Vol:2, Pg: 35-67, Yr: 2021-AJABTR

# Determination & expression of chimeric protein in colorectal cancer: A bioinformatics Approach

[1]Gargi Bhattacharyya, [1]Amit Chattopadhay*

[1] Serampore college, Department of Zoology

8, William Carey Sarani, Maniktala, Serampore, West Bengal 712201, India

*Corresponding: amitchattopadhyay1906@gmail.com

## Abstract

Colorectal cancer (CRC) accounts for about 10% of cancer-related mortality in western countries. Increasing ageing population, undesirable modern dietary and high-risk factors like smoking, obesity and low exercise. Chromosomal instability (CIN), CpG island methylator phenotype (CIMP), and microsatellite instability are the three different mechanism that give rise to CRC. It often grows slowly, and customarily doesn't produce symptoms until reaching a substantial size of several centimeters, which can block the passage of feces and cause cramping, pain, or bleeding which can present as visible bleeding with bowel movements or, rarely, dark "tarry" stools. Most colon tumors develop via a several different processes involving a series of histological, morphological, and genetical changes that accumulate over time to time. New treatments for primary and metastatic colorectal cancer have emerged, like in variety of therapeutic process by preparing chimeric proteins that triggers the cells and stop it from getting severe.

## 1. Introduction

Colorectal cancer (CRC), also called bowel cancer, carcinoma, or rectal cancer, is associated with colon or rectum (parts of the massive intestine). CRC is an irregular development of cells which will attack or spread to different pieces of the body. Risk factors include diet, obesity, smoking, and lack of physical activity. Dietary factors that increase the prospect include chicken, processed meat, and alcohol. Another risk factor is inflammatory bowel disease, which incorporates regional ileitis and inflammatory bowel diseases.

Abdominal pain, change in bowel habit, and rectal bleeding or anemia are the most common presenting symptoms of colorectal cancer but these symptoms also commonly occur in other gastrointestinal conditions hence makes its identification difficult. CRC often grows slowly, and customarily doesn't produce symptoms until reaching a substantial size of several centimeters, which can block the passage of feces and cause cramping, pain, or bleeding that may present as visible bleeding with bowel movements or, rarely, dark "tarry" stools. There are various steps that involve a series of histological, morphological, and genetical changes that are developed over time to time. Dysplastic adenomatous polyps is the initial stage from where colon cancer develops. A multistep process involves the inactivation of a series of genes that suppress tumors and repair DNA and also the simultaneous activation of oncogenes. Today, CRC is one in all leading causes of morbidity and represents a formidable health burden. Therefore, understanding the molecular and genetic features of its onset and progression is crucial. Here we focused on several novel cellular signaling pathways related to CRC metastasis and role of fusion proteins in CRC. Chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) pathways are three major pathways within the genetic instability of CRC.

Fusion proteins or chimeric proteins can be obtained by joining of two or more genes that originally coded for separate proteins. Fusion proteins are commonly found in cancer cells, where they are going to function as oncoproteins. Via in vitro process chimeric proteins can easily be prepared by fusing the structural genes of the proteins during an appropriate expression vector. Colorectal cancer (CRC) could also be a mode of cancer in humans that lands up in high mortality and morbidity. Here we mainly discussed CD166 and CD326 because they're immunoglobulins that are related to cell migration. These molecules are included in tumorigenesis of CRC and serve a superb marker of CRC stem cells. CD166 and CD326 are frequently found to be overexpressed in tumor cells. Both of these molecules are projected as the potential targets for diagnostics and therapy of CRC. CD166, the Activator Leukocyte Cell Adhesion Molecule (ALCAM), is a member of the immunoglobulin superfamily. This molecule is also an important factor not only for cell survival, motility, and cell growth but also for invasion during tumor progression and

metastases. CD326 is a member of a subgroup of transmembrane glycoproteins within the immunoglobulin superfamily and is additionally mentioned because the cell adhesion molecule (EpCAM). CD326 is expressed at low levels within the healthy epithelial cells but highly expressed within the cancerous epithelial cells like CRC cells, where it preforms essential functions like an epithelial-specific intracellular cell-adhesion activity. Thus, these two Chimeric proteins are useful in colorectal cancer.

Chemotherapy and actinotherapy are developed for the treatment of cancer. However, these methods often cause undesirable side effects. Recently, antibody-based therapy is hugely used in case of tumor allergens for nursing patients with tumors. The bioinformatics approaches can cause a giant reduction in time, expense, and failure in experimental attempts. During this regard, with the advancement of software and continuous information regarding the link between the structure and performance of the protein, they could play a vital responsibility in vaccine design, development of a protein suitable for antigen preparation employed in immunoassay, structural studies, drug-protein, and protein-protein interaction analysis. The present study aims at the evaluation of the ability of chimeric protein composing of V1-domain of the CD166 and epitopes of CD326 as a brand-new antitumor candidate. By using different bioinformatics tools prediction of varied properties of chimeric protein and might possibly be utilized to provide CRC diagnostic kits and develop a protective vaccine against CRC.

## 2. Risk factors and causes of colorectal cancer

Most colorectal cancers are caused due to lifestyle factors, with only a little number of cases because of underlying genetic disorders. [1]

Diet, obesity, smoking, and lack of physical activity comprises of risk factors. Dietary factors that increase the danger include beef, processed meat, and alcohol. Another risk factor is inflammatory bowel disease, which has colitis and inflammatory bowel disease. A number of the inherited genetic disorders that may cause colorectal cancer include familial adenomatous polyposis and hereditary non-polyposis colon cancer; however, these represent but 5% of cases. It typically starts as a nonmalignant tumor, often within the type of a polyp, which over time becomes cancerous. [1]

## 3. Symptoms of colorectal cancer

Abdominal pain, change in bowel habit, and rectal bleeding or anemia are the most common presenting symptoms of colorectal cancer but these symptoms also commonly occur in other gastrointestinal conditions. A change in bowel habit could be a more common presenting symptom for left sided cancers caused by a progressive narrowing of the bowel lumen, with diarrhea, a change in stool form, and eventually obstruction. About 10% of patients with iron deficiency anemia have colorectal cancer, most typically on the proper side, and thus iron deficiency in men, and ladies who aren't menstruating, is a sign for urgent referral and investigation. [2]

**4. Mechanism that causes colorectal cancer**

The environmental and genetic factors that cause colorectal cancer do so by [3] progressive accumulation of genetic and epigenetic alterations that activate oncogenes and inactivate tumour suppressor genes. The loss of genomic and/or epigenomic stability has been observed within the majority of early neoplastic lesions within the colon (namely, aberrant crypt foci, adenomas and serrated polyps) and is probably causing central molecular and pathophysiological event within the initiation and formation of colorectal cancer [4]. The loss of genomic and epigenomic stability accelerates the buildup of mutations and epigenetic alterations in tumor suppressor genes and oncogenes, which drive the malignant transformation of colon cells through rounds of clonal expansion that select for those cells with the foremost aggressive and malignant behavior [5]. A prevailing paradigm is that the cell of origin of most colorectal cancers may be a vegetative cell or stem cell-like cell that resides within the base of the colon crypts [6]. During this model, mutations in oncogenes and tumors suppressor genes in these cells result in the formation of cancer stem cells, which are essential for the initiation and maintenance of tumors. The colon, the evolution of normal epithelial cells to adenocarcinoma follows a predictable progression of histological and concurrent epigenetic and genetic changes (Figure 1). within the 'classic' colorectal cancer formation model, the overwhelming majority of cancers arise from a polyp beginning with an aberrant crypt, which then evolves into an early adenoma (<1 cm in size, with tubular or tubulovillous histology). The adenoma then progresses to a sophisticated adenoma (>1cm in measurement) at last forming a rectal cancer. This process is driven by

accumulation of mutations and epigenetic alterations and takes 10–15 years to occur but can progress sooner in certain settings (for example, in patients with Lynch syndrome) [6,7]. Notably, although the histology of conventional tubular adenomas is fairly homogeneous, the biological science of those polyps is heterogeneous, which could explain why some adenomas make colorectal cancer (approximately 10% of polyps) and a few don't [6-8].

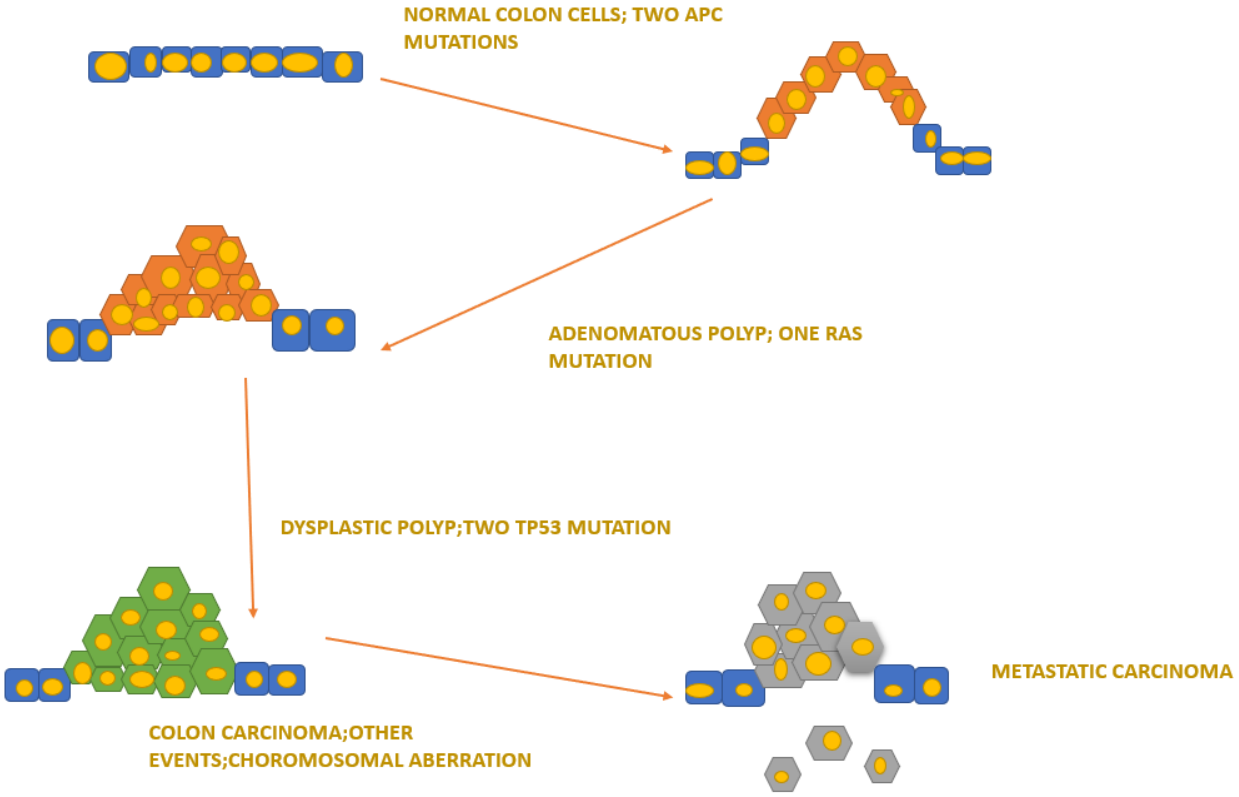



**FIG 1. The polyp to colorectal cancer sequences**

Up to 5–10 years earlier tubiform and tubulovillous adenomatous polyps were belief to be the one and only lesions able in development to cancer. Still, some rectal cancers have been displayed to emerge from a set of polyps called sessile serrated polyps, which counts for likely 5–10% of total polyps. These serrated polyps arise by molecular and histological events that are distinct from tubular adenomas [8] and are classified into three categories: hyperplastic polyps, sessile serrated adenomas and traditional serrated adenomas. The sessile serrated polyps have the potential to transform into colorectal cancers through the following sequence: hyperplastic polyp to

sessile serrated polyp to adenocarcinoma [Kambara, 2004]. Furthermore, serrated polyps that arise in the right colon (which includes the cecum, ascending colon and transverse colon) commonly show MSI (Microsatellite instability) and a form of epigenetic instability distinguish by too much anomalous CpG island DNA methylation, known as CpG Island Methylator Phenotype (CIMP). Polyps that develop in the left side of the colon (which consist of lower colon, s-shape colon and anus) are commonly microsatellite sturdy but regularly convey mutations in KRAS and a subset of these polyps have an attenuated form of the CIMP [9-10]. Given these molecular differences within the polyps and cancers they evolve into, a system for colorectal cancer has been proposed, with four subgroups of differing molecular features: microsatellite unsteady (Hyp-MSI), hyper muted-microsatellite steady (Hyp-MSS), microsatellite stable (MSS) or chromosome unstable (CIN) and CIMP tumors. The commonness of particular mutations may differ substantial in between the molecular middle classes, advising each have its own lay of collaborating drivers. However, the particular mutations and epigenetic alterations that outline these molecular subgroups are still being determined. Some mutations, like those in APC and SMAD4, are common among all the molecular subgroups — suggesting a central role in colorectal cancer normally — whereas others are restricted to 1 subgroup (for example, BRAF in CIMP colorectal cancers) [11]. In colorectal cancer, substantial heterogeneity within the specific mutations is clear between tumors, although the mutations seem to cluster in epistatically related groups (for example, genes involved in a very certain signaling pathway) [12]. The foremost common alterations seen in colorectal cancer include those in APC, catenin-beta1 (CTNNB1), KRAS, BRAF, SMAD4, transforming-growth factor-beta receptor 2 (TGFBR2), TP53, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit-alpha (PIK3CA), AT-rich interactive domain 1A (ARID1A), SRY (sex-determining region Y) box 9 (SOX9), family with sequence similarity 123B (FAM123B; also referred to as AMER1) and ERBB2, which promote tumorigenesis by perturbing the function of key signaling pathways, including the Wnt–β-catenin, epidermal protein (EGF)-mitogen-activated protein kinase (MAPK), phosphatidylinositol 3-kinase (PI3K) and TGF-β signaling pathways, or by affecting

genes that regulate central behaviors of cells, like DNA repair and proliferation [1, 13] (Table 1)(Table 2) and (Table 3). Colorectal cancer is often initiated by alterations that affect the Wnt signaling pathway, and therefore the ensuing neoplastic cells then progress upon deregulation of other signaling pathways, including the RAS–RAF–MAPK, TGF-β, and also the PI3K–AKT pathways. [1,5]

**TABLE1: GENES (TUMOUR SUPPRESSORS) RESPONSIBLE FOR COLORECTAL CANCER**

| **Gene (**Tumor suppressors) | **Chromosome** | **Function** | **Molecular lesion** | **Frequency (%)** | **Predictive?** | **Prognostic?** | **Diagnostic?** |
|---|---|---|---|---|---|---|---|
| APC | 5 | Regulates Wnt pathway | Inactivating mutations | 40–70 | No | No | Familial adenomatous polyposis |
| ARID1A | 1 | Member of SWI/SNF family, regulates chromatin structure. | Inactivating mutations | 15 | No | No | NA |
| CTNNB1 | 3 | Regulates Wnt signalling pathway | Activating mutations | 1 | No | No | No |
| DCC | 18 | Netrin receptor; regulates | Deletion/LOH | 9 (mutations) | No | Possible | No |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | apoptosis. | | | | | |
| FAM123B | X | Involved in Wnt pathway | Inactivating mutations | 10 | No | No | No |
| FBXW7 | 4 | Regulates proteasome mediated protein degradation | Inactivating mutations | 20 | No | No | No |
| PTEN | 10 | Regulates PI3K–AKT pathway | Inactivating mutations. | 10(mutation) 30 (loss of expression) | Possible | No | No |

**TABLE 2: GENES (PROTOONCOGENES) RESPONSIBLE FOR COLORECTAL CANCER**

| **Gene (Proto-oncogenes)** | **Chromosomes** | **Functions** | **Molecular lesion** | **Frequency%** | **Predictive?** | **Prognostic?** | **Diagnostic?** |
|---|---|---|---|---|---|---|---|
| BRAF | 7 | Take part in MAPK signaling pathway | V600E activating mutation | 8-28 | Probable | Probable | Lynch syndrome |
| ERBB2 | 17 | Involved in EGF–MAPK signaling pathway | Amplification | 35 | No | No | No |
| GNAS | 20 | Regulates G-protein signaling | Mutation | 20 | No | No | No |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| IGF2 | 11 | Regulates IGF signaling pathway | Copy number gain, loss of imprinting | 7(mutations) | No | No | No |
| NRAS | 1 | Regulates the MAPK pathway | Mutation in codon 12 or 13 | 2 | Yes | No | No |
| TCF7L2 | 10 | Regulates Wnt signaling | Gene fusion and translocation | 10 | No | No | No |
| KRAS | | Regulates intracellular signaling via MAPK. | Activating mutations in codon 12 or 13 | 40 | Yes | Possible | NA |

**TABLE 3: OTHER MOLECULAR ALTERATIONS**

| **Genes** | **Chromosomes** | **Functions** | **Molecular lesion** | **Frequency%** | **Predictive?** | **Prognostic?** | **Diagnostic?** |
|---|---|---|---|---|---|---|---|
| SEPT9 | 17 | NA | Methylation | >90 | Probable | No | Serum based assay for cancer detection |
| VIM NDRG4, BMP3 | 10,16 and 4 | NA | Methylation | 75 | No | No | Stool based test for early detection |
| 18Qloh | 18 | NA | Deletion of the long arm of chromosome 18 | 50 | No | Probable | No |
| CIN | N/A | NA | Aneuploidy | 70 | Probable | Probable | No |

APC, adenomatous polyposis coli; ARID1A, AT-rich interactive domain 1A; CTNNB1, catenin-β1; DCC, DCC netrin 1 receptor; FAM123B, family with sequence similarity 123B; FBXW7, F-box and WD repeat domain-containing 7, E3 ubiquitin protein ligase; GNAS, guanine nucleotide-binding protein, α-stimulating complex locus; IGF, insulin-like growth factor; TCF7L2, transcription factor 7-like 2; TGFβ, transforming growth factor-β; SEPT9, septin 9; VIM, vimentin; CIN, Chromosome Instability.

## 5. Molecular basis and pathways of colorectal cancer

Both genetic and epigenetic alterations are taking part in regulating tumorigenesis of CRC. The adenocarcinoma sequence was described much earlier, in 1980, when the transformation of normal colorectal epithelium to an adenoma and, finally, to an invasive and malignant tumor was elucidated. There are three major pathways involved within the genetic instability of CRC and its pathogenesis: chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP) pathways. [14,15]

### A. Chromosomal Instability (CIN) Pathway

Chromosomal instability refers to a major increase within the gain or loss of either the complete or the big portions of chromosomes and are responsible for the most typically occurring genetic instability in CRC. CIN is found in (KRAS and BRAF), inactivation of TSGs (APC and TP53), and a loss of heterozygosity for the long arm of chromosome 18 (18q LOH), thus, promoting CRC tumorigenesis. [16].

### A.1 Adenomatous Polyposis Coli (APC) Gene and Wnt Signaling Pathway

On chromosome 5q21-q22, the APC gene is located that encodes a 310 kDa protein and consists of 8535 nucleotides spanning 21 exons. Around 75% of the coding sequence is present on exon 15, the foremost frequent region for both germline and somatic mutations of APC. Furthermore, APC may be a multi-domain protein. From the N to C-terminus, armadillo repeat, 15- or 20-residue repeat, SAMP repeats, a basic domain, and C-terminal domains. Since APC interacts with various binding proteins through its different domains, APC plays a task in regulating cellular processes including chromosome segregation, cell migration, apoptosis, adhesion, proliferation, and differentiation. In both familial and sporadic CRCs, the APC/β-catenin/Wnt-Tcf pathway plays a serious role within the onset and progression of CRC carcinogenesis. The

transition of G0/G1 to the S phase of the cell cycle is inhibited by APC gene. The Wnt signaling pathway maintains the undifferentiated stem cells within the base of the colonic crypts, allowing survival of both normal and cancer stem cells. It's known that β-catenin is the main controller of the Wnt signaling pathway. The Wnt signaling is negatively control by wild-type APC protein by regulating proteasomal ubiquitin-mediated spreading of β-catenin transcription factor. Disruption of the APC protein ends up in enhanced Wnt signaling by intracellular β-catenin stabilization, which stimulates transcription of Wnt targeted genes and enhances TCF targets with increase in cell development, differing, spreading, and gluing of colon cells. Mutations in genes involved in APC/β-catenin/Tcf pathway in CRC cells without APC mutations are present in sporadic CIN tumors. Activating mutations within the gene for β-catenin (CTNNB1) block APC-regulated breakdown and are present in colorectal neoplasia. Although CTNNB1 mutations are more common in adenomas (12.5%) than invasive cancer (1.4%), they're found within the preliminary stages of CRC pathogenesis and plausibly substitute APC mutations in cancer onset and progression. Additionally, distinct units of the APC/β-catenin/Wnt pathway is either directly or indirectly changed, by constitutively triggering β-catenin or Tcf. Among the various regulatory genes that act with APC, the mitotic checkpoint protein BubR1 plays a significant role. BubR1 could be a part of the mitotic checkpoint machinery together with Bub1, Bub3, Mad1, Mad2, Mad3, Mps-1, CENP-E, and biological process cycle 20 (CDC20). The APC is blocked by binding of BubR1 to Cdc20 via targeting a 'pause anaphase' signal, that helps the polyploid cells, spreading cell lifespan, and cell proliferation, suggesting a plausible pathogenic mechanism within the initiation of CIN in CRC sporadic forms. Recently, activation of leucine-rich repeat-containing G-protein-coupled receptors (LGR-4 and LGR-5) triggered signaling by binding with proteins within the R-responding family. Moreover, cyclin D1 (CCND1) was implicated in APC signaling. Mutated APC cells activate downstream targets, like cyclin D1 and Myc. CCND1, together with other cyclin-dependent kinases (CDKs) that block cyclins, like p27 (CDKN1B) and p21 (CDKN1A), are vital for cell growth and apoptosis during cell cycle control, majorly during the transition from G1 to S phase. Prolonged activation of CCND1 by APC

mutation results in the onset of colonic neoplasia by allowing the cell to divert from apoptosis. The expression of CCND1 in colonic mucosa, adenoma, and adenocarcinoma is studied by Arber and colleagues that confirmed the high expression in carcinoma, therefore it indicates CCND1 is expression of an early event during multistage process of CRC tumorigenesis that will deregulate cell-cycle control in benign adenomas and stimulate tumor progression. On the other hand, β-catenin activity is indirectly triggered by aberrations in oncogenes controlling its activity at different levels. β-catenin mutually interacts with various members of the Notch pathway that are vital regulators of cell differentiation and play a task in CRC carcinogenesis. Kwon and colleagues showed that Notch1 stimulates the assembly of active β-catenin protein within the absence of ligand-receptor activation. Further genetic changes that modulate β-catenin activity include CDK8 (cyclin-dependent kinase-8) gene amplification and occurs in over 60% of CRC cases. Increased CDK8 enhances both β-catenin and Notch1, thus stimulating transcription and cell differentiation. [16].

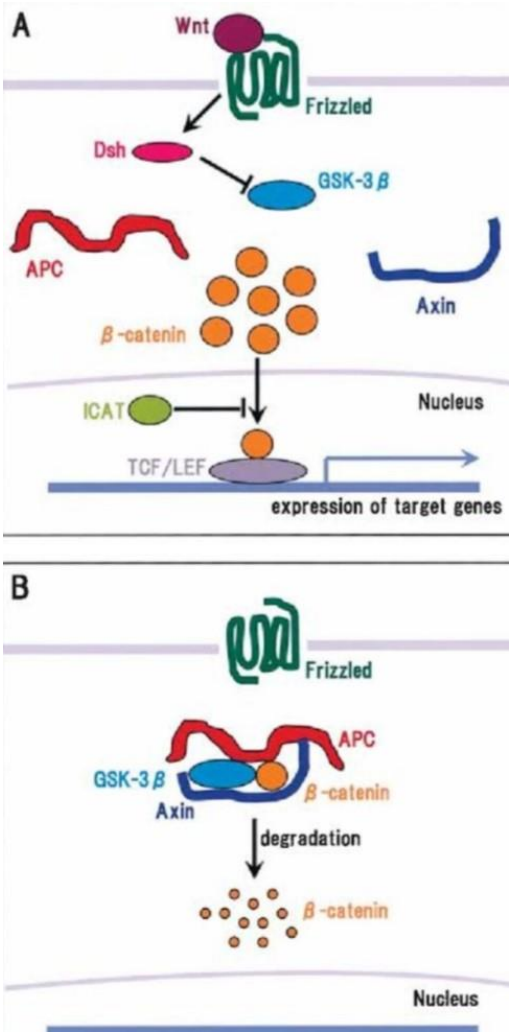



**FIG 2:** - **A**. Wnt signaling stabilizes -catenin by inhibiting GSK3.Accumulate -catenin is transported into nucleus, where it binds to the transcription factor TCF/LEF, leading to the expressing of target genes.

**B**. APC involved in -catenin and, in collaboration with axin, leads to phosphorylation of -catenin by GSK3. The phosphorylated -catenin undergoes proteasome-dependent degradation, resulting in inhibition of Wnt signaling pathway.

**A.2 TP53 Pathway**

On the short arm of chromosome 17, TP53 is located, and this is known as the "parents of the genome" and inscribe proteins that balance cell cycle, DNA repairing, aging, and cell-death. TP53 mutations or loss of function are reported in 50–75% of CRC cases; loss of p53-mediated pathways of apoptosis may be an important determinant of progression from adenoma to malignancy. High cell proliferation activities and uncontrolled cell cycle is enhanced by the loss of function of p53, that lead forward to colorectal carcinogenesis. Research showed missense mutations (48%) that substitute AT for GC because the foremost typical TP53 mutations in CRC, followed by point mutations (37.5%) with transitions at CpG sites. Commonly, p53 is taken into consideration a controller for BubR1 transcription and expression; loss of p53 expression downregulates BubR1, thus comprising checkpoint function to mitotic

aberrations leading to progression of CRC. Additionally, a wild sort of p53 is identified as an on-the-spot activator for WAD-1, a gene highly induced to suppress tumor cell growth within the p53 pathway. During CRC development and progression mutation in CDK, Phosphorylation of p53 and AMPK-dependent cell-cycle is induced by Adenosine monophosphate-activated protein kinase that updates the cyclin-dependent kinases inhibitor 1A (CDKN1A or p21). This eventually controls cell cycle regulation, cellular senescence, and cell aging. Furthermore, in CRC, the stimulation of p21 occurs in an exceedingly p53-dependent pathway; p21 also inhibits the activity of cyclin D1. Poor prognosis in CRC leads to loss of p21. Another sort of CDK associated with TP53 includes loss of function is CDK inhibitor 1B (CDKN1B or p27). CDK inhibitor protein is encoded by an enzyme p27 that is the main reason for controlling cell cycle development to S phase and its humiliation. It has been established that p27 expression is inversely related with the MSI-H and CIMP-H forms of CRC and TP53-negative cancers. Cyclooxygenase-2 (COX-2) interacts with p53 and is entailed in developing of swelling and CRC cell proliferation. Interestingly, COX-2-positive tumors are significantly linked with cancer-specific mortality independent of p53 status, thus suggesting COX-2 as an independent CRC prognostic factor. [16,17].

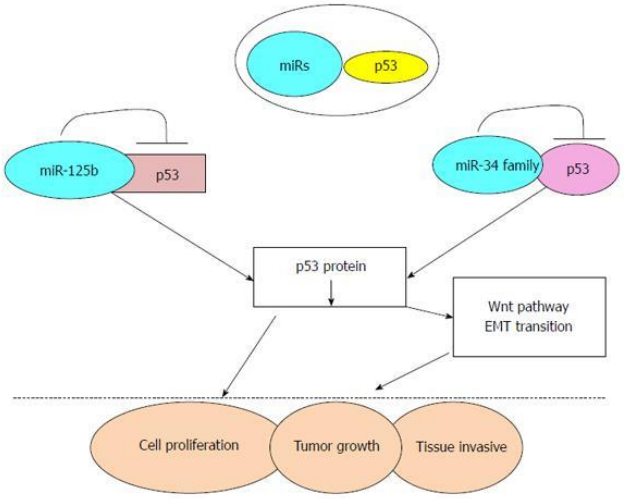



**FIG 3: - Tp53 pathway in colorectal cancer**

**A.3. The 18q Loss of Heterozygosity (LOH)**

Loss of heterozygosity (LOH) refers to the absence of one of the two copies or alleles of a gene, with the remaining allele

frequently being littered with mutation. LOH within the 18q region is most typically observed in advanced CRC, accounting for about 70% of the cases, and is related to poor prognosis in CRC. LOH at 18q indicates presence of several TSGs including deletion in Colorectal Carcinoma (DCC), SMAD2, and SMAD4; loss of expression of 18q LOH plays a big role in CRC pathogenesis. DCC, located on chromosome 18q21.2, encodes netrin-1 and is indicated as a plausible TSG; LOH within the DCC gene region is present in approximately 70% of CRCs. Moreover, some somatic mutations in DCC are found in CRC. Netrin-1 is made within the cryptos of colorectal mucosa; somatic cell differentiation ends up in loss of netrin-1 expression. Furthermore, mutation in DCC gene inhibits binding of netrin-1 to DCC transmembrane protein, resulting in abnormal cell survival. On the opposite hand, SMAD2 and SMAD4 are present on 18q21.1, the prevalent region lost during CRC progression, and correlate with adenoma development and adenocarcinoma progression in mice models, therefore advising a credible tumor silencer character for SMAD genes. Furthermore, immune-histochemical analysis reported a loss of SMAD4 expression in >50% of CRCs, which is related to lymphoid tissue metastases. Since the frequency of somatic mutations in SMAD2 and SMAD4 is relatively low in CRC, other TSGs may be liable for chromosome 18q loss. SMAD genes are inscribed for TGF-β. Dysregulation of TGF-β signaling occurs within the majority of CRCs. Inactivating mutations in receptor genes that consist of TGF-βR1, TGF-βR2, and TGF-β superfamily participants as Activin Receptor type 2 (ACVR2) appear in CRC. Functionally, marked mutations in TGF-βR2 are present in ~30% of all CRC cases and correlate with malignant transformation lately adenomas. TGF-βR2 mutations are most frequent in MSI tumors; however, they're also present in around 15% of MSS tumors. [16,17]

**B. The Microsatellite Instability Pathway**

Another sort of genomic instability in CRC is microsatellite instability (MSI), a particular characteristic of cancerous cells. MSI is that the hallmark of HNPCC or Lynch syndrome and occurs in >95% of HNPCC cases. However, within the majority of sporadic CRCs, the underlying mechanism for CIN remains nascent and MSI comprises merely 15–20% of all CRC cases. Small

insertions/deletions lead to frameshift mutations in tedious routes within the coding parts of TSGs, additionally contributes to tumorigenesis. Mori et al. performed large-scale genomic screening of the coding region of microsatellites and located mutations in nine loci (TGF-βR2, Bax, MSH3, ActRIIB, SEC63, AIM2, NADH-ubiquinone oxidoreductase, COBLL1, and EBP1) in >20% of tumors. TGF-βR2 was the foremost commonly mutated loci and instability within the poly-adenine tract of TGF-βR2 that is present in approximately 85% of MSI-H CRCs. Moreover, Bax, the opposite frequently mutated gene, was found to possess frameshift mutations within the polyguanine sequence in almost 50% of the MSI-H CRCs, leading to the inactivation of Bax and inhibition of apoptosis. MSI isn't found in polyps, except in Lynch syndrome. Furthermore, individuals with Lynch syndrome frequently develop MSI CRCs because of germline mutations in one of the MMR genes (MLH1, MSH2, MSH6, and PMS2); mutations in MLH1 or MSH2 gene result in an increased risk (70–80%) of developing cancer, while mutations within the MSH6 or PMS2 gene have a relatively lower risk (25–60%) of cancer development. On the contrary, sporadic MSI CRCs frequently display loss of MMR activity thanks to MLH1 silencing by aberrant DNA methylation. Furthermore, modeling studies and absence of carcinoma in individuals with biallelic germline mutations in MMR genes suggest that lack of MMR activity is insufficient to trigger polyp formation. Relevant to its clinical impact, there's considerable indirect data that polyps arising as a results of MMR activity loss encompasses a lesser transition interval from a polyp to colorectal cancer; polyps can turn into MSI CRCs within 2–3 years. Evidence suggesting lack of MMR progress and onset of MSI induces tumor development and progression relies on the actual fact that MSI is mostly observed in polyps adjacent to cancers and is infrequent in non-advanced polyps. It’s known that sporadic MSI CRCs correlate with the serrated neoplasia pathway and commonly carry BRAFV600E mutations; on the contrary, Lynch syndrome arises from MMR genes germline mutations and lacks mutated BRAF. Clinically, BRAF-mutated CRC correlates with poor prognosis and overall survival (OS) as compared to BRAF wild-type disease. Although it shows that BRAF-mutated CRC candidates had bad OS as in comparison to patients who carry RAS (KRAS and NRAS) mutations. Also,

mutation in BRAF may be a negative prognostic during stage II and III disease. On the contrary, a recent study, performed meta-analysis in 1164 MSI-H non-metastatic CRC patients showed that BRAF V600E mutation is related to worst OS, but not disease recurrence. Moreover, another meta-analysis in patients undergoing resection of liver metastasis showed that following mastectomy, OS was worst for BRAF-mutated metastatic CRC. NonV600E BRAF-mutated (BRAF codons 594 and 596) CRCs have good prediction as contrast to BRAF-mutated CRCs; BRAF 594 and 496 tumors are microsatellite firm, intestinal, no papillary with no abdom spread and have a significantly longer OS as compared to V600E BRAF-mutated tumors. Similarly, other studies in CRC patients showed, as compared to V600E BRAF-mutated tumors, non-V600E BRAF-mutated tumors are present within the younger population, lower grade tumors and also the median OS is contrast with two V600E BRAF-mutant and BRAF wild-type patients. [16-18].

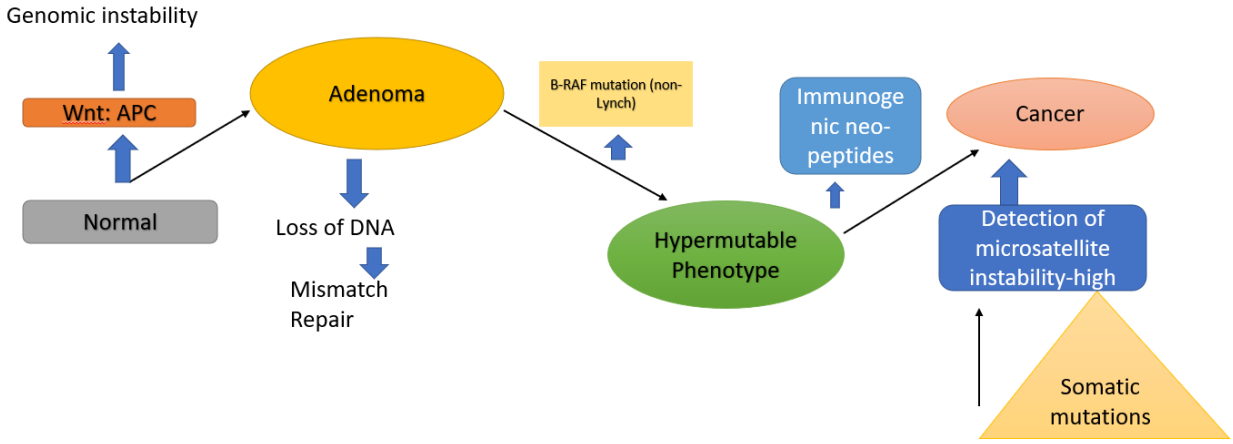
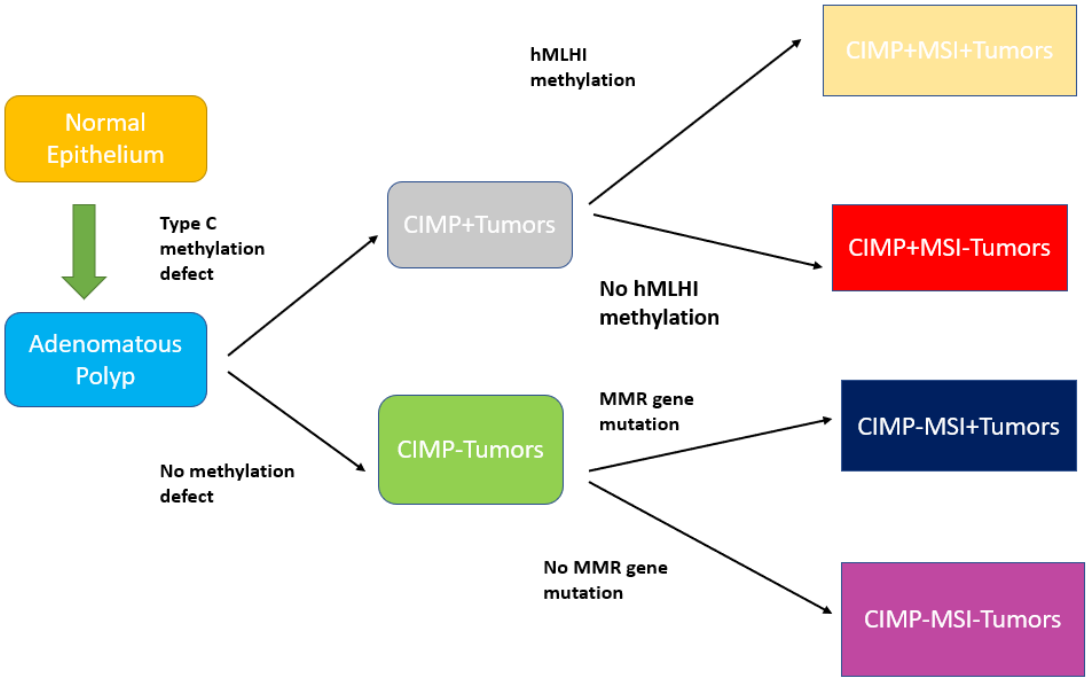



**FIG 4: - Microsatellite Instability pathways.**

### C. CpG Island Methylator Phenotype (CIMP) Pathway

DNA methylation is the addition of an alkyl to cytosine within the 50 -position that's catalyzed by DNA methyltransferases via covalent linkage within a CG dinucleotide sequence within the promoter region, termed CpG transcription. In normal cells, the majority of the CpG sites are heavily methylated, while CpG islands usually located within the promoter regions of genes, are unmethylated. However,

following cancer initiation, hypermethylation within the promoter region may cause inactivation of tumor suppressor genes, while global hypomethylation is expounded to genomic instability and chromosomal aberrations. Epigenetic instability in CRC is demonstrated as hypermethylation of loci that contain CpG islands. Change in methylation way can act on practically all signal pathways, that includes TP53, TGFβ/SMAD, Wnt, NOTCH and receptor tyrosine kinases is taken part in cell cycle process, transcription ordinance, DNA steadiness, cell death, cell-cell sticking, ontogeny, cell occupation and transition. various genes are recognizing to be methylated and muted in CRC, methylated ones include APC, MLH1, MGMT, SFRP1, SFRP2, CDKN2A, TIMP3, VIM, SEPT, CDH1 and HLTF. Additionally, there is a definite subset of CRCs, called the CpG island methylator phenotype (CIMP); CIMP tumors frequently carry BRAFV600E mutations. CIMP is further subclassified to support integrated genetic and epigenetic instability into CIMP2, CIMP-low, and CIMP-high. DNA methylation profiling revealed that approximately 20% of CRCs are CIMP tumors; CIMP tumors significantly correlated with age, female sex, proximal colon location, further as MSI, KRAS and BRAF mutation. The commonly used CIMP markers are MLH1, p16, MINT1, MINT2 and MINT31. Extra indicator include CACNA1G, CRABP1, IGF2, NEUROG1, RUNX3, SOCS1, HIC1 and IGFBP3 for practical CIMP recognition. Although, upregulated expression of the DNA methyltransferases (DNMT3B or DNMT1) is said to CIMP, the underlying mechanism(s) that promote CIMP are still unknown. One in every of the plausible underlying mechanisms relies on the silencing of barriers that inhibit methylation of normally unmethylated CpG islands. The alternative suggested mechanism is alterations within the chromatin structure and histone modification state of histone H3 end in the detection of aberrant DNA methylation in loci that obtain this alteration. PTEN, a TSG, manifest lessen methylation rates, TWIST1 gene is muted by promoter methylation in CRC. Except all of the above-mentioned mechanisms of DNA hypermethylation, global DNA hypomethylation frequently occurs at repetitive sequences including LINE-1 repeats, retrotransposons, introns and gene deserts. [16-19].




**6. FIG 5: - CpG island methylator phenotype in colorectal cancer.**

The epidermal protein receptor (EGFR) belongs to the ErbB/HER family and consists of 4 members; ErbB1 (EGFR/HER1), ErbB2 (Neu/HER2), ErbB3 (HER3), and ErbB4 (HER4). Activation of EGFR pathway, triggers several downstream intracellular signaling pathways, including the RAS/RAF/MEK/ERK, PI3K/AKT, and JAK/STAT3 pathways to manage cell growth, survival, and migration. Deregulated EGFR expression is present in various cancers including CRC; increased EGFR expression is present in 25–77% of CRC cancers. Activation of EGFR induces RAS-RAF activation, which ends up in phosphorylation of mitogen-activated protein kinase (MAPK or MEK) and activation of extracellular signal-related kinase (ERK). The MAPK alley consist of KRAS and BRAF; regulation of cell multiplication, discrimination, cell-death and aging. Activation of MAPK includes RAS, RAF and MEK; RAS stimulates the signaling cascade via the phosphoinositol kinases (PI3K) also as RAF. PI3K switching on prohibit cell-death, switching on of RAS arouse cellular multiplication, thus, encouraging cell life-span and tumor

occupation and transition. KRAS mutations include codons 12 and 13 on exon 2 and codon 61 on exon 3; codon 12 being the foremost frequently affected through missense mutations including substitution of glycine for aspartate (p.G12D and p.G13D), of which p.G13D account for 58% of the cases. KRAS together with NRAS and HRAS are oncogenes belonging to the RAS family. KRAS is often mutated in sporadic CRCs (35–45%) and is related to poor prognosis; in line with the adenocarcinoma sequence, KRAS mutations occur after APC mutations. On the opposite hand, BRAF, a member of RAF family of serine/threonine kinases regulates cellular responses to growth signals through the RAS-RAF-MAP kinase pathway. Activating mutations in BRAF are found in approximately 5–10% of metastatic CRC, however, they're rare in Lynch syndrome varieties of CRC. Moreover, BRAF mutations were identified in 40% of MSI-H and 4% of MSI-L tumors. The bulk of the BRAF mutations include the hotspot mutation, V600E (Val600Glu) and is found to correlate with poor prognosis in CRC patients. Angiogenesis, the event of latest blood vessels, is involved in tumor initiation, growth, and metastasis and involves several factors including vascular endothelial growth factors (VEGFs). In CRC, VEGF levels and VEGFR activity is enhanced and is related to poor prognosis. Elevated VEGF levels are seen in very early stages of colorectal neoplasia (adenoma); however, they were significantly elevated in an exceedingly later stage of cancer (metastatic stage). Aberrant KRAS and TP53 still as COX-2 expression regulate VEGF-VEGFR activity alteration, thus promoting cancer growth and migration. The molecular pathways involved within the pathogenesis of CRC are depicted in Figure 6. [16,17].

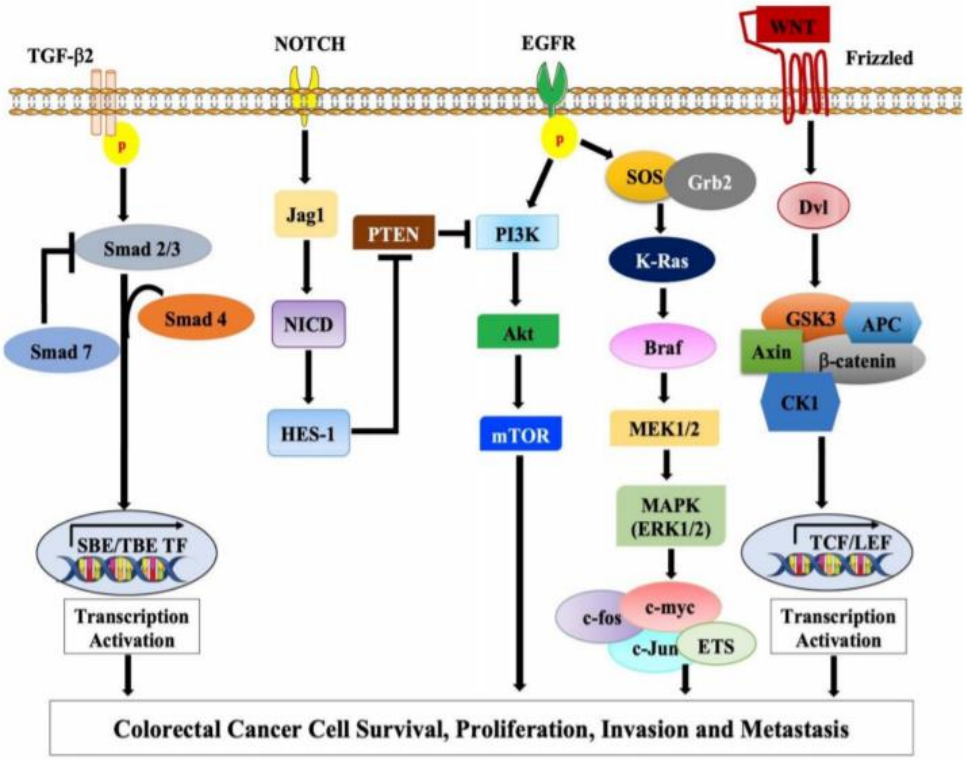



**FIG 6. Schematic representation of the molecular pathways involved in CRC pathogenesis**

APC: Adenomatous polyposis coli; BRAF: Serine/threonine-protein kinase B-Raf; EGFR: Epidermal growth factor receptor; GSK3: Glycogen synthase kinase 3; MAP: MUTYH-associated polyposis; MAP/MEK: Mitogen-activated protein kinase; PI3K: Phosphatidylinositol 3-kinase; SMAD: Small mothers against decapentaplegic; TGF-β: Transforming growth factor-β; TP53: Tumor protein 53; Wnt: Wingless

## 7. Development of colorectal cancer

From dysplastic adenomatous polyps' colorectal cancers arise in many cases. A multistep process involves the inactivation of a spread of genes that suppress tumors and repair DNA and therefore the simultaneous activation of oncogenes. This confers a selective growth advantage to the colonic somatic cell and drives the transformation from normal epithelium to polyp to invasive colorectal cancer [18]. Germline (hereditable) mutations underlie the well described inherited carcinoma syndromes whereas sporadic cancers arise from a stepwise accumulation of somatic genetic mutations. One germline mutation within the APC tumor factor is chargeable for the dominantly inherited

syndrome, familial adenomatous polyposis coli. It is characterized by the event of hundreds to thousands of adenomatous polyps within the colon and development of colorectal cancer and other cancers within the third and fourth decade of life.

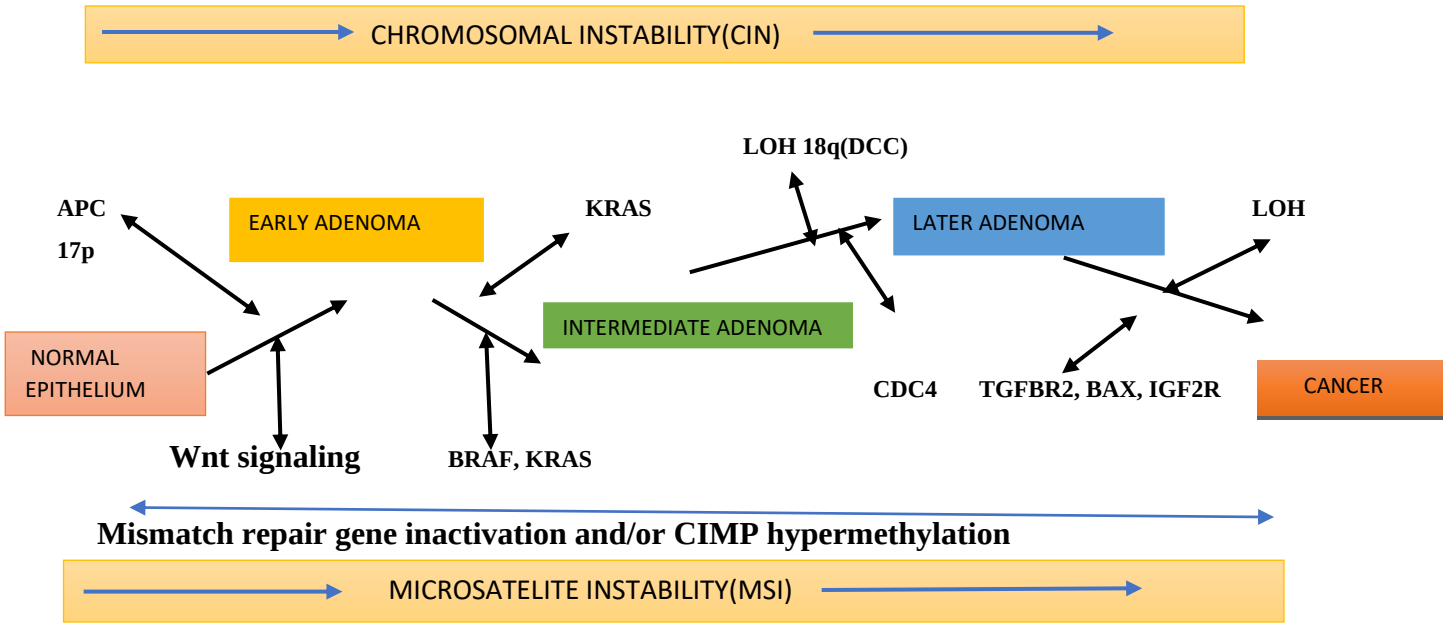



**FIG 7: - Flowchart showing development of colorectal cancer**

## 8. Chimeric protein

Chimeric or fusion proteins are proteins created through the joining of two or more genes that originally coded for separate proteins.

Chimeric or chimera usually designate hybrid proteins product of polypeptides having different functions or physico-chemical patterns. Chimeric mutant proteins occur naturally when a posh mutation, like a chromosomal translocation, tandem duplication, or retro transposition creates a complete unique coding sequence containing parts of the coding sequences from two different genes. Fusion proteins are commonly found in cancer cells, where they'll function as oncoproteins. Here are some listed fusion proteins present in cancer cells (Table 4).

**TABLE 4: - DIFFERENT FUSION PROTEINS IN DIFFERENT TYPES OF CANCER.**

| TYPE OF CANCER | NAME OF FUSSION PROTEIN | ROLE OF THE PROTEIN | REFERENCE |
|---|---|---|---|
| Adenoid cystic carcinoma | MYB-NFIB | Overexpression of MYB leads to increase the MYB-activating genes as well as genes associated with cell proliferation and angiogenesis. This fusion is highly beneficial for therapeutic target in ACC. | Persson et al.,2009 |
| Mucoepidermoid carcinoma | CRTC1-MAML2 | The fusion is found to induce Notch signaling and cause cellular transformation of RK3E epithelial cells. MAML2 used to distinguish mucoepidermoid carcinoma from other oncolytic lesions | Garcia et al.,2011 |
| Follicular thyroid carcinoma | PAX8-PPARG | PAX8 encodes a nuclear protein involved in follicular cell development in thyroid. PPARG regulates fatty acid storage and glucose metabolism. PAX8 promoter causes high expression of fusion and induces tumorigenesis. | Powell et al.,2004 |
| Breast carcinoma | ETV6-NTRK3 | This fusion originally characterized in congenital fibrosarcoma, it promotes oncogenesis via activation of the RAS-MAPK and PI3K-AKT pathways. | Wai et al.,2000 |

| | | | |
|---|---|---|---|
| Ewing sarcoma | EWSR1-FLI1 | It induces a TP53 dependent growth arrest in fibroblast supporting the importance of TP53 loss in Ewing tumors, it also activates CASP3 and promote apoptosis. | Brittany and Wei Zhang, 2013 |
| Synovial sarcoma | SS18-SSX1 | SS18-SSX1 promotes tumorigenesis by increasing the expression ofSHL-SH2 domain binding protein, which normally act as a tumor promoting factor. | Brittany and Wei Zhang, 2013 |
| Lung cancer | EMLA-ALK | The oncogenic function of this fusion protein leads to constitutive activation of downstream signaling cascades, including Akt, MAPK, and signal transducer and activator of transcription 3(STAT3). | Takezawa et al.,2011 |
| Bladder cancer | FGFR3-TACC3 | These fusion proteins span the cell membrane, TACC3 encodes a microtubule associated protein, this fusion helps to promote MAPK signaling and increase cell proliferation and transformation. | Williams et al.,2013 |
| Prostate cancer | TMPRSS2-ERG/ETV1/ETV4 | Benign prostrate cell overexpression ERG directly engages component of plasminogen activation pathway to mediate cellular invasion, potentially representing a downstream ETS target susceptible to therapeutic intervention. | Brittany and Wei Zhang, 2013 |

| | | | |
|---|---|---|---|
| Colorectal cancer | PTPRK-RSPO3 | RSPO3 reduces stem cell function in PTPRK-RSPO3 colon tumors, it regulates its function. RSPO are involved in cellular proliferation, differentiation, and maintenance of stem cell by modulating Wnt pathway. | Kazanskaya et al.,2004 |

EWSR1, Ewing sarcoma breakpoint region 1; ETS, E-twenty six; FLI1, friend leukemia virus integration 1; NTRK3, neurotrophic tyrosine receptor kinase, type 3; PAX8, paired box gene 8; PPARG, peroxisome proliferator-activated receptor gamma; MAML2, mastermind-like protein 2; TMPRSS2, transmembrane protease, serine 2; ALK, anaplastic lymphoma receptor tyrosine kinase; NFIB, nuclear factor 1 B-type; FGFR3, fibroblast growth factor receptor 3; RSPO3, R-spondin family protein 3; SSX, synovial sarcoma X

## 9. Formation of chimeric protein: -

### 9.1. by recombinant fusion proteins

Chimeric proteins can easily be prepared by recombinant means in vitro by fusing the structural genes of the proteins in question during a suitable expression vector. The translational 3′-terminus of the primary gene is deleted, as is that the promoter of the 5′-terminus of the second factor. The 2 genes are then ligated in-frame and expressed in an appropriate host. The foremost frequently used hosts are bacteria like Escherichia, but plant, mammalian, and bug cells have also been used. After transcription and translation, the cell will produce one single polypeptide chain with the properties of both the first gene products. The fusion is made at either or both termini of a protein. Whether one end is more favorable for the biological activity of the protein or not should be evaluated for every fusion construct. The DNA molecules to be fused may be short synthetic oligonucleotides or full-length structural genes. The increasing number of sequenced genes and genomes, together with polymerase chain reaction, provides many combinations for possible fusion of structural genes from a spread of sources. There are just endless numbers of possible combinations of fusion partners

have turned this method into a flexible and valuable tool within many areas of biochemistry and biotechnology. [19,20].

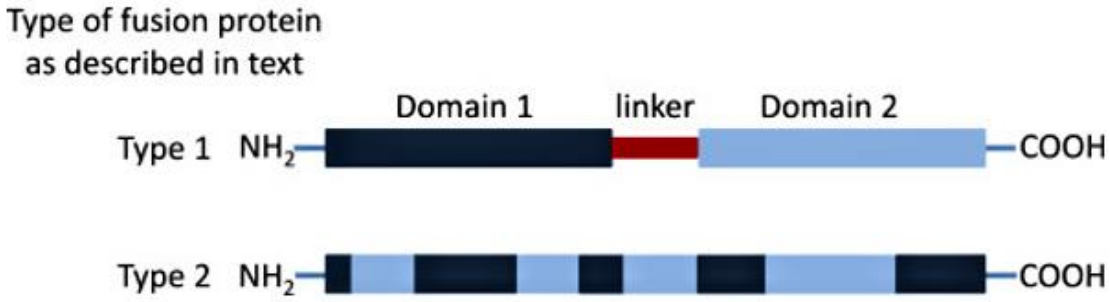



**FIG 8: -** Two basic types of fusion, or chimeric protein, the first of which consists of two proteins or proteins subunits fused end to end and usually linked by a linker, and the second, in which amino acids from both donors are interspersed.

### 9.2. Chimeric proteins in colorectal cancer

Colorectal cancer (CRC) could be a sort of cancer in humans that results in high mortality and morbidity. CD166 and CD326 are immunoglobulins that are related to cell migration. These molecules are included in tumorigenesis of CRC and serve a good marker of CRC stem cells. The V 1-domain of the CD166 and two epitopes of CD326 is a unique chimeric protein that is used in diagnostic or therapeutic applications. The chimeric protein is also useful as a CRC diagnostic tool and for developing a protective vaccine against CRC. CD166 and CD326 are frequently found to be overexpressed in tumor cells. Both of those molecules are projected as the potential targets for diagnostics and therapy of CRC. CD166, the Activator Leukocyte Cell Adhesion Molecule (ALCAM), may be a member of the immunoglobulin superfamily. This molecule may be a vital factor not just for cell survival, motility, and cell growth but also for invasion during tumor progression and metastases. CD166 may be a glycoprotein that was initially discovered as an MHC-I for the cell surface receptor of lymphocyte (CD6) It contributes

to 1) heterotypic adhesion to the lymphocyte cell-surface receptor (CD6). It consist of five extracellular immunoglobulin domains (VVC2C2C2) supported the functional mapping analyses, the V domain within the extracellular region of CD166 is crucial for both styles of cell-cell adhesion. The V domain is identified because the main MHC-I binding domain at the N-terminal immunoglobulin domain. It comprised of two parts as V1 and V2 with 93 and 110-amino acids length, respectively. CD326 could be a member of a subgroup of transmembrane glycoproteins within the immunoglobulin superfamily and is additionally referred to as the somatic cell adhesion molecule (EpCAM). At low levels within the healthy epithelial cells CD326 but within the cancerous epithelial cells like CRC cells is expressed highly, where it performs functions like an epithelial-specific intracellular cell-adhesion activity. CD326 expression has been related to CRC carcinogenesis, and its expression can be a beneficial biomarker for the clinical diagnosis of CRC. This implies that CD326 may be a possible target for the immunotherapeutic treatment of CRC. [21]

## 10. Determining the properties of chimeric proteins that are usefull in colorectal cancer by using bioinformatics tools.

### 10.1. The Secondary and Tertiary Structures of the Chimeric Protein:

The tertiary structure of the fusion/chimeric protein can be established by using homology modeling. The modeling of the chimeric protein can be done with Modeller9.20. Since the chimeric protein possesses domains from two proteins (CD326 and CD166), the templates 4MZV A and 5A2F A can be utilized to make the whole assembly of the chimeric protein using Modeller9.20. Finally, to gauge the stereochemistry of the modeled protein, the PDB structure of the chimeric protein was submitted to the Procheck_NT, followed by Ramachandran plot for structural stability. By using the Predict Protein server the

solvent accessibility, alpha-helix, random coil, and beta sheets structures can be analyzed. By using the Protparam server for predicting the steadiness and other physiochemical properties of the chimeric protein. [22] So, The PORTER and SOPMA online web-based servers help in predicting the secondary structure of the chimeric protein. It shows that the structure consists of 252 amino acids that are made of an alpha helix (43.25%) and random coil (40.48%) (Figure 9). By operating homology modeling the tertiary structure of the chimeric protein was prepared. By using pBLAST against the PDB database the template for homology modeling was attained. 5A2F "A" (1–127) and 4MZV "A" (121–246) with identities of 74.63% and 72.73%, respectively are the two templates were obtained. Considering both the templates, advanced modeling was dispensed using both the templates to make a full-length model of a chimeric protein (Figure 10) [22].

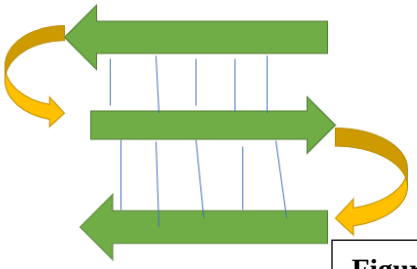
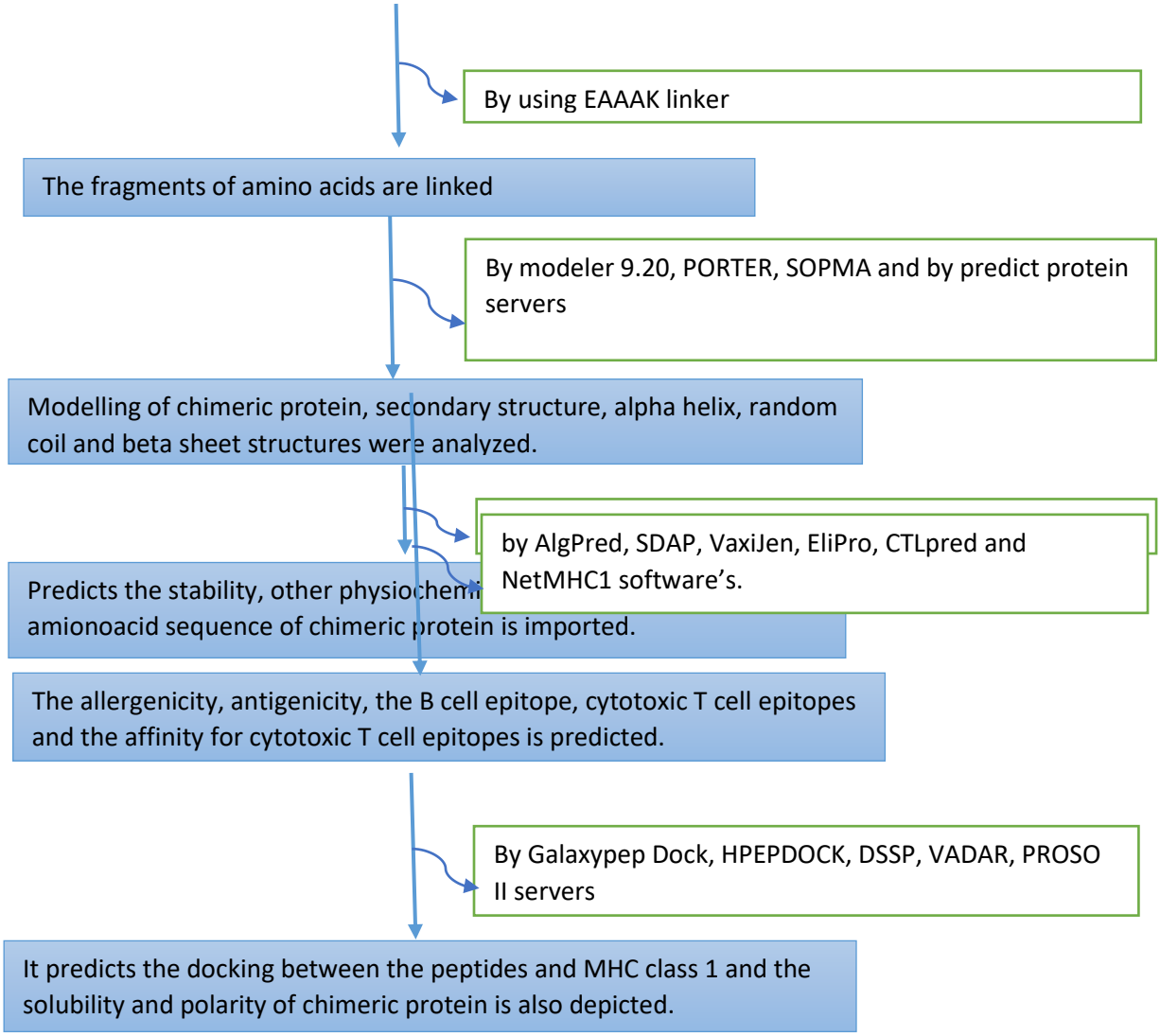



**Figure 9:** Prediction of the secondary structure of the chimeric protein.


**Figure 10:** The tertiary structure of the chimeric protein.

**FLOWCHART DEPICTING VARIOUS SERVERS AND SOFTWARES FOR STRUCTURE PREDICTIONS AND HOMOLOGY MODELLING OF CHEMIRIC PROTEIN.**

Sequence of CD326 and V1 domain of CD166 are selected.




## 11. Discussion

From above we came to grasp about what's colorectal cancer, its causes, risk factors, symptoms, mechanism how it causes and the way it develops. So, early diagnosis is its key to its cure. If the tumor spread to the lymph nodes, a patient's chance of living a minimum of five year drops to 40-60%. In regulating tumorigenesis of CRC both genetic and epigenetic alterations are involved. The adenocarcinoma sequence was described much earlier, in 1980, when the transformation of normal colorectal epithelium to an adenoma and, finally, to an invasive and tumor was elucidated. There are three major pathways involved within the genetic instability of CRC and its pathogenesis: chromosomal instability (CIN), microsatellite instability (MSI), and CpG island methylator phenotype (CIMP)

pathways is additionally discussed. So, here we determine, express and study a unique chimeric protein for cancer immunotherapy. Here three different genetic pathways that are involved in genetic instability of CRC are explained, they're chromosomal instability (CIN), (MSI). Multiple different properties are shown to activate the system against CRC. Generally we had studied many alternative ways to forestall CRC but by therapeutic way chimeric protein plays a crucial role. So, here we've taken a chimeric protein composing of V1-domain of the CD166 and epitopes of CD326 as a brand-new antitumor candidate. CD166, the Activator Leukocyte Cell Adhesion Molecule (ALCAM), could be a member of the immunoglobulin superfamily. This molecule could be a vital factor not just for cell survival, motility, and cell growth but also for invasion during tumor progression and metastases. CD326 is a member of a subgroup of transmembrane glycoproteins within the immunoglobulin superfamily and is additionally called the somatic cell adhesion molecule (EpCAM). Within the healthy epithelial cells CD326 is expressed at low levels but within the cancerous epithelial cells like CRC cells is highly expressed, where it performs important functions like an epithelial-specific intracellular cell-adhesion activity. Thus, these two Chimeric proteins are useful in colorectal cancer. Since it had been important to determine the structure-function relation of chimeric protein before starting experimental studies, the chimeric protein needs to be analyzed by various tools and softwares. The designed chimeric protein needs to retain high stability and same immunogenicity as of the first proteins. The advancement in immune-informatics has made possible styling of recent molecules and predicting its functionality. These findings will intensify efforts to develop a vaccine against CRC and should also suggest this synthetic chimeric protein could help to diagnosis of CRC. Thus, the chimeric protein can possibly be utilized to provide CRC diagnostic kits supported the ELISA technique and develop a protective vaccine against CRC.

**References**

1. "Colon Cancer Treatment (PDQ®)". NCI. May 12, 2014. Archived from the original on July 5, 2014. Retrieved June 29, 2014.

2. Bosman FT (2014). "Chapter 5.5: Colorectal Cancer". In Stewart BW, Wild CP (eds.). World Cancer Report. the International Agency for Research on Cancer, World Health Organization. pp. 392–402. ISBN 978-92-832-0443-5.

3. Rockey DC, Cello JP. Evaluation of the gastrointestinal tract in patients with iron-deficiency anemia. N Engl J Med 1993; 329:1691-5.

4. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell. 2000; 100:57–70.

5. Colussi D, Brandi G, Bazzoli F, Ricciardiello L. Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention. Int J Mol Sci. 2013; 14:16365–85.

6. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell. 1990;61:759–
67. [PubMed] [Google Scholar]

7. Zeki SS, Graham TA, Wright NA. Stem cells and their implications for colorectal cancer. Nat.Rev.Gastroenterol Hepatol. 2011;8:90–100.

8. Jones S, et al. Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci. 2008; 105:4283–4288.

9. Luo Y, et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. Gastroenterology. 2014; 147:418–29. e8.

10. Goldstein NS. Serrated pathway and APC (conventional)-type colorectal polyps: molecular-morphologic correlations, genetic pathways, and implications for classification. Am J Clin Pathol. 2006; 125:146–53.

11. Kambara T, et al. BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut. 2004; 53:1137–44.

12. Jass JR. Hyperplastic polyps and colorectal cancer: is there a link? Clin Gastroenterol Hepatol. 2004; 2:1–8.

13. Chittenden TW, et al. Functional classification analysis of somatically mutated genes in human breast and colorectal cancers. Genomics. 2008; 91:508–11.

14. Parsons DW, et al. Colorectal cancer: mutations in a signaling pathway. Nature. 2005; 436:792.

15. Brennan CW, et al. The somatic genomic landscape of glioblastoma. Cell. 2013; 155:462–77.

16. Malki, A.; ElRuz, R.A.; Gupta, I.; Allouch, A.; Vranic, S.; Al Moustafa, A.-E. Molecular Mechanisms of Colon Cancer Progression and Metastasis: Recent Insights and Advancements. Int. J. Mol. Sci. 2021, 22, 130. http:// dx.doi.org/10.3390/ijms22010130

17. Arnold CN, Goel A, Blum HE, Boland CR. Molecular pathogenesis of colorectal cancer. Cancer 2005; 104:2035-47.

18. Persson M, Andren Y, Mark J, et al. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. Proc Natl Acad Sci U S A. 2009; 106:18740–18744.

19. Garcia JJ, Hunt JL, Weinreb I, et al. Fluorescence in situ hybridization for detection of MAML2 rearrangements in oncocytic mucoepidermoid carcinomas: Utility as a diagnostic test. Hum Pathol. 2011; 42:2001–2009.

20. Powell JG, Wang XY, Allard BL, et al. The PAX8/PPARgamma fusion oncoprotein transforms immortalized human thyrocytes through a mechanism probably involving wild-type PPARgamma inhibition. Oncogene. 2004; 23:3634–3641.

21. Wai DH, Knezevich SR, Lucas T, et al. The ETV6-NTRK3 gene fusion encodes a chimeric protein tyrosine kinase that transforms NIH3T3 cells. Oncogene. 2000; 19:906–915.

22. Takezawa K, Okamoto I, Nishio K, et al. Role of ERK-BIM and STAT3-survivin signaling pathways in ALK inhibitor-induced apoptosis in EML4-ALK-positive lung cancer. Clin Cancer Res. 2011; 17:2140–2148.