IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Neural Network Prediction Of Genome Wide Transcriptome Signatures In Adenocarcinoma

Sudheer Menon[1], Shanmughavel Piramanayagam[2], Gopal Prasad Agarwal[3], Amanda Ramer Tait[4], Michelle Ramsay[5], Scott Hazelhurst[6], 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.

Abstract

The microarray method is used to investigate the transcriptomic formation of a tumor and the transcriptome does not tell biology because of transcription and translational post-modifications, alternative splicing, and the effects of pathological conditions. Due to this reason fusing more than one source of genome-wide data like, transcriptome, proteome, epigenome, and genome become important. A diagnostic approach includes information about the transcriptional status of all genes across multiple tissue types in cancer diagnosis. It is identified that two p53-regulated long non-coding RNAs (lncRNAs) are required for the effective binding of p53 with some of its target genes. There are 18 high-confidence lncRNAs that are transcriptional targets and they are all identified with the help of p53 ChIP-sequence analysis. It is also demonstrated that expression of p53-lncRNAs is lower in colorectal cancer samples having a tumor suppressor with very high diagnostic power.

Multi-omics sequence is becoming the clinical routine and transformation of precision oncology and it is also low-cost. Transparent and interpretable computational methods have been proposed which are called integral genomic signature analysis. With the help of a genomic dataset of chemical disruption, a battery of integral genomic signature analysis (iGenSig) models for the prediction of cancer drug response was developed. iGenSig gives a computational framework for the facilitation to adapt cancer therapy based upon multi-omics data. Understanding cancer requires the discovery of cancer driver mutations that confer a proliferative

advantage; however, due to the difficulty of modeling the somatic mutation rates seen across tumor genomes, searches have frequently been restricted to protein-coding sequences and particular noncoding elements (for example, promoters).

Keywords : Microarray, iGenSig, lncRNAs, Kernel Approach, p53 genes, LS-SVM, SVM, Tumor

Introduction

For the integration of heterogenous data and complex structured data, the Kernel technique is ideally best. Here, for supervised classification, a weighted least squares support vector machine (LS-SVM) was used (Suykens JAK et.al, 1999, Cawley GC, 2006). In the least squares support vector machine (LS-SVM), the quadratic programming problem is changed into a linear problem. So, it is an easier and faster method for dimensional data as compared to a support vector machine (SVM). LS-SVM help in clinical decision support as compared to clinical risk factors, because the information that comes out through clinical risk factors such as the size of the tumor, number of positive lymph nodes, etc is not reliable. Patients having the same pathological and clinical characteristics but with different clinical results can be distinguished with microarray methods.

To cure cancer, determination of the origin of the site which causes the tumor is very necessary and this site is used to treat cancer. Nowadays, pathology methods have been used for the diagnosis of cancer. Pathology methods use histochemistry and morphology for diagnosis and also to determine the drug eligibility for clinical trials. SCOPE (Supervised Cancer Origin Prediction Using Expression) is a set of neural networks which is used the complete transcriptome for the identification of the closest tumor match. It was determined that genes were weighted heavily for making the decision and SCOPE is used to place the genes related to each class without having previous information about genes.P53 is a transcriptional factor that is a tumor suppressor in human beings. It is vital for cellular responses such as DNA damage to maintain the integrity of cells' genome through activation of gene expression. This activation of gene expression helps to eliminate the damaged cells by apoptosis (programmed cell death). the downstream effects of p53 genes are arbitrated by its intrinsic properties as a transcription factor (Dinger M.E. et.al, 2008).

Precision oncology is used for patient care, it is a molecular method that provides the profiling of tumors (Frohlich H. et.al, 2018). Precision oncology is also used for the detection of mutations through genetic testing such as EGFR Mutation, ALK Rearrangements, etc, it is used to identify the small-sized prognostic or predictive gene signatures with the help of targeted expression assays. With the advancement in the low-cost genome sequencing method, precision oncology is at that point of deep transformation through the advantage of big data to give a wide array of clinical decision support. The computational methods which help these big data to ease the clinical decisions and to adapt health care are very demanding. For example, therapies prescribed for metastatic lung cancer have only low quality-adjusted life years (Vasconcellos, V. F. et.al,

2018).On contrary to it, the iGenSig (Integral Genomic Signature) approach uses the correlation intensities of vital genomic characteristics that were detected in specific samples.

Material and Methods

Data Collection and preparation

Plasma and tissue samples were collected before treatment, before surgery, and, after dose but before radiotherapy. All these experimental methods were done according to the laboratory procedures. By using Affymetrix human U1332.0 plus gene chip, those tissue samples that were frozen were hybridized. Then, firstly, the data was pre-processed for each point (Irizarry R.A. et.al, 2003). After pre-processing of data, the median of all the probes taken that have the same genes and the number of features were minimized from 54.613 subjects to 27,650 genes. Due to cross-hybridization chances, probes having multiple genes were eliminated. It was determined that about 96 proteins were involved in cancer with the help of plasma samples using the Luminex 100 instrument.

From the Genentech mesothelioma cohort 211 untreated mesothelioma cancers subjects were tested (Bueno R et.al, 2016). out of these 211 subjects, 126 subjects were classified as epithelioid mesotheliomas and 85 were sarcomatoid variants. From a personalized Oncogenomics study, adult metastatic disease and 15 CUPs were procured at BC Cancer (Laskin J et.al, 2015). 168 out of 201 metastates biopsy samples were collected from the site of metastasis and 33 samples were collected from the origin site.

Mesothelioma Subtype	Total Cases With Subtype, No.	Precision	Recall	F_1 Score	Predicted	Predicted Cases, No.
Biphasic epithelioidlike	72	1	1	1	Epithelioid mesothelioma	72
Epithelioid	54	1	0.98	0.99	Epithelioid mesothelioma	53
Sarcomatoid	29	NA	NA	NA	Sarcomatoid mesothelioma	18
		NA	NA	NA	Epithelioid mesothelioma	5
		NA	NA	NA	Sarcoma	4
		NA	NA	NA	Other	2
Biphasic sarcomalike	56	NA	NA	NA	Epithelioid mesothelioma	38
		NA	NA	NA	Sarcomatoid mesothelioma	17
		NA	NA	NA	Other	1

Table 1. SCOPE Performance on Genentech of primary Mesotheliomas

Diagnosed Type[a]	Total Cases, No.	Cohort Metrics[b]									Cases Predicted, No.[c]			
		TPR	FPR	TP	TN	FP	FN	Precision	Recall	F_1 Score	Diagnosis	Biopsy Site	Organ System	Other
Metastatic Site Biopsies														
Mesothelioma	1	1.00	0.00	1	130	0	0	1.00	1.00	1.00	1	0	0	0
Colorectal AC	21	0.81	0.00	17	114	0	4	1.00	0.81	0.89	17	1	2	1
UCEC	5	0.40	0.00	2	129	0	3	1.00	0.40	0.57	2	0	1	2
Uterine carcinosarcoma	4	0.25	0.00	1	130	0	3	1.00	0.25	0.40	1	0	2	1
Breast carcinoma	65	0.97	0.03	63	68	2	2	0.97	0.97	0.97	63	1	0	1
LNG_group	14	1.00	0.01	14	117	1	0	0.93	1.00	0.97	14	0	0	0
Sarcoma	17	0.53	0.01	9	122	1	8	0.90	0.53	0.67	9	1	0	7
Ovarian carcinoma	7	0.86	0.01	6	160	1	1	0.86	0.86	0.86	6	0	0	1
Pancreatic AC	9	0.33	0.01	3	158	1	6	0.75	0.33	0.46	3	1	4	1
MISC_group	9	0.88	0.00	8	125	1	1	0.73	0.88	0.77	8	0	0	1
Cholangio-carcinoma	5	0.80	0.02	4	127	2	1	0.67	0.80	0.73	4	0	1	0
GEJ_group	11	0.29	0.01	3	151	5	8	0.31	0.29	0.26	3	3	1	4
Primary Site Biopsies														
CNS_group	6	1.00	0.00	6	23	0	0	1.00	1.00	1.00	6	0	0	0
Breast carcinoma	4	1.00	0.00	4	25	0	0	1.00	1.00	1.00	4	0	0	0
Colorectal AC	1	1.00	0.00	1	28	0	0	1.00	1.00	1.00	1	0	0	0
GEJ_group	1	1.00	0.00	1	28	0	0	1.00	1.00	1.00	1	0	0	0
MISC_group	2	1.00	0.00	2	27	0	0	1.00	1.00	1.00	2	0	0	0
Pancreatic AC	2	1.00	0.00	2	27	0	0	1.00	1.00	1.00	2	0	0	0
Uterine carcinosarcoma	1	1.00	0.00	1	28	0	0	1.00	1.00	1.00	1	0	0	0
Sarcoma	6	0.83	0.00	5	24	0	1	1.00	0.83	0.91	5	0	0	1
Mesothelioma	4	0.75	0.00	3	26	0	1	1.00	0.75	0.86	3	0	0	1
LNG_group	5	0.88	0.02	4	23	1	1	0.75	0.88	0.76	4	0	1	0
UCEC	1	0.00	0.00	0	29	0	1	0.00	0.00	0.00	0	0	1	0
Total	201	0.76	0.005	160	3128	19	41	0.86	0.76	0.79	160	7	13	21

Table 2. Metastatic cohorent Performance

Kernel Approach and weighted LS-SVM

Kernel approaches are the set of algorithms,s that use a very large number of data types like sequences, vectors, networks, etc. In this data map of x from the original inputs to high dimensional features with mapping function φ(x). this embedded data is done by a mathematical equation k (xk, xi) which is called a kernel function. By using the number of data items, the size of the matrix could be determined and the complexity or nature does not affect it at all. For example, a set of 100 patients having 6913 gene expression values is represented by a 100x100 kernel matrix (Scholkopf B et.al, 2005). Vapnik developed the support vector machine (SVM) for the supervised classification of kernel approach/kernel algorithms (Vapnik V 1998). Support vector machine (SVM) can equip high dimensional data such as microarray data. Least Squares Support Vector Machine (LS-SVM) is the modified version of Support Vector Machine that was developed by Suykens and coworkers (Suykens JAK et.al, 1999). For high dimensional data sets, LS-SVM is the best option and it is much faster and is based on the linear system instead of quadratic programming.

Association of p53 with lncRNA and regulation of lncRNA

Transcripts with changed expression were determined by gene expression analysis. These changes in expression were due to the damage of DNA which consists of p53 response. To determine p53 direct targets, ChIP-Seq was performed on untreated HCT116 cells and also DNA-damaged cells. A total of 3617 p53 peaks were identified with the help of ChIP-Seq analysis in the DNA-damaged treated cells. It was also found that p53 was bound to 1481 sites without having DNA damage which shows the basal activity of p53 in the cells. In many cases, it was seen that the peaks detected without treatment which is to known bonafide p53 genes targets such as BAX, BBC3, and, CDKN1A. there were a total of 582 gene loci that bound top53, out of which 260 genes loci are protein-coding genes, 80 genes are lncRNAs and 155 are determined as unassigned genomic regions. p53 binding peaks revealed that 60% are localized with 10kb from TSS nearest gene, and the remaining 40% exist between more than 10kb apart.

For the identification of the p53 binding effect on gene expression, p53 ChIP-Seq data were compared with RNA-Seq analysis. So, it was revealed that p53 bound to 109 genes in the DNA-damaged treated cells, which means at least one transcript expressed differently when proximal peaks of p53 were considered. It was estimated that 75% of p53 directly upregulated the genes upon DNA damage while 25% of genes were downregulated. With the help of genomic sequences that were associated with p53 binding peaks, de novo motifs analysis was performed for the confirmation of p53 binding motifs. It was found that the p53 motif was highly supplemented across p53-bound loci, which confirms the presence of p53REs in genomic regions.

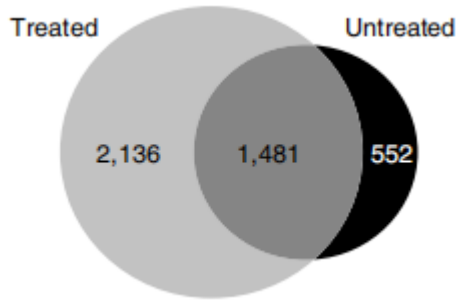

Fig 1. total no. of p53 binding loci determined by p53 ChIP-Seq Analysis

Integral genomic signature approach development

Gene expression profiling data, exome sequence data, and, drug sensitivity measurements of chemical perturbation of 989 cancer cell lines were used for the development of the integral Genomic Signature (iGenSig) approach (Yang W et.al, 2013). A high act area was used for the measurement of drug responses, as the area above the fitted dose response was considered a sensitive drug response, and the area under the dose

curve was considered a resistant drug response. For the uniformity of multi-OMIC characteristics, a Genomic feature Matrix Transposed (GMT) format was formed (Liberzon A 2016 and Chi X et.al, 2019). Using GMT format, the expression profiling data of upregulated, downregulated, mutation hotspots and, mutated genes were analyzed. With the help of 12 overlapping levels of expressed genes, de novo characteristics were generated to increase the cross-data applicability of iGenSig.

For 364 drugs, the iGenSig approach was done which showed a negatively biased drug response distribution in cancer cell lines. These cancer cell lines indicated the narrow effect of responses and 20 out of 364 sensitive cell line subjects have the availability of responders. Many top-performing drugs are FDA-approved targeted therapy agents for the treatment of cancer. These include Lapatinib, Venetoclax, Vincristine, Afatinib, Epirubicin and, Niraparib etc.

Discussions

The integration method was applied for two patient data sets, each having high throughput data sources. The Least Squares Support Vector Machine (LS-SVM) approach was used to build all data sets individually. Then, integrated data were measured manually at multiple points with the help of a change in expression between points. A cancer type miscellanea was presented that supported the whole genome expression of a tumor specimen. Our approaches acquired about 97% of accuracy.p53 ectopic expression was used to assess the activity of p53 non-coding regions. The performance of p53 to the expression of lncRNAs from coding and non-coding regions of the genomes is still absent. Also, it was shown the correlation of p53 with the DNA damage and expression of lncRNAs in human cancer cell lines.

iGenSig approach as high dimensional genomic characteristics to maximize the transfer of multi-OMICS based methods for precision oncology. The integral Genomic Signatures approach was designed to show the cross-data set applicability, transparency, and interpretability problems. It also demonstrated the improved performance of cross-applicability in clinical trials. iGenSig approach performance depends upon the availability of genomic correlates.

Conclusions

Integration methods on experimental data sets in the genome expression could improve the decision support performance. Also, the Kernel approach will be compared with the Bayesian network integration approach in the future. With the advancement in computational approaches, high-dimensional data could be manipulated easily and effectively that used to explore a machine learning method across multiple tumor types.

References

1. Beckerman, R. & Prives, C. Transcriptional regulation by p53. Cold Spring Harb. Perspect. Biol. 2, a000935 (2010).
2. Bender RA, Erlander MG. Molecular classification of unknown primary cancer. Semin Oncol. 2009;36 (1):38-43. doi:10.1053/j.seminoncol.2008.10.002.
3. Bueno R, Stawiski EW, Goldstein LD, et al. Comprehensive genomic analysis of malignant pleural mesothelioma identifies recurrent mutations, gene fusions and splicing alterations. Nat Genet. 2016;48(4): 407-416. doi:10.1038/ng.3520.
4. Cabili, M. N. et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specifific subclasses. Genes Dev. 25, 1915–1927 (2011).
5. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic minority over-sampling technique.J Artif Intell Res. 2002;16:321-357. doi:10.1613/jair.953.
6. Sudheer Menon (2020) "Preparation and computational analysis of Bisulphite sequencing in Germfree Mice" International Journal for Science and Advance Research In Technology, 6(9) PP (557-565).
7. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal (2021) "Computational identification of promoter regions in prokaryotes and Eukaryotes" EPRA International Journal of Agriculture and Rural Economic Research (ARER), Vol (9) Issue (7) July 2021, PP (21-28).
8. Sudheer Menon (2021) "Bioinformatics approaches to understand gene looping in human genome" EPRA International Journal of Research & Development (IJRD), Vol (6) Issue (7) July 2021, PP (170-173).
9. Sudheer Menon (2021) "Insilico analysis of terpenoids in Saccharomyces Cerevisiae"international Journal of Engineering Applied Sciences and Technology, 2021 Vol. 6, Issue1, ISSN No. 2455-2143, PP(43-52).
10. Sudheer Menon (2021) "Computational analysis of Histone modification and TFBs that mediates gene looping" Bioinformatics, Pharmaceutical, and Chemical Sciences (RJLBPCS), June 2021, 7(3) PP (53-70).
11. Cheng DT, Mitchell TN, Zehir A, et al. Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology.J Mol Diagn. 2015;17(3):251-264. doi:10.1016/j.jmoldx.2014.12.006
12. Cherniack AD, Shen H, Walter V, et al; Cancer Genome Atlas Research Network. Integrated molecular characterization of uterine carcinosarcoma. Cancer Cell. 2017;31(3):411-423. doi:10.1016/j.ccell.2017.02.010
13. Chi, X. et al. Universal concept signature analysis: genome-wide quantifification of new biological and pathological functions of genes and pathways. Brief Bioinform. 21, 1717–1732 (2019).
14. Clark AM, Ma B, Taylor DL, Griffith L, Wells A. Liver metastases: microenvironments and ex-vivo models. Exp Biol Med (Maywood). 2016;241(15):1639-1652. doi:10.1177/1535370216658144
15. Sudheer Menon Shanmughavel piramanayakam, Gopal Prasad Agarwal (2021) "FPMD-Fungal promoter motif database: A database for the Promoter motifs regions in fungal genomes" EPRA International Journal of Multidisciplinary research,7(7) PP (620-623).
16. Sudheer Menon, Shanmughavel Piramanayakam and Gopal Agarwal (2021) Computational Identification of promoter regions in fungal genomes, International Journal of Advance Research, Ideas and Innovations in Technology, 7(4) PP (908-914).
17. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) Bioinformatics methods for identifying hirschsprung disease genes, International Journal for Research in Applied Science & Engineering Technology (IJRASET), Volume 9 Issue VII July, PP (2974-2978).
18. Sudheer Menon, (2021), Bioinformatics approaches to understand the role of African genetic diversity in disease, International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET), 4(8), PP 1707-1713.
19. Sudheer Menon (2021) Comparison of High-Throughput Next generation sequencing data processing pipelines, International Research Journal of Modernization in Engineering Technology and Science (IRJMETS), 3(8), PP 125-136.

20. Dinger, M. E. et al. Long noncoding RNAs in mouse embryonic stem cell pluripotency and differentiation. Genome Res. 18, 1433–1445 (2008).
21. Forbes SA, Beare D, Boutselakis H, et al. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res. 2017;45(D1):D777-D783. doi:10.1093/nar/gkw1121
22. Grewal JK, Eirew P, Jones M, et al. Detection and genomic characterization of a mammary-like adenocarcinoma. Cold Spring Harb Mol Case Stud. 2017;3(6):a002170. doi:10.1101/mcs.a002170
23. Gröschel S, Bommer M, Hutter B, et al. Integration of genomics and histology revises diagnosis and enables effective therapy of refractory cancer of unknown primary with PDL1 amplification. Cold Spring Harb Mol Case Stud. 2016;2(6):a001180. doi:10.1101/mcs.a001180
24. Guttman, M. et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458, 223–227 (2009).
25. Hamblin A, Wordsworth S, Fermont JM, et al. Clinical applicability and cost of a 46-gene panel for genomic analysis of solid tumours: retrospective validation and prospective audit in the UK National Health Service. PLoS Med. 2017;14(2):e1002230. doi:10.1371/journal.pmed.1002230
26. Sudheer Menon (2021) Evolutionary analysis of SARS-CoV-2 genome and protein insights the origin of the virus, Wuhan, International Journal of Creative Research Thoughts (IJCRT), 9 (8), PP b696-b704.
27. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) A step-by-step work flow of Single Cell RNA sequencing data analysis, International Journal for Scientific Research and Development (IJSRD), 9(6) PP 1-13.
28. Sudheer Menon (2021) Computational characterization of Transcription End sites in Human Genome, International Journal of All Research Education and Scientific Methods (IJRESM), 9(8), PP 1043-1048.
29. Sudheer Sivasankaran Menon and Shanmughavel Piramanayakam (2021) 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), 2(4), PP 709-714.
30. Sudheer Menon, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) Bioinformatics tools and methods to analyze single cell RNA sequencing data, International Journal of Innovative Science and Research Technology, (IJISRT), 6(8), PP 282-288.
31. Haury AC, Gestraud P, Vert JP. The influence of feature selection methods on accuracy, stability and interpretability of molecular signatures. PLoS One. 2011;6(12):e28210. doi:10.1371/journal.pone.0028210
32. Hudson TJ, Anderson W, Artez A, et al; International Cancer Genome Consortium. International network of cancer genome projects [published correction appears in Nature. 2010;465(7300):966]. Nature. 2010;464 (7291):993-998. doi:10.1038/nature08987
33. Jang, I. S., Neto, E. C., Guinney, J., Friend, S. H. & Margolin, A. A. Systematic assessment of analytical methods for drug sensitivity prediction from cancer cell line data. Pac. Symp. Biocomput. 63–74 (2014).
34. Khalil, A. M. et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. Proc. Natl Acad. Sci. USA 106, 11667–11672 (2009).
35. Khan J, Wei JS, Ringnér M, et al. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. Nat Med. 2001;7(6):673-679. doi:10.1038/89044
36. Sudheer Menon (2021) Computational genome analysis for identifying Biliary Atresia genes, International Journal of Biotechnology and Microbiology, (IJBM), 3(2), PP 29-33.
37. Sudheer Menon (2021) Recent Insilco advancements in genome analysis and characteristics of SARS-Cov2. International Journal of Biology Research, (IJBR), 6(3), PP 50-54.
38. Sudheer Menon (2021) Bioinformatics methods for identifying Human disease genes, International Journal of Biology Sciences, (IJBR), 3(2), PP 1-5.
39. Sudheer Menon (2021) SARS-CoV-2 Genome structure and protein interaction map, insights to drug discovery, International Journal of Recent Scientific Research, (IJRSR), 12(8), PP 42659-42665.
40. Sudheer Menon (2021) Insilico Insights to Mutational and Evolutionary aspects of SARS-Cov2, International Journal of Multidisciplinary Research and Development, (IJMRD) 8(8), 167-172.

41. Laskin J, Jones S, Aparicio S, et al. Lessons learned from the application of whole-genome analysis to the treatment of patients with advanced cancers. Cold Spring Harb Mol Case Stud. 2015;1(1):a000570. doi:10.1101/ mcs.a000570
42. Liberzon, A. A description of the molecular signatures database (MSigDB) web site. Methods Mol. Biol. 1150, 153–160 (2014).
43. Ma XJ, Patel R, Wang X, et al. Molecular classification of human cancers using a 92-gene real-time quantitative polymerase chain reaction assay. Arch Pathol Lab Med. 2006;130(4):465-473.
44. Meiri E, Mueller WC, Rosenwald S, et al. A second-generation microRNA-based assay for diagnosing tumor tissue origin. Oncologist. 2012;17(6):801-812. doi:10.1634/theoncologist.2011-0466
45. Menendez, D., Inga, A. & Resnick, M. A. The expanding universe of p53 targets. Nat. Rev. Cancer 9, 724–737 (2009).
46. Sudheer Menon (2021) 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 9(3), PP 95-99.
47. Sudheer Menon (2021) 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), 4(2), PP 6-13.
48. Sudheer Menon (2021) Computing for Genomics and Proteomics in Microbial biotechnology in the identification of SARS COV2 Variants, International Journal of Scientific Engineering and Applied Science (IJSEAS), 7(9), PP 147-175.
49. Sudheer Menon (2021) Functional Genome analysis and drug screening by Bioinformatics methods in SARS-CoV-2. International Journal of Academic Research and Development (IJARD), 6(5), PP 33-39.
50. Sudheer Menon (2021) Advances in Systems biology and immunological data analysis for drug and vaccine development, International Journal of Educational Research and Studies (IJERS), 3 (3), PP 53-64.
51. Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests: a vignette study.JAMA Intern Med. 2013;173(21):1952-1958. doi:10.1001/jamainternmed.2013.10081
52. Monzon FA, Medeiros F, Lyons-Weiler M, Henner WD. Identification of tissue of origin in carcinoma of unknown primary with a microarray-based gene expression test. Diagn Pathol. 2010;5:3. doi:10.1186/1746-1596-5-3
53. Rapin N, Bagger FO, Jendholm J, et al. Comparing cancer vs normal gene expression profiles identifies new disease entities and common transcriptional programs in AML patients. Blood. 2014;123(6):894-904. doi:10.1182/ blood-2013-02-485771
54. Sudheer Menon (2021) Whole genome sequencing and Insilico approaches to understand the genetic basis of Rare Diseases, International Journal of Advanced Scientific Research (IJASR), 6(5), PP 7-13.
55. Sudheer Menon (2021) Bioinformatics approaches to identify neurodegenerative diseases by Next generation sequencing data, International Journal of Educational Research and Development (IJERD), 3(3), PP 70-74.
56. Sudheer Menon (2021) Bioinformatics approaches for cross-species liver cancer analysis based single cell RNA sequencing data, International Journal of Molecular Biology and Biochemistry (IJMBB), 3(1), PP 05-13.
57. Sudheer Menon (2021) 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), 6(4), PP 14-21.
58. Sudheer Menon (2021) Differential Gene expression analysis by Single Cell RNA sequencing reveals insights to genetic diseases, International Journal of Multidisciplinary Education and Research (IJMER), 6(4), PP 1-7.
59. Sudheer Menon (2021) 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), 8(9), PP 36-55.
60. Sudheer Menon and Binu Thomas (2021) An insights to Bioinformatics methods in Natural Product Drug Discovery, International Journal of Clinical Biology and Biochemistry (IJCBB), 3(1), PP 26-37.

61. Sudheer Menon and Binu Thomas (2021) 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), 6(5), PP 19-29.
62. Sudheer Menon and Binu Thomas (2021) Gene Expression Analysis, Functional Enrichment, and Network Inference in disease prediction, International Journal of Advanced Educational Research (IJAER), 6(5), PP 6-18.
63. Sudheer Menon and Binu Thomas (2021) Vaccine informatics, systems biology studies on neuro-degenerative diseases, International Journal of Multidisplinary Educational Research IJMER, 10(9-4), PP 85-101.
64. Rinn, J. L. & Chang, H. Y. Genome regulation by long noncoding RNAs. Annu. Rev. Biochem. 81, 145–166 (2012).
65. Rinn, J. L. et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 129, 1311–1323 (2007).
66. Robinson DR, Wu YM, Lonigro RJ, et al. Integrative clinical genomics of metastatic cancer. Nature. 2017;548 (7667):297-303. doi:10.1038/nature23306
67. Sudheer Menon (2021) Computational approaches to understand the role of epigenetics and microRNA in Cancer among different populations, International Journal of Multidisciplinary Research Review (IJMDRR) 7(9), PP 69-81.
68. Sudheer Menon and Binu Thomas (2021) Recent advances, methods and opportunities of Deep learning in computational and systems biology International Journal of Applied and Advanced Scientific Research (IJAASR) 6(2), PP 12-25.
69. Sudheer Menon (2021) Computational genome based studies of Complex Genetic Diseases, International Journal of Multidisciplinary Research and Modern Education (IJMRME) 7(2) PP 28-39.
70. Sudheer Menon (2021) 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), 6(2) PP 10-33.
71. Sudheer Menon, Amanda Ramer Tait, Shanmughavel Piramanagakam, Vincent Chi Hang Lui and Paul Kwong Hang Tam (2021) Bisulphite seq data reveals host-microbiota interaction in germfree and conventional mice, International Journal of Bioinformatics and Biological Sciences, 9(1), PP 1-8.
72. Sudheer Menon (2021) Insilico identification of Transcription End Sites in Eukaryotes, International Journal of Pharmacy and Biological Sciences (IJPBS), 11(3), PP-104-115.
73. Sudheer Menon (2021) 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), 7(5), PP 68-83.
74. Sudheer Menon (2021) 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), 7(6), PP 86-96.
75. 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), PP 433-445.
76. Binu Thomas, Mahesh Mohanan, and Sudheer Menon (2021), Role of edible Plants to balance dietary requirements in human diet, Devagiri Journal of Science, 7(1), PP-27-40
77. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), 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, 7(1), PP 1-20.
78. Shawe-Taylor J, Cristianini N: Kernel Methods for Pattern Analysis. Cambridge: Cambridge University Press; 2004.
79. Stefanovic S, Wirtz R, Deutsch TM, et al. Tumor biomarker conversion between primary and metastatic breast cancer: mRNA assessment and its concordance with immunohistochemistry. Oncotarget. 2017;8(31):51416-51428. doi:10.18632/oncotarget.18006 .
80. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), CDKN2A gene expression/mutation prognosis by computational methods in the identification of nodular melanoma, International Journal of Scientific Research and Modern Education (IJSRME), 7(1), PP 11-22.

81. Sudheer Menon , Vincent Chi Hang Lui1, Paul Kwong Hang Tam (2022), 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, 9(4), PP 1-9.
82. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), 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), 7(1), PP 9-29.
83. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), 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), 7(1), PP 21-38.
84. Sudheer Menon, Shanmughavel Piramanayagam and Gopal Prasad Agarwal (2022), Identification of immune gene–related IncRNA signature to detect brain tumors, International Journal of Current Research and Modern Education (IJCRME), 7(1), PP 6-20.
85. Sudheer Menon, Shanmughavel piramanayakam and Gopal Prasad Agarwal (2023), "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), 10 (12), PP- g398-g416.
86. Sudheer Menon (2023), Dynamical Modeling of the Core Gene Network and Mutation in Transitional Cell Carcinoma, Journal of Emerging Technologies and Innovative Research (JETIR), 10 (12), PP-h326-h338.
87. 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.
88. Ulitsky, I. & Bartel, D. P. lincRNAs: genomics, evolution, and mechanisms. Cell 154, 26–46 (2013).
89. Varadhachary GR, Raber MN, Matamoros A, Abbruzzese JL. Carcinoma of unknown primary with a coloncancer profile-changing paradigm and emerging definitions. Lancet Oncol. 2008;9(6):596-599. doi:10.1016/ S1470-2045(08)70151-7
90. Vennalaganti P, Kanakadandi V, Goldblum JR, et al. Discordance among pathologists in the United States and Europe in diagnosis of low-grade dysplasia for patients with Barrett's esophagus. Gastroenterology. 2017;152(3): 564-570.e4. doi:10.1053/j.gastro.2016.10.041
91. Wang HL, Kim CJ, Koo J, et al. Practical immunohistochemistry in neoplastic pathology of the gastrointestinal tract, liver, biliary tract, and pancreas. Arch Pathol Lab Med. 2017;141(9):1155-1180. doi:10.5858/arpa.2016- 0489-RA
92. Wang, K. C. & Chang, H. Y. Molecular mechanisms of long noncoding RNAs. Mol. Cell 43, 904–914 (2011).
93. Yang, W. et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 41, D955–D961 (2013).
94. Yu, L., Zhou, D., Gao, L. & Zha, Y. Prediction of drug response in multilayer networks based on fusion of multiomics data. Methods 192, 85–92 (2020).
95. Zoon CK, Starker EQ, Wilson AM, Emmert-Buck MR, Libutti SK, Tangrea MA. Current molecular diagnostics of breast cancer and the potential incorporation of microRNA. Expert Rev Mol Diagn. 2009;9(5):455-467. doi:10. 1586/erm.09.25