'Tumor Hypoxia Diagnosis' using Deep CNN Learning strategy a theranostic pharmacogenomic approach

Vaisali B[1], Parvathy C R [2], Hima Vyshnavi A M [3] and Krishnan Namboori P K [4]

[1,2,3,4]*Computational Chemistry Group (CCG), Amrita School of Engineering, Coimbatore*
Amrita Vishwa Vidyapeetham, India

b_vaisali@cb.students.amrita.edu
cb.en.p2bme16011@cb.students.amrita.edu
cb.en.d.cen17001@cb.students.amrita.edu
n_krishnan@cb.amrita.edu

ABSTRACT

Tumor hypoxia results in most of the anticancer drugs becoming ineffective. However, due to lack of proper signaling in the hypoxic micro environment, the condition cannot be detected in advance, leading into unnecessary delay in the diagnosis and treatment. The main objective of the work is to identify the 'hypoxia prone SNPs to help the patients to predict their possibility of hypoxia formation and to Design and develop a machine helping in diagnosing the hypoxia from pathological images using deep learning with 'convolution neural network'. The genetic signatures corresponding to 'tumor hypoxia development' have been identified by pharmacogenomic method, comprising of genomics, epigenomics, metagenomics and environmental genomics. All the common hypoxia related mutations have been included in the study. The formation of the hypoxia condition has to be carefully identified and monitored during the process of treatment to ensure that the right drug is being administered. In the present manuscript, a novel method of elucidating the condition using 'deep convolution network' from simple pathological image has been suggested. The efficiency of the suggested machine is found to be 92.8% making it as a potential device for prediction of hypoxia mutation and thereby helping us to monitor the hypoxic conditions effectively. Thus, the hypoxia prone SNPs corresponding to common mutations have been identified. The patients having the hypoxia prone SNPs are advised to guard against hypoxia formation with the help of diagnostic tests using the machine. The machine helps to warn the patients against the respective mutations from simple pathological image of the tumor cells.

1. INTRODUCTION

Cancer is one of the terminal diseases leading to large scale of mortality every year all around the world (Fitzmaurice et al., 2015). The normal symptoms of cancer include increased proliferation, decreased apoptotic pathway functions, deregulated metabolism and depletion in the cellular oxygen content.

Although treatment strategy for cancer involves multiple protocols such as chemotherapy, radiation therapy, surgical intervention etc., the failure rate remains unexpectedly high. One of the major reasons behind this is the development of a cellular condition known as hypoxia in cancer cells along with prolonged treatment of cancer, making these cells highly resistant to most of the anti-cancer drugs. The condition is characterized by maintaining insufficient oxygen within the cancer cells leading into a decrease in the cellular metabolic rate. This results in a different cellular environment, where most of the anti-cancer drugs fail to function, providing a natural 'drug resistance'. The 'hypoxia condition' can be considered as a ‘protective adaptation’ by the cancer cells especially solid tumor cells to increase anti-cancer drug resistance (Sriraman, Aryasomayajula & Torchilin, 2014). In most cases, this unfavorable condition of cancer cells triggers off extensive metastasis and accelerated malignant progression. The hypoxia in sarcomas leads to distant metastasis while hypoxia in cervical cancer results in local and regional spreading of cancer

Another major challenge associated with hypoxia is the difficulty in its early detection as the condition does not support proper signaling for recognizing the reactive oxygen species (ROS) (Fleet, 2006) , used for the diagnosis of hypoxia. This may lead the tumor cells to over-populate and promote metastasis (Brown & Wilson, 2004), (Wilson & Hay, 2011) excessively. Even an effective drug delivery system may fail to reach the region of hypoxia because of the poorly developed blood vessels, deregulated metabolism and increased drug resistance.

Though oxygen sensors such as Eppendorf needle are suggested to monitor hypoxia condition, due to the operational difficulty in introducing individual needle sensors, the technique is not widely accepted. The non- invasive analysis using indirect assays, studying the hypoxia inducible factors (HIF), bio-reductive metabolism, etc. has been suggested to measure and monitor the condition. Few imaging technique such as blood oxygen level-dependent magnetic resonance imaging (BOLD-MRI), phosphorescence have been introduced. The major disadvantage of using BOLD-MRI is that it measures only deoxyhemoglobin concentrations. The toxicity of phosphorescence dye used in the analysis and the inability to assess deeper tissues are the drawbacks of the phosphorescence based imaging technique. Moreover, the pharmacogenomic individual variations seen in the diagnostic finger prints of patients demand a 'person-specific diagnostic system' incorporating the attributes such as genomics, epigenomics, metagenomics, environmental genomics and drug genomics (HimaVyshnavi et al., 2017), (Iyer, Karthikeyan, Sanjay Kumar & Krishnan Namboori, 2017), (Iyer, Palayat, Shanmugam & Namboori, 2017). The early diagnosis of hypoxia condition associated with cancer is still a challenge.

The 'deep convolution neural network (CNN)' based learning environment has been reported as a novel efficient theranostics technique to get biological functional information from cellular images (Rawat & Wang, 2017). Well established 'tensor flow Convolution Neural Networks for CIFAR-10' as shown in ("TensorFlow Tutorial | Deep Learning Using TensorFlow | Edureka", 2018) has been widely used to address biological functionalities including molecular biology and genomic imprinting.

The CNN has been used in the analysis of histopathological images for a few biological conditions and found to be very efficient in retrieving diagnostic and prognostic information about the disease conditions (Khosravi, Kazemi, Imielinski, Elemento & Hajirasouliha, 2018), (Komura & Ishikawa, 2018), (Qu et al., 2018). In the present work, a novel approach has been used in studying the possibility of using this technique in identifying the ‘hypoxia condition’ associated with cancer treatment by incorporating deep CNN learning environment and correlating the same with pharmacogenomic variants.

2. MATERIALS AND METHODS

2.1. Pharmacogenomics

The SNPs have been identified as the most effective genetic variant incorporating pharmacogenomic finger prints (Sherry, 2001). The hypoxia or de-oxygenation causing mutations, ABCC1, ABCB1, MTHFR, RFC1, HPRT1, CYP2B6, CYP2C8, CYP2C9, ADAM17, CYP2C19, CYP2D6, HIF2A, HIF1A, CYP1A1, CYP1A2, CASP1, AKR1C1, AKR1C2, PTGS2, CASP6 and their genomic and epigenomic contributions have been identified. The genetic signatures (SNPs) behind all the relevant pharmacophoric attributes genomics, epigenomics, metagenomics, environmental genomics and drug genomics have been listed out.

These SNPs are further characterized and classified into damaging, tolerated, benign, possibly damaging, probably damaging and deleterious using Sorting Intolerant from Tolerant (SIFT) and Polymorphism Phenotyping (Polyphene) analysis (Ng, 2003), (Adzhubei, Jordan & Sunyaev, 2013). The probable metagenomic contribution in the formulation of hypoxia is identified by comparing the microbial genome with the genes responsible for hypoxia using Basic Local Alignment Search Tool (BLAST) of National Center for Biotechnology Information (NCBI) (Ye, Ma, Madden & Ostell, 2013).

The environmental factors causing the mutations have been noted down from the 'Comparative Toxicogenomics Database (CTD)' (Davis et al., 2016). The anti-tumor drugs were taken from the drug bank and coding SNPs that are most likely to have an impact on biological function were identified from LS-SNP/PDB tool. Solid tumor related genes have been taken up for the analysis and their epigenetic contributions towards variations have been studied (Bock, Walter, Paulsen & Lengauer, 2007), (Dworkin, Huang & Toland, 2009). The DNA methylation is found to be the most prominent epigenetic mechanism responsible for causing variation in CpG dinucleotide (Thienpont et al., 2016). The attributes corresponding to methylation of DNA have been identified.

2.2. Data Collection

The tumor pathological images corresponding to the 'hypoxia causing mutations' namely ABCC1, ABCB1, MTHFR, RFC1, HPRT1, CYP2B6, CYP2C8, CYP2C9, ADAM17,CYP2C19, CYP2D6, HIF2A, HIF1A, CYP1A1, CYP1A2, CASP1, AKR1C1, AKR1C2, PTGS2, CASP6 have been acquired from 'Human Protein Atlas'. To support the classification need, non-hypoxic tumor cell-images corresponding to the mutations BRCA1, BRCA2, BCAR3, BRMS1 and BCAS1 have been collected.

2.3. Deep learning Implementation

Totally, 300 breast cancer sample images were included in the present analysis, among which 150 were hypoxia positive breast cancer images and the remaining were hypoxia negative breast cancer images to avoid class imbalance. From the 300 samples ‘pathology images’, 80% have been considered as the training set and the remaining 20% as the testing set. The labels are encoded and used for training purpose and the python library tensor flow is used for performing deep CNN in the model ("The Human Protein Atlas", 2018).

2.4. Algorithm and optimization of the conditions

The structure of the prediction model is depicted in Figure.1. In this manuscript, the pathological images collected are of larger pixel values. In order to avoid higher weights in the initial hidden layers, the CNN is not fully connected instead is attached to few regions of the layer to avoid over fitting by changing the hyper parameters such as the filter size (to 7), epochs (to 50) and increasing the number of layers.

In the optimized process, there are totally 23 layers consisting of alternating convolution layer, maxpooling layer and rectified

linear unit activation layer along with two fully connected layers, where the output of alternative individual layers is stacked together. The cross entropy is calculated and an 'adaptive moment estimation' has been used to optimize the network weights by an iterative method. The deep neural net has been trained to different epochs to converge the results and provide maximum prediction accuracy (Figure 2).

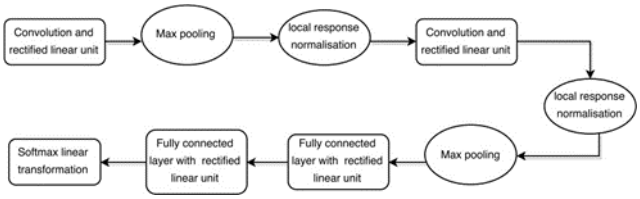

Figure 1. The model structure of tensor flow convolution neural network

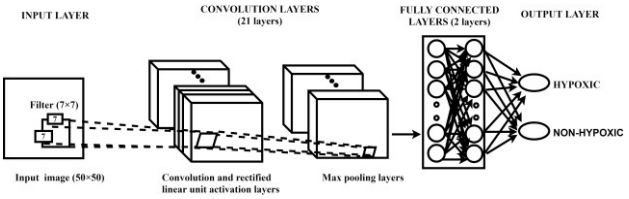

Figure 2. Optimized model structure

3. RESULTS AND DISCUSSION

Through classification, the upregulation of hypoxia gene mutations has been identified. The hypoxia related SNPs have been identified from the pharmacogenomic analysis through online database such as dbSNP – NCBI and SNP Nexus. The pharmacogenomic and deep learning analysis have been carried out parallelly and a correlation has been set up between the results. Thus, deep CNN is used for image classification task to identify the specific mutations responsible for hypoxia. Early detection of the development of hypoxia is made possible through the deep CNN model, while the proneness of hypoxia formulation has been made possible through the pharmacogenomic model (Namboori et al., 2011).

3.1. Genomics

The primary mutations responsible for setting up of bio reductive conditions are ABCC1, ABCB1, MTHFR, RFC1, HPRT1, CYP2B6, CYP2C8, CYP2C9, ADAM17, CYP2C19, CYP2D6, HIF2A, HIF1A, CYP1A1, CYP1A2, CASP1, AKR1C1, AKR1C2, PTGS2, and CASP6. The major SNPs responsible for the variations are included in Table 1. Obviously, the people with the SNPs expressed in their respective genes are more prone to the mutations.

Sl. no.	Genes	Total no. of SNPs	Frequent deleterious SNPs
1	ABCC1	28895	rs183032276, rs186193767, rs201020041
2	ABCB1	24289	rs1128501, rs137996914, rs139820108
3	MTHFR	10624	rs121434296, rs200138092, rs116620395
4	RFC1	19539	rs12502450, rs147804632, rs190369900
5	HPRT1	3353	rs137852480, rs137852493, rs137852496
6	CYP2B6	3948	rs117872433, rs12721655, rs138594605
7	CYP2C8	4512	rs369552457, rs141209951, rs150733212
8	CYP2C9	12018	rs150663116, rs200382419, rs28371687
9	ADAM1 7	8194	rs370514738, rs142946965, rs201123474
10	CYP2C1 9	14307	rs148593307, rs267602634, rs41291556
11	CYP2D6	1714	rs369390846, rs369772253, rs1058172
12	HIF2A	12595	rs191706577, rs28940297, rs119476044
13	HIF1A	6635	rs28940297, rs28940298, rs142376463
14	CYP1A1	1304	rs367604147, rs371662141, rs373568981
15	CYP1A2	1488	rs200571120, rs201537008, rs28399424
16	CASP1	1800	rs368177280, rs371404424
17	AKR1C1	1920	rs142200840, rs368611374, rs372782197
18	AKR1C2	2431	rs201515806
19	PTGS2	1663	rs148160346, rs201588411, rs371762608
20	CASP6	2153	rs144996365, rs1042887

Table 1. Genes with total no. of Single nucleotide polymorphisms and deleterious SNPs

3.2. Epigenomics

While considering the attributes corresponding to epigenetic variations, it has been found that the repetitive DNA, evolutionary history, transcriptome, epigenome & chromatin structure are the factors contributing most towards 'hypoxia micro environment (Figure 3). The SNPs in the 'methylation prone region' called Methylation prone SNPs (MeSNPs) have been identified (Table 2). The persons with these SNPs are more inclined to epigenomic variations.

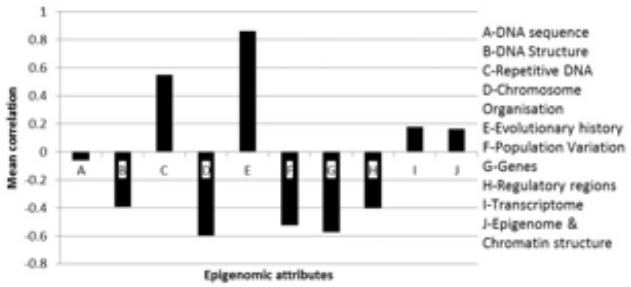

Figure 3. Epigraph of attributes compared to average mean correlation of the mutation genes

Sl. no.	Genes	MeSNPs
1	CYP1A1	rs146622566, rs149687459, rs151244239
2	CYP1A2	rs376179316, rs45565238, rs55802037
3	AKR1C1	rs142200840, rs370027719
4	ABCC1	rs200922662, rs201533167
5	ABCB1	rs199551851, rs201564736
6	MTHFR	rs373747884, rs200379144
7	RFC1	rs12502450, rs147804632, rs199688793
8	CYP2B6	rs142421637, rs144760726, rs148009906
9	CYP2C8	rs113008582, rs188111115, rs369552457
10	CYP2C9	rs28371674, rs12414460, rs141489852
11	CYP2C19	rs148593307, rs149590953, rs17884712
12	CYP2D6	rs1058172, rs138229048, rs139441693
13	CASP6	rs200347007, rs370198937
14	HIF2A	rs139763806, rs149676792
15	ADAM17	rs370064783

Table 2. Major hypoxia genes and corresponding methylated SNPs (MeSNPs)

3.3. Metagenomics

The metagenomic analysis describes the influence of microbes living in our body in supporting the mutations. The microbes having genome similar to the hypoxia genes have been identified as most influencing in causing the mutation (Banerjee, Mishra & Dhas, 2015). The corresponding SNPs have been included in Table 3, suggesting the people with these SNPs are more susceptible towards metagenomic variations.

Sl. no.	Genes	Microbe	Deleterious SNPs
1	HIF1A	Mycobacterium tuberculosis	rs1015149462, rs778992658
2	CYP1A1	Rhodoferax ferrireducens	rs146622566, rs577523247
3	CYP1A2	Rhodoferax ferrireducens	rs59410695, rs139412032
4	AKR1C1	Sulfurospirillum multivorans	rs756699697, rs916659284
5	AKR1C2	Mycobacterium tuberculosis	rs11818926, rs34371823
6	PTGS2	Klebsiella pneumoniae	rs572342120, rs775143480
7	ABCC1	Klebsiella pneumoniae	rs113607327, rs746604799
8	ABCB1	Haemophilus influenzae	rs886871489, rs894658296
9	MTHFR	Mycobacterium tuberculosis	rs2066462, rs17854808
10	RFC1	Niabella ginsenosidivorans	rs923604401, rs977954175
11	HPRT1	Clostridium perfringens	rs956631698
12	CYP2B6	Mycobacterium tuberculosis	rs747003572, rs995224549
13	CYP2C8	Klebsiella pneumoniae	rs61757319, rs372960611
14	CYP2C9	Ehrlichia ruminantium	rs143715236
15	CYP2C19	Klebsiella pneumoniae	rs966182584, rs984474043
16	CYP2D6	Mycobacterium	rs77449786, rs368135263
17	CASP6	Mycobacterium tuberculosis	rs74973078, rs369682210
18	HIF2A	Xanthomonas citri	rs888134923, rs993940618

Table 3. Hypoxia genes and proneness to microbial influence.

3.4. Environmental factors

Many epidemiological studies prove that the environmental factors also contribute towards mutations (Boffetta & Nyberg, 2003). The mutagens such as resveratrol, deferoxamine, quercetin, benzopyrenes, methylcholanthrene, beta-naphthoflavone, estradiol, dinoprostone, tetrachloro dibenzodioxin, celecoxib, indomethacin etc. are found to be most influential in causing the mutations. The chemicals interacting with the genes and having toxic effect are included in the table 4

Sl. no.	Genes	Chemicals involved
1	HIF1A	Resveratrol, Deferoxamine, Quercetin
2	CYP1A1	Methylcholanthrene, Benzo(a)pyrene, resveratrol, Estradiol
3	PTGS2	Dinoprostone, resveratrol, Tetrachlorodibenzodioxin, Celecoxib, Indomethacin, nimesulide
4	ABCC1	Vincristine, Etoposide
5	ABCB1	Verapamil, Paclitaxel, Cyclosporine, valspodar, Vinblastine
6	MTHFR	Folic Acid, Fluorouracil, Arsenic
7	CYP2D6	Tamoxifen

Table 4. Genes and Environmental factors influencing mutation.

3.5. Drug genomics

The drug genomics study correlates between the genetic signatures for various proteins specific to hypoxia genes and anti-cancer drugs that could be effective when advised under identification of the corresponding SNPs (Table 5)

Sl. no.	Drug	Proteins	SNPs
1	Flurouracil	1OG2, 1OG5, 1R9O, 4NZ2	rs142123260, rs148615754
2	Irinotecan	3IBD	rs138264188, rs183427203
3	Oxaliplatin	4I8V	rs112517897, rs142388113
4	Regorafenib	1PQ2, 2NNH, 3IBD	rs150733212, rs138264188
5	Paclitaxel	4C3Z	rs186193767, rs200085313
6	Cyclophosphamide	2F9Q, 3QM4	rs1058172, rs141289473

Table 5. Hypoxia associated proteins and proneness to drug action.

3.6. Deep CNN

The predictive machine attained convergence in 33 iterations with an accuracy of 92.8%, trained to 50 epochs. The parameters have been defined and the machine is set to give a paramount accuracy in detecting the hypoxic condition from a given pathological image (Figure 4).

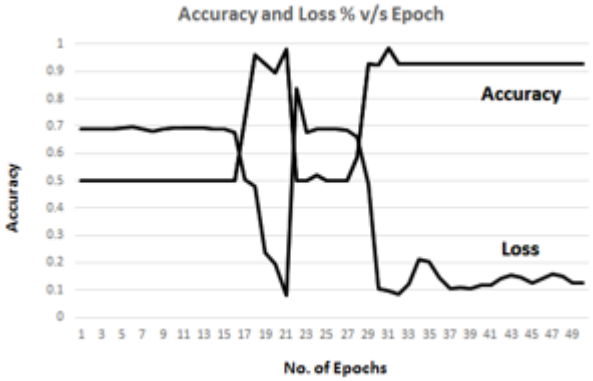

Figure 4. The accuracy and loss % versus epochs where the accuracy and loss tend to converge at epochs 33 and 37 respectively.

4. CONCLUSION

The hypoxia leading mutations, ABCC1, ABCB1, MTHFR, RFC1, HPRT1, CYP2B6, CYP2C8, CYP2C9, ADAM17, CYP2C19, CYP2D6, HIF2A, HIF1A, CYP1A1, CYP1A2, CASP1, AKR1C1, AKR1C2, PTGS2, CASP6 have been included in the analysis. The genetic signatures corresponding to proneness of these mutations have been listed. Moreover, the epigenetic, metagenomic and environmental factors and their SNPs leading into hypoxic conditions have been computed. Early identification and a continuous monitoring of hypoxia is essential in making the cancer treatment effective. This can be made using the 'deep CNN based image processing' of pathological images. The predictive model designed has been identified as a potential tool for identifying the tumour hypoxia and the mutation behind it. The machine gives a predictive accuracy of 92.8%, suggesting the tool as a useful device for tumor hypoxia prediction. The tool helps in incorporating the theranostic pharmacogenomic approach for the early detection and continuous monitoring of tumor hypoxia.

The deep CNN based support systems have been identified as potential theranostic devices in modern diagnostic and prognostic scenario especially in the pharmacogenomic approach, where a person specific continuous monitoring system is highly appreciated. The strategy followed in the paper helps in making similar correlations with all diseases and conditions and

providing 'simple, effective and economic theranostic' devices. This may further be extended to address critical biological and medical conditions that are expensive and time consuming to detect in the early stage and providing a 'personalized' strategy. The individual mutations can be identified specific to the setting up of tumor micro environment.

ACKNOWLEDGEMENT

The authors express their gratitude towards the officials and members of Amrita Molecular Modeling and Synthesis (AMMAS) Research Lab at Amrita Vishwa Vidyapeetham, Coimbatore for their support. The author, Vaisali B expresses her gratitude to AMRITA Vishwa Vidyapeetham for meeting her fellowship and Hima Vyshnavi A M expresses her gratitude to the Ministry of Electronics and Information Technology, Government of India for her Visvesvaraya Scholarship while doing the project.

REFERENCES

Fitzmaurice, C., Dicker, D., Pain, A., Hamavid, H., Moradi-Lakeh, M., MacIntyre, M. F., & Hamadeh, R. R. (2015). The global burden of cancer 2013. *JAMA oncology*, *1*(4), 505-527.

Sriraman, S. K., Aryasomayajula, B., & Torchilin, V. P. (2014). Barriers to drug delivery in solid tumors. *Tissue barriers*, *2*(3), e29528.

Fleet, A. (2006). Radiobiology for the Radiologist: Eric J. Hall, Amato J. Giaccia, Lippincott Williams and Wilkins Publishing; ISBN 0-7817-4151-3; 656 pages; 2006; Hardback; £ 53. *Journal of Radiotherapy in Practice*, *5*(4), 237-237.

Brown, J. M., & Wilson, W. R. (2004). Exploiting tumour hypoxia in cancer treatment. *Nature Reviews Cancer*, *4*(6), 437.

Wilson, W. R., & Hay, M. P. (2011). Targeting hypoxia in cancer therapy. *Nature Reviews Cancer*, *11*(6), 393.

Hima Vyshnavi, A.M., Anand, C.L., Deepak, O.M Namboori, P. K. (2017). Evaluation of Colorectal Cancer (CRC) Epidemiology A Pharmacogenomic Approach. *Journal of Young Pharmacists*, *9*(1), 36.

Iyer, P. M., Karthikeyan, S., Kumar, P. S., & Namboori, P K. (2017). Comprehensive strategy for the design of precision drugs and identification of genetic signature behind proneness of the disease—a pharmacogenomic approach. *Functional & integrative genomics*, *17*(4), 375-385.

M Iyer, P., K Palayat, S., Shanmugam, K., & Namboori, K. (2016). Retrometabolic Approach for Designing Personalized Anti-Cancer Drug Molecules for Controlling Breast Cancer Resulted by BRCA1 Mutations. *Current Pharmacogenomics and Personalized Medicine (Formerly Current Pharmacogenomics)*, *14*(1), 56-64.

Rawat, W., & Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural computation*, *29*(9), 2352-2449.

Bakshi, A. (2017). Deep Learning Tutorial: Artificial Intelligence Using Deep Learning. *Deep learning tutorial series.*

Khosravi, P., Kazemi, E., Imielinski, M., Elemento, O., & Hajirasouliha, I. (2018). Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous Digital Pathology Images. Ebiomedicine, 27, 317-328.

Komura, D., & Ishikawa, S. (2018). Machine Learning Methods for Histopathological Image Analysis. Computational And Structural Biotechnology Journal, 16, 34-42. doi: 10.1016/j.csbj.2018.01.001

Qu, J., Hiruta, N., Terai, K., Nosato, H., Murakawa, M., & Sakanashi, H. (2018). Gastric Pathology Image Classification Using Stepwise Fine-Tuning for Deep Neural Networks. *Journal Of Healthcare Engineering*, 2018, 1-13.

Namboori, P., Vineeth, K., Rohith, V., Hassan, I., Sekhar, L., Sekhar, A., & Nidheesh, M. (2011). The ApoE gene of Alzheimer's disease (AD). Functional & Integrative Genomics, 11(4), 519-522.

Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic acids research*, *29*(1), 308-311.

Dworkin, A. M., Huang, T. H. M., & Toland, A. E. (2009, June). Epigenetic alterations in the breast: Implications for breast cancer detection, prognosis and treatment. In *Seminars in cancer biology* (Vol. 19, No. 3, pp. 165-171). Academic Press.

Thienpont, B., Steinbacher, J., Zhao, H., D'Anna, F., Kuchnio, A., Ploumakis, A., & Hermans, E. (2016). Tumour hypoxia causes DNA hypermethylation by reducing TET activity. *Nature*, *537*(7618), 63.

The Human Protein Atlas. (2018). Retrieved from https://www.proteinatlas.org/

Banerjee, J., Mishra, N., & Dhas, Y. (2015). Metagenomics: A new horizon in cancer research. *Meta gene*, *5*, 84-89.

Boffetta, P., & Nyberg, F. (2003). Contribution of environmental factors to cancer risk. *British medical bulletin*, *68*(1), 71-94.

Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. *Nucleic acids research*, *31*(13), 3812-3814.

Adzhubei, I., Jordan, D., & Sunyaev, S. (2013). Predicting Functional Effect of Human Missense Mutations Using PolyPhen-2. *Current Protocols in Human Genetics*, *76*(1), 7.20.1-7.20.41. doi: 10.1002/0471142905.hg0720s76.

Ye, J., Ma, N., Madden, T. L., & Ostell, J. M. (2013). IgBLAST: an immunoglobulin variable domain sequence analysis tool. *Nucleic acids research*, *41*(W1), W34-W40.

Davis, A. P., Grondin, C. J., Johnson, R. J., Sciaky, D., King, B. L., McMorran, R., ... & Mattingly, C. J. (2016). The comparative toxicogenomics database: update 2017. *Nucleic acids research*, *45*(D1), D972-D978.

Bock, C., Walter, J., Paulsen, M., & Lengauer, T. (2007). CpG Island mapping by epigenome prediction. *PLoS computational biology*, *3*(6), e110.

BIOGRAPHIES

Dr. P. K. Krishnan Namboori currently serves as an Associate Professor at the Center for Excellence in Computational Engineering and Networking (CEN) and the head of the Computational Chemistry Group and Amrita Molecular Modeling and Synthesis (AMMAS) Research Lab, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore campus. He is also an Adjunct Faculty at the School of Biotechnology at Amritapuri Campus, Amrita Vishwa Vidyapeetham. He is presently working on pharmacogenomics and 'design and development of personalized drugs' for diseases such as cancer, life style diseases etc. A number of international journal publications, conference proceedings, four published patents and one international book are to his credit.

B. Vaisali received her M. Tech degree in Bio-Medical Engineering from Department of Electronics and Communication Engineering at Amrita School of engineering, Amrita Vishwa Vidyapeetham, India in the year 2018. She is currently working as a research scholar towards her Ph. D. degree in Amrita Molecular Modelling and Synthesis (AMMAS) Research lab, Amrita Vishwa Vidyapeetham. Her project is entitled 'Anti Hypoxic potentiation for anti-cancer drugs'. The main objective of the project is to identify a pharmacogenomic drug for suppressing hypoxia tumor microenvironment inducing genes. Her areas of interest include Drug design and delivery systems, Cancer research.

Parvathy. C. R received her M. Tech degree in Bio-Medical Engineering from Department of Electronics and Communication Engineering at Amrita School of engineering, Amrita Vishwa Vidyapeetham, India in the year 2018. Associating with 'Amrita Molecular Modeling And Synthesis (AMMAS)' research lab, Amrita Vishwa Vidyapeetham, Coimbatore, she completed the project entitled 'Design & Biological Evaluation of a Dual Specificity Kinase Inhibitor for Breast Cancer'. The main objective of the project is to identify a pharmacogenomic drug for mischievous kinase by considering binding specificity and resistance of kinase inhibitors with breast cancer target proteins. She is currently serving as Director-Research in NEO Research consultancy Pvt. Ltd., DLF Cyber City, Gurgaon, New Delhi. Her areas of interest include Biomedical sciences, Cancer research and Computational chemistry

Hima Vyshnavi A. M, after completion of her M. Tech in Computational Engineering and Networking, is currently pursuing her Ph. D in 'design and development of precision drugs for breast cancer move towards P5 pharmacogenomic approach' at AMMAS research lab, Amrita School of Engineering, Amrita Vishwa Vidyapeetham. She has two conference publications and two peer reviewed journal publications to her credit. She is the recipient of the prestigious 'Visvesvaraya Scholarship' of the 'Ministry of Electronics and Information Technology', Government of India. Her areas of interest include Bioinformatics, Cancer Research and Life Sciences.