Functional Analysis of Subpopulation-Specific DNA Ligase 1 Gene Polymorphisms in Human Cancers

Abigail Weir[1,], Katherine Marlow[2], Razan Waliagha[3], and Amit Kumar Mitra[3,4]*

[1] Department of Chemistry and Biochemistry, COSAM, Auburn University
[2] Department of Curriculum and Teaching, College of Education, Auburn University
[3] Department of Drug Discovery and Development Harrison, College of Pharmacy, Auburn University
[4] Center for Pharmacogenomics and Single-Cell Omics initiative (AUPharmGx), Auburn University

Pharmacogenomics leverages the relationship between a person's genome and its specific effect on therapeutic treatments [8]. The use of pharmacogenomics has been a growing topic in the pharmaceutical industry because it determines how the unique genetic makeup of each person determines the best course of therapeutic treatment. The link between the genome and treatments can be seen in the use of single nucleotide polymorphisms (SNPs). SNPs are single base pair changes of DNA, or point mutations, and are one of the sources of variations between human genomes. These variations can act as genetic markers of how an individual's genome responds to a therapeutic treatment [7].

Global projects, like the Human Genome Project, looked to understand the implication of genetic variations on human health. The Human Genome Project has come to understand much about common and rare diseases using a full human genome sequence [1]. However, this project had a broad focus on global populations. SNPs can vary by region, and projects like the Human Genome Project, do not consider the higher frequency of SNPs and genetic heterogeneity in certain areas.

India is a very diverse region. It takes up 2.4 % of the world's land area and has over 17.5 % of the world's population. India also contains a wide range of subpopulations divided by factors such as language, religion, geographic region, etc. The goal of the Indian Genome Variation Consortium was to perform an in-depth study of SNPs and repeats specific to Indian subpopulations [2]. Our primary focus was to identify functionally relevant SNPs involved in DNA repair and replication, such as DNA Ligase 1 [4].

DNA Ligase 1 (LIG1) is a replicative ligase involved in multiple mechanisms of DNA repair and replication. LIG1 functions by reforming the phosphodiester bond of a nicked DNA strand. This enzyme works to join Okazaki fragments in DNA replication as well as functions in DNA repair pathways such as base excision repair and nucleotide excision repair [9].

This project was planned in 3 phases. Phase 1: SNPs in LIG1 were evaluated based on the predicted functional effect. The collected data was used to determine the SNP predicted to be the most damaging. Phase 2: Determine the difference in expression of LIG1 and verify the presence of the SNP in different prostate cancer cell lines. Phase 3: Investigate the role of the top functional SNPs on the DNA repair and replication system by introducing the SNP into human cancer cell lines.

Successful completion of this project will determine deleterious SNPs in LIG1 associated with DNA repair and replication defects and eventually resistance to anti-cancer drugs. Ultimately, this project will assist with a personalized medicine approach to DNA-repair-deficiency-associated cancers specific to world subpopulations like India.

The SNP discovery and validation were based on determination of minor allele frequencies (MAF) in 55 different subpopulations of India representing all major linguistic subgroups and ethnicities. Further SNP evaluations were conducted using the National Center for Biotechnology Information (NCBI). NCBI data provides information about the SNP's chromosome position, alleles, clinical significance, gene consequence,

* Corresponding author: aew0106@auburn

and aggregate allele frequency. Criteria for SNP selection involved gene importance in cancer mechanisms, consequence of the variation, and frequency of the SNP.

For pharmacogenomic evaluation, we selected five genes. These genes are involved in DNA repair and replication, LIG1 APEX1, and drug metabolism, CYP2E1, CYP2B6, and PTGIS. The reference sequence for each gene was collected and annotated from the NCBI gene database (https://www.ncbi.nlm.nih.gov/gene/). Established relationships between SNPs in certain genes and drug treatments were explored using drug label annotations and clinical annotations from PharmGKB (https://www.pharmgkb.org/) as well as GeneCards (https://www.genecards.org/). Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines for gene-drug relationships were noted if available [6]. LIG1 was selected for further study due to the number of DNA repair and replication mechanisms it is involved in and its unexplored link to current chemotherapy treatments.

The gene consequence of the SNPs was evaluated to determine the predicted damaging effect of each variation. SNPs which were predicted to be missense mutations and, therefore, change the structure and function of the produced enzyme, were selected for further study. The predicted gene consequence was determined using NCBI SNP database (https://www.ncbi.nlm.nih.gov/snp/), SIFT prediction (https://sift.bii.a-star.edu.sg/www/SIFT_dbSNP.html), and the HumDiv and HumVar models from PolyPhen2 (http://genetics.bwh.harvard.edu/pph2/) (Fig. 1). Within LIG1 there was predicted to be five deleterious mutations: rs3731003, rs4987070, rs4987181, rs11666150, and rs12981963.

In comparison to allele frequencies from Indian subpopulation data, the SNPs of interest are the ones in which the global NCBI frequency is much lower than that of the subpopulations (Fig. 2). These SNPs are likely overlooked by global genome studies and require a more specific study. There were three SNPs within LIG1 which had a higher MAF in the Indian subpopulation as compared to the global population: rs4987070, rs4987181, and rs11666150. Based on the criteria, the SNP list was narrowed down to rs4987181 from LIG1.

Primers were designed and ordered for the SNP using Primer3 (https://bioinfo.ut.ee/primer3-0.4.0/). The section of the gene chosen was ensured to be adequate for testing through PCR and DNA Sanger sequencing. The gene of interest was amplified by PCR in six DNA samples isolated from human cancer (multiple myeloma) cell lines using an optimized annealing temperature of 57.9 °C. The samples used included U266P, JIM3, U266VR, RPMI8226VR, MMISVR, and UTMC2. There was successful amplification of LIG1 in all the samples except UTMC2.

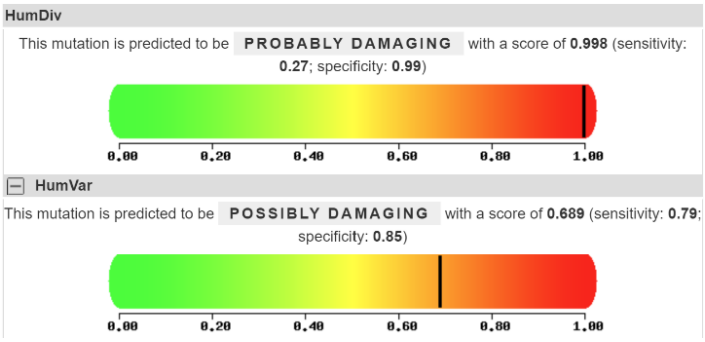

Fig. 1 PolyPhen-2 data for SNP rs4987181 in LIG1.

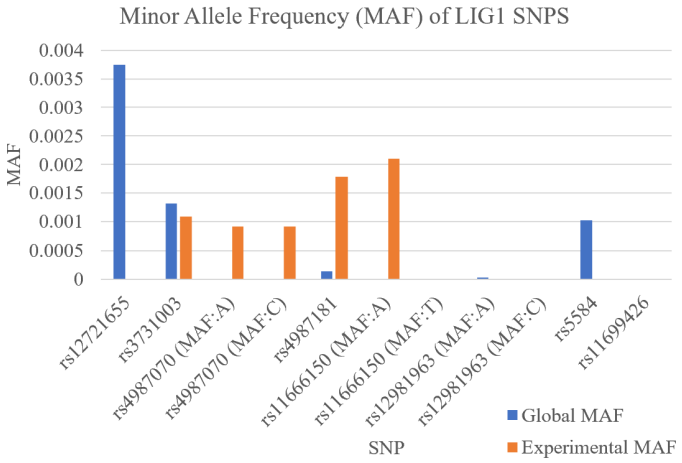

Fig 2. Minor allele frequency of SNPs in LIG1 in global population data (blue) and Indian subpopulation data (orange).

Bidirectional sanger sequencing was performed on the five successful samples using the SeqStudio Genetic Analyzer System with SmartStart (Applied Biosystems, Foster City, USA). Results from sequencing were analyzed using DNASTAR software for sequence alignment (DNASTAR, Madison, WI, USA). The sequence alignment confirmed the primers accurately targeted the region of LIG1 surrounding the SNP of interest. Although, the SNP of interest was not present in any of the multiple myeloma samples tested. This is because all the cancer cell lines belonged to non-Indian ances-

tries (non-Hispanic whites, African Americans, and East Asians), further emphasizing the importance of subpopulation-specific functional analysis. Within the targeted region, two other SNPs were present in several of the samples: rs3730861 and rs3730862. These SNPs are both in intronic regions and have no linked clinical significance. This indicates that they should not interfere with further testing in this region of LIG1.

Throughout global populations, prostate cancer (PC) remains one of the leading causes of cancer related death in males [3,5]. Its growing prevalence in the world necessitates the need for better and more personalized treatments. Recently approved PC therapies have focused on targeting the DNA repair pathways. Therefore, in our SNP evaluation, established prostate cancer cell lines were used for differential gene expression analysis. First, we compared the Castration-resistant prostate cancer cell line DU145 and the clonally derived taxane resistant line DU145-TxR (DUTXR). Thus, DUTXR differs from DU145 in its resistance to taxanes, a type of chemotherapy. Since both cell lines are androgen receptor negative, they will continue to grow in the absence of male hormones [3,5].

Relative gene expression in prostate cancer cell lines were determined using quantitative PCR (qPCR) through the CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). LIG1 was measured in both cell lines in comparison to the housekeeping gene Beta-Actin (ACTB) using the ΔΔCt method (Fig. 3). It was determined that there was a higher relative LIG1 expression in the DU145 cell line compared to the DUTXR cell line. Other genes measured to confirm the AR status of the cell lines were Androgen Receptor (AR), Kallikrein Related Peptidase 3 (KLK3) or Prostate-Specific Antigen, and Mitogen-Activated Protein Kinase 14 (MAPK14). These genes all had similar expression between the two cell lines indicating they were both AR negative (Fig. 3).

In conclusion, the evaluation of the SNPs revealed multiple SNPs that were uniquely significant to India and SNPs that were unique to specific subpopulations, such as rs4987181. The SNP evaluation also revealed that multiple SNPs are uniquely prevalent in these subpopulations and have deleterious effects. This alludes to the high diversity of subpopulations in India and the need for subpopulation-specific studies in relation to genetic variations. Furthermore, we validated the differential expression of LIG1 between taxane-sensitive and taxane-resistant AR negative cell lines. Comparison of established prostate cancer lines suggests that DUTXR could be more affected by a mutation in LIG1. This might be of significance for chemotherapies targeting DNA repair in these aggressive forms of cancer.

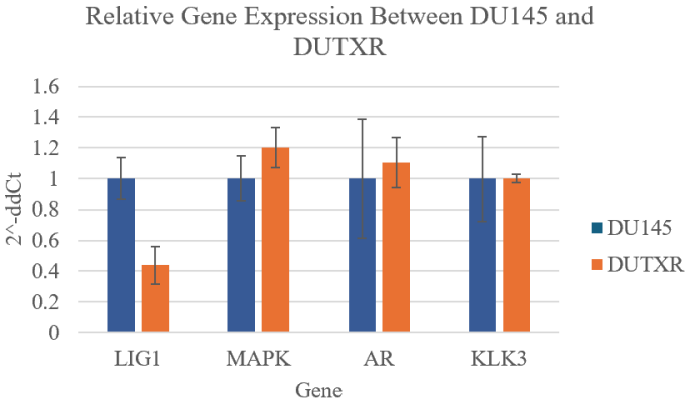

Fig. 3 Relative gene expression of LIG1, MAPK, AR, and KLK3 between DU145 and DUTXR cell lines.

Future work in this project will measure other relative gene levels in other established prostate cancer cell lines. These cell lines include an AR-positive line, LnCAP, and an AR-negative cell line with high metastatic potential, PC3M. The results of the SNP evaluation also motivated the exploration of how rs4987181 affects cancer cell lines. This will be studied using an in vitro model. Therefore, once all cell lines have been assessed for LIG1 levels, the SNP rs4987181 will be introduced to the cell line using site-directed mutagenesis (recombinant NA technologies). A cytotoxicity study for both the wild-type cell line and mutated cell line will be conducted using standard and novel chemotherapeutic agents to investigate the effect of LIG1 SNPs on drug efficacy.

Statement of Research Advisor

Abigail's research involves the functional analysis of the DNA repair gene LIG1 mutations. Her work incorporates bioinformatics and molecular biology approaches, i) to identify novel and reported SNPs in LIG1, ii) to predict the functional relevance of top LIG1 SNPs, iii) to investigate the effect of LIG1 SNPs on differential gene regulation in cancers, iv) to perform functional genomics studies to study the impact of differential LIG1 expression on DNA repair mechanisms vis-à-vis

cancer chemotherapy. We are planning to further this work with a particular focus on world populations and underserved subpopulations.
- Amit K. Mitra, Department of Drug Discovery and Development, Harrison College of Pharmacy

References

[1] Birney, E. The International Human Genome Project. Hum Mol Genet 30, R161-R163 (2021). https://doi.org/10.1093/hmg/ddab198

[2] Indian Genome Variation, C. The Indian Genome Variation database (IGVdb): a project overview. Hum Genet 118, 1-11 (2005). https://doi.org/10.1007/s00439-005-0009-9

[3] Mazumder, S. et al. Integrating Pharmacogenomics Data-Driven Computational Drug Prediction with Single-Cell RNAseq to Demonstrate the Efficacy of a NAMPT Inhibitor against Aggressive, Taxane-Resistant, and Stem-like Cells in Lethal Prostate Cancer. Cancers (Basel) 14 (2022). https://doi.org/10.3390/cancers14236009

[4] Mitra, A. K., Singh, A., Indian Genome Variation, C. & Rath, S. K. LIG1 polymorphisms: the Indian scenario. J Genet 93, 459-469 (2014). https://doi.org/10.1007/s12041-014-0415-7

[5] Mitra Ghosh, T. et al. Metronomic Administration of Topotecan Alone and in Combination with Docetaxel Inhibits Epithelial-mesenchymal Transition in Aggressive Variant Prostate Cancers. Cancer Res Commun 3, 1286-1311 (2023). https://doi.org/10.1158/2767-9764.CRC-22-0427

[6] Relling, M. V. et al. The Clinical Pharmacogenetics Implementation Consortium: 10 Years Later. Clin Pharmacol Ther 107, 171-175 (2020). https://doi.org/10.1002/cpt.1651

[7] Relling, M. V. & Evans, W. E. Pharmacogenomics in the clinic. Nature 526, 343-350 (2015). https://doi.org/10.1038/nature15817

[8] Roden, D. M. et al. Pharmacogenomics. Lancet 394, 521-532 (2019). https://doi.org/10.1016/S0140-6736(19)31276-0

[9] Sallmyr, A., Rashid, I., Bhandari, S. K., Naila, T. & Tomkinson, A. E. Human DNA ligases in replication and repair. DNA Repair (Amst) 93, 102908 (2020). https://doi.org/10.1016/j.dnarep.2020.102908

Authors Biography

Abigail Weir is a senior-year undergraduate pursuing a BS degree in Biochemistry at Auburn University. As an undergraduate research trainee in Dr. Mitra's lab (DDD), she has developed the expertise molecular biology tools (cloning, Sanger DNA sequencing, Real-time qPCR, etc.) to perform mutational analysis in DNA repair genes such as LIG1 and MSH2.

Katherine G. Marlow is a Masters student pursuing an M.S. degree in General Science Education at Auburn University. She holds a B.S. in Genetics from Auburn University. She has played key research roles in understanding pharmacogenomic biomarkers in drug metabolism genes and DNA repair genes.

Razan Waliagha, MS is a current third year dual degree PharmD./PhD. student at Auburn University Harrison College of Pharmacy. Her current work involves understanding the novel mechanisms of resistance to anti-cancer therapies in solid tumors

Dr. Amit Kumar Mitra is an Assistant Professor at the Harrison College of Pharmacy at Auburn University. He is also the founding Director of the Center for Pharmacogenomics and Single-Cell Omics (AU-PharmGx). Dr. Mitra's current research involves integrating in vitro and ex vivo drug response modeling, single-cell multi-omics, and functional genomics with translational bioinformatics to investigate inter-tumor and intra-tumor (subclonal) heterogeneity in human cancers.