves towards polyploidy, by

duplication of all remaining chromosomes. The second pattern, called “trisomic type”, was characterised by the gain of several chromosomes: 7, 12, X, 5 and 8. The chromosomal anomalies observed seemed at that time to have no topological relationships with oncogenes [65]. Indeed subsequently it became clearer that the cancer-specific aneuploidies generate complex, malignant phenotypes, through the abnormal dosages of the thousands of genes [66].

In general, genomic copy number changes are frequently found in different types of cancer and are believed to contribute to their development and progression through inactivation of tumour suppressor genes, activation of oncogenes, or more subtlly through gene dosage changes [67]. Chromosomal instability is an efficient mechanism for causing the physical loss of a wild-type copy of a tumor-suppressor gene, such as in CRC: *APC*, *TP53*, and SMAD family member 4 (*SMAD4*), whose normal activities oppose the malignant phenotype [68]. The first cytogenetic studies on primary CRC tumors suffered by limitations due to the poor quality of the preparations. The improvement of molecular cytogenetic studies, performed subsequently with the application of fluorescence in situ hybridization (FISH) based techniques such as comparative genomic hybridization (CGH) and spectral karyotyping (SKY), allowed many groups to show that colorectal carcinomas are characterized by multiple patterns of chromosomal imbalances which sequentially accumulate during adenoma to carcinoma progression [69].

Among the first findings, Fearon et al. [70] confirmed a frequent genomic loss in CRC at chromosome 18q21. By using comparative genomic hybridization (CGH) Korn et al. [71] found that most frequent gains are at 20q13 and with the same technique De Angelis et al. [72] determined that gains of 20q and losses of 18q were the most frequent aberrations.

Accumulation of losses in 8p21-pter, 15q11-q21, 17p12-13, and 18q12-21, and gains in 8q23-qter, 13q14-31, and 20q13 were found strongly associated with adenoma-to-carcinoma progression, independent of the degree of dysplasia [69].

Diep et al. by applying different statistical analyses and combining these on a large series of genome profiles from reported CRC, were able to identify specific chromosomal alterations linked to the different stages of tumor progression. They found that losses at 17p and 18 and gains of 8q, 13q, and 20 occur early in the establishment of primary CRC whereas loss of 4p is associated with the transition from Dukes' A to B–D stages. They observed that deletion of genes located at chromosome 4 might therefore contribute to increased aggressiveness of the tumor, enabling it to penetrate the muscular layer, and contribute to the establishment of advanced stages; moreover loss of chromosome 4 has been shown to be associated with poor clinical outcome in a large series of CRC. The transition from primary tumor to liver metastasis results correlated with the deletion of 8p and gains of 7p and 17q, whereas losses of 14q and gains of 1q, 11, 12p, and 19 are late events.

In a pilot study by Fensterer and coworkers [67] using macrodissected paraffin-embedded tissue samples, matrix-CGH was performed. The majority of advanced tumours displayed 13q-gain and 18q-loss. In locally restricted tumours, only half tumours showed a gain on 13q and 7/12 tumours showed a loss on 18q. Interphase-FISH and high-resolution arraymapping of the gain on 13q confirmed the validity of the arraydata and narrowed the chromosomal interval containing potential oncogenes. The amplification on 13q appeared to harbour candidate

genes that might confer a more aggressive phenotype to colorectal cancer cells [67].
Finally, to summarize the numerous investigations performed in the last decade in primary colon cancer tumors, including the recent study by Knutsen and co-workers, comparing the CGH patterns in all of these reports, gains were most frequent for 3/3q, 5/5p/5q, 7, 8q, 20/20q, 13, and the X, and losses were most frequent for 8p and 18/18q [73]. Searching for the most frequent structural and numerical aberrations found in the large intestine adenocarcinoma specimens in the Mitelman Database of Chromosome Aberrations in Cancer (http://cgap.nci.nih.gov/Chromosomes/Mitelman), among unbalanced Chromosomal Abnormalities there are 8q10 i(8)(q10) and 17q10 i(17)(q10); among numerical trisomies +7, +13, +20, and +X; among monosomies −4, −5, −8, −10, −14, −15, −17, −18, −21, −22, and −Y. The most recurrent aberration found in all cytogenetic studies performed, either in primary tumors or in colon cancer cell lines or in fixed colorectal cancer tissue blocks is thus 18q (**Table Ia)**. LOH of this region indicates an unfavorable outcome in patients with stage II CRC [74]. In the 18q21-18q21.1 region several tumor suppressor genes have been mapped, including *SMAD4/DPC4*. However, since microsatellite instability was inversely correlated with loss of heterozygosity for chromosomes 5q, 17p, and 18q, it has been hypothesized that some colorectal cancers (MSI in particular) may arise through a mechanism that does not necessarily involve loss of heterozygosity [75].

Table Ia *Most frequent Chromosomal aberrations found in CRC*

Chromosome loss	Chromosome gain	Reference
18, 17p, 1p, 4, 14, 5q, 21	7, 12, X, 5, 8	Dutrillaux, 1988
18q21		Fearon et al., 1990
	20q13	Korn et al., 1999
18q	20q	De Angelis et al., 1999
18p21-pter, 15q11-q21, 17p12-13, 18q12-21	8q23-ter, 13p14-31, 20q13	Hermsen et al., 2002
4, 18p, 14q	17p, 17q, 1q11, 12p, 19	Diep et al., 2006
8p, 18q, 1p22, 4q26, 15q21	20, 8q, 8q28, 16q24.3, 20q13	Camps et al., 2006
18q	13q	Fensterer et al., 2007
18, 17p, Y, 1p3, 8p	13, 20, 7, X, 12, 6	Muleris et al., 2008
8p, 18, 18q	3, 3q, 5, 5p, 5q, 7, 8q, 20, 20q, 13, X	Knutsen et al., 2010
4, 5, 8, 10, 14, 15, 17, 18, 21, 22, Y, 18q10 [i(8) (q10)], 17q10 [i (17) (q10)]	7, 13, 20, X	Mitelman Database online

1.1.2.5 LOH and CIN

LOH is considered to be a hallmark of CIN-positive tumors. An average 25%–30% of alleles are lost in tumours; it is not unusual to observe losses in 75% of alleles in tumour samples. Surprisingly, there are relatively few insights into the mechanisms underlying the generation of LOH. Nevertheless, several pathways have been implicated, including mitotic nondisjunction, recombination between homologous

chromosomes, and chromosomal deletion. Thiagalingam *et al* performed a detailed molecular and cytogenetic analysis of LOH in the five chromosomes (1,5, 8, 17, and 18) most frequently lost in human CRC [76]. The majority of losses on chromosome 18 involved the whole chromosome and appeared to be caused by mitotic nondisjunction. However, when the losses were limited to a part of a chromosome, these appeared to be a consequence of interchromosomal recombinations and deletions associated with DNA double-strand breaks rather than to mitotic recombination, break-induced replication, or gene conversion. The specific genetic defects that underlie these LOH events still remain unidentified.

1.1.2.6 Chromosomal Instability Pathway

Coupled with the typical karyotypic abnormalities observed in CIN tumours is the accumulation of a set of mutations in specific tumour suppressor genes and oncogenes. These mutations activate oncogenic pathways that are critical for CRC pathogenesis. It is not clear whether CIN creates the appropriate environment for the accumulation of these mutations or *vice versa*, although a role for *APC* in the establishment of the CIN phenotype has been recently proposed. Nevertheless, the combination of these mutations in CRC in the setting of CIN has been historically designated as the "chromosomal instability pathway."

APC/β-Catenin

The earliest genetic event in colorectal tumorigenesis is the activation of Wnt signalling through the genetic disruption of *APC* on 5q21.[77] Germline mutations in *APC* are responsible for familial adenomatous polyposis.[78] Somatic *APC* mutations are observed in 5% of dysplastic aberrant crypt foci, 30%–70% of sporadic adenomas, and in as many as 72% of sporadic tumours, indicating that functional loss of *APC* is an early event in tumour initiation. [79,80,81] Although germline-inactivating mutations are distributed throughout the entire gene, somatic mutations are clustered in the mutation cluster region between codons 1286 and 1513. [82] Hypermethylation of the *APC* promoter has been also reported in 18% of primary CRC and colonic adenomas, representing an alternative mechanism for *APC* gene inactivation. [83]

The APC gene product is a large protein with multiple functional domains that regulates complex cellular processes including differentiation, adhesion, polarity, migration, development, apoptosis, and even chromosomal segregation. A critical function is the interaction between APC and glycogen synthase kinase-3β (GSK-3β) and β-catenin, each an essential component of the Wingless/Wnt signalling pathway. APC binds to β-catenin, GSK-3β, and casein kinase 1α/ε on an axin-conductin scaffold; the subsequent phosphorylation of β-catenin by GSK-3β leads to proteasome-dependent degradation and suppression of the Wnt signal. [84] Mutant APC disrupts the complex formation, and the increased cytoplasmic levels of β-catenin can translocate to the nucleus, where it drives the transcription of multiple genes implicated in tumour growth and invasion through its interaction with the T-cell factor/lymphoid enhancer factor family. [85]

Gain-of-function mutations in β-catenin (*CTNNB1*) have been identified in as many as 50% of CRC with intact *APC*, reflecting the importance of the Wnt pathway. [86] Samowitz *et al* found a significantly higher frequency of *CTNNB1* mutations in a sub-groups of small adenomas (12.5%) *versus* large adenomas (2.4%) and invasive cancers (1.4%), suggesting that adenomas with alterations in β-catenin are less likely to progress to malignancy and that *APC* and *CTNNB1* mutations are not functionally equivalent.[87] Mutations in the *AXIN* and *AXIN2/conductin* genes

have also been reported, but only in CRCs with defects in the mismatch repair system.[88,89]

K-RAS

KRAS is mutated in 30%–50% of CRCs.[90] Single nucleotide point mutations that occur in codons 12 and 13 of exon 2, and to a lesser extent in codon 61 of exon 3, lock the enzyme in the guanosine triphosphate–bound, activated form, leading to constitutive activation of RAS downstream signalling. *KRAS* mutations have been observed in 60%–95% of nondysplastic or hyperplastic aberrant crypt foci, indicating that these microscopic lesions are unlikely to be precursors of adenomas and cancer. [91,92]

Activated RAS regulates multiple cellular functions through well-described effectors The best characterized effector is the Raf–mitogen-activated protein kinase kinase (MEK)–extracellular signal-regulated kinase (ERK) pathway. The Raf family includes 3 serine/threonine kinases (ARAF, BRAF, and RAF1) that activate MEK1 and MEK2, which in turn phosphorylate ERK1 and ERK2. ERK then phosphorylates cytosolic and nuclear substrates, including JUN and ELK1, which regulate enzymes such as cyclin D1, which is involved in the control of cell cycle progression.[93] The significance of ERK activation in CRC pathogenesis still remains unexplained. Whereas MEK is frequently activated in human CRC, recent studies have indicated that inhibitors of MEK are ineffective therapeutics.

RAS–guanosine triphosphate also binds the catalytic subunit of type I phosphatidylinositol 3 kinases, which translocate to the plasma membrane to generate phosphoinositol lipids. Of note, *PIK3CA* is mutated in almost 20% of CRC; mutations that occur in the hotspots located on exons 9 (E542K, E545K) and 20 (H1047R) are oncogenic in cellular models of CRC. [94] AKT/PKB is a critical downstream target of phosphatidylinositol 3 kinase that regulates a cascade of responses including cell growth, proliferation, and survival by inactivating several proapoptotic proteins such as BAD and Forkhead (FKHR) transcription factors.[95, 96] Phosphatidylinositol 3 kinase also activates Rac, a Rho family guanosine triphosphatase, which is important for RAS-mediated transformation in some cellular contexts.[97]

TP53

TP53 is located on the short arm of chromosome 17; it encodes a 393 aminoacid transcription factor that is a tumour suppressor and a central coordinator of cellular responses to stress, including DNA damage, aberrant proliferative signals, and oxidative stress.[101] *TP53* is induced by several oncogenic proteins, such as c-Myc, RAS, and the adenovirus E1A. [102] Under normal conditions, p53 is negatively regulated by MDM2, E3-ubiquitin ligase, and the related protein MDM4 (also known as MDMX), which bind to the transactivation domain of p53 and target it for degradation by ubiquitination. In cells with a high level of stress, the interactions between MDM2, MDM4, and p53 are disrupted, allowing activated p53 to exert its transcriptional activity. Defined as the “guardian of the genome” p53 is a master regulator that controls the transcription of hundreds of genes involved in DNA metabolism, apoptosis, cell cycle regulation, senescence, angiogenesis, immune response, cell differentiation, motility, and migration. Some of the best-studied targets of p53 are CIP1/WAF1, the p21 cell cycle inhibitor, GADD45 and 14-3-3, which contribute to G2 arrest, and BAX, FAS (APO1), PIG3, and KILLER (DR5), which regulate caspase activation and apoptosis. [103,104]

p53 dysfunction is the commonest hallmark of human tumours and its loss of function has been reported in 4%–26% of colon adenomas, 50% of adenomas with invasive foci, and in 50%–75% of CRC, suggesting its key-role in the transition from an adenoma to carcinoma.[105] The majority (~80%) of *TP53* mutations are missense: GC to AT transitions that occur principally in 5 hotspot codons (175, 245, 248, 273, and 282).[106] These mutations lead to the synthesis of an inactive protein with an abnormally long half-life that can be recognized by immunohistochemistry. p63 and p73 proteins are functionally and structurally related to p53, and they may play a role in tumor suppression. Although loss of p63 and p73 has been recently found in many tumour types, their involvement in CRC pathogenesis has not yet been determined.[107]

18q Loss

Allelic loss at chromosome 18q has been identified in as many as 70% of primary CRC, particularly in advanced stages. The major tumour suppressor genes localized on 18q are *SMAD2* and *SMAD4*, which are intracellular mediators of the transforming growth factor-β (TGF-β) pathway that are involved in the regulation of cell growth, differentiation, and apoptosis. However, *SMAD4* and *SMAD2* mutations have been found in <20% and 10% of colon cancers, respectively.[108,109]

Cyclooxygenase-2

Aberrant overexpression of cyclooxygenase-2 (COX-2) is thought to have an important role in development of CRC. *COX-2*, an early-response gene induced by growth factors, cytokines, inflammatory mediators, and tumour promoters, is overexpressed in as many as 43% of adenomas and 86% of carcinomas.[116] Oshima *et al* first highlighted the role of COX-2 in CRC.[91] They showev that in *APC*^{716} knockout mice, the number of intestinal polyps was reduced by 34% when one allele of *COX-2* was knocked out and by 86% when both alleles were deleted.

The tumorigenic effects of COX-2 has been mainly attributed to the production of prostaglandin E2. Increased levels of prostaglandin E2 have been reported in colorectal adenomas as well as carcinomas.[118,119] It has been demonstrated that COX-2 and prostaglandin E2 regulate proliferation, survival, migration, and invasion in CRC.[120] It is also known that COX-2 is a regulator of the angiogenic process: overexpression of COX-2 induces the production of proangiogenic factors including the vascular endothelial growth factor (VEGF) and basic fibroblast growth factor (bFGF).[121] *In vivo* studies have shown that homozygous deletion of *COX-2* not only impairs the growth of tumour xenografts but also reduces tumour vascularity.[122]

1.1.2.7 Timing of CIN: The Chicken or the Egg?

The question of whether CIN is a cause or a consequence of the malignant process remains unanswered. The stage of tumorigenesis at which the CIN phenotype arises is controversial Some Authors suggest that CIN initiates tumorigenesis, while others contended that CIN is acquired and has a role in maintaining the tumorigenic process.

Several studies have demonstrated that CIN, measured as allelic imbalance in specific chromosomal regions, occurs at very early stages of tumorigenesis.[123,124,125] Shih *et al* analyzed 32 sporadic colorectal adenomas (with an average diameter of 2 mm) for allelic imbalances using digital single nucleotide polymorphism polymerase chain reaction.[125] A relatively high frequency of allelic imbalances on chromosomes 5q (55%), 1p (10%), 8p (19%),

15q (28%), and 18q (28%) was identified, with >90% of the adenomas exhibiting allelic imbalance of at least one chromosomal arm. Cardoso *et al* investigated the CIN status of polyps from patients with germline mutations in *APC* or *MYH*.[126] As many as 60% and 80% of the polyps exhibited aneuploid changes, respectively, with gains of chromosome 7 and 13, and losses of chromosomes 17p, 19q, and 22q, being the most prevalent aberrations. These findings support the view that chromosomal abnormalities can occur during very early stages of tumorigenesis. Whether these abnormalities truly constitute an underlying CIN and whether they occur before or after *APC* inactivation are questions difficult to answer. Nowak *et al* developed a stochastic mathematical model to define under what conditions CIN would likely be the initial event in colorectal carcinogenesis or the second event following mutation of the first copy of the *APC* gene [127]. Their model has shown that for a wide range of realistic parameter values, CIN mutations have a potential role in the initiation of CRC.

1.1.3 MSI as a Unique Mechanism in Tumour Development

After the seminal proposal of the multistep genetic model of colorectal carcinogenesis, a number of investigators began searching the genome for novel tumor suppressor genes using even more powerful techniques. In 1992, Manuel Perucho used an arbitrarily primed polymerase chain reaction (PCR), extracting DNA from colonic tissues and amplifying thousands of sequences using a small number of randomly chosen PCR primers. This technique yielded unique but reproducible electrophoretic signatures from each sample[128]. Perucho *et al* used arbitrarily primed PCR to amplify matched samples of DNA from CRC and adjacent normal colonic tissue, subsequently they separated the PCR products electrophoretically, and compared the results to identify differences in PCR amplicons between normal and tumour tissues-DNA missing from the tumour would presumably represent somatic cell deletions that were candidate regions for new tumour suppressor genes.

Deleted DNA bands were found in most CRC samples.[129] However, upon careful analysis of the comparative autoradiograms, the Authors noted that 12% of ~130 tumours had bands that were not actually deleted, but were shorter in length and migrated slightly further down the gel. Therefore, they analyzed the sequences in these bands and discovered that they contained simple repetitive sequences (*i.e,* microsatellites), principally in polyadenine (A_n) tracts associated with *Alu* sequences. The DNA obtained from these tumours had undergone somatic deletion of ≥1 of the adenine residues. The method was unbiased and quantitative, permitting the estimate that each tumour contained ~10^5 such mutations. The tumours with these “ubiquitous somatic mutations at simple repeated sequences” had unique clinical and pathological characteristics. First, the tumours with these deletion mutations were: a) significantly more likely to arise in the proximal colon, b) less likely to be invasive, c) less likely to have mutations in *KRAS* or *p53*, d) more likely to appear poorly differentiated, and e) came from younger patients. Based upon these findings, the Authors concluded that these deletions represented a unique pathway to tumour development and predicted that the “catastrophic loss of fidelity in the replication machinery of normal cells” that caused them might be hereditary, although they had no evidence for this [129].Simultaneously, Thibodeau *et al* analyzed dinucleotide repeat sequences (a type of microsatellite), and search novel tumour suppressor genes on

chromosomes 5q, 15q, 17p, and 18q in CRC. They observed deletion mutations in the $[CA]_n$ sequences (called “C A repeats”) in these regions and coined the term *microsatellite instability* (which they indicated as *MIN*). They detected MIN in 25 of 90 (28%) CRCs, and noted that these aberrations were heterogeneous in different tumours. They referred to a large deletion or expansion within the $[CA]_n$ tract as a “type I mutation,” and called a single 2-base-pair repeat change a minor or “type II mutation.” The implications of type I *versus* II mutations have never been elucidated, but they might reflect the cumulative nature of mutations in dinucleotide repeat sequences as they occur over time[130].Importantly, Thibodeau *et al* found that 89% of tumours with MIN were in the proximal colon and that the patients with MIN CRCs had a better prognosis. Both Perucho *et al* and Thibodeau *et al* recognized that this represented a unique pathway for tumour development that “does not involve loss of heterozygosity.”[75]

Studies from an international consortium led by Bert Vogelstein in the United States and Albert de la Chapelle in Finland helped elucidate the clinical implications of MSI, using microsatellites (mainly, but not exclusively dinucleotide repeats) as tools for genome-wide linkage analysis of hereditary CRC. Aaltonen *et al* found significant linkage at chromosome 2p using the marker D2S123 in two large kindred with Lynch syndrome [then called hereditary nonpolyposis colorectal cancer (HNPCC)].[131] The authors used D2S123 to look for LOH at that locus in the CRC DNA, testing the hypothesis that this was the site of the tumour suppressor gene that caused hereditary CRC and that the second hit to the gene would be LOH. Instead, they observed MSI with the marker and then noted that deletion mutations in microsatellite sequences were widespread in hereditary CRC. They referred to this as the *replicative error phenotype*. More importantly, they found MSI in 13% of sporadic CRC and correctly reasoned that the hereditary tumours and a subset of sporadic tumours shared a unique but common pathway of tumour development.[132]

It became apparent that a subset of CRC (at least 12% and perhaps as much as 28%) were characterized by a large number of mutations at microsatellite sequences, formed in the proximal colon, had unique clinical features (such as a better prognosis than other types of CRC), and were hereditary. MSI was the first DNA marker available to identify hereditary CRC, although none of the investigators who discovered MSI in CRC understood mechanisms by which it developed.

A series of observations led to the realization that MSI arises from defects in the DNA mismatch repair (MMR) system and the identification of the four genes that cause Lynch syndrome. Additionally, microbial geneticists predicted that defects in the MMR system were probably responsible for MSI.

1.1.3.1 DNA MMR System

The “mutator phenotype” was discovered in bacteria in the 1970s and 1980s and was well-characterized by the beginning of the 1990s.[133] It is caused by mutational inactivation of genes involved in DNA repair. The first human disease that was clearly associated with defects in DNA repair was xeroderma pigmentosa, a rare autosomal recessive disease caused by biallelic inactivating mutations in genes involved in nucleotide excision repair. Many other DNA repair systems had been characterized that included MMR, base excision repair, and a variety of nucleases and other DNA excision repair enzymes, yet no human diseases were associated with these.

In prokaryotes, the MMR system consists of a family of enzymes that detect S-phase DNA replication errors *i.e.,* those that result in mismatches between the 2 strands of DNA, in which the newly synthesized strand has incorporated the wrong nucleotide. DNA polymerase sometimes makes errors incorporating the correct number of bases during replication of long repetitive DNA sequences, such as microsatellites. Slippage during replication of a repetitive sequence creates a temporary insertion-deletion loop (IDL) that can be recognized and repaired by the MMR system, along with single base-pair mismatches. If these are not repaired, during the second round of replication the original parental strand is copied correctly, but the erroneously synthesized daughter strand (with the mismatch or IDL) contains a mutation. Single base-pair mismatches result in point mutations, whereas IDLs result in frame-shift mutations that usually lead to a downstream nonsense mutation; this results in production of a truncated and therefore nonfunctional protein.

1.1.3.2 The basis of MSI

It is well known that MMR is more complex in yeast than eukaryotes. The yeast homologues of the bacterial *mutS* and *mutL* genes were cloned and given the names *Mut S homologue* (*MSH*) and *Mut L homologue* (*MLH*). Subsequently, additional homologous copies of these genes were cloned from yeast, giving rise to the acronyms *MSH1* though *MSH6*, and *MLH1* through *MLH3*. Another MutL homologue, called *post-meiotic segregation-1* (*PMS1*), was also identified in yeast. Each of these genes had diverged from the initial *mutS* and *mutL* sequences and has been associated with specific repair functions in nuclear DNA, during meiosis, or in mitochondrial DNA. Phylogenetic analysis indicates that *MSH1* was the founding member of the Mut S family and that the homologues diverged from this progenitor gene, but no highly homologous *MSH1* gene is present in higher organisms.[134]

More importantly, in the yeast and other eukaryotes, MutS and MutL proteins no longer function as homodimers. Instead, MSH2 forms a heterodimer with MSH6 or MSH3, giving rise to MutSα or MutSβ, respectively. [135] These heterodimers have different abilities to bind to DNA mismatches and, as a result, yeast and other eukaryotes have a broader ability to recognize and repair different types of DNA misincorporation. For example, MutSα has a higher affinity for recognizing single base-pair mismatches. In fact, MSH6 was initially called the "GT-binding protein."[136,137] It was later discovered that the *MSH3* gene encoded another MutS homologue that dimerized with MSH2 to create a complex with increased ability to bind to larger IDLs.[138] The evolution of diverse homologues of *mutS* increased the cell's ability to recognize and repair synthetic errors in DNA and increased replication fidelity in higher organisms.

Mammals have four homologues of the prokaryotic *Mut L* gene: *MLH1*, *MLH3, PMS1*, and *PMS2*. *PMS1* was the first *mutL* homologue cloned in yeast; it was given a unique name because of its functional role in meiosis. Tracing the genes phylogenetically is complicated because the yeast *PMS1*, *MLH1*, *MLH2*, and *MLH3* genes correspond to human *PMS2*, *MLH1*, *PMS1*, and *MLH3* genes, respectively. The function of the products of the *MutL* homologues are not as clear as those of the *MutS* homologues, but the encoded proteins function as heterodimers. MLH1 and PMS2 form MutLα, which mediates the interaction between the MutS proteins and enzymes involved in long-patch excision in postreplication mismatch repair. However, *MLH1* (the major *MutL* homologue in humans) encodes a product that

can dimerize with PMS1 to form MutLβ, which suppresses mutagenesis in yeast but has an uncertain function in humans. Moreover, the dimer of MLH1 and MLH3 (MutLγ) helps suppress IDL mutations and functions during meiosis in yeast, but its function in humans is unknown. Loss of MLH1 results in total loss of MMR activity, but loss of PMS2 can be partially compensated by MLH3. It is not known how PMS1 fits into this model, but the MLH1-PMS1 heterodimer is not part of the canonical human MMR system. **Figure 1.2** shows this model system.

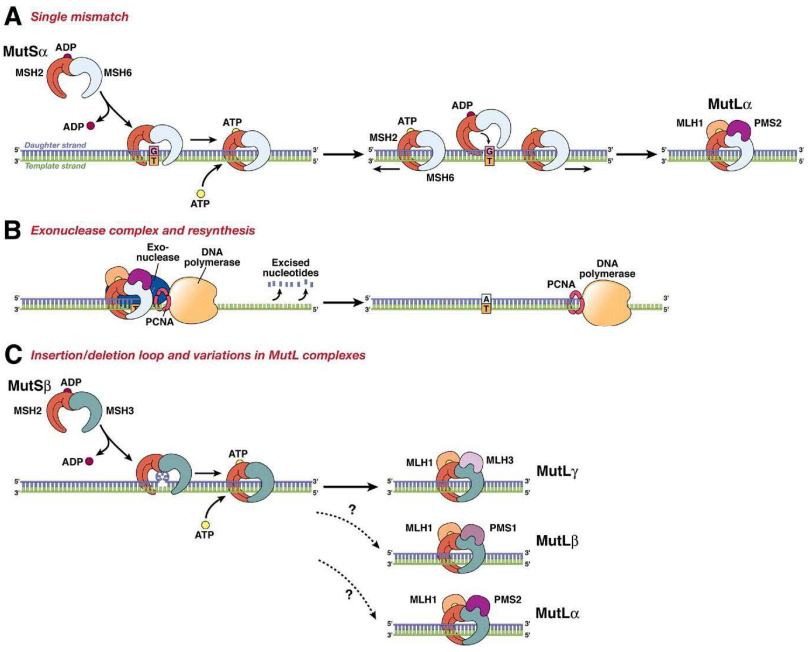


Figure 1.2 *The DNA MMR system functions through a series of steps. (A) MSH2–MSH6 (MutSα) recognizes single base-pair mismatches, in which the DNA polymerase has matched the wrong base (G) with the T on the template (shown on left), and creates a sliding clamp around the DNA. This step that requires the exchange of adenosine triphoshpate (ATP) for adenosine diphosphate (ADP) (by MSH2, but not MSH6 or MSH3). The complex diffuses away from the mismatch site, which is then bound by the MLH1-PMS2 (MutLα) complex (right). This "matchmaker" complex moves along the new DNA chain until it encounters the DNA polymerase complex. (B) The DNA MMR protein sliding clamp interacts with exonuclease-1, proliferating cell nuclear antigen (PCNA), and DNA polymerase. This complex excises the daughter strand back to the site of the mismatch (shown on left). Eventually, the complex falls off the DNA and resynthesis occurs, correcting the error. (C) Variations on the DNA MMR theme. Whereas MSH2–MSH6 recognizes single pair mismatches and small IDLs, MSH2–MSH3 (MutSβ) complements this by also recognizing larger IDLs (shown on left). The right side shows the possible interactions with different MutL dimers,as MLH1 can dimerize with PMS2, PMS1, or MLH3. The preferred interaction with MSH2–MSH3 is*

MLH1–MLH3 (MutLγ, but the precise roles of the other MutL heterodimers in this reaction are not entirely understood.

1.1.3.3 MSI is Caused by Deficiencies in MMR

By 1993, it was recognized that about 15% of colorectal tumors have a unique mechanism of pathogenesis; they might have been first solid tumors with subclasses that had features so distinct they could be considered as separate diseases. To identify the different pathogenic mechanisms, researchers analyzed the PCR products of colorectal tumors with MSI.

Microsatellite sequences are abundant throughout the genome; they are polymorphic among individuals but are unique and uniform in length in every tissue in each person. The heterogeneity of dinucleotide repeats made them valuable for forensic, gene mapping, and allele discrimination analyses. **Figure 1.3** illustrates amplification of dinucleotide and mononucleotide repeats in colonic tissues using autoradiography and gel electrophoresis.

The initial published examples of MSI were autoradiograms of radiolabeled PCR products; findings in yeast indicated that defects in genes that encode MMR factors might be responsible for the MSI observed, although no human MMR genes had been cloned.

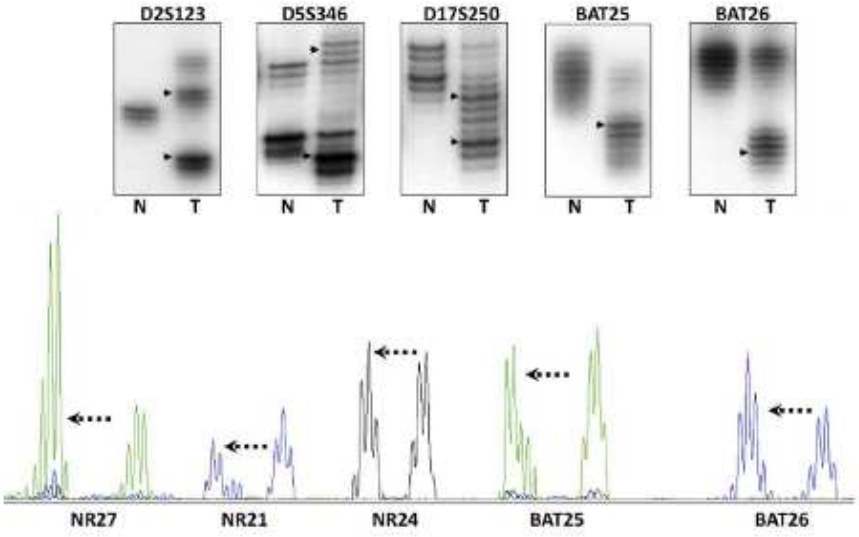


Figure 1.3 *MSI was initially discovered by autoradiography analyses of the PCR products after separation by gel electrophoresis (upper panel). N refers to DNA from the normal colon, and T from the tumor. The DNA polymerase used in PCR also has difficulty with the accurate amplification of templates, which is thought to be the explanation for the "ladder" of DNA bands that can be seen in the lanes for normal and tumor DNA. The upper panel illustrates the use of the 5 markers recommended by the National Cancer Institute consensus group; these consist of 3 dinucleotide repeats and 2 mononucleotide repeats (BAT25 and BAT26). In each instance, the DNA in the tumor has undergone somatic mutations (frequently, but*

not always, deletions), and the PCR product migrates to a different position on the gel, as indicated by the arrowheads. The lower panel shows the PCR products as they are analyzed by most laboratories using automated DNA sequencing with fluorescent primers. In this instance, 5 mononucleotide repeats have been analyzed, and in each instance, the mutations consist of deletions with different electrophoretic mobility (mutant alleles indicated by the arrows).

1.1.3.4 MMR Genes, HNPCC, and Lynch Syndrome

As much of the attention was focused on MSI in hereditary CRC, several investigators set out to determine whether germline mutations in *MMR* gene*s* were responsible for Lynch syndrome. Fishel *et al* cloned the human *MSH2* gene based on its homology to the yeast sequence, mapped it to human chromosome 2p22–21 (close to the locus implicated earlier in the year[132]), and found a sequence variation in patients with familial CRC[143] at the -6 position of intron 13 of *MSH2.* It was proposed that this variant altered splicing, inactivated the gene product, and was a germline mutation that mediated CRC pathogenesis. In fact, this is a common intronic polymorphism, but it was linked (in *cis*) to a true inactivating mutation that was found shortly thereafter. The association between *MSH2* and hereditary CRC explained the MSI phenotype observed in CRC; *MSH2* is located near the locus described in the first reported linkage study.

Leach *et al* confirmed the role of *MSH2* mutations in hereditary CRC, identifying germline mutations in three kindreds with familial CRC: one leucine to proline missense mutation, one large in-frame deletion, and one nonsense mutation.[144] These were found in multiple affected members of the family but not in individuals without cancer. Parsons *et al* showed that the CRC cell line HCT116 which had MSI at $[CA]_n$ sequences, was deficient in MMR activity.[145] Seven months after the discovery of MSI in CRC, MMR activity was linked to a class of genes that had not previously been of interest to cancer biologists. It was recognized that CRC might develop through loss of a DNA repair mechanism and that germline mutations in at least one of these genes mediated HNPCC.[146]

MLH1, PMS2, PMS1, and MSH6

Briefly, after the 2p locus was linked with familial CRC, CRC in three Swedish families was linked to chromosome 3p21-23.[147] The association between *MSH2* and Lynch syndrome suggested that other genes that encode MMR factors might be located on 3p, which would represent a second familial CRC locus. Two groups found *MLH1* on 3p21.[148] Bronner *et al* identified the human homologue of yeast MLH1 and found germline mutations in multiple members of a CRC family with significant linkage to the 3p locus.[149]

In the same time, Vogelstein *et al* and de la Chapelle *et al* scanned a DNA database for sequences homologous to the yeast *MLH1*, mapped one to chromosome 3p, and cloned the gene. Papadopoulos *et al* identified germline mutations in seven members of a large Finnish CRC kindred that had been linked to 3p, found mutations in three more CRC families, and reported a mutation in this gene in a CRC cell line.[150] Using the same strategy of searching databases for homologous sequences, the human *PMS1* and *PMS2* genes were implicated in familial CRC shortly thereafter.[151] The association between *PMS2* and CRC has been established, but the role of *PMS1* in sporadic or familial CRC remains unidentified. From December 1993 to September 1994, three more genes were

linked to familial CRC; it took another several years to identify the relationship between the *MSH6* gene[137] and Lynch syndrome.[152] *MSH6* is associated with somewhat atypical CRC kindreds with delayed onset of the cancer, which obscured the initial efforts at discovery. Once it was realized that *MSH5* families did not fit the classical expectations of the Lynch Syndrome phenotype, the case for its involvement was secured.[153]

Lynch Syndrome

The focus of MSI in CRC was immediately aimed at Lynch syndrome because of its inherited and unique features.[154] These patients develop tumours at early ages, often between 20 and 30 years old. They frequently have multiple tumours, including colon, rectum, endometrium, stomach, ovary, urinary tract, small intestine, and other sites, but no increase in the frequency of cancers of the breast, lung, or prostate.[155]

Families with clusters of CRC were therefore evaluated in linkage studies to identify loci of putative hereditary CRC genes. The Amsterdam Criteria were developed in 1991 so that research groups would have a uniform collection of families to study.[156] These criteria called for three CRC cases in a family in which one individual was a first-degree relative of the other two, CRC occurred in at least two generations (in which familial polyposis was excluded), and one affected family member was younger than 50 years of age. The Amsterdam II Criteria were introduced in 1999 to acknowledge the involvement of noncolonic tumors in the phenotype.[156]

Studies of these families facilitated the identification of genes responsible for this disease. Lynch syndrome is the hereditary disease caused by inactivating germline mutations in genes that encode MMR factors.[157] There are four definite genes involving in the Lynch syndrome: *MSH2*, *MLH1*, *MSH6*, and *PMS2*.[158] Other two genes *(MLH3*[158] and *Exo1*[159]) that are candidates for Lynch syndrome, have been only rarely associated with familial CRC or the reported associations were unexpected. There has been no independent confirmation of a role for *MLH3* in CRC and there has been a contradictory report on the role of *Exo1* in this disease.[160] However, both genes are an integral part of the MMR system, so there is rationale for considering these genes.

It is indubitable that the Amsterdam Criteria are rigorous; However, Although, it was surprising that 40%–60% of families that met the Amsterdam Criteria did not have a germline mutation in an *MMR* gene and the tumours did not have MSI.[161,162]. Families with non-Lynch syndrome CRC have been termed *familial colorectal cancer, type X* by one large collaborative study.[161] These families are characterized by microsatellite-stable tumours, a lower relative risk for CRC, the absence of excess tumours outside of the colon, and later onset of CRC. The genetic basis of this group is not known.

1.1.3.5 MMR Gene Mutations in Lynch Syndrome

To diagnose Lynch syndrome, MMR genes are analyzed, preferably using exon-by-exon sequencing with PCR primers that include relevant portions of the intron–exon boundaries; these detect point mutations and small insertion-deletion mutations that involve DNA sequences between the primers but not larger deletions or rearrangements, such as DNA sequence that are rejoined after a break, because the breakpoints are usually located deep within introns or completely outside of the coding regions. When DNA sequencing is performed, the tracing is examined for a double signal, *i.e.*, the simultaneous presence of a wild-

type and mutant nucleotide in the same sequence position. When there is a deletion of an entire exon in which the breakpoints are outside of the primers for that exon, the analysis gives normal results because only one strand is sequenced, and conventional techniques cannot determine whether one or two alleles are detected. Large deletions can be detected by multiplex ligation-dependent probe amplification, which quantifies the number of alleles at each exon.[163] Large deletions account for >33% of all *MSH2* mutations associated with HNPCC[164] *MSH2* is located within a dense cluster of *Alu* sequences, which are prone to internal recombination followed by excision of the DNA loop. Large rearrangements also occur with *MLH1* and *MSH6.*[165]

Analysis for mutations in DNA MMR genes does not end when a sequence variation is found; interpretation is another issue. Most deletions and nonsense mutations have pathologic consequences, but missense mutations are not always interpretable. It is important to determine how well the sequence has been preserved (indicating its functional significance) and if it is located in a region where the MMR proteins interact. Disruption of regions of protein-protein interaction by a nonconservative amino acid change could disrupt complex formation and function. The mutant allele might need to be cloned and its product tested in a functional assay.[166] A proportion of mutations identified in genetic testing are reported as "variants of uncertain significance," until the better ability to predict protein function based on its structure. It is interesting that there is a missense mutation in *MLH1* (c415G→C; D132H) that affects adenosine triphosphate binding of the mutant gene product, increases risk for CRC, but does not always cause MSI in CRC cells.

1.1.4 MSI in Sporadic CRC

Two of the three initial descriptions of MSI were made in specimens from sporadic CRC, rather than tumours from patients with familial CRC. Much of the subsequent attention to MSI was directed to familial CRC, but only about 3% of all CRC come from Lynch syndrome families. Approximately 12%-17% of all CRC have MSI, depending upon the methods used to detect it;[167,168] most CRC with MSI are sporadic.

The majority of CRC with MSI have lost expression of MLH1 and PMS2 protein. In 1997, Kane *et al* have shown that *MLH1* was silenced by methylation in most of these tumors.[169] The characteristic features of sporadic CRC with MSI include the absence of significant familial clustering, biallelic methylation of the *MLH1* promoter[170]; absence of MLH1 and PMS2 proteins; and frequent mutation (usually V600E) in *BRAF*[171]. The tumours are frequently diploid (74%) and patients with sporadic CRC with MSI have a better prognosis than those with non-MSI tumors.[172] Whereas patients with Lynch syndrome are younger than those with sporadic CRC, those with sporadic CRCs with MSI are older. It is known that loss of *MLH1* expression increases with age. Moreover, the gene is lost in ~50% of CRC patients older than 90 years of age.

1.1.4.1 Sporadic MSI and CpG Island Methylator Phenotype

MSI-associated sporadic CRCs arise through a process that involves the CpG island methylator phenotype (CIMP).[174] About half of the genes in the human genome have promoters that are embedded in clusters of cytosine-guanosine residues called CpG islands.[175] Cytosines in these regions can be methylated by DNA methyltransferases. Methylation is an epigenetic phenomenon by which a cell

permanently silences genes, and it remains in the progeny of each cell. *In vitro*, it is possible to demethylate at least some promoters using the DNA methyltransferase inhibitor 5-azacytidine; however, upon removal of this agent, remethylation occurs. Histone modifications indicate which CpG sites are marked for permanent methylation[176].

A subset of CRC has been identified as characterized by CIMP.[179,180] These tumours progress by methylating tumour suppressor genes and have distinctive clinical features compared to non-CIMP CRC.[180] Some of the commonly methylated promoters have no apparent role in tumour development and have been labelled "methylated in tumor" or MINT genes, but methylation also occurs in promoters of known tumour suppressor genes, including *p16* and *insulin-like growth factor 2*, and DNA repair genes such as *methylguanine methyltransferase* and *MLH1*.

The region of the *MLH1* promoter in which methylation mediates gene silencing is the 3′ end, close to the start codon. The 5′ end of the promoter is also prone to methylation, but this is not of functional importance unless the methylation extends to the critical 3′ region.[181] Therefore, specific CpG residues are more important than others in mediating gene silencing. Importantly, most of the CRCs with sporadic MSI come from a CIMP background, which creates an important distinction from Lynch syndrome tumours.[182,183] Consequently, it is essential to know more about CRC than whether there is MSI or not. It is remarkable that Lynch syndrome tumours are associated with germline mutations in DNA MMR genes, occur in younger individuals, can have *KRAS* mutations (but never *BRAF* mutations), and are associated with a better prognosis than non-MSI tumors.[184] Conversely, most sporadic CRC with MSI occur in older individuals, have *BRAF* mutations in about half of the cases (V600E), are associated with a background of CIMP and a reduced mortality.

1.1.4.2 Germline Epimutations

To aide in the diagnosis of Lynch syndrome, several investigators have searched for germline mutations in a coding regions, splice sites, or promoter regions of MMR genes. However, instead of a genetic mutation, some cases of Lynch syndrome are caused by epigenetic inactivation of genes by promoter methylation. These events, originally described as *germline epimutations*, are better referred to as *constitutional epimutations*, to reflect the aberrant silencing of a gene that is normally active in somatic cells in the absence of a sequence mutation. Constitutional epimutations of *MLH1* [185,186,187]and *MSH2*[188,189] have been identified in families with Lynch syndrome who have no apparent MMR gene sequence mutations. *MSH2* methylation can arise from a deletion in the 3′ end of the gene just upstream of *MSH2* (*TACSTD1*). The loss of the stop codon somehow results in methylation of the CpG island downstream, which leads to constitutional silencing of *MSH2*. Similar to germline mutations, constitutional epimutations in *MSH2* demonstrate classic autosomal dominant inheritance and are associated with a 50% risk of transmission to the offspring.[189,190]

1.1.5 Pathophysiology of Colorectal Carcinogenesis with MSI

CRCs with MSI were discovered, in part, because they were different from most other CRCs; the tumoural cells had a tendency to be diploid and had less LOH. The mutational patterns were of interest because of the unique clinical features of

these tumours. In 1995, Markowitz *et al* provided an important insight into the colorectal carcinogenetic pathway.[17]

It has been showed that TGF-β signalling inhibits proliferation in the colonic epithelium. Markowitz *et al* showed that the TGF-β type II receptor (TGFβR2) was not expressed in colorectal cell lines with MSI, but not without MSI, or in xenografts of these cells grown in nude mice.

TGFβR2 includes two microsatellite sequences: from nucleotides 1931 to 1936, there is a 6-base pair $[GT]_3$ sequence (the inverse of a C–A repeat) and from nucleotides 709 to 718 there are 10 consecutive adenines (A_{10}). Deletion mutations in these microsatellite sequences create frame-shifts that inactivate the gene product. These were found predominantly in the colorectal MSI cell lines. Cells with mutant forms of TGFβR2 did not slow proliferation in response to TGF-β. The most frequent mutation was a single base-pair deletion in the A_{10} sequence. Markowitz *et al* reported that this mutation made the protein nonfunctional.[191] Furthermore, they identified mutations in the A_{10} tract of 90% of 111 CRC found to have MSI, proving that this was a relevant site of this form of genomic instability.[192] In a new paradigm, repetitive DNA sequences were proposed to be particularly sensitive to the loss of DNA MMR activity, resulting in frame-shift mutations that led to premature stop codons and gene inactivation.

A number of other genes affected by MSI have been identified that encoded regulators of cell proliferation, including *GRB1, TCF-4, WISP3, activin receptor-2, insulin-like growth factor-2 receptor, axin-2*, and *CDX*, the cell cycle or apoptosis such as *BAX, caspase-5, RIZ, BCL-10, PTEN, hG4-1*, and *FAS*, and DNA repair comprising *MBD-4, BLM, CHK1, MLH3, RAD50, MSH3*, and *MSH6* [193]. Remarkably, every human MMR gene except *MLH1* includes a mononucleotide repeat of at least A_7,[194] so the process of MMR could become increasingly defective with cumulative losses of components on the system.[195] However, it is not clear how many of the mutations at these loci are of functional significance (which has been determined for *TGFβR2*), or whether some are simply markers of MSI, because biallelic inactivation of these genes has not been documented in all the tumours.

The identification of the proximate targets of carcinogenesis in the setting of defective MMR activity indicates that MSI-associated colorectal tumorigensis occurs through a different biological pathway compared to sporadic tumours. CRC cells with MSI can activate or inactivate the same signalling pathways as those without MSI, but different proteins within these pathways are involved. Genes are altered by different mechanisms in CRC cells with MSI compared to those without MSI. For example, sporadic CRC arise from a combination of mutations and LOH, resulting in biallelic inactivation of *APC*. CRC with MSI have mainly an increase in the number of point mutations compared to cancer cells without MSI, are more likely to be diploid, and do not have experienced widespread allelic imbalance or LOH. A substantial portion of CRC with MSI have normal expression of *APC*, but have mutations in β-catenin that make it unable to interact with APC protein and undergo degradation,[196,197] the biological equivalent of having no APC protein. Moreover, some colorectal tumours with MSI that have neither inactivated *APC* nor mutated *β-catenin* could have frame-shift mutations in factors further downstream in the WNT pathway, such as in the A_9 sequence of *TCF-4*.

Of equal importance are the differences in the pathogenesis of CRC with MSI that come from patients with Lynch syndrome compared with sporadic CRC cells with CIMP.

1.1.6 Models of MSI

Studies of MMR expanded from *Escherichia coli* and *Saccharomyces cerevisiae* to human cells and diseases in the mid-1990s. Parsons et al discovered that HCT116 cells had defects in MMR activity and MSI25 that were associated with biallelic mutations in *MLH1*. Stable transfer of human chromosome 3 into HCT116 (HCT116-ch3 cells) restored *MLH1* and created a model to study the role of MMR in the response to DNA damage and regulation of the cell cycle.[206] MMR activity was restored in HCT116 -ch3 cells, which had reduced tolerance to DNA alkylation by *N*-methyl-*N*=-nitro-*N*-nitrosoguanidine. The model also demonstrated that MMR regulates passage through the G2/M cell cycle checkpoint and the response to 6-thioguanine, which is incorporated into DNA as guanine but acts like a mismatch in newly synthesized DNA.92 MMR activity is required to restrain clonal expansion of cells exposed to *N*-methyl-*N*=-nitro- *N*-nitrosoguanidine and prevent proliferation of cells with damaged DNA.[208] Cultured CRC cells with intact MMR activity were significantly more sensitive to therapeutic concentrations of 5-fluorouracil (5-FU; 5–10 _M) than DNA MMR–deficient cells,[209] as well as other chemotherapeutics. [210,211] This provided the theoretical basis for later observations that patients with MSI CRCs mightnot respond to traditional 5-FU–based chemotherapy. Cell models were subsequently developed to study other genes in the DNA MMR family. Although stable transfer of human chromosome 3 corrected some DNA MMR activity, HCT116-ch3 cells were still devoid of *MSH3* activity. Therefore, the HCT116 -3 - 5 cell line was created by stable transfer of human chromosome 5 into HCT116 -ch3 cells, permitting an exploration of the effects of *MSH3* on MMR.[212] The CRC cell line HCT15 does not express *MSH6* and the endometrial cancer cell line HEC59 does not express *MSH2* or *MSH6*. Stable transfer of human chromosome 2 (which contains *MSH2* and *MSH6*) into these cells created cell lines that permitted analysis of the role of the MutS system in human cells.[213] In the modified HCT15 and HEC59 cell lines, *MSH2* and *MSH6* were re-expressed, MSI and the hypermutable phenotype were corrected (reducing mutations at the *HPRT* locus by 96% in HCT15), and the cells became more sensitive to cell death after exposure to mutagens. Similarly, in LoVo cells, which have defects at the *MSH2* locus, MMR activity and sensitivity to mutagenic agents were restored following transfer of a fragment of chromosome 2.[214] Lynch syndrome is an autosomal dominant disease, associated with the inheritance of a single mutated gene.
However, in cultured cells, defects in MMR are recessive; a single copy of the gene restores DNA MMR activity. MMR activity is lost in cells that lack *MSH2* or *MLH1*, but only partially disrupted in cells without *MSH6* or *MSH3*.

1.1.7 Diagnosis of MSI

Patients with colorectal tumors with MSI have longer survival times than patients with tumors without MSI.[75] A study of 175 patients with Lynch syndrome (120 of whom had Lynch syndrome-MLH1 type), compared to more than 14,000 population-based patients with CRC, showed that the 5-year cumulative relative survival for patients with Lynch syndrome was 65%, compared to 44% of patients with sporadic CRC who were older than 65 years of age.[231] A pooled analysis of

MSI that included 32 studies and 7642 CRC cases found an overall hazard ratio of 0.65 for patients with tumors that have MSI. Patients with Lynch syndrome have lower-stage disease when they are identified compared to patients with other types of CRC, and it is less common for patients with Lynch syndrome to present with metastatic disease[232-234]. Gryfe et al reported that 17% of 607 CRC patients who were younger than 50 years old had MSI; the hazard ratio for patients with MSI-associated tumors was 0.42, compared with patients in the same cohort with non- MSI tumors. Patients with tumors with MSI had lower mortality rates when patients were stratified by tumor stage, including patients with stage IV cancer.[235]

Tumors with MSI have greater numbers of tumor-infiltrating lymphocytes that are activated and cytotoxic;[236] the lymphocytic reaction is independently associated with longer survival.[237] Studies report that colorectal tumors with lymphocytic infiltrates are associated with longer survival times; in fact, lymphocyte infiltration may be a better prognostic factor than routine pathology staging. [238] Abnormal results from immunohistochemical analyses of MSH2 and MLH1 independently predict better outcomes in patients with stage II–T3 colorectal tumors. [239] Therefore, detection of MSI in a patient with CRC is a positive prognostic factor, particularly among young patients. However, not all studies have confirmed the value of MSI detection in prognosis.[240,241] This might be because there are technical challenges in identifying tumors with MSI. Some studies could have included CRC cases that were falsely reported to have MSI-positive tumors, which might have altered the modest differences in predicted patient outcomes.[241,242] Second, MSI analyses include 2 groups of patients: those with Lynch syndrome (who tend to be younger) and those with acquired methylation of the *MLH1* gene (who are older). If Lynch syndrome patients are excluded from a cohort, the average age of patients with MSI is older—older patients have shorter overall survival times regardless of whether or not they have CRC. Most evidence, however, indicates that MSI predicts positive outcomes in patients with CRC. The principal use of MSI testing in the clinic is to identify patients with Lynch syndrome. Approximately 15% of all colorectal tumors have MSI, and 75%–80% of this group have acquired methylation of *MLH1*; only 2%–3% of all CRCs have germline mutations in one of the MMR genes.49 Although MSI analysis is the first approach to identifying patients with Lynch syndrome, immunohistochemical analysis is equivalent in levels of sensitivity and specificity and is accessible to most pathology laboratories.

Several panels of microsatellites have been used to diagnose MSI. The first consensus meeting recommended a panel of 3 dinucleotide repeats and 2 mononucleotide repeats that had been validated by a German consortium. This panel required that normal tissue be compared with tumor tissue, so many groups use a panel of 5 mononucleotide repeats that can be amplified and analyzed in a single assay. The location and nature of each of the microsatellite targets used for diagnostic purposes are listed in Table 2. The Pentaplex panel is as sensitive and specific as the initial panel proposed by the National Cancer Institute, but has some advantages that improve specificity. The T25 repeat in the 3'-untranslated region of *caspase-2* is useful in detecting loss of *MSH6*.

1.1.8 Clinical Implications of CIN and MSI

The insights into the genetic basis of CRC have allowed the identification of new prognostic and predictive molecular biomarkers as well as novel therapeutic strategies. Many Authors have attempted to determine whether *KRAS, TP53*, or 18q alterations could serve as prognostic markers. Some data have indicated an increased risk of relapse and death among patients with a codon 12 mutation in *KRAS*, but other studies have failed to confirm this association.[231,232,233,234] Similarly, some studies suggested that a *TP53* mutation is associated with a higher risk of death, but this appears to be confined to those who already have the worst prognosis (Dukes' stage D).[243,235] A number of compelling studies associated allelic deletion of chromosome 18q with poor outcomes.[236,237] However, recent studies have failed to validate these earlier observations.[238,239]] Consequently, none of proposed individual markers are actually used as molecular prognostic factors.[240] However, data indicate that the overall CIN phenotype is associated with a less favourable outcome in patients compared with those with tumours that exhibit MSI. Patients with CIN consistently exhibited low rates of overall and progression-free survival, irrespective of ethnicity, anatomic location, or treatment with 5-fluorouracil, compared to patients with MSI tumours. Additionally, prognosis is not influenced by adjuvant therapy in patients with stage II–III CRC.[241]

Therapeutic targeting of pathways that directly initiate and perpetuate CIN has now reached the clinical practice. Small-molecule inhibitors of Aurora kinases, Plks, the spindle motor protein Eg5, and CENP-E have shown anti-tumour activity in preclinical models and are currently being evaluated in phase I and II clinical trials for the treatment of solid tumours. Of particular interest, is the observation by Swanton *et al* that CIN-positive tumours are intrinsically resistant to taxanes due to the similarity between pathways that regulate the separation of chromosomes at mitosis and those implicated in the response to taxanes.[242] This serves as the basis for the Chromosomal Instability and Anti-Tubulin Response Assessment (CINATRA) trial to assess whether patients with near diploid MSI-positive CRC derive benefit from EPO906, a new microtubule stabilizer, relative to patients with CIN-positive cancers.

The time course of progression from adenoma to cancer provides a window of opportunity for chemoprevention. COX-2 inhibitors successfully prevent polyp recurrence. Three prospective, randomized, placebo-controlled, multicenter trials of secondary prevention of colorectal adenomas [Adenomatous Polyp Prevention on Vioxx (APPROVe), Adenoma Prevention with Celecoxib (APC), and Prevention of Sporadic Adenomatous Polyps (PreSAP) trials] demonstrated that selective COX-2 inhibitors reduced polyp recurrence, with a greater effect on formation of advanced adenomas.[242,243]

In the future, attempts to better define the mechanisms that initiate CIN, the relationship between CIN and tumour progression, and the feasibility of targeting chromosomally unstable cells will be critical to advance our understanding of the most common form of genetic instability in CRC.

1.2 EPITHELIAL TO MESENCHYMAL TRANSITION

1.2.1 What is epithelial-mesenchymal transition?

Epithelial-to-mesenchymal transition (EMT) refers to a biologic dynamical process that allows a polarized *epithelial cell*, which normally interacts with basement membrane via its basal surface, to undergo multiple biochemical changes that enable it to assume a *mesenchymal cell* phenotype, which includes enhanced migratory capacity, invasiveness, elevated resistance to apoptosis, and greatly increased production of extra-cellular matrix (ECM) components [246]. The completion of an EMT is signalled by the degradation of underlying basement membrane and the formation of a mesenchymal cell that can migrate away from the epithelial layer in which it originated. A number of distinct molecular pathways are engaged to initiate the EMT and enable it to reach completion. These include activation of transcription factors, expression of specific cell-surface proteins, reorganization and expression of cytoskeletal proteins, production of ECM-degrading enzymes, and changes in the expression of specific microRNAs. In many cases, the involved factors are used as biomarkers to demonstrate the dynamics of EMT (**Figure 1.4**).

The pioneering study of Elizabeth Hay first described an "epithelial-mesenchymal transformation" using a model of chick primitive streak formation [247]. Subsequently, the term "transformation" has been replaced with "transition," reflecting in part the reversibility of the process and the fact that it is distinct from neoplastic transformation [246]. The phenotypic plasticity afforded by an EMT is revealed by the occurrence of the reverse process, called mesenchymal-to-epithelial transition (MET), which involves the reversal of a mesenchymal cell to its originating epithelial cell.. Today, a few data are still available about this process. The most-studied example is the MET associated with kidney formation, which is driven by genes including paired box 2 (*Pax2*), bone morphogenetic protein 7 (*Bmp7*), and Wilms tumor 1 (*Wt1*) (248-250).

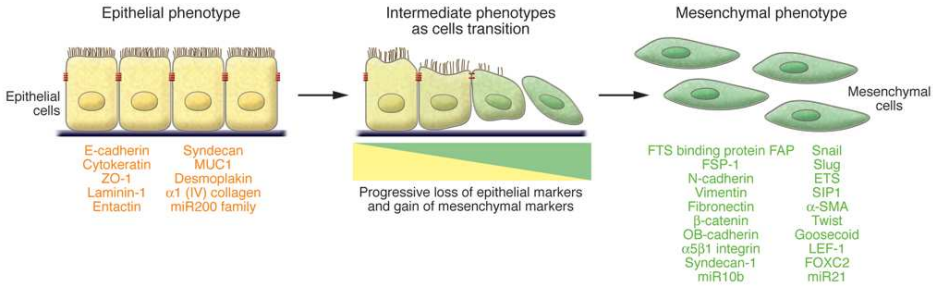


Figure 1.4 *EMT involves a functional transition of polarized epithelial cells into mobile and ECM component–secreting mesenchymal cells. The epithelial and mesenchymal cell markers commonly used by EMT researchers are listed. Colocalization of these two sets of distinct markers defines an intermediate phenotype of EMT, indicating cells that have passed only partly through an EMT. Detection of cells expressing both sets of markers makes it impossible to identify all mesenchymal cells that originate from the epithelia via EMT, as many mesenchymal cells likely shed all epithelial markers once a transition is completed. For this reason, most studies in mice use irreversible epithelial cell–lineage tagging*

to address the full range of EMT-induced changes. ZO-1, zona occludens 1; MUC1, mucin 1, cell surface associated; miR200, microRNA 200; SIP1, survival of motor neuron protein interacting protein 1; FOXC2, forkhead box C2.

1.2.2 Why does EMT occurs?

The concept of cell division as a way to generate more cells and expand tissue size emerged about 150 years ago [250]. Central to this concept was an understanding that all cells in a body derive from other cells and the resulting deduction that ultimately all are derived from a single cell, the fertilized egg. An additional level of complexity came from the realization that cells can assume various phenotypic states during their development. In other words, they can undergo the process of differentiation. As was learned more recently, during specific steps of embryogenesis, the cells within certain epithelia appear to be *plastic* and thus able to move back and forth between epithelial and mesenchymal states via the processes of EMT or MET [251]. Upon completion of the development of human organs, the epithelial cells typically exert tissue-specific functions, while the mesenchymal cells play a supporting role. This notion implies that a state of terminal differentiation is necessary to carry out such specialized functions and that cells are maintained in a permanent state of differentiation once development is complete.

This concept has been challenged by several observations that the cells within a terminally differentiated epithelium can indeed change their phenotype through activation of an EMT program, which enables transdifferentiation, resulting in the conversion of epithelial cells to mesenchymal derivatives during development and adulthood. These programs can also be activated in association with tissue repair and pathological stresses, including those involving various types of inflammation and high-grade carcinomas.

The above considerations make it possible to,state that EMT is the main mechanism involved to disperse cells in embryos, generate mesenchymal cells in injured tissues, and promote the invasive and metastatic behaviour of epithelial cancers.

1.2.3 Classification of EMT into three different subtypes

EMTs are encountered in three distinct biological settings that carry very different functional consequences (**Figure 1.5**). While the specific signals that delineate the EMT in the three discrete settings are not yet clear, it is now well accepted that functional distinctions are apparent. A proposal to classify EMT into three different biological subtypes based on the biological context in which they occur was discussed at a 2007 Meeting on EMT in Poland and a subsequent Meeting in March 2008 at Cold Spring Harbor Laboratories.

The EMTs that are associated with implantation, embryo formation, and organ development are organized to generate diverse cell types that share common mesenchymal phenotypes.This class of EMT termed as “type 1,” neither causes fibrosis nor induces an invasive phenotype resulting in systemic spread via the circulation. Among other outcomes, type 1 EMT can generate mesenchymal cells (primary mesenchyme) that have the potential to subsequently undergo a MET to generate secondary epithelia.

The EMT associated with wound healing, tissue regeneration, and organ fibrosis is defined as ”type 2”. In type 2 EMT, the program begins as part of a repair-associated event that normally generates fibroblasts and other related cells to reconstruct tissues following trauma and inflammatory injury. However, in contrast to type 1 EMT, type 2 EMT is associated with inflammation and cease once inflammation is attenuated, as is seen during wound healing and tissue regeneration. In the setting of organ fibrosis, type 2 EMT can continue to respond to ongoing inflammation, leading eventually to organ destruction. Tissue fibrosis is in essence an unabated form of wound healing due to persistent inflammation.
Type 3 EMT occurs in neoplastic cells that have previously undergone genetic and epigenetic changes, specifically in genes that favour clonal outgrowth and the development of localized tumours. These changes, notably affecting oncogenes and tumour suppressor genes, conspire with the EMT regulatory circuitry to produce outcomes far different from those observed in the other two types of EMT. Cancerous cells undergoing a type 3 EMT may invade and metastasize and thereby generate the final, life-threatening manifestations of cancer progression. Importantly, cancer cells may pass through EMTs to differing extents, with some cells retaining many epithelial traits while acquiring some mesenchymal ones and other cells shedding all vestiges of their epithelial origin and becoming fully mesenchymal. It still remains unclear what specific signals induce type 3 EMT in carcinoma cells. It is known that such signals may originate in the tumour stroma that is associated with many primary carcinomas.
While these three classes of EMT represent distinct biological processes, a common set of genetic and biochemical elements appears to underlie and thus enable these outwardly diverse phenotypic programs.

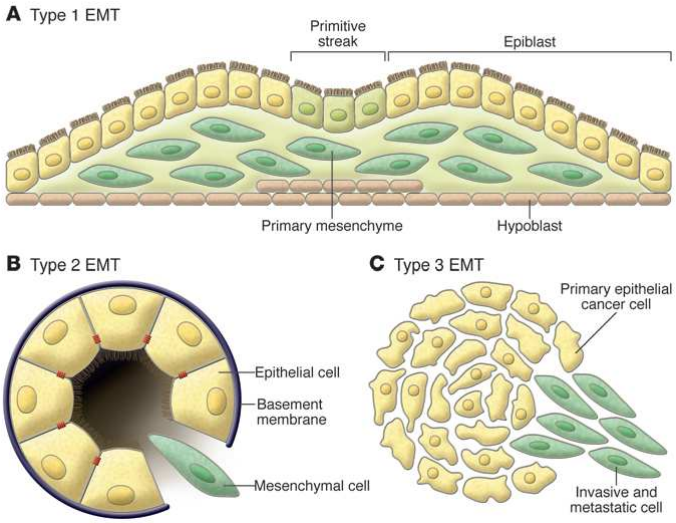


Figure 1.5 *Different types of EMT. (**A**) Type 1 EMT is associated with implantation and embryonic gastrulation and gives rise to the mesoderm and endoderm and to mobile neural crest cells. The primitive epithelium, specifically the epiblast, gives rise to primary mesenchyme via an EMT. This primary mesenchyme can be re-induced to form secondary epithelia by a MET. It is speculated that such secondary*

*epithelia may further differentiate to form other types of epithelial tissues and undergo subsequent EMT to generate the cells of connective tissue, including astrocytes, adipocytes, chondrocytes, osteoblasts, and muscle cells. (**B**) EMTs are re-engaged in the context of inflammation and fibrosis and represent the type 2 EMTs. Unlike the type 1 EMT, the type 2 EMT is expressed over extended periods of time and can eventually destroy an affected organ if the primary inflammatory insult is not removed or attenuated. (**C**) Finally, the secondary epithelia associated with many organs can transform into cancer cells that later undergo the EMTs that enable invasion and metastasis, thereby representing type 3 EMTs.*

1.2.3.1 Type 1 EMT: EMT during implantation, embryogenesis, and organ development

Following the earliest stages of embryogenesis, the implantation of the embryo and the initiation of placenta formation are both associated with an EMT that involves the parietal endoderm [253]. In particular, the trophoectoderm cells, which are precursors of the cytotrophoblast, undergo an EMT to facilitate invasion of the endometrium and the subsequent proper anchoring of the placenta, enabling its function in nutrient and gas exchange [254,255]. As described briefly below, this is only the first of many type 1 EMT that accompany and underlie embryonic morphogenesis; details regarding some of the EMT that occur during development are discussed elsewhere [256].

A fertilized egg undergoes *gastrulation*, generating three germ layers. Initially, a primitive streak is generated in the epiblast layer [257]. The epithelial cells in this tissue express E-cadherin and exhibit apical-basal polarity. Formation of the primitive streak is considered to be the first sign of gastrulation, which leads in turn to the formation of the three germ layers that generate all tissue types of the body. The primitive streak is formed from a furrowed invagination in the midline of the epiblast layer that forms initially at the lower extremity of the embryo and, later on, extends in the direction of the future head. The epithelial-like cells of the epiblast undergo programmed changes dictated by specific expression of proteins associated with cell migration and differentiation [258].

Once formed, the primitive streak, acting via invagination or ingression, generates the mesendoderm, which subsequently separates to form the mesoderm and the endoderm via an EMT (also known as epiblast-mesoderm transition) [247] by replacing the hypoblast cells, which presumably either undergo apoptosis or contribute to the mesoderm layer via an EMT. The embryonic mesoderm that forms between the epiblast and hypoblast eventually gives rise to primary mesenchyme associated with the axial, paraxial, intermediate, and lateral plate mesodermal layers [259]. Cells of the primary mesenchyme exhibit enhanced migratory properties when compared with those of the epiblast and the hypoblast.

At the biochemical level, the EMT associated with gastrulation is dependent on and orchestrated by canonical Wnt signaling, and embryos deficient in Wnt3 cannot undergo the EMT associated with gastrulation [260,261] The subsequent formation of the primitive streak is associated with expression of Wnt8c, and ectopic expression of Wnt8c in embryos leads to multiple primitive streaks [262,263]. TGF-β superfamily proteins, notably Nodal and Vg1, mediate the action of Wnts, and their deficiencies can lead to mesodermal defects due to the absence of functional EMTs [264,265,266,267,268]. Wnts also cooperate with FGF receptors to help

regulate an EMT associated with gastrulation [269-272]. The Snail, Eomes, and Mesps transcription factors orchestrate the EMT associated with gastrulation (273-275]. It has been shown that Snail represses E-cadherin and induces EMT mediated by cell adhesion molecules, such as occludins and claudins, and by polarity genes, including Discs large (Dlg) and Crumbs homolog 3 (Crb3) [276-278].

During embryonic development, an EMT involving the epithelial cells of the neuroectoderm gives rise to migratory neural crest cells [279]. Initially, the premigratory neural crest cells express genes such as Sox, Snail, Slug, and forkhead box D3 (FoxD3), and these cells subsequently undergo an EMT [280,281]. As a consequence, they then dissociate from the neural folds, become motile, and disperse to the different parts of the embryo, where they undergo further differentiation into, among other cell types, the melanocytes that provide pigment to the skin.

This EMT occurring in the neural crest is triggered by signalling pathways similar to those orchestrating the EMT associated with gastrulation. Thus, signalling pathways mediated by Wnts, FGFs, BMPs, c-Myb, and msh homeobox 1 (Msx-1) conspire to induce EMT [282-283]. Among them, BMP most prominently induces the

migratory property of neural crest cells. Noggin, an inhibitor of BMP, is essential in negatively regulating this activity, and its expression places a hold on EMT [284,285]. Additionally, E-cadherin and N-cadherin need to be repressed in order for neural crest EMT to occur [287]. By some estimates, EMT programs are deployed during several subsequent phases of embryogenesis, indicating the continued involvement of these programs later in development, such as the endothelial-mesenchymal transition (EndMT) that takes place during heart valve formation.

1.2.3.2 Type 2 EMT: EMT associated with tissue regeneration and organ fibrosis

Organ fibrosis, which occurs in a number of epithelial tissues, is mainly mediated by inflammatory cells and fibroblasts that release a variety of inflammatory signals as well as components of a complex ECM that includes collagens, laminins, elastin, and tenacins. More specifically, such EMT are found to be associated with fibrosis occurring in kidney, liver, lung, and intestine [288-291].

Some of the earliest proof of this came from the study of transgenic mice bearing germ-line reporter genes whose expression was driven by epithelial cell-specific promoters. The behaviour of these reporters provided direct evidence for epithelial cells serving, via EMTs, as important precursors of the fibroblasts that arise during the course of organ fibrosis [292-294].

Fibroblast-specific protein 1 (FSP1; also known as S100A4 and MTS-1), an S100 class of cytoskeletal protein, alpha-smooth muscle antigen (α-SMA), and collagen I have provided reliable markers to characterize the mesenchymal products generated by the EMT that occur during the development of fibrosis in various organs [292-293]. These markers, along with discoidin domain receptor tyrosine kinase 2 (DDR2), vimentin, and desmin, have been largely used to identify epithelial cells of the kidney, liver, lung, and intestine that are in the midst of undergoing an EMT associated with chronic inflammation. Such cells continued to exhibit epithelial-specific morphology and molecular markers, such as cytokeratin and E-cadherin, but showed concomitant expression of the FSP1 mesenchymal

marker and α-SMA. Such cells are likely to represent the intermediate stages of EMT, when epithelial markers continue to be expressed but new mesenchymal markers have already been acquired. The behaviour of these cells provided one of the first indications that epithelial cells under inflammatory stresses can advance to various extents through an EMT, creating the notion of "partial EMTs" (**Figure 1.6**).
It has been shown that these cells leave the epithelial layer, negotiate their way through the underlying basement membrane, and accumulate in the interstitium of the tissue [296], where they ultimately shed all of their epithelial markers and gain a fully fibroblastic phenotype.
Recent data in mice have demonstrated that endothelial cells associated with the microvasculature can also contribute to the formation of mesenchymal cells during the course of fibrosis, doing so via an analogous process like EndMT.

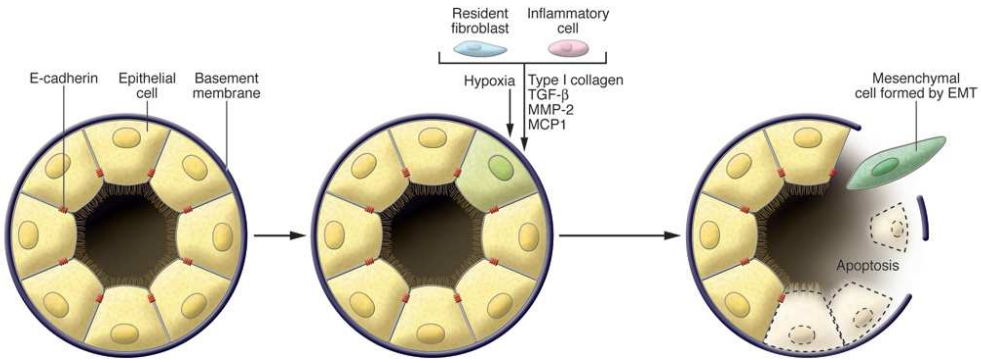


Figure 1.6 *EMT and fibrosis. The EMTs associated with fibrosis are associated with inflammation and the generation of numerous types of molecules by inflammatory cells and resident activated fibroblasts (myofibroblasts). These molecules cause disruption of the epithelial layers via degradation of the basement membrane. The epithelial cells lose polarity and either undergo apoptosis (the majority of cells) or EMT (the minority of cells). MCP1, monocyte chemoattractant protein 1*

This pathologic process echoes a similar normal process occurring during development [288]. Thus, during embryogenesis, an EndMT occurs during the organization of the endocardial cushion and the heart valves. In cardiac fibrosis associated with post-ischemic injury of the heart, EndMT, which involves both the endocardium (the inner endothelial layer of the heart) and the microvascular endothelium of the heart, has been demonstrated to play a key role in contributing to the emergence of newly formed fibroblasts [289] (**Figure 1.6**). In tissue culture models of EndMT, TGF-β1 induces EndMT of capillary endothelial cells and the loss of endothelial markers, such CD31 and integrin αVβ3, as well as the acquisition of fibroblast- and myofibroblast-specific markers such as FSP1, α-SMA, DDR2, collagen I, and vimentin. While not yet documented, it is plausible that many of the molecular regulators of EMT also play critical roles in orchestrating EndMTs.
In one analysis [246], lineage-tagging experiments and bone marrow transplant studies demonstrated that during the course of kidney fibrosis in mice, about 12%

of fibroblasts are derived from the bone marrow and about 30% are derived via EMT from the tubular epithelial cells of the kidney. In addition, it has been shown more recently that in kidney fibrosis about 35% of fibroblasts are derived via EndMT from the endothelial cells normally residing within the kidney. The remaining portions are speculated to arise via activation of resident fibroblasts or other mesenchymal cells, such as perivascular smooth muscle cells/pericytes and fibrocytes [297] in the circulation. It is likely that these proportions may vary dramatically, depending on the stage of fibrosis, the organ, and the particular experimental model being studied.

Inflammatory injury can result in the recruitment of a diverse array of cells that can trigger an EMT through their release of growth factors, such as TGF-β, PDGF, EGF, and FGF-2 [301]. In kidney, most prominent among these cells are macrophages and activated resident fibroblasts that accumulate at the site of injury and release these growth factors. In addition, these cells release chemokines and MMPs, notably MMP-2, MMP-3, and MMP-9. Epithelial cells come under the influence of these signalling molecules and, acting together with the inflammatory cells, induce basement membrane damage and focal degradation of type IV collagen and laminin [301]. Delaminated epithelial cells may then migrate toward the interstitial area (the space between epithelial layers) under the influence of gradients of growth factors and other chemoattractants [301]. This initial recruitment of epithelial cells into an EMT can be inhibited by blocking the expression of MMP-9 through the disruption of tissue plasminogen activator (tPA) [302].

Other studies have also demonstrated that Hepatocyte-growth factor (HGF) can decrease levels of TGF-β, restore TGF-β–mediated loss of E-cadherin, and potentially decrease amounts of active MMP-9 [303]. β1 integrin and integrin-linked kinase (ILK) are also identified as important mediators of the TGF-β-induced EMT associated with tubular epithelial cells [304]. TGF-β induces EMT via both a Smad2/3-dependent pathway and a MAPK-dependent pathway. Recently it has been also demonstrated a key role for the E-cadherin/β-catenin signaling axis for EMT involving epithelial cells [305,306].

The significance of TGF-β-induced EMT for progression of organ fibrosis has been showed using BMP-7, an antagonist of TGF-β signalling, in mouse models of kidney, liver, billiard tract, lung, and intestinal fibrosis [307]. BMP-7 functions as an endogenous inhibitor of TGF-β-induced EMT [307]. Among other effects, it reverses the TGF-β-induced loss of the key epithelial protein, E-cadherin [307]. Restoration of E-cadherin levels by BMP-7 is mediated via its cognate receptors, activinlike kinase-2/3/6 (ALK-2/3/6), and the Smad4/5 downstream transcription factors. Systemic administration of recombinant BMP-7 to mice with severe fibrosis resulted in reversal of EMT and repair of damaged epithelial structures, with repopulation of healthy epithelial cells, all presumably mediated via a MET. This reversal was also associated with restoration of organ function, a substantial decrease in $FSP1^+$ and $\alpha\text{-}SMA^+$ interstitial fibroblasts, and the *de novo* activation of BMP-7 signalling [307].

In a study of 133 patients with kidney fibrosis, an EMT was demonstrated in a substantial number of the specimens, as evaluated using double labelling of the tubular epithelial cells with cytokeratin, vimentin, α-SMA, or zona occludens 1 (ZO-1). Similarly, in patients with Crohn disease, an EMT was demonstrated in areas of fibrosis in the colon [308]. The above studies suggest the necessary confidence to

develop novel therapeutic interventions to suppress EMTs and potentially reverse organ fibrosis.

1.2.3.3 Type 3 EMT: EMT associated with cancer progression and metastasis

Excessive epithelial cell proliferation and angiogenesis represent the hallmark of the initiation and early growth of primary epithelial cancers [309]. The subsequent acquisition of invasiveness, initially manifested by invasion through the basement membrane, is thought to herald the onset of the last stages of the multi-step process that leads eventually to metastatic dissemination, with life-threatening consequences. The genetic and biochemical mechanisms underlying the acquisition of the invasive phenotype and the subsequent systemic spread of the cancer cell have been areas of intensive research. In many of these studies, activation of an EMT program has been proposed as a critical mechanism for the acquisition of malignant phenotypes by epithelial cancer cells [310].

Many studies conducted in mouse models as well as *in vitro* have shown that cancer cells can acquire a mesenchymal phenotype and express mesenchymal markers such as α-SMA, FSP1, vimentin, and desmin [311]. These cells are mainly located at the invasive front of the tumour, and are considered to be the cells that might enter in the subsequent steps of the invasion-metastasis cascade, that involves intravasation, transport through the circulation, extravasation, colonization of distant organs, formation of micrometastases and growth of small colonies into macroscopic metastases [310,311,312].

An apparent paradox comes from the observation that the EMT-derived migratory cancer cells typically establish secondary colonies at distant sites that morphologically resemble the primary tumour; accordingly, they no longer exhibit the mesenchymal phenotypes ascribed to metastasizing carcinoma cells. Reconciling this behaviour with the proposed role of EMT as a facilitator of metastatic dissemination requires the additional notion that metastasizing cancer cells must shed their mesenchymal phenotype via a MET during the course of secondary tumour formation [313]. The tendency of disseminated cancer cells to undergo MET likely reflects the local microenvironments that they encounter after extravasation into the parenchyma of a distant organ, quite possibly the absence of the heterotypic signals they experienced in the primary tumour that were responsible for inducing the EMT in the first site [314,315]. These considerations indicate that induction of an EMT is likely to be a centrally important mechanism for the progression of carcinomas to a metastatic stage and implicates MET during the subsequent colonization process (**Figure 1.7**). However, many steps of this mechanistic model still require direct experimental validation. Moreover, it remains unclear whether these phenomena and molecular mechanisms relate to and explain the metastatic dissemination of non-epithelial cancer cells and the full spectrum of signalling agents that contribute to EMT.. One suggestion is that the genetic and epigenetic alterations undergone by cancer cells during the course of primary tumour development render them especially responsive to EMT-inducing heterotypic signals originating in the tumour-surrounding stroma.

In the case of many carcinomas, EMT-inducing signals emanating from the tumor-associated stroma, notably HGF, EGF, PDGF, and TGF-β, appear to be responsible for the induction or functional activation in cancer cells of a series of EMT-inducing transcription factors, notably Snail, Slug, zinc finger E-box binding homeobox 1 (ZEB1), Twist, Goosecoid, and FOXC2 [319,322]. Once expressed and activated, each of these transcription factors can act pleiotropically to

choreograph the complex EMT program, more often than not with the help of other members of this cohort of transcription factors. The actual implementation by these cells of their EMT program depends on a series of intracellular signalling networks involving, among other signal transducing proteins, ERK, MAPK, PI3K, Akt, Smads, RhoB, β-catenin, lymphoid enhancer binding factor (LEF), Ras, and c-Fos as well as cell surface proteins such as β4 integrins, α5β1 integrin, and αVβ6 integrin [323]. Activation of EMT programs is also facilitated by the disruption of cell-cell adherens junctions and the cell-ECM adhesions mediated by integrins [324-329].

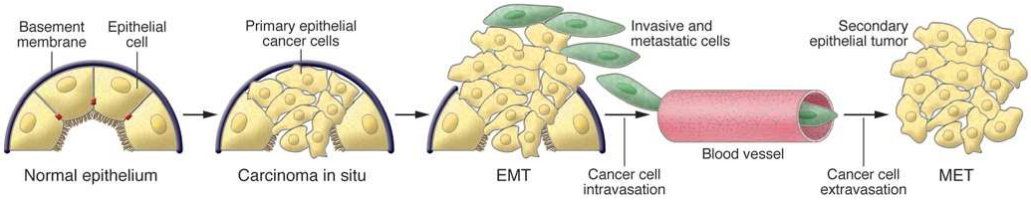


Figure 1.7 *Contribution of EMT to cancer progression. Progression from normal epithelium to invasive carcinoma goes through several stages. The invasive carcinoma stage involves epithelial cells losing their polarity and detaching from the basement membrane. The composition of the basement membrane also changes, altering cell-ECM interactions and signaling networks. The next step involves EMT and an angiogenic switch, facilitating the malignant phase of tumor growth. Progression from this stage to metastatic cancer also involves EMTs, enabling cancer cells to enter the circulation and exit the blood stream at a remote site, where they may form micro- and macro-metastases, which may involve METs and thus a reversion to an epithelial phenotype.*

1.2.4 TGF-β inducers of EMT

TGF-β is an important suppressor of epithelial cell proliferation and thus primary tumorigenesis. However, it is now clear that in certain contexts it can also serve as a positive regulator of tumour progression and metastasis [330-331]. Thus, *in vitro* studies have demonstrated that TGF-β can induce an EMT in certain types of cancer cells [333]. Two possible signalling pathways have been identified as mediators of TGF-β-induced EMT. The first of these involves Smad proteins, which mediate TGF-β action to induce EMTs via the ALK-5 receptor [334-338]. Smad-mediated signalling induced by TGF-β facilitates motility. Inhibitory Smads modulate differential effects of relevant transcription factors and cytoplasmic kinases and induce the autocrine production of TGF-β, which can further reinforce and amplify the EMT program [339,340]]. Signalling pathways that mediate the action of β-catenin and LEF also cooperate with Smads [341] in inducing an EMT [342-343]. In this regard, the involvement of LEF and β-catenin in PDGF-induced EMT was recently described [341]. These studies collectively demonstrate that the TGF-β/Smad/LEF/PDGF axis is an important inducer of an EMT phenotype in cancer.

Evidence for the involvement of a second TGF-β-induced pathway in EMT is also compelling. More specifically, some data indicate that p38 MAPK and RhoA mediate an autocrine TGF-β-induced EMT in NMuMG mouse mammary epithelial cells [339]. This process also requires integrin β1-mediated signalling and the

activation of latent TGF-β by αVβ6 integrin. Fibulin-5, an ECM molecule, augments TGF-β-induced EMT in a MAPK-dependent mechanism [340]. TGF-β can induce an EMT in Ras-transformed hepatocytes, mammary epithelial cells (via MAPK), and MDCK cells; at the same time, Ras-activated PI3K inhibits TGF-β-induced apoptosis to facilitate this transition [341-344].
Evidence for these connections comes from observations that ERK/MAPK and PI3K/Akt pathways mediate Ras mutant-induced EMT, and that such an EMT is reversed by either wild-type Ras or MAPK kinase 1 (MEK1) inhibitors [344]. In this regard, Raf also mediates TGF-β-induced EMT and promotes invasiveness of cancer cells. In mouse models of skin carcinoma and human CRC, the absence of TGF-β receptor expression actually confer better prognosis [345,346]. The connection between inflammation and EMT has been demonstrated by the fact that COX-2 inactivate Smad signaling and enhance EMT stimulated by TGF-β through a PGE2-dependent mechanism [. Changes in the expression of certain cell polarity proteins may also play an important role in TGF-β–induced EMT, since evidence of a role for partitioning detective protein 6 (Par6) in this process is emerging [353].
The connection between loss of E-cadherin expression by cancer cells and passage through an EMT has been established by many studies [354-355]. For example, induction of the c-Fos oncogene in normal mouse mammary epithelial cell lines induces an EMT and is associated with a decrease in E-cadherin expression [342]. Moreover, epithelial cell adhesion complexes reorganize and cell proliferation is suppressed when the full-length or the cytoplasmic portion of E-cadherin (containing the β-catenin binding site) is ectopically expressed in cells that have passed through an EMT, causing such cells to lose their mesenchymal phenotype [356]. Sequestration of β-catenin in the cytoplasm is important for the preservation of epithelial features of cancer cells, and acquisition of the mesenchymal phenotype correlates with the movement of β-catenin to the nucleus, where it becomes part of Tcf/LEF complexes [357]. Such β-catenin accumulation in the nucleus, which is often associated with loss of E-cadherin expression, correlates with susceptibility to enter into an EMT and acquisition of an invasive phenotype. Thus, cells that lose E-cadherin become more responsive to induction of an EMT by various growth factors.
Several Authors have demonstrated that the epigenetic control of E-cadherin and β-catenin/LEF activity is important in establishing the metastatic potential of cancer cells [358-360]. In addition, it has also been shown that cell lines that lack E-cadherin show increased tumorigenicity and metastasis when transferred into immunodeficient mice [361]. E-cadherin expression levels vary dramatically in different human tumours, and an inverse relationship between levels of E-cadherin and patient survival has been documented [360]. In this regard, mutations in the E-cadherin gene have been identified in cancer cells, making them more susceptible to EMT and metastasis. A more thorough analysis of such mutations and their correlation to metastatic progression is still required.
The central role played by E-cadherin loss in the EMT program is further suggested by the actions of several EMT-inducing transcription factors that facilitate acquisition of a mesenchymal phenotype, such as Snail and Slug, as well as those encoding two key zinc finger-containing basic helix-loop-helix transcription factors, survival of motor neuron protein interacting protein 1 (SIP1) and E12 (also indicated as E47-E2A). These transcription factors are induced by TGF-β exposure and, once expressed, repress E-cadherin expression. Snail also facilitates an

invasive phenotype in mice [275]. Loss of E-cadherin promotes Wnt signaling and is associated with high levels of Snail in the nucleus [362]. SIP1 represses E-cadherin expression and binds, along with Snail, to the E-cadherin promoter in an overlapping fashion [363]. The expression of Snail and E-cadherin correlates inversely with the prognosis of patients suffering from breast cancer or oral squamous cell carcinoma [365]. The use of gene expression analyses to compare expression of genes in metastatic and non-metastatic mouse breast cancer cell lines has led to the identification of Twist and Goosecoid as important genes that facilitate EMT and induce metastasis[325] . It has been reported that matrix-degrading enzymes such as MMP-3 facilitate EMT by inducing genomic instability via Rac1b and ROS [366].

Noncoding microRNAs are also components of the cellular signalling circuitry that regulates the EMT program. microRNA 200 (miR200) and miR205 inhibit the repressors of E-cadherin expression, ZEB1 and ZEB2, and thereby help in maintaining the epithelial cell phenotype [367-372]. In breast carcinoma, a loss of miR200 correlates with increased expression of vimentin and a decrease in the levels of E-cadherin in cancer cells. Acting in the opposite direction, miR21 is up-regulated in many cancers and facilitates TGF-β-induced EMT [372]. Interestingly, CD44hiCD20lo cells purified from normal and malignant breast cancer tissue exhibit features of an EMT, and human cancer cells induced to undergo EMT exhibit stem cell-like properties and increased metastatic potential [327]. Therefore, EMT may play a role in the generation of high-grade invasive cells with stem cell-like features, and the latter phenotype, which includes self-renewal potential, may facilitate the formation of secondary tumours by disseminating cancer cells, a notion that still requires direct demonstration.

1.2.5 EMT Biomarkers

Table Ib shows the commonest markers, some of which are acquired and other are attenuated or repressed during the EMT. All of the available biomarkers applied to investigate the subtypes of EMT can be grouped in four classes: a) cell-surface markers, b) Cytoskeletal markers, and c) extra-cellular proteins and d) transcription factors.

Cell-surface markers of EMT. E-cadherin is the prototypical epithelial cell marker. It is expressed in epithelial cells, and its expression is decreased during EMT in embryonic development, tissue fibrosis, and cancer. Moreover, loss of E-cadherin function promotes EMT [381]. In recent years, changes in the level of expression of different cadherins, so-called “cadherin switches”, have been increasingly used to monitor EMT. Indeed, the cadherin switch from E-cadherin to N-cadherin, which is expressed in mesenchymal cells, fibroblasts, cancer cells, and neural tissue, has often been used to monitor the progression of EMT during embryonic development and the advancement of cancer. In addition, because OB-cadherin is a more definitive marker for activated fibroblasts, an E-cadherin-OB-cadherin switch is of interest for type 2 EMT associated with fibrogenesis.

EMT is associated classically with a relocation of cells from a basement membrane microenvironment into a fibrillar ECM. A change in the level of expression of different integrins *i.e.* an integrin switch often reflects alterations in cell-ECM interactions. In addition, integrin signalling facilitates EMT [383], and various integrins are expressed on both epithelial and mesenchymal cells.

Usually, integrins have limited utility as biomarkers for EMT. There are, however, examples where integrins can be used as EMT markers in a context-dependent manner.

In CRC, only cancer cells that have undergone type 3 EMT to a metastatic phenotype express high levels of β6 integrin (normal epithelial cells and non-invasive cancer cells have low-level expression) [384]. During gastrulation, type 1 EMT is associated with *de novo* expression of α5β1, which is a receptor for fibronectin [385]. Similarly, type 2 EMT in experimental kidney fibrosis is associated with increased α5 integrin expression [386]. Increased expression of α5 integrin also correlates with the metastatic potential of B16F10 melanoma cells and EMT [387], suggesting that α5 integrin plays a role in each subtype of EMT.

Another marker that reflects adaptation to the altered ECM microenvironment associated with EMT is the collagen-specific receptor tyrosine kinase (DDR2, discoidin domain receptor tyrosine kinase 2) [388]. Upon binding to type I or type X collagen, DDR2 mediates up-regulation of MMP1 and cell motility [389]. In adult tissues, expression of DDR2 is confined to subsets of fibroblasts and vascular smooth muscle cells, whereas *de novo* expression of DDR2 in endothelial cells is associated with type 2 EMT [390].

DDR2 expression in cancer cells correlates with increased invasiveness, demonstrating its utility in identifying type 3 EMT [391].

Cytoskeletal markers of EMT. FSP1 is a member of the family of $Ca2^{+}$-binding S100 proteins. It is a prototypical fibroblast marker for detecting EMT in cancer and fibrogenesis [373,377]. Most epithelial cells undergoing type 2 EMT express FSP1 early in transition to fibroblasts, and lineage tagging in transgenic reporter mice reveals that more than one-third of all $FSP1^{+}$ fibroblasts in fibrotic kidneys and livers are EMT derived [377,378]. In addition, in models of cancer, metastatic cells often express FSP1 as a part of the molecular program of type 3 EMT. Ectopic expression of FSP1 itself facilitates EMT in adult epithelial cells and cancer cells [379]. In embryonic development, FSP1 is expressed from E8.5 in the mouse, well after mesenchymal cells first appear. Because FSP1 is mainly confined to transitioning epithelial cells and fibroblasts as opposed to primitive mesenchymal cells, it has limited utility for detection of type 1 EMT.

A controversial marker of EMT is the intermediate filament vimentin, which is expressed in various cells, including fibroblasts, endothelial, hematopoietic and glial cells [392-393]. During the mouse embryonic development, vimentin is first expressed at E6.5–E7.0, and it is used as a marker for type 1 EMT during the gastrulation phase [394,395]. However, because adult epithelial cells transiently express vimentin in response to various insults [396], vimentin should not be considered a marker of type 2 EMT in the setting of fibrosis. By contrast, vimentin is commonly used to identify cells undergoing type 3 EMT in cancers [397].

This is based on a positive correlation of vimentin expression with increased invasiveness and metastasis [398].

α-SMA is one of six actin family members. In the adult, prominent α-SMA expression can be found in vascular smooth muscle cells and myoepithelial cells [399]. During embryonic development, the EMT that gives rise to the endocardial cushion is characterized eventually by *de novo* α-SMA expression [400]. Type 2 EMT is also sometimes associated with cells that eventually express α-SMA as myofibroblasts. Although some of these cells that express both α-SMA and FSP1 are known as myofibroblasts, most $\alpha\text{-SMA}^{+}$ cells express only α-SMA and are

distinct from EMT-derived fibroblasts. In cancer, evidence that type 3 EMT is associated with α-SMA is mostly confined to breast cancer [401]here α-SMA largely is detected in breast tumours of the “basal phenotype” [402]

β-Catenin is a cytoplasmic plaque protein that plays a dual role in EMT: it links cadherins to the cytoskeleton and serves as a co-transcriptional activator together with T cell factor (TCF)/LEF [403]atenin activity is mainly regulated by mechanisms controlling the level of β-catenin in the cytoplasm, through either its recruitment to cadherin-binding partners or ubiquitination and subsequent degradation [404] β-catenin/TCF/LEF complex directly controls gene expression associated with EMT, particularly Snail1 [405]Catenin has been used as a marker of EMT in embryonic development, cancer, and fibrosis [406]. Although β-catenin is localized to cell membranes in normal epithelial cells and non invasive tumoural cells, it is located either in the cytoplasm (as reflective of dissociation from E-cadherin) or in the nucleus (as reflective of its role as a transcriptional activator) in cells that undergo EMT [374].

Extracellular proteins. Fibronectin is a high-molecular weight glycoprotein that serves as a scaffold for fibrillar ECM [407]. Because it is one of the first molecules to

appear when the fibrillar ECM is formed, it has been used as an indicator of type 1 EMT associated with gastrulation, palate fusion, and neurulation. Even though fibronectin is an integral constituent of the fibrotic ECM associated with fibrosis and the desmoplastic stroma in tumours, the utility of fibronectin as a type 2 and type 3 EMT biomarker is limited, in part, because it is produced by various cell types, including fibroblasts, mononuclear cells, and epithelial cells [408, 409]. Both type 2 and type 3 EMT, however, are associated with increased fibronectin expression *in vitro* [410].

Type IV collagen, laminin, nidogen, and sulphated proteoglycans are the principal basement membrane constituents that during EMT are downregulated.

Among these laminin is the better established biomarker of the EMT. Laminins are heterotrimeric glycoproteins composed of one α chain, one β chain, and one γ chain [411]. Currently, 15 different heterotrimers are known [411]. With regard to EMT, most studies focused on laminin1 (α1β1γ1), which is present in the embryo from the peri-implantation period forward [412]. Type 1 EMT during gastrulation and palate fusion is marked by loss of laminin 1 [413]. Both type 1 and type 2 EMT are associated with downregulation of laminin 1 *in vitro* and disruption and loss of laminin 1 *in vivo* [414]. By contrast, up-regulation of laminin 5 (α3β3γ2) is associated with type 3 EMT in cancer and type 2 EMT. Laminin 5 is detected in the discontinuous laminin patterns associated with invasive cancers, and its expression is linked to type 3 EMT in breast carcinomas of the ductal type [415], hepatocellular carcinoma [416], and oral squamous carcinoma [417].

Similarly, laminin 5 is part of the fibrotic ECM associated with idiopathic pulmonary fibrosis [418]. A correlation between laminin 5 and EMT in the embryo has not yet been reported.

Transcription factors. Fibroblast transcription site-1 (FTS-1) is a *cis*-acting regulatory element present in the promoter region of various EMT-associated genes, including those that encode FSP1, Twist, Snail1, high mobility group AT-hook 2 (HMGA2), LEF1, Ets-1, E-cadherin, β-catenin, ZO-1, α-SMA, and vimentin.

CBF-A and KRAB-associated protein 1 (KAP-1) form a complex with FTS-1 to modulate gene activity. Formation of the CBF-A/KAP-1/FTS-1 complex induces

type 2 EMT in kidney tubular epithelial cells. *De novo* expression of CBF-A, a member of the heterogeneous nuclear ribonucleoprotein A/B (hnRNP A/B) family, can be observed in type 2 and type 3 EMT associated with onset of fibrosis and metastatic tumour formation. However, its expression is not confined to EMT, and it is involved in various other processes, including stimulation of Ig transcription [419]. KAP-1 is also a ubiquitously expressed transcriptional co-repressor or activator that is well known to bind KRAB domains on zinc finger proteins [419]. In EMT, under the limited circumstances where it has been studied, KAP-1 seems to be an activator and, in the CBF-A/KAP-1/FTS-1 complex, a proximal master activator of EMT-associated genes.

An increasing number of environmental factors are now known to induce EMT, including growth factors, ECM constituents, proteases, and hypoxia. Despite the different nature of known stimulants, the EMT response is relatively uniform. This raises interest in potential key regulators of EMT that are at the intersection of various signalling pathways and control the EMT response. Snail transcription factors are one prominent example of a common downstream target of various signalling pathways that regulates EMT [428]. Of the three vertebrate Snail family members of zinc finger proteins (Snail1, Snail2, and Snail3), the functionally equivalent Snail1 and Snail2 (which was formerly known as Slug) mediate EMT [428]. In fact, all known EMT events during development, cancer, and fibrosis appear to be associated with Snail activation [428].

Snail is most widely recognized as a suppressor of E-cadherin expression, although it regulates various other aspects of the EMT phenotype, including the increased expression of mesenchymal cell/fibroblast markers *i.e.* fibronectin and vitronectin, decreased expression of epithelial markers *i.e.* claudins, occludins, and cytokeratins, inhibition of proliferation through suppression of cyclin D proteins and cyclin-dependent kinase 4 (CDK4), increased MMP expression, and protection from cell death (through suppression of expression of caspases, DNA fragmentation factor, and Bcl-interacting death agonist). Snail transcriptional activity is regulated by control of its subcellular localization. Phosphorylation of Snail causes its export from the nucleus into the cytoplasm, resulting in its inactivation as a transcription factor. In human renal biopsies, fibrosis is associated with increased Snail1 expression [420], and Snail activation in tamoxifen-inducible *Snail1*-transgenic mice results in type 2 EMT and renal fibrosis.

Twist is a basic helix-loop-helix protein that is transcriptionally active during lineage determination and cell differentiation. It is up-regulated during early embryonic morphogenesis [421,422], fibrosis [423], and cancer metastasis [424]. In the development of metastatic cancer cells by type 3 EMT, Twist can act independently of Snail to repress E-cadherin [6] and to up-regulate fibronectin and N-cadherin [410].

Another transcription factor is called Forkhead box C2 (FOXC2). It acts as a pleiotropic inducer of EMT. During embryonic development, FOXC2 is widely expressed and is required for angiogenesis, musculogenesis, and organogenesis of the kidney, heart, and urinary tract [425]. Postnatally, expression of FOXC2 is normally restricted to adipocytes [426]. However, FOXC2 is also expressed in ductal breast cancers and metastatic breast cancer cell lines [427]. Over-expression of any one of the EMT inducers TGF-β1, Snail, Twist, or Goosecoid increases FOXC2 expression, and over-expression of FOXC2 itself induces EMT,

suggesting an important role for FOXC2 in type 3 EMT [427]. A role for FOXC2 in EMT associated with embryogenesis or fibrosis has not been established yet.

Table Ib *EMT Biomarkers*

Name	Acquired EMT type	Name	Attenuated EMT type
Cell-surface proteins			
N-cadherin	1, 2	E-cadherin	1, 2, 3
OB-cadherin	3	ZO-1	1, 2, 3
α5β1 integrin	1, 3		
αVβ6 integrin	1, 3		
Syndecan-1	1, 3		
Cytoskeletal markers			
FSP1	1, 2, 3	Cytokeratin	1, 2, 3
α-SMA	2, 3		
Vimentin	1, 2		
β-Catenin	1, 2, 3		
ECM proteins			
α1(I) collagen	1, 3	α1(IV) collagen	1, 2, 3
α1(III) collagen	1, 3	Laminin 1	1, 2, 3
Fibronectin	1, 2		
Laminin 5	1, 2		
Transcription factors			
Snail1 (Snail)	1, 2, 3		
Snail2 (Slug)	1, 2, 3		
ZEB1	1, 2, 3		
CBF-A/KAP-1 complex	2, 3		
Twist	1, 2, 3		
LEF-1	1, 2, 3		
Ets-1	1, 2, 3		
FOXC2	1, 2		
Goosecoid	1, 2		
MicroRNAs			
miR10b	2	Mir-200 family	2
miR-21	2, 3		

1.3 TWIST 1: A HOMOLOG OF DROSOPHILA TWIST

TWIST1 belongs to the basic helix-loop-helix (bHLH) class of transcriptional regulators that recognize a consensus DNA element called the E box.

1.3.1 Cloning

By PCR of a placenta cDNA library using primers based on mouse Twist, Bourgeois *et al.* in 1996 [429] first cloned human TWIST. The deduced 206-amino acid protein contains a central DNA-binding basic region followed by a helix-loop-helix domain. Mouse and human TWIST share 96.6% amino acid identity. *In vitro*-translated TWIST had an apparent molecular mass of 25 kD.
The protein contains 201 amino acids and has a calculated molecular mass of 20.9 kD. It has a hydrophilic N terminus, followed by a bHLH DNA-binding and dimerization motif. It also contains several potential phosphorylation sites, including two in the loop region of the dimerization domain that are conserved among all species examined. TWIST orthologs were detected in all mammalian species examined, and within mammals, the DNA-binding region showed 100% sequence conservation. A possible ortholog was also detected in chicken, but not in yeast. Northern blot analysis detected a 1.6-kb transcript that was highly expressed in placenta. Lower expression was detected in adult heart and skeletal muscle, and weak expression was found in kidney and pancreas, but not in brain. Twist 1 expression was also detected in endometrial fibroblasts, peritoneal mesothelial cells, and fetal lung fibroblasts, but not in other cell lines examined.

1.3.2 Gene Function

Studies in Drosophila by Shishido *et al* (1993) [435] indicated that Twist may affect the transcription of fibroblast growth factor receptors (FGFRs), a gene family implicated in craniosynostosis. The emerging cascade of molecular components involved in craniofacial and limb development included TWIST, which may function as an upstream regulator of FGFRs.
Histone acetyltransferases (HATs) play a critical role in transcriptional control by relieving repressive effects of chromatin. Hamamori *et al* (1999) showed that Twist directly binds two independent HAT domains of acetyltransferases, p300 and p300/CBP-associated factor (PCAF), and directly regulates their HAT activities [436]. The N terminus of Twist is a primary domain interacting with both acetyltransferases, and the same domain is required for inhibition of p300-dependent transcription by Twist. Adenovirus E1A protein mimicked the effects of Twist by inhibiting the HAT activities of p300 and PCAF.
Using electrophoretic mobility shift assays, El Ghouzzi *et al* (2001) demonstrated that the TWIST-E12 (TCF3; 147141) dimer specifically recognizes the CATATG motif [431].
Using a murine breast tumour model, Yang *et al* (2004) determined that Twist plays an essential role in metastasis [6]. Suppression of Twist expression in highly metastatic mammary carcinoma cells specifically inhibited their ability to metastasize from the mammary gland to the lung. Ectopic expression of Twist resulted in loss of E-cadherin mediated cell-cell adhesion, activation of mesenchymal markers, and induction of cell motility, suggesting that Twist contributes to metastasis by promoting an epithelial-mesenchymal transition. In human breast cancers, high Twist expression correlated with invasive lobular

carcinoma, a highly infiltrating tumour type associated with loss of E-cadherin expression.

1.3.4 Mapping

Bourgeois *et al* (1996)[429] used isotopic in situ hybridization to map the TWIST1 gene to chromosome 7p21. The murine gene had been mapped to bands B-C1 of chromosome 12 by Mattei *et al* (1993)[433].

1.3.5 Molecular Genetics

Twist is required in head mesenchyme for cranial neural tube morphogenesis in mice. While homozygous Twist-null murine embryos exhibit failure of neural tube closure, heterozygosity for Twist-null mutations results in a moderate phenotype including minor skull and limb anomalies consistent with those of the Saethre-Chotzen syndrome (SCS). Furthermore, the clinical phenotype of SCS was mapped to 7p22-p21 by linkage analysis; Bourgeois *et al* (1996) and Howard *et al* (1997) mapped the human TWIST gene to the same region of 7p. This prompted Howard *et al* (1997) and El Ghouzzi *et al* (1997) to seek mutations in the TWIST gene in SCS[429,432]. Both groups found a number of mutations occurring within the basic DNA binding, helix I, and loop domains resulting in substitutions or premature termination of the protein.

1.3.6 Animal Model

Heterozygous-null Twist mice have variable expression of craniofacial and limb defects consistent with the phenotype and variability seen in SCS (Bourgeois *et al*, 1998)[437].

Mice heterozygous for the ethylmethanesulfonate (EMS)-induced polydactyly ems (pde) mutation show preaxial polydactyly of the hindlimbs. Browning *et al* (2001) determined that the pde mutation maps to chromosome 12 and is an allele of Twist [430]. However, sequencing the Twist protein-coding region and several hundred basepairs upstream of the coding region failed to reveal a disease-associated mutation. Browning *et al* (2001) concluded that the lesion may be in a regulatory element of the gene.

Pan *et al* (2009) found that homozygous Twist1 knockout in mice was embryonic lethal [434].

2. SCIENTIFIC ISSUES and AIMS

The main reasons for poor prognosis in CRC are delayed diagnosis, at a metastatic stage (International Union Against Cancer, stage IV) and disease spreading to regional lymph-nodes (stage III). Of locally advanced disease (stage II, T3 and T4,) a considerable fraction (up to 30%) can eventually undergo metastatic evolution after surgery. Very few molecular markers have been proved to correlate with colon cancer prognosis. Among these, the presence of microsatellite instability (MSI) and thus of mismatch repair defects (10%-15% of all colon cancers) is the most frequently cited and has even been proposed for clinical purposes. The main determinant of the prognostic advantage of MSI cancers is their lower metastatic potential, as compared to MS-stable (MSS) ones. Yet, the molecular understanding of the favourable prognostic implications of MSI remains largely a matter of speculation.

As to functional models explaining for the metastatic behaviour of cancer, in epithelial malignancies in which the MSI phenotype does not occur, it has recently been suggested that epithelial to mesenchymal transition (EMT) is one of the determinants of cancer-cell invasion and metastasis. Although EMT cannot be thought to be the only determinant of metastatic spread, nor a mandatory signature of cancers, it provides a model to approach mechanistic studies on metastasis development. In a breast cancer, syngenic animal model (BALB/c), the mesenchymal development driving gene *Twist1* has been shown to correlate with the invasive and metastic properties of an orthotopic mice breast cancer tumor according to an EMT model. In human breast cancer, *Twist1* expression correlates with the degree of chromosomal abnormalities (chromosomal instability), invasiveness and angiogenesis. Twist1 expression has also been reported in other invasive/metastatic human cancers. Among the gastrointestinal cancers, Twist1 expression seems to occur in gastric and liver tumors, but data on CRC are scarce. Colon cancer provides the chance to test Twist1 expression in relation to the presence/absence of the molecular phenotype of MSI, and in the light of the classic genetic model of cancer progression through the accumulation of genetic damage. Relying upon the breast cancer published data we propose to investigate Twist1 expression in CRC according their MS-status. Accordingly, we plan not only to verify whether *Twist1* constitutive expression will ultimately result in the transition to mesenchymal features, coupled to invasive properties in CRC cells, but also to test whether specific*Twist1* inducers can induce EMT in CRC cell-lines, with and without MSI. Moreover we aim to explore EMT driven by *Twist1* in a mouse model of CRC.

3. MATERIALS AND METHODS

3.1 Cell culture

Human Colon Cancer cell lines: CoLo115, GP5D, HCT116, SW48, DLD1, LoVo, CoLo741, HT29, CoLo205, SW403, CaCo2, CX-1, LS174T, SW480, SW620, SW837, LS180, LS411, CT26 were purchased from American Type Colture Collection were cultured in RPMI with 10% fetal bovine serum (FBS),100mg/ml of penicyllin G, 50mg/ml of Streptomicin and Ultraglutamine 2mM and growth at 37℃, 5% CO_2 incubator.

3.2 Gene expression profile microarray

Microarray analysis
The transcriptional profile of the CoLo115, GP5D, HCT116, SW48, DLD1, LoVo, CoLo741, HT29, CoLo205, SW403, CaCo2, CX-1, LS174T, SW480, SW620, SW837, LS180, LS411 colorectal cancer cell lines: as affymetrix raw data files [cell intensity (Cell) files] was kindly given by Giancarlo Marra (University of Zurich). Each preparation derived from a different single donor using the Human Genome U133 Plus 2.0 Array (Affymetrix, Santa Clara, CA, USA), containing a total of ~55,000 transcripts. Cell files were processed using a Robust Multi-Array Analysis (RMA) procedure, an algorithm that is freely available at http://www.bioconductor.org. The RMA method was used to convert the intensities from the multiple probes of a probe set in a single expression value with greater precision and reduced background noise (relying on perfect the perfect match probes only and thus ignoring the mismatch probes) and then to normalize by sketch quintile normalization. Significance analysis of Microarray (SAM), Principal Component Analysis (PCA) of variance and Hierarchical Clustering (HCL), after mean scaling and $\log_2$ transformation, were performed using a software tool of The Institute for Genomic Research (TIGR) MeV: http://www.tigr.org/software/th4/mev.html.
For the analysis of pathways was used Ingenuity System Pathways Analysis software (IPA).

3.3 RNA extraction

3.3.1 Mammalian Cell lines

Cells were detached with trypsin 1%, washed with physiologic water, then centrifuged at 1550 rpm for 10 min, disrupted in 600 ml of Buffer RLT and finally homogenized. The same volume of Ethanol is then added to the lysate, creating conditions that promote selective binding of RNA to the RNeasy membrane. The sample is then applied to the RNeasy Mini spin column. Total RNA binds to the membrane, contaminants are efficiently washed away, and RNA is eluted in 50 ml of RNase-free water. All bind, wash, and elution steps are performed by centrifugation in a microcentrifuge.

3.3.2 Tissues

78 CRC tissues and their normal mucosa counterpart were used to study Twist1 mRNA expression. RNA from was purified using Trizol reagent (Applied Biosystem).
100 mg of tissue samples were homogenized in 1 ml of Trizol reagent using Tissue ruptor (Qiagen).
Samples were incubated for 5 min at 15 to 30℃ to permit the complete dissociation of nucleoprotein complexes. 0.2 ml of chloroform was added per 1 ml of TRIZOL® and tubes were shacked vigorously by hand for 15 sec and incubated at 15 to 30℃ for 2 to 3 min. Then samples were centrifuged at 12,000*g* for 15 min at 4℃. Following centrifugation, the aqueous phase was transferred to a fresh tube. RNA was precipitated from the aqueous phase by mixing with 600 ml of isopropyl alcohol and tubes incubated at room temperature for 10 min and centrifuged at 12,000*g* for 30 min.
The supernatant removed and the RNA pellet washed two times with1 ml of 75% ethanol. After a centrifugation at 12,000 x g the RNA pellet was suspended with 100 ml of nuclease free water.

3.4 DNase treatment

To remove genomic DNA contamination, all samples were treated with TURBO DNA free kit (Applera).
Each sample was incubated for 1h at 37 C after the addition of 0.1 Volume of 10x Turbo DNase free buffer and 2 ml of DNase enzyme.
Dnase was then neutralized using *DNase Inactivation Reagent.*

3.5 Reverse-Trascription Polymerase Chain Reaction (RT-PCR)

Total RNA derived from cell lines or tissue specimens was reverse-transcripted to single-stranded cDNA using High Capacity cDNA Reverse Transcription Kit (Applied Biosystems). Reaction mixture was prepared addicting to 2 mg of total RNA 5ml of 10x RT Buffer, 5ml of 10x RT Random Primers, 2 of dNTP Mix (100mM), 2,5 ml of MultiScribe™ Reverse Transcriptase and 1ml RNase Inhibitor.
Thermal cycler () was programmed following these optimized conditions:
Step 1: 25℃ for 10 min
Step 2: 37℃ for 120 min
Step 3: 85℃ for 5 min

3.6 Quantitative Real-time Polymerase Chain Reaction (qRT-PCR)

Quantitative Real-Time PCR was performed using the HT 7900 Fast system Instrument (Applied Biosystem, California, USA) and SYBR Green (Applied Biosystem, California, USA) as intercalating fluorescent dye.
All primer sequences were designed with Primer 3 program (Rozen and Skaletsky). One µl of sample cDNA (40 ng) was added to the PCR reaction mixture containing 2X Power SYBR Green Master Mix and forward and reverse primers (5 µM) in a final volume of 25 µl. All analyses were performed in triplicate. The PCR amplification profile has been employed as follows: initial denaturation at 95℃ for 10min, followed by 40 cycles of denaturation at 95°C for 15 s, annealing for 40 s at

60℃ and extension at 72℃ for 20 s. Detection of t he fluorescent product was carried out during the annealing period. Data analysis was performed using the comparative Delta Delta Ct (ΔΔCt) method (Livak and Schmittgen) The relative gene copy number was calculated using the expression 2-ΔΔCt, in order to establish the range of values for each group. The standard deviations (SD) from the mean values were also calculated.

3.7 Cell lines treatment

3.7.1 Cytokines & Hypoxia

To induce Twist1 expression the human colon cancer Cell lines HT29, SW480 and Colo741 were plated into six well-plates at the density of 70-80% of confluence and the subsequent day treated with TGF-β, IL-6, TNF-α, IFN-γ and IL-1β at a final concentration of 20 ng/ml for 6 h. Another plate was also incubated in hypoxia condition for 6 h using an incubator with controlled atmosphere 1% O_2, 5% CO_2 and 94% N.

3.7.2 2'-deoxy-5-azacytidine (DAC)

Colorectal cell lines HCT116 and CoLo741 were treated with 4μM of the de-methylator agent 2'-deoxy-5-azacytidine (DAC) for 4 days. During the treatment every 24 h the medium was changed with fresh one added with 4μM of DAC, this procedure was necessary due to the toxicity of DAC.

3.7.3 Emetine and Actionomycin D

HCT116 and CoLo741 cell lines once reached the 70% of confluence in a six-plate dish were treated with Emetine and Actinomicycin D at final concentration of 100 μg/ml and 2 μg/ml from 1 to 3 h.

3.8 Twist1 sequencing

The coding region of *TWIST* gene was amplified and sequenced (see **Table IIIb** for primer sequences).
PCRs were performed in 40μl volumes that contained 200 ng genomic DNA, 4 μl High Fidelity PCR buffer, 0.2 mM each dATP, dCTP and dTTP, 0.15 mM 7-deaza-dGTP, 0.05 mM dGTP, 0.25 μM each primer, 2mM MgSO4, and 1U Platinum *Taq* High Fidelity (Invitrogen). PCR parameters were: 94℃ for 6 min; 38 cycles of 94℃ for 1 min, 66℃ for 15 sec, and 68℃ for 45 sec; fo llowed by 68℃ for 7 min.
PCR products were purified with Qiagen PCR purification kit (Qiagen, Milan, Italy). To perform the cycle sequencing reaction, 2 μl of purified DNA fragment was blended with each primer (0.8 mmol/l) in a Terminator Ready Reaction Mix containing Big Dye Terminators (Applied Biosystems, Foster City, California, USA), denatured for 5 min at 95℃ and submitted to 30 cyc les at 95℃ for 30 s, 50℃ for 10 s, and 60℃ for 4 min. A second purification wit h DyeEx 2.0 Spin Kit (Qiagen) was performed for Big Dye removal. Five μl of marked and purified DNA were submitted to sequencing analysis on an ABI PRISM 310 Genetic Analyser (Applied Biosystems).

3.9 Twist1 transfection in SW480 cell line

Twist1 stable transfection in SW480 cell line was performed using a full length cDNA cloned in pCMV6-AC vector (Origene) (see Figure 2.1 for plasmid scheme). Before transfection in mammalian cells, the plasmidic DNA was propagated in Top10 E. Coli chemically competent cells.

3.9.1 DNA Plasmid amplification in Top10 E.Coli

10 µL of chemically competent cells (Top10, E. Coli) was thawed on ice and then 5 µL (5ng) of expression plasmid was added and incubated on ice for 30 min.
The heat shock procedure was performed by incubating the mixture of DNA and cells at 42^{o}C for 30 sec, and then they were removed to ice immediately.
After heat-shock, 250 uL of recovery medium (SOC) was added to the cells, and incubated at 37^{o}C for 1 h with agitation. Two different dilutions of the transformation mixture were plated in separate LB-amp agar plates (10% and 50% of the transformation reaction on separate plates, each diluted up to 100 uL in SOC). The plates were incubated overnight at 37^{o} C. The following day, 5 single bacterial colonies from each transformation were inoculated into individual sterile culture tubes containing 50 mL of liquid medium with 100 ug/mL ampicillin (LB-amp) and incubated overnight at 37^{o} C with agitation.

3.9.3 Plasmid purification

Cells were pelleted using centrifugation at 5,000 × *g* for 10 min and the supernatant discarded.
Pellets were resuspended in 3 mL of cell re-suspension solution and lysed in 3 mL of cell lysis solution, after 3 min of incubation at room temperature was added 3 ml of Neutralization Solution to the lysed cells, the tubes capped and mixed by gently inverting 10 times.
The lysate were centrifuged at room temperature, 15,000 × *g* for 15 min. This centrifugation pelleted the bulk of the cellular debris. Remaining debris was removed using the PureYield™ Clearing Columns.
The cleared lysate was decanted into the PureYield™ Clearing Column and applying vacuum the lysate passed through the clearing membrane in the PureYield™ Clearing Column, and the DNA remained bound to the binding membrane in the PureYield™ Binding Column.
Columns were washed with 5 mL of Endotoxin Removal Wash and 20ml of Column Wash Solution. Membranes were dried for 30 sec by applying a vacuum. After drying, the columns were placed into a new 50 ml disposable plastic tube tip of the column on a paper towel to remove excess ethanol. Place the column into a new 50ml disposable plastic tube (e.g., Corning or Falcon™).
Elution was done by vacuum using the Eluator™ Vacuum Elution Device and 600µl of Nuclease-Free Water applying maximum vacuum for 1 min. DNA was then quantified using the nanodrop.

3.9.3 Plasmid sequencing

DNA sequencing from the 5` end of the cDNA insert was performed with VP1.5 (5`-GGACTTTCCAAAATGTCG-3`) whose priming site is located ~120 bp upstream of the polylinker. DNA sequencing from the 3` end of the cDNA insert was performed

with XL39 (5`-ATTAGGACAAGGCTGGTGGG-3`) whose priming site is located ~70 bp downstream of the polylinker.

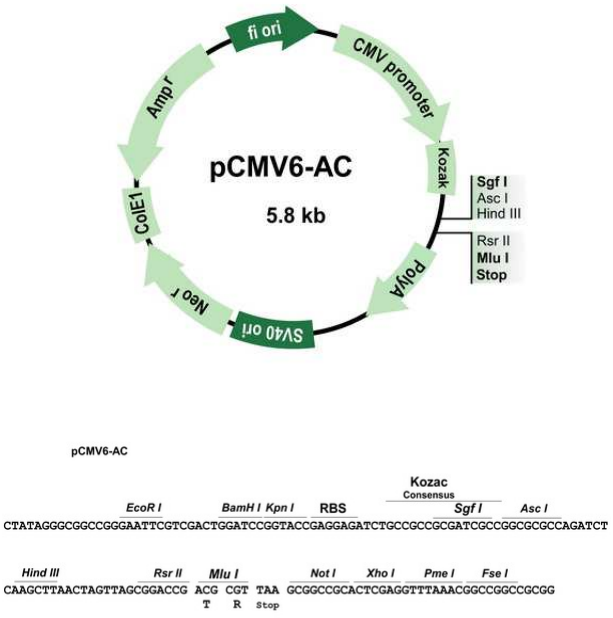


Figure 3.1 *Schematic of pCMV6-AC plasmid and its multiple cloning sites*

3.9.4 Mammalian cells transfection

The day before transfection, $1x10^6$ SW480 cells were plated into each vessel of a 6 well plate. Cells were grown overnight at 37oC in a 5% CO2 incubator. The following day was prepared the transfection mixture as follow: 4 ng of plasmidic DNA was diluted into 250 ml serum-free medium OptiMEM without antibiotics same was done with 4 ml of the transfection reagent Lipofectamine 2000 (invitrogen) . After 5 minutes of incubation the diluted DNA and Lipofectamine were combined and incubated for 20 minutes at room temperature .Then 500 ml of the complexes were added to each well containing cells and medium. Cells were incubated at 37℃ 5% CO2 incubator for 48h then pas sed 1:10 and for 2 weeks treated with selective medium containing 0.5mg/ml of Neomicyn.

3.10 Twist1 silencing

Twist1 knockdown in Colo741 cell line was performed using a lentiviral based ShRNA vector. 5 different shRNA were tested.
The ShRNA vectors were purchased as twist homolog 1 (acrocephalosyndactyly 3; Saethre-Chotzen syndrome) (Drosophila) MISSION® shRNA Bacterial Glycerol Stock from Sigma.
The ShRNA was provided as frozen bacterial glycerol stocks (Luria Broth, carbenicillin at 100 µg/ml and 10% glycerol) in Escherichia coli for propagation and downstream purification of the shRNA clones (pLKO.1) which contains the puromycin selection marker for stable transfection.

Five different shRNA were tested, below the sequences:

shRNA 1 Region: 3UTR
Sequence: CCGGGCATTCTGATAGAAGTCTGAACTCGAGTTCAGACTTCTATCAGAATGCTTTTT
shRNA 2 Region: codifing sequence
Sequence:: CCGGCCTGAGCAACAGCGAGGAAGACTCGAGTCTTCCTCGCTGTTGCTCAGGTTTTT
shRNA 3 Region: codifing sequen
Sequence: CCGGCGCCTTCTCGGTCTGGAGGATCTCGAGATCCTCCAGACCGAGAAGGCGTTTTT
shRNA 4 Region: codifing sequen
Sequence: CCGGTCCGCAGTCTTACGAGGAGCTCTCGAGAGCTCCTCGTAAGACTGCGGATTTTT
shRNA 5 Regione: codifing sequen
Sequence: CCGGGCTGGACTCCAAGATGGCAAGCTCGAGCTTGCCATCTTGGAGTCCAGCTTTTT

Bacteria were first amplified (as described before) and to reach a good quantity of ShRNA, plasmidic DNA was purified using MIDI prep protocol (as described before).

3.10.1 Lentiviral supernatant production

To produce lentiviral surnatant HEK293T cells were used as packaging cells. HEK293T cells originally derived from a continuous human embryonic kidney cell line. 293T were seeded the first day at 40% of confluence in complete medium (RPMI), when they reached the 70% of confluence were transfected with 3 lentiviral vectors (packaging plasmid, envelope plasmid, transfer vector) as follow:
MIX 1:

- VSV G-expressing plasmid: 5 ug,
- pCMVdeltaR8.74: 5 ug
- Transfer vector: 15 ug
- Optimem: 1.5 ml

MIX 2:

- Lipofectamine: 60 ul
- Optimem 1.5 ml

MIX 2 was incubated for 5 minutes and then was added to MIX 1 for 20 minutes at RT and added to 293T.
After 18 hours, transfection medium was removed and added complete medium, for viral harvest.
Cells were incubated for 24 hours then the viral supernatant recovered and harvested and meium was replaced with fresh complete medium and the procedure repeated after 24 hours. The viral harvests were then pooled and spinned at 4000 rpm 2 minutes to pellet packaging cells that were collected during harvesting. The supernatant was stored to a sterile polypropylene tube, till the infection of Colo741.
Packaging plasmid (pCMV-R8.74): second generation vector, it contains *gag*, *pol* and *rev* gene.
Envelope plasmid: contains *VSV-G* gene (it's a vector pseudotyped with the vesicular stomatitis virus glycoprotein, to allow entry in a broad spectrum of cell types).

3.10.2 Infection

The infection was checked after 48 and72 hours with qRT-PCR and western blotting. After infection, cells should be kept in P3 room for at least 2 passages and supernatant produced should be checked on independent cells for absence of viral activity.
All the procedures were be carried out in accordance with biosafety requirements of the host institution.

3.11 Immunofluorescence

Glass slides coated with poly-L-lisine were used as bulk to grow the cells.
Cells were plated at density of 4x104 in 500 ml of culture medium. The following day medium was aspirated and cells were washed with physiologic and fixed with PFA 4% in PBS for 15 min at room temperature. The glasses were washed two times with PBS BSA 2% and incubated in PBS and BSA 2% for 30 min to block the aspecific sites.
After blocking, cells were permeabilized with triton 0.3% in PBS, and than saturated with PBS BSA 2%, Glycin 0.1% and normal goat Serum 5% for 1h at room temperature.
Staining was than performed by treating the cells with the primary antibody (see **Table IIIa** for Antibodies dilutions).
The following day after 2 washes with PBS, glasses were incubated with the secondary antibody Alexa Fluor® 647 or 488 goat anti-mouse or anti-rabbit IgG (H+L) 2 mg⁄mL (Invitrogen) (dilution 1:2000) for 1h at room temperature. Nuclei were couterstaining with DAPI (Invitrogen). Glasses were than mounted on slides using Fluor Save Iinvitrogen.
All the images were acquired using a confocal microscopy (Olympus) and a CellR microscopy (Olympus)

3.12 Western blotting

Colorectal cancer cell lines were lysed in chilled lysis buffer supplemented with a complete protease inhibitor cocktail (Roche). Whole cell protein extracts were used for protein determination using a Coomassie Bradford assay kit (BioRad). Equal amounts of proteins (30µg) for each sample were loaded onto 10% polyacrilamide gels. Proteins were then separated by SDS gel electrophoresis and transferred onto nitrocellulose membrane (BioRad). The membranes were incubated 1 h with PBS 5% skimmed milk and then were incubated with primary antibodies diluted in PBS 3% skimmed milk.
Membranes were then incubated 1 h with anti-mouse or anti-rabbit secondary antibody conjugated to horseradish peroxidase (Amersham) diluted with PBS 3% skimmed milk. Antibody binding was visualized by enhanced chemiluminescence ECL (Amersham) immunoblotting detection system and chemiluminescence was detected by autoradiography.

3.13 Wound Healing assay

Cells were seeded in multi-well plates and cultured until reach the confluence. Using a pipette tip it was made a straight scratch, simulating a wound. After 18 h, 32 h and 44 h the wound enclosure was photographed in phase contrast using Cell

R Olympus microscope. The enclosure dimension was measured using Cell R software and expressed as percentage of enclosure. The experiment was done in triplicate, and for each wound 3 images were captured.

3.14 Invasion assay

To assess the invasive properties of cells *in vitro* BioCoat GFR Matrigel Invasion Chambers were used. The BioCoat GFR Matrigel Invasion Chamber consists of Falcon Cell Culture Inserts with 8 micron pore size PET membrane. Each membrane has a thin layer of GFR Matrigel Basement Membrane Matrix that serves to reconstitute the basement membrane *in vitro*. The layer occludes the pores of the membrane, blocking non-invasive cells from migrating through the membrane. However, cells that migrate are able to detach from the insert and invade through the GFR Matrigel matrix and the 8 micron membrane pores.
To rehydrate the membrane 0.5 ml of warm (37°C) culture medium were added to the interior of the inserts.
After rehydration were prepared cell suspensions in culture medium containing 0.1% FBS at 5 x 10^4 cells/ml and 0.5 ml were added to the inside of the inserts while in the lower chamber were added 0.5 ml of culture medium containing 10% FBS.
All incubated for 24 h in a humidified tissue culture incubator at 37°C, 5% CO_2 atmosphere.
After incubation, the non-invading cells were removed from the upper surface of the membrane by "scrubbing" with cotton tipped swab. The cells on the lower surface of the membrane were stained with Diff-Quick stain. The Diff-Quick kit contains a fixative and two stain solutions. The cell nucleus stained purple and the cytoplasm stained pink. Cells were counted by photographing the central fields (triplicate) of the membrane. Data were expressed as mean of the number of cells that invaded the GFR Matrigel™ Matrix.

3.15 Patient and Samples

3.15.1 Study population and retrieval of clinico-pathological data

The study included 201 unselected Caucasian patients undergwent to resective surgery for colorectal cancer at our Institution between 1997 and 2006. Patients were excluded from the study if (*a*) pathologic examination did not confirm a tumour invading at least the sub-mucosa (pT_1 or higher); (*b*) the tumour was a local recurrence of previously resected colorectal cancer. The study was approved by the Ethical Committee of the Istituto Clinico Humanitas, and the informed consent of the patient to the treatment of their personal data was obtained by the referring physician or by other clinicians involved in the study.
A clinical database was prepared by investigators blind to the results of cancer genetic testing. Pathologic tissue specimens were reviewed by a single pathologist who was also unaware of molecular data. Tumour pathologic staging, histopathological typing, tumour grade, and presence or absence of extramural vein invasion were assessed. Tumour clinico-pathologic staging was finally assessed by combining histopathologic findings, surgical records (including intraoperative liver ultrasonography), and perioperative imaging (abdominal computerized tomography and chest X-rays in all patients). Demographics and

complete clinical data at diagnosis were made available at hospital Intranet resources.

3.15.2 Assesment of MSI

Coded sections of paraffin-embedded colorectal cancer tissue were sent from the pathologist to the Research Laboratory. MSI assignment was based on the analysis of repeats in mononucleotides. Using dinucleotide markers *BAT25-BAT26*. After DNA extraction by proteinase-K digestion and phenol-chloroform purification, amplification of the *BAT25* and *BAT26* loci with fluoresceinated primers (see **Table IIIb** for primer sequences) was followed by capillary-gel electrophoresis (ABI PRISM 310 DNA Sequencer, Perkin-Elmer).

3.16 Immunohistochemistry

201 CRC specimens of 2 µm thick sections were cut and processed for immunohistochemistry for Twist1. After deparaffining and rehydration, the sections were immersed in an antigen retrieval bath, incubated with 3% H_2O_2 for 15 min to quench endogenous peroxidase activity, and then treated for two h at room temperature with primary antibodies raised against TWIST1 or with mouse IgG (Dako, Milan, Italy) as a negative control, followed by 30 min' incubation with the DAKO Envision system (Dako). 3,3-diaminobenzidine tetrahydrochloride (Dako) was used as a chromogen to yield brown reaction products. The nuclei were lightly counterstained with hematoxylin solution (DAKO). Twist1 staining extent was rated according to the percentage of positive cells. Samples with no stained cells were rated as 0, those with <10% stained cells were rated as 1, those with 10-50% of stained cells were rated as 2, those with > 50% were rated as 3.
In order to verify whether TWIST1 was co-expressed with epithelial, mesenchymal markers and CD antigens (see **Table IIIa**) 2 µm thick sections were processed for double-immunohistochemical assay. After deparaffining and rehydration, the sections were incubated with 3% H_2O_2 for 15 min to quench endogenous peroxidase activity and then treated for two h at room temperature with primary antibodies raised against TWIST1, followed by 30 min' incubation with the DAKO Envision system (Dako). 3,3-diaminobenzidine tetrahydrochloride (Dako) was used as a chromogen to yield brown reaction products. The slides were then incubated for 60 min with primary monoclonal/policlonal antibodies (see **Table IIIa** for Antibody used and dilutions), followed by 30 min' incubation with secondary antibody (Envision, Dako, and Mach4 Biocare) with BAJORAN PURPLE (BioCare Medical) or AEC (Dako) used as a chromogens to yield a purple or red reaction products. The nuclei were lightly counterstained with hematoxylin solution (DAKO).

3.17 iFish-IHC

To perform FISH on paraffin embedded tissues, was used the Zymed Hotspot-light kit (Invitrogen).
Paraffin embedded (FFPE) sections of 4-5 µm were stained as previous described for Twist1 (see immunohystochemistry paragraph), without perform Hematoxilin staining. After wash with BioClear solution, slides were immersed in 100% Ethanol solution three times for three min and then air dryed. Slides were placed on slides rack and boiled at 98°C for 20 min in Heat pre-trea tment Solution (Reagent A,Invitrogen). After the heat pre-treatment slides were placed immediantly in dH2O

at RT (15-30℃) and washed three times. Tissues were then subjected to Enzymatic digestion (Reagent B, Invitrogent) for 14 min at 37℃. After digestion slides were washed three times in dH2O for 2 min and dehydratated in graded ethanol series: 70% EtOH 2 min, 85% EtOH 2 min., 95 % EtOH 2 min., 100% EtOH 2 min., 100% EtOH 2 min. After Air-drying 6 μl of CEP7 spectrum green probe (Vysis) was added to the center of the tissue section, than the Coverslip was placed and sealed using rubber cement. Probe denaturation for 5 min at 85℃ and incubation for 20 h at 37℃ was performed using Hibridizer (Dako). The following day were prepared two Coplin jars containing SSC buffer and 0.07% of NP-40 where the slides were placed and washed for 5 min at 73℃ two times (stringent was). After stringent wash slides were placed in dH2O and than 100% Ethanol. After air-drying , slides were mounted using 5 ml of DAPI (Vectastain) and preserved at -20℃. FISH stainig were observed at Fish station microscope (Olympus). Images were acquired using Cytovision software (Olympus).

3.18 Animal model

3.18.1 Retroviral GFP transfection of CT26

Retroviral transfection of CT26 was performed as mentioned before .. using as lentiviral transfection vector pRRL-sin-PPT-hCMV-GFP-WPRE.

3.18.2 Mice, Transanal injection, necropsy and tissue preparation

Male Balb/c mouse were purchased from Charles River. The mice were housed and maintained in specific pathogen-free conditions.
The mice were used in accordance with institutional guidelines when they were 8-12 weeks old.
Murine colon cancer CT26-GFP was harvested from near-confluent culture by a brief exposure to 0.5% trypsin and 0.02% ethylenediaminetetraacetic acid (EDTA) (Lonza).
Cells were concentrated and re-suspended in DMEM containing 10% FBS.
Five Balb/c mice were anesthetized with 100 mg/Kg Ketamine and 50mg/Kg xylazine. The mice the received gentle anal dilatation using blunt-tipped forceps at the anal opening. Using a 29-gauge syringe, murine colon cancer CT26-GFP transfected were injected sub-mucosal into the distal-posterior rectum at concentrations of $1x10^3$ in a volume of 100 μl.
The mice were killed by CO_2 on day 28. After dissection and removal of tumor (rectum), mesenteric lymphnodes, liver and kidneys, tissue samples were prepared for histological analysis. For Hematoxylin and Eosin staining and immunohystochemistry analysis the tissue samples were fixed in formalin and embedded in paraffin.

3.18.3 Immunohistochemistry on mouse tissues

Immunohystochemistry on mouse tissues was performed as previous described (**section 3.16**).
When was used a primary antibody of Mouse origin to prevent a-specific signals, the primary antibody was prior biotinylate by using Mouse on Mouse Biotinylation System (BioCare Medicals).
5 to 10 μl of antibody is mixed with a biotinylation reagent for 30 min.

The biotinylated primary antibody complex can be immediately applied to the tissue specimen. The tissue is then incubated with streptavidin-peroxidase, and the reaction is developed by a diaminobenzidine/hydrogen peroxide chromogen-substrate (DAB).

Table IIIa: *List of Antibodies used for Immunohistochemistry (IHC-P), Immunofluorescence (IF) and Western Blotting (WB)*

Antibody	Species	Application	Dilution	Producer
TWIST1	mouse	IHC-P IF WB	1:30 1:30 1:100	ABCAM
E-CADHERIN	mouse	IHC-P IF WB	1:100 1:100 1:500	DAKO
N-CADHERIN	rabbit	IF WB	1:400 1:1000	ABCAM
VIMENTIN	mouse	IHC-P IF WB	1:100 1:100 1:1000	DAKO
CD54	mouse	IHC-P	1:100	DAKO
α-SMA	mouse	IHC-P	1:100	DAKO
CD45RO	mouse	IHC-P	1:100	DAKO
CD138	mouse	IHC-P	1:100	DAKO
CD68	mouse	IHC-P	1:100	DAKO
CD20	mouse	IHC-P	1:200	DAKO
FSP-1	rabbit	IHC-P	1:50	DAKO
ACTIN	mouse	WB	1:10000	ABCAM
CD133	mouse	IF	1:1	Miylteni
GFP	rabbit	IHC-P	1:200	INVITROGEN

Table IIIb: *Primer sequences used for quantitative Real-Time PCR (qRT-PCR) and for fragment analysis and sequencing (sequencing)*

Gene	Species	Assay	Sequence
18s	Human	qRT-PCR	Forward CGC CGC TAG AGG TGA AAT TCT Reverse CTT TCG CTC TGG TCC GTC TT
TWIST1	Human	qRT-PCR	Forward AGC AAG ATT CAG ACC CTC AAG CT Reverse CCT GGT AGA GGA AGT CGA TGT ACC T
SIP1/ZEB2	Human	qRT-PCR	Forward GCT ACA CGT TTT GCC TAC CGC Reverse CGA TTA CCT GCT CCT TTG GGT T
SNAI1	Human	qRT-PCR	Forward CTT CCA GCA GCC CTA CGA C Reverse CGG TGG GGT TGA GGA TCT
SNAI2/SLUG	Human	qRT-PCR	Forward TGT TTG CAA GAT CTG CGG C Reverse TGC AGT CAG GGC AAG AAA AA
ZEB1	Human	qRT-PCR	Forward GAA AGT CAT CCA GCC AAA TGG Reverse ACT TGG TTC TCA GCT TGG GGA ATC A
E-cadherin	Human	qRT-PCR	Forward GGA ACT ATG AAA AGT GGG CTT G Reverse AAA TTG CCA GGC TCA ATG AC
N-cadherin	Human	qRT-PCR	Forward GGT GGA GGA GAA GAA GAC CAG Reverse GGC ATC AGG CTC CAC AGT
Vimentin	Human	qRT-PCR	Forward TGG TCC TAC CCA CGC AGA TT Reverse GGC CAA CCC AGA AGT TGG AA
VEGFA	Human	qRT-PCR	Forward TAC CTC CAC CAT GCC AAG TG Reverse ATG ATT CTG CCC TCC TCC TTC
IL-10	Human	qRT-PCR	Forward TTA AGG GGT TAC CTG GGT TGC CAA GC Reverse ACT TGG TTC TCA GCT TGG GGC ATC A
CD133	Human	qRT-PCR	Forward ACA AGC CAG AAA CTG TAA TCT TAGG Reverse AGG CAT CAG AAT AAT AAA CAG CAG
TWIST1-20T	Human	Sequencing	Forward TCA CAC TCT GCA TTC TGA TAG AAG Reverse CAC GAC CTC TTG AGA ATG CATG
TWIST1 codifing seq	Human	Sequencing	Forward GAG GCG CCC CGC TCT TCT CC Reverse AAT CGA GGT GGA CTG GGA ACC G
TGFBRII	Human	Sequencing	Forward AGG ACT GCC CAT CCA CTG AGA CAC Reverse GGA GCT TGG GGT CAT GGC AAA CT
BAT25	Human	Sequencing	Forward TCG CCT CCA AGA ATG TAA GTG Reverse TCT GCA TTT TAA CTA TGG CTC
BAT26	Human	Sequencing	Forward TGA CTA CTT TTG ACT TCA GCC Reverse AAC CAT TCA ACA TTT TTA ACC C
18s	Mouse	qRT-PCR	Forward CGC CGC TAG AGG TGA AAT TCT Reverse CTT TCG CTC TGG TCC GTC TT
Twist1	Mouse	qRT-PCR	Forward AGC GGG TCA TGG CAT GGC TAA CG Reverse CAG CGT GGG GAT GAT CTT G
SIP1	Mouse	qRT-PCR	Forward AGA GTG TTG TGG AGC ACA G Reverse TGC ATC ATA ATC TTC CTT CAT
SNAI1	Mouse	qRT-PCR	Forward CTT CCA GCA GCC CTA CGA C Reverse CGG TGG GGT TGA GGA TCT
SLUG	Mouse	qRT-PCR	Forward TGT TTG CAA GAT CTG CGG C Reverse TGC AGT CAG GGC AAG AAA AA
E-cadherin	Mouse	qRT-PCR	Forward AAA ATG AAA AGG GCG AAT TT Reverse GTA GAA AAC CTT TGT TTC TTT GTC
N-cadherin	Mouse	qRT-PCR	Forward GGT GGA GGA GAA GAA GAC CAG Reverse GGC ATC AGG CTC CAC AGT
Vimentin	Mouse	qRT-PCR	Forward AAC CTG GCC GAG GAC ATC Reverse CTT TCG GCT TCC TCT CTC TCT G

4. RESULTS

4.1 Identification of Colon Cancer cells in EMT status expressing Twist1

The main feature of neoplastic epithelial cells going through a mesenchymal transition is the switch from a round-shape to an elongated fibroblast-like morphology. Accordingly, to identify CRC cell-lines in EMT-status we moved from a simple morphological observation. Out of 17 human CRC cell lines [9 MSS ones, Colo741, HT29, Colo205, SW403, CaCo2, CX-1, SW480, SW620, SW837, and 8 MSI ones, GP5D, HCT116, SW48, DLD1, LoVo, LS174T, LS180 and LS411], the CoLo741 cells, derived from a peritoneal metastasis, qualified for their fibroblast-like appearance(as shown in **Figure 4.1**).

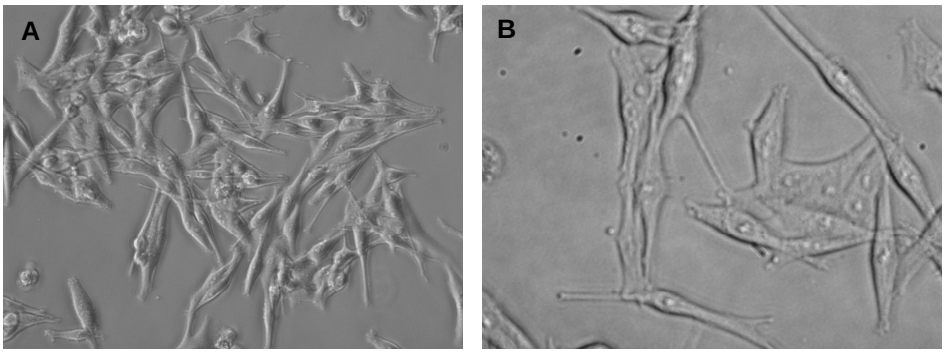


Figure 4.1: **A, B**. *Colo741 CRC cells showing the typical elongated mesenchymal morphology (20x objective and 40x objective, respectively)*

In parallel, the molecular phenotype was explored by gene-expression microarray. The analysis focused on the expression of *Twist1* the EMT master regulator, plus the other transcription factors involved in the transition: *SNAI1, SIP-1, Zeb1*, and SLUG. Microarray results indicated that *Twist1* expression was a rare phenomenon in cultured immortal CRC cells, as only two cell lines (12% of tested ones) expressed significant *Twist1* mRNA levels (as shown in **Table IVa**). Besides the MSS/CIN CoLo741 cells with a morphology consistent with the EMT status, the other was the MSI cell-line HCT116.
Specifically, the Colo741 cells were the only ones to simultaneously express 4 out of 5 EMT regulators: *Twist1, Zeb1, Slug* and *SIP1*, while all the other cell lines, expressed no EMT player (LS180), only one (*ZEB1*, in HT29, DLD1, LS174T, SW480 and CX-1, or SNAI1 in LOVO, SW403 and LS411), two (*SNAI1* and *ZEB1* in CO115, GP5D, SW48, CaCo2 and SW837, or *Twist1* and *ZEB1* in HCT116) or at most three (*SNAI1, SLUG,* and *ZEB1* in SW620) (see **Table IVa**). Consistently, only Colo741 cells showed the simultaneous absence of *E-cadherin* coupled with the over-expression of *N-cadherin* (for all the genes up/down expressed by CoLo741 see **Appendix I-II**).

Table IVa. *mRNA expression of EMT-associated transcription factors in CRC cell lines (by microarray analysis)*

	Transcription factors				
CRC Cell-Line	***ZEB1***	***SNAI1***	***SNAI2/SLUG***	***SIP1/ZEB2***	***Twist1***
LS180*	-	-	-	-	-
HT29	+	-	-	-	-
DLD1*	+	-	-	-	-
LS174T*	+	-	-	-	-
SW480	+	-	-	-	-
CX-1	+	-	-	-	-
LOVO*	-	+	-	-	-
SW403	-	+	-	-	-
LS411*	-	+	-	-	-
CO115	+	+	-	-	-
GP5D*	+	+	-	-	-
SW48*	+	+	-	-	-
CaCo2	+	+	-	-	-
SW837	+	+	-	-	-
HCT116*	+	-	-	-	+
SW620	+	+	+	-	-
CoLo741	+	-	+	+	+

*MSI CRC cell-lines, as opposed to MSS/CIN ones

4.2 Genes exclusively up- and down-regulated in CoLo741 cells

By analysis of the microarray data, a total of 345 genes were found to be differently expressed with respect to the other cells. Of these genes, 323 were expressed only by Colo741, while 22 were not expressed (see **Appendix I-II**).
Analysis with IPA software revealed that these genes mediate several biological functions as summarized below and reported in the detailed table below.

Diseases and Disorders	p-value	# Molecules
Genetic Disorder	3,34E-18 - 6,95E-03	198
Cardiovascular Disease	3,10E-16 - 7,91E-03	102
Neurological Disease	9,15E-16 - 7,81E-03	137
Skeletall/Muscular Disorders	1,69E-13 - 7,81E-03	112
Immunological Disease	7,31E-11 - 7,81E-03	91

Molecular and Cellular Functions	p-value	# Molecules
Cellular Development	2,79E-11 - 7,95E-03	99
Cellular Movement	1,58E-10 - 7,81E-03	69
Cellular Growth/Proliferation	2,64E-09 - 7,63E-03	95
Cell Death	2,77E-08 - 7,88E-03	91
Cell Morphology	8,24E-07 - 7,88E-03	59

System Development /Function	p-value	# Molecules
Tissue Development	4,89E-11 - 7,95E-03	75
Nervous Development	1,59E-05 - 7,81E-03	66
Hematological Development	4,73E-05 - 6,41E-03	42
Cardiovascular Development	5,91E-05 - 7,53E-03	34
Organismal Development	5,91E-05 - 6,17E-03	37

Canonical Pathway	p-value	Ratio
IL-8 Signaling	2,8E-05	12/186 (0,065)
Hepatic Fibrosis	3,51E-04	9/135 (0,067)
Huntington's Disease Signaling	1,01E-03	11/239 (0,046)
Thrombin Signaling	1,48E-03	10/205 (0,049)
CRC Metastasis Signaling	1,82E-03	11/249 (0,044)

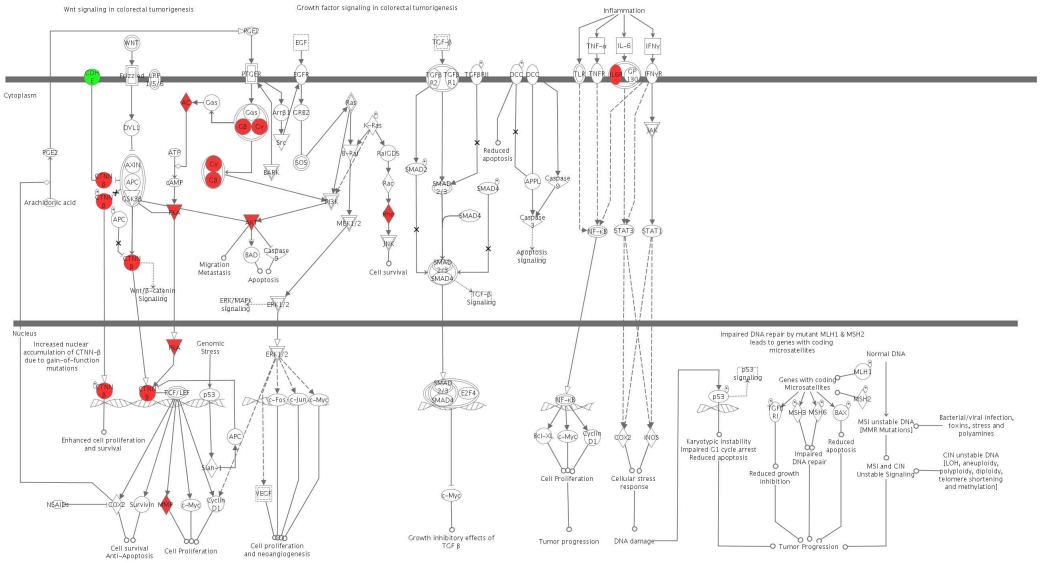


Figure 4.2 *Schematic representation by IPA software of CRC mestatasis signalling activated in CoLo741 cells. Green balloons indicate genes exclusively down-regulated while Red balloons indicate gene exclusively up-regulated*.

4.3 Spheroid pattern of growth of CRC cell-lines is restricted to epithelial cells, independently from the presence of CD133+ (stem) cells

In line with their phenotype and adhesive molecular pattern, epithelial cell-lines are able to grow in low adhesive plates to form three dimensional structures mimicking the structure of a primary glandular tumor. The spheroid formation allows to summarize cell-lines' adhesiveness and cohesive status, plus eventually their stem-cell like features.
After 15 days of culture all the cell-lines with an epithelial phenotype (HT29, HCT116 and SW480) formed tumor spheroids of different dimension, as shown in **Figure 4.3**, while CoLo741 cells remained attached to the plate in a monolayer, the same behavior displayed by fibroblasts cultured in these condition.
With respect to other CRC cell-lines, the lack of the capability to form spheroids further supports the unique mesenchymal phenotype of these cells, which lost the ability to adhere each other, rather tending to disseminate into the culture plate. This feature indicates that a stable EMT status confers to the Colo741 cells the capability to detach from other tumor cells, coupled to a pronounced tendency to spread.

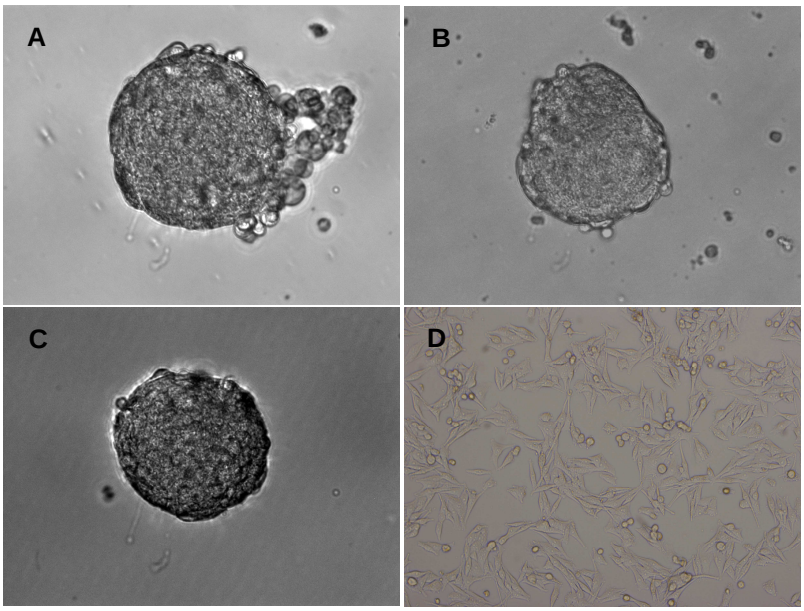


Figure 4.3 Example of *Spheroids derived from HT29 (***A***, upper-left panel), HCT116 (***B***,upper-right panel) and SW480 cells* (**C**,*down-left panel). Mesenchymal CoLo741 cells (***D***) remain attached to the plate.*

Additionally, in HT29 and SW480 cells forming spheroids we analyzed (by qRT PCR and by immunofluorescence) the expression of the CD133 stem-cell marker, thought to be related to the formation of colon-spheres. From qRT PCR analysis performed from mRNA extracted from the monolayer cultured cells resulted that CD133 gene was expressed only by HT29 (AQ 1,77E-05). As both HT19 CD133+ cells and SW480 CD133- cells were able to form spheroids, our results revealed that immortalized cells can growth into spheroids independently from CD133 expression (**Figure 4.4**). Finally, CD133 immunopositivity in HT29 cells was restricted to the spheroid core.
These results support the notion that within an epithelial cell population sharing a clonal origin, CD133 expression is dynamically modulated by factors which act differently depending upon cell location.

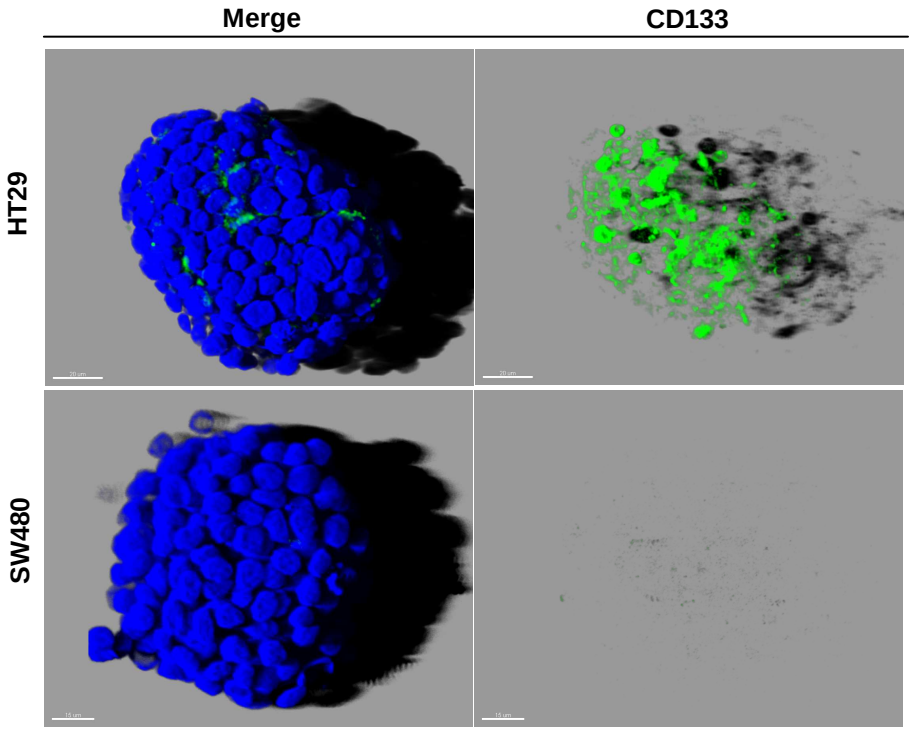


Figure 4.4 *Three-dimensional re-building with Imaris software of immunofluorescence staining for CD133. Examples of spheroids derived from HT29 CD133+ cells and SW480 CD133- cells (upper-left and down-left merge DAPI and CD133, upper-right and down-right CD133 staining)*

4.4 Aneuplody of chromosome 7 in CIN CoLo741 cells expressing Twist1

Considering the known MSS/CIN status of Colo741cells, we decide to approach their karyotyping by classical cytogenetic analysis. The purpose of these experiments was to identify a genetic hallmark of these Twist1 positive cells, to be later used for the FISH analysis of stromal Twst1 positive cells in human CRC specimens.
Chromosome count revealed that Colo741 cells comprise heterogeneous populations with different chromosomal number anomalies, ranging from 46 to 67 chromosomes (as shown in **Figure 4.5**).

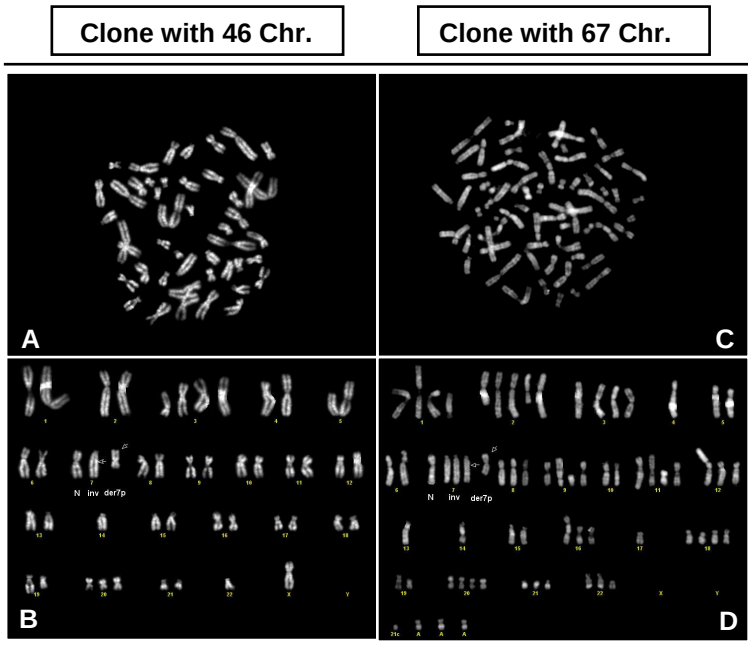


Figure 4.5 *Classical Cytogenetic analysis of CoLo741 cells;* **A,B** *CoLo741 clone with 46 chromosomes of which one is derived from rearrangements of chromosome 7.* **C,D** *CoLo741 clone with chromosomal anapleudy, 67 chromosomes with 4 copies of chromosome 7 of which 1 normal, 3 inverted and 1 derivative.*

Several chromosomes as shown in **Figure 4.5** were affected by translocation and different derivatives chromosome were seen. We identified by Q banding a different number of chromosome 7, from 2 to 4.
When thereafter was performed a FISH by staining for Twist1, position 7p21.1, and for CEP7 probes an abnormal number of Chr. 7 was confirmed. Additionally, by FISH we also identified a derivative chromosome 7 containing only Twist1 but not its centromere, and moreover a chromosome 7 with the inversion of a large part of

the p arm containing Twist1 (as shown in **Figure 4.6**). Of great interest, chromosome 7 derivative and inversion were common to all the nuclei studied, independently from their chromosomal content, confirming the same clonal derivation of the different cell populations.
The fact that the more hyperploid population can derive from the almost-diploid one is reinforced, besides the shared anomalies of the chromosome 7, by the fact the second showed a heterozygous *BRAF* mutation, while we detected a hemyzygous mutation in the hyperploid cells (data not shown).

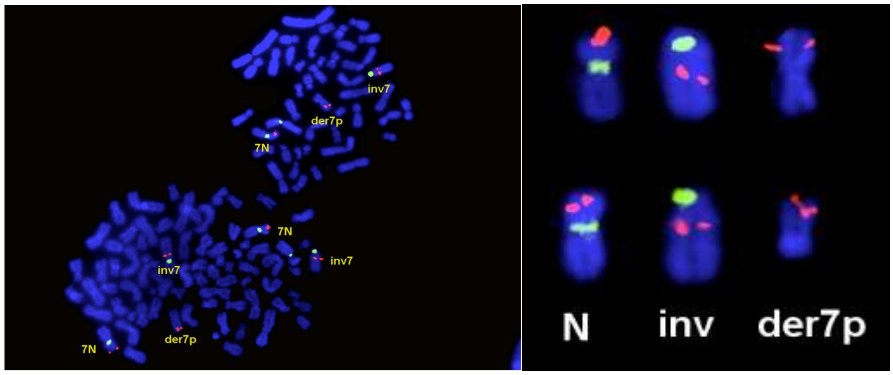


Figure 4.6 *In Situ Hybridization for CEP7 (green) and Twist1 (red) of CoLo741 chromosomes (left panel). Magnification and details of Chromosome 7 anomalies: a normal copy, an inversion of the p-arm containing Twist1 and a derivative of Chromosome 7 which lost the centromere and contained Twist1 signals (right panel).*

4.5 CIN CRC cells (CoLo741) but not MSI CRC cells (HCT116) express Twist1 protein and show a full mesenchymal phenotype

Once identified a cell-line with morphology and an expression pattern demonstrating its stabilized transition in a mesenchymal state, we focused on Twist1 as the master regulator of the epithelial to mesenchymal transition. As previously described Microarray analysis identified two cell lines expressing Twist1: CoLo741 and HCT116.
The two cell lines differ for the type of genetic instability: Colo741 are colon cancer cells with chromosomal instability (CIN/MSS), while HCT116 are CRC cells with microsatellite instability (MSI).
Twist1 expression microarray data were confirmed by qRT-PCR, western blot and immunofluorescence. As well as Twist1, a panel of other genes implicated in the transition (N-cadherin, Vimentin and E-cadherin) was studied to evaluate the activation of EMT according to the morphology.
Interestingly, despite the presence of *Twist1* transcript in HCT116, although less than in the Colo741 cells, they did not express any other marker of EMT (Vimentin, N-cadherin, as shown in **Figure 4.7**), retaining E-cadherin expression and maintaining a visible epithelioid morphology. In contrast CoLo741, the MSS counterpart, manifested all the features of the transition: a spindle-like morphology

and the mRNA expression of the mesenchymal markers proper to the EMT, such as N-cadherin, Vimentin (as shown in **Figure 4.7**), Fibronectin (see **Appendix I**), and three of the other transcription factor related to the phenomenon: SIP1, Slug and Zeb1 and contemporary absence of the epithelial marker E-cadherin.
To better understand the discrepancy between *Twist1* mRNA expression and the absence of other EMT hallmarks in HCT116, we decided to investigate its protein status compared to Colo741 and to HT29 cell line as epithelial control.
Despite *Twist1* mRNA expression in HCT116 the protein was undetectable (as shown in **Figures 4.7-4.8**). As to the expression of the epithelial marker E-cadherin (), we found that there was an overlapping pattern of both epithelial and mesenchymal protein markers, between the epithelial cells HT29 and the *Twist1*-expressing HCT116 cells. The absence of expression of N-cadherin and Vimentin and the expression of E-cadherin were common to the two cell-lines (as shown in **Figure 4.7-4.8**) while an opposite situation was seen in CoLo741, where according to the literature the presence of Twist1 protein come together with the repression of E-cadherin and an enhanced activation of N-cadherin and Vimentin.
These results sustained the possibility that Twist1 impaired expression in HCT116 is mainly due to its MSI type of genetic instabilty.

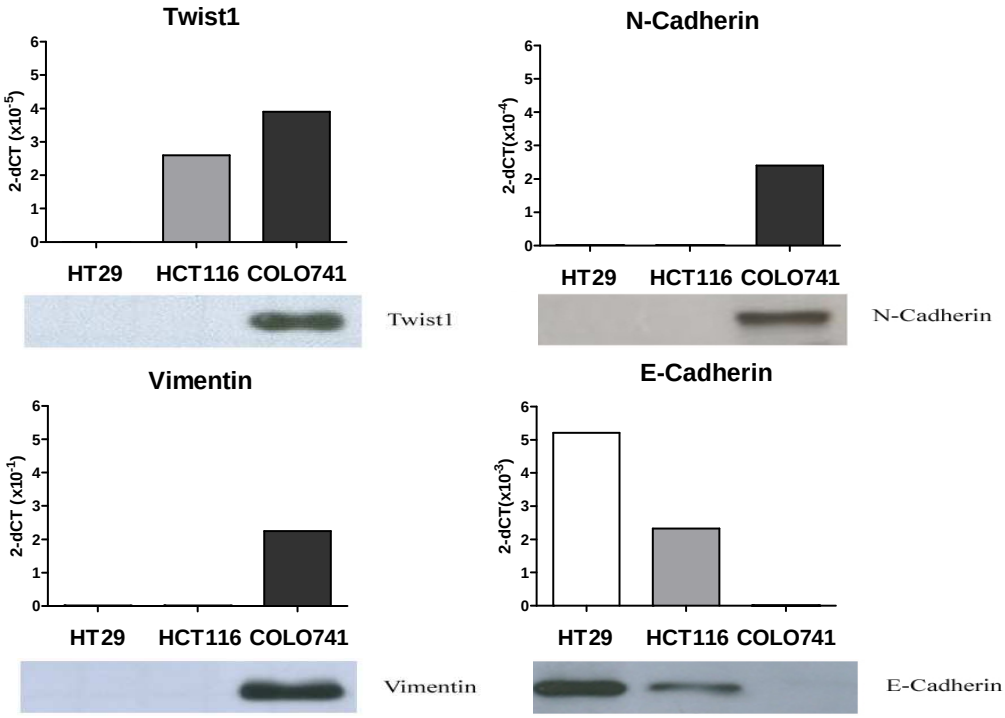


Figure 4.7 *mRNA expression as Absolute Quantification (AQ, 2-dCT method) and protein expression of mesenchymal markers: Twist1, N-Cadherin and Vimentin, and Epithelial marker E-cadherin in HT29, HCT116 and CoLo741 CRC cells,, by qRT-PCR and Western Blot.*

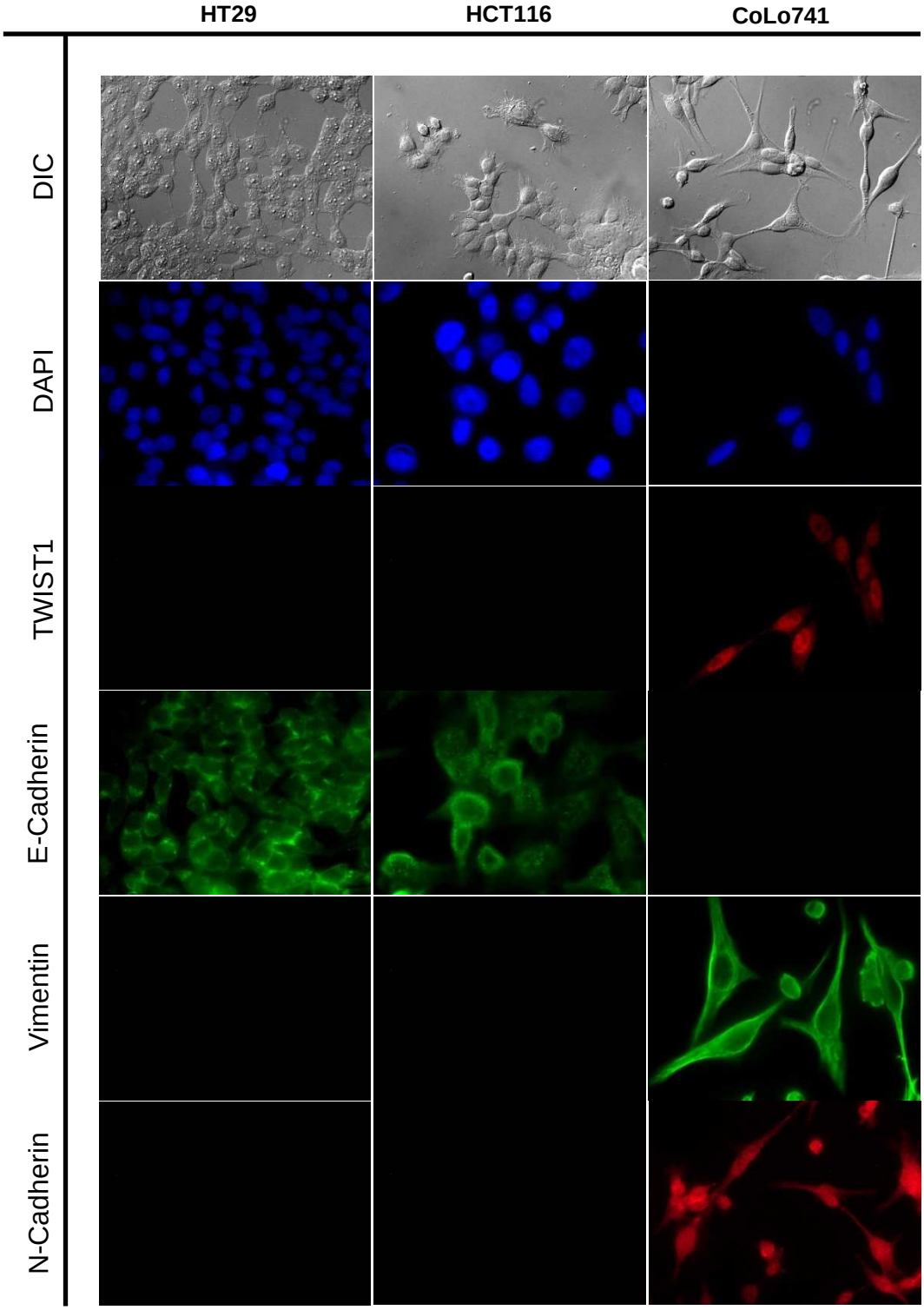


Figure 4.8 *Different protein expression of mesenchymal and epithelial markers in HT29, HCT116 and CoLo74 cells, by immunofluorescence (40x objective)*

4.6 DNA methylation affects Twist1 mRNA but does not account for the lack of protein in MSI CRC cells

A plausible explanation of the lack of *Twist1* protein in HCT116 might be related to the epigenetic phenomen of methylation. For this reason, we performed an analysis of Twist1 expression in HCT116 and Colo741 after treatment with the de-methylator agent 2'-deoxy-5-azacytidine (DAC).
In HCT 116 treated with 4µM of DAC for 78 h, a 3-fold increase of *Twist1* transcript was detected as compared to the untreated cells (**Figure 4.9**). Differently, the levels of *Twist1* mRNA in Colo741 were no different before and after treatment.
Following DNA de-methylation in HCT116, coupled with *Twist1* mRNA increase, we also observed the appearance of the Twist1 protein expression, as shown in **Figure 4.9**. Noticeably, while the amount of *Twist1* transcript in de-methylated HCT116 was identical to that found in Colo741 cells in basal condition (HCT116 AQ mean±SD 1.06 E-06 ± 7.28 E-08, AQ mean±SD CoLo741 1.65 E-06 ± 7.71 E-08) the protein levels remained lower by 2-fold than in Colo741 cells (**Figure 4.9**).
Moreover, in HCT-116 *N-cadherin* mRNA and protein was not affected by the restoration of a minimal quote of *Twist1* protein, and the concomitant stimulation with TGF-β for 6h did not further affect Twist1 expression.
Our results suggest that methylation of *Twist1* promoter in HCT116 partially determines the low transcript levels (as compared to Colo741). In meantime when similar levels were reached, using DAC, protein level remained low, and N-cadherin expression absent.

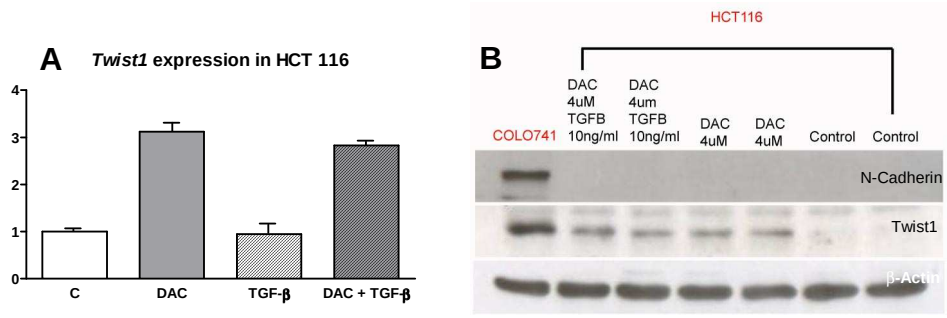


Figure 4.9 A *Twist1 mRNA expression in HCT116 cells after 78h of treatment with DAC, TGF-β or DAC+TGF-β;* **B** *Twist1 and N-Cadherin protein expression in HCT116 cells, pre- and post- DAC/TGF-β treatment compared to CoLo741 cells (by Western Blot)*

4.7 Analysis of Twist1 sequence in CRC cell-lines and tissues specimens according to MS- status.

To explain for the differential expression of Twist1 protein among the MSS and MSI CRC cell lines, we looked for the presence of frameshift mutations in Twist1 coding sequence (cds).

Twist1 cds was first entirely sequenced and compared in HCT116 and CoLo741.

Although the codifying sequence contained a non perfect tri-nucleotide repeats $(CGG)_{11}$, we did not found any genetic damage in MSI CRC specimens. However, exploring the 3'-UTR mRNA, we found a $(T)_{20}$ tract which was affected by shortening in MSI cell line HCT116. Moving from this finding we explored the presence of this shorted untranslated-region in other 5 MSI cell lines, of which one derived from a primary prostate cancer (DU145), compared to 5 MSS cell lines and in 61 mRNA derived from thirty MSI CRC and thirty-one MSS CRC tissues. None of the transcripts of the 5 MSS cell lines (0/5) showed a shortening of T20 tract, while five of the six MSI cell lines, including HCT116, had the $(T)_{20}$ tract in the 3-UTR reduced from 5 to 6 bases (5/6; 84%, p= 0.002), and only DU-145 did not show 3'-UTR shortening (as shown in **Figure 4.10**).

The same behaviour was seen in CRC tissues: 27 of 30 (90%) MSI CRC displayed the reduction of the 3-UTR region, but none of the MSS tissues (0/31; p p<0.0001) did (examples given in **Figure 4.10**).

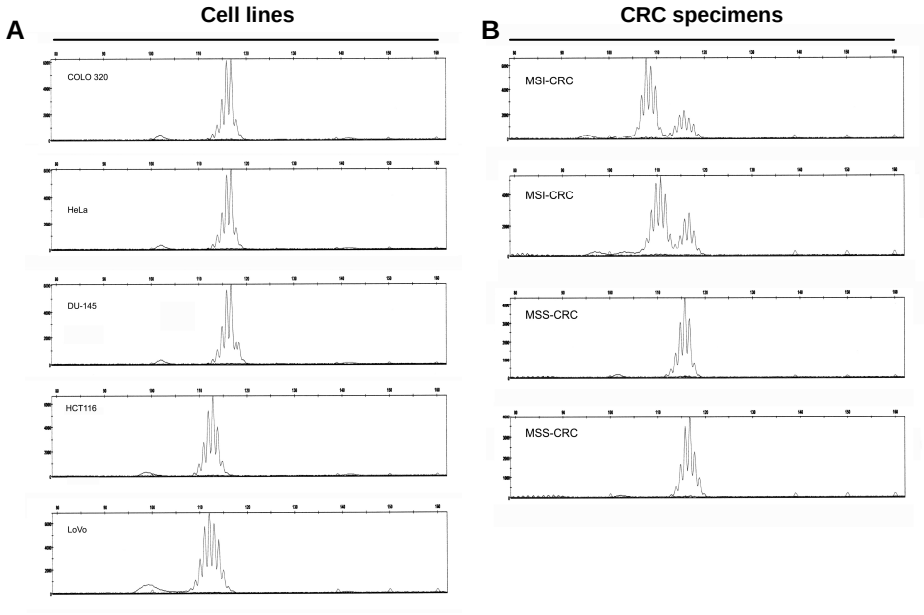


Figure 4.10 A. *Chromatograms of cancer cell lines analyzed for the 3'-UTR status (COLO320 and HeLa MSS cell lines with wild type 3'-UTR, DU145 MSI cell line with wild type 3'-UTR and HCT116, LOVO MSI cell lines with shortned 3'-UTR;* **B***. Chromatograms of 4 CRC specimens analyzed for their 3'-UTR status according to the MS-status (MSI CRC with a shortened 3'-UTR while MSS CRC with wild-type 3'-UTR)*

4.8 Frameshifts shortening Twist1 3-UTR as cause of mRNA instabilty in MSI cancer cells

The presence of a mutated 3'-UTR tract in mRNA is a well documented mechanism inducing instability of the transcript and impaired protein expression.
To study Twist1 mRNA stability according to the 3-UTR shorted tract, HCT116 and CoLo741 cells were treated in a time course experiment from 1h to 3h with Emetin and Actinomicin D, the first stabilizing the mutant transcript through the inhibition of the non sense mediated decay complex (NMD), implicated in the degradation of unstatble mRNA, and the second inhibits RNA transcription by binding DNA.
As internal control was used TGFBRII, a gene mutated in MSI [17] and degraded by NMD complex.
After 3h of treatment with Actinomicin D and Emetin, we saw an increase of ~35 fold of TGFBRII mRNA in HCT116, not in CoLo741. By analyzing Twist1 behaviour in the same conditions, we found that the transcript increased of ~3 fold in HCT116 respect to the untreated cells and CoLo741 (as shown in **Figure 4.11**), indicating an instability of Twist1 transcript.

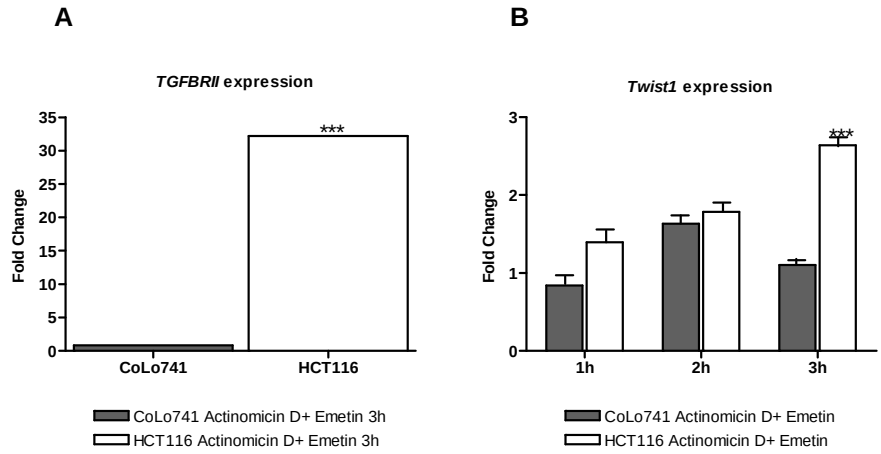


Figure 4.11 ***A****. mRNA expression of TGFBRII after 3h of treatment with Actinomicin D and Emetin in Colo741 CIN/MSS cells and HCT116 MSI TGFBRII mutated cells;* **B.** *Twist1 mRNA expression in CoLo741 and HCT116 cells after 1,2 and 3 h of treatment with Actinomycin-D and Emetin. A significant increase after 3h of Twist1 mRNA in HCT116 MSI cell line with shortened 3-UTR region respect to CoLo741 cells is shown.*

4.9 Twist1 expression allows to complete EMT by enhancing migration and invasiveness in CRC cells

Twist1, as transcription factor, plays a pivotal role in the modulation of the gene set which prime EMT. In breast cancer models, Twist1 is directly involved in the repression of E-cadherin, and in the transcriptional activation of N-cadherin[6]. However, its role in EMT gene activation or silencing in CRC remains unexplored.

To unravel the functions that Twist1 specifically exerts during invasion and migration of cancer cells in the colon, we analyzed the major target genes downstream Twist1 in CRC cellular models selected and modified on the basis of our microarray-based exploratory analysis of EMT gene expression in CRC cell-lines.

In a first model, *ZEB1* expressing SW480 cells, the transfection of *Twist1* (SW480 pCMV6.0-Twist1) resulted in its ectopic, constitutive expression at both mRNA and protein level (as shown in **Figure 4.13**). In SW480-pCMV6.0-Twist1 the two mesenchymal markers *N-cadherin* and *Vimentin* were up-regulated, respectively 3 and 6 fold, but the mRNA levels of *E-cadherin*, the epithelial gene selectively repressed by Twist1, did not differ with respect to control cells (SW480-pCMV6.0-empty vector).

In line with the partial mesenchymal switch documented by the mRNA pattern of SW480-Twist1 transfected, the analysis of the migration rate by wound-healing assay, showed an increase of the migration capability, as percentage of wound closure after 18h, 30h and 42h, of SW480-pCMV6.0-Twist1 compared to the SW480-pCMV6.0. At 30h, we recorded the largest difference in migration capability (**Figure 4.14**), when the closure of the wound break accomplished by SW480-Twist1 transfected cells was 22% ahead of the repair reached by the control. Results were similar when an invasion assay was perfomed on a matrigel coated chamber: the index of invasion, calculated based on the number of cells that invaded the matrigel membrane, was about two fold higher in SW480-Twist1 cells than in the control cells (mean±SD of SW480 pCMV6.0 invading cells after 24h, 33±3.6 versus 58±12.1 of SW480 pCMV6.0-Twist1 invading cells, $p<0.03$).

As a second explorative model of Twist1 role in CRC cells, we performed the stable knock-down of *Twist1* gene in our CRC cell line model of EMT. By silencing Twist1 with a short hairpin RNA in Colo741 (Colo741pLK0.1-shRNA-Twist1) we observed a change from dispersed cell growth to tightly packed colonies, coupled with a cell morphology close to an epithelial one (**Figure 4.12**).

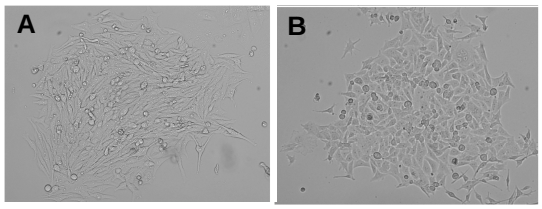


Figure 4.12 A *CoLo741-pLK0.1 showing a mesenchymal morphology,* ***B*** *CoLo741-pLK0.1-shRNA Twist1 with a more pronunciated epithelial morphology and a cobble –stone growth.*

The partial loss of the spindle like shape was associated with the *de novo* expression of the E-cadherin, at the mRNA and protein level, with unmodified

mRNA levels of N-cadherin (**Figure 4.13**). In keeping with the partial reversion from a mesenchymal to an epithelial-like phenotype (MET), a 47% decrease of the number of cells invading the matrigel membrane was seen in an invasion assay (mean±SD of Colo741-pKL0.1 invading cells after 24h was 102.3±24 versus 55±16 of Colo741-pLK0.1-shRNATwist1 invading cells, $p<0.05$).
Results suggest that Twist1 expression by itself is not sufficient *per se* for the switch from the expression of epithelial (E-cadh) to mesenchymal (N-cadh) markers (EMT), nor for the reverse process (MET). Differently, Twist1 expression confers enhanced motility, coupled with fibroblastoid-like shape and growth changes which lead to EMT completation in the background of N-Cadherin expression.
Twist1 alone can induce a partial switch to the mesenchymal phenotype and at the same time its knocking down a partial reversion to the epithelial morphology, Twist1 expression is a necessary condition to the EMT program and in particular its expression orchestrate CRC cells migration and invasiveness.

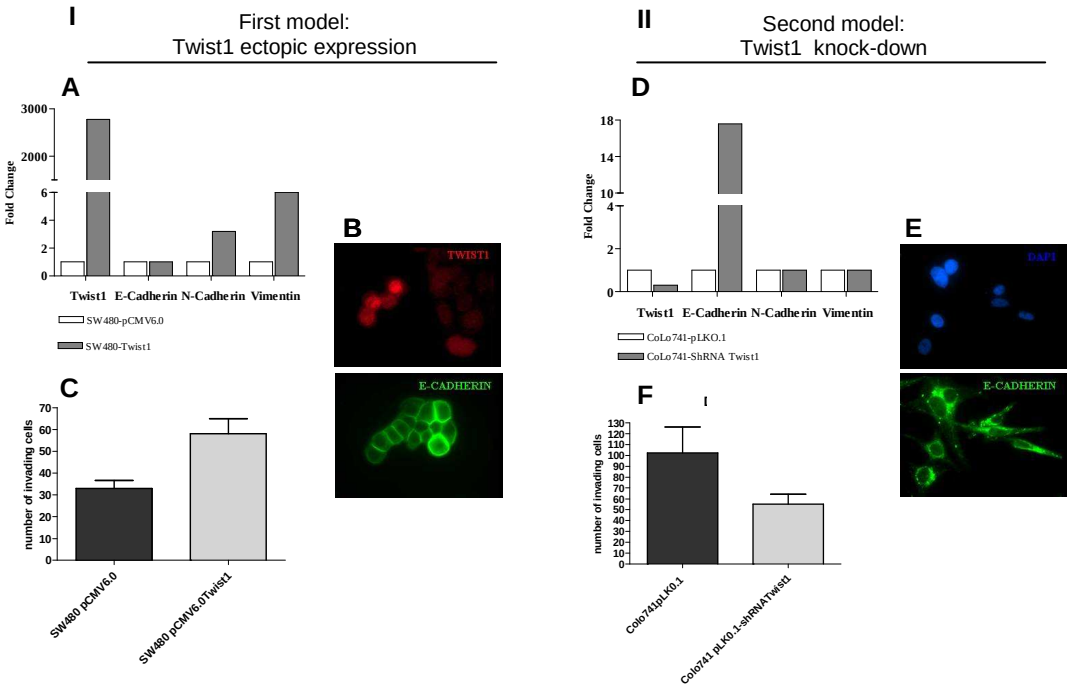


Figure 4.13 Panel I *First model: ectopic expression of Twist1 in SW480 CRC cells induces mRNA up-regulation of N-cadherin and Vimentin (***A***), not affects E-cadherin mRNA and protein expression (***B***) but increases cells invasion (***C***) (by Matrigel invasion assay).*
Panel II *Second model: Twist1 knock-down in CoLo741 CRC cells induces the expression of E-Cadherin at both mRNA (***D***) and protein level (***E***) and contemporary inhibits cell invasiveness (***F***).*

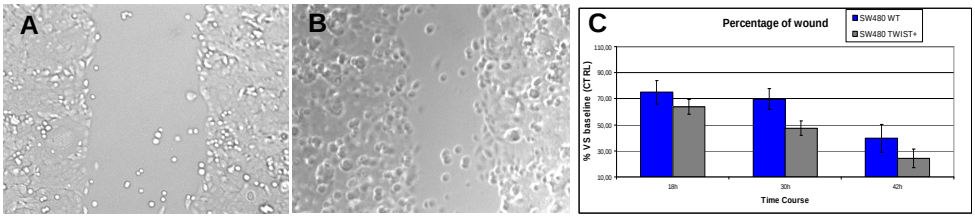


Figure 4.14 *SW480 migration assay:* **A** *wound enclosure of SW480-pCMV6.0 after 30h from the stretch;* **B** *wound enclosure of SW480-pCMV6.0-Twist1 after 30h from the stretch;* **C** *Graph indicating the percentage of wound enclosure of SW480 and SW480 Twist1 transfected with SD (error bars) measured at 18, 30 and 42 ho. Data shown are the result of three independent experiments (**, P=0.0034; ***, P < 0.0001).*

4.10 Effects of microenvironment modulators on Twist1 expression

It is known that factors of the tumor microenvironment may induce the perturbation of the epithelial state of cancer cells leading to a partial or complete EMT through Twist1 up-regulation. The two main micro environmental factors driving the EMT are hypoxia [424] and the soluble TGF-β [330].
To test whether hypoxia and/or TGF-β can induce *de novo* Twist1 expression in HT29, or up-regulate its expression in HCT116 and/or CoLo741 cell-lines, we performed stimulation studies in which the mRNA of the downstream targets of HIF-1alfa (*VEGF*) and TGF-β (*IL-10*) were also assessed.
Six hours after the treatment with TGF-β we found that only HT29 and Colo741 were responsive to the stimulus (IL-10 increase 3.7 and 2 fold respectively, Figure). HCT116 were unable to respond as predicted by their MSI status associated with the lack of functional TGFBRII due to its frameshifted coding sequence [439].
Opposite to the un-responsiveness of the other cell lines, only Colo741 cells up-regulated Twist1 (≈2-fold increase, **Figure 4.15**).

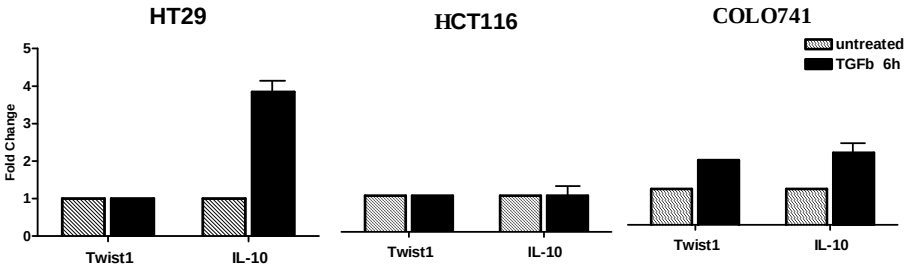


Figure 4.15 *Twist1 and IL-10 expression after six 6h of treatment with TGF-β in HT29, HCT116 and CoLo741 CRC cells.*

Similar results were observed when the cells were cultured in a hypoxic environment. Six hours of oxygen deprivation induced the up-regulation of HIF-1α

which consequently increased the expression of *VEGF* in all of the cell lines (mean increase: 5-fold, as shown in **Figure 4.16**). Even though HIF-1α is a direct regulator of *Twist1* transcription and all the treated cell lines were sensitive to the hypoxic condition, only CoLo741 up-regulated *Twist1* mRNA.
To expand these observations aimed to explore the microenvironmental capability to induce and/or sustain EMT in CRC cells, we also performed stimulation with other cytokines (namely: IL-6, IL-10, TNF-α and INF-γ) which have not yet directly related to the EMT (data not shown). After ten hours of treatment, none of these stimuli induced or up-regulated Twist-1 expression in all the cell lines. These results indicate that short microenvironment stimuli push EMT through *Twist1* activation only in neoplastic clones which are already in the EMT status.

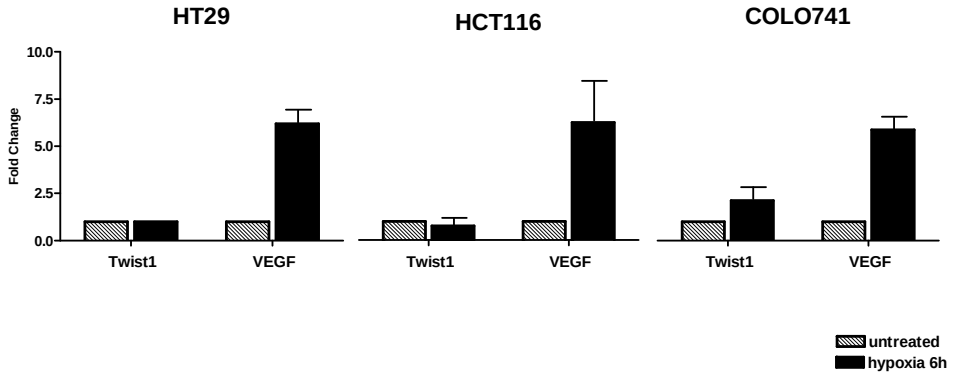


Figure 4.16 *Twist1 and VEGF expression in HT29, HCT116 and CoLo741 CRC cells cultured for 6h in Hypoxic conditions.*

4.11 TWIST1 mRNA levels are increased in Colorectal Cancer and are significantly associated with local invasion, nodal and distant metastasis

To elucidate the involvement of Twist1 in CRC we analyzed the mRNA levels in 78 CRC specimens compared to their normal colorectal mucosa, using quantitative real-time PCR. 18S ribosomal RNA was used as the internal control and the analysis of gene expression was performed using the absolute quantification (AQ) 2-DCt method (normal mucosa AQ and tumour AQ). The AQ expression pattern between normal and tumour tissues was found to be significantly different (P<0.001, Student's t-test). The mean value of normal AQ was 3.3 E-07± 8.5 E-07, whereas tumour tissues displayed a mean AQ of 2.7E-06±6.9E-06 (**Figure 4.17**). Our results indicates that Twist1 mRNA is up-regulated in CRC tissues compared with their normal counterparts. Furthermore, concerning the MS-status of tumoral specimens, no differences in Twist1 mRNA levels between MSS (n=65, 83.3%; AQ mean: 2.7E-06±7.5E-06) and MSI cancers (n=13, 16.7%; AQ mean: 3.0E-06±3.6E-06, p=0.5) was found. Considering that the *Twist1* transcript is unstable in the MSI CRC cell-line HCT116 with impaired EMT, despite the presence of *Twist1* mRNA in MSI CRC, the analysis of clinico-pathological features was performed only for MSS CRC.

mRNA level of *Twist1* was significantly correlated with CRC stage at diagnosis. As shown in **Figure 4.17**, according to TNM staging system, Twist1 expression increased with the extension of local invasion (expression in pT3 and pT4 CRC *versus* expression in pT1-pT2 CRC, p=0.001 and p=0.01, respectively). Moreover, Twist1 expression was significantly up-regulated in patients with regional nodal invasion (Stage III) and metastasis in distant organs (Stage IV) than in those who did not present lymph node or distant organ metastases (AQ mean stage III/IV: 4±1.4E-6 *vs* AQ mean stage I/II: 1.1±1.4E-6; p=0.04). These findings suggest that *Twist1* is involved in the progression of CRC.

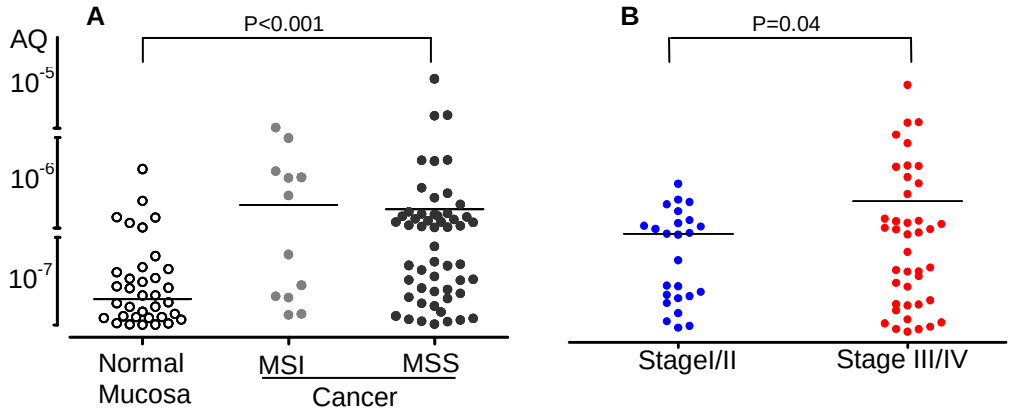


Figure 4.17 A *Twist1 mRNA expression in CRC specimens compared to their Normal counterpart and according to MS-status;* **B** *Twist1 mRNA expression according to the Stage at diagnosis.*

4.12 Immunolocalization and characterization of Twist 1+ cells in CRC tissues

By using an immunohistochemical approach, the expression of Twist1 was evaluated in 201 specimens of surgically resected primary CRC. Only cells with nuclear reactivity were considered as Twist1 immunopositive cells. For each CRC specimen included in the study, two sections were analyzed. Twist1 was detected in 165 of the 201 tumor tissues analysed (82%) but not in normal colonic mucosa. Representative stained tissues are shown in **Figure 4.18**.

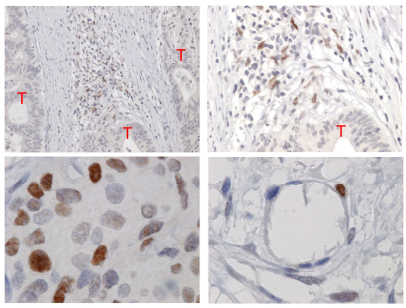
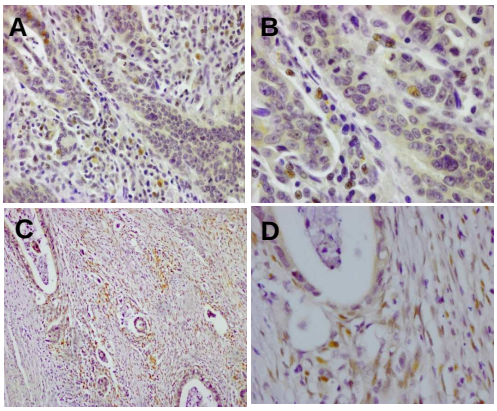


Figure 4.18 *Rappresentative staining of Twist1 in tumor stroma (upper-left and right, 10x and 40x objective, respectively),in a tumor gland (down-left picture, 100x objective) and endhothelial cells (down-right, 40x objective)*

The number of immunopositive cells detected in CRC specimens with Twist1+ immunostaining was variable, ranging from few (Score 1), to widely diffuse immunopositive Twist1+ cells (**Figure 4.19**, score 2-3). Nuclear Twist1 expression was recognizable in different tumor compartments: in the neoplastic glands and in the stromal compartment (Tumor and Stroma pattern), where Twist1+ cells had a spindle fibroblast-like morphology (as shown in **Figure 4.19**). In several cases Twist1 was detected only in the stromal, not in tumor, cells (Stromal pattern). In the latter cases, immunopositive cells were located close to tumor glands (**Figure 4.19**), at the invasive front, in areas where the tumour was losing its glandular epithelial structure (**Figure 4.19**).

Figure 4.19 A,B *Twist1 positive cells (nuclei brown) in CRC glands and in the stromal compartment (20x and 40x objective, respectively);* **C,D** *Twist1 positive cells with fibroblast-like morphology in the stroma, detaching from a degenerating gland (20x and 40x objectives, respectively)*



By double immuhystrochemistry (Twist1 and CD54) we found a high density of Twist1 positive cells in proximity of blood vessels (**Figure 4.20**), even resembling intravasation. At the same time, occasionally, endothelial cells were found immunopositive (**Figure 4.18**).

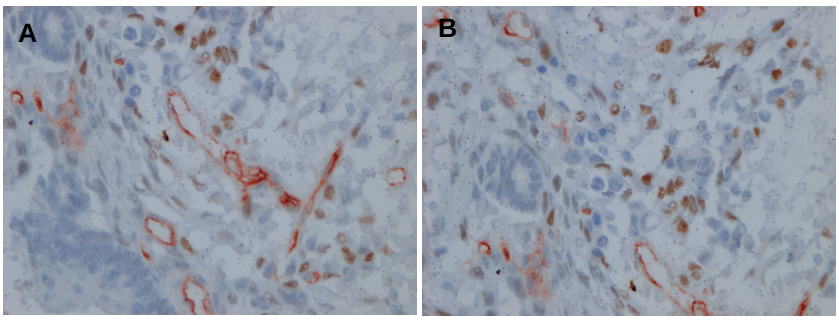


Figure 4.20 A,B *Examples of Twist1 positive cells (Brown) located in proximity of blood vessels (Red) (40x objective), (by double immunohistochemestry).*

To asses the presence of Twist1+ cells in the metastasis of CRC patients, immunohistochemistry was also performed in three metastasis (liver, uterus and lymph-node) derived from the same Twist1+ primary CRC. Few Twist1+ cells were seen in the liver and in the lymph-node metastasis, while in the uterus metastasis Twist1+ cells were abundant in the stromal compartment (**Figure 4.21**).

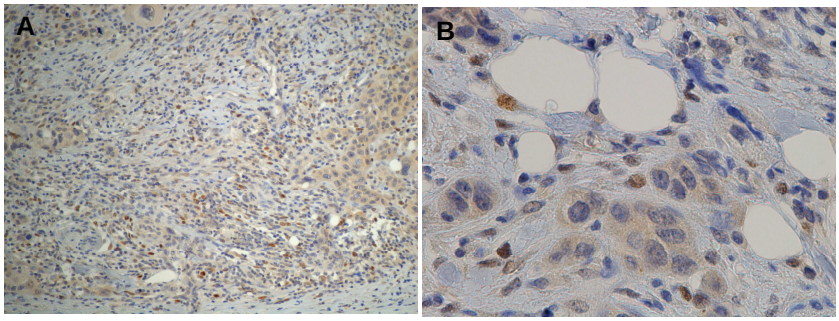


Figure 4.21 A,B *Twist1 immunopositive cells in the stromal compartment of a uterus CRC metastasis (10x and 40x objectives, respectively)*

To better characterize the phenotype of Twist1 immunopositive cells doubles staining were performed. Antibodies against both epithelial (E-cadherin) and mesenchymal (Vimentin, α-Smooth Muscle Antigen (α-SMA) antigens were used. As shown in **Figure 4.21** all the Twist1-immunopositive cells located in the tumor gland and those in the stromal compartment, lack the expression of E-cadherin, while a number of cells expressed Vimentin or α-SMA. However, in all the sections examined, a variable amount of Twist1 immunopositive cells were found immunonegative for both epithelial and mesenchymal markers.

Moreover, different attempts with antibodies for Cluster of Differentiation (CD), proper of sub-population of the immune system were also tried (CD45RO, CD138, CD68, CD20), to draw an immunologic profile of Twist1+ cells. As shown in Figure X none of the Twist1 immunopositive cells co-expressed these antigens. Additionally to exclude the fibroblast origin of Twist1 immunopositive cells a double staining with antibodies raised against fibroblast-specific-protein-1 (FSP1) and Desmin were performed. All the Twist1 immunopositive cells were immunonegative for both.

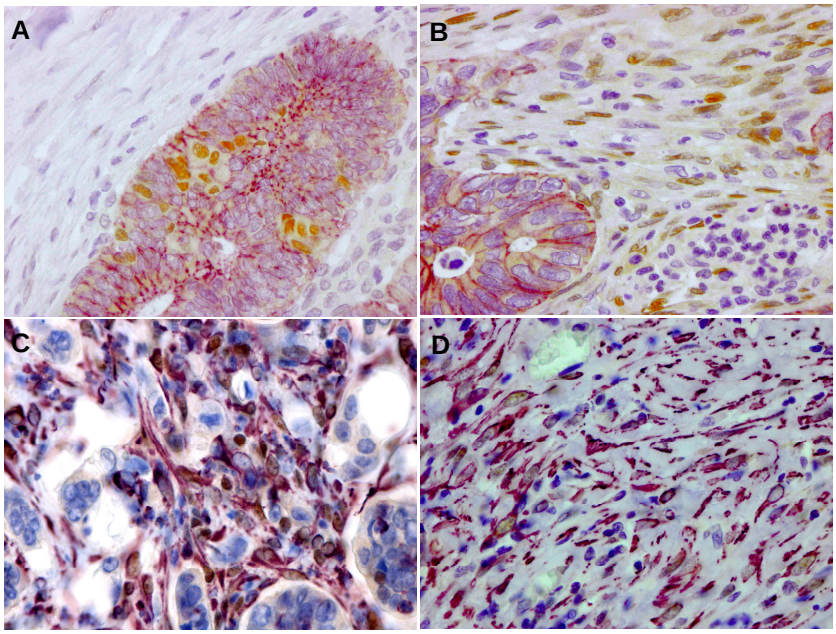


Figure 4.21 A,B *Double immunohystochemistry for Twist1 and E-cadherin. Twist1+ (Brown) cells are located in both tumor glands and stroma and do not express epithelial marker E-cadherin (Purple), while Twist1 positive cells (Brown) in the stroma are positive for Vimentin (Purple) (***C***) and Smooth-muscle-antigen (Purple) (***D***).*

4.13 Twist1 protein expression in CRC specimens and patients' clinical-pathological features

Resembling Twist1 expression pattern in our series, tissue specimens were categorized in three classes, on the basis of the localization of Twist1 immunopositive cells: a) tissues completely immunonegative for Twist1 (negative pattern), b) tissues with immunopositive cells located in both the stroma and in tumor glands (stromal and tumor positive pattern), and c) tissues containing immunopositive cells located only in the stroma (stromal positive pattern).
Thirty-six of 201 (18 %) tissues were negative for Twist1 expression. Twist1 immunoreactivity was mainly observed in the stromal compartment (116/201, 57.8%, see **Table IVb**.). In 49 of 201 (24.4%) tissues Twist 1 immunopositive cells were recognized in both the stroma and the tumor glands. Considering the MS-status, stromal-only Twist+ specimens were more frequent in MSS (109/174; 62.6%) than in MSI (7/27; 25.9%; p=0.015 by Pearson's test) CRC (**Figure 4.22**).

		Twist1-	Twist1+ T&S	Twist1+ S	P
All patients	201 (100%)	36 (18%)	49 (24.4%)	116 (57.8%)	
Gender					
M	114 (56.7%)	20 (17.5%)	28 (24.6%)	66 (57.9%)	0,98
F	87 (43.3%)	16 (18.4%)	21 (24.1%)	50 (57.5%)	
Age (mean±STD)	64.8±11.9	65.0±11.9	64.9±11.9	64.9±12	0,98
Microsatellite status					
MSS	174 (86.6%)	29 (16.7%)	36 (20.7%)	109 (62.6%)	0,001
MSI	27 (13.4%)	7 (25.9%)	13 (48.1%)	7 (25.9%)	
Tumor site					
Proximal colon	70 (35,0%)	15 (21,4%)	19 (27,2%)	36 (51,4%)	
Distal colon	89 (44.5%)	12 (12,4%)	22 (24.7%)	56 (62.9%)	0,34
Rectum	41 (20.5%)	10 (24.4%)	8 (19.5%)	23 (56.1%)	
Tumor stage					
I	27 (13.4%)	10 (37.0%)	7 (26.0%)	10 (27.0%)	
II	69 (34.3%)	16 (23.2%)	28 (26.1%)	35 (50.7%)	0,008
III	60 (29,9%)	7 (11.7%)	16 (26.6%)	37 (61.7%)	
IV	45 (22.4%)	3 (6.7%)	8 (17.8%)	34 (75.6%)	
Tumor invasion					
pT1	8 (4.0%)	3 (37.5%)	3 (37.5%)	2 (25.0%)	
pT2	25 (12.4%)	7 (28.0%)	6 (24.0%)	12 (48.0%)	0,31
pT3	127 (63.2%)	20 (15.7%)	32 (25.2%)	75 (59,1%)	
pT4	41 (20.4%)	6 (14.6%)	8 (19.5%)	27 (65.9%)	
Tumor grade					
G1	21 (10.4%)	6 (28.6%)	6 (28.6%)	9 (42.8%)	
G2	122 (60.7%)	23 (18.9%)	26 (21.3%)	73 (59.8%)	0,33
G3	58 (28.9%)	7 (12.1%)	17 (29.3%)	34 (58.6%)	
Nodal Metastasis					
N0	99 (49.3%)	27 (27.3%)	25 (25.3%)	47 (47.4%)	0.001
N+	102 (50.7%)	9 (8,8%)	24 (23.5%)	69 (67.7%)	
Metacronous Metastasis					
M0	155 (76.7%)	34 (21.9%)	40 (25.8%)	81 (52.3%)	0.014
M+	47 (23.3%)	3 (6.4%)	9 (19.1%)	35 (74.5%)	
Angioinvasion					
Yes	151 (75.1%)	30 (21.9%)	40 (25.8%)	81 (52.3%)	
No	47 (23.4%)	6 (12.8%)	9 (19.1%)	32 (68.1%)	0,21
N.A.	3 (23.3%)	0	0	3 (100%)	

Table IVb *Association of Clinical-pathological features of CRC patients with Twist1 protein expression (for P-value was used Chi-Square Test or Pearson's Test, when appropriated).*

In MSI CRC Twist1 expression was not significant correlated with any clinico-pathological features.

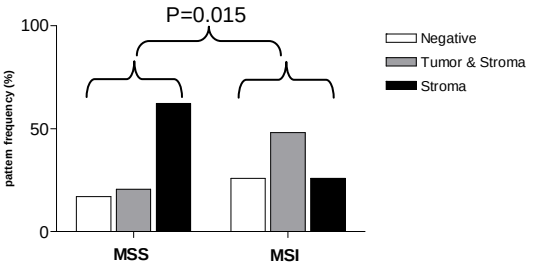


Figure 4.22 Twist1 protein expression in CRC tissues according to MS-status. MSS CRC show a significantly higher frequency of stromal pattern respect to MSI ones.

Expression of Twist1 was also evaluated according to the stage at diagnosis in MSS-CRC. Twist1-negative tissues decreased from stage I/II (21/80, 26.2%) to stage III (5/51, 9.8%, p=0.07 vs stage I/II) and IV (3/43, 7%; p=0.03 *versus* stage I/II), see **Figure 4.23**. Conversely, stromal-only TWIST1 immunopositive cancers increased from stage I/II (41/80, 51.2%) to III (35/51, 68.6% p=0.07) and IV (33/43, 76.7%; p=0.007 vs stage I/II). Furthermore, Twist1 stromal expression was associated with more locally invasive tumors (pT1/pT2, 13/29, 44.8% vs pT3/pT4, 96/145, 66.2%, p=0.04, **Figure 4.23**).
A correlation between Twist1-expression and lymph-node status was found: the rate of stromal positive pattern versus negative pattern significantly increased from patients without to those with nodal metastasis (p<0.001, **Figure 4.23**). Similarly, the rate of CRC with stromal positive pattern versus those with a negative one increased from patients without distant metastasis to that with secondary lesions (p<0.03, **Figure 4.23**).

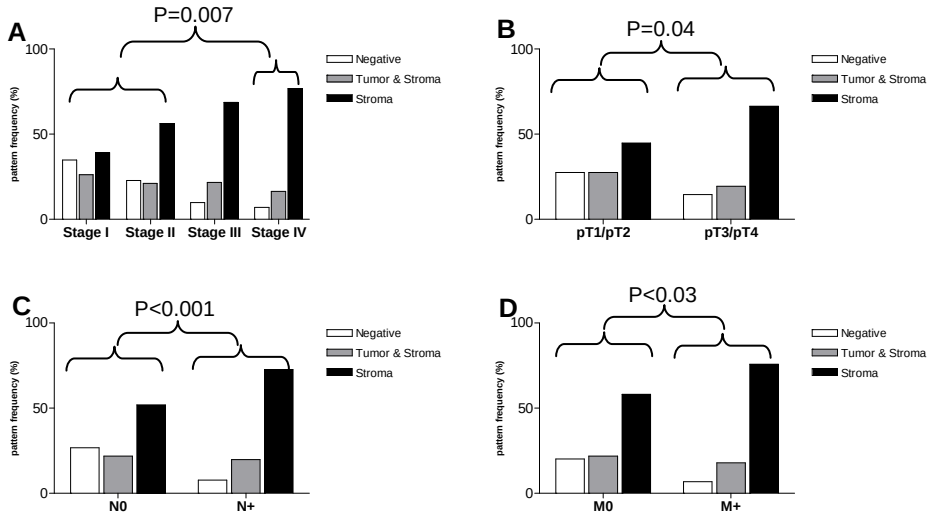


Figure 4.23 *Twist1 protein expression in MSS CRC.* **A***, the rate of stromal–negative Twist1 expression increase from Stage I/II to stage III and significantly to stage IV;* **B** *the rate of stromal–negative Twist1 expression significantly increase with local invasion;* **C,D** *Twist1 stromal expression significantly increase with nodal and distant metastasis.*

The correlation between patients' survival was also assed. Survival analysis was performed for Disease Specific Survival (DSS) and Disease Free Survival (DFS).
As shown in **Figure 4.24**, the Kaplan-Meier curves displayed that patients with MSS TWIST1-negative and MSI CRC had similar DSS and DFS, while the DSS and DFS of MSS CRC patients were worse for those with stromal-only TWIST1^{+} than for those with Twist1^{-} cancers (p=0.01 and 0.02, respectively). Moreover, also the presence of Twist1 in both tumor and stroma correlated with a poor DSS as compared to Twist1-negative CRC (p=0.03 DSS), although the difference as to disease recurrence did not reach the statistical significance (DFS, p=0.11).

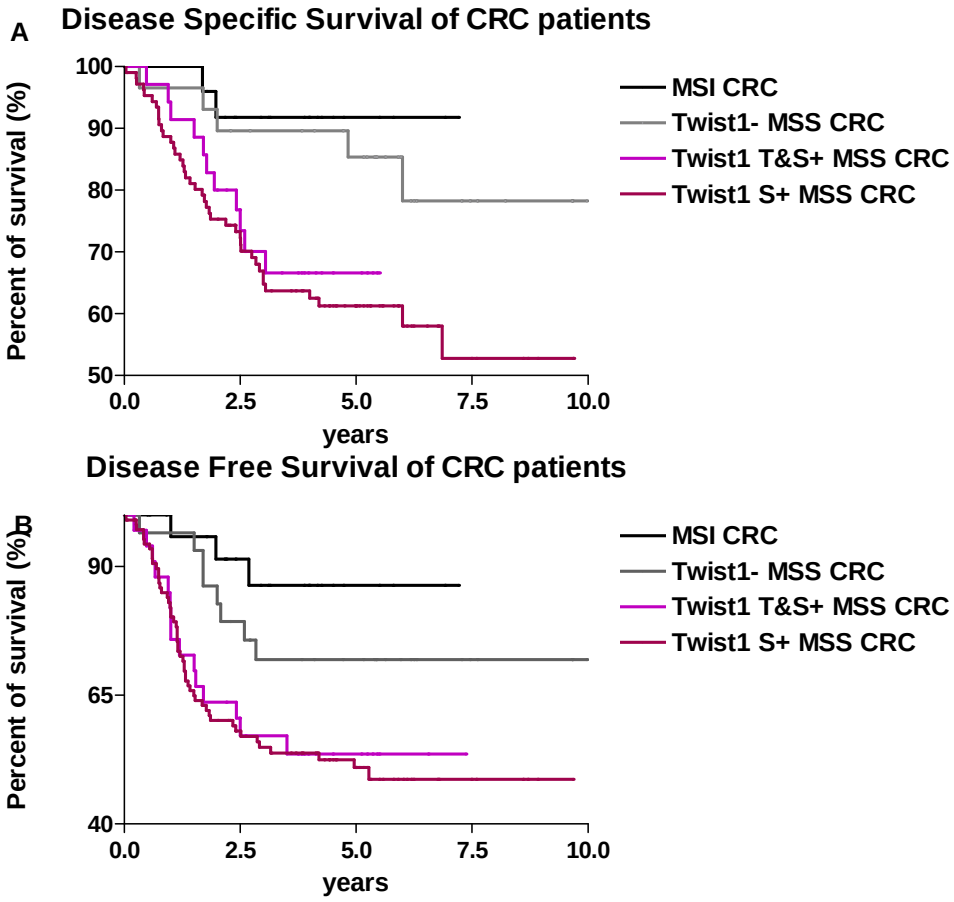


Figure 4.24 A *Disease Specific Survival (DSS) of CRC patients according to Twist1 expression (T&S Tumor and Stroma positivity, S Stromal positivity). Patients with T&S and S only positivity have a significant poor survival respect to those Twist1- (p=0.02 and p=0.01 respectively);* ***B*** *Disease Free Survival (DFS) of CRC patients according to Twist1 expression (T&S Tumor and Stroma positivity, S Stromal positivity). Patients with S Twist1+ have a significant lower survival than those Twist1- (p=0.03).*

4.14 Chromosome 7 aneuploidy can be traced in both epithelial cancer cells and in stromal, Twist1 positive cells. A pilot study.

Following immunohistochemistry for Twist1, we moved to perform fluorescent *in situ* ibridization (FISH) in CRC specimens by molecular cytogenetic analysis by using a specific DNA probe complementary to the centromer of chromosome 7.
Of two Twist1 positive, MSS CRC samples analyzed, one (pT3N1) was identified as unstable for the chromosome 7.
FISH revealed that the glandular neoplastic cells had different number of signals for chr 7, ranging from 1 to 4 as shown in **Figure 4.25**. Of ten fields evaluated focusing our attention on stromal Twist1 positive cells, we found that one Twist1 immunopositive cell in the stromal compartment, far from the neoplastic epithelial cells, had chr 7 trisomy, resembling a pattern saw in the neoplastic cells **(Figure 4.25**). None of the other stromal cells positive or negative for Twist1 expression shared the same chromosomal alteration.
Besides intrinsic limitations, this is the first experiment ever depicting the same cytogenetic clonal alteration in cancer cells and in the surrounding microenvironmental cells.
Twist1 positive cells in the tumor microenvironment are likely a mixture of activated mesenchymal cells with a normal kariotype hiding neoplastic cells undergone EMT and Twist1 activation.

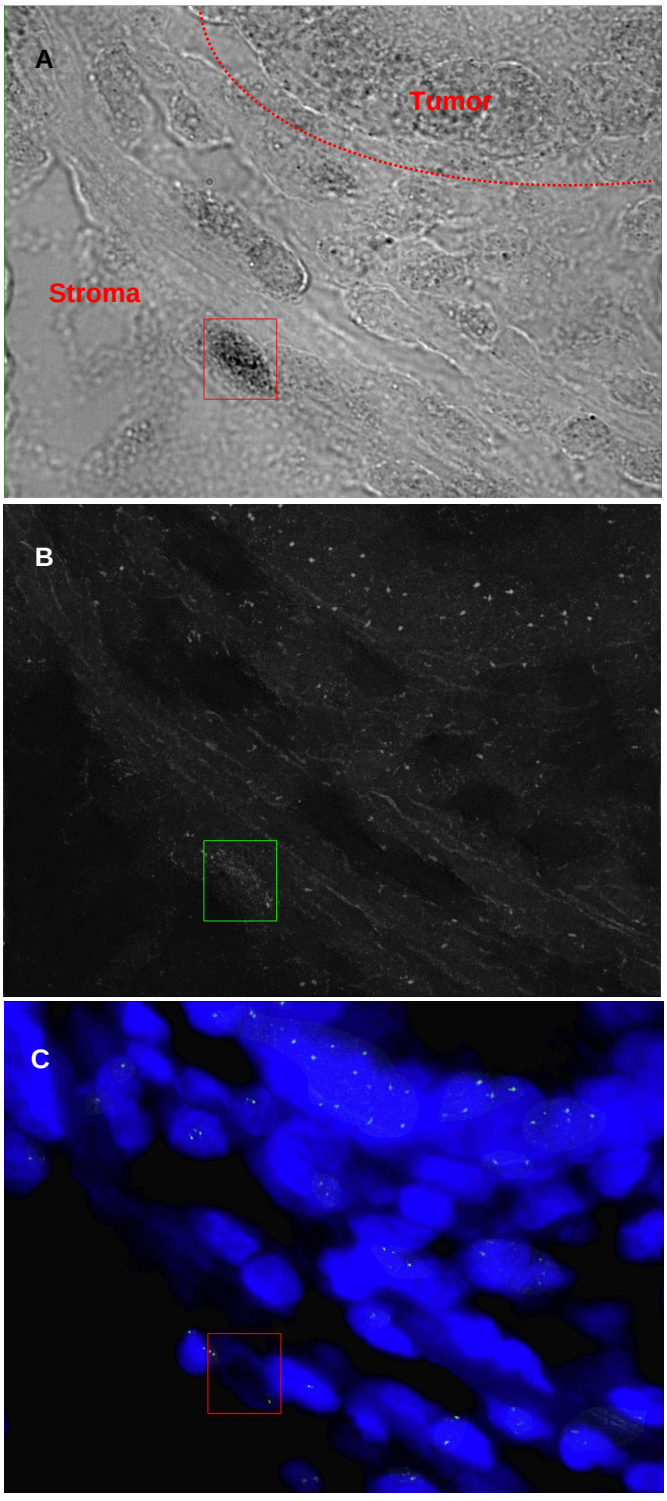


Figure 4.25 A *Brighfield image showing Twist1+ cells (dark grey, Red Box) in the stromal compartment;* **B** *Spectrum Green image showing signals for CEP7 (white points), Green box highlights Twist1+ cells in the stromal compartment with 3 signals for CEP7;* ***C*** *Merge image (DAPI (blue) and CEP7 (green)). Neoplastic cells in the glandular structure show aneuplody for Chromosome 7 (from 1 to 4 signals of CEP7,green), the same Chromosome 7 umbalance is shared by the Nucleus of Twist1+ cell in the stromal compartment (3 signals for CEP7 (red box)).*

4.15 CT26 mouse cells expressing Twist1: a murine model to study in vivo EMT

One of the cardinal points to study EMT *in vivo* is the identification of a suitable murine cell line that has a clear mesenchymal phenotype, consistent with the EMT program. Looking for such a model, we focused on CT26 cell-line, which is syngenic top Balb/c mice, has a high metastatic potential and a marked fibroblastoid morphology (**Figure 4.26**).

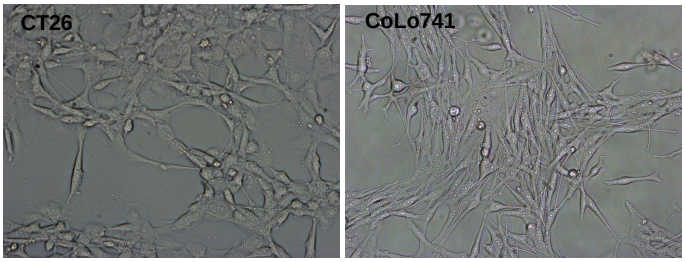


Figure 4.26 *Murine CRC CT26 and human CRC CoLo741 cells share the same mesenchymal morphology (20x Objective)*

The mesenchymal signature of CT26 was documented by qRT-PCR, showing also in CT26 cells the expression of the same markers associated with the mesenchymal phenotype of the human CRC cell line Colo741, coupled with the absence of the master epithelial marker E-cadherin. Both cell lines expressed Twist1, SIP-1, N-cadherin and Vimentin. Eventually, the levels of Twist1 and N-Cadherin mRNA detected in CT26 were higher than those found in Colo741, and at the meantime a lower expression of SIP1 and Vimentin was proper to CT26 (**Figure 4.27**). By W.B analysis we confirmed the presence of Twist1 and N-cadherin and Vimentin and contemporary the absence of E-cadherin protein.

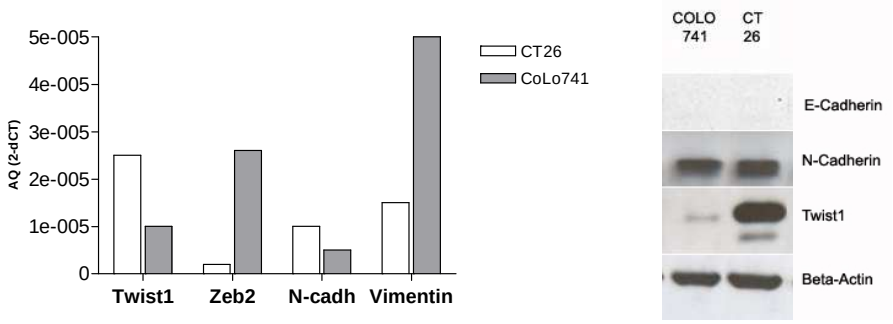


Figure 4.27 A *mRNA expression of EMT associated markers in CT26 and CoLo741 cells (by qRT-PCR);* **B** *Protein expression pattern of mesenchymal and epithelial markers in CT26 and CoLo741 cells (by Western Blotting)*

In vitro, GFP-transfected CT26 cell were morphologically indistinguishable from Balb/c mouse embryonic fibroblast (MEF), but by taking advantage of GFP expression, we can distinguish the two cell lines (**Figure 4.28**).

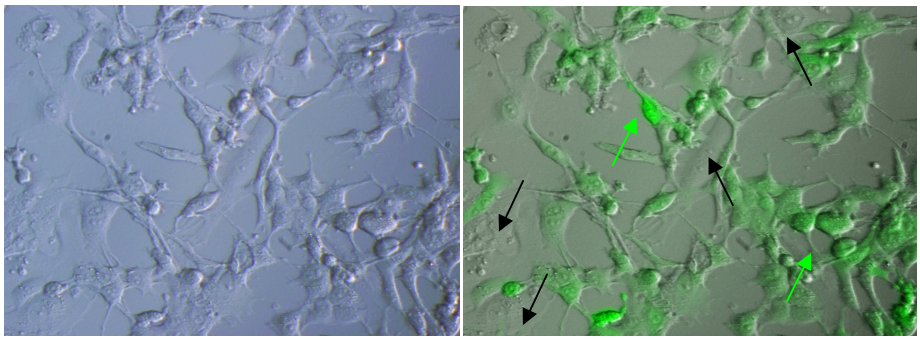


Figure 4.28 A *Phase Contrast microscopy image of CT26-GFP co-cultured with MEF;* **B** *GFP image of co-cultured CT26-GFP (green arrow) with MEF (black arrow). Both two cells of embryonic different origin share the same mesenchymal morphology*.

4.16 Tracking local invasion and lymph-node metastasis of GFP-transfected CT26 cells in an orthotopic murine model of EMT colorectal cancer

For our study we maximized the resemblance of the human disease by orthotopic rectal inoculation of the mesenchymal CRC murine cell-line CT26, GFP-transfected, in syngenic Balb/c mice.
All the five mice injected with CT26-GFP cells developed rectal adenocarcinoma of about 0.5 cm of dimension. After 28 days from the trans-anal injection, mice were sacrificed and tumors collected, together with liver, kidneys and the mesenteric lymph-nodes. Hematoxylin-eosin histological analysis of the tumor samples showed a poorly differentiated adenocarcinoma growing in the sub-mucosa (**Figure 4.29**), invading the muscular layer, and extending to the perivisceral adipose tissue.

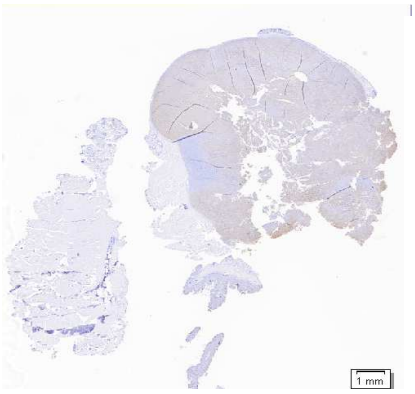


Figure 4.29 *Rappresentative CT26 Tumor mass growing in the sub-mucosa of mouse rectum, stained for GFP (by immunohystochemistry)*

By immuno-histochemical staining for GFP, the local invasive properties of mesenchymal cancer cells was magnified (**Figure 4.30**), revealing the presence of GFP+ CT26 cells in the adipose tissue.
In addition elongated GFP-positive neoplastic cells were identified in the stromal compartment (**Figure 4.30**) recapitulating the same location of Twist1+ cells in human malignancies.

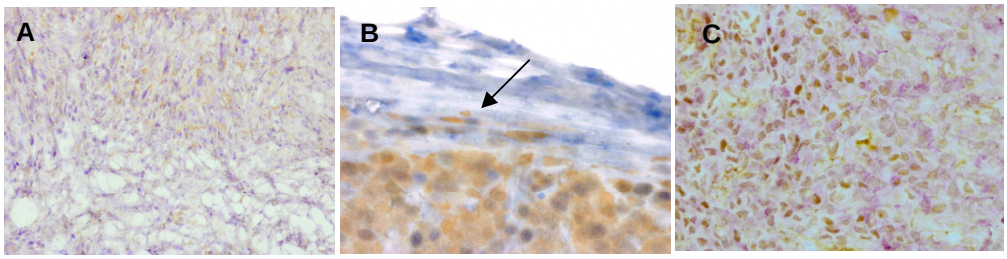


Figure 4.30 A *CT26-GFP cells (Brown) invading the adipose tissue (20x objectives);* **B** *CT26-GFP with an elongated cytoplasm and nucleus (black arrow) in the tumor stromal compartment (40x objective),* **C** *CT26 GFP+ (purple) expressing Twist1 (Brown)*

In accordance with their *in vitro* features, CT26-GFP cells maintained Twist1 expression while growing and invading nearby tissues *in vivo (***Figure 4.30***)*.
One of the six lymph-nodes collected displayed an altered morphology. By immunohistochemistry neoplastic GFP immunopositive cells as shown in **Figure 4.31** were localized within the deranged lymph-node. At 28th day none of the distant organs were affected by neoplastic colonization.

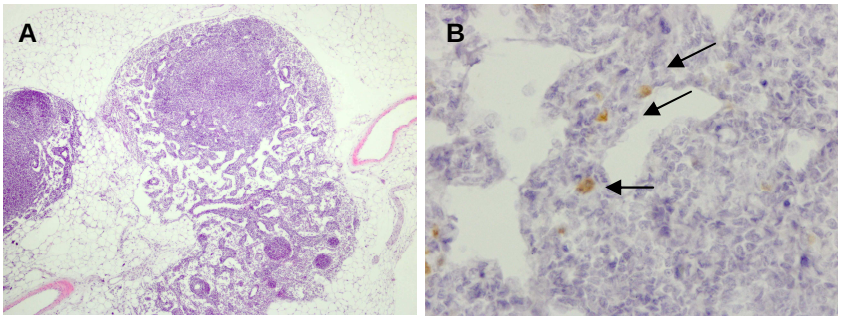


Figure 4.31 ***A****, Haematoxilin-eosin staining of a deranged mesenteric lymph-node;* ***B*** *CT26-GFP+ cells in the deranged mesenteric lymph-node (black arrow), (by immunohistochemistry)*

5. DISCUSSION and CONCLUSIONS

Metastatic progression is the main cause of death in CRC cancer patients. Metastasis is the final point of a multistep-process in which neoplastic cells become able first to invade the surrounding tissues (invasion), second to reach and trespass into vessels (intravasation), third to travel in the blood-stream, to come out from the endothelium (extravasation), to reach and colonize a secondary organ (colonization), and finally to growth therein as a distant metastasis (metastasis).

The successful ending of each step is necessary to achieve the dissemination of tumor cells from the primary site to a distant organ. Today several models imply that neoplastic epithelial cells are able to cross all these events by acquiring a new morphology and phenotype through the early developmental mechanism of epithelial to mesenchymal transition (EMT)[8,247,248,250,251,253,310]. EMT is a mechanistic approach to explain a phenomena derived from murine model, but till now only postulated in human carcinogenesis. The term EMT refers to a complex molecular and cellular programme by which epithelial cells shed their differentiated characteristics, including cell adhesion, planar and apical-basal polarity, and lack of motility, to acquire mesenchymal features, including motility, invasiveness and a heightened resistance to apoptosis[6,11,318,256]. The EMT inducers are a class of pleyotropic transcription factors whose main role is the direct and indirect regulation of the E-cadherin expression coupled with the transcriptional activation of mesenchymal programs affecting proliferation, motility and survival.[258,265, 274]

Most trascriptional inducers of EMT, identified as *bona fide*; belong to the Snail, Zeb and bHLH families (SNAI1, SLUG or SNAI2, TWIST1, ZEB1, ZEB2 or SIP1, TCF3, TCF4, Goosecoid, KLF8 and FOXC2)[6,275,321,325,328,363,380]. It was observed that these transcription factors cooperate into the induction of EMT, despite their redundant/different roles in distinct cellular models and in tumors arising in different organs. Accumulating observations display that some of the EMT inducers alone are responsible of the phenotypical switch by up-regulating the other markers, as demonstrated for SNAI1 at the onset of transition induced by TGF-beta [306,321,382,406]. However, specific and non redundant roles for each transcription factor in a hierarchical EMT functional cascade remain to be identified, as it remains their association to a potentially stabilized EMT state. Finding of an EMT stable cell-line to study the network between the major transcription factor regulators of the EMT, is fundamental to our better understanding of the background of epithelial to mesenchymal-transition. In fact, the limitations of the full comprehension of the EMT process is mainly due to the lack of an *in vitro* model of a natural and stable EMT state . Currently, all the experimental approaches, as recapitulated by Moreno-Bueno et Coll. [438], rely on the study of A) cell-culture systems of monolayer-3D cultured cells in a transient induced status of EMT or B) of cells in a stable state of EMT promoted by the ectopic expression of one of the EMT transcription factors. Clearly, both these approaches cannot fully replicate what happens *in vivo*.

First section: Cellular and functional studies. With our work, to our knowledge, we identify and characterize for the first time a mesenchymal CRC cell-line arising from an epithelial tumor, CoLo741, which is in a stable state of EMT. Moreover our model, as cultured in basal condition, contrast the vision that EMT could be only induced by interactions between cancer cells and environmental elements, and

support the hypothesis that also EMT arises as a manifestation of genetic or epigenetic alterations, other than of a paracrine or autocrine stimuli.
Human CRC cancer cells CoLo741 cells reflect the theoretical model of EMT by switching all the epithelial markers in mesenchymal ones and retaining a spindle-like morphology, more similar to a fibroblast than to an epithelial tumor cell line. Colo741 cells lost the ability to form spheroid structure, as they prefer to detach each other, rather adhering to plates, as done by fibroblasts. According to the concept of cancer EMT as a transposition of the embryiogenic program of differentiation, resulting in the conversion of epithelial cells to mesenchymal derivatives during development and adulthood, appears that CoLo741 cells have reached the final/stable step of the transdifferentiation as in the organs development.
Comparing Colo741 to other 16 epithelial CRC cells we can approach the modelling of the activation of the EMT program according to their gene-expression profile. While EMT is till now considered as activated by the expression of one of the previous described transcription factors, here we report that most CRC cell-lines expressed at least one of these transcription factors, without changing their epithelial expression profile and morphology.
In our cell panel, reaching of the mesenchymal state occurs only by the simultaneously expression of at least 4 of the transcriptional regulator (Twist1, Zeb1, Sip1 and Slug). Till now the proposed EMT hierarchy supposes that SNAI as master regulator and inducer of the other factors. Our results in Colo741 re-appraise this model, since these cells did not express SNAI1.
Moreover SNAI1 was one of the genes more frequently expressed among by the other epithelial cell-lines. Our systematic assessment of known EMT markers in CRC cell-lines, shows that is not merely the qualitative expression of a given gene, but rather the sum of a set of transcriptional regulator to determine the revertion to a mesenchymal phenotype. Obviously we think that further studies are necessary to better explain the inter-relationships of each transcription factor within the EMT program, as well as their pleyotropic effects. Additionally, coming back to our cellular model we report that the gene expression profile appears to be different from the other cells. In fact, several genes directly or indirectly modulated by the EMT transcription factors, but absolutely related to this state, were up/down expressed only mirroring biological functions strictly related to the embryogenic development (i.e. ALX1-RUNX3-FOXD2 expression), to colon cancer metastasis (i.e. SPARC, MMP1), and improved resistance to cell-death (i.e. BCL2 and BCL2A1 expression).
Despite our study is not extended to the evaluation of the post-trascriptional activation of SNAI1 protein, the absence of the transcript in Colo741, as paradigm of EMT, is a proof of evidence of the inefficacy or of the secondary role of SNAI1 as main inducer of the EMT switch-on, a reason why our studies were focused on the transcriptional regulator Twist1, which can act independently of SNAI1 to repress E-cadherin [6] and to up-regulate N-cadherin [410] .
Twist1 was for the first time identified as master regulator of EMT in a murine model of Breast cancer [6]. Its role in cancer metastatic progression was later demonstrated in several malignancies [410,424,325].
Starting from the gene expression profile we confirmed by qRT-PCR its significant expression in Colo741 and HCT116. These cell lines represent two models of different genetic instability in CRC. Colo741 is a prototypic cells with chromosomal

instability (CIN or microsatellite stable, MSS) while HCT116 of microsatellite instability (MSI). As stated, a mesenchymal phenotype was restricted to Colo741, despite Twist1 mRNA expression was seen in HCT116 too. Evaluation of Twist1 protein revealed that its expression is restricted to Colo741.
The discrepancy between gene expression and the absence of Twist1 and associated mesenchymal markers in HCT116 is mainly related to the different genetic and epigenetic mutational pattern of MSI CRC with respect to their MSS/CIN counterpart.
From the epigenetic point of view we demonstrate that a partial methilation of Twist1 affected HCT116. Moreover , MSI CRC due to the lack of mismatch repair system accumulate frameshift mutation inside genes with a short repetitive tract, resulting in unstable RNA transcripts, targeted for degradation through the nonsense mediated decay (NMD) machinery.
In our study we found a shortened 3'-UTR in Twist1 is associated with an unstable transcript, by a mechanism involved in the regulation of translation of mRNA into protein in MSI cancer cells. In line with these data, instability of *Twist1* transcript in HCT116 resulted in a lower level (2-fold) of protein even when we reproduced the same Twist1 mRNA level of CoLo741 by DNA de-methilation.
Currently, we propose the impaired Twist1 expression in HCT116 as a paradigm of inefficient EMT in MSI cancers.
Besides our discovery about epigenetic and genetic regulation of Twist1 expression, several studies reported the activation of Twist1 by microenviromental stimuli and hypoxic conditions. Cytokines such as trasforming growth factor beta and tumor necrosis factor are considered environmental inducers of EMT.
With this work we demonstrate that a short stimulation with Hypoxia and TGF-beta increase Twist1 expression only in Colo741, but not in other cell-lines indicating that *in vitro* only clones prone to EMT are able activate the trascriptional machinery in response to microenvironment conditions. In other words, only cells with mesenchymal features are able to increase the EMT transcriptional pattern. Moreover other tested citokynes (such as IL-6, IL-1b, TNF-alfa and IFN-g) are not involved in the activation or enhancing of EMT in our CRC models.
Taking together these results is conceivable that *in vivo* EMT is a phenomena restricted only to some clone which shared a mixture of epigenetic, genetic condition with an inclination to respond to external stimuli. In this view, not all the cells were able to undergone EMT.
With respect to the MSI cells HCT116, our data are in agreement with those by Pino et al. In fact, besides our original observation concerning Twist1 DNA methylation and mRNA instability, *TGFBRII* frameshift mutation leads to the un-responsiveness towards its ligand TGFbeta [439].
One of the central observations of our study is that, once we ectopically activated Twist1, its *de novo* expression in the context of an epithelial cell, such as SW480, was able, to enhance migratory and invasive capability, together with a partial switch to the molecular mesenchymal phenotype, as testified by the up-regulation of N-cadherin and Vimentin. At the same time we demonstrated that its ablation in mesenchymal cell Colo741 re-established an epithelial morphology, by directly affecting the expression of the epithelial cadherin. We also found that Twist1 ablation induced the up-regulation of Zeb2 gene, likely a compensatory mechanism, but despite this increase migratory and invasive capability were at least 2-fold reduced as compared to the control. This experimental evidence

suggests that Twist1 expression is a signature of the mesenchymal activation, that its expression is necessary to maintain, or to successfully induce, the transition, and that its role principally affects the migratory and invasive capability of CRC cells. Ideally, acquisition of migratory capability should advantage Twist1 expressing cells in dissociating and disseminating in the surrounding tissues, the prerequisite for complete at least the invasion step along the metastatic process.
In fact when we analyze CRC specimens according to Twist1 protein expression we found not only the presence of immune positive cells in the tumor gland but also in the stromal compartment. Stromal Twist1 positives, share the expression of mesenchymal markers and are mainly located at the margin of degenerate tumor glands underlying degradation of basal membrane and indicating cells that can migrate away from the epithelial layer in which they originate.
Second section: Human Tissues studies. Our study is the first that evaluates Twist1 expression in CRC, from transcript to protein expression, according to the genetic instability of CRC.
Genetic instability, may influence the prognosis. The clinical behaviour and enhanced survival benefit of MSI tumors is distinctive and not attributable to differences in therapeutic response. The molecular basis for the prognostic advantage due to the MSI status is not clearly established. Indeed it is proved that MSI cancers have a lower rate of metastastic progression compared to the MSS counterpart. For this reason appeared of interest to study the EMT, as mechanism of metastatic spread, in MSI and MSS CRC, taking advantage of Twist1 signature for the mesenchymal switch.
The first analysis performed by qRT-PCR in 78 CRC tissues compared to their normal mucosa, revealed that Twist1 expression was significantly higher in tumor samples and as our in vitro model showed (CoLo741 and HCT116), that Twist1 was expressed at transcript level by both MSS and MSI CRC specimens.
Higher mRNA levels of MSS CRC were associated with more local invasive tumors and, as previously reported by Valdes-Mora Twist1 mRNA is significantly associated with nodal and distant metastasis [15].
Other associations with clinical and pathological features were also found by studying Twist1 protein expression. As previously described in cellular models, mRNA was not sufficient to indicate the presence and efficacy of Twist1 protein. Considering tumor specimens as complex tissues, composed by many cells of different clonal origin, we found expression of Twist1 protein in different compartments. Clonality of Twist1 positive cells in the stroma does not reflect an epithelial origin as they express exclusively mesenchymal markers (Vimentin and α-SMA), but not FSP-1 marker and Desmin. Thus they are un-likely to represent a phenotypic product of fibroblasts, as expected to originate by the EMT that occur during the development of fibrosis in chronic inflammation (i.e., these cells are not expression of a so-called type 2-EMT).
Twist1 immunoreactivty in stroma may reflect an activated compartment. Therefore, Twist1 positive cells either contribute to create a better microenviroment for the progression of colon cancer, either reflect the presence of neoplastic cells invading nearby tissues. These possibilities are not mutually exclusive and probably represent different stages of CRC evolution. In fact Twist1 stromal expression in MSS CRC was associated with local invasion, nodal and distant metastasis. Accordingly, even the rate of Twist1 positive cases in the stromal compartment was significantly higher in MSS than in MSI cancers. This lower rate

of stromal expression goes hand in hand with a better survival, and a similar survival of MSS patients which were negative for Twist1 protein and of those with MS CRC was detected. Due to the roles that Twist1 plays in CRC cells as to our findings, the association of its expression with a poor disease specific survival and disease free survival in MSS CRC fits with the current EMT model. In line with the model, Twist1 low expression rate in MSI CRC, due to different concomitant genetic and epigenetic alterations, would help to explain the low metastatic potential of these subset of cancers.

It remains to be clarified whether Twist1 positive cells in the stroma compartment are tumor cells. This central point till now remains largely postulated, for this like for the other EMT main players (e.g, SNA1, ZEB1). None of the studies regarding EMT in human malignancies described or identify the clonal origin of stromal cells expressing EMT markers. Snai1 protein at tumor-stroma interface was associated with activated mesenchymal cells promoting the conversion of carcinoma cells to stromal cells [13]. The same was postulated for Twist1 stromal expression without give any experimental evidence [14].

The lack of a suitable methodology to track the EMT neoplastic cells in the tumor tissues or in blood samples is currently the main and strongest criticism concerning the true existence of cancer EMT in humans.

Coupled to our previous results, the cornerstone of our findings is that by the combination of immunohistochemistry and molecular cytogenetic (iFISH for CEP7) we can track stromal Twist1 immunopositive cells as of neoplastic origin. Twist1 positive cells could be hide by the acquired morphology more similar to a fibroblast than an epithelial CRC cells. We have a snap-shot of the EMT status in a precise spatial and temporary context.

Here we demonstrate with a preliminary study that at least one of the stromal cells expressing Twist1 shared the same chromosomal instabilty, unbalance of chromosome 7, of the tumor gland cells and also of our mesenchymal CRC model CoLo741. Chromosome 7 gain is a well documented phenomena associated with liver metastasis formation [72].

It remains to be more precisely established the role of these cells, according to EMT model as it has been proposed so far. To the best of our experimental evidences, we can suppose that epithelial cancer cells expressing Twist1 play a crucial role in the local tumor invasion. In fact, the enhanced migratory capability conferred by Twist1 expression could help these cells to complete invasion as first step of the metastatic process. At the same time the chance of finding these cells in the stroma is likely correlated to the number of Twist1 positive cells in this compartment. As a general notion, the larger the number of Twist1 stromal cells, the higher the probability to identify neoplastic cells undergone to EMT and prone to metastasize. This reasoning would be in agreement with our associative evidence demonstrating the association between stromal Twist1 immunopositivity and poor survival in MSS CRC.

Third Section: Animal model. In parallel our EMT murine model recapitulates the same behaviour. Once identified a mouse cell lines sharing the same EMT transcriptional background of mesenchymal CoLo741, this mesenchymal EMT cell once transplanted into BalB/c mouse invaded the surrounding tissues and the stroma compartment showing an elongated fibroblast-like morphology. As in human malignancies this EMT model of neoplastic cells gave nodal infiltration and (in the literature) liver metastasis[440].

In **summary**, we demonstrate through Twist1 mesenchymal signature the presence of EMT neoplastic cells in the stroma of CRC tumors and we give a method to identify cells undergoing to EMT able to invade surrounding tissues. Altogether our results readdress the study of the stromal compartment from that of a mere recipient of non-neoplastic cells to an incubator of cancer EMT cells able to disrupt the surrounding tissues mimicking activated fibroblast. This finding is the experimental demonstration of postulated phenomena regarding the presence of EMT cells in human malignancies.

Accordingly, Twist1 stromal expression in CRC as a signature of activated neoplastic EMT program correlates with cancer progression and poor patients' survival. These findings re-address the study of tumor specimens from a classic epithelial view point to a mesenchymal and stromal one. From a clinical perspective, targeting Twist1 neoplastic cells in the stroma might give the chance to reduce or inhibit the metastatic process in human CRC, and to do that as first approach we provide a suitable murine EMT model to test the pharmacological interferences with this mechanism.

References

1. Laghi L, Bianchi P, Miranda E, Balladore E, Pacetti V, Grizzi F, Allavena P, Torri V, Repici A, Santoro A, Mantovani A, Roncalli M, Malesci A. CD3+ cells at the invasive margin of deeply invading (pT3-T4) colorectal cancer and risk of post-surgical metastasis: a longitudinal study.Lancet Oncol. 2009;10(9):877-84.
2. Murdoch C, Muthana M, Coffelt SB, Lewis CE. The role of myeloid cells in the promotion of tumour angiogenesis. Nat Rev Cancer 2008; 618-31.
3. Ishiguro K, Yoshida T, Yagishita H, Numata Y, Okayasu T. Epithelial and stromal genetic instability contributes to genesis of colorectal adenomas. Gut. 2006 May;55(5):695-702.
4. Bardi G, Johansson B, Pandis N, Heim S, Mandahl N, Andrén-Sandberg A, Hägerstrand I, Mitelman F.Trisomy 7 in short-term cultures of colorectal adenocarcinomas.Genes Chromosomes Cancer. 1991 Mar;3(2):149-52
5. Patocs A, Zhang L, Xu Y, Weber F, Caldes T, Mutter GL, Platzer P, Eng C. Breast-cancer stromal cells with TP53 mutations and nodal metastases. N Engl J Med. 2007 Dec 20;357(25):2543-51.
6. Yang, J., et al. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. Cell. 117:927–939.
7. Yang MH, Hsu DS, Wang HW, Wang HJ, Lan HY, Yang WH, Huang CH, Kao SY, Tzeng CH, Tai SK, Chang SY, Lee OK, Wu KJ. Bmi1 is essential in Twist1-induced epithelial-mesenchymal transition. Nat Cell Biol. 2010 Oct;12(10):982-92.
8. Micalizzi DS, Farabaugh SM, Ford HL. Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression.J Mammary Gland Biol Neoplasia. 2010 Jun;15(2):117-34
9. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB, Evans KW, Hollier BG, Ram PT, Lander ES, Rosen JM, Weinberg RA, Mani SA. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes.Proc Natl Acad Sci U S A. 2010 Aug 31;107(35):15449-54.
10. Schmalhofer O, Brabletz S, Brabletz T. E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer. Cancer Metastasis Rev. 2009 Jun;28(1-2):151-66.
11. Yang, J., and Weinberg, R.A. 2008. Epithelialmesenchymal transition: at the crossroads of development and tumor metastasis. Dev. Cell. 14:818–829.
12. Soini Y, Tuhkanen H, Sironen R, Virtanen I, Kataja V, Auvinen P, Mannermaa A, Kosma VM. Transcription factors zeb1, twist and snai1 in breast carcinoma.BMC Cancer. 2011 Feb 16;11:73.
13. Francí C, Gallén M, Alameda F, Baró T, Iglesias M, Virtanen I, García de Herreros A Snail1 protein in the stroma as a new putative prognosis marker for colon tumours. PLoS One. 2009;4(5):e5595.
14. Valdés-Mora F, Gómez del Pulgar T, Bandrés E, Cejas P, Ramírez de Molina A, Pérez-Palacios R, Gallego-Ortega D, García-Cabezas MA, Casado E, Larrauri J, Nistal M, González-Barón M, García-Foncillas J, Lacal JC. TWIST1 overexpression is associated with nodal invasion and

male sex in primary colorectal cancer. Ann Surg Oncol. 2009 Jan;16(1):78-87.
15. Morson B. President's address: the polyp-cancer sequence in the large bowel. Proc R Soc Med 1974;67:451–457.
16. Fearon E, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759–767.
17. Markowitz S, Wang J, Myeroff L, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability.Science 1995;268:1336–1338.
18. Thiagalingam S, Lengauer C, Leach FS, et al. Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. Nat Genet 1996;13:343–346.
19. Samuels Y, Velculescu VE. Oncogenic mutations of PIK3CA in human cancers. Cell Cycle 2004;3:1221–1224.
20. Baker S, Fearon E, Nigro J, et al. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. Science 1989; 244:217–221.
21. Wood LD, Parsons DW, Jones S, et al. The genomic landscapes of human breast and colorectal cancers. Science 2007;318: 1108–1113.
22. Leary R, Lin J, Cummins J, et al. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. Proc Natl Acad Sci U S A 2008;105:16224–16229.
23. Haigis KM, Kendall KR, Wang Y, et al. Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. Nat Genet 2008;40:600–608.
24. Albertini R, Nicklas J, O'Neill J, et al. In vivo somatic mutations in humans: measurement and analysis. Annu Rev Genet 1990; 24:305–326.
25. Loeb LA, Loeb KR, Anderson JP. Multiple mutations and cancer. Proc Natl Acad Sci U S A 2003;100:776–781.
26. Lengauer C, Kinzler K, Vogelstein B. Genetic instabilities in human cancers. Nature 1998;396:643–649.
27. Rowan A, Halford S, Gaasenbeek M, et al. Refining molecular analysis in the pathways of colorectal carcinogenesis. Clin Gastroenterol Hepatol 2005;3:1115–1123.
28. Weber JC, Meyer N, Pencreach E, et al. Allelotyping analyses of synchronous primary and metastasis CIN colon cancers identified different subtypes. Int J Cancer 2007;120:524–532.
29. Cheng YW, Pincas H, Bacolod MD, et al. CpG island methylator phenotype associates with low-degree chromosomal abnormalities in colorectal cancer. Clin Cancer Res 2008;14:6005–6013.
30. Toyota M, Ahuja N, Ohe-Toyota M, et al. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999; 96:8981–8986.
31. Sinicrope F, Rego R, Halling K, et al. Prognostic impact of microsatellite instability and DNA ploidy in human colon carcinoma patients. Gastroenterology 2006;131:729–737.
32. Shen L, Toyota M, Kondo Y, et al. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci U S A 2007;104:18654–18659.
33. Boveri T. Zur Frage der Entstehung Maligner Tumoren. Jena, Germany: Gustav Fisher, 1914:1–64.

34. Herzog F, Primorac I, Dube P, et al. Structure of the anaphasepromoting complex/cyclosome interacting with a mitotic checkpoint complex. Science 2009;323:1477–1481.
35. Cahill D, Lengauer C, Yu J, et al. Mutations of mitotic checkpoint genes in human cancers. Nature 1998;392:300–303.
36. Bardelli A, Cahill D, Lederer G, et al. Carcinogen-specific induction of genetic instability. Proc Natl Acad Sci U S A 2001;98:
37. Li Y, Benezra R. Identification of a human mitotic checkpoint gene: hsMAD2. Science 1996;274:246–248.
38. Jin D, Spencer F, Jeang K. Human T cell leukemia virus type 1 oncoprotein Tax targets the human mitotic checkpoint protein MAD1. Cell 1998;93:81–91.
39. Wang Z, Cummins JM, Shen D, et al. Three classes of genes mutated in colorectal cancers with chromosomal instability. Cancer Res 2004;64:2998–3001.
40. Weaver BA, Silk AD, Montagna C, et al. Aneuploidy acts both oncogenically and as a tumor suppressor. Cancer Cell 2007; 11:25–36.
41. Ganem N, Godinho S, Pellman D. A mechanism linking extra centrosomes to chromosomal instability. Nature 2009;460: 278–282.
42. Basto R, Brunk K, Vinadogrova T, et al. Centrosome amplification can initiate tumorigenesis in flies. Cell 2008;133:1032–1042.
43. Anand S, Penrhyn-Lowe S, Venkitaraman AR. AURORA-A amplification overrides the mitotic spindle assembly checkpoint, inducing esistance to Taxol. Cancer Cell 2003;3:51–62.
44. Ewart-Toland A, Briassouli P, de Koning JP. Identification ofStk6/STK15 as a candidate low-penetrance tumor-susceptibility gene in mouse and human. Nat Genet 2003;34:403–412.
45. Katayama H, Ota T, Jisaki F, et al. Mitotic kinase expression and colorectal cancer progression. J Natl Cancer Inst 1999;91: 1160–1162.
46. Macurek L, Lindqvist A, Lim D. Polo-like kinase-1 is activated by aurora A to promote checkpoint recovery. Nature 2008;455: 119–123.
47. Seki A, Coppinger JA, Jang CY, et al. Bora and the kinase Aurora A cooperatively activate the kinase Plk1 and control mitotic entry. Science 2008;320:1655–1658.
48. Takahashi T, Sano B, Nagata T, et al. Polo-like kinase 1 (PLK1) is overexpressed in primary colorectal cancers. Cancer Sci 2003;94:148–152.
49. O'Hagan RC, Chang S, Maser RS, et al. Telomere dysfunction provokes regional amplification and deletion in cancer genomes. Cancer Cell 2002;2:149–155.
50. Rudolph K, Millard M, Bosenberg M, et al. Telomere dysfunction and evolution of intestinal carcinoma in mice and humans. Nat Genet 2001;28:155–159.
51. Plentz R, Wiemann S, Flemming P, et al. Telomere shortening of epithelial cells characterises the adenoma-carcinoma transition of human colorectal cancer. Gut 2003;52:1304–1307.
52. Engelhardt M, Drullinsky P, Guillem J, et al. Telomerase and telomere length in the development and progression of premalignant lesions to colorectal cancer. Clin Cancer Res 1997;3: 1931–1941.

53. Takagi S, Kinouchi Y, Hiwatashi N, et al. Telomere shortening and the clinicopathologic characteristics of human colorectal carcinomas. Cancer 1999;86:1431–1436.
54. Katayama S, Shiota G, Oshimura M, et al. Clinical usefulness of telomerase activity and telomere length in the preoperative diagnosis of gastric and colorectal cancer. J Cancer Res Clin Oncol 1999;125:405–410.
55. Nakamura K, Furugori E, Esaki Y, et al. Correlation of telomere lengths in normal and cancers tissue in the large bowel. Cancer Lett 2000;158:179–184.
56. Gertler R, Rosenberg R, Stricker D, et al. Telomere length and human telomerase reverse transcriptase expression as markers for progression and prognosis of colorectal carcinoma. J Clin Oncol 2004;22:1807–1814.
57. Chadeneau C, Hay K, Hirte H, et al. Telomerase activity associated with acquisition of malignancy in human colorectal cancer. Cancer Res 1995;55:2533–2536.
58. Tatsumoto N, Hiyama E, Murakami Y, et al. High telomerase activity is an independent prognostic indicator of poor outcome in colorectal cancer. Clin Cancer Res 2000;6:2696–2701.
59. Khanna K, Jackson S. DNA double-strand breaks: signaling, repair and the cancer connection. Nat Genet 2001;27:247–254.
60. Bassing CH, Suh H, Ferguson DO, et al. Histone H2AX: a dosage-dependent suppressor of oncogenic translocations and tumors. Cell 2003;114:359–370.
61. Celeste A, Difilippantonio S, Difilippantonio MJ, et al. H2AX haploinsufficiency modifies genomic stability and tumor susceptibility. Cell 2003;114:371–383.
62. Zha S, Sekiguchi J, Brush JW, et al. Complementary functions of ATM and H2AX in development and suppression of genomic instability. Proc Natl Acad Sci U S A 2008;105:9302–9306.
63. Peddibhotla S, Lam MH, Gonzalez-Rimbau M, et al. The DNAdamage effector checkpoint kinase 1 is essential for chromosome segregation and cytokinesis. Proc Natl Acad Sci U S A 2009;106:5159–5164.
64. H. Rajagopalan, M. A. Nowak, B. Vogelstein, and C.Lengauer, “The significance of unstable chromosomes incolorectal cancer,” *Nature Reviews Cancer*, vol. 3, no. 9, pp.695–701, 2003.
65. B. Dutrillaux, “Recent data on the cytogenetics of colorectal adenocarcinoma,” *Bulletin du Cancer*, vol. 75, no. 6, pp. 509–516, 1988.
66. P. Duesberg, R. Li, A. Fabarius, and R. Hehlmann, “Aneuploidy and cancer: from correlation to causation,” *Contributions to Microbiology*, vol. 13, pp. 16–44, 2006.
67. H. Fensterer, B. Radlwimmer, J. Str¨ater et al., “Matrixcomparative genomic hybridization from multicenter formalin-fixed paraffin-embedded colorectal cancer tissue blocks,” *BMC Cancer*, vol. 7, article 58, 2007.
68. S. D. Markowitz and M. M. Bertagnolli, “Molecular origins of cancer: molecular basis of colorectal cancer,” *New England Journal of Medicine*, vol. 361, no. 25, pp. 2449–2460, 2009.
69. M. Hermsen, C. Postma, J. Baak et al., “Colorectal adenoma to carcinoma progression follows multiple pathways of chromosomal instability,” *Gastroenterology*, vol. 123, no. 4,pp. 1109–1119, 2002.

70. E. R. Fearon, K. R. Cho, J. M. Nigro et al., “Identification of a chromosome 18q gene that is altered in colorectal cancers,”*Science*, vol. 247, no. 4938, pp. 49–56, 1990.
71. W. M. Korn, T. Yasutake,W. L. Kuo et al., “Chromosome arm 20q gains and other genomic alterations in colorectal cancer metastatic to liver, as analyzed by comparative genomic hybridization and fluorescence in situ hybridization,” *Genes Chromosomes and Cancer*, vol. 25, no. 2, pp. 82–90, 1999.
72. P. De Angelis, O. P. F. Clausen, A. Schjølberg, and T. Stokke, “Chromosomal gains and losses in primary colorectal carcinomas detected by CGH and their associations with tumour DNA ploidy, genotypes and phenotypes,” *British Journal of Cancer*, vol. 80, no. 3-4, pp. 526–535, 1999.
73. T. Knutsen, H. M. Padilla-Nash, D. Wangsa et al., “Definitive molecular cytogenetic characterization of 15 colorectal cancer cell lines,” *Genes Chromosomes and Cancer*, vol. 49, no. 3, pp. 204–223, 2010.
74. E. Martinez-Lopez, A. Abad, A. Font et al., “Allelic loss on chromosome 18q as a prognosticmarker in stage II colorectal cancer,” *Gastroenterology*, vol. 114, no. 6, pp. 1180–1187, 1998.
75. S. N. Thibodeau, G. Bren, and D. Schaid, “Microsatellite instability in cancer of the proximal colon,” *Science*, vol. 260, no. 5109, pp. 816–819, 1993.
76. Thiagalingam S, Laken S, Willson J, et al. Mechanisms underlying losses of heterozygosity in human colorectal cancers. Proc Natl Acad Sci USA 2001;98:2698–2702.
77. Powell S, Zilz N, Beazer-Barclay Y, et al. APC mutations occur early during colorectal tumorigenesis. Nature 1992;359:235–237.
78. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. Cell 1996;87:159–170.
79. Otori K, Konishi M, Sugiyama K, et al. Infrequent somatic mutation of the adenomatous polyposis coli gene in aberrant crypt foci of human colon tissue. Cancer 1998;83:896–900.
80. Miyaki M, Konishi M, Kikuchi-Yanoshita R, et al. Characteristics of somatic mutation of the adenomatous polyposis coli gene in colorectal tumors. Cancer Res 1994;54:3011–3020.
81. Cottrell S, Bicknell D, Kaklamanis L, et al. Molecular analysis of APC mutations in familial adenomatous polyposis and sporadic colon carcinomas. Lancet 1992;340:626–630.
82. Miyoshi Y, Nagase H, Ando H, et al. Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. Hum Mol Genet 1992;1:229–233.
83. Esteller M, Sparks A, Toyota M, et al. Analysis of adenomatous polyposis coli promoter hypermethylation in human cancer. Cancer Res 2000;60:4366–4371.
84. Rubinfeld B, Albert I, Porfiri E, et al. Binding of GSK3b to the APC-b-catenin complex and regulation of complex assembly. Science 1996;272:1023–1026.

85. Mann B, Gelos M, Siedow A, et al. Target genes of beta-catenin-T-cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. Proc Natl Acad Sci U S A 1999;96:1603–1608.
86. Sparks A, Morin P, Vogelstein B, et al. Mutational analysis of the APC/beta-catenin/Tcf pathway in colorectal cancer. Cancer Res 1998;58:1130–1134.
87. Samowitz W, Powers M, Spirio L, et al. Beta-catenin mutations are more frequent in small colorectal adenomas than in larger adenomas and invasive carcinomas. Cancer Res 1999;59: 1442–1444.
88. Shimizu Y, Ikeda S, Fujimori M, et al. Frequent alterations in the Wnt signaling pathway in colorectal cancer with microsatellite instability. Genes Chromosomes Cancer 2002;33:73–81.
89. Liu W, Dong X, Mai M, et al. Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating _-catenin/ TCF signalling. Nat Genet 2000;26:146–147.
90. Santini D, Loupakis F, Vincenzi B, et al. High concordance of KRAS status between primary colorectal tumors and related metastatic sites: implications for clinical practice. Oncologist 2008;13:1270–1275.
91. Takayama T, Ohi M, Hayashi T, et al. Analysis of K-ras, APC, and beta-catenin in aberrant crypt foci in sporadic adenoma, cancer, and familial adenomatous polyposis. Gastroenterology 2001;121:599–611.
92. Pretlow TP, Pretlow TG. Mutant KRAS in aberrant crypt foci (ACF): initiation of colorectal cancer? Biochim Biophys Acta 2005;1756:83–96.
93. Pruitt K, Der C. Ras and Rho regulation of the cell cycle and oncogenesis. Cancer Lett 2001;171:1–10.
94. Samuels Y, Wang Z, Bardelli A, et al. High frequency of mutations of the PIK3CA gene in human cancers. Science 2004;304: 554.
95. Vivanco I, Sawyers C. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. Nature Rev Cancer 2002;2:498–501.
96. Hennessy B, Smith D, Ram P, et al. Exploiting the PI3K/AKT pathway for cancer drug discovery. Nature Rev Drug Discov 2005;4:988–1004.
97. Malliri A, van der Kammen R, Clark K, et al. Mice deficient in the Rac activator Tiam1 are resistant to Ras-induced skin tumours. Nature 2002;417:867–871.
98. González-García A, Pritchard C, Paterson H, et al. RalGDS is required for tumor formation in a model of skin carcinogenesis. Cancer Cell 2005;7:219–226.
99. Barbie D, Tamayo P, Boehm J, et al. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. Nature 2009;462:108–112.
100. Meylan E, Dooley A, Feldser D, et al. Requirement for NF-kappaB signalling in a mouse model of lung adenocarcinoma. Nature 2009;462:104–107
101. Vogelstein B, Lane D, Levine A. Surfing the p53 network. Nature 2000;408:307–310.
102. Levine A. p53, the cellular gatekeeper for growth and division. Cell 1997;88:323–331.
103. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. Nat Med 2004;10:789–799.

104. Menendez D, Inga A, Resnick MA. The expanding universe of p53 targets. Nat Rev Cancer 2009;9:724–737.
105. Leslie A, Carey F, Pratt N, et al. The colorectal adenoma-carcinoma sequence. Br J Surg 2002;89:845–860.
106. Béroud C, Soussi T. The UMD-p53 database: New mutations and analysis tools. Hum Mutat 2003;21:176–181.
107. Deyoung MP, Ellisen LW. p63 and p73 in human cancer: defining the network. Oncogene 2007;26:5169–5183.
108. Keino-Masu K, Masu M, Hinck L, et al. Deleted in colorectal cancer (DCC) encodes a netrin receptor. Cell 1996;87:175–185.
109. Fazeli A, Dickinson S, Hermiston M, et al. Phenotype of mice lacking functional deleted in colorectal cancer (Dcc) gene. Nature 1997;386:796–804.
110. Cho K, Oliner J, Simons J, et al. The DCC gene: structural analysis and mutations in colorectal carcinomas. Genomics 1994;19:525–531.
111. Takagi Y, Kohmura H, Futamura M, et al. Somatic alterations of the DPC4 gene in human colorectal cancers in vivo. Gastroenterology 1996;111:1369–1372.
112. Takagi Y, Koumura H, Futamura M, et al. Somatic alterations of the SMAD-2 gene in human colorectal cancers. Br J Cancer 1998;78:1152–1155.
113. Wu C, Kirley S, Xiao H, et al. Cables enhances cdk2 tyrosine 15 phosphorylation by Wee1, inhibits cell growth, and is lost in many human colon and squamous cancers. Cancer Res 2001; 61:7325–7332.
114. Park do Y, Sakamoto H, Kirley S, et al. The Cables gene on chromosome 18q is silenced by promoter hypermethylation and allelic loss in human colorectal cancer. Am J Pathol 2007;171: 1509–1519.
115. Kirley S, D'Apuzzo M, Lauwers G, et al. The Cables gene onchromosome 18Q regulates colon cancer progression in vivo. Cancer Biol Ther 2005;4:861–863.
116. Eberhart C, Coffey R, Radhika A, et al. Up-regulation of cyclooxygenase 2 gene expression in human colorectal adenomas and adenocarcinomas. Gastroenterology 1994;107:1183–1188.
117. Oshima M, Murai N, Kargman S, et al. Chemoprevention of intestinal polyposis in the Apcdelta716 mouse by rofecoxib, a specific cyclooxygenase-2 inhibitor. Cancer Res 2001;61:1733–1740.
118. Pugh S, Thomas G. Patients with adenomatous polyps and carcinomas have increased colonic mucosal prostaglandin E2. Gut 1994;35:675–678.
119. Rigas B, Goldman I, Levine L. Altered eicosanoid levels in human colon cancer. J Lab Clin Med 1993;122:518–523.
120. Greenhough A, Smartt HJ, Moore AE, et al. The COX-2/PGE2 pathway: key roles in the hallmarks of cancer and adaptation to the tumour microenvironment. Carcinogenesis 2009;30:377–386.
121. Fukuda R, Kelly B, Semenza GL. Vascular endothelial growth factor gene expression in colon cancer cells exposed to prostaglandin E2 is mediated by hypoxia-inducible factor 1. Cancer Res 2003;63:2330–2334.
122. Williams C, Tsujii M, Reese J, et al. Host cyclooxygenase-2 modulates carcinoma growth. J Clin Invest 2000;105:1589–1594.

123. Bardi G, Parada L, Bomme L, et al. Cytogenetic comparisons of synchronous carcinomas and polyps in patients with colorectal cancer. Br J Cancer 1997;76:765–769.
124. Stoler D, Chen N, Basik M, et al. The onset and extent of genomic instability in sporadic colorectal tumor progression.Proc Natl Acad Sci U S A. 1999;96:15121–15126.
125. Shih I, Zhou W, Goodman S, et al. Evidence that genetic instability occurs at an early stage of colorectal tumorigenesis. Cancer Res 2001;61:818–822.
126. Cardoso J, Molenaar L, de Menezes R, et al. Chromosomal instability in MYH- and APC-mutant adenomatous polyps. CancerRes 2006;66:2514–2519.
127. Nowak M, Komarova N, Sengupta A, et al. The role of chromosomalinstability in tumor initiation. Proc Natl Acad Sci U S A 2002;99:16226–16231.
128. Vogelstein B, Fearon ER, Kern SE, et al. Allelotype of colorectal carcinomas. Science 1989;244:207–211.
129. Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 1993;363:558–561.
130. Blake C, Tsao JL, Wu A, et al. Stepwise deletions of polyA sequences in mismatch repair-deficient colorectal cancers. Am J Pathol 2001;158:1867–1870.
131. Peltomaki P, Lothe RA, Aaltonen LA, et al. Microsatellite instability is associated with tumors that characterize the hereditary non-polyposis colorectal carcinoma syndrome. Cancer Res 1993; 53:5853–5855.
132. Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. Science 1993;260:812–816.
133. Grilley M, Holmes J, Yashar B, et al. Mechanisms of DNA-mismatch correction. Mutat Res 1990;236:253–267.
134. Culligan KM, Meyer-Gauen G, Lyons-Weiler J, et al. Evolutionary origin, diversification and specialization of eukaryotic MutS homolog mismatch repair proteins. Nucleic Acids Res 2000;28:463–471.
135. Acharya S, Wilson T, Gradia S, et al. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proc Natl Acad Sci U S A 1996;93:13629–13634.
136. Aquilina G, Hess P, Branch P, et al. A mismatch recognition defect in colon carcinoma confers DNA microsatellite instability and a mutator phenotype. Proc Natl Acad Sci U S A 1994;91:8905–8909.
137. Palombo F, Gallinari P, Iaccarino I, et al. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. Science 1995;268:1912–1914.
138. Strand M, Earley MC, Crouse GF, et al. Mutations in the MSH3 gene preferentially lead to deletions within tracts of simple repetitive DNA in Saccharomyces cerevisiae. Proc Natl Acad Sci U S A 1995;92:10418–10421.
139. Kunkel TA, Erie DA. DNA mismatch repair. Annu Rev Biochem 2005;74:681–710.

140. Acharya S, Foster PL, Brooks P, et al. The coordinated functions of the E. coli MutS and MutL proteins in mismatch repair. Mol Cell 2003;12:233–246.
141. Jascur T, Boland CR. Structure and function of the components of the human DNA mismatch repair system. Int J Cancer 2006;119: 2030–2035.
142. Boland CR, Fishel R. Lynch syndrome: form, function, proteins, and basketball. Gastroenterology 2005;129:751–755.
143. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 1993;75:1027–1038.
144. Leach FS, Nicolaides NC, Papadopoulos N, et al. Mutations of mutS homolog in hereditary nonpolyposis colorectal cancer. Cell 1993;75:1215–1225.
145. Parsons R, Li GM, Longley MJ, et al. Hypermutability and mismatch repair deficiency in RER tumor cells. Cell 1993;75: 1227–1236.
146. Lynch HT, Smyrk TC, Watson P, et al. Genetics, natural history, tumor spectrum, and pathology of hereditary nonpolyposis colorectal cancer: an updated review. Gastroenterology 1993;104: 1535–1549.
147. Lindblom A, Tannergard P, Werelius B, et al. Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. Nat Genet 1993;5:279–282.
148. Bronner CE, Baker SM, Morrison PT, et al. Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. Nature 1994;368:258–261.
149. Papadopoulos N, Nicolaides NC, Wei YF, et al. Mutation of a mutL homolog in hereditary colon cancer. Science 1994;263:1625–1629.
150. Nicolaides NC, Papadopoulos N, Liu B, et al. Mutations of two PMS homologues in hereditary nonpolyposis colon cancer. Nature 1994;371:75–80.
151. Miyaki M, Konishi M, Tanaka K, et al. Germline mutation of MSH6 as the cause of hereditary nonpolyposis colorectal cancer. Nat Genet 1997;17:271–272.
152. Kolodner RD, Tytell JD, Schmeits JL, et al. Germ-line msh6 mutations in colorectal cancer families. Cancer Res 1999;59: 5068–5074.
153. Vasen HF. Clinical description of the Lynch syndrome [hereditary nonpolyposis colorectal cancer (HNPCC)]. Fam Cancer 2005;4: 219–225.
154. Watson P, Vasen HF, Mecklin JP, et al. The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. Int J Cancer 2008;123:444–449. Vasen HF, Mecklin JP, Khan PM, et al. The International Collaborative Group on Hereditary Non-Polyposis Colorectal Cancer (ICG-HNPCC). Dis Colon Rectum 1991;34:424–425.
155. Vasen HF, Watson P, Mecklin JP, et al. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. Gastroenterology 1999;116:1453–1456.
156. Boland CR. Evolution of the nomenclature for the hereditary colorectal cancer syndromes. Fam Cancer 2005;4:211–218.
157. Peltomaki P. Lynch syndrome genes. Fam Cancer 2005;4:227–232.

158. Lipkin SM, Wang V, Jacoby R, et al. MLH3: a DNA mismatch repair gene associated with mammalian microsatellite instability. Nat Genet 2000;24:27–35.
159. Wu Y, Berends MJ, Post JG, et al. Germline mutations of EXO1gene in patients with hereditary nonpolyposis colorectal cancer (HNPCC) and atypical HNPCC forms. Gastroenterology 2001;120:1580–1587.
160. Alam NA, Gorman P, Jaeger EE, et al. Germline deletions of EXO1 do not cause colorectal tumors and lesions which are null for EXO1 do not have microsatellite instability. Cancer Genet Cytogenet 2003;147:121–127.
161. Lindor NM, Rabe K, Petersen GM, et al. Lower cancer incidence in Amsterdam-I criteria families without mismatch repair deficiency: familial colorectal cancer type X. JAMA 2005;293:1979–1985.
162. Llor X, Pons E, Xicola RM, et al. Differential features of colorectal cancers fulfilling Amsterdam criteria without involvement of the mutator pathway. Clin Cancer Res 2005;11:7304–7310.
163. Sellner LN, Taylor GR. MLPA and MAPH: new techniques for detection of gene deletions. Hum Mutat 2004;23:413–419.
164. Wijnen J, van der Klift H, Vasen H, et al. MSH2 genomic deletions are a frequent cause of HNPCC. Nat Genet 1998;20:326–328.
165. van der Klift H, Wijnen J, Wagner A, et al. Molecular characterization of the spectrum of genomic deletions in the mismatch repair genes MSH2, MLH1, MSH6, and PMS2 responsible forhereditary nonpolyposis colorectal cancer (HNPCC). Genes Chromosomes Cancer 2005;44:123–138.
166. Raevaara TE, Korhonen MK, Lohi H, et al. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. Gastroenterology 2005;129:537–549.
167. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 2005;23:609–618.
168. Ward R, Meagher A, Tomlinson I, et al. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. Gut 2001;48:821–829.
169. Kane MF, Loda M, Gaida GM, et al. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Res 1997;57:808–811.
170. Veigl ML, Kasturi L, Olechnowicz J, et al. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998;95: 8698–8702.
171. Wang L, Cunningham JM, Winters JL, et al. BRAF mutations in colon cancer are not likely attributable to defective DNA mismatch repair. Cancer Res 2003;63:5209–5212.
172. Sinicrope FA, Rego RL, Halling KC, et al. Prognostic impact of microsatellite instability and DNA ploidy in human colon carcinoma patients. Gastroenterology 2006;131:729–737.
173. Kakar S, Burgart LJ, Thibodeau SN, et al. Frequency of loss of hMLH1 expression in colorectal carcinoma increases with advancing age. Cancer 2003;97:1421–1427.
174. Toyota M, Ahuja N, Ohe-Toyota M, et al. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A 1999; 96:8681–8686.

175. Issa JP. CpG island methylator phenotype in cancer. Nat Rev Cancer 2004;4:988–993.
176. Kondo Y, Issa JP. Epigenetic changes in colorectal cancer. Cancer Metastasis Rev 2004;23:29–39.
177. Issa JP, Baylin SB, Belinsky SA. Methylation of the estrogen receptor CpG island in lung tumors is related to the specific type of carcinogen exposure. Cancer Res 1996;56:3655–3658.
178. Issa JP, Ahuja N, Toyota M, et al. Accelerated age-related CpG island methylation in ulcerative colitis. Cancer Res 2001;61: 3573–3577.
179. Goel A, Nagasaka T, Arnold CN, et al. The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer. Gastroenterology 2007;132:127–138.
180. Ogino S, Nosho K, Kirkner GJ, et al. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. Gut 2009;58:90–96.
181. Deng G, Nguyen A, Tanaka H, et al. Regional hypermethylation and global hypomethylation are associated with altered chromatin conformation and histone acetylation in colorectal cancer. Int J Cancer 2006;118:2999–3005.
182. Nagasaka T, Koi M, Kloor M, et al. Mutations in both KRAS and BRAF may contribute to the methylator phenotype in colon cancer. Gastroenterology 2008;134:1950–1960.
183. Boland CR, Shin SK, Goel A. Promoter methylation in the genesis of gastrointestinal cancer. Yonsei Med J 2009;50:309–321.
184. Gryfe R, Swallow C, Bapat B, et al. Molecular biology of colorectal cancer. Curr Probl Cancer 1997;21:233–300.
185. Gazzoli I, Loda M, Garber J, et al. A hereditary nonpolyposis colorectal carcinoma case associated with hypermethylation of the MLH1 gene in normal tissue and loss of heterozygosity of the unmethylated allele in the resulting microsatellite instabilityhigh tumor. Cancer Res 2002;62:3925–3928.
186. Suter CM, Martin DI, Ward RL. Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet 2004;36:497–501.
187. Hitchins M, Williams R, Cheong K, et al. MLH1 germline epimutations as a factor in hereditary nonpolyposis colorectal cancer. Gastroenterology 2005;129:1392–1399.
188. Chan TL, Yuen ST, Kong CK, et al. Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat Genet 2006;38:1178–1183.
189. Ligtenberg MJ, Kuiper RP, Chan TL, et al. Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3= exons of TACSTD1. Nat Genet 2009;41:112–117.
190. Morak M, Laner A, Scholz M, et al. Report on de-novo mutation in the MSH2 gene as a rare event in hereditary nonpolyposis colorectal cancer. Eur J Gastroenterol Hepatol 2008;20:1101–1105.
191. Wang J, Sun L, Myeroff L, et al. Demonstration that mutation of the type II transforming growth factor beta receptor inactivates its tumor suppressor activity in replication error-positive colon carcinoma cells. J Biol Chem 1995;270:22044–22049.

192. Parsons R, Myeroff LL, Liu B, et al. Microsatellite instability and mutations of the transforming growth factor beta type II receptor gene in colorectal cancer. Cancer Res 1995;55:5548–5550.
193. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. Cancer Res 2002;62:2447–2454.
194. 79. Chang DK, Metzgar D, Wills C, et al. Microsatellites in the eukaryotic DNA mismatch repair genes as modulators of evolutionary mutation rate. Genome Res 2001;11:1145–1146.
195. Perucho M. Microsatellite instability: the mutator that mutates the other mutator. Nat Med 1996;2:630–631.
196. Mirabelli-Primdahl L, Gryfe R, Kim H, et al. Beta-catenin mutations are specific for colorectal carcinomas with microsatellite instability but occur in endometrial carcinomas irrespective of mutator pathway. Cancer Res 1999;59:3346–3351.
197. Miyaki M, Iijima T, Kimura J, et al. Frequent mutation of betacatenin and APC genes in primary colorectal tumors from patients with hereditary nonpolyposis colorectal cancer. Cancer Res 1999;59:4506–4509.
198. Rajagopalan H, Bardelli A, Lengauer C, et al. Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status. Nature 2002; 418:934.
199. Chang CL, Marra G, Chauhan DP, et al. Oxidative stress inactivates the human DNA mismatch repair system. Am J Physiol Cell Physiol 2002;283:C148–C154.
200. Lee SH, Chang DK, Goel A, et al. Microsatellite instability and suppressed DNA repair enzyme expression in rheumatoid arthritis. J Immunol 2003;170:2214–2220.
201. Gasche C, Chang CL, Rhees J, et al. Oxidative stress increases frameshift mutations in human colorectal cancer cells. Cancer Res 2001;61:7444–7448.
202. Hofseth LJ, Khan MA, Ambrose M, et al. The adaptive imbalance in base excision-repair enzymes generates microsatellite instability in chronic inflammation. J Clin Invest 2003;112:1887–1894.
203. Suzuki H, Harpaz N, Tarmin L, et al. Microsatellite instability in ulcerative colitis-associated colorectal dysplasias and cancers. Cancer Res 1994;54:4841–4844.
204. Brentnall TA, Crispin DA, Bronner MP, et al. Microsatellite instability in nonneoplastic mucosa from patients with chronic ulcerative colitis. Cancer Res 1996;56:1237–1240.
205. Lyda MH, Noffsinger A, Belli J, et al. Microsatellite instability and K-ras mutations in patients with ulcerative colitis. Hum Pathol 2000;31:665–671.
206. Koi M, Umar A, Chauhan DP, et al. Human chromosome 3 corrects mismatch repair deficiency and microsatellite instability and reduces N-methyl-N=-nitro-N-nitrosoguanidine tolerance incolon tumor cells with homozygous hMLH1 mutation. Cancer Res 1994;54:4308–4312.
207. Hawn MT, Umar A, Carethers JM, et al. Evidence for a connection between the mismatch repair system and the G2 cell cycle checkpoint. Cancer Res 1995;55:3721–3725.

208. Carethers JM, Hawn MT, Chauhan DP, et al. Competency in mismatch repair prohibits clonal expansion of cancer cells treated with N-methyl-N=-nitro-N-nitrosoguanidine. J Clin Invest 1996;98:199–206.
209. Carethers JM, Chauhan DP, Fink D, et al. Mismatch repair proficiency and in vitro response to 5-fluorouracil. Gastroenterology 1999;117:123–131.
210. Aebi S, Kurdi-Haidar B, Gordon R, et al. Loss of DNA mismatch repair in acquired resistance to cisplatin. Cancer Res 1996;56:3087–3090.
211. Fink D, Nebel S, Aebi S, et al. The role of DNA mismatch repair in platinum drug resistance. Cancer Res 1996;56:4881–4886.
212. Haugen AC, Goel A, Yamada K, et al. Genetic instability caused by loss of MutS homologue 3 in human colorectal cancer. Cancer Res 2008;68:8465–8472.
213. Umar A, Koi M, Risinger JI, et al. Correction of hypermutability, N-methyl-N=-nitro-N-nitrosoguanidine resistance, and defective DNA mismatch repair by introducing chromosome 2 into human tumor cells with mutations in MSH2 and MSH6. Cancer Res 1997;57:3949–3955.
214. Watanabe Y, Haugen-Strano A, Umar A, et al. Complementation of an hMSH2 defect in human colorectal carcinoma cells by human chromosome 2 transfer. Mol Carcinog 2000;29:37–49.
215. Boland CR, Koi M, Chang DK, et al. The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch Syndrome: from bench to bedside. Fam Cancer 2008;7:41–752.
216. Chang DK, Ricciardiello L, Goel A, et al. Steady-state regulation of the human DNA mismatch repair system. J Biol Chem 2000; 275:29178.
217. Arnold CN, Goel A, Boland CR. Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines. Int J Cancer 2003;106:66–73.
218. Gasche C, Chang CL, Natarajan L, et al. Identification of frameshift intermediate mutant cells. Proc Natl Acad Sci U S A 2003;100:1914–1919.
219. Reitmair AH, Schmits R, Ewel A, et al. MSH2 deficient mice are viable and susceptible to lymphoid tumours. Nat Genet 1995; 11:64–70.
220. de WN, Dekker M, Berns A, et al. Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. Cell 1995;82:321–330.
221. Edelmann W, Cohen PE, Kane M, et al. Meiotic pachytene arrest in MLH1-deficient mice. Cell 1996;85:1125–1134.
222. Yao X, Buermeyer AB, Narayanan L, et al. Different mutator phenotypes in Mlh1- versus Pms2-deficient mice. Proc Natl Acad Sci U S A 1999;96:6850–6855.
223. Edelmann W, Yang K, Umar A, et al. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. Cell 1997;91:467–477.
224. Reitmair AH, Cai JC, Bjerknes M, et al. MSH2 deficiency contributes to accelerated APC-mediated intestinal tumorigenesis. Cancer Res 1996;56:2922–2926.
225. Boland CR, Thibodeau SN, Hamilton SR, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248–5257.

226. Dietmaier W, Wallinger S, Bocker T, et al. Diagnostic microsatellite instability: definition and correlation with mismatch repair protein expression. Cancer Res 1997;57:4749–4756.
227. Laiho P, Launonen V, Lahermo P, et al. Low-level microsatellite instability in most colorectal carcinomas. Cancer Res 2002;62: 1166–1170.
228. Halford SE, Sawyer EJ, Lambros MB, et al. MSI-low, a real phenomenon which varies in frequency among cancer types. J Pathol 2003;201:389–394.
229. Ahrendt SA, Decker PA, Doffek K, et al. Microsatellite instability at selected tetranucleotide repeats is associated with p53 mutations in non-small cell lung cancer. Cancer Res 2000;60: 2488–2491.
230. Jass JR, Young J, Leggett BA. Biological significance of microsatellite instability-low (MSI-L) status in colorectal tumors. Am J Pathol 2001;158:779–781.
231. Sankila R, Aaltonen LA, Jarvinen HJ, et al. Better survival rates in patients with MLH1-associated hereditary colorectal cancer. Gastroenterology 1996;110:682–687.
232. Watson P, Lin KM, Rodriguez-Bigas MA, et al. Colorectal carcinoma survival among hereditary nonpolyposis colorectal carcinoma family members. Cancer 1998;83:259–266.
233. Aarnio M, Mustonen H, Mecklin JP, et al. Prognosis of colorectal cancer varies in different high-risk conditions. Ann Med 1998; 30:75–80.
234. Malesci A, Laghi L, Bianchi P, et al. Reduced likelihood of metastases in patients with microsatellite-unstable colorectal cancer. Clin Cancer Res 2007;13:3831–3839.
235. Gryfe R, Kim H, Hsieh ET, et al. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N Engl J Med 2000;342:69–77.
236. Phillips SM, Banerjea A, Feakins R, et al. Tumour-infiltrating lymphocytes in colorectal cancer with microsatellite instability are activated and cytotoxic. Br J Surg 2004;91:469–475.
237. Ogino S, Nosho K, Irahara N, et al. Lymphocytic reaction to colorectal cancer is associated with longer survival, independent of lymph node count, microsatellite instability, and CpG island methylator phenotype. Clin Cancer Res 2009;15:6412–6420.
238. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. Science 2006;313:1960–1964.
239. Parc Y, Gueroult S, Mourra N, et al. Prognostic significance of microsatellite instability determined by immunohistochemical staining of MSH2 and MLH1 in sporadic T3N0M0 colon cancer. Gut 2004;53:371–375.
240. Barnetson RA, Tenesa A, Farrington SM, et al. Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer. N Engl J Med 2006;354:2751–2763.
241. Kim GP, Colangelo LH, Wieand HS, et al. Prognostic and predictive roles of high-degree microsatellite instability in colon cancer: a National Cancer Institute-National Surgical Adjuvant Breast and Bowel Project Collaborative Study. J Clin Oncol 2007;25:767–772.

242. Boland CR. Clinical uses of microsatellite instability testing in colorectal cancer: an ongoing challenge. J Clin Oncol 2007;25: 754–756.
243. Hendriks YM, Wagner A, Morreau H, et al. Cancer risk in hereditary nonpolyposis colorectal cancer due to MSH6 mutations: impact on counseling and surveillance. Gastroenterology 2004; 127:17–25.
244. Shia J. Immunohistochemistry versus microsatellite instability testing for screening colorectal cancer patients at risk for hereditary nonpolyposis colorectal cancer syndrome. Part I. The utility of immunohistochemistry. J Mol Diagn 2008;10:293–300.
245. Suraweera N, Duval A, Reperant M, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. Gastroenterology 2002; 123:1804–1811.
246. Kalluri, R., and Neilson, E.G. 2003. Epithelial mesenchymal transition and its implications for fibrosis. J. Clin. Invest. 112:1776–1784.
247. Hay, E.D. 1995. An overview of epithelio-mesenchymal transformation. Acta Anat. (Basel). 154:8–20. Lipschutz, J.H. 1998. Molecular development of the kidney: a review of the results of gene disruption studies. Am. J. Kidney Dis. 31:383–397.
248. Rothenpieler, U.W., and Dressler, G.R. 1993. Pax-2 is required for mesenchyme-to-epithelium conversion during kidney development. Development. 119:711–720.
249. Hogan, B.L., and Kolodziej, P.A. 2002. Organogenesis: molecular mechanisms of tubulogenesis. Nat. Rev. Genet. 3:513–523.
250. Virchow, R. 1858. Die Cellularpathologie in ihrer Begruendung auf physiologische und pathologische Gewebelehre. August Hirschwald. Berlin, Germany.
251. Lee, J.M., Dedhar, S., Kalluri, R., and Thompson, E.W. 2006. The epithelial-mesenchymal transition: new insights in signaling, development, and disease. J. Cell Biol. 172:973–981.
252. Zeisberg, M., and Neilson, E.G. 2009. Biomarkers for epithelial-mesenchymal transitions. J. Clin. Invest. 119:1429–1437.
253. Vicovac, L., and Aplin, J.D. 1996. Epithelialmesenchymal transition during trophoblast differentiation. Acta Anat. (Basel). 156:202–216.
254. Aplin, J.D., Haigh, T., Vicovac, L., Church, H.J., and Jones, C.J. 1998. Anchorage in the developing placenta: an overlooked determinant of pregnancy outcome? Hum. Fertil. (Camb.). 1:75–79.
255. Bischof, P., et al. 2006. Implantation of the human embryo: research lines and models. From the implantation research network ‘Fruitful’. Gynecol. Obstet. Invest. 62:206–216.
256. Acloque, H., Adams, M.S., Fishwick, K., Bronner-Fraser, M., and Nieto, M.A. 2009. Epithelial- mesenchymal transitions: the importance of changing cell state in development and disease. J. Clin. Invest. 119:1438–1449.
257. Hay, E.D. 1990. Role of cell-matrix contacts in cell migration and epithelial-mesenchymal transformation. Cell Differ. Dev. 32:367–375.
258. Thiery, J.P., and Sleeman, J.P. 2006. Complex networks orchestrate epithelial-mesenchymal transitions. Nat. Rev. Mol. Cell Biol. 7:131–142.
259. Hay, E.D. 2005. The mesenchymal cell, its role in the embryo, and the remarkable signaling mechanisms that create it. Dev. Dyn. 233:706–720.

260. Skromne, I., and Stern, C.D. 2001. Interactions between Wnt and Vg1 signalling pathways initiate primitive streak formation in the chick embryo.Development. 128:2915–2927.
261. Liu, P., et al. 1999. Requirement for Wnt3 in vertebrate axis formation. Nat. Genet. 22:361–365.
262. Popperl, H., et al. 1997. Misexpression of Cwnt8C in the mouse induces an ectopic embryonic axis and causes a truncation of the anterior neuroectoderm. Development. 124:2997–3005.
263. Thomas, P., Brickman, J.M., Popperl, H., Krumlauf, R., and Beddington, R.S. 1997. Axis duplication and anterior identity in the mouse embryo. Cold Spring Harb. Symp. Quant. Biol. 62:115–125.
264. Chea, H.K., Wright, C.V., and Swalla, B.J. 2005. Nodal signaling and the evolution of deuterostome gastrulation. Dev. Dyn. 234:269–278.
265. Skromne, I., and Stern, C.D. 2002. A hierarchy of gene expression accompanying induction of the primitive streak by Vg1 in the chick embryo. Mech. Dev. 114:115–118.
266. Collignon, J., Varlet, I., and Robertson, E.J. 1996.Relationship between asymmetric nodal expression and the direction of embryonic turning. Nature. 381:155–158.
267. Varlet, I., Collignon, J., and Robertson, E.J. 1997. nodal expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. Development. 124:1033–1044.
268. Shah, S.B., et al. 1997. Misexpression of chick Vg1 in the marginal zone induces primitive streak formation. Development. 124:5127–5138.
269. Ciruna, B., and Rossant, J. 2001. FGF signaling regulates mesoderm cell fate specification and morphogenetic movement at the primitive streak. Dev. Cell. 1:37–49.
270. Rossant, J., Ciruna, B., and Partanen, J. 1997. FGF signaling in mouse gastrulation and anteroposterior patterning. Cold Spring Harb. Symp. Quant. Biol. 62:127–133.
271. Zhao, H., et al. 2006. Fibroblast growth factor receptor 1 (Fgfr1) is not essential for lens fiber differentiation in mice. Mol. Vis. 12:15–25.
272. Perea-Gomez, A., et al. 2002. Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. Dev. Cell. 3:745–756.
273. Lindsley, R.C., et al. 2008. Mesp1 coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. Cell Stem Cell. 3:55–68.
274. Arnold, S.J., Hofmann, U.K., Bikoff, E.K., and Robertson, E.J. 2008. Pivotal roles for eomesodermin during axis formation, epithelium-to-mesenchyme transition and endoderm specification in the mouse. Development. 135:501–511.
275. Nieto, M.A. 2002. The snail superfamily of zincfinger transcription factors. Nat. Rev. Mol. Cell Biol. 3:155–166.
276. Ikenouchi, J., Sasaki, H., Tsukita, S., and Furuse, M. 2008. Loss of occludin affects tricellular localization of tricellulin. Mol. Biol. Cell. 19:4687–4693.

277. Cavatorta, A.L., Giri, A.A., Banks, L., and Gardiol, D. 2008. Cloning and functional analysis of the promoter region of the human Disc large gene. Gene. 424:87–95.
278. Whiteman, E.L., Liu, C.J., Fearon, E.R., and Margolis, B. 2008. The transcription factor snail represses Crumbs3 expression and disrupts apico-basal polarity complexes. Oncogene. 27:3875–3879.
279. Duband, J.L., and Thiery, J.P. 1982. Appearance and distribution of fibronectin during chick embryo gastrulation and neurulation. Dev. Biol. 94:337–350.
280. Sauka-Spengler, T., and Bronner-Fraser, M. 2008. A gene regulatory network orchestrates neural crest formation. Nat. Rev. Mol. Cell Biol. 9:557–568.
281. Knecht, A.K., and Bronner-Fraser, M. 2002. Induction of the neural crest: a multigene process. Nat. Rev. Genet. 3:453–461.
282. Liem, K.F., Jr., Jessell, T.M., and Briscoe, J. 2000. Regulation of the neural patterning activity of sonic hedgehog by secreted BMP inhibitors expressed by notochord and somites. Development. 127:4855–4866.
283. Villanueva, S., Glavic, A., Ruiz, P., and Mayor, R. 2002. Posteriorization by FGF, Wnt, and retinoic acid is required for neural crest induction. Dev. Biol.241:289–301.
284. Karafiat, V., Dvorakova, M., Pajer, P., Cermak, V., and Dvorak, M. 2007. Melanocyte fate in neural crest is triggered by Myb proteins through activation of c-kit. Cell Mol. Life Sci. 64:2975–2984.
285. Burstyn-Cohen, T., Stanleigh, J., Sela-Donenfeld, D., and Kalcheim, C. 2004. Canonical Wnt activity regulates trunk neural crest delamination linking BMP/noggin signaling with G1/S transition. Development. 131:5327–5339.
286. Sela-Donenfeld, D., and Kalcheim, C. 2002. Localized BMP4-noggin interactions generate the dynamic patterning of noggin expression in somites. Dev. Biol. 246:311–328.
287. Thiery, J.P. 2003. Epithelial-mesenchymal transitions in development and pathologies. Curr. Opin. Cell Biol. 15:740–746.
288. Potenta, S., Zeisberg, E., and Kalluri, R. 2008. The role of endothelial-to-mesenchymal transition in cancer progression. Br. J. Cancer. 99:1375–1379.
289. Zeisberg, E.M., et al. 2007. Endothelial-tomesenchymal transition contributes to cardiac fibrosis. Nat. Med. 13:952–961.
290. Zeisberg, M., et al. 2007. Fibroblasts derive from hepatocytes in liver fibrosis via epithelial to mesenchymal transition. J. Biol. Chem. 282:23337–23347.
291. Kim, K.K., et al. 2006. Alveolar epithelial cell mesenchymal transition develops in vivo during pulmonary fibrosis and is regulated by the extracellular matrix. Proc. Natl. Acad. Sci. U. S. A. 103:13180–13185.
292. Strutz, F., et al. 1995. Identification and characterization of a fibroblast marker: FSP1. J. Cell Biol. 130:393–405.
293. Okada, H., Danoff, T.M., Kalluri, R., and Neilson, E.G. 1997. Early role of Fsp1 in epithelial-mesenchymal transformation. Am. J. Physiol. 273:F563–F574.

294. Rastaldi, M.P., et al. 2002. Epithelial-mesenchymal transition of tubular epithelial cells in human renal biopsies. Kidney Int. 62:137–146.
295. Zeisberg, M., et al. 2003. BMP-7 counteracts TGFbeta1-induced epithelial-to-mesenchymal transition and reverses chronic renal injury. Nat. Med. 9:964–968.
296. Okada, H., Strutz, F., Danoff, T.M., Kalluri, R., and Neilson, E.G. 1996. Possible mechanisms of renal fibrosis. Contrib. Nephrol. 118:147–154.
297. Zavadil, J., and Bottinger, E.P. 2005. TGF-beta and epithelial-to-mesenchymal transitions. Oncogene. 24:5764–5774.
298. Ekblom, P. 1996. Genetics of kidney development. Curr. Opin. Nephrol. Hypertens. 5:282–287.
299. Aufderheide, E., Chiquet-Ehrismann, R., and Ekblom, P. 1987. Epithelial-mesenchymal interactions in the developing kidney lead to expression of tenascin in the mesenchyme. J. Cell Biol. 105:599–608.
300. Ivanova, L., Butt, M.J., and Matsell, D.G. 2008. Mesenchymal transition in kidney collecting duct epithelial cells. Am. J. Physiol. Renal Physiol. 294:F1238–F1248.
301. Strutz, F., et al. 2002. Role of basic fibroblast growth factor-2 in epithelial-mesenchymal transformation. Kidney Int. 61:1714–1728.
302. Yang, J., et al. 2002. Disruption of tissue-type plasminogen activator gene in mice reduces renal interstitial fibrosis in obstructive nephropathy. J. Clin.Invest. 110:1525–1538.
303. Yang, J., and Liu, Y. 2002. Blockage of tubular epithelial to myofibroblast transition by hepatocyte growth factor prevents renal interstitial fibrosis.J. Am. Soc. Nephrol. 13:96–107.
304. Li, Y., Yang, J., Dai, C., Wu, C., and Liu, Y. 2003. Role for integrin-linked kinase in mediating tubular epithelial to mesenchymal transition and renal interstitial fibrogenesis. J. Clin. Invest. 112:503–516.
305. Kim, K., Lu, Z., and Hay, E.D. 2002. Direct evidence for a role of beta-catenin/LEF-1 signaling pathway in induction of EMT. Cell Biol. Int. 26:463–476.
306. Nawshad, A., Lagamba, D., Polad, A., and Hay, E.D. 2005. Transforming growth factor-beta signaling during epithelial-mesenchymal transformation: implications for embryogenesis and tumor metastasis. Cells Tissues Organs. 179:11–23.
307. Zeisberg, M., et al. 2003. Bone morphogenic protein- 7 inhibits progression of chronic renal fibrosis associated with two genetic mouse models. Am. J.Physiol. Renal Physiol. 285:F1060–F1067.
308. Bataille, F., et al. 2008. Evidence for a role of epithelial mesenchymal transition during pathogenesis of fistulae in Crohn's disease. Inflamm. Bowel Dis.14:1514–1527.
309. Hanahan, D., and Weinberg, R.A. 2000. The hallmarks of cancer. Cell. 100:57–70.
310. Thiery, J.P. 2002. Epithelial-mesenchymal transitions in tumour progression. Nat. Rev. Cancer. 2:442–454.
311. Fidler, I.J., and Poste, G. 2008. The "seed and soil" hypothesis revisited. Lancet Oncol. 9:808.

312. Brabletz, T., et al. 2001. Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. Proc. Natl. Acad. Sci. U. S. A. 98:10356–10361.
313. Zeisberg, M., Shah, A.A., and Kalluri, R. 2005. Bone morphogenic protein-7 induces mesenchymal to epithelial transition in adult renal fibroblasts and facilitates regeneration of injured kidney. J. Biol. Chem. 280:8094–8100.
314. Jechlinger, M., Grunert, S., and Beug, H. 2002. Mechanisms in epithelial plasticity and metastasis: insights from 3D cultures and expression profiling. J. Mammary Gland Biol. Neoplasia. 7:415–432.
315. Bissell, M.J., Radisky, D.C., Rizki, A., Weaver, V.M., and Petersen, O.W. 2002. The organizing principle: microenvironmental influences in the normal and malignant breast. Differentiation. 70:537–546.
316. Smit, M.A., and Peeper, D.S. 2008. Deregulating EMT and senescence: double impact by a single twist. Cancer Cell. 14:5–7.
317. Ansieau, S., et al. 2008. Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. Cancer Cell. 14:79–89.
318. Weinberg, R.A. 2008. Twisted epithelial-mesenchymal transition blocks senescence. Nat. Cell Biol.10:1021–1023.
319. Shi, Y., and Massague, J. 2003. Mechanisms of TGFbeta signaling from cell membrane to the nucleus. Cell. 113:685–700.
320. Niessen, K., et al. 2008. Slug is a direct Notch target required for initiation of cardiac cushion cellularization. J. Cell Biol. 182:315–325.
321. Medici, D., Hay, E.D., and Olsen, B.R. 2008. Snail and Slug promote epithelial-mesenchymal transition through beta-catenin-T-cell factor-4-dependent expression of transforming growth factorbeta3. Mol. Biol. Cell. 19:4875–4887.
322. 79. Kokudo, T., et al. 2008. Snail is required for TGF{beta}-induced endothelial-mesenchymal transition of embryonic stem cell-derived endothelial cells. J. Cell. Sci. 121:3317–3324.
323. Tse, J.C., and Kalluri, R. 2007. Mechanisms of metastasis: epithelial-to-mesenchymal transition and contribution of tumor microenvironment. J. Cell. Biochem. 101:816–829.
324. Gupta, P.B., Mani, S., Yang, J., Hartwell, K., and Weinberg,R.A. 2005. The evolving portrait of cancer metastasis. Cold Spring Harb. Symp. Quant. Biol. 70:291–297.
325. Yang, J., Mani, S.A., and Weinberg, R.A. 2006. Exploring a new twist on tumor metastasis. Cancer Res. 66:4549–4552.
326. Mani, S.A., et al. 2007. Mesenchyme Forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. Proc. Natl. Acad. Sci. U. S. A. 104:10069–10074.
327. Mani, S.A., et al. 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell. 133:704–715.
328. Hartwell, K.A., et al. 2006. The Spemann organizer gene, Goosecoid, promotes tumor metastasis. Proc. Natl. Acad. Sci. U. S. A. 103:18969–18974.
329. Taki, M., Verschueren, K., Yokoyama, K., Nagayama, M., and Kamata, N. 2006. Involvement of Ets-1 transcription factor in inducing matrix

metalloproteinase- 2 expression by epithelial-mesenchymal transition in human squamous carcinoma cells. Int.J. Oncol. 28:487–496.
330. Bierie, B., and Moses, H.L. 2006. Tumour microenvironment: TGFbeta: the molecular Jekyll and Hyde of cancer. Nat. Rev. Cancer. 6:506–520.
331. Oft, M., Heider, K.H., and Beug, H. 1998. TGFbeta signaling is necessary for carcinoma cell invasiveness and metastasis. Curr. Biol. 8:1243–1252.
332. Hata, A., Shi, Y., and Massague, J. 1998. TGF-beta signaling and cancer: structural and functional consequences of mutations in Smads. Mol. Med. Today 4:257–262.
333. Song, J. 2007. EMT or apoptosis: a decision for TGF-beta. Cell Res. 17:289–290.
334. Miyazono, K., ten Dijke, P., and Heldin, C.H. 2000. TGF-beta signaling by Smad proteins. Adv. Immunol. 75:115–157.
335. Derynck, R., Akhurst, R.J., and Balmain, A. 2001. TGF-beta signaling in tumor suppression and cancer progression. Nat. Genet. 29:117–129.
336. Heldin, C.H., Miyazono, K., and ten Dijke, P. 1997. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature. 390:465–471.
337. Roberts, A.B., et al. 2006. Smad3 is key to TGFbeta-mediated epithelial-to-mesenchymal transition, fibrosis, tumor suppression and metastasis. Cytokine Growth Factor Rev. 17:19–27.
338. Piek, E., Moustakas, A., Kurisaki, A., Heldin, C.H., and ten Dijke, P. 1999. TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NmuMG breast epithelial cells. J. Cell. Sci. 112:4557–4568.
339. Bhowmick, N.A., et al. 2001. Transforming growth factor-beta1 mediates epithelial to mesenchymal transdifferentiation through a RhoA-dependent mechanism. Mol. Biol. Cell. 12:27–36.
340. Saika, S., et al. 2004. Transient adenoviral gene transfer of Smad7 prevents injury-induced epithelial-mesenchymal transition of lens epithelium in mice. Lab. Invest. 84:1259–1270.
341. Yang, L., Lin, C., and Liu, Z.R. 2006. P68 RNA helicase mediates PDGF-induced epithelial mesenchymal transition by displacing Axin from beta-catenin. Cell. 127:139–155.
342. Eger, A., Stockinger, A., Schaffhauser, B., Beug, H., and Foisner, R. 2000. Epithelial mesenchymal transition by c-Fos estrogen receptor activation involves nuclear translocation of beta-catenin and upregulation of beta-catenin/lymphoid enhancer binding factor-1 transcriptional activity. J. Cell Biol. 148:173–188.
343. Stockinger, A., Eger, A., Wolf, J., Beug, H., and Foisner, R. 2001. E-cadherin regulates cell growth by modulating proliferation-dependent beta-catenin transcriptional activity. J. Cell Biol. 154:1185–1196.
344. Bhowmick, N.A., Zent, R., Ghiassi, M., McDonnell, M., and Moses, H.L. 2001. Integrin beta 1 signaling is necessary for transforming growth factor-beta activation of p38MAPK and epithelial plasticity. J. Biol. Chem. 276:46707–46713.
345. Lee, Y.H., Albig, A.R., Maryann, R., Schiemann, B.J., and Schiemann, W.P. 2008. Fibulin-5 initiates epithelial-mesenchymal transition (EMT) and

enhances EMT induced by TGF-beta in mammary epithelial cells via a MMP-dependent mechanism. Carcinogenesis. 29:2243–2251.
346. Lehmann, K., et al. 2000. Raf induces TGFbeta production while blocking its apoptotic but not invasive responses: a mechanism leading to increased malignancy in epithelial cells. Genes Dev. 14:2610–2622.
347. Gotzmann, J., et al. 2002. Hepatocytes convert to a fibroblastoid phenotype through the cooperation of TGF-beta1 and Ha-Ras: steps towards invasiveness. J. Cell. Sci. 115:1189–1202.
348. Oft, M., et al. 1996. TGF-beta1 and Ha-Ras collaborate in modulating the phenotypic plasticity and invasiveness of epithelial tumor cells. Genes Dev. 10:2462–2477.
349. Janda, E., et al. 2002. Ras and TGF[beta] cooperatively regulate epithelial cell plasticity and metastasis: dissection of Ras signaling pathways. J. Cell Biol. 156:299–313.
350. Cui, W., et al. 1996. TGFbeta1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice. Cell. 86:531–542.
351. Watanabe, T., et al. 2001. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. N. Engl. J. Med. 344:1196–1206.
352. Neil, J.R., Johnson, K.M., Nemenoff, R.A., and Schiemann, W.P. 2008. Cox-2 inactivates Smad signaling and enhances EMT stimulated by TGF-beta through a PGE2-dependent mechanisms. Carcinogenesis. 29:2227–2235.
353. Thiery, J.P., and Huang, R. 2005. Linking epithelialmesenchymal transition to the well-known polarity protein Par6. Dev. Cell. 8:456–458.
354. Tepass, U., Truong, K., Godt, D., Ikura, M., and Peifer, M. 2000. Cadherins in embryonic and neural morphogenesis. Nat. Rev. Mol. Cell Biol. 1:91–100.
355. Edelman, G.M., Gallin, W.J., Delouvee, A., Cunningham, B.A., and Thiery, J.P. 1983. Early epochal maps of two different cell adhesion molecules. Proc. Natl. Acad. Sci. U. S. A. 80:4384–4388.
356. Reichmann, E., et al. 1992. Activation of an inducible c-FosER fusion protein causes loss of epithelial polarity and triggers epithelial-fibroblastoid cell conversion. Cell. 71:1103–1116.
357. Gottardi, C.J., Wong, E., and Gumbiner, B.M. 2001. E-cadherin suppresses cellular transformation by inhibiting beta-catenin signaling in an adhesionindependent manner. J. Cell Biol. 153:1049–1060.
358. Muta, H., et al. 1996. E-cadherin gene mutations in signet ring cell carcinoma of the stomach. Jpn. J. Cancer Res. 87:843–848.
359. Saito, A., et al. 1999. Disruption of E-cadherinmediated cell adhesion systems in gastric cancers in young patients. Jpn. J. Cancer Res. 90:993–999.
360. Hirohashi, S. 1998. Inactivation of the E-cadherin mediated cell adhesion system in human cancers. Am. J. Pathol. 153:333–339.
361. Birchmeier, W., and Behrens, J. 1994. Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. Biochim. Biophys. Acta. 1198:11–26.

362. Blanco, M.J., et al. 2002. Correlation of Snail expression with histological grade and lymph node status in breast carcinomas. Oncogene. 21:3241–3246.
363. Comijn, J., et al. 2001. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. Mol. Cell. 7:1267–1278.
364. Van de Putte, T., et al. 2003. Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. Am. J. Hum. Genet. 72:465–470.
365. Yokoyama, K., et al. 2001. Reverse correlation of Ecadherin and snail expression in oral squamous cell carcinoma cells in vitro. Oral Oncol. 37:65–71.
366. Radisky, D.C., et al. 2005. Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability. Nature. 436:123–127.
367. Korpal, M., Lee, E.S., Hu, G., and Kang, Y. 2008. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. J. Biol. Chem. 283:14910–14914.
368. Park, S.M., Gaur, A.B., Lengyel, E., and Peter, M.E. 2008. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. Genes Dev. 22:894–907.
369. Gregory, P.A., et al. 2008. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat. Cell Biol. 10:593–601.
370. Paterson, E.L., et al. 2008. The microRNA-200 family regulates epithelial to mesenchymal transition. Scientific World Journal. 8:901–904.
371. Gregory, P.A., Bracken, C.P., Bert, A.G., and Goodall, G.J. 2008. MicroRNAs as regulators of epithelialmesenchymal transition. Cell Cycle. 7:3112–3118.
372. Zavadil, J., Narasimhan, M., Blumenberg, M., and Schneider, R.J. 2007. Transforming growth factorbeta and microRNA:mRNA regulatory networks in epithelial plasticity. Cells Tissues Organs. 185:157–161.
373. Xue, C., Plieth, D., Venkov, C., Xu, C., and Neilson, E.G. 2003. The gatekeeper effect of epithelial mesenchymal transition regulates the frequency of breast cancer metastasis. Cancer Res. 63:3386–3394.
374. Brabletz, T., et al. 1998. Nuclear overexpression of the oncoprotein beta-catenin in colorectal cancer is localized predominantly at the invasion front. Pathol. Res. Pract. 194:701–704.
375. Ramaswamy, S., Ross, K.N., Lander, E.S., and Golub, T.R. 2003. A molecular signature of metastasis in primary solid tumors. Nat. Genet. 33:49–54.
376. Zeisberg, E.M., Potenta, S., Xie, L., Zeisberg, M., and Kalluri, R. 2007. Discovery of endothelial to mesenchymal transition as a source for carcinoma associated fibroblasts. Cancer Res. 67:10123–10128.
377. Iwano, M., et al. 2002. Evidence that fibroblasts derive from epithelium during tissue fibrosis. J. Clin. Invest. 110:341–350.
378. Zeisberg, M., et al. 2007. Fibroblasts derive from hepatocytes in liver fibrosis via epithelial to mesenchymal transition. J. Biol. Chem. 282:23337–

23347 Okada, H., et al. 2000. Progressive renal fibrosis in murine polycystic kidney disease: an immunohistochemical observation. Kidney Int. 58:587–597.
379. Okada, H., Danoff, T.M., Kalluri, R., and Neilson, E.G. 1997. Early role of Fsp1 in epithelial- mesenchymal transformation. Am. J. Physiol. 273:F563–F574.
380. Venkov, C.D., et al. 2007. A proximal activator of transcription in epithelial-mesenchymal transition. J. Clin. Invest. 117:482–491. Barrallo-Gimeno, A., and Nieto, M.A. 2005. The Snail genes as inducers of cell movement and survival: implications in development and cancer. Development. 132:3151–3161.
381. Hay, E.D., and Zuk, A. 1995. Transformation between epithelium and mesenchyme: normal, pathological, and experimentally induced. Am. J. Kidney Dis. 26:678–690.
382. Huber, M.A., Kraut, N., and Beug, H. 2005. Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr. Opin. Cell Biol. 17:548–558.
383. Li, Y., Yang, J., Luo, J.H., Dedhar, S., and Liu, Y. 2007. Tubular epithelial cell dedifferentiation is driven by the helix-loop-helix transcriptional inhibitor Id1. J. Am. Soc. Nephrol. 18:449–460.
384. Bates, R.C., et al. 2005. Transcriptional activation of integrin beta6 during the epithelialmesenchymal transition defines a novel prognostic indicator of aggressive colon carcinoma. J. Clin. Invest. 115:339–347.
385. Davidson, L.A., Marsden, M., Keller, R., and Desimone, D.W. 2006. Integrin alpha5beta1 and fibronectin regulate polarized cell protrusions required for Xenopus convergence and extension. Curr. Biol. 16:833–844.
386. White, L.R., et al. 2007. The characterization of alpha5-integrin expression on tubular epithelium during renal injury. Am. J. Physiol. Renal Physiol. 292:F567–F576.
387. Qian, F., Zhang, Z.C., Wu, X.F., Li, Y.P., and Xu, Q 2005. Interaction between integrin alpha(5) and fibronectin is required for metastasis of B16F10 melanoma cells. Biochem. Biophys. Res. Commun. 333:1269–1275.
388. Vogel, W., Gish, G.D., Alves, F., and Pawson, T. 1997. The discoidin domain receptor tyrosine kinases are activated by collagen. Mol. Cell. 1:13–23.
389. Leitinger, B., and Kwan, A.P. 2006. The discoidin domain receptor DDR2 is a receptor for type X collagen. Matrix Biol. 25:355–364.
390. Goldsmith, E.C., et al. 2004. Organization of fibroblasts in the heart. Dev. Dyn. 230:787–794. Evtimova, V., Zeillinger, R., and Weidle, U.H. 2003.
391. Identification of genes associated with the invasive status of human mammary carcinoma cell lines by transcriptional profiling. Tumour Biol. 24:189–198.
392. Franke, W.W., Schmid, E., Osborn, M., and Weber, K. 1978. Different intermediate-sized filaments distinguished by immunofluorescence microscopy. Proc. Natl. Acad. Sci. U. S. A. 75:5034–5038.
393. Dellagi, K., Vainchenker, W., Vinci, G., Paulin, D., and Brouet, J.C. 1983. Alteration of vimentin intermediate filament expression during differentiation of human hemopoietic cells. EMBO J. 2:1509–1514.

394. Colucci-Guyon, E., et al. 1994. Mice lacking vimentin develop and reproduce without an obvious phenotype. Cell. 79:679–694.
395. Flechon, J.E., Degrouard, J., and Flechon, B. 2004. Gastrulation events in the prestreak pig embryo: ultrastructure and cell markers. Genesis. 38:13–25.
396. Witzgall, R., Brown, D., Schwarz, C., and Bonventre, J.V. 1994. Localization of proliferating cell nuclear antigen, vimentin, c-Fos, and clusterin in the postischemic kidney. Evidence for a heterogenous genetic response among nephron segments, and a large pool of mitotically active and dedifferentiated cells. J. Clin. Invest. 93:2175–2188.
397. Boyer, B., Tucker, G.C., Valles, A.M., Gavrilovic, J., and Thiery, J.P. 1989. Reversible transition towards a fibroblastic phenotype in a rat carcinoma cell line. Int. J. Cancer Suppl. 4:69–75.
398. Raymond, W.A., and Leong, A.S. 1989. Vimentin-- a new prognostic parameter in breast carcinoma? J. Pathol. 158:107–114.
399. Gabbiani, G., Kapanci, Y., Barazzone, P., and Franke, W.W. 1981. Immunochemical identification of intermediate-sized filaments in human neoplastic cells. A diagnostic aid for the surgical pathologist. Am. J. Pathol. 104:206–216.
400. Nakajima, Y., Mironov, V., Yamagishi, T., Nakamura, H., and Markwald, R.R. 1997. Expression of smooth muscle alpha-actin in mesenchymal cells during formation of avian endocardial cushion tissue: a role for transforming growth factor beta3. Dev. Dyn. 209:296–309.
401. Damonte, P., Gregg, J.P., Borowsky, A.D., Keister, B.A., and Cardiff, R.D. 2007. EMT tumorigenesis in the mouse mammary gland. Lab. Invest. 87:1218–1226.
402. Sarrio, D., et al. 2008. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. Cancer Res. 68:989–997.
403. Bienz, M. 2005. beta-Catenin: a pivot between cell adhesion and Wnt signalling. Curr. Biol. 15:R64–R67.
404. Gavert, N., and Ben-Ze'ev, A. 2007. beta-Catenin signaling in biological control and cancer. J. Cell. Biochem. 102:820–828.
405. Yook, J.I., et al. 2006. A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells. Nat. Cell Biol. 8:1398–1406.
406. Medici, D., Hay, E.D., and Goodenough, D.A. 2006. Cooperation between snail and LEF-1 transcription factors is essential for TGF-beta1-induced epithelial-mesenchymal transition. Mol. Biol. Cell. 17:1871–1879.
407. Hynes, R.O., and Yamada, K.M. 1982. Fibronectins: multifunctional modular glycoproteins. J. Cell Biol. 95:369–377.
408. Dvorak, H.F. 1986. Tumors: wounds that do not heal. Similarities between tumor stroma generation and wound healing. N. Engl. J. Med. 315:1650–1659.
409. Zeisberg, M., Strutz, F., and Muller, G.A. 2001. Renal fibrosis: an update. Curr. Opin. Nephrol. Hypertens. 10:315–320.
410. Yang, Z., et al. 2007. Up-regulation of gastric cancer cell invasion by Twist is accompanied by N-cadherin and fibronectin expression. Biochem. Biophys. Res. Commun. 358:925–930.
411. Colognato, H., and Yurchenco, P.D. 2000. Form and function: the laminin family of heterotrimers. Dev. Dyn. 218:213–234.

412. Miner, J.H., Li, C., Mudd, J.L., Go, G., and Sutherland, A.E. 2004. Compositional and structural requirements for laminin and basement membranes during mouse embryo implantation and gastrulation. Development. 131:2247–2256.
413. Zagris, N., Chung, A.E., and Stavridis, V. 2005. Entactin and laminin gamma 1-chain gene expression in the early chick embryo. Int. J. Dev. Biol. 49:65–70.
414. Zeisberg, M., Maeshima, Y., Mosterman, B., and Kalluri, R. 2002. Renal fibrosis. Extracellular matrix microenvironment regulates migratory behavior of activated tubular epithelial cells. Am. J. Pathol. 160:2001–2008.
415. Carpenter, P.M., Wang-Rodriguez, J., Chan, O.T., and Wilczynski, S.P. 2008. Laminin 5 expression in metaplastic breast carcinomas. Am. J. Surg. Pathol. 32:345–353.
416. Giannelli, G., Bergamini, C., Fransvea, E., Sgarra, C., and Antonaci, S. 2005. Laminin-5 with transforming transforming growth factor-beta1 induces epithelial to mesenchymal transition in hepatocellular carcinoma. Gastroenterology. 129:1375–1383.
417. Marinkovich, M.P. 2007. Tumour microenvironment: laminin 332 in squamous-cell carcinoma. Nat. Rev. Cancer. 7:370–380.
418. Chilosi, M., et al. 2006. Migratory marker expression in fibroblast foci of idiopathic pulmonary fibrosis. Respir. Res. 7:95.
419. Aranburu, A., Liberg, D., Honore, B., and Leanderson, T. 2006. CArG box-binding factor-A interacts with multiple motifs in immunoglobulin promoters and has a regulated subcellular distribution. Eur. J. Immunol. 36:2192–2202.
420. Boutet, A., et al. 2006. Snail activation disrupts tissue homeostasis and induces fibrosis in the adult kidney. EMBO J. 25:5603–5613.
421. Yu, W., Kamara, H., and Svoboda, K.K. 2008. The role of twist during palate development. Dev. Dyn. 237:2716–2725.
422. Castanon, I., and Baylies, M.K. 2002. A Twist in fate: evolutionary comparison of Twist structure and function. Gene. 287:11–22.
423. Kida, Y., Asahina, K., Teraoka, H., Gitelman, I., and Sato, T. 2007. Twist relates to tubular epithelialmesenchymal transition and interstitial fibrogenesis in the obstructed kidney. J. Histochem. Cytochem. 55:661–673.
424. Yang, M.H., et al. 2008. Direct regulation of TWIST by HIF-1alpha promotes metastasis. Nat. Cell Biol. 10:295–305.
425. Kume, T., Deng, K., and Hogan, B.L. 2000. Murine forkhead/winged helix genes Foxc1 (Mf1) and Foxc2 (Mfh1) are required for the early organogenesis of the kidney and urinary tract. Development. 127:1387–1395.
426. Hayashi, H., and Kume, T. 2008. Forkhead transcription factors regulate expression of the chemokine receptor CXCR4 in endothelial cells and CXCL12-induced cell migration. Biochem. Biophys. Res. Commun. 367:584–589.
427. Cederberg, A., et al. 2001. FOXC2 is a winged helix gene that counteracts obesity, hypertriglyceridemia, and diet-induced insulin resistance. Cell.106:563–573.

428. Barrallo-Gimeno, A., and Nieto, M.A. 2005. The Snail genes as inducers of cell movement and survival: implications in development and cancer.*Development.* 132:3151–3161.
429. Bourgeois, P., Stoetzel, C., Bolcato-Bellemin, A. L., Mattei, M. G., Perrin-Schmitt, F. The human H-twist gene is located at 7p21 and encodes a b-HLH protein that is 96% similar to its murine M-twist counterpart. Mammalian Genome 7: 915-917, 1996.
430. Browning, V. L., Chaudhry, S. S., Planchart, A., Dixon, M. J., Schimenti, J. C. Mutations of the mouse Twist and sy (fibrillin 2) genes induced by chemical mutagenesis of ES cells. Genomics 73: 291-298, 2001.
431. El Ghouzzi, V., Le Merrer, M., Perrin-Schmitt, F., Lajeunie, E., Benit, P., Renier, D., Bourgeois, P., Bolcato-Bellemin, A.-L., Munnich, A., Bonaventure, J. Mutations of the TWIST gene in the Saethre-Chotzen syndrome. Nature Genet. 15: 42-46, 1997.
432. Howard, T. D., Paznekas, W. A., Green, E. D., Chiang, L. C., Ma, N., Ortiz De Luna, R. I., Delgado, C. G., Gonzalez-Ramos, M., Kline, A. D., Jabs, E. W. Mutations in TWIST, a basic helix-loop-helix transcription factor, in Saethre-Chotzen syndrome. Nature Genet. 15: 36-41, 1997.
433. Mattei, M. G., Stoetzel, C., Perrin-Schmitt, F. The B-HLH protein encoding the M-twist gene is located by in situ hybridization on murine chromosome 12. Mammalian Genome 4: 127-128, 1993.
434. Pan, D., Fujimoto, M., Lopes, A., Wang, Y.-X. Twist-1 is a PPAR-delta-inducible, negative-feedback regulator of PGC-1-alpha in brown fat metabolism. Cell 137: 73-86, 2009.
435. Shishido, E., Higashijima, S., Emori, Y., Saigo, K. Two FGF-receptor homologues of Drosophila: one is expressed in mesodermal primordium in early embryos. Development 117: 751-761, 1993.
436. Hamamori, Y., Sartorelli, V., Ogryzko, V., Puri, P. L., Wu, H.-Y., Wang, J. Y. J., Nakatani, Y., Kedes, L. Regulation of histone acetyltransferases p300 and PCAF by the bHLH protein Twist and adenoviral oncoprotein E1A. Cell 96: 405-413, 1999.
437. Bourgeois, P., Bolcato-Bellemin, A.-L., Danse, J.-M., Bloch-Zupan, A., Yoshiba, K., Stoetzel, C., Perrin-Schmitt, F. The variable expressivity and incomplete penetrance of the twist-null heterozygous mouse phenotype resemble those of human Saethre-Chotzen syndrome. Hum. Molec. Genet. 7: 945-957, 1998.
438. Moreno-Bueno G, Peinado H, Molina P, Olmeda D, Cubillo E, Santos V, Palacios J, Portillo F, Cano A. The morphological and molecular features of the epithelial-to-mesenchymal transition.Nat Protoc. 2009;4(11):1591-613. Epub 2009 Oct 15.
439. Pino MS, Kikuchi H, Zeng M, Herraiz MT, Sperduti I, Berger D, Park DY, Iafrate AJ, Zukerberg LR, Chung DC. Epithelial to mesenchymal transition is impaired in colon cancer cells with microsatellite instability. Gastroenterology. 2010 Apr;138(4):1406-17.
440. Donigan M, Loh BD, Norcross LS, Li S, Williamson PR, DeJesus S, Ferrara A, Gallagher JT, Baker CH. A metastatic colon cancer model using nonoperative transanal rectal injection. Surg Endosc. 2010 Mar;24(3):642-7.

Appendix I *Gene exclusively expressed by CoLo741 (by Microarray analysis)*

Symbol	Entrez Gene Name	Location	Type(s)
ABCA5	ATP-binding cassette, sub-family A (ABC1), member 5	Plasma Membrane	transporter
ABCB5	ATP-binding cassette, sub-family B (MDR/TAP), member 5	Unknown	other
ABL2	v-abl Abelson murine leukemia viral oncogene homolog 2 (arg, Abelson-related gene)	Cytoplasm	kinase
ACP5	acid phosphatase 5, tartrate resistant	Cytoplasm	phosphatase
ADAMTS1	ADAM metallopeptidase with thrombospondin type 1 motif, 1	Extracellular Space	peptidase
ADCY1	adenylate cyclase 1 (brain)	Plasma Membrane	enzyme
AKAP12	A kinase (PRKA) anchor protein 12	Cytoplasm	transporter
AKT3	v-akt murine thymoma viral oncogene homolog 3 (protein kinase B, gamma)	Cytoplasm	kinase
ALS2CR4	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 4	Unknown	other
ALX1	ALX homeobox 1	Nucleus	transcription regulator
AMOTL1	angiomotin like 1	Plasma Membrane	other
AMPH	amphiphysin	Plasma Membrane	other
ANGPT1	angiopoietin 1	Extracellular Space	growth factor
ANTXR1	anthrax toxin receptor 1	Plasma Membrane	other
ARHGAP15	Rho GTPase activating protein 15	Unknown	other
ARHGEF9	Cdc42 guanine nucleotide exchange factor (GEF) 9	Cytoplasm	other
ASB4	ankyrin repeat and SOCS box-containing 4	Nucleus	transcription regulator
ATP6V0A4	ATPase, H+ transporting, lysosomal V0 subunit a4	Cytoplasm	transporter
ATP8B2	ATPase, class I, type 8B, member 2	Plasma Membrane	transporter
BBS5	Bardet-Biedl syndrome 5	Cytoplasm	other
BCHE	butyrylcholinesterase	Plasma Membrane	enzyme
BCL2	B-cell CLL/lymphoma 2	Cytoplasm	other
BCL2A1	BCL2-related protein A1	Cytoplasm	other
BEX4	brain expressed, X-linked 4	Cytoplasm	other
BICD1	bicaudal D homolog 1 (Drosophila)	Cytoplasm	other
BNC2	basonuclin 2	Nucleus	other
C10ORF90	chromosome 10 open reading frame 90	Unknown	other
C13ORF1	chromosome 13 open reading frame 1	Unknown	other
C14ORF132	chromosome 14 open reading frame 132	Unknown	other
C1S	complement component 1, s subcomponent	Extracellular Space	peptidase
C6ORF142	chromosome 6 open reading frame 142	Unknown	other
C7ORF38	chromosome 7 open reading frame 38	Unknown	other
C9ORF93	chromosome 9 open reading frame 93	Unknown	other
CA8	carbonic anhydrase VIII	Cytoplasm	enzyme
CADM1	cell adhesion molecule 1	Plasma Membrane	other
CAPN3	calpain 3, (p94)	Cytoplasm	peptidase
CARD16	caspase recruitment domain family, member 16	Unknown	other
CASP1	caspase 1, apoptosis-related cysteine peptidase (interleukin 1, beta, convertase)	Cytoplasm	peptidase
CCDC132	coiled-coil domain containing 132	Unknown	other
CCDC50	coiled-coil domain containing 50	Unknown	other
CCDC52	coiled-coil domain containing 52	Unknown	other

CCPG1	cell cycle progression 1	Unknown	other
CD276	CD276 molecule	Plasma Membrane	other
CDH2	cadherin 2, type 1, N-cadherin (neuronal)	Plasma Membrane	other
CDK5R1	cyclin-dependent kinase 5, regulatory subunit 1 (p35)	Nucleus	kinase
CHFR	checkpoint with forkhead and ring finger domains	Nucleus	enzyme
CHST10	carbohydrate sulfotransferase 10	Cytoplasm	enzyme
CNPY3	canopy 3 homolog (zebrafish)	Unknown	other
CNTN3	contactin 3 (plasmacytoma associated)	Plasma Membrane	other
COL4A1	collagen, type IV, alpha 1	Extracellular Space	other
COL5A2	collagen, type V, alpha 2	Extracellular Space	other
CPB2	carboxypeptidase B2 (plasma)	Extracellular Space	peptidase
CPEB1	cytoplasmic polyadenylation element binding protein 1	Cytoplasm	other
CRISPLD1	cysteine-rich secretory protein LCCL domain containing 1	Unknown	other
CSGALNACT1	chondroitin sulfate N-acetylgalactosaminyltransferase 1	Cytoplasm	enzyme
CTNNB1	catenin (cadherin-associated protein), beta 1, 88kDa	Nucleus	transcription regulator
CTSF	cathepsin F	Cytoplasm	peptidase
CTSK	cathepsin K	Cytoplasm	peptidase
CYBRD1	cytochrome b reductase 1	Cytoplasm	enzyme
D4S234E	DNA segment on chromosome 4 (unique) 234 expressed sequence	Cytoplasm	other
DAB2	disabled homolog 2, mitogen-responsive phosphoprotein (Drosophila)	Plasma Membrane	other
DCLK1	doublecortin-like kinase 1	Cytoplasm	kinase
DDR2	discoidin domain receptor tyrosine kinase 2	Plasma Membrane	kinase
DGKI	diacylglycerol kinase, iota	Cytoplasm	kinase
DNAJB9	DnaJ (Hsp40) homolog, subfamily B, member 9	Nucleus	other
DNER	delta/notch-like EGF repeat containing	Plasma Membrane	transmembrane receptor
DOK5	docking protein 5	Plasma Membrane	other
DPYD	dihydropyrimidine dehydrogenase	Cytoplasm	enzyme
DPYSL4	dihydropyrimidinase-like 4	Cytoplasm	enzyme
DSCR8	Down syndrome critical region gene 8	Unknown	other
DUSP10	dual specificity phosphatase 10	Nucleus	phosphatase
DYNC2H1	dynein, cytoplasmic 2, heavy chain 1	Cytoplasm	other
EDNRB	endothelin receptor type B	Plasma Membrane	G-protein coupled receptor
ELOVL2	elongation of very long chain fatty acids (FEN1/Elo2, SUR4/Elo3, yeast)-like 2	Cytoplasm	enzyme
EMP1	epithelial membrane protein 1	Plasma Membrane	other
ENG	endoglin	Plasma Membrane	other
ENTHD1	ENTH domain containing 1	Unknown	other
EPB41L3	erythrocyte membrane protein band 4.1-like 3	Plasma Membrane	other
EPHA4	EPH receptor A4	Plasma Membrane	kinase
EPM2A	epilepsy, progressive myoclonus type 2A, Lafora disease (laforin)	Cytoplasm	phosphatase
EVI2A	ecotropic viral integration site 2A	Plasma Membrane	transmembrane receptor
F2R	coagulation factor II (thrombin) receptor	Plasma Membrane	G-protein coupled receptor
FAH	fumarylacetoacetate hydrolase (fumarylacetoacetase)	Cytoplasm	enzyme
FAM101B	family with sequence similarity 101, member B	Unknown	other
FAM120C	family with sequence similarity 120C	Unknown	other

FAM126A	family with sequence similarity 126, member A	Cytoplasm	other
FAM70A	family with sequence similarity 70, member A	Unknown	other
FAP	fibroblast activation protein, alpha	Cytoplasm	peptidase
FBLN5	fibulin 5	Extracellular Space	other
FBXO32	F-box protein 32	Cytoplasm	enzyme
FCRLA	Fc receptor-like A	Plasma Membrane	other
FERMT3	fermitin family homolog 3 (Drosophila)	Cytoplasm	enzyme
FGD4	FYVE, RhoGEF and PH domain containing 4	Cytoplasm	other
FHL1	four and a half LIM domains 1	Cytoplasm	other
FHOD3	formin homology 2 domain containing 3	Unknown	other
FLRT2	fibronectin leucine rich transmembrane protein 2	Plasma Membrane	other
FMN1	formin 1	Nucleus	transporter
FMN2	formin 2	Unknown	other
FN1	fibronectin 1	Plasma Membrane	enzyme
FOXF2	forkhead box F2	Nucleus	transcription regulator
FSTL5	follistatin-like 5	Extracellular Space	other
GAS7	growth arrest-specific 7	Cytoplasm	transcription regulator
GCNT2	glucosaminyl (N-acetyl) transferase 2, I-branching enzyme (I blood group)	Cytoplasm	enzyme
GHR	growth hormone receptor	Plasma Membrane	transmembrane receptor
GJC1	gap junction protein, gamma 1, 45kDa	Plasma Membrane	transporter
GLI3	GLI family zinc finger 3	Nucleus	transcription regulator
GLUD1	glutamate dehydrogenase 1	Cytoplasm	enzyme
GNB4	guanine nucleotide binding protein (G protein), beta polypeptide 4	Plasma Membrane	enzyme
GNG2	guanine nucleotide binding protein (G protein), gamma 2	Plasma Membrane	enzyme
GPM6B	glycoprotein M6B	Plasma Membrane	other
GPR143	G protein-coupled receptor 143	Plasma Membrane	G-protein coupled receptor
GPR37	G protein-coupled receptor 37 (endothelin receptor type B-like)	Plasma Membrane	G-protein coupled receptor
GPR85	G protein-coupled receptor 85	Plasma Membrane	G-protein coupled receptor
GPRC5B	G protein-coupled receptor, family C, group 5, member B	Plasma Membrane	G-protein coupled receptor
GREB1	GREB1 protein	Cytoplasm	other
GTF2H1	general transcription factor IIH, polypeptide 1, 62kDa	Nucleus	transcription regulator
GYG2	glycogenin 2	Unknown	enzyme
GYPC	glycophorin C (Gerbich blood group)	Plasma Membrane	other
HEY1	hairy/enhancer-of-split related with YRPW motif 1	Nucleus	transcription regulator
HHAT	hedgehog acyltransferase	Cytoplasm	enzyme
HIST1H4B	histone cluster 1, H4b	Nucleus	other
HIVEP2	human immunodeficiency virus type I enhancer binding protein 2	Nucleus	transcription regulator
HLA-DMA	major histocompatibility complex, class II, DM alpha	Plasma Membrane	transmembrane receptor
HOXC9	homeobox C9	Nucleus	transcription regulator
HRASLS	HRAS-like suppressor	Cytoplasm	other
HRK	harakiri, BCL2 interacting protein (contains only BH3 domain)	Cytoplasm	other
HS6ST2	heparan sulfate 6-O-sulfotransferase 2	Extracellular Space	enzyme
HSD17B14	hydroxysteroid (17-beta) dehydrogenase 14	Unknown	enzyme
ICAM1	intercellular adhesion molecule 1	Plasma Membrane	transmembrane receptor

ICK	intestinal cell (MAK-like) kinase	Cytoplasm	kinase
IFFO1	intermediate filament family orphan 1	Unknown	other
IFI16	interferon, gamma-inducible protein 16	Nucleus	transcription regulator
IFI44	interferon-induced protein 44	Cytoplasm	other
IGF1R	insulin-like growth factor 1 receptor	Plasma Membrane	transmembrane receptor
IGSF11	immunoglobulin superfamily, member 11	Unknown	other
IKIP	IKK interacting protein	Cytoplasm	other
IL17D	interleukin 17D	Extracellular Space	other
IL24	interleukin 24	Extracellular Space	cytokine
IL6R	interleukin 6 receptor	Plasma Membrane	transmembrane receptor
IRF4	interferon regulatory factor 4	Nucleus	transcription regulator
JAM3	junctional adhesion molecule 3	Plasma Membrane	other
KATNAL1	katanin p60 subunit A-like 1	Unknown	enzyme
KIF13A	kinesin family member 13A	Cytoplasm	transporter
KYNU	kynureninase (L-kynurenine hydrolase)	Cytoplasm	enzyme
LAMA1	laminin, alpha 1	Extracellular Space	other
LAMA4	laminin, alpha 4	Extracellular Space	enzyme
LAMP2	lysosomal-associated membrane protein 2	Plasma Membrane	enzyme
LDOC1	leucine zipper, down-regulated in cancer 1	Nucleus	other
LHFP	lipoma HMGIC fusion partner	Unknown	other
LIN7B	lin-7 homolog B (C. elegans)	Cytoplasm	other
LINS1	lines homolog 1 (Drosophila)	Unknown	other
LOC100128202	hypothetical protein FLJ25694	Unknown	other
LONRF2	LON peptidase N-terminal domain and ring finger 2	Unknown	other
LRGUK	leucine-rich repeats and guanylate kinase domain containing	Unknown	other
LUM	lumican	Extracellular Space	other
LY96	lymphocyte antigen 96	Plasma Membrane	other
LYPD1	LY6/PLAUR domain containing 1	Plasma Membrane	G-protein coupled receptor
LYST	lysosomal trafficking regulator	Cytoplasm	transporter
LZTS1	leucine zipper, putative tumor suppressor 1	Nucleus	transcription regulator
MAFB	v-maf musculoaponeurotic fibrosarcoma oncogene homolog B (avian)	Nucleus	other
MAGEA1	melanoma antigen family A, 1 (directs expression of antigen MZ2-E)	Plasma Membrane	other
MAGEA10	melanoma antigen family A, 10	Unknown	other
MAGEC2	melanoma antigen family C, 2	Plasma Membrane	other
MAGEE1	melanoma antigen family E, 1	Plasma Membrane	other
MATN2	matrilin 2	Extracellular Space	other
MCC	mutated in colorectal cancers	Unknown	other
MCOLN3	mucolipin 3	Cytoplasm	ion channel
MCTP1	multiple C2 domains, transmembrane 1	Unknown	other
MCTP2	multiple C2 domains, transmembrane 2	Unknown	other
MDFIC	MyoD family inhibitor domain containing	Nucleus	other
MDFIC	MyoD family inhibitor domain containing	Nucleus	other
MEF2A	myocyte enhancer factor 2A	Nucleus	transcription regulator
MEF2D	myocyte enhancer factor 2D	Nucleus	transcription regulator

MGC9913	hypothetical protein MGC9913	Unknown	other
MLANA	melan-A	Plasma Membrane	other
MMP1	matrix metallopeptidase 1 (interstitial collagenase)	Extracellular Space	peptidase
MOXD1	monooxygenase, DBH-like 1	Cytoplasm	enzyme
MSRB3	methionine sulfoxide reductase B3	Cytoplasm	enzyme
MTMR9	myotubularin related protein 9	Unknown	phosphatase
MYLK	myosin light chain kinase	Cytoplasm	kinase
MYO5A	myosin VA (heavy chain 12, myoxin)	Cytoplasm	enzyme
NAV2	neuron navigator 2	Nucleus	other
NAV3	neuron navigator 3	Unknown	other
NDN	necdin homolog (mouse)	Nucleus	other
NEDD4L	neural precursor cell expressed, developmentally down-regulated 4-like	Cytoplasm	enzyme
NEFL	neurofilament, light polypeptide	Cytoplasm	other
NLGN1	neuroligin 1	Plasma Membrane	enzyme
NNMT	nicotinamide N-methyltransferase	Cytoplasm	enzyme
NOV	nephroblastoma overexpressed gene	Extracellular Space	growth factor
NOX4	NADPH oxidase 4	Cytoplasm	enzyme
NR3C1	nuclear receptor subfamily 3, group C, member 1 (glucocorticoid receptor)	Nucleus	ligand-dependent nuclear receptor
NR4A2	nuclear receptor subfamily 4, group A, member 2	Nucleus	ligand-dependent nuclear receptor
NR4A3	nuclear receptor subfamily 4, group A, member 3	Nucleus	ligand-dependent nuclear receptor
NRCAM	neuronal cell adhesion molecule	Plasma Membrane	other
OLFML2A	olfactomedin-like 2A	Unknown	other
OLIG2	oligodendrocyte lineage transcription factor 2	Nucleus	transcription regulator
OSTM1	osteopetrosis associated transmembrane protein 1	Plasma Membrane	other
PAG1	phosphoprotein associated with glycosphingolipid microdomains 1	Plasma Membrane	other
PAX3	paired box 3	Nucleus	transcription regulator
PCDH20	protocadherin 20	Unknown	other
PCDHB5	protocadherin beta 5	Plasma Membrane	other
PCLO	piccolo (presynaptic cytomatrix protein)	Cytoplasm	transporter
PCTK2	PCTAIRE protein kinase 2	Cytoplasm	kinase
PDGFD	platelet derived growth factor D	Extracellular Space	growth factor
PDK4	pyruvate dehydrogenase kinase, isozyme 4	Cytoplasm	kinase
PDZRN3	PDZ domain containing ring finger 3	Extracellular Space	other
PELO	pelota homolog (Drosophila)	Nucleus	other
PFTK1	PFTAIRE protein kinase 1	Nucleus	kinase
PGBD1	piggyBac transposable element derived 1	Unknown	enzyme
PHACTR1	phosphatase and actin regulator 1	Cytoplasm	other
PKIA	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	Unknown	other
PKNOX2	PBX/knotted 1 homeobox 2	Nucleus	other
PLD1	phospholipase D1, phosphatidylcholine-specific	Cytoplasm	enzyme
PLEKHA8	pleckstrin homology domain containing, family A member 8	Unknown	other
PLEKHO1	pleckstrin homology domain containing, family O member 1	Plasma Membrane	other
PLSCR4	phospholipid scramblase 4	Plasma Membrane	enzyme
PLTP	phospholipid transfer protein	Extracellular Space	other

PLXNC1	plexin C1	Plasma Membrane	other
PNMA2	paraneoplastic antigen MA2	Nucleus	transporter
POU3F2	POU class 3 homeobox 2	Nucleus	transcription regulator
POU3F2	POU class 3 homeobox 2	Nucleus	transcription regulator
PPM1E	protein phosphatase 1E (PP2C domain containing)	Nucleus	phosphatase
PPP2R3A	protein phosphatase 2 (formerly 2A), regulatory subunit B'', alpha	Nucleus	phosphatase
PREX1	phosphatidylinositol-3,4,5-trisphosphate-dependent Rac exchange factor 1	Cytoplasm	other
PRKACB	protein kinase, cAMP-dependent, catalytic, beta	Cytoplasm	kinase
PRKD1	protein kinase D1	Cytoplasm	kinase
PRUNE2	prune homolog 2 (Drosophila)	Unknown	other
PTPRM	protein tyrosine phosphatase, receptor type, M	Plasma Membrane	phosphatase
PTPRM	protein tyrosine phosphatase, receptor type, M	Plasma Membrane	phosphatase
QKI	quaking homolog, KH domain RNA binding (mouse)	Nucleus	other
RAB32	RAB32, member RAS oncogene family	Cytoplasm	other
RAB38	RAB38, member RAS oncogene family	Cytoplasm	enzyme
RHOBTB1	Rho-related BTB domain containing 1	Unknown	enzyme
RHOJ	ras homolog gene family, member J	Cytoplasm	enzyme
RHOQ	ras homolog gene family, member Q	Plasma Membrane	enzyme
RIMKLB	ribosomal modification protein rimK-like family member B	Unknown	other
RNF182	ring finger protein 182	Unknown	other
ROPN1B	ropporin, rhophilin associated protein 1B	Cytoplasm	other
RRP15	ribosomal RNA processing 15 homolog (S. cerevisiae)	Nucleus	other
RUNX3	runt-related transcription factor 3	Nucleus	transcription regulator
S100B	S100 calcium binding protein B	Cytoplasm	other
SAMD9L	sterile alpha motif domain containing 9-like	Unknown	other
SCFD2	sec1 family domain containing 2	Unknown	transporter
SEC22A	SEC22 vesicle trafficking protein homolog A (S. cerevisiae)	Cytoplasm	transporter
SEC22C	SEC22 vesicle trafficking protein homolog C (S. cerevisiae)	Cytoplasm	other
SEMA5A	sema domain,(semaphorin) 5A	Plasma Membrane	transmembrane receptor
SEMA6A	sema domain, (semaphorin) 6A	Plasma Membrane	other
SEMA6D	sema domain,(semaphorin) 6D	Plasma Membrane	other
SGCD	sarcoglycan, delta (35kDa dystrophin-associated glycoprotein)	Cytoplasm	other
SH3TC1	SH3 domain and tetratricopeptide repeats 1	Unknown	other
SHOX2	short stature homeobox 2	Nucleus	transcription regulator
SLAIN1	SLAIN motif family, member 1	Unknown	other
SLAMF7	SLAM family member 7	Plasma Membrane	other
SLAMF9	SLAM family member 9	Extracellular Space	other
SLC25A13	solute carrier family 25, member 13 (citrin)	Cytoplasm	transporter
SLC47A1	solute carrier family 47, member 1	Unknown	other
SLFN12	schlafen family member 12	Nucleus	enzyme
SMCR8	Smith-Magenis syndrome chromosome region, candidate 8	Unknown	other
SNAI2	snail homolog 2 (Drosophila)	Nucleus	other
SNAP25	synaptosomal-associated protein, 25kDa	Plasma Membrane	transporter
SNCA	synuclein, alpha (non A4 component of amyloid precursor)	Cytoplasm	other

SNCA	synuclein, alpha (non A4 component of amyloid precursor)	Cytoplasm	other
SNRPN	small nuclear ribonucleoprotein polypeptide N	Nucleus	other
SOX10	SRY (sex determining region Y)-box 10	Nucleus	transcription regulator
SOX6	SRY (sex determining region Y)-box 6	Nucleus	transcription regulator
SPARC	secreted protein, acidic, cysteine-rich (osteonectin)	Extracellular Space	other
SPESP1	sperm equatorial segment protein 1	Cytoplasm	other
SPG20	spastic paraplegia 20 (Troyer syndrome)	Cytoplasm	other
SPP1	secreted phosphoprotein 1	Extracellular Space	cytokine
SPRYD5	SPRY domain containing 5	Unknown	other
SRPX	sushi-repeat-containing protein, X-linked	Cytoplasm	other
SSH1	slingshot homolog 1 (Drosophila)	Cytoplasm	phosphatase
SSX2IP	synovial sarcoma, X breakpoint 2 interacting protein	Plasma Membrane	other
ST3GAL6	ST3 beta-galactoside alpha-2,3-sialyltransferase 6	Cytoplasm	enzyme
ST6GALNAC3	ST6	Cytoplasm	enzyme
STK32B	serine/threonine kinase 32B	Unknown	kinase
STX7	syntaxin 7	Plasma Membrane	transporter
SULT1C2	sulfotransferase family, cytosolic, 1C, member 2	Cytoplasm	enzyme
SYNM	synemin, intermediate filament protein	Cytoplasm	other
SYT14	synaptotagmin XIV	Unknown	transporter
TAF9B	TAF9B RNA polymerase II, TATA box binding protein (TBP)-associated factor, 31kDa	Nucleus	transcription regulator
TBX2	T-box 2	Nucleus	transcription regulator
TCHP	trichoplein, keratin filament binding	Cytoplasm	other
TIMP3	TIMP metallopeptidase inhibitor 3	Extracellular Space	other
TM4SF18	transmembrane 4 L six family member 18	Unknown	other
TMEM108	transmembrane protein 108	Unknown	other
TMOD2	tropomodulin 2 (neuronal)	Cytoplasm	other
TMSB15A	thymosin beta 15a	Cytoplasm	other
TNFAIP3	tumor necrosis factor, alpha-induced protein 3	Nucleus	other
TNFAIP3	tumor necrosis factor, alpha-induced protein 3	Nucleus	other
TNS1	tensin 1	Plasma Membrane	other
TPM3	tropomyosin 3	Cytoplasm	other
TRIB2	tribbles homolog 2 (Drosophila)	Plasma Membrane	kinase
TRIM63	tripartite motif-containing 63	Nucleus	enzyme
TRPM1	transient receptor potential cation channel, subfamily M, member 1	Plasma Membrane	ion channel
TTYH2	tweety homolog 2 (Drosophila)	Unknown	ion channel
TUSC3	tumor suppressor candidate 3	Extracellular Space	enzyme
TWISTNB	TWIST neighbor	Unknown	other
TYR	tyrosinase (oculocutaneous albinism IA)	Cytoplasm	enzyme
UGGT2	UDP-glucose glycoprotein glucosyltransferase 2	Cytoplasm	enzyme
VANGL2	vang-like 2 (van gogh, Drosophila)	Plasma Membrane	other
VEPH1	ventricular zone expressed PH domain homolog 1 (zebrafish)	Nucleus	other
VLDLR	very low density lipoprotein receptor	Plasma Membrane	transporter
VPS53	vacuolar protein sorting 53 homolog (S. cerevisiae)	Cytoplasm	other
WASF3	WAS protein family, member 3	Cytoplasm	other

WDR91	WD repeat domain 91	Unknown	other
WIPF1	WAS/WASL interacting protein family, member 1	Cytoplasm	other
ZBTB38	zinc finger and BTB domain containing 38	Nucleus	transcription regulator
ZEB2	zinc finger E-box binding homeobox 2	Nucleus	transcription regulator
ZFP82	zinc finger protein 82 homolog (mouse)	Nucleus	other
ZNF134	zinc finger protein 134	Nucleus	transcription regulator
ZNF148	zinc finger protein 148	Nucleus	transcription regulator
ZNF280B	zinc finger protein 280B	Nucleus	other
ZNF329	zinc finger protein 329	Unknown	other
ZNF331	zinc finger protein 331	Nucleus	other
ZNF347	zinc finger protein 347	Nucleus	other
ZNF354C	zinc finger protein 354C	Nucleus	other
ZNF426	zinc finger protein 426	Nucleus	other
ZNF480	zinc finger protein 480	Nucleus	other
ZNF512B	zinc finger protein 512B	Nucleus	other
ZNF542	zinc finger protein 542	Unknown	other
ZNF654	zinc finger protein 654	Unknown	other
ZNF655	zinc finger protein 655	Nucleus	other

Appendix II *Genes exclusively not expressed by CoLo741 (By Microarray analysis)*

Symbol	Entrez Gene Name	Location	Type(s)
ABLIM1	actin binding LIM protein 1	Cytoplasm	other
AKR1A1	aldo-keto reductase family 1, member A1 (aldehyde reductase)	Cytoplasm	enzyme
ANXA3	annexin A3	Cytoplasm	enzyme
ASS1	argininosuccinate synthetase 1	Cytoplasm	enzyme
C19ORF33	chromosome 19 open reading frame 33	Nucleus	other
CD24	CD24 molecule	Plasma Membrane	other
CDH1	cadherin 1, type 1, E-cadherin (epithelial)	Plasma Membrane	other
CENPV	centromere protein V	Nucleus	other
DSG2	desmoglein 2	Plasma Membrane	other
FUT4	fucosyltransferase 4 (alpha (1,3) fucosyltransferase, myeloid-specific)	Cytoplasm	enzyme
KLF4	Kruppel-like factor 4 (gut)	Nucleus	transcription regulator
KRT19	keratin 19	Cytoplasm	other
MALL	mal, T-cell differentiation protein-like	Plasma Membrane	other
ME1	malic enzyme 1, NADP(+)-dependent, cytosolic	Cytoplasm	enzyme
MYO5C	myosin VC	Unknown	other
PERP	PERP, TP53 apoptosis effector	Plasma Membrane	other
PLA2G16	phospholipase A2, group XVI	Nucleus	other
SLC27A2	solute carrier family 27 (fatty acid transporter), member 2	Cytoplasm	transporter
SLC38A1	solute carrier family 38, member 1	Plasma Membrane	transporter
ST14	suppression of tumorigenicity 14 (colon carcinoma)	Plasma Membrane	peptidase
TSTD1	thiosulfate sulfurtransferase (rhodanese)-like domain containing 1	Unknown	other
VAMP8	vesicle-associated membrane protein 8 (endobrevin)	Plasma Membrane	other