Pharmaceutical Sciences Asia

Pharm Sci Asia 2025; 52(1), 11-29

DOI:10.29090/psa.2025.01.24.1659

Research Article

Unveiling Potential Therapeutic Targets for Colon Cancer: A Comprehensive Bioinformatics Analysis of miRNA-Mediated Regulation

Fachrur Rizal Mahendra[1,2], Gusnia Meilin Gholam[1,2], Muhammad Musthofa[1], Muhammad Marsha Azzami Hasibuan[1], Akbar Rafiqi[1], Hafizh Zahra[1], Rahadian Pratama[1], Laksmi Ambarsari[1], I Made Artika[1], Waras Nurcholis[1,3*]

1 Department of Biochemistry, Faculty of Mathematics and Natural Sciences, Bogor 8 Agricultural University, Bogor 16680, Indonesia
2 Bioinformatics Research Center, Indonesian Institute of Bioinformatics (INBIO), Malang, 10 Indonesia
3 Tropical Biopharmaca Research Center, Bogor Agricultural University, Kampus IPB Taman 12 Kencana, Bogor, 16128, Jawa Barat, Indonesia

ABSTRACT

Cancer remains a significant health challenge in Asia, with lung, breast, and colorectal cancers ranking among the most prevalent and lethal. Colon cancer, when detected early, offers the prospect of successful treatment; however, late-stage diagnosis often limits curative interventions. Recent advances have proposed microRNAs (miRNAs) as promising elements in cancer therapy, acting as either tumor suppressor genes or inhibitors of oncogenic miRNAs. This study employed topology and centrality analyses to identify pivotal genes in colon cancer and explore dysregulated miRNAs with high binding probability to their target proteins. This investigation extends to the assessment of miRNA-mRNA interactions and duplex complexes with argonaute (AGO) proteins, shedding light on potential therapeutic avenues. Notably, our bioinformatics predictions potential upregulated genes in colon cancer of Asia population such CAV1, KLF4, and TAGLN and the potential of miRNA suppressing these genes are hsa-miR-429, hsa-miR-200c, and hsa-miR-22 respectively, as robust candidates for suppressing oncogene expression in colon cancer. Detailed scrutiny of the binding affinity, conformation, and conformational stability through molecular dynamics enhances the precision of these findings. This study provides valuable insights into the selection of optimal miRNA candidates for targeted utilization in colon cancer therapy.

Keywords:
Colon Cancer; miRNA; Network Analysis; In silico

1. INTRODUCTION

By 2020, lung, breast, and colorectal cancers were expected to be the top three cancers in Asia[1]. Geographical factors are one of the determinants of a high number of patients with cancer. For example, prostate cancer is among the top three cancers in North America and Oceania, and it ranks eighth in Asia. Most cancer-related deaths in Asia are caused by digestive cancers (liver and stomach cancers)[2]. The prevalence of *Helicobacter pylori* is generally higher in Asian countries/regions than in Western countries, which may explain the higher burden of digestive cancers, such as colon cancer, in Asia. Other aspects, including ethnic differences, smoking, alcohol consumption, diet, and sanitation measures, may also contribute to the wide geographical variation in the cancer burden[3]. Regional estimates show that among half of the new cases, deaths and 5-year prevalent cases were found in Asia[4]. Almost 90% of patients with cancer can be cured by

***Corresponding author:**
***** Waras Nurcholis Email: wnurcholis@apps.ipb.ac.id

surgery if they are still in the early stages[5]. However, most patients can only be diagnosed with colon cancer during the severe phase, when symptoms already appear. Unfortunately, it seems that the current prognosis does not unfold in a manner that was initially anticipated [6].

Cancer treatment includes therapy, surgery, and transplantation. Some cancers require a variety of surgeries and therapies, and there are cancers for which therapy is sufficient as a curative measure. Some types of therapies include chemotherapy using drugs, radiation therapy using high doses of radiation, and photodynamic therapy using light-activated drugs. Neither therapy can differentiate between normal and cancer cells; therefore, normal cells are also affected. In contrast, the shortcomings of photodynamic therapy cannot reach deep cells[7]. Another way to treat cancer is to use miRNAs as tumor suppressor genes or inhibit oncogenic miRNAs. This method has been developed for the treatment of breast, liver, lung, prostate, leukemia, and gastric cancers[8]. Enhancing sensitivity to chemotherapy or other therapeutic medications can be achieved by modifying the expression levels of miRNA[9]

miRNAs were thought to be insignificant until recently, when it was discovered that they were involved in the development of cancer and that their abnormal expression was linked to a variety of clinical events. They were discovered to have significant roles as tumor suppressors via controlling oncogenes or as oncomiRs by blocking several tumor-suppressing genes[10,11]. Our recent study aimed to determine the most influential genes in colon cancer, based on topology and centrality analyses. We investigated dysregulated miRNAs with a high probability of binding to their target proteins. We also analyzed the affinity values of miRNA-mRNA interactions and duplex complexes with argonaute (AGO) proteins. In addition, this study will provide information on the best miRNA candidates for miRNA utilization against colon cancer, and it is hoped that this study will be the first step in determining the analysis of miRNA utilization against colon cancer and the first step in the research process of miRNA-based colon cancer preventive measures.

2. MATERIALS AND METHODS

2.1. Data Collection

Samples were obtained from the GDC Portal database (https://portal.gdc.cancer.gov/) and then the filtering process was carried out on clinical parameters by selecting the TCGA program, Colon Adenocarcinoma (COAD), primary diagnosis including adenocarcinoma, mucinous adenocarcinoma, and papillary adenocarcinoma. In addition, for cancer samples using all classes of cancer development stages. Then in the demographic type parameter, certain racial types were selected in the Asian population, with a total of 11 positive colon cancer patients. On the other hand, as normal control data using RNA-Seq data derived from gene expression quantification data in the white population of 13 people (This was done due to limited samples in the Asian population). Other parameters are set based on dafult. After that, the selection was made from the aspect of data experiment strategy in the form of RNA-Seq (STAR - Counts Workflow and reference genome GRCh38.p0) being read sequences obtained from the Illumina Sequencing platform, using the transcriptome profiling category, and the type of gene expression quantification data. The data obtained was then exported using GDC Data Transfer Tools. The obtained gene data were used for subsequent analyses Ramanto et al[12].

2.2. Differentiation Genes Expression

Analysis of differentiation of expression of genes (DEG) using Empirical Analysis of Digital Gene Expression Data in R (EdgeR) based on[13], using input gene expression data on Asian colon cancer populations and controls with unstranded Transcripts Per Million (TPM) format. The data was then analyzed by making a count table by converting the data to DGElist and removing genes with zero counts in all samples, then the normalization process was carried out using the Trimmed Mean of M-values (TMM) method in overcoming gene expression data bias, followed by analysis of gene variability similarity assumptions using the estimate common and tagwise dispersion method. The next step, an independent filtering process based on the Global Jaccard Index was carried out by removing genes with insignificant expression. Thus, Log Fold Change (LogFC) (Cancer/Normal), P-Value, and False Discovery Rate (FDR) data were obtained. Furthermore, gene expression data was analyzed by volcano plot (LogFC ± 2, -Log(FDR) = 3) where gene data was only taken in conditions of significant up-regulation of expression (LogFC > 2, -Log(FDR) = 3)[14].

2.3. Gene Enrichment Analysis

Enrichment analysis using genes classified in protein-coding genes involved in cancer progression cases furthermore in colon cancer. Analysis using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) databases integrated on STRING-DB webserver (https://string-db.org/) and ShyGO 8.0 (http://bioinformatics.sdstate.edu/go/) by filtering based on hallmarks of colon cancer and phenotypes that correlate with colon cancer conditions[15,16].

2.4. Protein Topology Analysis and Protein Target Centrality

Protein topology was determined using the STRING-DB webserver with the input of 863 up regulate protein coding genes with queries of Homo sapiens species, using a required score of 0.70 and a false discovery rate (FDR) of 5%. STRING data contain nodes (representing proteins) and edges (representing associations between proteins). Then the processed protein gene analysis was carried out by removing proteins that did not have interactions or outliers from the network and continued with centrality analysis. Protein network topology was analysed using Cyctoscape 3.8.3 with the CytoNCA feature to obtain eight centrality parameters, namely eigenvector centrality (EC), betweenness centrality (BC), closeness centrality (CC), degree centrality (DC), information centrality (IC), local average connectivity (LAC), network centrality (NC), and subgraph centrality (SC) (Table 1)[17]. Calculation of overall centrality (OC) using the method (see equations 1 - 8), based on the Gan SL and Djauhari [18] method, principle component analysis (PCA) approach in determining eigenvalue with packages ggplot2, ggrepel, FactoMineR, and factoextra in R Studio. The overall centrality (OC) data obtained is gene sequence data with the most central network (see equations 9),. The data is then filtered based on OC > 0.

$$BC_{(U)} = \sum_{s \neq u \neq t} \frac{\rho(s,u,t)}{\rho(s,t)} \quad (Equations\ 1)$$

ρ(s,t) is the closest distance that connects node p to t, ρ(s,u,t) is the number of paths that pass through node u.

$$DC_{(U)} = |Nu| \quad (Equations\ 2)$$

|Nu| represents the number of nodes neighbouring nodes u

$$CC_{(U)} = \frac{|Nu| - 1}{\sum_{v \in V} d(u,v)} \quad (Equations\ 3)$$

|Nu| represents the number of nodes neighbouring nodes u, d(u,v) represents the closest distance connecting nodes u and v.

$$LAC_{(U)} = \frac{\sum_{\omega \in Nu} deg_{Cu}(\omega)}{|Nu|} \quad (Equations\ 4)$$

Cu is the induced subgraph, $deg_{cu}(\omega)$ is the number of nodes directly connected to Cu.

$$EC_{(U)} = \alpha_{max}(u) \quad (Equations\ 5)$$

The largest eigenvalue of a neighbouring matrix 'A', represented by α_{max} eigenvector amax.

$$SC_{(U)} = \sum_{I=0}^{\infty} \frac{\mu_I(u)}{I!} = \sum_{v=1}^{N} [\alpha_v(u)]^2 e^{\lambda_v} \quad (Equations\ 6)$$

where $\mu_I(u)$ denotes the number of closed loops of length l that start and end at node u. α_v refers to (α_1, α_2, ... , α_N) is the orthonormal basis of R^N composed of the eigenvalue vectors of A associated with the eigenvalues λ_1, λ_2; ...; λ_N , where $\alpha_v(u)$ is the u-th component of v.

$$IC_{(U)} = \left[\frac{1}{N}\sum_{v} \frac{1}{I_{uv}}\right]^{-1} \quad (Equations\ 7)$$

The variables are derived from $I_{UV} = (C_{uu} + C_{vv} - C_{uv})$, where the matrix $C = |D - A + J|^{-1}$, D is the diagonal degree of each node while J is a matrix with all elements equal to 1.

$$ECC_{(u,v)} = \frac{Z_{u,v}}{min(d_u - 1, d_v - 1)}$$

$$NC_{(U)} = \sum_{v \in N_u} ECC\ (u,v) = \sum_{v \in N_u} \frac{Z_{u,v}}{min(d_u - 1, d_v - 1)} \quad (Equations\ 8)$$

We see how many other nodes join nodes u and v for an edge $E_{u,v}$ that connects nodes u and v. The variables $Z_{u,v}$ are the sequence of triangle numbers that span the edge network,du and dv are the degrees of nodes u and v respectively. The value min (d_u-1, d_v-1) is the number of edges at u and v that have the maximum possible participation.

$OC(u) = v_1BC(u) + v_2CC(u) + v_3DC(u) + v_4SC(u) + v_5EC(u) + v_6LAC(u) + v_7IC(u) + v_8NC(u)$

$(Equations\ 9)$

2.5. miRNA – Target Gene Screening

The central gene with the highest overall centrality value was then manually curated based on its role in the pathways of the mechanism of increasing colon cancer malignancy or oncogene. The list of oncogenes obtained was then screened based on the method of Tastsoglou et al[19], against the type of micro-RNA regulating gene through binding to the miRNA Response Element (MRE) region on mRNA using the MicroT

webserver(https://dianalab.ece.uth.gr/microt_webserver/) by integrating the MirGeneDB 2.1 and miRBase v22.1 database sources. in Homo sapiens species and a minimum threshold score > 0.7. A list of miRNAs was selected based on the overlap of results from both databases and used in subsequent analyses.

2.6. Structure Prediction of miRNA and mRNA Interaction Model

Structure prediction analysis of miRNA and gene/mRNA molecular interactions using a webserver (http://huanglab.phys.hust.edu.cn/hnadock/) based on He et al[20] , so as to obtain the structure of the RNA double strand/duplex (miRNA - mRNA) using the input of the mature sequence of the miRNA and the untranslated region-3 (UTR-3) sequence of the mRNA. In addition, a list of nitrogenous base interactions between mRNA and miRNA molecules was inputted.

2.7. Docking of miRNA-mRNA molecules with AGO Protein

The AGO-2 structure preparation began by removing water atoms and aliphatic hydrogen in the structure with Discovery Studio Visualizer 2017 Client software, followed by energy minimization of protein and ligand structures using the YASARA Structure tool to obtain a stable structure with minimal energy[21]. Analisis interaksi dupleks RNA dengan situs pengikatan protein Argonaute-2 (AGO-2) menggunakan webserver HDOCK (http://hdock.phys.hust.edu.cn/) menggunakan input stuktur protein AGO-2 (PDB: 3F73) yang terdeposit pada protein databank (https://www.rcsb.org/).

Molecular docking between ligand (miRNA-mRNA) and AGO protein was performed using the HDOCK webserver with a hybrid algorithm based on template-based modeling and ab initio free docking. The process began by entering the AGO-2 protein and dupleks RNA (miRNA – mRNA) as a ligand with file with ekstensi *. pdb. In addition, the binding sites were determined for the active site of protein AGO2 at amino acid sequences Tyr171, Arg172, Leu173, Arg194, Thr201, Leu281, Asp478, Asp546, Arg548, Arg580, Asp660, dan Leu279[22].

Selain itu, kami juga mempertimbangkan untuk melakukan analisis interaksi pada residu asam amino protein AGO2 yang berinteraksi pada ligand co-kristal yaitu Gln48, Arg52, Try17, Arg172, Ile173, Arg192, Arg194, Tyr197, Arg200, Leu217, Pro218, Tyr226, His227, Arg232, Pro255, His256, Leu265, Leu267, Leu281, Arg286, Pro412, Met413, Arg418, Lys422, Ile434, Asn436, Arg444, Arg446, Asn449, Lys457, Asp478, Asp546, Arg548, Lys575, Ser576, Arg580, Thr613, Arg615, Pro650, Arg651, Asp660, dan Arg661, and the binding sites on ligands in nucleotide sequence. The results of molecular docking are in the form of docking parameters, including docking score, confidence score, ligand RMSD, and duplex miRNA-mRNA complex with AGO protein in *. pdb extension[23].

2.7.1. Interaction Analysis and Visualisation

Structural interaction analysis was performed using Discovery Studio Visualizer 2017 Client software to determine the physical and electrostatic bonding parameters. Structure visualization using Discovery Studio and analisis interaksi menggunakan webserver PDBSum (https://www.ebi.ac.uk/thornton-srv/databases/pdbsum/) then visualize the Arc diagram using the tidyverse, ggraph, and igraph packages in R-Studio[24,25].

2.8. Molecular Dynamics Simulation

Preparation of molecular dynamics based on[26], by setting several parameters in the md_run.mcr file. Simulation parameters used were a temperature of 310 K dengan konsekuensi dilakukan aktivasi Manometer1D pada pressure control, pH 7.4, H_2O solvent, 0.9% NaCl (standard physiological solution), cell shape is a cube with size 10 extension. ForceField was set using AMBER14 with Cutoff of 8, and system set with the model cell boundary periodic. In addition, the simulation speed was set with normal speed (maximize performance with 2 x 2.5 fs timestep and constraints), simulation time by setting the duration with duration 30000 ps or 30 ns, and using snapshot interval with command for fast speed, snapshots are taken every 250.000 fs.

2.8.1. Molecular Dynamics Trajectory Analysis

The trajectory analysis was conducted based on methods described in[26,27]. Several parameters were set in the md_analysis.mcr macro file using the default *.sim trajectory files generated by the YASARA software. The analysis of molecular dynamics (MD) trajectories included calculating root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration (Rg), solvent-accessible surface area (SASA), and solute-solvent hydrogen bonds.

Additionally, to examine conformational changes in the AGO-2 protein structure, principal component analysis (PCA) and dynamic cross-correlation matrix (DCCM) analyses were performed. The *.sim trajectories were first converted to *.xtc format (GROMACS trajectory) using commands within the md_convert.mcr macro. Subsequently, the *.xtc files were converted into *.dcd format (compatible with NAMD/CHARMM) using Visual Molecular Dynamics (VMD) software. PCA and DCCM analyses were conducted using the bio3d package in RStudio, with *.dcd and *.pdb files as inputs.

3. RESULTS

3.1. Significant Protein Screening in Colon Cancer

There are 60660 genes covering various types of genes (coding and regulatory) derived from each sample of patients with colon cancer (TCGA-COAD) totaling 11 patients in the Asian population and 17 normal patients as a comparison taken from the TCGA database. Using differentiation expression genes (DEGs) analysis with the EdgeR analysis method, the process of removing genes with low expression and zero expression was carried out, resulting in 18292 genes. Furthermore, only genes that encode protein coding were taken, resulting in 13876 genes. The obtained genes were then subjected to further sorting analysis using volcano plot analysis by setting the expression threshold of Log10(FC) > 2 (Up-regulated) and the significance of -Log10(FDR) < 3. The data obtained in the form of significant gene upregulation data on the colon cancer phenotype were 863 genes (Figure 1). Research states that genes that are significantly increased in malignancy conditions have a major role in the growth and development of cancerous tissue.

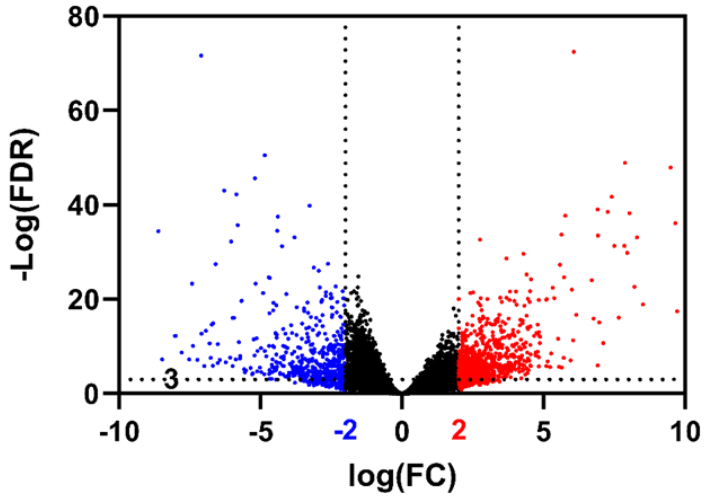

Figure 1. Volcano Plot analysis in assessing significance related to colon cancer gene expression compared to normal conditions. Blue plots indicate down-regulated expression and red plots indicate up-regulation, expression thresholds are set at Log10FC -2 and 2 and significance in -LogFDR > 3 (0.05).

Enrichment analysis was used to validate the role of upregulated genes in the developmental phenotype of colon tumors (Figure 2(a-c)), where patients with colon cancer phenotypes showed increased expression of genes that directly or indirectly contribute to colon cancer proliferation, such as increased activity of cellular signaling cascades and cellular proliferation, increased activity of cellular morphogenesis and differentiation, and increased activity of carbohydrate metabolism. Furthermore, some genes are also responsible for the mechanism of tumor development and progression through the organization of extracellular matrix components, which can affect the formation of cancer formation, the degree of invasion and migration, and treatment resistance.

Based on the association of these genes to roles in colon cancer-related pathways such as several types of complex signaling pathways such as PPAR, AMPK, GPCR, and PI3K-AKT. These pathways are responsible for the regulation of proliferation, apoptosis, metabolism, and migration of cancer cells. The family of nuclear receptors known as PPARs (Peroxisome Proliferator-Activated Receptors) regulate the expression of genes related to cell proliferation, lipid metabolism, and differentiation. The PPAR-γ subtype, for example, has the ability to inhibit colon cancer cell proliferation and promote differentiation and apoptosis, while PPAR-δ also has the ability to act as an oncogene by promoting cancer cell proliferation. AMPK (AMP-Activated Protein Kinase) activation is a cellular energy sensor that regulates energy homeostasis and acts as a tumor inhibitor in colon cancer by inhibiting the mTOR pathway, which is essential for cell growth. The antitumor effects of AMPK are demonstrated by reducing cell proliferation and increasing apoptosis.

A family of membrane receptors known as G Protein-Coupled Receptors (GPCRs) regulate various extracellular signals, including hormones and growth factors. Activation of GPCRs can enhance signaling pathways that promote tumorigenesis, such as PI3K-AKT and MAPK, and regulate the tumor microenvironment to support tumor growth and metastasis. Some GPCRs, such as prostaglandin E2 receptor (EP2), are involved in the proliferation, migration, and angiogenesis of colon cancer cells. The PI3K-AKT pathway, also known as Phosphatidylinositol 3-Kinase/Protein Kinase B, is a major signaling pathway that regulates cell growth, proliferation, motility, survival, and metabolism. Colon cancer often causes mutations and amplification

in components of the PI3K-AKT pathway, leading to increased AKT activity and promoting tumor cell growth and survival. PI3K and AKT inhibitors are being developed and clinically tested as potential therapies to halt the development and progression of colon cancer. The PPAR, AMPK, GPCR, and PI3K-AKT pathways are interconnected and form a complex signaling network. For example, AMPK activation can inhibit the PI3K-AKT pathway by suppressing mTOR, while GPCRs can activate PI3K-AKT and affect lipid metabolism through PPARs. Better mapping and understanding of these interactions may provide important insights for more effective therapeutic approaches to colon cancer.

Pathways such as L1 Cell Adhesion Molecule (L1CAM) and Neural Cell Adhesion Molecule 1 (NCAM1) are pathways initiating the metastatic process in colon cancer that correlate with the regeneration process of colon epithelial tissue Although L1CAM is not an expression product of colon epithelial tissue, L1CAM has a crucial role in regeneration after colitis, appendicitis, and organoid formation in colon cancer so that it has a critical role in enhancing invasion and migration mechanisms in other tissues[28].

There is also an increase in pathways associated with membrane receptors such as P2Y and extracellular matrix organization that contribute to tumor progression through the activation of various ligand molecules such as chemokines and interleukins. Furthermore, pathways associated with transport mechanisms such as SLC mediated transmembrane transport can increase the viability of cancer cells by increasing the supply of macronutrients (glucose, fatty acids, amino acids, and nucleic acids) and micronutrients (ions, vitamins, and metabolites), which are often used in direct and indirect cancer treatment targets[29].

The mechanism of carcinogenic compound induction consists of several stages: initiation, promotion, and progression. At each of these stages, substances or environmental conditions have the potential to initiate and/or accelerate the carcinogenic process. Environmental stimulation of intracellular and extracellular receptors is one of the non-genotoxic methods of carcinogenic compounds in the process of cancer development.

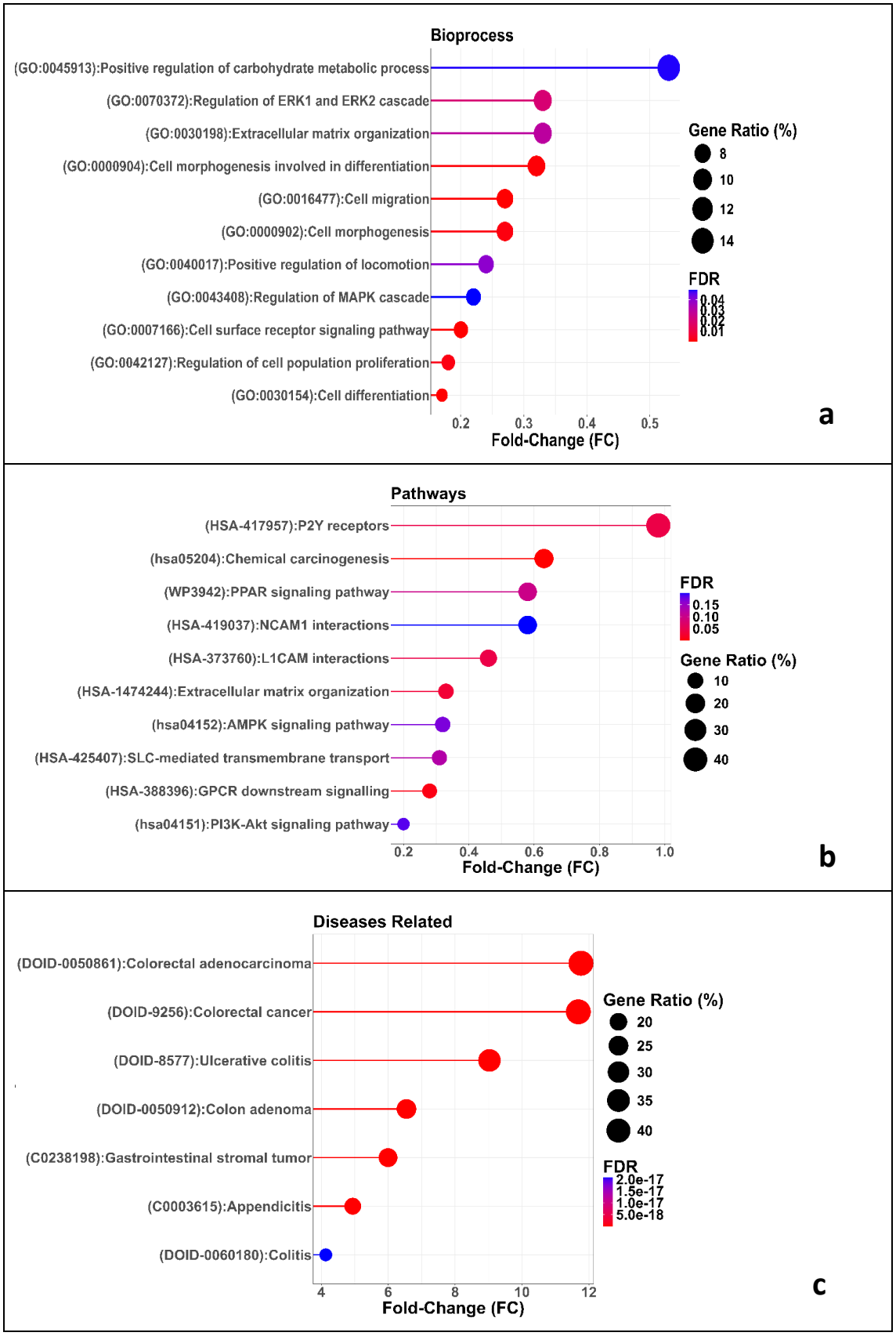

Figure 2. Dot plot of up-regulate gene enrichment in colon cancer incidence of Asian populations**.**

3.2. Gene Network Associate with Colon Cancer

A total of 863 genes that experienced significant upregulation were analyzed using network pharmacology by assessing protein coding genes (nodes) and interactions (edges). So that it can be used in determining the central role of the type of protein encoding gene in a system (Figure 3). Gene with high overall centrality indicates that the gene has a crucial role in a system, where damage or decreased expression in the gene will cause alteration of the network[30].

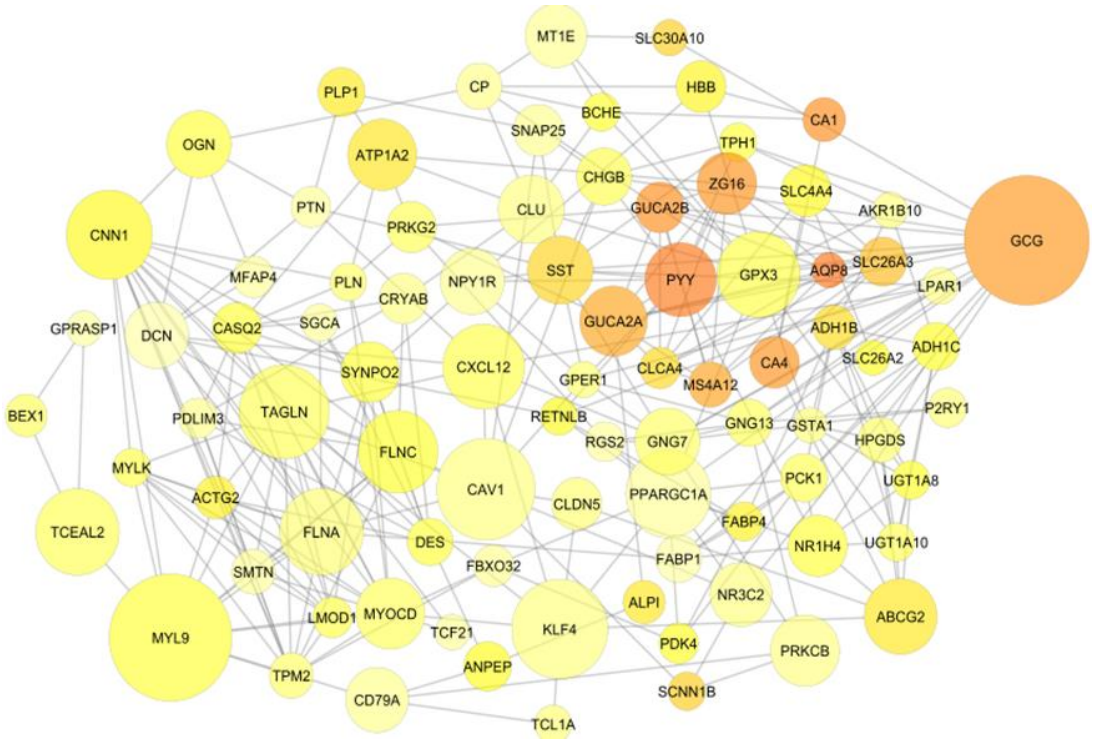

Figure 3. Gene network associated with colon cancer

Based on the most central colon cancer upregulation genes curated for their activity in regulating cancer, the genes obtained were filtered by selecting genes that have potential activity as oncogenes, resulting in 20 central oncogenes associated with colon cancer pathogenicity with high centrality (Table 1). Genes with high centrality indicate their crucial role in the pathogenicity system of colon cancer, where there is an assumption that if there is an abnormality in gene expression, it will result in pathway alteration[30]. In addition, gene centrality assessment is often used in the selection of potential multi-targets of a disease especially diseases with complex signaling such as cancer in specific tissues[31].

Table 1. 20 Central ongenes associated with colon cancer of Asian populations.

Genes	Betweenness	Closeness	Degree	Eigenvector	Information	LAC	Network	Subgragh	Overall
CAV1	3.64	0.64	1.88	1.24	1.49	-0.22	0.21	0.39	21.84
KLF4	3.58	0.64	1.64	0.78	1.40	-0.48	-0.54	0.07	21.08
TAGLN	2.51	0.64	2.85	4.71	1.75	2.83	3.13	5.57	19.01
FLNA	2.17	0.62	2.61	4.19	1.70	1.98	2.12	4.45	16.68
GPX3	2.83	0.62	1.40	-0.17	1.31	0.24	0.35	-0.07	16.35
CXCL12	2.61	0.68	1.64	0.71	1.40	0.14	0.34	0.12	15.99
ABCG2	2.15	0.54	1.64	-0.11	1.40	0.14	1.06	-0.01	12.96
NR1H4	1.18	0.56	2.12	0.04	1.57	0.97	2.07	0.35	8.42
SLC4A4	0.91	0.35	0.68	-0.37	0.93	-0.11	-0.12	-0.28	5.37
CASQ2	0.51	0.42	1.16	1.66	1.20	0.23	0.06	0.68	5.09
CRYAB	0.63	0.43	0.92	0.70	1.07	0.07	0.05	0.00	4.94
SLC26A3	0.64	0.30	1.40	-0.32	1.31	1.25	1.12	-0.05	4.50
FABP1	0.57	0.50	0.43	0.17	0.76	0.53	0.25	-0.10	3.99
CP	0.64	0.44	0.19	-0.29	0.56	0.35	0.09	-0.30	3.77

3.3. miRNAs Targeting Genes Associated in Colon Cancer

MicroRNAs, or miRNAs, are small non-coding RNA molecules that regulate gene expression by targeting mRNAs for degradation or inhibiting translation, thereby affecting biological processes such as cellular differentiation, development, and proliferation[32]. In cancer, miRNAs can function as oncogene suppressors (onco-miRNAs), thus controlling the expression of genes involved in processes such as cell proliferation, apoptosis, differentiation, and metastasis in cancer. On the other hand, there are also miRNAs that suppress the expression of tumor suppressor genes (suppressor-miRNAs), thus implicated in increased pathogenicity of various cancers and increased activity of opposing oncogenic genes[33]. However, some findings suggest alteration of miRNA expression in various cancers and decreased expression of suppressor-miRNAs or increased onco-miRNAs has an important role in cancer initiation and progression[34]. Alterations in miRNAs are caused by various internal and external factors, but are dominated by external epigenetic factors such as exposure to carcinogenic substances and oxidative stress. Epigenetic alterations can occur, among others, due to the induction of carcinogenic compounds and free radicals on histone regulation and changes in methylation patterns on DNA[35].

In this study, we selected miRNAs regulating genes (124 miRNAs) that overlap in both miRBase and miRGene databases for all selected oncogenes (Figure 4). In the development mechanism of colon cancer pathogenicity, a gene tends to be regulated by more than one type of miRNA molecule (Table 2). The CAV1 gene has multi-complementary with many miRNAs on its UTR-3 side with the best interaction in the 8-mer form with hsa-miR-651-3p with interaction score 0.97. Gene KLF4 has the best in the 9-mer form interaction with hsa-miR-32-5p with interaction score 0.99. TAGLN gene has the best in the 8-mer form interaction with hsa-miR-185-5p with interaction score 0.81. Where the more the interaction score approaches 1.0, the more valid the interaction [19], Where the assessment of the interaction between miRNA and mRNA is influenced by various factors such as the physical nature of the bond such as the type of hydrogen bond formed, the complementary base sequence (k-mer), and the presence of base insertion (mismatch) (Table 3). The thermodynamic assessment aspect of the role of free energy (ΔG) is related to the spontaneity of miRNA-mRNA interactions related to aspects of bond stability, where increasingly negative free energy indicates an increase in interaction spontaneity, in addition to the Root Mean Square Deviation (RMSD) value indicates the deviation of the initial position of the three-dimensional interaction spontaneity on the RNA duplex (miRNA - mRNA).

In this study, the best type of miRNA-mRNA interaction based on the type of seed k-mer is by selecting 8-mer (an exact match to positions 2-8 of the mature miRNA (the seed + position 8) followed by an 'A') and 9-mer (The 9-mer miRNA seed, specifically nucleotides 3-9), based on increased target recognition with high base pairing, where the k-mer is sorted based on binding efficacy, namely 8mer> 7mer> 7mer-A1> 6mer[36,37].

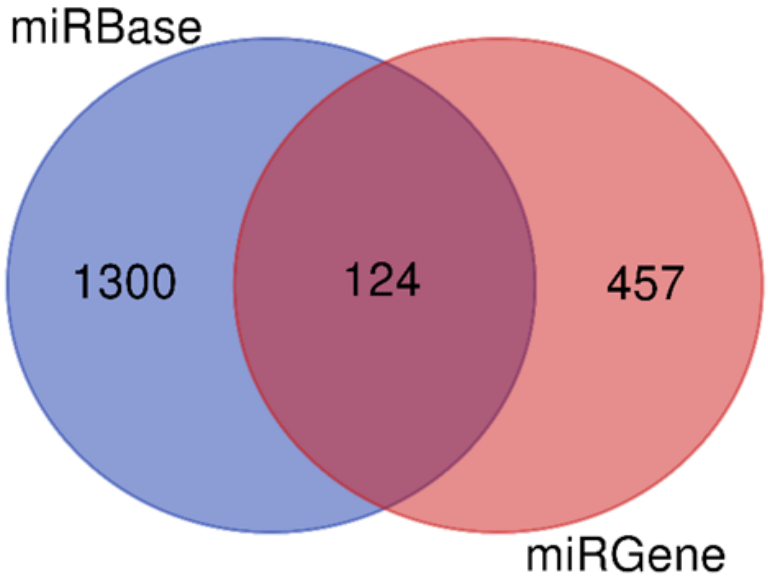

Figure 4. Venn diagram of the number of miRNAs that interact with oncogenes in colon cancer.

Table 2. Multi target miRNAs.

Gene Symbol	Gene Esembl ID	miRNAs name	Interaction score
CAV1	ENSG00000105974	hsa-miR-651-3p	0.97
		hsa-miR-1911-3p	0.96
		hsa-miR-1910-3p	0.91
TAGLN	ENSG00000149591	hsa-miR-185-5p	0.81
		hsa-miR-3173-5p	0.76
		hsa-miR-1225-3p	0.76
KLF4	ENSG00000136826	hsa-miR-32-5p	0.99
		hsa-miR-4677-5p	0.96
		hsa-miR-556-3p	0.90

Table 3. Potential interactions of oncogenes and miRNAs.

mRNA-miRNA	MRE Binding Area	Binding type	Docking Score	Ligand RMSD (Å)
CAV1 - hsa-miR-651-3p	5'-AUAGA U CUGAUUUU AG AUAC UUUUCCUUU-3' transcript UC UAUG GAAAGGAAA-5' miRNA 3'- A C U	8-mer	-177.74	358.28
TAGLN - hsa-miR-185-5p	5'-UCCAAAG A UC C-3' transcript CAG G UUGCUUUUCUCUCC GUC C GACGGAAAGAGAGG 3'- A UU U-5' miRNA	8-mer	-191.45	57.32
KLF4 - hsa-miR-32-5p	5'-UUA G UUCA UGC ACU UGGU GAUGUGCAAUA-3' transcript 3'- ACG UGA AUCA UUACACGUUAU-5' miRNA U	9-mer	-158.07	109.07

3.4. Molecular Interaction of miRNA-mRNA Complexes with AGO Enzymes

All the three-duplex miRNA-mRNA complexes with AGO-2 protein (Figure 6) showed docking parameter results (Table 3), with docking scores ranging from -197.65 to 248.13 (kcal/mol), with confidence values >0.7, and RMSD below 100 Å for all complexes except the TAGLN - hsa-miR-185-5p complex. 7, and RMSD below 100 Å for all complexes except the TAGLN-hsa-miR-185-5p complex which was slightly above 100 Å. HDOCK used hybrid docking approach that combines template-based modeling and free docking was used to determine the RMSD values that were retrieved from the HDOCK server. The accuracy of the predictions can be improved by using this technique to incorporate experimental data on small-angle X-ray scattering and the protein-protein binding site during the docking process. A positive docking score resulted in a more plausible binding model and negative score suggested no link between the ligand and protein[23,38]. Besides that, protein-RNA complexes typically have a docking score of 200 (kcal/mol) or higher[38]. The root-mean-square deviation (RMSD) of the ligand was calculated by comparing the coordinate of ligand in the docking model with the initial coordinate input or modeled structure. A smaller RMSD indicates that the complex is close to the experimental state[39].

The amino acid residue interaction data for the duplex miRNA-mRNA complexes are shown in (Figure 5). The results show that only the CAV1-hsa-miR-651-3p and KLF4-hsa-miR-32-5p complexes have direct interactions with the catalytic active side of the AGO-2 protein, where the most interactions are generated in the AGO-2 complex with CAV1-hsa-miR-651-3p, resulting in good catalytic stability in the AGO-2 receptor. Furthermore, it was found that on the catalytically active side of AGO-2 there was an increase in hydrophobic amino acids compared to other parts of the protein (Figure 6). This can be attributed to several factors such as increased thermal stability of the protein-nucleic acid complex due to hydrophobic interactions, especially in the hydrophobic core that can protect the complex from the aqueous environment[40]. Hydrophobic amino acids are essential for DNA and RNA recognition in the interaction of nucleic acids with hydrophobic residues in proteins such as Argonaute (AGO). These amino acids often interact through van der waal forces that expand on the DNA molecule, causing DNA-binding proteins to recognize certain sequences[41]. In protein-DNA complexes, hydrophobic interactions that often involve contacts between bases on nucleic acids and hydrophobic amino acids are significant. Furthermore, hydrophobic side chains on certain DNA-binding proteins insert between DNA base pairs, resulting in kinks or curvatures in the DNA structure[42].

Table 4. Docking results between AGO-2 and RNA duplex (miRNA-mRNA).

Receptor	miRNA-mRNA complex	Docking Scores (kcal/mol)	RMSD (Å)
AGO-2 (PDB : 3F73)	Co-crystals	-197.65	24.67
	CAV1 - hsa-miR-651-3p	-248.13	76.39
	TAGLN - hsa-miR-185-5p	-206.84	102.74
	KLF4 - hsa-miR-32-5p	-223.3	78.19

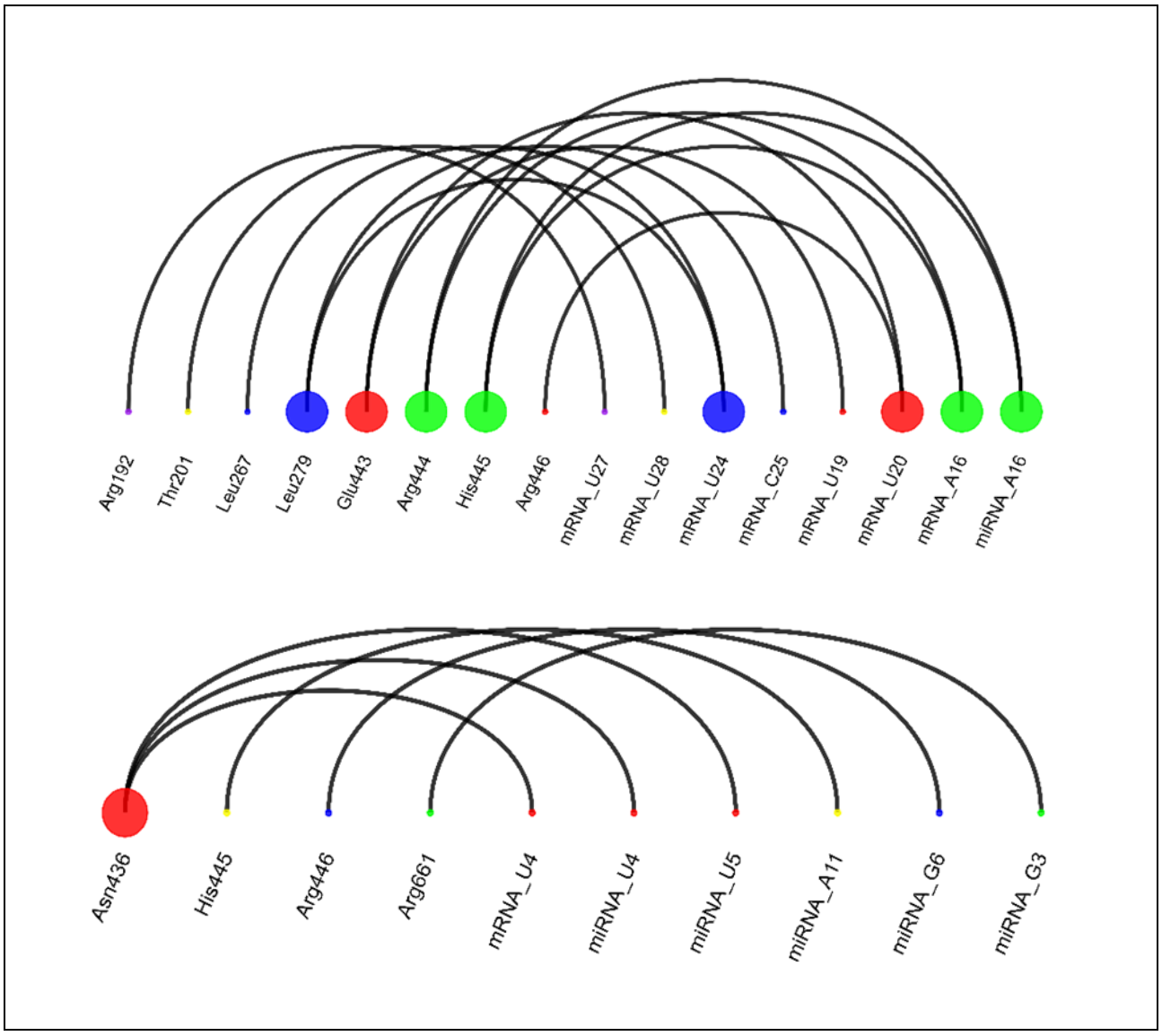

Figure 5. Arc plot of RNA duplex interaction (mRNA-miRNA) with the catalytic active side on Argonaute-2 Protein (AGO-2) against the complexes of (a) CAV1 - hsa-miR-651-3p, (b) KLF4 - hsa-miR-32-5p.

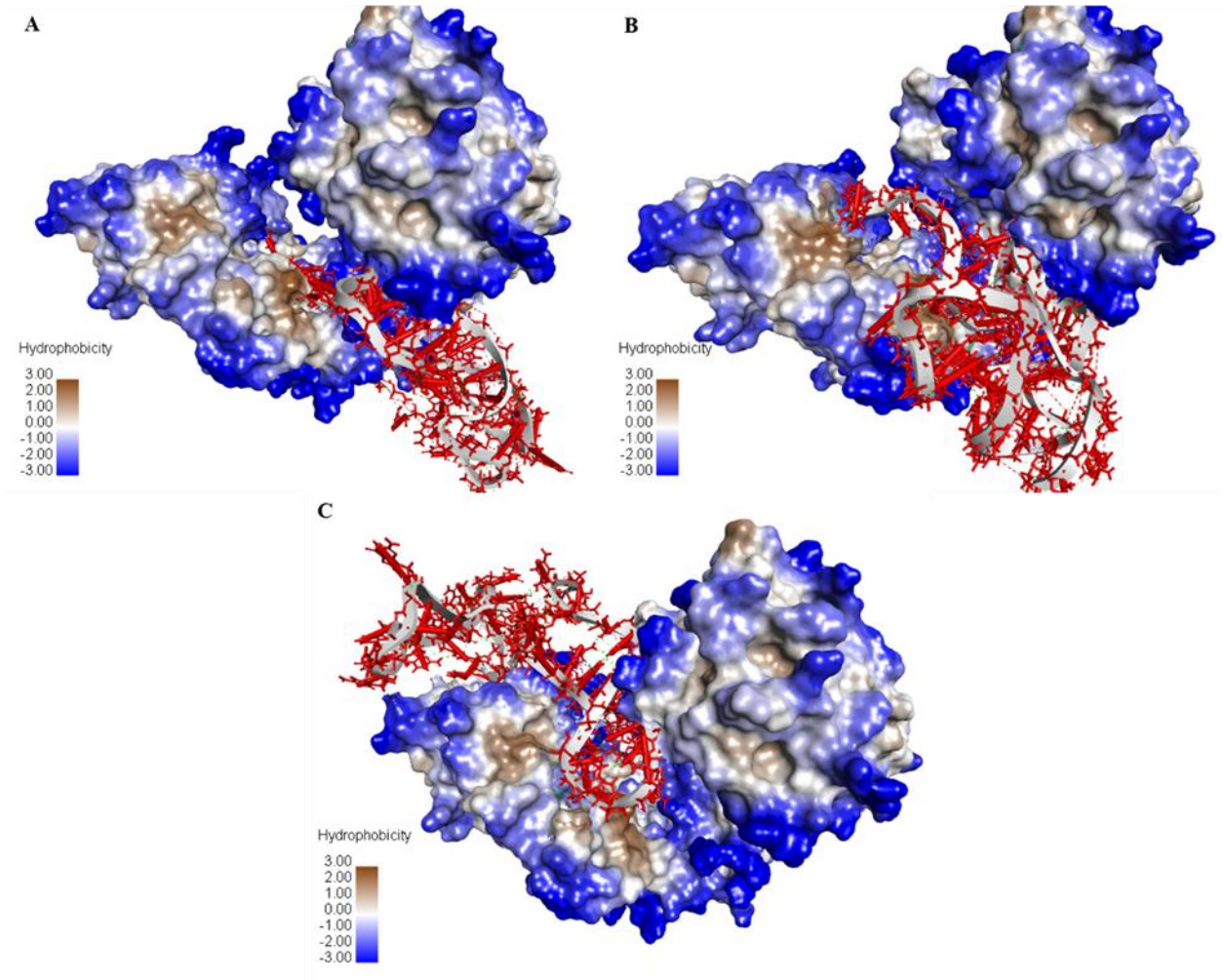

Figure 6. A closer look into the protein-molecule complex. A) AGO2-CAV1-miR-651-3p. B) AGO2-KLF4-miR-32-5p. C) AGO2-TAGLN-miR-185-5p. Structure using accessible surface area model with varying levels of surface hydrophobicity.

3.5. Molecular Dynamics Stability of miRNA-miRNA Complexes with AGO Enzymes

This study analyzed molecular dynamics (MD) trajectories by looking at RMSD, RMSF, Rg, SASA, and h-bond parameters. RMSD was calculated to determine the conformational changes of AGO2, AGO2-CAV1-miR-651-3p, AGO2-KLF4-miR-32-5p, and AGO2-TAGLN-miR-185-5p. Based on the analysis, the AGO2-CAV1-miR-651-3p complex starts the trajectory with a value of 0.438 Å, the AGO2-KLF4-miR-32-5p complex starts the trajectory with a value of 0.428 Å, and the AGO2-TAGLN-miR-185-5p complex starts the trajectory with a value of 0.414 Å. The start of the trajectory of each complex shows a good start. However, this AGO2-CAV1-miR-651-3p complex touched a value of more than 5 Å, for the first time on a 7 ns trajectory. After that trajectory, fluctuations between 3-5 Å occurred continuously until the 30 ns trajectory ended. complex AGO2-CAV1-miR-651-3p had an average RMSD of 4.63 Å. The AGO2-TAGLN-miR-185-5p complex has fluctuations ranging from 1.7 to 4.8 Å, with an average over the 30 ns trajectory of 3.1 Å. On the other hand, after a good start, the AGO2-KLF4-miR-32-5p complex has a fluctuation range between 1.85 to 3.5 Å with an average RMSD value over a 30 ns trajectory of 2.74 Å. This complex gives the most minimal graph when compared to the other two complexes and even better when compared to the MD confromation of AGO2 protein. AGO2 is an individual protein that is not given any molecules, this is to check what the protein looks like alone in a 30 ns trajectory in MD simulations and to compare the values of all parameters analyzed, whether it gets better when there is a molecule or the higher the fluctuation. Thus, when the RMSD value is smaller, it indicates good stability. The correlation of events in this study is evident in that the AGO2-KLF4-miR-32-5p complex has good stability and the conformational changes are not large. All 30 ns trajectory graphs with RMSD analysis are presented in Figure 7A.

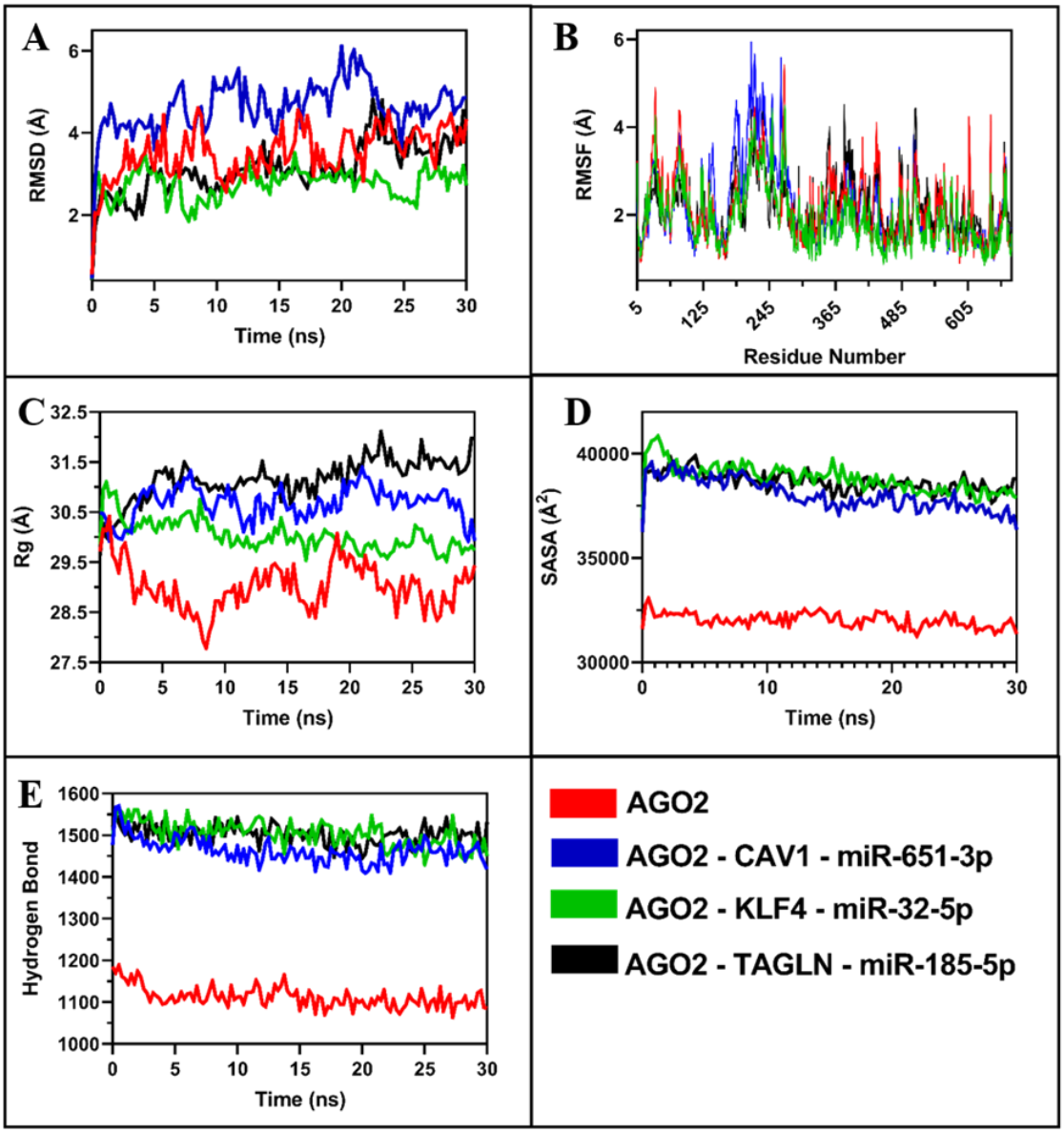

Figure 7. Analysis of MD simulation with 30 ns trajectory. A) RMSD. B) RMSF. C) Rg. D) SASA. E) H-bonds.

We evaluated the flexibility of CAV1-miR-651-3p, KLF4miR-32-5p, and TAGLN-miR-185-5p toward AGO2 by considering their binding environments and taking into account their interaction values with amino acid residues during a 30 ns trajectory. Therefore, we used root-mean-square fluctuations (RMSF) to consider the occurrence of complex fluctuations in this study. The analysis results show that there are high fluctuations between residue positions 125 to 305. In fact, the prediction in the area of residues 185 to 260 fluctuations resulted in RMSF values that touched more than 4 Å. Nevertheless, it was observed that the AGO2-KLF4-miR-32-5p complex gave the feel of a more stable complex when compared to the others. We indicated this from the graphs produced in Figure 7B.

Our conclusion in MD analysis, through RMSD and RMSF analysis and visual observation which can be observed in (Figure 7), the most stable conformational complex during MD simulation is AGO2-KLF4-miR-32-5p complex. Protein compactness and stability were assessed by radius of gyration (RG). We observe again that individual AGO2 has the lowest value when compared to AGO2 that has CAV1-miR-651-3p, KLF4-miR-32-5p, and TAGLN-miR-185-5p molecules. From the point of view of Rg, it is interesting to observe through the graph that has been produced, the KLF4-miR-32-5p molecule gives hope that there will be the smallest and minimal conformational changes to AGO2, when compared to the CAV1-miR-651-3p and TAGLN-miR-185-5p molecules. we present the graph in Figure 7C.

We evaluated the complexes by considering the solvent-accessible protein surface (SASA) parameter because this parameter gives us insight into whether CAV1-miR-651-3p, KLF4miR-32-5p, and TAGLN-miR-185-5p are retained within the AGO2 binding pocket or are removed from the binding cavity during the 30 ns trajectory. We looked at the SASA values of individual AGO2, it was observed that individual AGO2 had much lower values than the other complexes. This result is very justified because there are no molecules that bind to AGO2, so it is natural to have low values in the SASA analysis. Therefore, if the SASA value is lower, there is good binding to CAV1-miR-651-3p, KLF4-miR-32-5p, and TAGLN-miR-185-5p. From the SASA point of view, the AGO2-CAV1-miR-651-3p complex is more minimal in that the CAV1-miR-651-3p molecule is ejected from the molecular binding site in the AGO2 cleft. We present the SASA graph in Figure 7D.

We perform PCA for multivariate analysis and can be used to evaluate the conformational stability of structures and even to analyze the thermodynamic stability of protein complexes during certain trajectories in MD simulations (Figure 8). On the other hand, we evaluated the conformational stability factor using PCA by calculating the distribution of Cα carbon movement in the protein. We again used individual AGO2 as a comparison of conformational assessment with AGO2 bound to CAV1-miR-651-3p, KLF4miR-32-5p, and TAGLN-miR-185-5p. PCA results show that the score plot in red indicates a conformation that has reached stability. The results of the PC2 calculation which has the lowest value is the AGO2-KLF4miR-32-5p complex with a PC2 value of 16.1%. all complete data we present in Figure 3. the value is smaller than the other complexes and even close to the PC2 value on individual AGO2. The smaller the variation value, the more rigid and more stable the conformation. The result of PCA analysis correlates with the analysis in MD parameters, that AGO2- KLF4-miR-32-5p complex has good conformational stability.

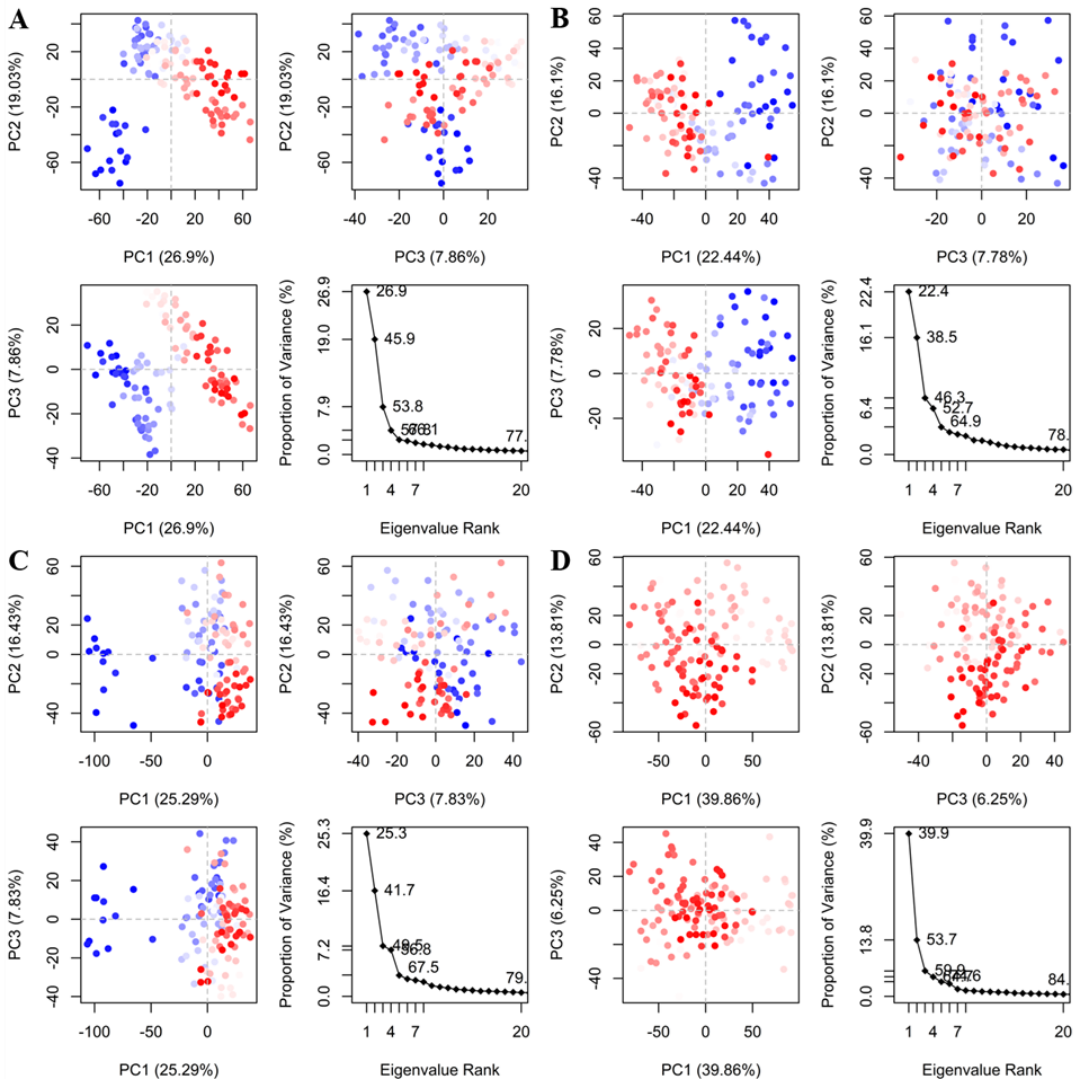

Figure 8. Principle component analysis (PCA) of system conformational changes versus time for the three complexes and one individual AGO2 protein. A) AGO2-TAGLN-miR-185-5p. B) AGO2-KLF4-miR-32-5p. C) AGO2-CAV1-miR-651-3p. D) AGO2 (Apo).

We consider that when there is movement and fluctuation of Cα atoms, it is indicated that changes occur and affect the functionality of a protein. Therefore, DCCM is present to assess and evaluate the movement changes in Cα. In the explanation of the DCCM graph as we present in (Figure 9), it can be seen that there are numbers 1 and -1, meaning that if the value touches the number 1 then the correlation is positive (parallel movement), while the value that touches the number -1 then there is an anticorrelation (antiparallel movement). In other words, when the DCCM result is redder, it is predicted that the confromation is more stable and the result correlates with the RMSD parameter in the evaluation of MD simulation results. From the DCCM results, we can provide an understanding that conformations with positively correlated parallel movements occur in the AGO2-KLF4miR-32-5p complex. The DCCM plot of the complex is close to the similarity with the DCCM plot of individual AGO2. Thus, it has been confirmed that the AGO2- KLF4miR-32-5p complex is a complex with minimal non-binding interactions (white in color).

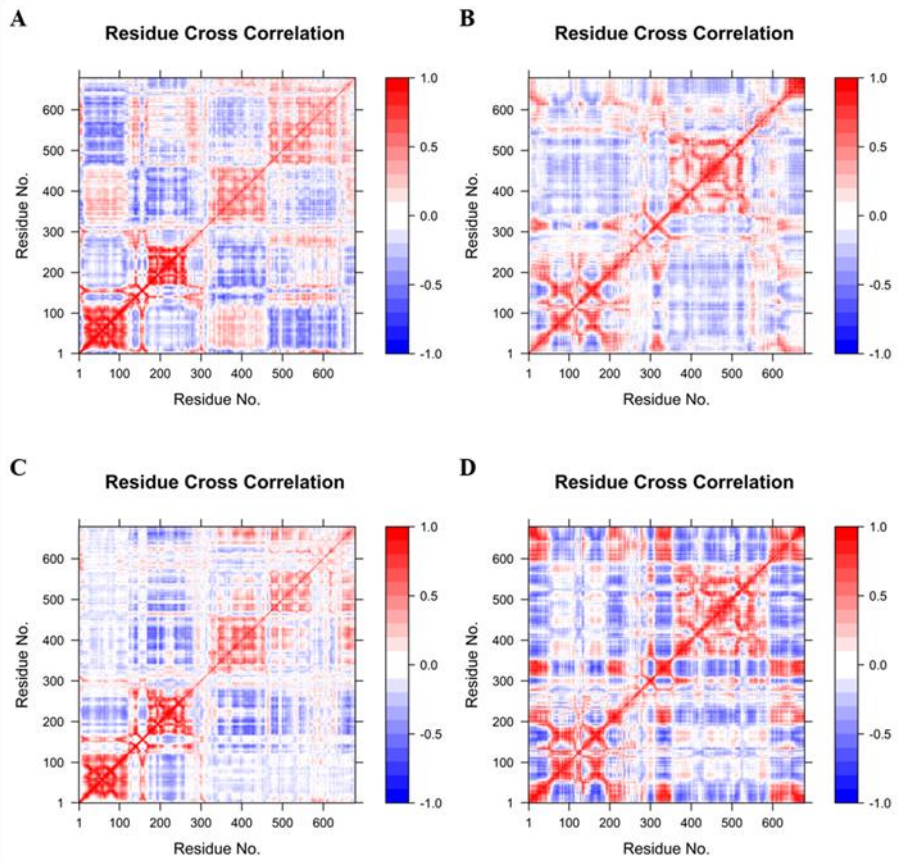

Figure 9. Calculation analysis of dynamic cross-correlation matrix of Cα atoms in three AGO2 complexes and one individual AGO2 in MD simulation for 30 ns. A) AGO2-TAGLN-miR-185-5p. B) AGO2-KLF4-miR-32-5p. C) AGO2-CAV1-miR-651-3p. D) Individual AGO2 protein.

4. DISCUSSION

Micro-RNA (miRNA)-based therapeutic agents are a breakthrough in the field of pharmacology in the treatment of various malignancies, where miRNAs can efficiently act on various types of targets, especially transcription factors that are more difficult to access by conventional drugs[43]. This study found that several transcription factors contribute to the pathogenicity of colon cancer in Asian populations.

The CAV1 gene encodes the caveolin-1 protein, a transcription factor essential for maintaining the invaginated structure of caveolae often found in fibroblasts, adipocytes, smooth muscle cells, and vascular epithelial cells[44]. CAV1 is a protein that has a complex system which under certain circumstances can play a role in tumor suppression in several types of cancer such as in melanoma by repressing metastatic and migratory conditions by regulating the Integrin / Src / FAK Signaling Pathway complex, CAV1 also acts as a tumor suppressor via c-Myc-mediated metabolic reprogramming[45]. On the other hand, in other cancer conditions CAV1 is correlated in the activation of cancer progression of Ewing's sarcoma family tumors (ESFT) and correlated to the worst prognosis in prostate cancer[46,47]. P Studies such as He et al[48], mentioned that overexpression of CAV1 and CAV2 in squamous cancer leads to poor prognosis. CAV1 gene expression

has a function in regulating cell differentiation by regulating cyclin D1 complex and ERK-1/2 signaling complex. In the progressivity of colon cancer, CAV1 will be degraded through CMA by HSPA8, CAV1 contributes to the development of colon cancer by releasing catenin into the nucleus to activate Wnt/catenin signaling, which causes metastasis in BRAF V600E CRC[49]. Previous studies have shown an association between treatment resistance and increased CAV1 expression in colon cancer cells, leading to increased expression and metastasis rates[50]. In addition, the abnormal increase of CAV1 is highly correlated with hypomethylation of promoter CpG sites in colon cancer conditions. In the event of decreased CAV1 expression, it will increase the activity of AMPK which will directly restrain the cell cycle and activate the autophagy process. On the other hand, overexpression of CAV1 leads to increased cellular glucose consumption and increased glycolysis by stimulating GLUT3 transcription through the HMGA1 binding site in the GLUT3 promoter[51]. This is in line with the incidence of metabolic abnormalities such as hyperglycemia, hyperinsulinemia, and adipokine imbalance associated with the incidence of prediabetes and colorectal cancer (CRC) and supports the onset and progression of CRC, and changes in glycolysis play an important role in cancer development[52]. The CAV1 gene, which is associated with carbohydrate metabolism syndrome and susceptibility to several cancers in Asian populations particularly in Arabs, East Asians, and South Asians[53]. In addition, the CAV1 G14713A polymorphism is associated with cancer risk in Asians, especially digestive system-related cancers[54].

KLF4 is also a zinc finger transcription factor that plays an important role in the development of a number of cancers, such as lung, breast, colorectal and pancreatic cancer. However, it is not known whether KLF4 suppresses or favors cancer. KLF4 dual effects in cancer usually depend on the type of cancer and molecular regulation of signaling pathways[55]. KLF4 functions as a tumor suppressor for cutaneous squamous cell carcinoma and gastrointestinal cancer, but also functions as a pro-cancer factor for cutaneous melanoma, breast cancer, and osteosarcoma pathogenesis[56]. The effects of KLF4 also differ based on the stage of cancer development. KLF4 protein is decreased in human pancreatic cancer and human pancreatic cancer cell lines, but increased in early lesions such as ADM and PanIN and promotes the formation of precancerous lesions in mouse models[55,57]. In relation to the pathogenesis of colon cancer, KLF4 has a role in the development of colon cancer where knockdown can significantly reduce the potential for cancer tissue development in the colon, prevent the process of migration and invasion, and reduce treatment resistance[58]. Based on the results of Taracha-Wisniewska's research[59], mentioned that mutations in KLF4 p.A472D can cause resistance to cetuximab, which targets the extracellular domain of the epidermal growth factor receptor (EGFR) and is useful in treating patients with metastatic colorectal cancer who have RAS wild-type and BRAF V600E wild-type. Furthermore, recent studies have revealed that adherent cells contain higher levels of KLF4. KLF4 is forced to be co-expressed with OCT3/4, SOX2, and c-MYC, which enhances cancer cell stem properties including tumorigenicity, spheroid formation, and resistance to chemotherapy[60]. Mesenchymal phenotype and increased stemness caused by TGFβ-1 pathway activated by KLF4 overexpression in certain CSCs[61]. The WNT/β-catenin pathway was responsible for the observed increased levels of KLF4 and other stem markers in CD133+ CRC stem cells. Moreover, KLF4 levels were also increased even in CD133 metastatic cells[62]. Tumors from metastatic patients had higher levels of embryonic stem cell markers, such as KLF4, implying that the WNT/β-catenin and Hedgehog/GLI pathways may be involved in metastatic reprogramming[60]. These results demonstrate the context-dependent function of KLF4 in colorectal cancer and offer a prospective new target for CRC therapy[63].

TAGLN or transgelin gene expressing cellular signaling proteins has dualism of action as an oncogene or tumor suppressor depending on the conditions and tissue type in which it is expressed[64]. TAGLN2 is associated with inhibition of invasion and migration in three cervical cancer cell lines (HeLa, SiHa, and C-33A), and overexpression of TAGLN2 in HeLa cells can inhibit cell viability, migration, and invasion. It is proposed that this can be achieved by downregulating chemokine receptor C-X-C type 4, matrix metalloproteinases (MMP-2, MMP-9), p50, transcription factor p65, and by increasing the expression levels of E-cadherin and inhibitor of nuclear factor κ-light chain-enhancer of activated B cells (NF-κB) (IκB)[65]. TAGLN, on the other hand, plays a role in many malignancies and is critical for the development of some malignancies, such as in triple-negative breast cancer (TNBC), TAGLN is abundantly expressed and stimulates cell growth, invasion, migration, and multiplication[66]. TAGLN may be a potential target identified as a key player in the regulation of ovarian cancer progression through its involvement in responding to the rigidity of the matrix extracellular environment and activation of the RhoA/ROCK specific pathway, so as to transform the primary tumor towards metastasis and invasion, in addition TAGLN is a potential target for therapy because it functions as a mechanosensitive gene in ovarian cancer[67]. In colon cancer conditions, TAGLN interacts with HMGA2 through a TGF-β regulatory mechanism by promoting epithelial-mesenchymal transition (EMT) in colorectal cancer to promote cell invasion and migration, which aids metastasis as well as tumor progressivity and angiogenesis formation[68,69]. TAGLN protein expression and nucleus translocation

can be induced by TGF-β to significantly enhance CRC cell migration. Overexpression of HMGA2 restored the effect of TGF-β on HT29 and HCT116 cells that was attenuated by TAGLN inhibition, confirming the interaction of HMGA2 and TAGLN. TAGLN knockdown can also prevent the loss of expression of metastatic markers induced by TGF-β[70].

5. CONCLUSIONS

Quantifying gene centrality and its significance in colon cancer incidence is helpful in alternative therapeutic therapies. Therapy using epigenetic miRNAs as regulators of gene expression in colon cancer has the potential to prevent and treat colon cancer patients. Bioinformatics prediction shows several miRNAs that can suppress oncogene expression in colon cancer with parameters of binding affinity, conformation, and conformational stability in molecular dynamics. In silico prediction results showed that CAV1, KLF4, and TAGLN are potential colon cancer treatment targets in Asian populations based on epigenetic inhibition of miRNAs hsa-miR-651-3p, hsa-miR-32-5p, and hsa-miR-185-5p, respectively.

6. ACKNOWLEDGEMENTS

We extend our sincere gratitude to Indonesia's Ministry of Education, Culture, Research, and Technology for their financial support of our fundamental regular research, provided under contract number 027/E5/PG.02.00.PL/2024.

Author contribution

All authors read, discussed, and contributed to the writing and finalizing of this article. F.R.M.: conceptualizing, designing content, writing original draft, drawing figures, editing; G.M.G., M.M., M.M.A.H., A.R., R.P.: conceptualizing, editing, writing, and reviewing; H.Z., L.A., I.M.A., W.N.: writing, reviewing, and editing. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

none to declare.

Funding

Indonesia's Ministry of Education, Culture, Research, and Technology funded this research under contract number 027/E5/PG.02.00.PL/2024.

thics approval

none to declare.

Article info:
Received February 1, 2024
Received in revised form July 28, 2024
Accepted August 21, 2024

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021 May;71(3):209–49.
2. Huang J, Ngai CH, Deng Y, Tin MS, Lok V, Zhang L, et al. Cancer Incidence and Mortality in Asian Countries: A Trend Analysis. Cancer Control J Moffitt Cancer Cent. 2022;29:10732748221095956.
3. Luo G, Zhang Y, Guo P, Wang L, Huang Y, Li K. Global patterns and trends in stomach cancer incidence: Age, period and birth cohort analysis. Int J Cancer. 2017 Oct 1;141(7):1333–44.
4. Wong MC, Ding H, Wang J, Chan PS, Huang J. Prevalence and risk factors of colorectal cancer in Asia. Intest Res [Internet]. 2019 May 20 [cited 2024 Jul 19];17(3):317–29. Available from:http://www.irjournal.org/journal/view.php?doi=10.5217/ir.2019.00021
5. Chakrabarti S, Peterson CY, Sriram D, Mahipal A. Early stage colon cancer: Current treatment standards, evolving paradigms, and future directions. World J Gastrointest Oncol [Internet]. 2020 Aug 8 [cited 2024 Jul 19];12(8):808. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7443846/
6. Holtedahl K, Borgquist L, Donker GA, Buntinx F, Weller D, Campbell C, et al. Symptoms and signs of colorectal cancer, with differences between proximal and distal colon cancer: a prospective cohort study of diagnostic accuracy in primary care. BMC Fam Pract. 2021 Jul 8;22(1):148.
7. CDC. Cancer Diagnosis and Treatment [Internet]. Cancer Survivors. 2024 [cited 2024 Jul 19]. Available from: https://www.cdc.gov/cancer-survivors/patients/index.html
8. Pekarek L, Torres-Carranza D, Fraile-Martinez O, García-Montero C, Pekarek T, Saez MA, et al. An Overview of the Role of MicroRNAs on Carcinogenesis: A Focus on Cell Cycle, Angiogenesis and Metastasis. Int J Mol Sci [Internet]. 2023 Apr 14 [cited 2024 Jul 19];24(8):7268. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10139430/
9. Chandraprabha Vineetha R, Anitha Geetha Raj J, Devipriya P, Sreelatha Mahitha M, Hariharan S. MicroRNA-based therapies: Revolutionizing the treatment of acute myeloid leukemia. Int J Lab Hematol [Internet]. 2024 [cited 2024 Jul 19];46(1):33–41. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ijlh.14211
10. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer. 2006 Apr;6(4):259–69.
11. Sassen S, Miska EA, Caldas C. MicroRNA: implications for cancer. Virchows Arch Int J Pathol. 2008 Jan;452(1):1–10.
12. Ramanto KN, Widianto KJ, Wibowo SSH, Agustriawan D. The regulation of microRNA in each of cancer stage from two different ethnicities as potential biomarker for breast cancer. Comput Biol Chem. 2021 Aug;93:107497.
13. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics [Internet]. 2010 Jan 1 [cited 2024 Jul 17];26(1):139–40. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2796818/

14. Inayatullah M, ANSARI MDA, Ayachit G, Gandhi M, Sharma P, Bhairappanavar S, et al. Differential gene expression analysis of HNSCC tumors deciphered tobacco dependent and independent molecular signatures. Oncotarget. 2019 Nov 6;10:6168–83.
15. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res [Internet]. 2019 Jan 8 [cited 2024 Jul 17];47(Database issue):D607–13. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6323986/
16. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants. Bioinformatics [Internet]. 2020 Apr 15 [cited 2024 Apr 7];36(8):2628–9. Available from: https://doi.org/10.1093/bioinformatics/btz931
17. Tang Y, Li M, Wang J, Pan Y, Wu FX. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. Biosystems. 2015 Jan;127:67–72.
18. Gan SL, Djauhari M. An Overall Centrality Measure: The Case of U.S Stock Market. Int J Basic Appl Sci. 2012 Dec 1;12:99–103.
19. Tastsoglou S, Alexiou A, Karagkouni D, Skoufos G, Zacharopoulou E, Hatzigeorgiou A. DIANA-microT 2023: including predicted targets of virally encoded miRNAs. Nucleic Acids Res. 2023 Apr 24;51.
20. He J, Wang J, Tao H, Xiao Y, Huang SY. HNADOCK: A nucleic acid docking server for modeling RNA/DNA–RNA/DNA 3D complex structures. Nucleic Acids Res [Internet]. 2019 Jul 2 [cited 2024 Mar 22];47(W1):W35–42. Available from: https://doi.org/10.1093/nar/gkz412
21. Krieger E, Joo K, Lee J, Lee J, Raman S, Thompson J, et al. Improving physical realism, stereochemistry and side-chain accuracy in homology modeling: four approaches that performed well in CASP8. Proteins [Internet]. 2009 [cited 2024 Jul 17];77(Suppl 9):114–22. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2922016/
22. Wang Y, Juranek S, Li H, Sheng G, Tuschl T, Patel DJ. Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex. Nature. 2008 Dec 18;456(7224):921–6.
23. Yan Y, Tao H, He J, Huang SY. The HDOCK server for integrated protein–protein docking. Nat Protoc [Internet]. 2020 May [cited 2022 Oct 20];15(5):1829–52. Available from: https://www.nature.com/articles/s41596-020-0312-x
24. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM. PDBsum: Structural summaries of PDB entries. Protein Sci Publ Protein Soc. 2018 Jan;27(1):129–34.
25. Moes D, Banijamali E, Sheikhhassani V, Scalvini B, Woodard J, Mashaghi A. ProteinCT: An implementation of the protein circuit topology framework. MethodsX [Internet]. 2022 Jan 1 [cited 2024 Jul 17];9:101861. Available from: https://www.sciencedirect.com/science/article/pii/S2215016122002400
26. Ezaj MMA, Junaid M, Akter Y, Nahrin A, Siddika A, Afrose SS, et al. Whole proteome screening and identification of potential epitopes of SARS-CoV-2 for vaccine design-an immunoinformatic, molecular docking and molecular dynamics simulation accelerated robust strategy. J Biomol Struct Dyn. 2022 Sep;40(14):6477–502.
27. Grant BJ, Rodrigues APC, ElSawy KM, McCammon JA, Caves LSD. Bio3d: an R package for the comparative analysis of protein structures. Bioinformatics [Internet]. 2006 Nov 1 [cited 2024 Feb 14];22(21):2695–6. Available from: https://doi.org/10.1093/bioinformatics/btl461
28. Ganesh K, Basnet H, Kaygusuz Y, Laughney AM, He L, Sharma R, et al. L1CAM defines the regenerative origin of metastasis-initiating cells in colorectal cancer. Nat Cancer. 2020 Jan;1(1):28–45.
29. Bharadwaj R, Jaiswal S, Velarde de la Cruz EE, Thakare RP. Targeting Solute Carrier Transporters (SLCs) as a Therapeutic Target in Different Cancers. Diseases [Internet]. 2024 Mar [cited 2024 Jul 17];12(3):63. Available from: https://www.mdpi.com/2079-9721/12/3/63
30. Özgür A, Vu T, Erkan G, Radev DR. Identifying gene-disease associations using centrality on a literature mined gene-interaction network. Bioinformatics [Internet]. 2008 Jul 1 [cited 2024 Jul 19];24(13):i277–85. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2718658/
31. Chu XD, Zhang YR, Lin ZB, Zhao Z, Huangfu SC, Qiu SH, et al. A network pharmacology approach for investigating the multi-target mechanisms of Huangqi in the treatment of colorectal cancer. Transl Cancer Res [Internet]. 2021 Feb [cited 2024 Jul 19];10(2):681–93. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8798599/
32. Ivey KN, Srivastava D. microRNAs as Developmental Regulators. Cold Spring Harb Perspect Biol [Internet]. 2015 Jul [cited 2024 Jul 19];7(7):a008144. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4484971/
33. Lee SH, Brianna B. Therapeutic Targeting of Overexpressed MiRNAs in Cancer Progression. Curr Drug Targets. 2022;23(13):1212–8.
34. Sidorova EA, Zhernov YV, Antsupova MA, Khadzhieva KR, Izmailova AA, Kraskevich DA, et al. The Role of Different Types of microRNA in the Pathogenesis of Breast and Prostate Cancer. Int J Mol Sci [Internet]. 2023 Jan 19 [cited 2024 Jul 19];24(3):1980. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9916830/
35. Arif KMT, Elliott EK, Haupt LM, Griffiths LR. Regulatory Mechanisms of Epigenetic miRNA Relationships in Human Cancer and Potential as Therapeutic Targets. Cancers [Internet]. 2020 Oct [cited 2024 Jul 19];12(10):2922. Available from: https://www.mdpi.com/2072-6694/12/10/2922
36. Li Y, Zhang C yang. Analysis of microRNA-induced silencing complex-involved microRNA-target recognition by single-molecule fluorescence resonance energy transfer. Anal Chem. 2012 Jun 5;84(11):5097–102.
37. Riffo-Campos AL, Riquelme I, Brebi P. Tools for Sequence-Based miRNA Target Prediction: What to Choose? Int J Mol Sci. 2016 Dec 9;17:1987.
38. Yan Y, Zhang D, Zhou P, Li B, Huang SY. HDOCK: A web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res [Internet]. 2017 Jul 3 [cited 2022 Oct 19];45(Web Server issue):W365–73. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5793843/
39. Kufareva I, Abagyan R. Methods of protein structure comparison. Methods Mol Biol Clifton NJ [Internet]. 2012 [cited 2022 Oct 20];857:231–57. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4321859/
40. Desantis F, Miotto M, Di Rienzo L, Milanetti E, Ruocco G. Spatial organization of hydrophobic and charged residues affects protein thermal stability and binding affinity. Sci Rep [Internet]. 2022 Jul 15 [cited 2024 Jul 19];12(1):12087. Available from: https://www.nature.com/articles/s41598-022-16338-5
41. Faltejsková K, Jakubec D, Vondrášek J. Hydrophobic Amino Acids as Universal Elements of Protein-Induced DNA Structure Deformation. Int J Mol Sci. 2020 Jun 2;21(11):3986.
42. Das S, Roy S, Bhattacharyya D. DNA base sequence specificity through partial intercalation: DFT-D based energy analysis of molecular dynamics snapshots. J Mol Graph Model [Internet]. 2020 Dec 1 [cited 2024 Jul 19];101:107722. Available from: https://www.sciencedirect.com/science/article/pii/S1093326320305118
43. Chiarella E, Aloisio A, Scicchitano S, Bond HM, Mesuraca M. Regulatory Role of microRNAs Targeting the Transcription Co-

Factor ZNF521 in Normal Tissues and Cancers. Int J Mol Sci [Internet]. 2021 Aug 6 [cited 2024 Jul 19];22(16):8461. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8395128/
44. Kitowska A, Wesserling M, Seroczynska B, Szutowicz A, Ronowska A, Peksa R, et al. Differentiation of high-risk stage I and II colon tumors based on evaluation of CAV1 gene expression. J Surg Oncol. 2015 Sep;112(4):408–14.
45. Wang S, Wang N, Zheng Y, Yang B, Liu P, Zhang F, et al. Caveolin-1 inhibits breast cancer stem cells via c-Myc-mediated metabolic reprogramming. Cell Death Dis [Internet]. 2020 Jun 11 [cited 2024 Jul 18];11(6):1–16. Available from: https://www.nature.com/articles/s41419-020-2667-x
46. Tirado OM, Mateo-Lozano S, Villar J, Dettin LE, Llort A, Gallego S, et al. Caveolin-1 (CAV1) is a target of EWS/FLI-1 and a key determinant of the oncogenic phenotype and tumorigenicity of Ewing's sarcoma cells. Cancer Res. 2006 Oct 15;66(20):9937–47.
47. Guo L, Liu Y, Yang T, Wang G, Liu J, Li S, et al. CAV1 and KRT5 are potential targets for prostate cancer. Medicine (Baltimore) [Internet]. 2023 Dec 8 [cited 2024 Jul 18];102(49):e36473. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10713156/
48. He J, Ouyang S, Zhao Y, Liu Y, Liu Y, Zhou B, et al. Prognostic Value of CAV1 and CAV2 in Head and Neck Squamous Cell Carcinoma. Biomolecules [Internet]. 2023 Feb 6 [cited 2024 Jul 18];13(2):303. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9952890/
49. Li B, Ming H, Qin S, Zhou L, Huang Z, Jin P, et al. HSPA8 Activates Wnt/β-Catenin Signaling to Facilitate BRAF V600E Colorectal Cancer Progression by CMA-Mediated CAV1 Degradation. Adv Sci [Internet]. 2023 Nov 16 [cited 2024 Jul 18];11(3):2306535. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10797426/
50. Díaz-Valdivia NI, Calderón CC, Díaz JE, Lobos-González L, Sepulveda H, Ortíz RJ, et al. Anti-neoplastic drugs increase caveolin-1-dependent migration, invasion and metastasis of cancer cells. Oncotarget [Internet]. 2017 Dec 5 [cited 2024 Jul 19];8(67):111943–65. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5762371/
51. Ha TK, Her NG, Lee MG, Ryu BK, Lee JH, Han J, et al. Caveolin-1 increases aerobic glycolysis in colorectal cancers by stimulating HMGA1-mediated GLUT3 transcription. Cancer Res. 2012 Aug 15;72(16):4097–109.
52. Colloca A, Donisi I, Anastasio C, Balestrieri ML, D'Onofrio N. Metabolic Alteration Bridging the Prediabetic State and Colorectal Cancer. Cells [Internet]. 2024 Jan [cited 2024 Jul 18];13(8):663. Available from: https://www.mdpi.com/2073-4409/13/8/663
53. Al Madhoun A, Hebbar P, Nizam R, Haddad D, Melhem M, Abu-Farha M, et al. Caveolin-1 rs1997623 variant and adult metabolic syndrome—Assessing the association in three ethnic cohorts of Arabs, South Asians and South East Asians. Front Genet [Internet]. 2022 Oct 21 [cited 2024 Jul 19];13:1034892. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9634410/
54. Tang J, Wang G, Li Q, Song R. Quantitative assessment of caveolin-1 G14713A polymorphism and cancer susceptibility in the Asian population. OncoTargets Ther. 2016;9:1381–7.
55. He Z, He J, Xie K. KLF4 transcription factor in tumorigenesis. Cell Death Discov [Internet]. 2023 Apr 8 [cited 2024 Jul 18];9(1):1–13. Available from: https://www.nature.com/articles/s41420-023-01416-y
56. Malik D, Kaul D. KLF4 Genome: A double edged sword. J Solid Tumors [Internet]. 2015 Mar 26 [cited 2024 Jul 19];5(1):49 Available from:https://www.sciedupress.com/journal/index.php/jst/article/view/6165
57. Qi XT, Li YL, Zhang YQ, Xu T, Lu B, Fang L, et al. KLF4 functions as an oncogene in promoting cancer stem cell-like characteristics in osteosarcoma cells. Acta Pharmacol Sin. 2019 Apr;40(4):546–55.
58. Leng Z, Tao K, Xia Q, Tan J, Yue Z, Chen J, et al. Krüppel-Like Factor 4 Acts as an Oncogene in Colon Cancer Stem Cell-Enriched Spheroid Cells. PLOS ONE [Internet]. 2013 Feb 13 [cited 2024 Jul 18];8(2):e56082. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0056082
59. Taracha-Wisniewska A, Kotarba G, Dworkin S, Wilanowski T. Recent Discoveries on the Involvement of Krüppel-Like Factor 4 in the Most Common Cancer Types. Int J Mol Sci [Internet]. 2020 Jan [cited 2024 Jul 18];21(22):8843. Available from: https://www.mdpi.com/1422-0067/21/22/8843
60. G S, C B, M K, I SC, A C, A R i A. In vivo epigenetic reprogramming of primary human colon cancer cells enhances metastases. J Mol Cell Biol [Internet]. 2016 Apr [cited 2024 Jul 19];8(2). Available from: https://pubmed.ncbi.nlm.nih.gov/26031752/
61. Leng Z, Li Y, Zhou G, Lv X, Ai W, Li J, et al. Krüppel-like factor 4 regulates stemness and mesenchymal properties of colorectal cancer stem cells through the TGF-β1/Smad/snail pathway. J Cell Mol Med [Internet]. 2020 Jan [cited 2024 Jul 19];24(2):1866–77. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6991673/
62. Varnat F, Siegl-Cachedenier I, Malerba M, Gervaz P, Ruiz i Altaba A. Loss of WNT-TCF addiction and enhancement of HH-GLI1 signalling define the metastatic transition of human colon carcinomas. EMBO Mol Med [Internet]. 2010 Nov [cited 2024 Jul 19];2(11):440–57. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3394505/
63. Lee E, Cheung J, Bialkowska AB. Krüppel-like Factors 4 and 5 in Colorectal Tumorigenesis. Cancers [Internet]. 2023 Apr 24 [cited 2024 Jul 19];15(9):2430. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10177156/
64. Kim HR, Park JS, Karabulut H, Yasmin F, Jun CD. Transgelin-2: A Double-Edged Sword in Immunity and Cancer Metastasis. Front Cell Dev Biol [Internet]. 2021 Apr 8 [cited 2024 Jul 19];9. Available from: https://www.frontiersin.org/journals/cell-and-developmental-biology/articles/10.3389/fcell.2021.606149/full
65. Zhou Q, Jiang X, Yan W, Dou X. Transgelin 2 overexpression inhibits cervical cancer cell invasion and migration. Mol Med Rep [Internet]. 2019 Jun 1 [cited 2024 Jul 19];19(6):4919–26. Available from: https://www.spandidos-publications.com/10.3892/mmr.2019.10116
66. Dvořáková M, Jeřábková J, Procházková I, Lenčo J, Nenutil R, Bouchal P. Transgelin is upregulated in stromal cells of lymph node positive breast cancer. J Proteomics [Internet]. 2016 Jan 30 [cited 2024 Jul 19];132:103–11. Available from: https://www.sciencedirect.com/science/article/pii/S1874391915301986
67. Wei X, Lou H, Zhou D, Jia Y, Li H, Huang Q, et al. TAGLN mediated stiffness-regulated ovarian cancer progression via RhoA/ROCK pathway. J Exp Clin Cancer Res [Internet]. 2021 Sep 19 [cited 2024 Jul 19];40(1):292. Available from: https://doi.org/10.1186/s13046-021-02091-6
68. Neuzillet C, Tijeras-Raballand A, Cohen R, Cros J, Faivre S, Raymond E, et al. Targeting the TGFβ pathway for cancer therapy. Pharmacol Ther. 2015 Mar;147:22–31.
69. Colak S, Ten Dijke P. Targeting TGF-β Signaling in Cancer. Trends Cancer. 2017 Jan;3(1):56–71.
70. Zhou H, Li L, Xie W, Wu L, Lin Y, He X. TAGLN and High-mobility Group AT-Hook 2 (HMGA2) Complex Regulates TGF-β-induced Colorectal Cancer Metastasis. OncoTargets Ther. 2020;13:10489–98.