EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Volume: 12
Issue: 1
Year: 2025

GENE EXPRESSION AND ASSOCIATED PATHWAYS INVOLVED IN THE PROGRESSION OF DIFFERENT STAGES OF COLON CANCER BY NEXT-GENERATION SEQUENCING DATA

Sudheer Menon*[1], Shanmughavel Piramanayagam[2], Gopal Prasad Agarwal[3], Amanda Ramer Tait[4], Michelle Ramsay[5], Scott Hazelhurst[6] and Keiko Ozato[7]

[1]Department of Public Health, College of Applied Medical Sciences in Al-Namas, the Bisha University, Saudi Arabia.
[2]Department of Bioinformatics Bharathiar University, Coimbatore, India.
[3]Department of Biochemical Engineering and Biotechnology, Indian Institute of Technology, Hauz Khaz, New Delhi, India.
[4]Food Technology Department, University of Nebraska, Lincoln, United States of America.
[5,6]Wits Bioinformatics, Witwatersrand University, Johannesburg, South Africa.
[7]Keiko Ozato, NICHD, National Institutes of Health (NIH), United States of America.

***Corresponding Author: Sudheer Menon**
Department of Public Health, College of Applied Medical Sciences in Al-Namas, the Bisha University, Saudi Arabia.

Article Received on 04/11/2024 **Article Revised on 25/11/2024** **Article Accepted on 15/12/2024**

ABSTRACT
Colorectal cancers (CRC) is with about 1 million instances the 0.33 maximum not unusual place most cancers worldwide. Extensive studies are ongoing to decipher the underlying genetic styles with the wish to enhance early most cancers analysis and remedy. In this direction, the latest development in next-technology sequencing technology has revolutionized the sphere of most cancers genomics. However, one caveat of those research stays a big range of genetic versions diagnosed and their interpretation. We finish that with deep sequencing of tumor exomes one can be capable of expecting the microsatellite fame of CRC and similarly perceive probably clinically applicable mutations. **Aims and Objectives:** The foremost intention of this studies observe is to assess the gene expression and pathways which are related to the development of various tiers of Colon Cancer. This will assist keep away from all the matters which can worsen colon most cancers chances. It will even assist within the analysis of Colon Cancer.

INTRODUCTION
Colorectal most cancers are the 0.33 maximum, not unusual place most cancers with approximately 1 million instances worldwide. Over the ultimate decades, it has grown to be clear that CRC evolves via more than one pathway and that those pathways may be kind of described on the premise of molecular styles which includes the integrity of the mismatch restore system (MMR) or mutational and epigenetic styles. Deficiency within the MMR is meditated in DNA microsatellite instability (MSI) which has additionally been related to remedy outcome, however, which desires to be in addition established in extra medical research[1],[2],[3],[4],[5],[6] High-throughput Sanger sequencing research then again has proven that the mutation frequency of candidate most cancers genes is probably an awful lot better than predicted and that the specific mixture of mutations may affect the tumor's properties.[7],[8],[9],[10],[11] With the improvement of next-technology sequencing (NGS) technology, the sequencing throughput has dramatically extended and the prices have decreased. In addition, and especially crucial for medical settings, NGS may be implemented to formalin-constant and paraffin-embedded FFPE tissue fabric in addition to particularly degraded DNA that's mechanically organized in pathology departments or determined in historical DNA.[12],[13]

Several researches have used NGS technology for the identity of the underlying mutation in monogenetic diseases.[14],[15] However, best a restrained range of research document on next-technology sequencing to perceive new candidate most cancers genes; one of the earliest research-tested cytogenetically ordinary acute myeloid leukemia, and breast most cancers genomes.[16],[17] In addition, research on cancer and small-molecular lung most cancers molecular strains have furnished the primary insights into genomic changes precipitated with the aid of using ultraviolet mild publicity or tobacco smoke.[18],[19]

Methodology/Principal Findings

Here we provide the primary paintings on whole-exome NGS of number one colon cancers. We executed 454 whole-exome pyrosequencing of tumor in addition to adjoining now no longer affected ordinary colonic tissue from microsatellite strong (MSS) and microsatellite instable (MSI) colon most cancers sufferers and diagnosed extra than 50,000 small nucleotide versions for every tissue. According to predictions primarily based totally on MSS and MSI patho-mechanisms, we diagnosed 8 instances of extra somatic non-synonymous versions in MSI cancers than in MSS and we had been capable of reproducing the bring about 4 extra CRCs. Our bioinformatics filtering method narrowed down the price of maximum considerable mutations to 359 for MSI and 45 for MSS CRCs with anticipated altered protein functions. In each CRCs, MSI, and MSS, we determined somatic mutations within the intracellular kinase area of bone morphogenetic protein receptor 1A, BMPR1A, a gene in which to this point germline mutations are related to juvenile polyposis syndrome and display that the mutations functionally impair the protein function. To advantage perception into the genomes of microsatellite strong and unstable colorectal cancers and to perceive purposeful applicable mutational styles we used a hybridization-primarily based whole-exome DNA shooting method observed with the aid of using 454 next-technology sequencings.[20]

Applying stringent bioinformatics analyses, we narrowed down the quantity of functionally considerable somatic mutations in MSI to 359 and forty-five in MSS cancers, for that reason highlighting precise mutation styles relying on the microsatellite fame. We had been capable of affirming our consequences with the aid of using sequencing the exomes of 4 extra CRC instances (one MSI, 3 MSS) the use of exceptional enrichment and sequencing technology. Among those mutations are BRAF within the MSI most cancers and KRAS and TP53 within the MSS most cancers, in addition underscoring the validity of our choice method.[21]

Further purposeful characterizations diagnosed recurrent somatic mutations in BMPR1A, a protein that has been related to this point with juvenile polyposis syndrome, a most cancers predisposition syndrome.

Sequence-precise enrichment and sequencing strategy

We sequenced tumors and paired ordinary colon tissues from sufferers with excessive-grade adenocarcinoma of the colon (G3), affected person 1 with a microsatellite instable and affected person 2 with a microsatellite strong tumor (Table 1, Figure S1). For the willpower of germline mutations, we sequenced similarly to the tumor tissues from every affected person adjoining now no longer affected ordinary colonic tissue. Using Illumina sequencing and SNP arrays we decided that the tumor of affected person 1 is replica range strong while affected person 2 confirmed versions which we used for the re-assessment of diagnosed excessive stringency mutations (Table S3).

Table 1: Colon Cancer patient selected for NGS.

	Patient 1	Patient 2
Age	59	65
Gender	Male	Male
Grade	G-3	G-3
Localization	Proximal CRC	Proximal CRC
MS status	MSI	MSS
CNV	No	Yes

We analyzed the whole exomes of extra than 135,000 exons with single-examine shotgun 454 sequencings (Figure 1, Figure S1, Figure S2, Table 2). To examine the impact of insurance intensity at the sensitivity and specificity of series version detection, genotype calls of the Affymetrix SNP array 6.0 had been as compared step-sensible to the referred to as nucleic acid positions and ended in the accuracy of extra than 99% (Figure S3). In addition to the SNP array, we used Sanger sequencing to affirm 23 decided-on mutations (Table S5, Figure S5).

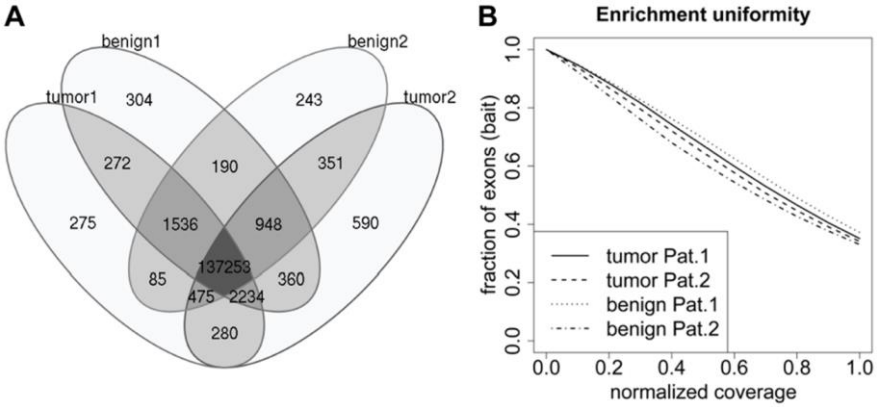

Figure 1: (A) Venn diagram of captured exons of ordinary and tumor samples. Captured exons with as a minimum one study had been counted. (B) Representative normalized insurance-distribution plot. The fraction of bait-included exons withinside the genome attaining coverages the same or decrease than the normalized insurance is indicated at the x-axis. The imply insurance consistent with exon turned into divided via way of means of the imply insurance of all exons.

Table 2: Tumor and normal genome coverages from MSI and MSS cancer patients.

	Patient 1		Patient 2	
	normal tissue	tumor tissue	normal tissue	tumor tissue
mapped reads (% of all reads)	5,659,707 (99.67%)	5,569,487 (97.21%)	2,425,905 (97.17%)	4,624,656 (96.71%)
unique mapped reads (% of all reads)	5,180,233 (91.23%)	5,285,822 (92.26%)	2,304,598 (92.38%)	4,367,855 (91.34%)
unique mapped bases (bp) (% of all bases)	1,978,702,340 (92.18%)	2,045,499,143 (88.22%)	883,388,420 (94.62%)	1,916,322,803 (94.53%)
median read length (bp)	393	418	418	483
unique reads in target region (% of all reads)	4,501,660 (79.28%)	4,477,985 (78.16%)	1,919,239 (67.88%)	3,640,778 (76.14%)
Target Base Coverage (%)	95.58	94.82	93.79	94.96
regions hit (of 176,159)	150,763	149,121	142,982	143,424

doi:10.1371/journal.pone.0015661.t002

Identification of somatic mutations in coding sequences for an MSI CRC

Searching for versions in coding areas we determined 12,767 and 12,518 small nuclear versions in 6,428 and 6,205 genes, for tumor and ordinary respectively. Of those versions, 1,428 for manipulating and 2,404 for tumor have common heterozygosity or minor allele frequency decrease than 1% or have now no longer been formerly stated in dbSNP or the one thousand Genomes Project (Figure 2). Since indels (small insertions and deletions) at homopolymeric websites are a primary supply of sequencing mistakes of the 454 platform we unnoticed this sort of alteration in our analyses. Our somatic version detection approach becomes designed to limit false-nice somatic version calls in preference to deciding zygosity.

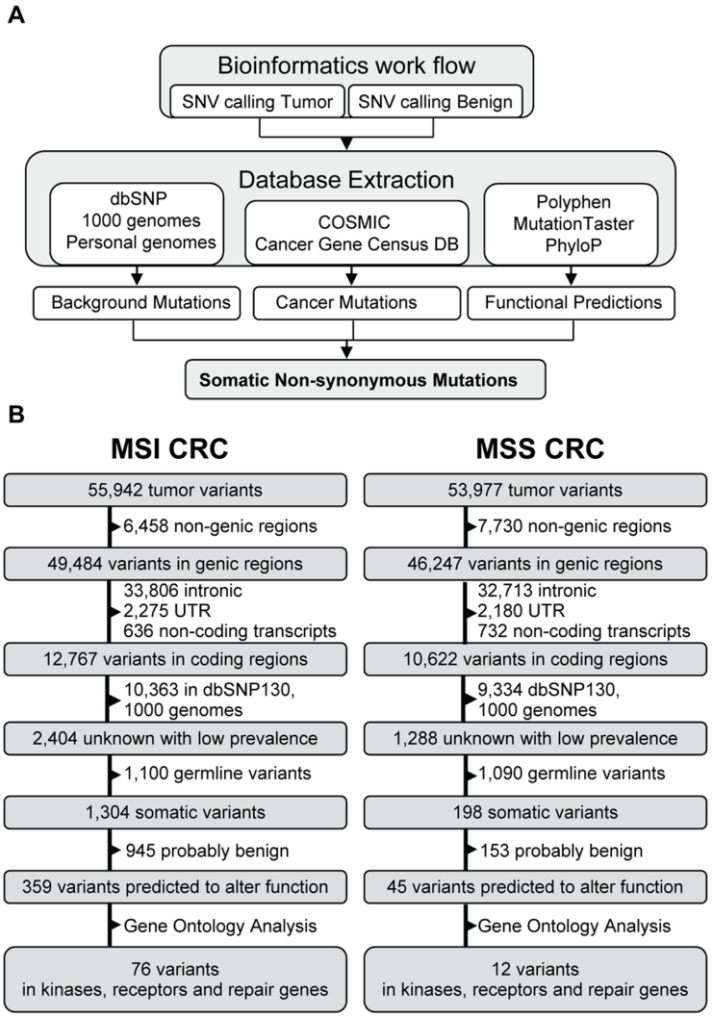

Figure 2: The identity technique of somatic applicable SNVs. (A) Schematic of the bioinformatics SNV detection workflow. (B) Extraction of functionally applicable somatic mutations for MSI and MSS colorectal cancers. Variants had been detected with the GS Reference Mapper earlier than they had been filtered for his or her localization, annotation in dbSNP130 or the 1000genomes, somatic and functionally impairment. From

dbSNP130 or the 1000genomes versions with frequencies above 1% had been used. For MSI CRC 359 versions and MSS CRC 45 with anticipated altered protein capabilities had been diagnosed.

In the primary step, tumor versions had been known as below stringent criteria, which had been decided with the aid of using assessment to the SNP genotyping array information. The 2nd step ascertained whether or not the tumor version become germline or somatic. To maintain the false-bad price within the benign tissue at much less than 10%, we set the insurance cut-off at 5-fold, under which no conclusions had been drawn concerning whether or not a version become somatic or germline. If the insurance cut-off becomes met, an unmarried study displaying the identical version in benign and tumor tissue resulted within the categorization of the version as germline. Using this approach we diagnosed 915 somatic non-synonymous mutations affecting 864 genes. The majority of somatic mutations had been missense mutations (65%). However, many (7%) are placed inside untranslated areas of genes and can consequently bring about altered expression or extended decay of mRNA species. Furthermore, about 0.5% of those mutations are determined at splice web websites and will affect splicing events, main to an altered transcriptome structure.

In addition, 3 somatic versions had been diagnosed in miRNA areas (Table S6). These mutations are of specific hobby due to the fact miRNAs had been implicated as grasp regulators of tumor homeostasis. Analysis of the unique styles of nucleic acid versions, which include regarded and unknown versions, confirmed the anticipated fees of nucleotide exchanges, as decided with the aid of using calculations the use of dbSNP130 (Figure S4).

Functional evaluation of mutations for an MSI colon most cancers

Since now no longer all the 1,304 somatic mutations are probable to be pathologically relevant, we sought to become aware of people who in all likelihood wreck protein feature or have an effect on extraordinarily conserved amino acids and can consequently be functionally important. We used Polyphen and MutationTaster class gear to are expecting the purposeful results of amino acid modifications or frameshift mutations and determined that 359 genes had as a minimum one probably unfavorable mutation (Table S1).[22],[23] Of the probably unfavorable somatic mutations, 309 had been placed in positions extraordinarily conserved in 44 unique species, which include opossum (Monodelphis domestica), chicken (Gallus gallus), and lamprey (Petromyzon marinus). Of those, 259 had been placed in genes expressed within the colon, of which 47 had been repairing, receptor, or kinase genes.

Visualization of decided-on mutations on protein systems suggests that those nucleotides are at the protein floor, probably ensuing in disrupted protein-protein interactions. The Catalog of Somatic Mutations in Cancer (COSMIC) database is a complete series of most cancer-associated mutations. Approximately 39% of our mutated genes had been already defined on this database, and 13% had been determined with the aid of using Wood et al.[10] As did those preceding databases, we determined the BRAF p.V600E mutation within the MSI case and we diagnosed KRAS and TP53 mutations within the MSS tumor. BRAF mutations are determined in about 10% of CRCs, predominantly MSI and 30 to 35% of all sufferers with sporadic colorectal cancers bring somatic KRAS and TP53 mutations. These findings in addition exhibit the sensitivity of our class approach.

Mutational landscape of an MSS colon most cancers

For the MSS colon most cancers, we diagnosed 10,622 small nuclear versions. After the identical filtering procedures as for the MSI most cancers the use of the dbSNP database and the information from the one thousand Genomes Project, 1,288 versions remained which both had low occurrence or had been unknown. Of those, 198 had been somatic and 45 had been anticipated to regulate gene feature primarily based totally on MutationTaster and Polyphen calculations (Figure 2B, Table S2).[22],[23] Regarding replicate a wide variety of versions 5 of the 46 diagnosed mutations are placed inside amplified areas, and, as anticipated, none in areas with deletions. The ratios of reads in regards series to mutated series aren't exceeding ratios in reproduction wide variety solid regions which helps the SNV-calling algorithm.

In the evaluation of 1,304 somatic mutations within the MSI tumor, we determined 198 somatic mutations within the MSS tumor which demonstrates that the faulty MMR device in MSI tumors outcomes in a large growth in mutation fees in colorectal most cancers. Furthermore, searching at intersections among each most cancer sorts we determined BMPR1A, WDTC1 (WD and tetratricopeptide repeat 1), and EHD3 (EH-area containing 3) mutated in each tumor. The choice becomes primarily based totally on purposeful impairment with the excessive possibility in Polyphen and MutationTaster.[22],[23] All mutations are placed at the floor of the protein and are extraordinarily conserved. Since we determined large most cancers-associated pathways related handiest with BMPR1A however now no longer with WDTC1 or EHD3, and in addition, germline mutations in BMPR1A are a regarded chance aspect for juvenile polyposis syndrome, we selected BMPR1A for added purposeful assays (Figure 3, Figure S5). Using reporter assays with wild-kind and mutated BMPR1A proteins we had been camping a position to reveal that the mutated proteins are strongly impaired of their signaling feature and that stimulation with BMP2 outcomes in a discounted most activity (Figure 3D).

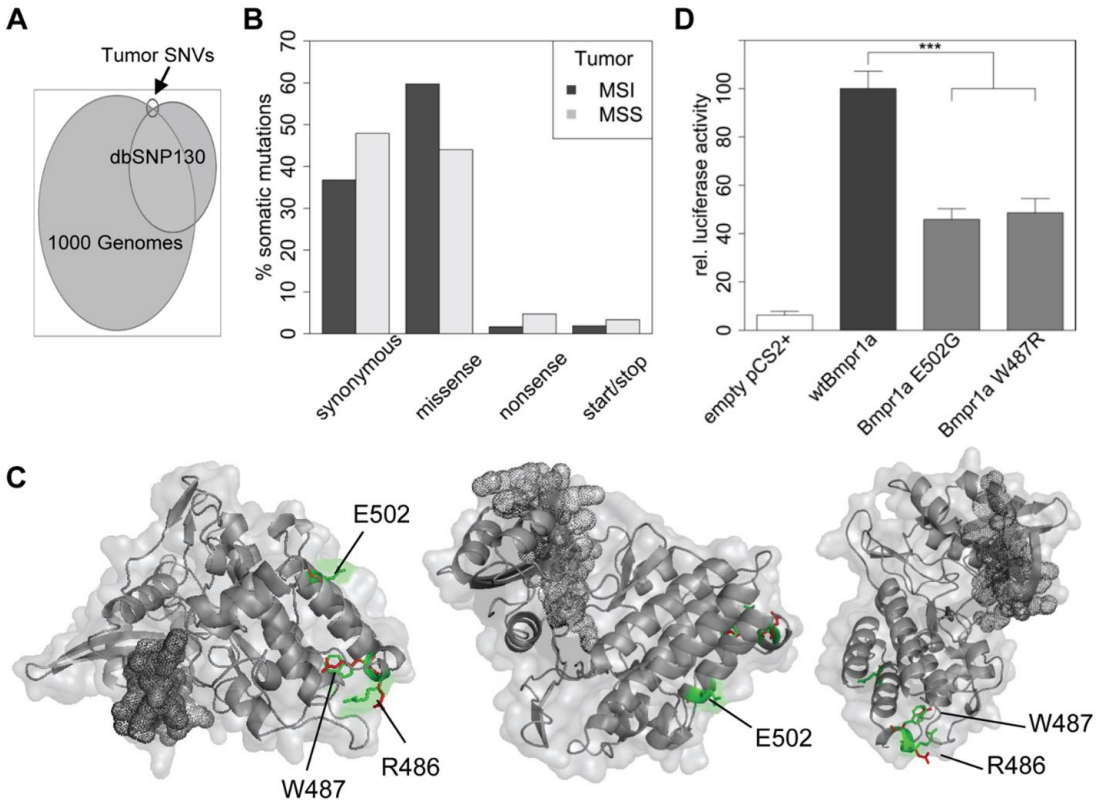

Figure 3: Characterization of number one diagnosed SNVs. (A) Proportional Venn diagram. Fractions are known as SNVs same as the Genomes Project information and dbSNP130. Only information for which the minor allele frequency or the common heterozygosity turned into acknowledged and beneath 1% had been used for comparison. (B) Distribution of synonymous, missense, nonsense, and mutations affecting the begin or prevent codon are proven on the subject of all somatic mutations. (C) BMPR1A mutations p.W487R and p.E502G are positioned on the protein kinase area of BMPR1A.

Reference amino acids are in green, the mutated paperwork is proven in red. The internet shape on the left decrease facet shows the ATP binding area. (D) BMPR1A mutations display reduced signaling activity. The activity of wt mBMPR1A, mBMPR1A E502G, and mBMPR1A W487R turned into decided in C2C12 cells the usage of a SMAD-responsive Luciferase reporter gene assay. Induced Luciferase hobby turned into normalized to Renilla activity. The hobby of untransfected cells turned into a set to 0% and the hobby of wt mBmpr1a turned into a set to 100%. Significant variations had been calculated with a two-tailed t-check and marked as: * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

MSI colorectal cancers harbor as much as 8-fold extra coding somatic mutations than MSS cancers

The analyses supplied to date had been primarily based totally on 454 complete-exome sequencings of 1 colorectal most cancers affected person for every microsatellite reputation. To similarly verify that the improved quantity of coding mutations in MSI cancers may be generalized and isn't always because of the era used ('array' enrichment and 454 sequencing) we sequenced the exomes of 4 extra colorectal cancers, every with matching ordinary tissues. This time we used 'in answer hybridization' for taking pictures of DNA accompanied through SOLiD sequencing. After the equal filtering strategies as defined for the primary sufferers, we once more decided as much as 8-fold better mutation quotes for the MSI colorectal most cancers than for MSS cancers (Table 3). In this regard, we observed 532 non-synonymous somatic SNVs within the extra MSI CRC and the handiest 65, 74, and seventy-six within the 3 MSS CRC instances.

Thus, the variations in mutation quotes are reproducible and unbiased of the sequencing era used. Discussion Using next-era sequencing, we sequenced the exomes of MSS and MSI colon most cancers sufferers, with imply coverages of about 20-fold. We carried out a -sided type set of rules to discover functionally applicable mutations. Using this approach, we reveal for the primary time that an array-seize NGS one-step workflow is an effective device for deep characterization of stable tumors and display that MSI tumors convey 8 instances extra useful applicable mutations than MSS tumors (Figure 2).

Table 3: Distribution of SNVs in MSI and MSS cancers.

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6
MS status	MSI	MSS	MSS	MSS	MSS	MSI
enrichment technology	array	array	in solution	In solution	in solution	in solution
NGS technology	454	454	SOLiD	SOLiD	SOLiD	SOLiD
number of mutations*	897	124	65	74	76	532

*non-synonymous somatic mutations, not annotated in dbSNP.
doi:10.1371/journal.pone.0015661.t003

The useful effect of the somatic versions became anticipated the use of useful prediction algorithms, Polyphen and Mutation Taster, and we observed 359 somatic mutations for the MSI and 45 for the MSS most cancers this is extraordinarily probable to purpose useful impairment.[22],[23] The heterogeneity of mutated genes shows that now no longer a particular gene consistent with se however the affected pathway performs a prime function in tumor development. In this regard, we discover a great enrichment of mutations in most cancer-associated pathways together with most cancers, mobile development, DNA replication, recombination, and repair (Table S5).

Interestingly, we discover 50% of the most importantly enriched pathways within the MSS most cancers additionally as notably enriched pathways for the MSI most cancers, indicating that although MSI cancers harbor an improved quantity of mutations each cancer may broaden thru overlapping pathomechanisms. Historically, microsatellite trying out in colorectal cancers became the primary predictive check for the identity of an underlying mismatch repair (MMR) mutation. Since extra than 90% of hereditary non-polyposis colorectal cancers (HNPCC) display MSI, the microsatellite reputation has ended up a not unusual place diagnostic marker of MMR. In addition, survival benefits and healing outcomes had been pronounced for sufferers with MSI tumors[5],[24] Using ultradeep sequencing of 1 conserved location, UCR41, de Grassi and co-workers display that this location has better mutation quotes in HNPCC samples than in healthful controls and advise that this is probably used as a touchy molecular assay of genomic instability.[24] We have prolonged their analyses to complete exomes and observed 8-fold variations within the numbers of somatic mutations of MSI and MSS colorectal cancers.

In comparison, a take a look at Greenman et al. which pronounced at the sequencing of 518 protein kinase genes in 210 various human cancers observed an about 25-fold better mutation fee for MMR-poor cancers.[9] However, those are extrapolations and in evaluation to our take a look at, they blanketed tumors from special origins and tested all somatic mutations no matter their useful relevance. With our sequencing approach, we additionally detected a somatic MLH3 mutation within the MSI tumor, which, even though MLH3 mutations do now no longer belong to the classical MMR mutations in CRC, may make contributions to the microsatellite instability phenotype.[25] Furthermore, the aggregate of MSI and BRAF mutation, as detected for the MSI tumor defined, is maximum regularly observed for CpG island methylation phenotype 1 (CIMP1) tumors that are related to MLH1 promoter methylations.[21],[26] The promoter methylation in flip is related to gene silencing mechanisms that are suggestive as a reason for the MSI reputation of the tumors.

On the opposite side, it's been proposed that chromosomal instability (CIN) and CIMP constitute unbiased and inversely associated mechanisms of instability.[27] CIMP-terrible instances are related to p53 (71%) and KRAS (33%) mutations however are not often observed with BRAF (2%) mutations. Since we observed huge replica quantity versions, in addition to KRAS and TP53 mutations within the MSS tumor, analyzed this tumor is maximum probable CIMP-terrible.[5],[24] The sequencing and evaluation method we've got supplied is probably the idea for destiny-type equipment for colorectal cancers due to the fact it can permit parallel detection of an improved mutation frequency in MSI tumors in addition to the detection of the underlying MMR defect. In addition, we had been capable of locating mutations of genes regularly related to sure subtypes of colorectal cancers together with BRAF, KRAS, and TP53.

Within our excessive precedence genes, encompassing all genes which skip all choice filters, BMPR1A stands proud as mutated in each instance. The usual shape well-known shows that the mutated amino acids are all placed on the C-terminal intracellular helix package on the protein kinase area and shows that protein-protein interactions are destroyed.[28] Germline BMPR1A mutations predispose to juvenile polyposis syndrome; however, our findings suggest that still, somatic mutations may play a critical function in sporadic colorectal most cancers development[29],[30],[31],[32],[33] Besides those mutations, we've got additionally recognized numerous mutated most cancers drug goals or genes which can be related to remedy outcomes, inclusive of BRAF, KRAS, FGFR2, and MTOR, which may assist to pick superior drug combinations. As such comparable focused re-sequencing methods and bioinformatics filtering techniques may end up a gold widespread for personally tailor-made colorectal most cancers remedy within the destiny.

RESULTS

Ethics Statement

The take a look at has been authorized through the Ethical Committee of the Medical University of Graz. For new samples, sufferers have given their written knowledgeable consent. For vintage samples (15 years vintage) no knowledgeable consent became available, consequently, all samples and clinical facts used on this take a look at were irreversibly anonymized.

Case presentation and tissue pattern collection

Patient 1 had an excessive-grade (G3) adenocarcinoma of the proximal colon, staged pT-3C, pN-zero, pM-X, pR-zero, microsatellite instable (Figure 1A). In addition, this situation became decided on due to its chromosomal stability, as decided the usage of genome-extensive next-technology sequencing (NGS) (Figure 1B). Patient 2 had an excessive-grade (G3) adenocarcinoma of the proximal colon, staged pT-4B, pN-2, pM-X, microsatellite stable. Human tissue received all through surgical operation became snap-frozen in liquid nitrogen. Cryosections (three µm thick) have been organized and stained with hematoxylin and eosin to assess tumor molecular content material.

Dissections have been executed below the microscope to acquire a tumor molecular content material of >80%. DNA isolation became executed by the usage of the QIAamp DNA Mini Kit (Qiagen), in line with the manufacturer's instructions.

DISCUSSION

Whole Exome DNA Enrichment and Genome Sequencer FLX sequencing

Genomic DNA of each tissue became subjected to entire-exome series seize the usage of Roche/NimbleGen's 2.1M Human Exome Array. This array is primarily based totally on construct 36.3 of the human genome series and captures the coding areas of 16,755 NCBI RefSeq genes (about 180,000 coding exons) in addition to 493 miRNA areas. Tumor and ordinary tissue DNA have been subjected to entire-exome series shooting in line with the manufacturer's protocol. DNA became sheared through nebulization to fragment sizes beneath neath 800bp, cleaned (Zymo Research), and end-polished the usage of T4 DNA Polymerase and T4 Polynucleotide Kinase. Linker adapters pSel3 (5′ – CTCGAG AAT TCT GGA TCC TC – three′) and pSel4-P (5′ – Phos/GAG GAT CCA GAA TTC TCG AGT T – three′) have been ligated and length choice became executed the usage of AMPure DNA Purification Beads (Agencourt). Quality became managed with the Bioanalyzer system.

LM-PCR became executed with LMPCR3 primers (5′ – CUC GAG AAU UCU GGA UCC UC – three′) earlier than the library became used for hybridization at 42°C for 72h. The arrays have been washed instances at 47.5°C, instances at room temperature, and instances at 42°C with washing buffers as recommended. Bound genomic DNA became eluted with 125 mM NaOH for 10 min at room temperature and amplified through LM-PCR the usage of primers LMPCR3. Captured amplified samples have been subjected to quantitative PCR to degree the relative enrichment. The enriched Nimblegen DNA became used to assemble unmarried-stranded Genome Sequencer FLX (454/Roche) libraries. After emulsion PCRs sequencing primers have been annealed to the template and beads have been incubated with Bst DNA polymerase, apyrase, and unmarried-stranded binding protein. Pyrosequencing became executed on a 70×75 mm picotiter plate in thirteen separate sequencing runs. After default uncooked facts processing, resequencing trimming clear out became used to boom the facts output. (Parameters used: doValleyFilterTrimBack = fake, vfBadFlowThreshold = 6, vfLastFlowToTest = 168, errorQscoreWindowTrim = zero.01).

For the sequencing, we executed thirteen Genome Sequencer FLX runs, which produced over 558 million bases and 1. 43million reads in line with a run. Reads have been aligned to the human reference genome, NCBI constructs 36 (http://hgdownload.cse.ucsc.edu/goldenPath/hg18/), the usage of GS Reference Mapper Version 2.zero.zero.12 (Roche). The first-rate fits within the genome have been used because of the place for the reads with a couple of fits. Only precise reads with a minimal duration of fifty bp have been used for similar analysis (see run data Tab. 2). Detection of variations became executed with the GS Reference Mapper Version 2.zero.zero.12 (Roche). Redundant reads have been subtracted earlier than version callings. Only the HCDiff (excessive self-belief differences) of the GS Mapper software program became used as the premise of version detection.[20] HCDiff callings presume as a minimum 3 reads with the version with each ahead of and opposite reads included; alternatively, the exceptional rankings on the variable positions ought to be over 20 (or over 30 if a homopolymer of 5 or greater bases is involved). As extra exceptional criteria, we used the most effective variations with insurance >10× of excellent reads.

Whole Exome ‘in solution’ DNA Enrichment and SOLiD sequencing

Enrichments and SOLiD library guidance have been executed in line with Agilent's SureSelect Target Enrichment protocol for the Applied Biosystems SOLiD system. In brief, entire genomic DNA became sheared and end-repaired. For adapter ligations, 30x extra of the adapters became used. Size choices for 150–two hundred bp DNA fragments have been executed observed through a nick-translation and amplification step with Platinum polymerase (Invitrogen) and Pfu-Polymerase (Fermentas). For hybrid choice, the libraries have been adjusted to 500 ng in three.four µl quantity and brought to the SureSelect Block solutions. Hybridizations have been executed for twenty-four h at 65°C, hybrids have been extracted with 500 ng M-280 streptavidin

Dynabeads (Invitrogen) and eventually eluted with 50 µl Elution buffer. After amplification with Platinum polymerase, the libraries have been quantified through qPCR and DNA awareness became titrated to acquire a fragment of 10–20% monoclonal template beads within the emulsion PCR the usage of in general 0.7 to one billion beads.

Successive bead enrichment and deposition of a hundred thirty million beads in line with area slide (quad) have been observed through popular 50 bp fragment runs. Each of the 4 affected person samples became analyzed on an unmarried quad. Mapping became executed with the Bioscope alignment pipeline the usage of the seed & expand set of rules with a mismatch penalty rating of −2.zero. Single Nucleotide Variants (SNV) have been referred to as with the DiBayes set of rules incorporated into the Bioscope package.

Single nucleotide version (SNV) detection

Tissue substances have been genotyped at the Affymetrix 6.0array, in line with the manufacturer's protocol. Array positions with an exceptional rating (p-value) <zero.1 have been used for contrast with the sequencing facts. Sequencing facts positions have been used if their insurance passed three-fold. This generated 46,000 and 49,000 positions for tumor and benign tissue, respectively, that have been eligible for contrast. To decide fake advantageous and fake terrible rates, we set the array facts as popular and prominent among reference names and SNP name dependence at the array facts.

For the detection of somatic variations, a bimodal approach became carried out with tumor variations referred to as beneath stringent standards, while variations on top of things tissue have been referred to as the use of much less stringent standards: A minimal threshold for reads became set with variations of 15% of all reads at a given role within the tumor. Less stringent standards have been used for calling manipulated tissue variations with at the very least one variation study and a minimal insurance cutoff of 5. For coverages above 30-fold, one variation study became accepted.

Capillary Sequencing

Follow-up affirmation of diagnosed SNVs became finished on an ABI 3730 (Applied Biosystems) capillary sequencing tool following popular procedures.

Determination of Copy quantity versions

Preparation of unmarried study libraries and sequencing have been finished with the use of the Solexa sequencing platform (GenomeAnalyzer IIx, Illumina) following the manufacturer's instructions. Image Analysis and base calling have been finished with the use of Firecrest 1.9.5_14 and Bustard 1.9.5_14 and reads have been aligned to the human genome (NCBI36) with the use of Bowtie zero.9.7.1.[34] Copy quantity evaluation became finished in R the use of the DNA replica package.[35] In short, DNA study frequencies have been decided for packing containers of fifty Kb. The log2 frequency ratio of corresponding packing containers became calculated for tumors as opposed to regular tissue. The median of ratios became targeted to 0 experiment-wise.

Log ratios have been smoothed through DNA replica the use of default values and duplicate quantity variant became detected through DNA replica the use of a threshold of popular deviations. Genotyping at the Affymetrix 6.0 array became finished consistent with the manufacturer's protocol. Regions of replica quantity advantage and loss have been decided through paired and evaluation the use of the Hidden Markov Model (HMM) of the Partek Genomics Suite software (Partek Inc, St.Louis, MO) with default parameter settings. For paired evaluation, replica quantity values have been generated through evaluating tumor and benign tissue profiles from identical patients.

Bioinformatics workflow

For every tissue, versions have been annotated the use of the gene fashions generated through Ensembl (Ensembl 54.36, www.ensembl.org). All versions have been mapped to all transcript fashions, which caused more than one annotation for numerous loci. For instance, a variation can cause an amino acid to alternate in a single transcript and seem within the UTR of another. Comparison of tumor and benign tissue variations to dbSNP130 and a thousand Genomes Project records became performed with the subset of dbSNP130 and a thousand Genomes Project positions with minor allele frequencies or common heterozygosity >zero.01. Variants have been subjected to many comparisons with outside records assets. Most records reassert are included in the US genome browser (http://hgdownload.cse.ucsc.edu/goldenPath/hg18/database/) or have been derived from websites just like a thousand Genomes Project (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/release/2009_04/) the gene ontology records (http://archive.geneontology.org/full/2009-10-01/go_200910-termdb.obo-xml.gz), the cosmic database model 46 (http://www.sanger.ac.uk/genetics/CGP/cosmic/) or the most cancers gene census database (http://www.sanger.ac.uk/genetics/CGP/census/).

Functional classifications have been finished with the use of Polyphen and MutationTaster class tools (http://genetics.bwh.harvard.edu/pph/, http://neurocore.charite.de/MutationTaster/).[22],[23] Base conservation amongst 44 species became examined the use of the phyloP music of the US browser (http://genome.cshlp.org/content/early/2009/10/26/gr.097857.109.abstract). Bases have been taken into consideration pretty conserved if their conservation rating became extra or identical to 2.0 (zero.975 amount of all conservation scores). For gene expression, wholesome manipulated samples from

http://www.ncbi.nlm.nih.gov/sites/GDSbrowser?acc=GDS2609 have been used. We calculated genetic imply expression values throughout all samples and used the primary quartile as a threshold to decide gene expression. Pathway analyses have been finished with the ingenuity pathway evaluation tool (http://www.ingenuity.com). All new records from this take a look at had been deposited on the NCBI dbSNP database of genetic variants (user-call MPIMGCancerogenomics). Accession numbers are blanketed as Table S4.

Protein shape modeling

Models for BMPR1A have been received from SwissModel and ModBase. Very comparable fashions besides for a few loops (overall Calpha rmsd 0.66) have been rendered in PyMol.[36],[37]

Reporter assays

The hobby of the wild-kind mouse protein (wtmBMPR1A) and its mutants became decided through measuring precipitated Luciferase hobby within the transiently transfected pre-myoblastic mouse molecular line C2C12 (ATCC). WtmBMPR1A became amplified from mouse cDNA and cloned within the expression vector pCS2+. Both mutations W487R and E502G have been inserted through Quikchange mutagenesis (Stratagene) the use of the subsequent primer pairs:

mBmpr1a_W487R_fwd
caatcgtgtctaaccgcCggaacagcgatgaatg;
mBmpr1a_W487R_rev
cattcatcgctgttccGgcggttagacacgattg and
mBmpr1a_E502G_fwd
gttttgaagctaatgtcagGatgttgggcccataatc;
mBmpr1a_E502G_rev
gattatgggcccaacatCctgacattagcttcaaaac.

C2C12 cells have been cultured in DMEM glucose 4,5 g/L with 10﷼S have been co-transfected with every Bmpr1a expression construct, a Smad Binding Element (SBE) luciferase construct[38] and the normalization vector pRL-Tk (Promega Corporation, Madison, WI, USA) the use of Turbofect (Fermentas GmbH, St. Leon-Rot, Germany). Luciferase hobby changed into decided as defined previously.[39] Supporting Information

REFERENCES

1. Wood LD, Parsons DW, Jones S, Lin J, Sjoblom T, et al. The genomic landscapes of human breast and colorectal cancers. Science, 2007; 318: 1108–1113.
2. The Cancer Genome Atlas Research Network. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature, 2008; 455: 1061–1068.
3. Stiller M, Knapp M, Stenzel U, Hofreiter M, Meyer M. Direct multiplex sequencing (DMPS)–a novel method for targeted high-throughput sequencing of ancient and highly degraded DNA. Genome Res., 2009; 19: 1843–1848.
4. Schweiger MR, Kerick M, Timmermann B, Albrecht MW, Borodina T, et al. Genome-wide massively parallel sequencing of formaldehyde fixed-paraffin embedded (FFPE) tumor tissues for copy-number- and mutation-analysis. PLoS ONE, 2009; 4: e5548.
5. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature, 2009; 461: 272–276.
6. Sudheer Menon "Preparation and computational analysis of Bisulphite sequencing in Germfree Mice" International Journal for Science and Advance Research In Technology, 2020; 6(9): 557-565.
7. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal "Computational identification of promoter regions in prokaryotes and Eukaryotes" EPRA International Journal of Agriculture and Rural Economic Research (ARER), July 2021; 9(7): 21-28.
8. Sudheer Menon "Bioinformatics approaches to understand gene looping in human genome" EPRA International Journal of Research & Development (IJRD), July 2021; 6(7): 170-173.
9. Sudheer Menon "Insilico analysis of terpenoids in Saccharomyces Cerevisiae"international Journal of Engineering Applied Sciences and Technology, 2021 Vol. 6, Issue1, ISSN No. 2455-2143, 2021; 43-52.
10. Sudheer Menon "Computational analysis of Histone modification and TFBs that mediates gene looping" Bioinformatics, Pharmaceutical, and Chemical Sciences (RJLBPCS), June 2021; 7(3): 53-70.
11. Sudheer Menon Shanmughavel piramanayakam, Gopal Prasad Agarwal "FPMD-Fungal promoter motif database: A database for the Promoter motifs regions in fungal genomes" EPRA International Journal of Multidisciplinary research, 2021; 7(7): 620-623.
12. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal, 2021.
13. Krawitz PM, Schweiger MR, Rodelsperger C, Marcelis C, Kolsch U, et al. Identity-by-descent filtering of exome sequence data identifies PIGV mutations in hyperphosphatasia mental retardation syndrome. Nat Genet, 2010; 42: 827–829.
14. Ley TJ, Mardis ER, Ding L, Fulton B, McLellan MD, et al. DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature, 2008; 456: 66–72.
15. Shah SP, Morin RD, Khattra J, Prentice L, Pugh T, et al. Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. Nature, 2009; 461: 809–813.
16. Pleasance ED, Cheetham RK, Stephens PJ, McBride DJ, Humphray SJ, et al. A comprehensive catalogue of somatic mutations from a human cancer genome. Nature, 2010; 463: 191–196.
17. Pleasance ED, Stephens PJ, O'Meara S, McBride DJ, Meynert A, et al. A small-cell lung cancer

genome with complex signatures of tobacco exposure. Nature, 2010; 463: 184–190.
18. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science, 1993; 260: 816–819.
19. Hedges DJ, Burges D, Powell E, Almonte C, Huang J, et al. Exome sequencing of a multigenerational human pedigree. PLoS ONE, 2009; 4: e8232.
20. Shen L, Toyota M, Kondo Y, Lin E, Zhang L, et al. Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci U S A., 2007; 104: 18654–18659.
21. Computational Identification of promoter regions in fungal genomes, International Journal of Advance Research, Ideas and Innovations in Technology, 7(4): 908-914.
22. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam, Bioinformatics methods for identifying hirschsprung disease genes, International Journal for Research in Applied Science & Engineering Technology (IJRASET), July, 2021; 9(VII): 2974-2978.
23. Sudheer Menon, Bioinformatics approaches to understand the role of African genetic diversity in disease, International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), 2021; 4(8): 1707-1713.
24. Sudheer Menon Comparison of High-Throughput Next generation sequencing data processing pipelines, International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), 2021; 3(8): 125-136.
25. Sudheer Menon Evolutionary analysis of SARS-CoV-2 genome and protein insights the origin of the virus, Wuhan, International Journal of Creative Research Thoughts (IJCRT), 2021; 9(8): b696-b704.
26. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam A step-by-step work flow of Single Cell RNA sequencing data analysis, International Journal for Scientific Research and Development (IJSRD), 2021; 9(6): 1-13.
27. Sudheer Menon Computational characterization of Transcription End sites in Human Genome, International Journal of All Research Education and Scientific Methods (IJRESM), 2021; 9(8): 1043-1048.
28. Sudheer Sivasankaran Menon and Shanmughavel Piramanayakam Insilico prediction of gyr A and gyr B in Escherichia coli insights the DNA-Protein interaction in prokaryotes, International Journal of Multidisciplinary Research and Growth Evaluation, (IJMRD), 2021; 2(4): 709-714.
29. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. A method and server for predicting damaging missense mutations. Nat Methods, 2010; 7: 248–249.
30. Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods, 2010; 7: 575–576.
31. De Grassi A, Segala C, Iannelli F, Volorio S, Bertario L, et al. Ultradeep sequencing of a human ultraconserved region reveals somatic and constitutional genomic instability. PLoS Biol., 8: e1000275.
32. Lipkin SM, Wang V, Stoler DL, Anderson GR, Kirsch I, et al. Germline and somatic mutation analyses in the DNA mismatch repair gene MLH3: Evidence for somatic mutation in colorectal cancers. Hum Mutat, 2001; 17: 389–396.
33. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, et al. CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci U S A., 1999; 96: 8681–8686.
34. Goel A, Nagasaka T, Arnold CN, Inoue T, Hamilton C, et al. The CpG island methylator phenotype and chromosomal instability are inversely correlated in sporadic colorectal cancer. Gastroenterology, 2007; 132: 127–138
35. Howe JR, Sayed MG, Ahmed AF, Ringold J, Larsen-Haidle J, et al. The prevalence of MADH4 and BMPR1A mutations in juvenile polyposis and absence of BMPR2, BMPR1B, and ACVR1 mutations. J Med Genet., 2004; 41: 484–491.
36. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam Bioinformatics tools and methods to analyze single cell RNA sequencing data, International Journal of Innovative Science and Research Technology, (IJISRT), 2021; 6(8): 282-288.
37. Sudheer Menon Computational genome analysis for identifying Biliary Atresia genes, International Journal of Biotechnology and Microbiology, (IJBM), 2021; 3(2): 29-33.
38. Sudheer Menon Recent Insilco advancements in genome analysis and characteristics of SARS-Cov2. International Journal of Biology Research, (IJBR), 2021; 6(3): 50-54.
39. Sudheer Menon Bioinformatics methods for identifying Human disease genes, International Journal of Biology Sciences, (IJBR), 2021; 3(2): 1-5.
40. Sudheer Menon SARS-CoV-2 Genome structure and protein interaction map, insights to drug discovery, International Journal of Recent Scientific Research, (IJRSR), 2021; 12(8): 42659-42665.
41. Sudheer Menon Insilico Insights to Mutational and Evolutionary aspects of SARS-Cov2, International Journal of Multidisciplinary Research and Development, (IJMRD), 2021 8(8): 167-172.
42. Sudheer Menon Computational biology, machine learning and reverse vaccinology detects the role of conserved Nsp3 Protein and its importance in Covid-19 vaccine development, European Journal of Biotechnology and Bioscience, 2021 9(3): 95-99.
43. Sudheer Menon Protein Evolutionary Analysis of SARS-CoV-2 Delta Plus and C.1.2 Insights Virulence and Host Immunity, International Journal of Multidisciplinary

Research and Publications (IJMRAP), 2021; 4(2): 6-13.
44. Markowitz S, Wang J, Myeroff L, Parsons R, Sun L, et al. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. Science, 1995; 268: 1336–1338.
45. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M., Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature, 1993; 363: 558–561
46. Howe JR, Bair JL, Sayed MG, Anderson ME, Mitros FA, et al. Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis. Nat Genet, 2001; 28: 184–187.
47. Howe JR, Roth S, Ringold JC, Summers RW, Jarvinen HJ, et al. Mutations in the SMAD4/DPC4 gene in juvenile polyposis. Science, 1998; 280: 1086–1088.
48. Hahn SA, Schutte M, Hoque AT, Moskaluk CA, da Costa LT, et al. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science, 1996; 271: 350–353.
49. Sudheer Menon, Computing for Genomics and Proteomics in Microbial biotechnology in the identification of SARS COV2 Variants, International Journal of Scientific Engineering and Applied Science (IJSEAS), 2021; 7(9): 147-175.
50. Sudheer Menon, Functional Genome analysis and drug screening by Bioinformatics methods in SARS-CoV-2. International Journal of Academic Research and Development (IJARD), 2021; 6(5): 33-39.
51. Sudheer Menon, Advances in Systems biology and immunological data analysis for drug and vaccine development, International Journal of Educational Research and Studies (IJERS), 2021; 3(3): 53-64.
52. Sudheer Menon, Whole genome sequencing and Insilico approaches to understand the genetic basis of Rare Diseases, International Journal of Advanced Scientific Research (IJASR), 2021; 6(5): 7-13.
53. Sudheer Menon Bioinformatics approaches to identify neurodegenerative diseases by Next generation sequencing data, International Journal of Educational Research and Development (IJERD), 2021; 3(3): 70-74.
54. Sudheer Menon Bioinformatics approaches for cross-species liver cancer analysis based single cell RNA sequencing data, International Journal of Molecular Biology and Biochemistry (IJMBB), 2021; 3(1): 05-13.
55. Sudheer Menon Genomic sequence of worldwide strains of SARS-CoV-2: insights the Role of Variants in Disease Epidemiology, International Journal of Advanced Research and Development (IJARD), 2021; 6(4): 14-21.
56. Sudheer Menon Differential Gene expression analysis by Single Cell RNA sequencing reveals insights to genetic diseases, International Journal of Multidisciplinary Education and Research (IJMER), 2021; 6(4): 1-7.
57. Hai-Bing, Y.; Sivasankaran, M.S.; Ottakandathil, B.R.; Zhong-Luan, W.; Man-Ting, S.; Ho-Yu, C.; Kak-Yuen,W.; Kwong-Hang, T.; Chi-Hang, L. Environmental Toxin Biliatresone-Induced Biliary Atresia-like Abnormal Cilia and Bile Duct Cell Development of Human Liver Organoids. Toxins, **2024;** 16: 144.
58. Sudheer Menon, "Neural network prediction of genome wide transcriptome signatures in Adenocarcinoma", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, December 2024; 12(12): b716-b726.
59. Sudheer Menon, "Computational Prediction of SARS-CoV-2 Genomic, Proteomic Mutation, and Variants by (NGS) Next-Generation Sequencing Data", International Journal of Science and Research (IJSR), December 2024; 13(12): 565-573.
60. Bodmer WF Cancer genetics: colorectal cancer as a model. J Hum Genet, 2006; 51: 391–396.
61. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol., 2009; 10: R25.
62. Venkatraman ES, Olshen AB. A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics, 2007; 23: 657–663.
63. Kiefer F, Arnold K, Kunzli M, Bordoli L, Schwede T. The SWISS-MODEL Repository and associated resources. Nucleic Acids Res., 2009; 37: D387–392.
64. Pieper U, Eswar N, Braberg H, Madhusudhan MS, Davis FP, et al. MODBASE, a database of annotated comparative protein structure models, and associated resources. Nucleic Acids Res., 2004; 32: D217–222.
65. Sudheer Menon Protein–protein interactions by exploiting evolutionary information insights the genes and conserved regions in the corresponding human and mouse genome, International Journal of Advanced Multidisciplinary Research (IJARM), 2021; 8(9): 36-55.
66. Sudheer Menon and Binu Thomas. An insights to Bioinformatics methods in Natural Product Drug Discovery, International Journal of Clinical Biology and Biochemistry (IJCBB), 2021; 3(1): 26-37.
67. Sudheer Menon and Binu Thomas. Use of multi omics data in precision medicine and cancer research with applications in tumor subtyping, prognosis, and diagnosis, International Journal of Advanced Education and Research (IJAER), 2021; 6(5): 19-29.
68. Sudheer Menon and Binu Thomas. Gene Expression Analysis, Functional Enrichment, and Network Inference in disease prediction, International Journal of Advanced Educational Research (IJAER), 2021; 6(5): 6-18.
69. Sudheer Menon and Binu Thomas. Vaccine informatics, systems biology studies on neuro-degenerative diseases, International Journal of Multidisplinary Educational Research IJMER, 2021; 10(9-4): 85-101.

70. Sudheer Menon. Computational approaches to understand the role of epigenetics and microRNA in Cancer among different populations, International Journal of Multidisciplinary Research Review (IJMDRR), 2021; 7(9): 69-81.
71. Sudheer Menon and Binu Thomas. Recent advances, methods and opportunities of Deep learning in computational and systems biology International Journal of Applied and Advanced Scientific Research (IJAASR), 2021; 6(2): 12-25.
72. Sudheer Menon Computational genome based studies of Complex Genetic Diseases, International Journal of Multidisciplinary Research and Modern Education (IJMRME), 2021; 7(2): 28-39.
73. Jonk LJ, Itoh S, Heldin CH, ten Dijke P, Kruijer W. Identification and functional characterization of a Smad binding element (SBE) in the JunB promoter that acts as a transforming growth factor-beta, activin, and bone morphogenetic protein-inducible enhancer. J Biol Chem., 1998; 273: 21145–21152.
74. Hampf M, Gossen M. A protocol for combined Photinus and Renilla luciferase quantification compatible with protein assays. Anal Biochem, 2006; 356: 94–99.
75. Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, et al. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. Clin Genet, 2009; 76: 1–18.
76. Fallik D, Borrini F, Boige V, Viguier J, Jacob S, et al. Microsatellite instability is a predictive factor of the tumor response to irinotecan in patients with advanced colorectal cancer. Cancer Res., 2003; 63: 5738–5744.
77. Sjoblom T, Jones S, Wood LD, Parsons DW, Lin J, et al. The consensus coding sequences of human breast and colorectal cancers. Science, 2006; 314: 268–274.
78. Sudheer Menon. Transcriptomic and metabolic pathway analysis reveals the mutated genes p53 and TP53 and their role in different types of cancer, International Journal of Engineering Research and Modern Education (IJERME), 2021; 6(2): 10-33.
79. Sudheer Menon, Amanda Ramer Tait, Shanmughavel Piramanagakam, Vincent Chi Hang Lui and Paul Kwong Hang Tam. Bisulphite seq data reveals host-microbiota interaction in germfree and conventional mice, International Journal of Bioinformatics and Biological Sciences, 2021; 9(1): 1-8.
80. Sudheer Menon Insilico identification of Transcription End Sites in Eukaryotes, International Journal of Pharmacy and Biological Sciences (IJPBS), 2021; 11(3): 104-115.
81. Sudheer Menon Computational identification of upregulated genes in Breast, Prostate, and Colorectal Cancer and the Cellular Signaling process, Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences, (RJLBPCS), 2021; 7(5): 68-83.
82. Sudheer Menon Structural and Functional Characteristics of miRNAs in colon cancer and the identification of targets by Insilco methods, Research Journal of Life Sciences, Bioinformatics, Pharmaceutical and Chemical Sciences, (RJLBPCS), 2021; 7(6): 86-96.
83. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021, Statistical approach for Predicting Genome Architecture and the major challenges, World Journal of Advanced Research and Reviews (WJARR), 2022; 13(1): 433-445.
84. Binu Thomas, Mahesh Mohanan, and Sudheer Menon, Role of edible Plants to balance dietary requirements in human diet, Devagiri Journal of Science, 2021; 7(1): 27-40.
85. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science, 2008; 321: 1801–1806.
86. Greenman C, Stephens P, Smith R, Dalgliesh GL, Hunter C, et al. Patterns of somatic mutation in human cancer genomes. Nature, 2007; 446: 153–158.
87. Aaltonen LA, Peltomaki P, Leach FS, Sistonen P, Pylkkanen L, et al. Clues to the pathogenesis of familial colorectal cancer. Science, 1993; 260: 812–816.
88. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal, Bioinformatics Analysis of Actin molecules and the actin gene by next-generation sequencing data insights cell mobility, International Journal of Advanced Trends in Engineering and Technology, IJATET, 2022; 7(1): 1-20.
89. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal, CDKN2A gene expression/mutation prognosis by computational methods in the identification of nodular melanoma, International Journal of Scientific Research and Modern Education (IJSRME), 2022; 7(1): 11-22.
90. Sudheer Menon, Vincent Chi Hang Lui1, Paul Kwong Hang Tam, comparison of computational tools for differential gene expression analysis of rna sequencing and single-cell rna sequencing data, European Journal of Biomedical AND Pharmaceutical sciences, 2022; 9(4): 1-9.
91. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal, computational characterization of genomewide dna methylation and hydroxymethylation in the mouse hypothalamus by next-generation sequencing data, International Journal of Applied and Advanced Scientific Research (IJAASR), 2022; 7(1): 9-29.
92. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal, Differential expression analysis of stressinducible candidate genes in response to cold and draught in Coffea Arabica, International Journal of Advanced Trends in Engineering and Technology (IJATET), 2022; 7(1): 21-38.

93. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal, Identification of immune gene–related IncRNA signature to detect brain tumors, International Journal of Current Research and Modern Education (IJCRME), 2022; 7(1): 6-20.
94. Sudheer Menon, Shanmughavel piramanayakam and Gopal Prasad Agarwal, "Identification of Molecular Markers Associated With liver steatosis and Treatment Based on Integrated Transcriptome Analysis by PNPLA3 gene", Journal of Emerging Technologies and Innovative Research (JETIR), 2023; 10(12): g398-g416.
95. Sudheer Menon, Dynamical Modeling of the Core Gene Network and Mutation in Transitional Cell Carcinoma, Journal of Emerging Technologies and Innovative Research (JETIR), 2023; 10(12): h326-h338.
96. Peinado MA, Malkhosyan S, Velazquez A, Perucho M. Isolation and characterization of allelic losses and gains in colorectal tumors by arbitrarily primed polymerase chain reaction. Proc Natl Acad Sci U S A, 1992; 89: 10065–10069. Sequencing data, European Journal of Biomedical AND Pharmaceutical sciences, 9(4): 1-9.