

Article

Evaluation of SYK Gene as a Prognostic Biomarker and Suggested Aromatic Potential Phytochemicals to Halt the Colorectal Cancer

Partha Biswas^{1, 2}, Dipta Dey³, Atikur Rahman⁴, Md. Aminul Islam¹, Tasmina Ferdous Susmi¹, Md. Abu Kaium¹, Md. Nazmul Hasan⁵, MD. Hasanur Rahman^{2, 6}, Shafi Mahmud⁷, Md. Abu Saleh⁷, Priyanka Paul³, Md Rezanur Rahman⁸, Md. Aatur Rahman^{9, 10, 11}, Han Gyeul Song¹⁰, Bonglee Kim^{10, 11*}

¹ Department of Genetic Engineering and Biotechnology, Faculty of Biological Science and Technology, Jashore University of Science and Technology (JUST), Jashore- 7408, Bangladesh

² 2ABEx Bio-Research Center, East Azampur, Dhaka 1230, Bangladesh.

³ Department of Biochemistry and Molecular Biology, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj-8100, Bangladesh.

⁴ Fermentation Engineering, School of Biotechnology, Jiangnan University, Wuxi, Jiangsu, China.

⁵ Pharmaceutical Biotechnology Laboratory, Department of Genetic Engineering and Biotechnology, Faculty of Biological Science and Technology, Jashore University of Science and Technology (JUST), Jashore- 7408, Bangladesh.

⁶ Department of Biotechnology and Genetic Engineering, Bangabandhu Sheikh Mujibur Rahman Science and Technology University, Gopalganj 8100, Bangladesh.

⁷ Department of Genetic Engineering and Biotechnology, University of Rajshahi, Rajshahi-6205, Bangladesh.

⁸ Department of Biochemistry and Biotechnology, Khwaja Yunus Ali University, Enayetpur, Sirajganj, Bangladesh.

⁹ Global Biotechnology & Biomedical Research Network (GBBRN), Department of Biotechnology and Genetic Engineering, Faculty of Biological Sciences, Islamic University, Kushtia 7003, Bangladesh.

¹⁰ Department of Pathology, College of Korean Medicine, Kyung Hee University, Seoul 02447, South Korea.

¹¹ Korean Medicine-Based Drug Repositioning Cancer Research Center, College of Korean Medicine, Kyung Hee University, Seoul 02447, Korea.

* Correspondence: Department of Pathology, College of Korean Medicine, Kyung Hee University, 26 Kyungheedaero, Dongdaemun-gu, Seoul 02447, South Korea. Bonglee Kim, Phone: +82-2-961-9217; E-mail: bongleekim@khu.ac.kr

Abstract: Background: Colorectal cancer is considered the third most fatal among all type of cancer. Spleen tyrosine kinase (SYK) is a non-receptor type tyrosine-protein that plays crucial role in signaling mediated via immune receptor. We adopted an onco-informatics analysis to evaluate the SYK expression and prognostic value of SYK in colorectal cancer, and identification of potential phytochemicals which may inhibit overexpression of SYK protein as well as minimized colorectal cancer. Materials & Methods: Differential expression of SYK gene was analyzed using the several transcriptomic databases including Oncomine, UALCAN, GENT2 and GEPIA2. The server, cBioPortal was used to analyze mutation and copy number alterations whereas GENT2, GEPIA, OncoLnc and Prognoscan were employed to examine the survival rate. A protein-protein interaction network of SYK and co-expressed genes of SYK was conducted via GeneMANIA. Considering SYK gene encoding protein as drug target, selected phytochemicals were assessed by molecular docking using PyRx 0.8 packages. YASARA molecular dynamics simulators were applied for the post validation of the molecular docking data. Results: We have observed significant overexpression of mRNA expression levels of SYK gene colorectal adenocarcinoma (COAD) samples compared with normal tissues. Significant methylation level and various genetic alterations are assembled in SYK gene which can lead to the development of colorectal cancer. As a result, lower level of SYK expression was related to the more chances of patients' survival by which all the outcomes from the multiple bioinformatics platforms and web resources have demonstrated the significant evidences that the SYK kinase can possess as a potential biomarker for the treatment of colorectal cancer. Here, aromatic phytochemicals namely, Kaempferol and Glabridin targeting SYK showed more stability compared to controls and may be useful for the treatment of colorectal cancer. Conclusion: Our

J.

study showed dysregulated expression of SYK in colorectal cancer and potentiality to act as a biomarker for the prognosis of CRC. Moreover, we have shown phytochemicals (Kaempferol and Glabridin) target SYK as potential treatment strategies and drug repositioning potentiality in colorectal cancer.

Keywords: Colorectal cancer, SYK, Prognostic Significance, Genetic Alteration, Molecular dynamics simulations.

1. Introduction

Cancer is a terrific health problem and its prevalence continues to accelerate every year throughout the world. According to WHO, cancer is a large group of diseases that occurred due to abnormally heightened cells grow uncontrollably, and invade to adjacent organs in the body [1]. Moreover, the Centers for Disease Control and Prevention (CDC) reported that cancer ranked second solely after coronary illness as the head reason for death in America and around the world [2]. The International Agency for Research on Cancer (IARC) gauges the worldwide malignant growth trouble ascends to 19.3 million new cases and 10.0 million disease passings in 2020. It likewise announced that overall one of every 5 individuals create malignant growth during their lifetime and one out of 8 men and one out of 11 ladies pass on from the illness [3]. Furthermore, the American Cancer Society (ACS) as of late distributed their Cancer Facts and Figures 2021 reports, where they expressed that in the United States 1,898,160 new malignant growth cases and 608,570 disease passings are projected to happen in 2021 [4]. All these cases attract both researchers and physicians all over the world to cast the utmost emphasis on cancer research. Among all the 39 different cancer types, colorectal cancer (CRC) is the third maximum commonly diagnosed and second maximum malignant growth-related demise worldwide in both sexes and all ages of patients [4]. According to GLOBOCAN 2020 data, around 2 million people diagnosed, comprising 10%, and around 1 million people died, comprising 9.4% in CRC of all cancer types in both sexes and all ages worldwide [5]. It's a matter of global concern that the burden of colorectal cancer is expected to escalate about 60% in countries with high or very high human development index (HDI) particularly in Eastern Europe, Asia, and South America by 2030 [6]. Thus, it is necessary to establish improved assessment and treatment strategies however, the therapy and prognosis fluctuate broadly with various malignancy stages and biological features. Rely upon tumor node metastasis (TNM) staging, CRC patients determined with stage-I have a five-year endurance pace of more than 90%, while in stage-IV it declines around 12% [7]. Colorectal cancer is a heterogeneous group of diseases caused by a series of genetic and epigenetic mutations in certain cells of the epithelium [8]. For the complexity of the disease, a more comprehensive method is required to gain high accuracy for CRC detection and diagnosis as a single test might provide low accuracy. In this regard, the multi-omic data mining approach can be a feasible alternative to anti-cancer therapies.

Spleen tyrosine kinase (SYK) encodes a non-receptor type tyrosine-protein which is mainly associated with the adaptive immune receptor signal [9]. This protein is involved in various signaling-mediated biological processes including innate immunity recognition, platelet activation, and vascular development along with cellular responses such as differentiation, proliferation, and phagocytosis [9]. In the human cell, the SYK gene is located at the q22.2 position of the 9th chromosomes [10]. As SYK is broadly communicated in hematopoietic cells, it is a capable modulator of epithelial cell development and a potential tumor target in human cancers [11]. The expression of the SYK gene is equally important for both normal and cancer cells whereas it contains tumor-promoting and tumor-suppressing capability too [12]. Various evidence suggests that SYK involves as the reason behind the formation of different cancers [13]. Mainly the growth of tumor and metastasis

J.

is developed and activated through the SYK signaling pathway as an intermediate [14]. SYK gene is considered a potent tumor suppressor gene in the human and the suppression of SYK activity increases breast cancer tumorigenicity [11,15]. Apart from breast cancer, lower SYK expression and methylation is connected with metastasis in other cancer including lung cancer, liver cancer, oral squamous cell cancer, pancreatic cancer, bladder cancer, gastric cancer and urinary cancer [16–18]. Besides colorectal cancer is also associated with this gene and the loss of its expression is found in several colorectal cancer tissues [19]. Thus SYK promotes the suppression of cancer formation as well as contributes to inhibiting cell proliferation and metastasis capacity [19].

Latest cancer therapies have severe side effects, high costs, and single survival thresholds [20]. In addition, the development of characteristics of drug resistance in cancer cells, the development of new therapeutic agents against cancer is a crying need in recent times. Several studies have shown that plant extracts or phytochemicals are purified and have executed significant positive results in cancer treatment by using directly or in combination with existing drugs [21]. Various well-known phytochemicals are found to exhibit a specific action against tumors, breast cancer, and prostate cancer [22]. It is also shown that various plant derived compounds have anti-proliferative action against cancer by modulation of cellular pathways that make the phytochemicals worthy as perfect drug candidates [23]. Plant bioactive compounds suppress the interaction of carcinogenic molecules, such as damage to DNA. The phytochemicals are usually considered safe, and easily available to consumers, so they are suggested as effective therapeutics for cancer treatment with fewer side effects in the human body.

In this study, we have performed an in-depth analysis to examine the role and clinical relevance of Spleen tyrosine kinase (SYK) in colorectal cancer (CRC) development by adopting various cancer databases. Here, we have figured out the expression patterns of the gene in different malignancy types, clinicopathological parameters, methylation status, genetic alteration frequencies, survival assay analysis, genes coexpressed with SYK, and the interaction network of these genes. Moreover, we have suggested few phytochemicals as a drug-like anticancer compound via molecular modeling targeting SYK by using in silico drug design approaches. Consequently, we have analyzed the pharmacokinetic properties (ADMET profiling) of the selected phytochemicals, protein-ligand docking, post docking analysis, and molecular dynamics simulation. These multi-omics data mining and in silico drug designing approaches will facilitate researchers in finding new anti-cancer therapies.

2. Materials and Methods

2.1 The Analysis of SYK gene Expression in Colorectal Cancer

The objectives of this research was to examine mRNA transcription levels in various cancers, including colorectal cancer, we employed Oncomine (<https://www.oncomine.org/>), UALCAN (<http://ualcan.path.uab.edu/>), GENT2 (<http://gent2.ap-pex.kr/gent2/>), GEPIA2 (<http://gepia2.cancer-pku.cn/#index>) databases. These are all publicly accessible interactive online platform which demonstrates the mRNA transcriptional levels in different malignancy tests with their comparable to typical control. Currently, Oncomine is the world's biggest oncogene chip information base and incorporated information mining stages containing 715 autonomous datasets and 86,733 examples [24]. SYK gene mRNA expression in clinical disease examples was contrasted and relating ordinary controls where the limit boundaries of p-value and fold change were settled as 0.0001 and 2, respectively. In-depth analysis of SYK gene in colon adenocarcinoma (COAD) using the TCGA database was done from UALCAN web portal, provides expression, methylation, pan-cancer view, survival, and correlation data [25]. Besides, DNA methylation levels at the promoter region of SYK gene based on various parameters are also picked out from UALCAN database. Gene Expression database of Normal and Tumor tissues 2 (GENT2) was also employed to retrieve SYK gene expression data across 72 different paired tissues [26]. Moreover, SYK gene mRNA expression and expression in cancer stages using the

J.

TCGA data in COAD is also further assessed from Gene expression profiling interactive analysis (GEPIA) 2 database [27]. It provides a gene-specific relative investigation of various malignant growth types utilizing a standard handling pipeline containing around 8,587 typical and 9,736 tumor samples.

2.2 Mutation and CNAs Determination of SYK gene

The cBioPortal for Cancer Genomics (<https://www.cbioportal.org/>) is an open-source interactive web portal for systematic exploration of multidisciplinary cancer genomic datasets currently containing 308 cancer studies [28,29]. It provides information from over 5000 tumor samples for mapping the recurrence of mutations and other hereditary modifications, atomic profiling of cell lines and cancer tissue by applying multifaceted cancer studies. We scrutinize cBioPortal to analyze mutation and copy number alterations of SYK gene in colorectal cancer.

2.3 Survival Assay Analysis

Survival plot analysis is the statistical view that provides the data of patients' survival rate against time for cancer patients. Various web-based tools were employed which includes GENT2, GEPIA, OncoLnc and UALCAN to examine the survival rate of colorectal cancer patients against the expression of SYK gene. The website, GENT2 (<http://gent2.appex.kr>) was used to identify the survival condition where the analysis was performed based on cancer subtype and prognosis condition for a specific tissue [30]. This database provides data for colon cancer based on 1146 sample by analyzing 5 subtypes including Molecular Subtype, AJCC Stage, Duke Stage, Grade and Histology. GEPIA (<http://gepia.cancer-pku.cn/>) is a database that also analyzes survival data for overall survival (OS) or disease-free survival (DFS) [31]. It uses log-rank test and performs based on gene expression by selecting any specific type of cancer [31]. Besides OncoLnc database was also used to investigate survival analysis. OncoLnc (<http://www.oncolnc.org/>) provides survival data for 8,647 patients by analyzing mRNA, miRNA, or lncRNA expression from 21 cancer 18 [32]. The database, UALCAN (<http://ualcan.path.uab.edu>) is an online resource that is also used for survival analysis of SYK gene. The database offers the relation between gene expression and patient survival which evaluates TCGA patient survival data for Kaplan–Meier survival analyses and generates overall survival plots [33].

2.4 Correlation Analysis and Interaction Network

It is essential to classify the associated genes of the target gene in order to better analyze gene expression. We accomplished this by looking at the SYK gene's associated genes on the UALCAN and GEPIA websites. GEPIA server was employed to detect similar genes for TCGA COAD tumors. Besides using the correlation segment of UALCAN database, positively correlated genes with SYK in COAD was identified where Pearson-CC >0.47 was considered as moderate significant from both websites. Then the identified similar or correlated genes were inputted into the GeneMANIA database to predict the protein-protein interaction network. GeneMANIA (<https://genemania.org/>) is an online database that predicts an interaction network for a set of genes with the inputted genes. The database provides the interaction with the association of protein and genetic interactions, pathways, co-expression, co-localization and protein domain similarity [34]. Besides GeneMania was also used to predict the protein-protein interaction network of SYK gene to express the relationship between the SYK gene with the gene in the list.

2.5 Target Identification of SYK gene for Colorectal Cancer

Spleen tyrosine kinase (SYK) is a non-receptor protein kinase of tyrosine, which has a part to perform when combined with different malignancies, thus this kinase has become an important target for the development of a therapeutic drug. It has been reported that the loss or decrease nature of SYK is expressed in different normal tissue and cancer cell types [35–38]. The loss of SYK expression is the main reason for the development of colorectal cancer. The expression could be restored in the colorectal cancer cell line by using several selected drugs or phytochemicals. As the cell proliferation is inhibited by the lower or overexpression of SYK gene, it could be used as a therapeutic target agent

J.

against colorectal cancer through targeting protein identification and thus the colorectal cancer disease can be overcome.

2.6 Pharmacokinetics analysis

All the selected aromatic compounds initially analyzed via the "Swiss ADME (<http://www.swissadme.ch/index.php>)" online server, in which evaluate the physiochemical properties of the lead molecules and every candidate follow the Lipinski's 5 rules. To better understanding of the pharmacokinetics characteristics (absorption, distribution, metabolism, excretion, and toxicity) of the extracted ligand was conducted by the "pKCSM as an online server (<http://biosig.unimelb.edu.au/pkcsml/prediction>)" [39].

2.7 Extracting lead molecule for Optimization

Phytochemicals were first retrieved from the PubChem data base to prepare the Ligand for the docking of protein-ligands and obtained the 3D phytochemical structure required for molecular docking. The compounds were prepared as a ligand by using UCSF Chimera 1.14 (<https://www.cgl.ucsf.edu/chimera/>) tool, where optimization of ligand, added hydrogen atoms, charges at zero, and the output file saves as PDB format [40].

2.8 Extracting targeted macromolecule for optimization

The tertiary, non-mutated structure of the targeted soluble SYK protein, which bearing the PDB ID at 4XG3 was collected from the online server PDB (Protein Data Bank) (<https://www.rcsb.org/>) and following the sequential steps for optimization of the protein, at first eliminating unnecessary things just like: water molecules, metal ions, ligands, heteroatoms as well as unwanted extra chains. Hydrogen bonds are added to the macromolecules for more precise docking and the overall optimization operation was conducted by the "UCSF Chimera 1.14" software [41].

2.9 Molecular Docking and Post Docking Data Visualization

Optimization of the selected ligand which proposed in the above for using as a drug target against the selected SYK (spleen tyrosine kinase) receptor, the protein-ligand docking was carried out through using the tool "PyRx software package version 0.8" (<https://pyrx.sourceforge.io/home>) that performs based on the Auto-Dock Vina Platform. In each docking, the top ten best binding affinities were displayed via (.csv) file [42].

Ligplot+ version 2.2 (<https://www.ebi.ac.uk/thornton-srv/software/LigPlus/>) have been conducted for analysis of the interaction between protein-ligand complex, where evaluate the non-bond and the non-covalent (polar, Hydrophobic) interaction. However, the Ligplot+ visualizing tool is effectively performed based on the java interface (Java SE Runtime Environment 8u271), where only the PDB file can be applied that are retrieved from the PYMOL visualizing tool [43].

2.10 Molecular Dynamics Simulation

The molecular dynamics complex simulation was carried out in YASARA dynamics where AMBER14 force field [44] was used [45]. The cubic simulation cell was created and the complexes were optimized and hydrogen bond networks were oriented [46]. The TIP3P water solvation model was used at (0.997 g/L-1, 25c, 1 atm) where steepest gradient algorithms were used by simulated annealing method to minimize the protein complexes. The simulation system was neutralized at 0.9% NaCl, 310K and pH 7.4. The techniques of Particle Mesh Ewald were used to calculate the electrostatic interaction with a radius of 8 Å cut off [47]. In either side of the systems the simulation cell was expanded to 20Å so that the protein could move freely inside the simulation cell [48]. In order to maintain the simulation temperature, the Berendsen thermostat was used. The simulation was conducted with a time of 1.25fs, and trajectories were preserved every 100ps [49]. The simulation was performed over 100 ns and trajectories analysis was carried out in order to measure the root mean square fluctuations, root mean square deviation, hydrogen bond, radius of gyration, and solvent accessible surface area.

3. Results

3.1 Expression Analysis of SYK gene

J.

Firstly, the differential expression patterns of SYK gene in various cancer types were analyzed by using Oncomine, GENT2, and GEPIA2 databases. We have found significantly differential expression patterns of SYK mRNAs in various cancers compared to healthy controls (Figure 1a). Notably, we have observed significant upregulation of SYK gene in colorectal cancer (p-value <0.05) compared to controls. In total, there are 12 databases for CRC on Oncomine. Statistical details of SYK gene expression in different subtypes of colorectal cancer from the Oncomine database were presented in supplementary table 1. According to the GEPIA2 database, SYK gene has been significantly upregulated in COAD comparing with their paired normal tissues (Figure 1c).

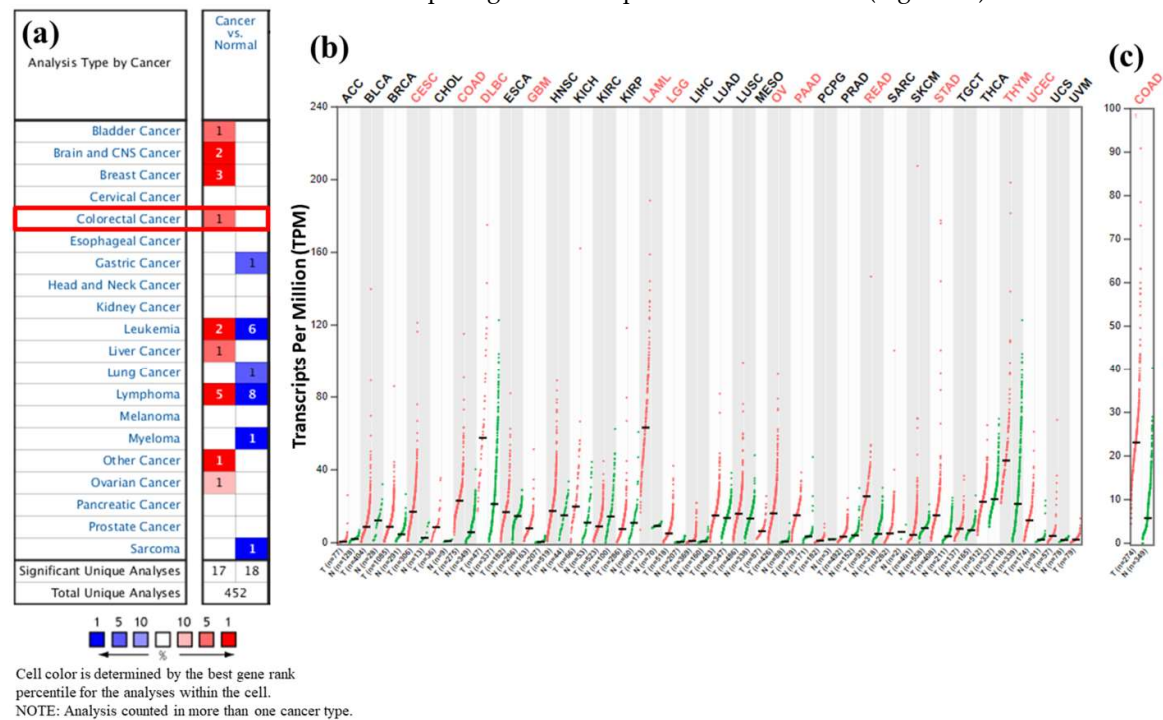


Figure 1. SYK gene mRNA expression profile across diverse cancer types; (a) mRNA expression of SYK in different cancers from Oncomine database where red cells with a number in it indicate the dataset with significant mRNA overexpression and blue cells with mRNA downregulation, (b) Pan-cancer expression profile analysis of SYK gene from GEPIA2 database in dot plot where each dots represent the expression of samples. mRNA expression is presented in transcript per million (TPM) scale across different cancer types, (c) Transcription level of SYK particularly in colorectal adenocarcinoma (COAD) patients with their corresponding normal samples from GEPIA2 database.

Moreover, SYK gene is also significantly upregulated in diverse cancer types like Invasive breast carcinoma (BRCA), Cervical squamous carcinoma and endocervical adenocarcinoma (CESC), cholangiocarcinoma (CHOL), squamous cell carcinomas of the head and neck (GBM), acute myeloids leukemia (LAML), rectum adenocartinoma (READ), adenocarcinoma of the stomach (STAD), ulterine corpus carcinoma carcinoma of the endometrium (UCEC) etc (Figure 1b). On the contrary, SYK gene is downregulated in Bladder Urothelial Carcinoma (BLCA), Kidney renal clear cell carcinoma (KIRC), Kidney renal papillary cell carcinoma (KIRP), Prostate adenocarcinoma (PRAD), Skin Cutaneous Melanoma (SKCM), Thyroid carcinoma (THCA), etc (Figure 1b). In addition, the results obtained from Oncomine and GEPIA2 databases are also cross-checked from GENT2 database. It provides SYK gene expression profile in boxplot across 72 different paired normal

J.

vs cancer tissues which also confirmed that this gene is significantly upregulated in various cancer types including colorectal cancer. Secondly, we analyzed SYK gene mRNA expression particularly in colorectal adenocarcinoma (COAD) tissues with their corresponding normal samples in light of various clinicopathological boundaries from UALCAN and GEPIA2 databases. The results obtained from UALCAN database indicated that SYK gene mRNA expression is significantly increased in TCGA COAD samples compared with their typical counterparts. The correlation between SYK gene mRNA expression and different clinicopathological parameters in TCGA COAD tissues demonstrates that SYK gene is significantly upregulated compared to the normal based on different variables such as sample types, individual cancer stages, patient's race, patient's gender, patient's weight, patient's age, histological subtypes, nodal metastasis status, and TP53 mutation status, etc. The data were presented in figure 2 and supplementary table 2.

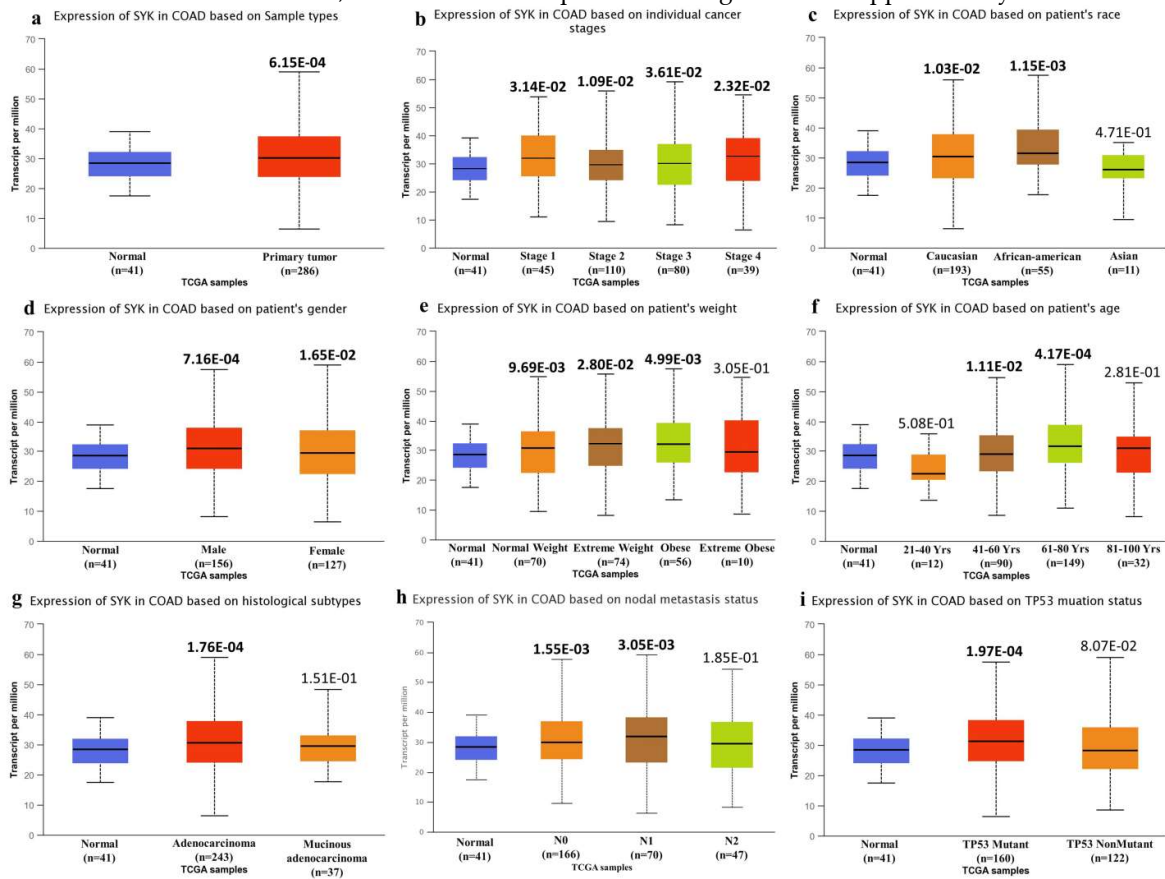


Figure 2. Expression profile analysis of SYK in TCGA colorectal adenocarcinoma (COAD) between normal vs different stages of cancer, based on different variables using the UALCAN database; (a) Expression of SYK gene in COAD dependent on sample types, (b) SYK expression in COAD dependent on individual cancer stages, (c) SYK expression in COAD dependent on patient's race, (d) SYK expression in COAD dependent on patient's gender, (e) SYK expression in COAD dependent on patient's weight, (f) SYK expression in COAD dependent on patient's age, (g) SYK expression in COAD dependent on histological subtype, (h) SYK expression in COAD dependent on nodal metastasis status, and (i) SYK expression in COAD dependent on TP53 mutation status. Statistically significant P-values are marked in bold.

Moreover, the data depicted that SYK gene overexpression is more statistically significant in male patients than the female and patients between 41-80 years old. Besides, this overexpression is more significant in African-American patients than the Caucasian and Asian. Likewise, SYK gene expression in box plot and based on pathological stages in

J.

colorectal adenocarcinoma is also retrieved from GEPIA2 database. This database also provides similar kind of results as per UALCAN for SYK gene expression. SYK gene is overexpressed in COAD samples than the normal in log2 scale where the p-value cutoff was 0.01 (Figure 3a).

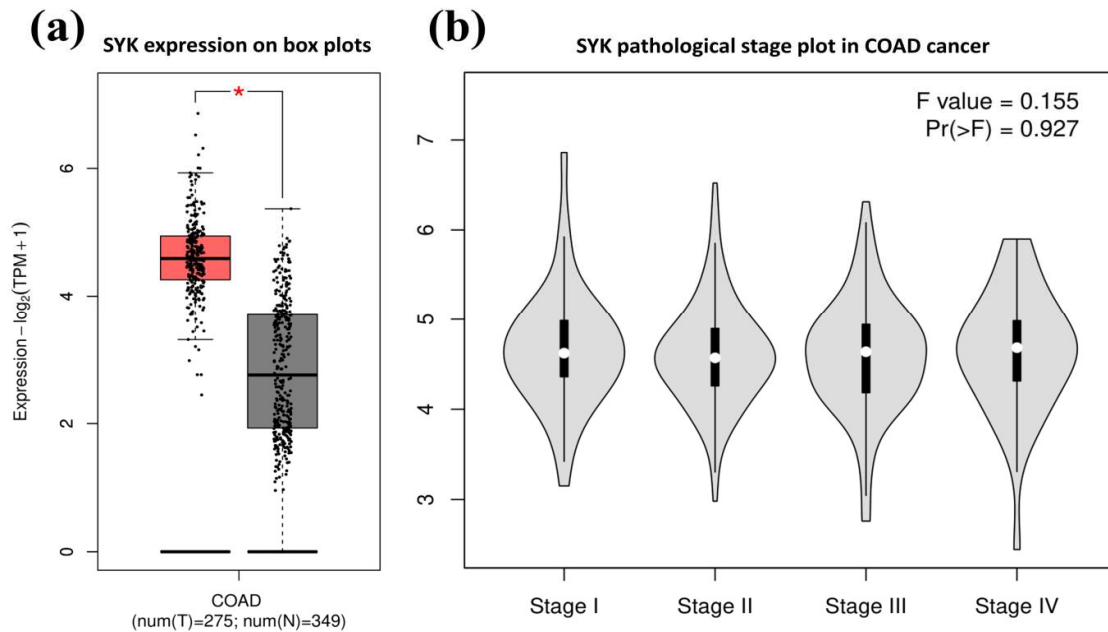


Figure 3. Tissue wise expression profile of SYK gene in COAD from GEPIA2 database; (a) SYK gene expression in COAD on box plot where red box indicates tumor samples and the black box indicates normal samples, (b) Expression of SYK in four tumor stages of COAD on pathological stage plot.

Similarly, overexpression of SYK gene is also evaluated in different pathological stages of COAD (Figure 3b). Thirdly, the promoter methylation level of SYK gene in COAD is determined from UALCAN database where the methylation profile is presented on the Beta value scale. The degree of promoter DNA methylation is shown by the Beta value going from 0 (unmethylated) to 1 (completely methylated).

J.

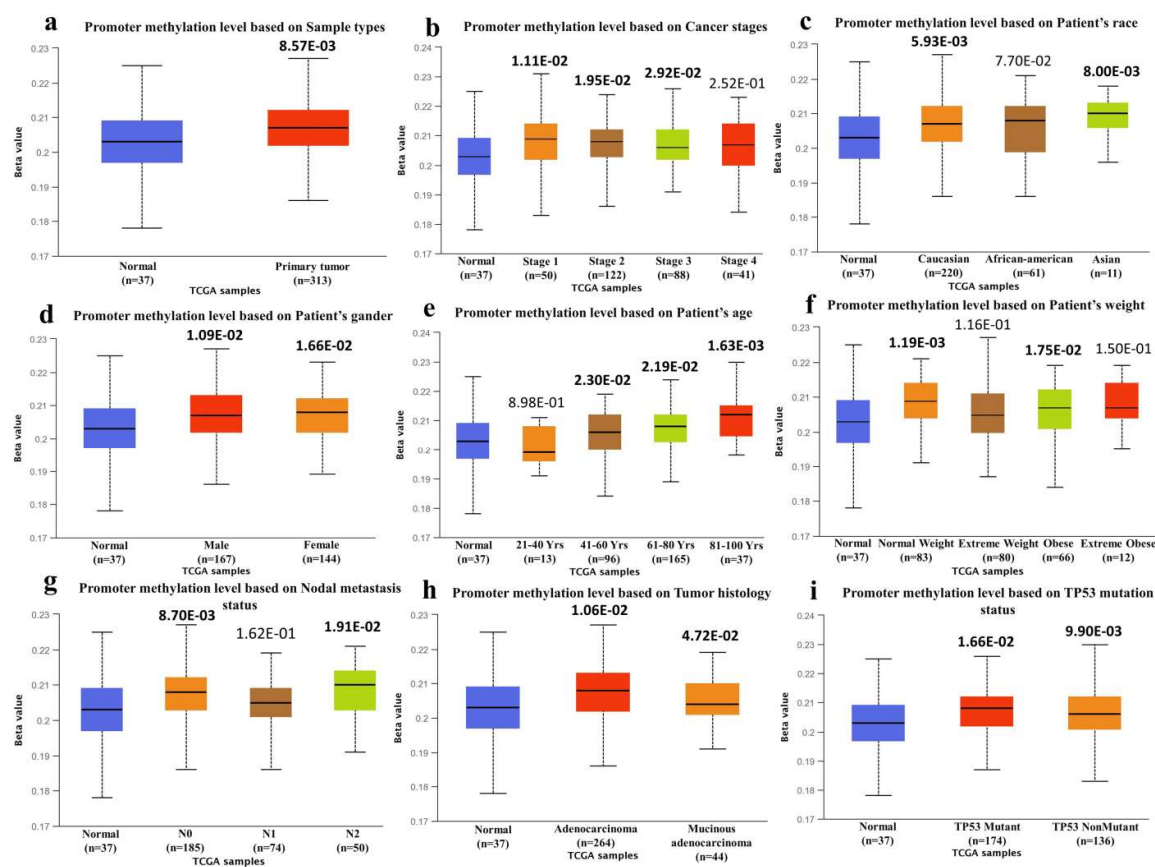


Figure 4. SYK gene promoter methylation level in TCGA colorectal adenocarcinoma (COAD) based on different variables by using UALCAN database; (a) Promoter methylation level of SYK in COAD based on sample types, (b) Promoter methylation level of SYK in COAD based on individual cancer stages, (c) Promoter methylation level of SYK in COAD based on patient's race, (d) Promoter methylation level of SYK in COAD based on patient's gender, (e) Promoter methylation level of SYK in COAD based on patient's age, (f) Promoter methylation level of SYK in COAD based on patient's weight, (g) Promoter methylation level of SYK in COAD based on nodal metastasis status, (h) Promoter methylation level of SYK in COAD based on tumor histology, and (i) Promoter methylation level of SYK in COAD based on TP53 Mutation status. According to UALCAN, the Beta value indicates the level of DNA methylation ranging from 0 (unmethylated) to 1 (fully methylated). P-value in bold denotes statistical significance.

The promoter methylation level of SYK gene is altogether expanded in COAD than the typical tissues dependent on sample types, individual cancer stages, patient's race, patient's gender, patient's age, patient's weight, tumor histology, and T53 mutation status, etc. The data of promoter methylation level of SYK in COAD is presented in Figure 4 and supplementary table 3 in both graphical and tabulate format, respectively.

3.2. Genetic Alteration Analysis in SYK Protein Sequences Associated with Colorectal Cancer Development

By using cBioPortal information, we generated multiple genetic alterations data to inspect the useful meaning of SYK gene in colorectal cancer advancement. At first, we queried for changes of SYK in this database utilizing 3953 samples of 3806 colorectal malignancy patients from 10 investigations. For these 635 amino acids long human SYK protein we found 57 mutations and the somatic mutation frequency in SYK is 1.3% (Table 1).

J.

Table 1. A rundown of hereditary shifts in SYK protein arrangements related with colorectal malignant growth advancement. In total, 57 mutations are detected including 7 duplicate mutations in patients with multiple samples.

Cancer Study	Sample Size	Protein Change	Mutation Type	Sample ID
Colorectal Adenocarcinoma (Genentech, Nature 2012)	74	A52T	Missense	587376
		V633M	Missense	587342
		P85L	Missense	587332
Colorectal Adenocarcinoma (DFCI, Cell Reports 2016)	619	A353V	Missense	coadread dfci 2016 68
		S511N	Missense	coadread dfci 2016 1762
		L143M	Missense	coadread dfci 2016 197
		N406Tfs*13	FS del	coadread dfci 2016 207430
		N406Tfs*13	FS del	coadread dfci 2016 2944
		R520H	Missense	coadread dfci 2016 2372
		P430L	Missense	coadread dfci 2016 3048
		K397E	Missense	coadread dfci 2016 3670
		R109Q	Missense	coadread dfci 2016 62
		M166Nfs*14	FS ins	TCGA-AG-A02N-01
Colorectal Adenocarcinoma (TCGA, Firehose Legacy)	640	K387N	Missense	TCGA-AF-3913-01
		R574*	Nonsense	TCGA-AG-A002-01
		M166Nfs*14	FS ins	TCGA-AG-A02N-01
Colorectal Adenocarcinoma (TCGA, Nature 2012)	276	K387N	Missense	TCGA-AF-3913-01
		R574*	Nonsense	TCGA-AG-A002-01
		A353T	Missense	TCGA-D5-6922-01
Colorectal Adenocarcinoma (TCGA, PanCancer Atlas)	594	M166Nfs*14	FS ins	TCGA-AG-A02N-01
		R574*	Nonsense	TCGA-AG-A002-01
		R574*	Nonsense	TCGA-F5-6814-01
		F549L	Missense	TCGA-F5-6814-01
		K105N	Missense	TCGA-AG-A00Y-01
		D344G	Missense	TCGA-F5-6814-01
		P119S	Missense	TCGA-A6-2686-01
		Y91C	Missense	TCGA-AY-6197-01
		T345R	Missense	TCGA-G4-6586-01
		E442Sfs*31	FS del	TCGA-WS-AB45-01
		R42C	Missense	P-0004602-T01-IM5
		R42C	Missense	P-0005230-T01-IM5
		R45C	Missense	P-0013492-T01-IM5
Metastatic Colorectal Cancer (MSKCC, Cancer Cell 2018)	1134	G33Afs*2	FS del	P-0006365-T01-IM5
		G33Afs*2	FS del	P-0013876-T01-IM5
		A286V	Missense	P-0010929-T01-IM5
		E442K	Missense	P-0010587-T01-IM5
		A282V	Missense	P-0005443-T01-IM5
		S84T	Missense	P-0005823-T01-IM5
		E144G	Missense	P-0010581-T01-IM5
		R367*	Nonsense	P-0006960-T01-IM5
		P411L	Missense	P-0006960-T01-IM5
		V560A	Missense	P-0006960-T01-IM5
		H62R	Missense	P-0005455-T01-IM5
		V433M	Missense	P-0005455-T01-IM5
		G185*	Nonsense	P-0002671-T01-IM3
		R175Gfs*4	FS del	P-0002671-T01-IM3
		Y47N	Missense	P-0000769-T01-IM3
		K509R	Missense	P-0001500-T03-IM5
		M166Nfs*14	FS ins	P-0007831-T01-IM5
		E230G	Missense	P-0013227-T01-IM5
		V633M	Missense	RC-MSK-008-pt
		V633M	Missense	RC-MSK-008-tm
Rectal Cancer (MSK,Nature Medicine 2019)	339	K509R	Missense	P-0001500-T03-IM5
		G33Afs*2	FS del	05CO041
		A83T	Missense	01CO014
		A412S	Missense	05CO015
Colon Cancer (CPTAC-2 Prospective, Cell 2019)	110	Y47C	Missense	11CO059

J.

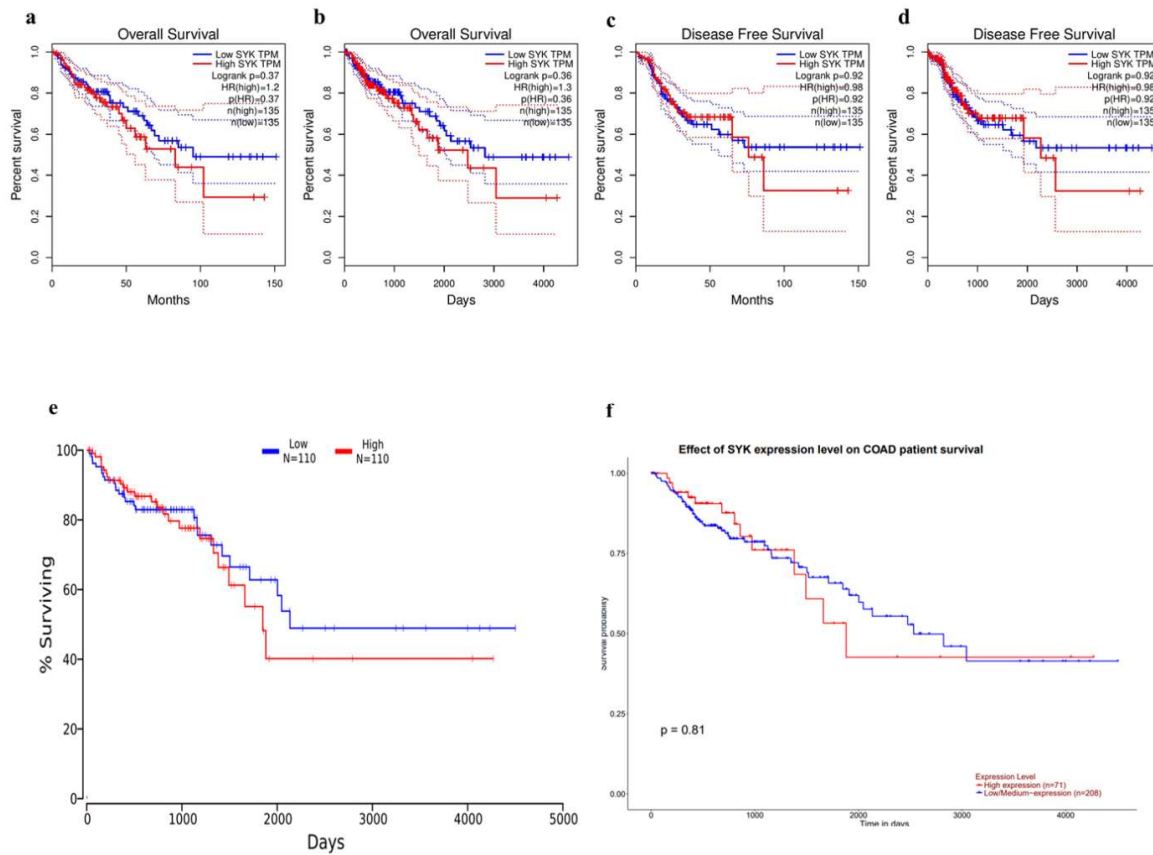


Figure 6. Survival Assay for SYK gene in colorectal cancer. Survival plot from GEPIA database where (a), (b) overall survival and (c), (d) Diseases free survival. (e) Survival plot from OncoLnc and (f) UALCAN database.

These survival plots were conducted using 50% median cut-off value, 95% Confidence Interval (CI) and Hazards Ratio (HR) from GEPIA. The database depicted that higher survival was observed for lower SYK level for overall survival and for diseases free survival condition (Figure 6a, 6b, 6c, 6d). It also seemed to be the higher survival for lower expression level (Figure 6a, 6b, 6c, 6d). Whereas to perform a Kaplan-Meier plot by analyzing a wide range of TCGA cancer studies from OncoLnc database, both upper percentile and lower percentile were selected 25. Then, the data of overall survival was analyzed using the patients' survival data of colorectal cancer with the comparison of gene expression data from UALCAN database. OncoLnc analyses 110 cancer sample for both low and high SYK expression on the contrary UALCAN analyses 71 COAD sample for higher expression and 208 for lower expression. According to the graph from OncoLnc and UALCAN databases, the down expression was observed for favorable prognosis (Figure 6e and 6f).

J.

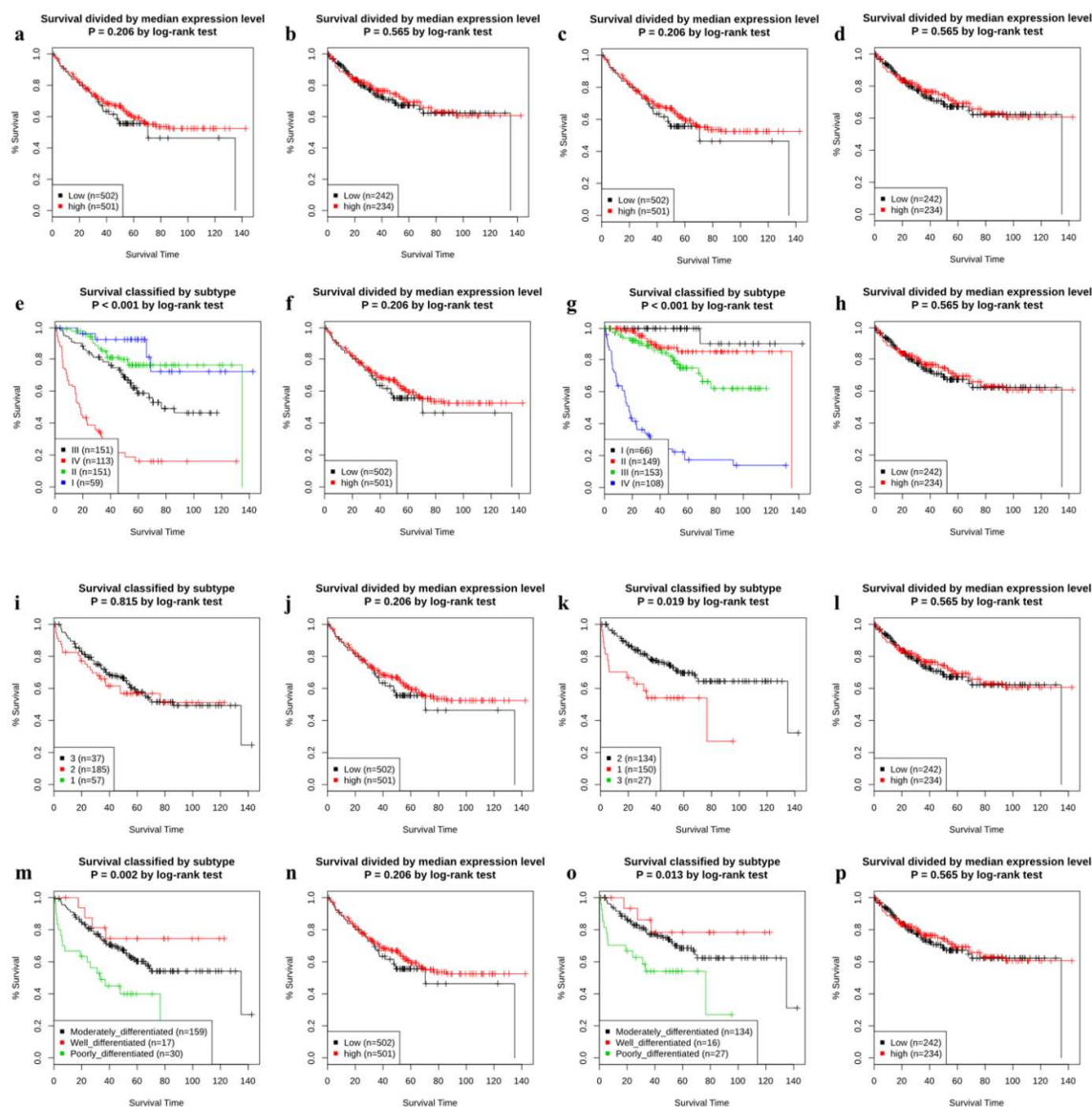


Figure 7. Prognostic value of SYK gene in human colorectal cancer from GENT2 database. (a) KM-plots divided by median cut-off- overall survival; (b) KM-plots divided by median cut-off- diseases specific survival; (c) KM-plots divided by median cut-off- overall survival for DUKE_Stage; (d) KM-plots divided by median cut-off- diseases specific survival for DUKE_Stage; (e) KM-plots divided by subtypes- overall survival for AJCC_Stage (f) KM-plots divided by median cut-off- overall survival for AJCC_Stage; (g) KM-plots divided by subtypes- diseases specific survival for AJCC_Stage; (h)) KM-plots divided by median cut-off- diseases specific survival for AJCC_Stage; (i) KM-plots divided by subtypes- overall survival for Grade; (j) KM-plots divided by median cut-off- overall survival for Grade; (k) KM-plots divided by subtypes- diseases specific survival for Grade; (l) KM-plots divided by median cut-off- diseases specific survival for Grade; (m) KM-plots divided by subtypes- overall survival for Histology; (n) KM-plots divided by median cut-off- overall survival for Histology; (o) KM-plots divided by subtypes- diseases specific survival for Histology; (p) KM-plots divided by median cut-off- diseases specific survival for Histology.

Besides, Kaplan-Meier plots are provided by GENT2 database using subtypes and median cut-off where colon tissue contains Molecular Subtype, AJCC Stage, Duke Stage, Grade and Histology categories for subtype analysis. Similar survival was detected for

J.

both lower and higher level of SYK expression (Figure 7). Overall a positive correlation was noticed between the overexpression of SYK gene and poor prognostic predictor for colon cancer. Lower level of SYK expression was related to the more chances of patients' survival. Therefore, by analyzing prognosis value and survival plot, it can be concluded that SYK can be considered as a tumor suppressor gene for colorectal cancer.

3.4 Analysis of Correlated genes and Preparing Interaction Network

For the identification of the correlated genes of SYK for COAD, we employed two web-utilities namely, UALCAN and GEPIA. This correlation was considered to be significant when the Pearson-correlation coefficient value is greater than 0.47 (Supplementary Table 4). Both the databases revealed that SECISBP2 showed the highest positive correlation with SYK in COAD. We extracted 20 correlated genes from both websites to prepare an interaction network. GeneMANIA web-tool was used to investigate the interaction network. Those 20 correlated similar genes (TCF7, CDC14B, FKTN, TBC1D13, GLE1, CCDC170, NAA35, TSTD2, GOLGA1, TMEM8B, SNX30, RNF20, EFCAB14, PCDH19, FGGY, SECISBP2, ZNF782, SMC5, C6orf97 and KIAA0494) along with SYK was the employed to build network (figure 8a). The network was built by automatically selected weighting method where it showed 79.13% of co-expression and 20.87% of co-localization in the PPI network.

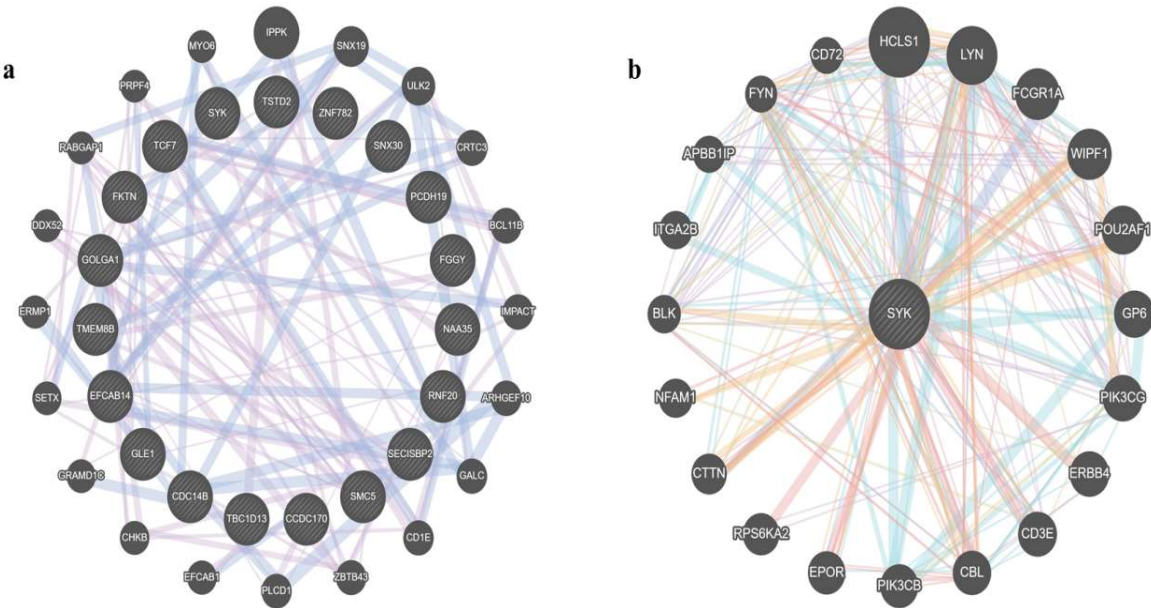


Figure 8. Protein-protein interaction network. (a) Interaction network of SYK with the positively correlated genes of COAD; (b) interaction network of SYK with the significantly interacted gene based on co-expression, genetic interactions, pathways, physical interactions, co-localization and predicted shared protein domains.

In addition, for the prediction of PPI network of SYK gene to express the relationship between SYK gene with the significantly interacted genes, GeneMania web-tool was also used. Automatically selected weighting method was selected for the network of 20 genes (HCLS1, LYN, FCGR1A, WIPF1, POU2AF1, GP6, PIK3CG, ERBB4, CD3E, CBL, PIK3CB, EPOR, RPS6KA2, CTTN, NFAM1, BLK, ITGA2B, APBB1IP, FYN, CD72) related with SYK through several attributes including Co-expression, Co-localization, Pathway, Physical Interactions, Shared protein domains, Predicted, and Genetic Interactions (figure 8b). This network shows 67.64% of physical interactions, 13.50% of co-expression, 6.35% of predicted, 6.17% of co-localization, 4.35% of pathway, 1.40% of genetic Interactions and 0.59% of shared protein domains.

3.5 Pharmacokinetics Analysis

J.

In this current investigation, six ligands (among them one used as control) were selected based on their ADMET results and the phytochemicals possess a better absorption rate on the human digestive system. In addition, the compounds exhibited a better excretion rate from the body after distribution and metabolism of the ligands along with it have no higher toxic effect on organs (mainly hepatocytes) (Table 2). Crystal structure and potential hydrogen and hydrophobic interactions of all ligands and proteins were displayed in Figure 9.

Table 2: The Complete Pharmacophore and Pharmacokinetics Profile of the Selected Ligand Molecules

Name of phytochemical s/ADMET Values	Physiochemical											Pharmacokinetics	
	Properties								Properties				
	M O	HBA	HBD	Log P	RB	IAb	TCI	LD50	HPT	AMT	MTHD	CaP	CTOR
01. Capecitabine (Control)	359.35 4	8	3	0.7602	6	68.027	1.054	2.459	Yes	No	1.051	0.255	2.401
02. Glabridin	324.37 6	4	2	4.0007	1	94.164	0.121	2.523	No	No	-0.395	1.284	1.148
03. Curcumin	368.38 5	6	2	3.3699	8	82.19	-0.002	1.833	No	No	0.081	-0.093	2.228
04. Kaempferol	286.23	6	4	2.2824	1	74.29	0.477	2.449	No	No	0.531	0.032	2.505
05. Quercetin	302.23 8	7	5	1.988	1	77.207	0.407	2.471	No	No	0.499	-0.229	2.612
06. Genistein	270.24	5	3	2.5768	1	93.387	0.151	2.268	No	No	0.478	0.9	2.189

Units of Physiochemical Properties: MO- Molecular weight (g/mol); HBA- Hydrogen bond acceptor; HBD- Hydrogen bond donor; LogP- Estimated octanol/ Coefficient of liquid fraction; RB- Rotatable bonds

Units of Pharmacokinetics Properties: IAb-Intestinal absorption (% absorbed); TCI-Total clearance (log ml/min/kg); LD50- Acute toxicity of the oral rat; HPT- Hepatotoxicity; AMT- AMES toxicity; MTHD- Maximum tolerated human dose (log mg/kg/day); CaP- Caco2 Permeability (Numeric as log Papp in 10⁻⁶ cm/s); CTOR- Cronic Toxicity in Oral Rat (Numeric as log mg/kg_bw/day)

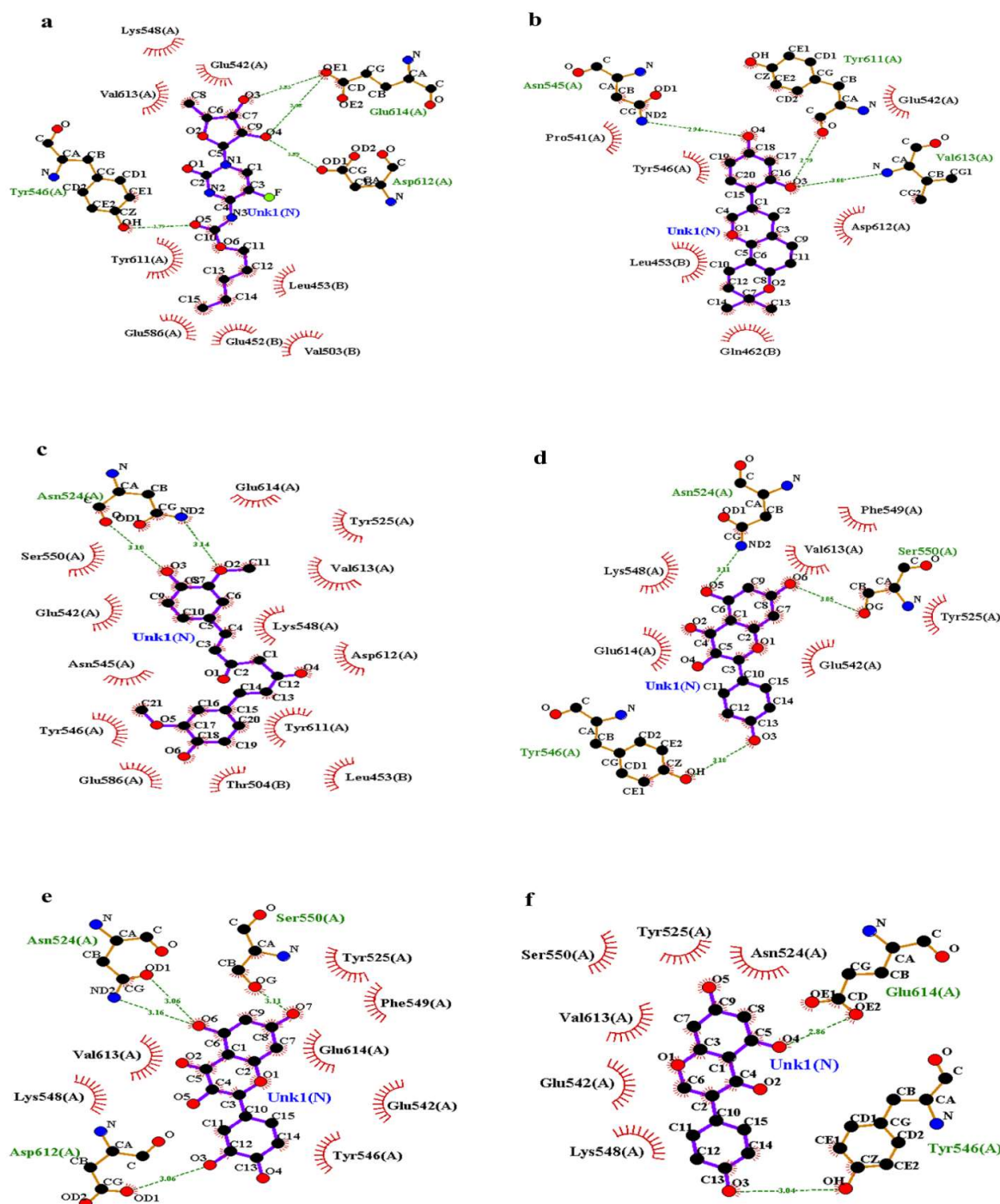


Figure 9. The docked conformation of the SYK-lead molecule complexes represent the potential hydrogen and hydrophobic interactions, with hydrogen bonds depicted as olive green dotted lines of a specific distance (Å) and hydrophobic interactions depicted as red color oval-shaped structures. Thin red lines with ellipses in protein residues are shown by the red mark, and hydrophobic interactions are shown by thin red lines with ellipses in protein residues. The protein residues that are primarily similar in 3D positions to the residues are marked by the red color spiked arc ellipses. Capecitabine (Control)-SYK complex (A), Glabridin-SYK complex (B), Curcumin-SYK complex (C), Kaempferol-SYK complex (D), Quercetin-SYK complex (E), and Genistein-SYK complex (F).

J.

3.6 Molecular Docking and Post Docking Data Analysis

To uncover the best ligand candidates among the selected aromatic phytochemicals by using molecular docking process through the PyRx 0.8 package platform. Basically, the PyRx 0.8 tools displayed the best docking score between macromolecule and ligands. Also, Table 2 displayed all the highest binding score. Here, the Capecitabine is considered as a control ligand that shows the docking score at -6.5 Kcal/mol with the grid center X=36.7526, Y=3.6328, Z=25.3921 along with the dimension of grid box at X x Y x Z at 35.9128 Å, 33.0428 Å, 25.0000 Å. The aromatic ligand Glabridin have displayed the best fitting score i.e., -8.2 Kcal/mol via using the control drug grid parameters. Furthermore, other potential phytochemicals for instant-Curcumin, Kaempferol, Quercetin, Genistein have been possessed the best docking score -8.0, -7.3, -7.2, -7.1 Kcal/mol respectively, which were represented by Table 3.

Ligplot+ V.2.2 showed all the fitting status of docked complex where non-covalent and hydrophobic interaction find out. The “SYK kinase domain-Capecitabine (control drug)” complex stabilized via four hydrogen bonds (with Asp612-2.89 Å, Glu614-2.85 Å, Glu614-3.08 Å, and Tyr546-2.79 Å) as well as eight hydrophobic bond interactions (with Glu452, Glu586, Glu542, Leu453, Lys548, Tyr611, Val613, Val503). Table 2 also represented that the “SYK kinase domain-Glabridin” complex stabilized through three hydrogen bonds interaction (Asn545- 2.94 Å, Tyr611- 2.79 Å, and Val613- 3.08 Å) and six hydrophobic bonds (Asp612, Gln462, Glu542, Leu453, Pro541, Tyr546) whereas two hydrogen bonds (Asn524- 3.10 Å and Asn524- 3.14 Å) and thirteen hydrophobic bonds (Asn545, Asp612, Glu542, Glu586, Glu614, Leu453, Lys548, Ser550, Thr504, Tyr525, Tyr546, Tyr611, Val613) were showed by Curcumin against the receptor. Kaempferol formed the four hydrogen bonds (Asn524-3.11 Å, Ser550- 3.05 Å and Tyr546-3.18 Å) and six hydrophobic bonds (Glu614, Glu542, Lys548, Phe549, Tyr525, Val613) interaction. Conversely, Quercetin stabilized by four hydrogen bonds (Asn524- 3.06 Å, Asn524- 3.16 Å, Asp612- 3.06 Å, and Ser550- 3.11 Å) and seven hydrophobic bonds (Glu542, Glu614, Lys548, Phe549, Tyr525, Tyr546, Val613). Finally, Genistein formed two hydrogen bonds (Glu614- 2.86 Å and Tyr546- 3.04 Å) and seven hydrophobic bonds (Asn524, Glu542, Lys548, Ser550, Tyr525, Val613) interaction against 4XG3 receptor molecule.

Table 3: Binding affinity of our interested ligands with 4GX3 receptor macromolecule and comprehensive intermolecular interaction between them as compared with control molecule.

Ligands Name	Binding Affinity (Kcal/mol)	Amino Acid Involved Interaction	
		Hydrogen Bond Interaction	Hydrophobic Bond Interaction
Capecitabine (Control)	-6.5	Asp612 (2.89 Å), Glu614 (2.85 Å), Glu614 (3.08 Å) and Tyr546 (2.79 Å)	Glu452, Glu586, Glu542, Leu453, Lys548, Tyr611, Val613, Val503
Glabridin	-8.2	Asn545 (2.94 Å), Tyr611 (2.79 Å) and Val613 (3.08 Å)	Asp612, Gln462, Glu542, Leu453, Pro541, Tyr546
Curcumin	-8.0	Asn524 (3.10 Å) and Asn524 (3.14 Å)	Asn545, Asp612, Glu542, Glu586, Glu614, Leu453, Lys548, Ser550, Thr504, Tyr525, Tyr546, Tyr611, Val613
Kaempferol	-7.3	Asn524 (3.11 Å), Ser550 (3.05 Å) and Tyr546 (3.18 Å)	Glu614, Glu542, Lys548, Phe549, Tyr525, Val613
Quercetin	-7.2	Asn524 (3.06 Å), Asn524 (3.16 Å), Asp612 (3.06 Å), and Ser550 (3.11 Å)	Glu542, Glu614, Lys548, Phe549, Tyr525, Tyr546, Val613
Genistein	-7.1	Glu614 (2.86 Å) and Tyr546 (3.04 Å)	Asn524, Glu542, Lys548, Ser550, Tyr525, Val613.

3.7 Molecular Dynamics Simulation

The molecular dynamics simulation trajectories were used to explore multiple simulation descriptors and understand the complexes flexibility. Figure 10 demonstrated that the root mean square deviation of the simulation complexes were initially experienced an

J.

upper trend in RMSD till 5ns and experienced a stable trend till 30ns. Therefore, the complexes were raised their RMSD profile again after 40ns but all complexes had steady profile. The complexes did not fluctuate much in RMSD except Curcumin and Quercetin. Therefore, the other four complexes had steady trend in RMSD which indicates the complexes rigidity. The complexes had RMSD below 2.5\AA except for Quercetin and Curcumin. The root mean square fluctuations of the amino acid residues from the protein-ligand complexes were checked to understand the flexibility across the amino acid residues. The maximum amino acid residues from the complexes had RMSF lower than 2.5\AA which indicates the lower flexibility Figure 10.

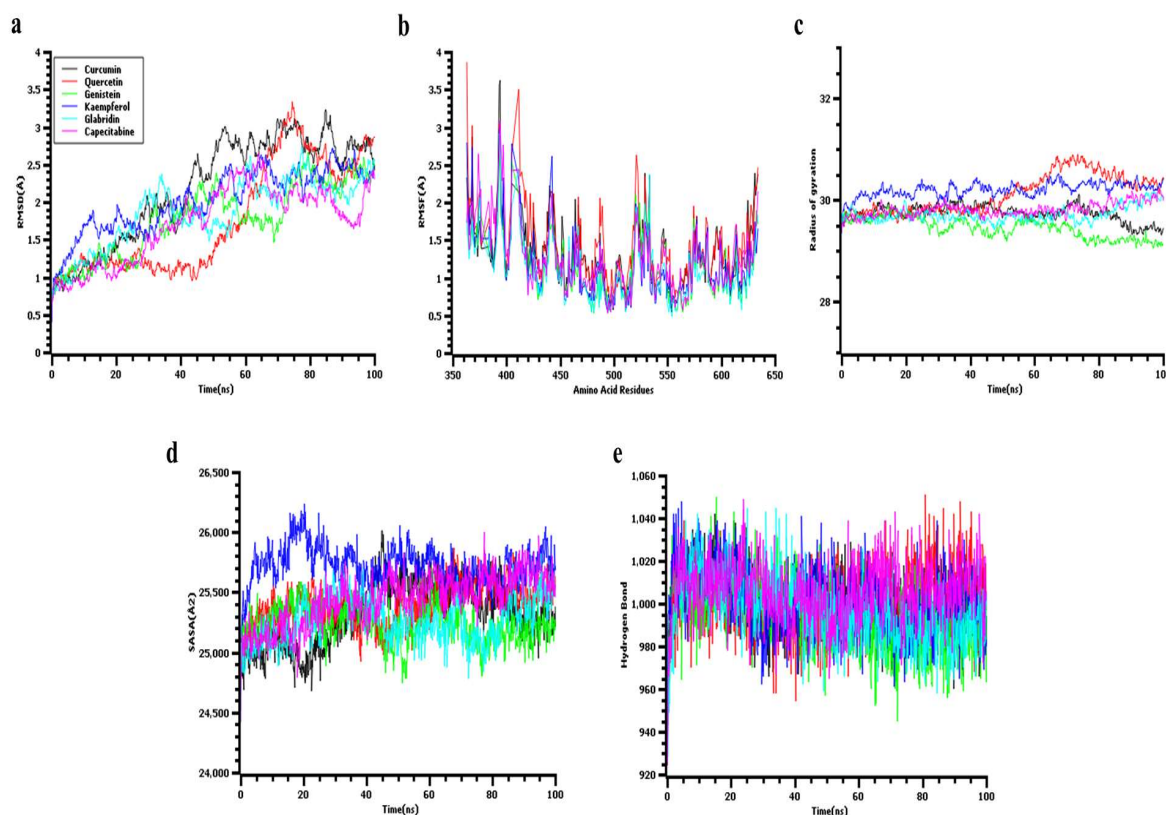


Figure 10. 10a (The RMSD values were obtained individually by the molecular dynamic simulation of each ligand with the spleen tyrosine kinase (SYK) targeted receptor); 10b (RMSF values obtained from a molecular dynamic simulation for each ligand when simulated precisely and independently with the target receptor); 10c (Gyration of radius (Rg) throughout the simulation); 10d (Available surface area of the solvent (SASA)); 10e (Maximum number of H-bonds during the simulation).

The solvent-accessible surface area from the simulation trajectories was analyzed to understand the change in the protein surface area. The Kaempferol complex had higher SASA trend at the beginning phase which indicates the expansion in the protein surface area. After 30ns, the Kaempferol complex had a similar SASA profile. The Glabridin complex had a comparatively lower SASA profile in other complexes, and this trend indicates the truncated nature of the complexes. The other complexes had steady profile and did not deviate much in SASA Figure 10.

Moreover, the radius of gyration of the protein complexes was checked to understand the protein mobility and rigidity. The protein-ligand complexes had a straight line in radius of gyration profile and observed little fluctuations. The two complexes; Quercetin and Kaempferol had a higher Rg trend after 60ns which might occur due to the labile nature of these complexes. Therefore, the hydrogen bond of the simulation systems was

J.

analyzed as they play an important role in stabilizing the complexes Figure 10. The six complexes had stable hydrogen bond trend in simulation and therefore no significant fluctuations were observed.

4. Discussion

Normal cell division is the result of strict control of DNA replication that maintains cellular growth, differentiation, and tissue homeostasis, and whenever this control breaks down it leads to uncontrolled cell growth i.e., tumor and/or cancer. Moreover, this multi-stage cell cycle process involves hundreds of distinct genes that cause autonomous growth of cancer cells due to mutations, amplifications, rearrangements, and modifications [50]. Cancer is a broad-spectrum lethal disease that is illustrated by uncontrolled cell division in the body along with the cancer cells cross their boundaries and invade the adjoining parts of the body, thus it spreads to other parts of the body from its origination [51]. Among the various types of cancers like bladder cancer, colorectal cancer, breast cancer, kidney cancer, lung cancer, and oropharyngeal cancer, colorectal cancer plays a significant cause of death in the entire world. Colorectal cancer, usually referred to as colon cancer, rectal or bowel cancer grows out of the colon or rectum [52]. It is recommended that the abnormal expression level of the SYK (Spleen tyrosine kinase) gene is associated with the development of colorectal cancer [53]. In recent times, Spleen tyrosine-protein kinases are prominent and effective drug targets due to their intimate connection with cell cycle regulation and association with tumorigenesis [54]. However, the distinct role of SYK in colorectal cancer is yet to be discovered. Therefore, in this study, we performed integrative bioinformatics analysis through several powerful publicly available datasets that indicated SYK may perform as a potential prognostic biomarker in the treatment of colorectal cancer.

The mRNA expression patterns of SYK gene across different types of cancer samples, particularly in colorectal cancer samples with their analogous normal control and SYK gene was upregulated in BRCA, CESC, CHOL, ESCA, GBM, HNSC, LAML, READ, STAD, UCEC, and down-regulated in BLCA, KIRC, KIRP, PRAD, SKCM, THCA, etc (Figure 1). Moreover, the SYK is significantly upregulated in COAD than the normal samples based on different clinicopathological parameters (Figure 2 and Figure 3). Thus, there is a positive correlation between SYK expression and increase risk of tumor metastasis formation. Several reports also demonstrate that the role of the SYK gene in diverse cancer types as a potential prognostic marker like in human breast carcinomas [55], cervical carcinogenesis [56], human lung cancer [57], ovarian cancer [58], and neck [59], glioblastoma [60], etc. Here, the divergent role of SYK in cancer may be due to phosphorylation or methylation at the promoter region. To examine this, we explored the SYK gene promoter methylation level across COAD based on different clinicopathological parameters (Figure 4). Promoter methylation level fluctuates significantly across different stages of COAD. The methylation level is higher in tumor samples than the normal samples. Zuli Yang et al., (2013) reported a correlation between clinical relevance and the methylation of SYK gene in colorectal cancer patients where they found SYK methylation in 48.6% CRC tissue samples and 57.1% in cell lines, respectively [61]. SYK expression could be resorted by the demethylation agent. However, this required further in-depth investigation. Furthermore, to investigate the functional relevance of the SYK gene in colorectal cancer development, we analyzed the mutations and copy number alterations of the SYK protein sequence based on 10 cancer studies. In total 57 mutations were found of which 40 were missense type and 17 were truncating type (Figure 5). Moreover, the highest alteration frequency of 4.17% was found in colorectal adenocarcinoma. Some researchers also reported that identifying genomic regions that undergo intermittent alteration might provide a powerful way to discover oncogenes in human cancers [62]. From this result, it is evident that SYK might play an active role in CRC development.

The prognostic relevance of the SYK gene for CRC by using KM plotter from various websites including GEPIA, OncoLnc, UALCAN, and GENT2 and analysis of survival

J.

curves expressed that lower expression of SYK is related to higher survival tendency for overall survival (OS) data. Besides diseases free survival (DFS) condition, it also seemed to be the higher survival for lower expression level (Figure 6). Further, down expression was observed for better prognosis provided by the data from different colon tissue subtypes which contains Molecular Subtype, AJCC Stage, Duke Stage, Grade, and Histology categories (Figure 7). It is necessary to identify the particular gene which is responsible for altered gene expression and survival probability differences because that gene can be a cancer biomarker [63]. The validation of potential biomarker is associated with the survival rate [64]. Even the SYK gene is a potential marker and tumor suppressor for the reduced expression in human breast cancer (BC) and human pancreatic ductal adenocarcinoma (PDAC) [65,66]. Like this, a positive correlation was noticed between the overexpression of the SYK gene and the poor prognostic probability for colorectal cancer. Besides, survival plots from various databases suggest that SYK expression can contribute for the development and prognosis of colorectal cancer which can also be considered as a tumor suppressor gene.

To find out the activity of the SYK gene, analyze the correlation and co-expression where Pearson correlation coefficients are used to measure the relation. Here searched the UALCAN and GEPIA database and identified positively correlated genes of SYK in COAD tissues. Twenty correlated genes from both websites were collected to prepare an interaction network from GeneMANIA where Pearson-CC >0.47 was considered as moderate significant. Both the database confirmed that the gene, SECISBP2 shares the highest Pearson correlation coefficients value with SYK. Even, SECISBP2 is considered a novel therapeutic target for diffuse large B-cell lymphoma, and mutations in this gene have been related to abnormal thyroid hormone metabolism. The interaction showed 79.13% of co-expression which means those genes are linked with each other by their expression level in similar across conditions and 20.87% of co-localization (Figure 8a) which indicates the linked genes are expressed in the same tissue or same cellular location. Besides to determine another interaction network, GeneMANIA was also employed and the biologically processed based 'automatically selected weighting method' was utilized for the network (Figure 8b). All the interacted twenty genes are directly or indirectly connected with different cancer. HCLS1 contains the potential prognostic characteristics with CRC [67], Lyn can be used as a prime target for the treatment of hormone-refractory human prostate cancer (PC), and Lyn is also involved in the activation of CRC [68,69], PTEN was identified as a tumor suppressor and a therapeutic target for cancer [70], WIPF1 is a candidate of the breast, brain, and colorectal cancer prognosis [71], as an ErbB subfamily, the ErbB4 gene plays a significant role in breast carcinogenesis with breast cancer prognostics and therapy [72], CBL-b can regulate cancer metastasis [73], EPOR contributes to breast cancer progression [74], CTTN can be used as new molecular therapeutic targets for Esophageal squamous cell carcinoma (ESCC) and CTTN also promotes the proliferation of CRC cells [75,76] and so on. Here, these 2 categories of protein-protein interaction networks depicted that SYK has interacted with those genes that are somehow involved in tumor suppression or tumor promotion.

There are many therapeutic agents available like chemotherapy, radiation therapy, immunotherapy, surgery, and stem cell transplantation, but all these methods process severe side effects and high cost [77]. So, the development of phytochemicals-based therapeutic drugs development is a crying need to treat colorectal cancer because the phytochemicals-based therapeutic agents possess fewer side effects on the human body. In recent times, in silico drug design method has become very popular because it increases the speed of drug development through analyzing the result of pharmacophore profiling, molecular docking, post-docking screening, molecular dynamic simulation, and prediction of noble compounds against various disease [78]. Here selected the six phytochemicals among them five phytochemicals (Figure 1) are functionally active and the remaining on phytochemical is used as a control which can be used as a perfect drug against colorectal cancer.

J.

ADMET profiling is an effective strategy for substantially saving the costs of drug development and offering 'fact tests' and secondary complimentary opinions for high-performance assays [79,80]. Through using pkCSM-pharmacokinetics & Swiss ADME online servers, the pharmacokinetic properties of those selected six phytochemicals and the data was characterized via Table 2 [81]. Under Pharmacokinetics study, the pharmacophore properties (Lipinski's rules of 5) of all selected ligands were carried out for instant-molecular weight, the number of rotatable bonds, log p-value, donors and acceptors of hydrogen bonds, and the degree of infringement (violation level). All selected phytochemicals showed no permeability to the blood-brain barrier, better human intestinal absorption, no carcinogenicity, no hepatotoxicity, and no AMES toxicity for both humans and mice, except for the control ligand. After their digestion, both substances excrete a small level of toxic waste.

Molecular docking is the process that estimates how two or more molecules bind together with the best structural confirmation as well as the lowest binding energy. More significant and efficient drug candidates were selected depending on the scoring value of molecular docking via the PyRx virtual screening tool [Version: 0.8] [82]. The docking operation for five phytochemicals as well as a control molecule with receptor 4GX3 and found the docking score of control was -6.5 Kcal/mol. Among the five phytochemicals, Glabridin showed the best binding energy -8.2 whereas Genistein was -7.1. Afterwards, an effective investigative tool Ligplot + [Version: 2.2] that generally runs depending on the Java interface was used to analyze the protein-ligand interaction from the 2D protein-ligand interaction scheme. Here, it is important to mention that Ligplot + only runs the PDB file that was retrieved from the PyMol visualizer tool [Version: 2.4.1].

Molecular dynamics simulation (MDS) has been demonstrated as a diverse technique for investigating biomolecular interactions and is used to evaluate the interaction between the protein structure and function for modern drug discovery and the performance data from dynamic trajectory [83,84]. In this current investigation, the dynamic simulator YASARA (Version 11.9.18) has retained all-physiological and physio-chemical parameters (Temp-310K, pH-7.4) and has to add the compressors (Na⁺, Cl⁻) in 100 ns [85] and the performance data assess the dynamic condition of the selected active compound's to suppress the soluble receptor molecules (SYK kinase). Most notably, the Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Gyration of Radius (Rg), Hydrogen Bond Number, and the Solvent-Accessible Surface Area (SASA) are also analyzed by this simulation tool [86]. To quantify the stability of the protein structure and to predict conformation shifts, the root mean square deviation (RMSD) [87] of the selected SYK kinase soluble protein backbone was used, most importantly the lower value shows the most stable complex. RMSD value of less than 1.5 Å typically shows greater accuracy in the docking concerning the effective binding position. In this present study, the RMSD values of protein-ligands interacted compound in the acceptable range, i.e. minimum values less than 2 Å (the lowest value for Capecitabine is approximately 1.0 Å), whereas maximum values did not exceed 3Å, indicating a better docking position alongside, the enzyme structure was not disrupted (Figure 10). The RMSF measures the mean protein molecule fluctuations from a reference location, and the residue level fluctuations are portrayed in the RMSF plots. Furthermore, it is essential to differentiate local changes along the protein chain [88]. The fluctuation of maximum ligands molecule within 1.5 to 2.5 Å, while the quercetin and curcumin molecule showed a fluctuation rate of not more than 4 Å, which indicates that the rate of flexibility is a significant amount (Figure 10). The gyration radius (Rg) measures the distance between the center of mass and both protein termini. Therefore, this parameter measures the compactness of the protein molecule and gives us a deeper understanding of the protein's folding properties [89]. Additionally, the higher Rg value displayed the slackpacking whereas the lower Rg value reveals the compact packing [90]. Figure 10 illustrates all Rg values among which the best matching capacity was 28.4 and Glabridin, 28.45 compared with the control ligand (Capecitabine) value A. Figure shows all Rg values. The compounds for instant Glabridin and Curcumin,

J.

alongside their enzymes, have better compactness, and the numerical values presented in both cases 28.45. The weaker binding capacity is 28.39 Quercetin and delta Genistein. Similarly, the value of Solvent-accessible surface area (SASA) determines the effectiveness of interaction among the macromolecule-ligand complex, whereas the interaction between enzyme surface area and water depicts the amount of energy within per area of ligand and macromolecule [91]. One of the highest SASA values complies with the unstable structure in which hydrophobic amino acid residues are in close contact with the water molecule [92]. According to the SASA result, the Kaempferol has been followed by the highest values from SASA (26,250 Å²), but Capecitabine, Kaempferol, and Glabridine have the higher values from SASA, however, both Turco and Genistein have the lowest value from SASA and bind most precisely to the receptor (Figure 10). The Conformational stability among the macromolecule and ligands was identified by the complete number of intermolecular hydrogen bonds [93] and the lowest number of hydrogen bonds in the versatile Kaempferol and Glabridin has determined the macromolecules-linking conformational stability of 1049 and 928 that is more stable than the control referred to in this research (Figure 10).

However, it's an in silico based research study there have no sufficient evidences via in silico research approaches on colorectal cancer by targeting the SYK kinase. Our research study will focus on a new insight on colorectal cancer treatment in which more wet lab and clinical based research studies will be required to valid these drug-like molecules for the treatment of colorectal cancer by targeting the SYK kinase.

5. Conclusion

In conclusion, we used multiple online bioinformatics platforms and web resources to perform a systematic analysis of SYK expression, methylation, mutations, and CNAs, associated genes, survival status, and prognostic value in various human cancers especially for colorectal cancer. Additionally, the current findings demonstrated the critical role of SYK expression and potential SYK-related pathways in the development of human colorectal cancer. Our research may shed new light on SYK kinase as a novel biomarker and therapeutic target for colorectal cancer, thereby assisting in the translation of genomic information into clinical practice. Furthermore, the ADMET analysis, cellular docking, and molecular dynamics are effectively implicated to find out the right phytochemicals against human colorectal cancer targeted spleen tyrosine kinase (SYK). Among the 500 phytochemicals tested, The strong interactions and binding affinity with all or at least one of the catalytic residues remain in human colorectal cancer spleen tyrosine kinase (SYK) are Capecitabine (Control), Glabridin, Curcumin, Kaempferol, Quercetin, and Genistein. These protein-ligand complexes also show several non-covalent interactions called hydrogen-binding, hydrophobic and electrostatic interactions. MDS findings suggest that the physiological environment has observed the most stable protein-ligands interaction and that they interact more frequently with spleen tyrosine kinase (SYK) through the hydrogen bonds. The pharmacokinetic and ADMET analyzes have shown their effectiveness as drug molecules with no cellular toxicity. It can be established that most of the bioactive compounds we have identified, will show significant efficacy and can be used for designing successful anti-cancer drugs for human colorectal cancer.

Author Contributions: Data curation, Md. Abu Kaium and Priyanka Paul; Formal analysis, Dipta Dey, Atikur Rahman, Shafi Mahmud and Md. Abu Saleh; Funding acquisition, Bonglee Kim; Investigation, Bonglee Kim; Methodology, Partha Biswas, Dipta Dey and Md. Nazmul Hasan; Resources, Md. Aminul Islam; Software, MD. Hasanur Rahman; Supervision, Bonglee Kim; Visualization, Han Gyeul Song; Writing – original draft, Tasmina Ferdous Susmi; Writing – review & editing, Md Rezanur Rahman and Md. Ataur Rahman.

Data Availability Statement: The datasets used and/or analyzed during this study are available from the corresponding authors upon request.

J.

Acknowledgments: This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R111A2066868), the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A5A2019413), a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HF20C0116), and a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number : HF20C0038)

Conflicts of Interest: The authors declare no conflict of interest.

References

1. W.H. Organization, WHO report on cancer: setting priorities, investing wisely and providing care for all, (2020).
2. S.J. Henley, E.M. Ward, S. Scott, J. Ma, R.N. Anderson, A.U. Firth, C.C. Thomas, F. Islami, H.K. Weir, D.R. Lewis, R.L. Sherman, M. Wu, V.B. Benard, L.C. Richardson, A. Jemal, K. Cronin, B.A. Kohler, Annual report to the nation on the status of cancer, part I: National cancer statistics, *Cancer*. 126 (2020) 2225–2249. <https://doi.org/10.1002/cncr.32802>.
3. W. Organization, Latest global cancer data: Cancer burden rises to 19.3 million new cases and 10.0 million cancer deaths in 2020, (n.d.). https://www.iarc.who.int/wp-content/uploads/2020/12/pr292_E.pdf.
4. R.L. Siegel, K.D. Miller, H.E. Fuchs, A. Jemal, *Cancer Statistics, 2021*, CA. *Cancer J. Clin.* 71 (2021) 7–33.
5. W. Organization, All cancers Source: Globocan 2020, (n.d.). <https://gco.iarc.fr/today/data/factsheets/cancers/39-All-cancers-fact-sheet.pdf>.
6. M. Arnold, M.S. Sierra, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global patterns and trends in colorectal cancer incidence and mortality, *Gut*. 66 (2017) 683–691.
7. K.D. Miller, L. Nogueira, A.B. Mariotto, J.H. Rowland, K.R. Yabroff, C.M. Alfano, A. Jemal, J.L. Kramer, R.L. Siegel, *Cancer treatment and survivorship statistics, 2019*, CA. *Cancer J. Clin.* 69 (2019) 363–385.
8. P. Rawla, T. Sunkara, A. Barsouk, *Epidemiology of colorectal cancer: Incidence, mortality, survival, and risk factors*, *Przeglad Gastroenterol.* 14 (2019) 89.
9. A. Mócsai, J. Ruland, V.L. Tybulewicz, The SYK tyrosine kinase: a crucial player in diverse biological functions, *Nat. Rev. Immunol.* 10 (2010) 387–402.
10. Y. Kuno, A. Abe, N. Emi, M. Iida, T. Yokozawa, M. Towatari, M. Tanimoto, H. Saito, Constitutive kinase activation of the tm-Syk fusion gene in myelodysplastic syndrome with t (9; 12)(q22; p12), *Blood*. 97 (2001) 1050–1055.
11. P.J. Coopman, M.T. Do, M. Barth, E.T. Bowden, A.J. Hayes, E. Basyuk, J.K. Blacato, P.R. Vezza, S.W. McLeskey, P.H. Mangeat, The Syk tyrosine kinase suppresses malignant growth of human breast cancer cells, *Nature*. 406 (2000) 742–747.
12. M.O. Krisenko, R.L. Geahlen, Calling in SYK: SYK's dual role as a tumor promoter and tumor suppressor in cancer, *Biochim. Biophys. Acta BBA-Mol. Cell Res.* 1853 (2015) 254–263.
13. J. Blacato, A. Graves, B. Rashidi, M. Moroni, L. Tchobe, M. Ozdemirli, B. Kallakury, K.H. Makambi, C. Marian, S.C. Mueller, SYK allelic loss and the role of Syk-regulated genes in breast cancer survival, *PloS One*. 9 (2014) e87610.
14. P.J. Coopman, S.C. Mueller, The Syk tyrosine kinase: a new negative regulator in tumor growth and progression, *Cancer Lett.* 241 (2006) 159–173.
15. K. Repana, K. Papazisis, P. Foukas, R. Valeri, A. Kortsaris, E. Deligiorgi, D. Kyriakidis, Expression of Syk in invasive breast cancer: correlation to proliferation and invasiveness, *Anticancer Res.* 26 (2006) 4949–4954.
16. S. Ogane, T. Onda, N. Takano, T. Yajima, T. Uchiyama, T. Shibahara, Spleen tyrosine kinase as a novel candidate tumor suppressor gene for human oral squamous cell carcinoma, *Int. J. Cancer*. 124 (2009) 2651–2657.
17. T. Layton, C. Stalens, F. Gunderson, S. Goodison, S. Silletti, Syk tyrosine kinase acts as a pancreatic adenocarcinoma tumor suppressor by regulating cellular growth and invasion, *Am. J. Pathol.* 175 (2009) 2625–2636.
18. R. Ren, K. Tyryshkin, C.H. Graham, M. Koti, D.R. Siemens, Comprehensive immune transcriptomic analysis in bladder cancer reveals subtype specific immune gene expression patterns of prognostic relevance, *Oncotarget*. 8 (2017) 70982.
19. Z. Yang, L. Huo, H. Chen, B. Ni, J. Xiang, L. Kang, L. Wang, J. Peng, Y. Yuan, J. Wang, Hypermethylation and prognostic implication of Syk gene in human colorectal cancer, *Med. Oncol.* 30 (2013) 586.
20. M.M. Manson, A. Gescher, E.A. Hudson, S.M. Plummer, M.S. Squires, S.A. Prigent, Blocking and suppressing mechanisms of chemoprevention by dietary constituents, *Toxicol. Lett.* 112 (2000) 499–505.
21. E.D. Owuor, A.-N.T. Kong, Antioxidants and oxidants regulated signal transduction pathways, *Biochem. Pharmacol.* 64 (2002) 765–770.
22. L. Gong, Y. Li, A. Nedeljkovic-Kurepa, F.H. Sarkar, Inactivation of NF-κ B by genistein is mediated via Akt signaling pathway in breast cancer cells, *Oncogene*. 22 (2003) 4702–4709.
23. H. Hussain, N. Hamdan, E.U.-H. Sim, Anticancer and antimicrobial peptides from medicinal plants of Borneo island in Sarawak, *Adv. Tradit. Med.* (2020). <https://doi.org/10.1007/s13596-020-00504-z>.
24. D.R. Rhodes, S. Kalyana-Sundaram, V. Mahavisno, R. Varambally, J. Yu, B.B. Briggs, T.R. Barrette, M.J. Anstet, C. Kincaid-Beal, P. Kulkarni, Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles, *Neoplasia*. 9 (2007) 166–180.

J.

24. D.S. Chandrashekar, B. Bashel, S.A.H. Balasubramanya, C.J. Creighton, I. Ponce-Rodriguez, B.V. Chakravarthi, S. Varambally, UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses, *Neoplasia*. 19 (2017) 649–658.
25. S.-J. Park, B.-H. Yoon, S.-K. Kim, S.-Y. Kim, GENT2: an updated gene expression database for normal and tumor tissues, *BMC Med. Genomics*. 12 (2019) 1–8.
26. Z. Tang, B. Kang, C. Li, T. Chen, Z. Zhang, GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis, *Nucleic Acids Res.* 47 (2019) W556–W560.
27. E. Cerami, J. Gao, U. Dogrusoz, B.E. Gross, S.O. Sumer, B.A. Aksoy, A. Jacobsen, C.J. Byrne, M.L. Heuer, E. Larsson, The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data, *AACR*, 2012.
28. J. Gao, B.A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S.O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal, *Sci. Signal*. 6 (2013) p11–p11.
29. S.-J. Park, B.-H. Yoon, S.-K. Kim, S.-Y. Kim, GENT2: an updated gene expression database for normal and tumor tissues, *BMC Med. Genomics*. 12 (2019) 1–8.
30. Z. Tang, C. Li, B. Kang, G. Gao, C. Li, Z. Zhang, GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses, *Nucleic Acids Res.* 45 (2017) W98–W102.
31. J. Anaya, OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs, *PeerJ Comput. Sci.* 2 (2016) e67.
32. D.S. Chandrashekar, B. Bashel, S.A.H. Balasubramanya, C.J. Creighton, I. Ponce-Rodriguez, B.V.S.K. Chakravarthi, S. Varambally, UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses, *Neoplasia N. Y.* N. 19 (2017) 649–658. <https://doi.org/10.1016/j.neo.2017.05.002>.
33. D. Warde-Farley, S.L. Donaldson, O. Comes, K. Zuberi, R. Badrawi, P. Chao, M. Franz, C. Grouios, F. Kazi, C.T. Lopes, A. Maitland, S. Mostafavi, J. Montojo, Q. Shao, G. Wright, G.D. Bader, Q. Morris, The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function, *Nucleic Acids Res.* 38 (2010) W214–220. <https://doi.org/10.1093/nar/gkq537>.
34. M. Fluck, G. Zurcher, A.-C. Andres, A. Ziemiecki, Molecular characterization of the murine syk protein tyrosine kinase cDNA, transcripts and protein, *Biochem. Biophys. Res. Commun.* 213 (1995) 273–281.
35. S. Tsuchida, S. Yanagi, R. Inatome, J. Ding, P. Hermann, T. Tsujimura, N. Matsui, H. Yamamura, Purification of a 72-kDa protein-tyrosine kinase from rat liver and its identification as Syk: involvement of Syk in signaling events of hepatocytes, *J. Biochem. (Tokyo)*. 127 (2000) 321–327.
36. J. Dejmek, K. Leandersson, J. Manjer, A. Bjartell, S.O. Emdin, W.F. Vogel, G. Landberg, T. Andersson, Expression and signaling activity of Wnt-5a/discoidin domain receptor-1 and Syk plays distinct but decisive roles in breast cancer patient survival, *Clin. Cancer Res.* 11 (2005) 520–528.
37. S. Okamura, C.C. Ng, K. Koyama, Y. Takei, H. Arakawa, M. Monden, Y. Nakamura, Identification of seven genes regulated by wild-type p53 in a colon cancer cell line carrying a well-controlled wild-type p53 expression system, *Oncol. Res. Featur. Preclin. Clin. Cancer Ther.* 11 (1999) 281–285.
38. M. Alagumuthu, V.P. Muralidharan, M. Andrew, M.H. Ahmed, S.K. Iyer, S. Arumugam, Computational approaches to develop isoquinoline based antibiotics through DNA gyrase inhibition mechanisms unveiled through antibacterial evaluation and molecular docking, *Mol. Inform.* 37 (2018) 1800048.
39. Z. Yang, K. Lasker, D. Schneidman-Duhovny, B. Webb, C.C. Huang, E.F. Pettersen, T.D. Goddard, E.C. Meng, A. Sali, T.E. Ferrin, UCSF Chimera, MODELLER, and IMP: an integrated modeling system, *J. Struct. Biol.* 179 (2012) 269–278.
40. S. Subasri, V. Viswanathan, K. Manish, D. Velmurugan, Phytochemical analysis, molecular docking and molecular dynamics simulations of selected phytoconstituents from four herbs as anti-dotes for snake bites, *Clin Proteom Bioinform.* 1 (2016) 1–13.
41. D. Prasanth, M. Murahari, V. Chandramohan, S.P. Panda, L.R. Atmakuri, C. Guntupalli, In silico identification of potential inhibitors from Cinnamon against main protease and spike glycoprotein of SARS CoV-2, *J. Biomol. Struct. Dyn.* (2020) 1–15.
42. R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand–protein interaction diagrams for drug discovery, *ACS Publications*, 2011.
43. C.J. Dickson, B.D. Madej, A. Skjevik, R.M. Betz, K. Teigen, I.R. Gould, R.C. Walker, Lipid14: the amber lipid force field, *J. Chem. Theory Comput.* 10 (2014) 865–879.
44. H. Land, M.S. Humble, YASARA: a tool to obtain structural guidance in biocatalytic investigations, in: *Protein Eng.*, Springer, 2018: pp. 43–67.
45. I.S. Yadav, P.P. Nandekar, S. Shrivastava, A. Sangamwar, A. Chaudhury, S.M. Agarwal, Ensemble docking and molecular dynamics identify knoevenagel curcumin derivatives with potent anti-EGFR activity, *Gene*. 539 (2014) 82–90.
A. Swargiary, S. Mahmud, M.A. Saleh, Screening of phytochemicals as potent inhibitor of 3-chymotrypsin and papain-like proteases of SARS-CoV2: an in silico approach to combat COVID-19, *J. Biomol. Struct. Dyn.* (2020) 1–15.
46. M.S. Islam, S. Mahmud, R. Sultana, W. Dong, Identification and in silico molecular modelling study of newly isolated *Bacillus subtilis* SI-18 strain against S9 protein of *Rhizoctonia solani*, *Arab. J. Chem.* 13 (2020) 8600–8612.
47. E. Krieger, G. Vriend, New ways to boost molecular dynamics simulations, *J. Comput. Chem.* 36 (2015) 996–1007.

J.

48. M. Tarabichi, A. Antoniou, M. Saiselet, J.M. Pita, G. Andry, J.E. Dumont, V. Detours, C. Maenhaut, Systems biology of cancer: entropy, disorder, and selection-driven evolution to independence, invasion and "swarm intelligence," *Cancer Metastasis Rev.* 32 (2013) 403–421.
A. Jemal, R. Siegel, E. Ward, Y. Hao, J. Xu, T. Murray, M.J. Thun, Cancer statistics, 2008, CA. *Cancer J. Clin.* 58 (2008) 71–96.
49. U.F. QURESHI, M.N. ASLAM, M.N. ANSARI, M. KHAN, ROLE OF ASPIRIN AS PROPHYLAXIS AGAINST COLORECTAL CANCER, *Pak. Postgrad. Med. J.* 29 (2018) 16–19.
50. P.J. Coopman, M.T. Do, M. Barth, E.T. Bowden, A.J. Hayes, E. Basyuk, J.K. Blancato, P.R. Vezza, S.W. McLeskey, P.H. Mangeat, The Syk tyrosine kinase suppresses malignant growth of human breast cancer cells, *Nature.* 406 (2000) 742–747.
51. M.O. Krisenko, R.L. Geahlen, Calling in SYK: SYK's dual role as a tumor promoter and tumor suppressor in cancer, *Biochim. Biophys. Acta BBA-Mol. Cell Res.* 1853 (2015) 254–263.
52. P.J. Coopman, M.T. Do, M. Barth, E.T. Bowden, A.J. Hayes, E. Basyuk, J.K. Blancato, P.R. Vezza, S.W. McLeskey, P.H. Mangeat, The Syk tyrosine kinase suppresses malignant growth of human breast cancer cells, *Nature.* 406 (2000) 742–747.
53. S. Zhao, G. Sun, P.W. Tony, D. Ma, C. Zhao, Expression and methylation status of the Syk gene in cervical carcinoma, *Arch. Gynecol. Obstet.* 283 (2011) 1113–1119.
54. S.W. Dong, L. Ma, N. Xu, H.Q. Yan, H.Y. Liu, Y.W. Li, P. Zhang, Research on the reactivation of Syk expression caused by the inhibition of DNA promoter methylation in the lung cancer, *Neoplasma.* 58 (2011) 89.
55. S. Zhang, S. Deen, S.J. Storr, A. Yao, S.G. Martin, Expression of Syk and MAP4 proteins in ovarian cancer, *J. Cancer Res. Clin. Oncol.* 145 (2019) 909–919.
56. S. Luangdilok, C. Box, L. Patterson, K. Harrington, L. Pitkin, P. Rhys-Evans, O. Pornchai, S. Eccles, Syk tyrosine kinase is linked to cell motility and progression in squamous cell carcinomas of the head and neck, *Cancer Res.* 67 (2007) 7907–7916.
57. G. Moncayo, M. Grzmil, T. Smirnova, P. Zmarz, R.M. Huber, D. Hynx, H. Kohler, Y. Wang, H.-R. Hotz, N.E. Hynes, SYK inhibition blocks proliferation and migration of glioma cells and modifies the tumor microenvironment, *Neuro-Oncol.* 20 (2018) 621–631.
58. Z. Yang, L. Huo, H. Chen, B. Ni, J. Xiang, L. Kang, L. Wang, J. Peng, Y. Yuan, J. Wang, Hypermethylation and prognostic implication of Syk gene in human colorectal cancer, *Med. Oncol.* 30 (2013) 586.
59. R. Beroukhi, C.H. Mermel, D. Porter, G. Wei, S. Raychaudhuri, J. Donovan, J. Barretina, J.S. Boehm, J. Dobson, M. Urashima, The landscape of somatic copy-number alteration across human cancers, *Nature.* 463 (2010) 899–905.
60. Gy\Horffy, P. Surowiak, J. Budczies, A. Lanczky, Online survival analysis software to assess the prognostic value of biomarkers using transcriptomic data in non-small-cell lung cancer, *PloS One.* 8 (2013) e82241.
61. C. Gy\Horffy, A. Lanczky, Z. Szallasi, Implementing an online tool for genome-wide validation of survival-associated biomarkers in ovarian-cancer using microarray data from 1287 patients., *Endocr. Relat. Cancer.* 19 (2012) 197–208.
62. T. Toyama, H. Iwase, H. Yamashita, Y. Hara, Y. Omoto, H. Sugiura, Z. Zhang, Y. Fujii, Reduced expression of the Syk gene is correlated with poor prognosis in human breast cancer, *Cancer Lett.* 189 (2003) 97–102.
63. T. Layton, C. Stalens, F. Gunderson, S. Goodison, S. Silletti, Syk tyrosine kinase acts as a pancreatic adenocarcinoma tumor suppressor by regulating cellular growth and invasion, *Am. J. Pathol.* 175 (2009) 2625–2636.
64. F. Zhou, D. Tang, Y. Xu, H. He, Y. Wu, L. Lin, J. Dong, W. Tan, Y. Dai, Identification of microRNAs and their Endonucleolytic Cleaved target mRNAs in colorectal cancer, *BMC Cancer.* 20 (2020) 1–15.
65. M. Goldenberg-Furmanov, I. Stein, E. Pikarsky, H. Rubin, S. Kasem, M. Wygoda, I. Weinstein, H. Reuveni, S.A. Ben-Sasson, Lyn is a target gene for prostate cancer: sequence-based inhibition induces regression of human tumor xenografts, *Cancer Res.* 64 (2004) 1058–1066.
66. N. Su, L. Peng, B. Xia, Y. Zhao, A. Xu, J. Wang, X. Wang, B. Jiang, Lyn is involved in CD24-induced ERK1/2 activation in colorectal cancer, *Mol. Cancer.* 11 (2012) 1–13.
67. L. M Dillon, T. W Miller, Therapeutic targeting of cancers with loss of PTEN function, *Curr. Drug Targets.* 15 (2014) 65–79.
68. Staub, J. Groene, M. Heinze, D. Mennerich, S. Roepcke, I. Klamann, B. Hinzmann, E. Castanos-Velez, C. Pilarsky, B. Mann, An expression module of WIPF1-coexpressed genes identifies patients with favorable prognosis in three tumor types, *J. Mol. Med.* 87 (2009) 633–644.
69. M. Sundvall, K. Iljin, S. Kilpinen, H. Sara, O.-P. Kallioniemi, K. Elenius, Role of ErbB4 in breast cancer, *J. Mammary Gland Biol. Neoplasia.* 13 (2008) 259–268.
70. M. Paolino, A. Choidas, S. Wallner, B. Pranjic, I. Uribealago, S. Loeser, A.M. Jamieson, W.Y. Langdon, F. Ikeda, J.P. Fededa, The E3 ligase Cbl-b and TAM receptors regulate cancer metastasis via natural killer cells, *Nature.* 507 (2014) 508–512.
71. K.K. Chan, K.B. Matchett, J.A. Coulter, H.-F. Yuen, C.M. McCrudden, S.-D. Zhang, G.W. Irwin, M.A. Davidson, T. Rullicke, S. Schober, Erythropoietin drives breast cancer progression by activation of its receptor EPOR, *Oncotarget.* 8 (2017) 38251.
72. P. Lu, J. Qiao, W. He, J. Wang, Y. Jia, Y. Sun, S. Tang, L. Fu, Y. Qin, Genome-wide gene expression profile analyses identify CTTN as a potential prognostic marker in esophageal cancer, *PloS One.* 9 (2014) e88918.
73. X. Zhang, K. Liu, T. Zhang, Z. Wang, X. Qin, X. Jing, H. Wu, X. Ji, Y. He, R. Zhao, Cortactin promotes colorectal cancer cell proliferation by activating the EGFR-MAPK pathway, *Oncotarget.* 8 (2017) 1541.
74. M.M. Manson, A. Gescher, E.A. Hudson, S.M. Plummer, M.S. Squires, S.A. Prigent, Blocking and suppressing mechanisms of chemoprevention by dietary constituents, *Toxicol. Lett.* 112 (2000) 499–505.

J.

72. I.M. Kapetanovic, Computer-aided drug discovery and development (CADD): in silico-chemico-biological approach, *Chem. Biol. Interact.* 171 (2008) 165–176.
73. R. Mannhold, H. Kubinyi, G. Folkers, *Pharmacokinetics and metabolism in drug design*, John Wiley & Sons, 2012.
74. R.E. Notari, *Pharmacokinetics Molecular Modification: Implications in Drug Design Evaluation*, *J. Pharm. Sci.* 62 (1973) 865–881.
75. M.M. Matin, M.H. Roshid, S.C. Bhattacharjee, A.K. Azad, PASS Predication, antiviral, in vitro antimicrobial, and ADMET studies of rhamnopyranoside esters, *Med. Res. Arch.* 8 (2020).
76. S. Subasri, V. Viswanathan, K. Manish, D. Velmurugan, Phytochemical analysis, molecular docking and molecular dynamics simulations of selected phytoconstituents from four herbs as anti-dotes for snake bites, *Clin Proteom Bioinform.* 1 (2016) 1–13.
77. C.G. Priya Doss, C. Chakraborty, L. Chen, H. Zhu, Integrating in silico prediction methods, molecular docking, and molecular dynamics simulation to predict the impact of ALK missense mutations in structural perspective, *BioMed Res. Int.* 2014 (2014).
78. C.G.P. Doss, B. Rajith, C. Chakraborty, V. Balaji, R. Magesh, B. Gowthami, S. Menon, M. Swati, M. Trivedi, J. Paul, In silico profiling and structural insights of missense mutations in RET protein kinase domain by molecular dynamics and docking approach, *Mol. Biosyst.* 10 (2014) 421–436.
- E. Krieger, G. Vriend, New ways to boost molecular dynamics simulations, *J. Comput. Chem.* 36 (2015) 996–1007.
- F. Krieger, G. Koraimann, G. Vriend, Increasing the precision of comparative models with YASARA NOVA—a self-parameterizing force field, *Proteins Struct. Funct. Bioinforma.* 47 (2002) 393–402.
79. K. Sargsyan, C. Grauffel, C. Lim, How molecular size impacts RMSD applications in molecular dynamics simulations, *J. Chem. Theory Comput.* 13 (2017) 1518–1524.
80. T. Joshi, T. Joshi, P. Sharma, S. Chandra, V. Pande, Molecular docking and molecular dynamics simulation approach to screen natural compounds for inhibition of *Xanthomonas oryzae* pv. *Oryzae* by targeting peptide deformylase, *J. Biomol. Struct. Dyn.* (2020) 1–18.
81. M.A. Alamri, A. Altharawi, A.B. Alabbas, M.A. Alossaimi, S.M. Alqahtani, Structure-based virtual screening and molecular dynamics of phytochemicals derived from Saudi medicinal plants to identify potential COVID-19 therapeutics, *Arab. J. Chem.* 13 (2020) 7224–7234.
82. K. Kousar, A. Majeed, F. Yasmin, W. Hussain, N. Rasool, Phytochemicals from Selective Plants Have Promising Potential against SARS-CoV-2: Investigation and Corroboration through Molecular Docking, MD Simulations, and Quantum Computations, *BioMed Res. Int.* 2020 (2020).
83. A.A. Elfiky, W.M. Elshemey, Molecular dynamics simulation revealed binding of nucleotide inhibitors to ZIKV polymerase over 444 nanoseconds, *J. Med. Virol.* 90 (2018) 13–18.
84. S. Mahmud, E. Rahman, Z. Nain, M. Billah, S. Karmakar, S.C. Mohanto, G.K. Paul, A. Amin, U.K. Acharjee, M.A. Saleh, Computational discovery of plant-based inhibitors against human carbonic anhydrase IX and molecular dynamics simulation, *J. Biomol. Struct. Dyn.* (2020) 1–17.
85. P. Carbone, F. Müller-Plathe, Molecular dynamics simulations of polyaminoamide (PAMAM) dendrimer aggregates: molecular shape, hydrogen bonds and local dynamics, *Soft Matter.* 5 (2009) 2638–2647.