2024 SICASH
SLIIT INTERNATIONAL CONFERENCE ON ADVANCEMENTS IN SCIENCES & HUMANITIES

In Silico Analysis of the Diversity of DPYD Gene Variants Affecting Fluoropyrimidine Toxicity: A Comparison of South Asians with Other World Populations

Nirmalee Perera*[1], Manura Weerasinghe[2], Jagath Kasturiarachchi[3], Priyanga Ranasinghe[4]
[1,2,3] *Institute of Information Technology, Sri Lanka*
[4]*University of Colombo, Sri Lanka*
Email address of the corresponding author - *nirmaleeshayamane97@gmail.com

Abstract

Fluoropyrimidine (FP) chemotherapy drug is utilized to treat colon, head, neck and breast cancers. Apart from its effectiveness, toxicity is a limitation. DPD (dihydropyrimidine dehydrogenase) enzyme, which aids in the FP metabolism is produced by the highly polymorphic *DPYD* gene. Mutations in the *DPYD* gene cause the deficiency or non-functionality of the DPD enzyme which varies among different populations. This research aimed to compare allele frequencies of common *DPYD* gene variants of South Asians (SAS) such as *DPYD*2A(rs3918290), DPYD*9(rs1801265), DPYD*5, rs2297595, DPYD*6, rs17376848, rs56038477, DPYD*4(rs1801158), rs67376798* and *rs75017182* with Africans (AFR), Amish (AMI), Latin Americans (AMR), Ashkenazi Jewish (ASJ), East Asians (EAS), Finnish (FIN) and Non-Finnish (NFE). Allele frequencies were obtained from the Genome Aggregation Database in the PharmGKB database. X^2 analysis was performed. $p<0.05$ was deemed to be statistically significant. The study found a significant difference between the SAS population and AFR, AMR, ASJ, EAS, FIN and NFE populations for the *DPYD*9A* gene variant, except for the AMI population. The distribution of the *DPYD*2A* gene variant of SAS was found to be significant in the AFR, ASJ, FIN and NFE populations, except for AMR and AMI. The prevalence of *DPYD*5, DPYD*6, rs17376848*, and *rs56038477* in the SAS significantly differed from all above-mentioned populations. The distribution of the *rs75017182* gene variant in SAS has shown significant differences with AFR, AMR, ASJ and EAS except for NFE and FIN. This study highlights the variations in pharmacogenomics data specific to populations that could lead to personalized medicine and the need for *DPYD* genotyping before cancer treatment, especially in SAS communities where clinically significant genetic variations and haplotypes occur. Study findings pinpoint the potential contribution of *DPYD* gene variations to individual variability in anti-cancer dosage requirements among SAS.

Keywords: Pharmacogenomics; Personalized medicine; Fluoropyrimidine; Fluoropyrimidine toxicity; South Asians

Introduction

FP is a chemotherapy drug discovered in the 1950s for cancer treatments. It disrupts the production and function of DNA and RNA by preventing cell division and cell death (Lamont & Schilsky, 1999). However, FP toxicity remains a significant concern in prescribing FP chemotherapy drugs, including 5-Fluorouracil and Capecitabine. Its efficacy is limited by chemotherapy resistance and causes various side effects including skin toxicity, mucositis, fatigue, hand-foot syndrome, multiorgan failure, diarrhea and

https://doi.org/10.54389/EHCV8962

myelosuppression. In some cases, side effects can be fatal. However, the percentage of patients who die due to FP-related toxicity is less than 1% (Lunenburg et al., 2020). *DPYD* gene has been the major focus of research on FP toxicity, while *TYMS* and *MTHFR* have been investigated for efficacy (Amirfallah et al., 2018). *DPYD* gene is located on chromosome 1p21 with 23 exons, it produces the DPD enzyme which metabolizes FPs. Several *DPYD* variations were found to be linked with FP toxicity (Hishinuma et al., 2020). Notably, low or insufficient DPD enzyme activity varies significantly across the population, with at least 3-5% of individuals (Amstutz et al., 2011). This variation in FP toxicity cases can be identified due to the differences in the distribution of SNPs linked to FP toxicity. This poses a significant challenge for healthcare professionals globally, in the realm of personalized medicine, when it comes to prescribing medication for cancer patients who experience FP toxicity.

For instance, *DPYD *2A (rs3918290), rs67376798, HapB3, rs56038477, rs75017182*, and *DPYD*13 (rs55886062)* gene variants have been found to cause issues in Caucasians (Henricks et al., 2018). These polymorphisms can predict FP toxicity among Caucasian carriers, but their impact on non-Caucasian carriers varies greatly. For instance, the rs5588602 gene variant is not present in SAS but in 0.2% of European Caucasians.

The current literature in Sri Lanka and the global context does not represent a comparison of *DPYD* gene variant frequencies of SAS with world populations, which could limit the understanding of *DPYD* variation worldwide. Therefore, this study's findings would unveil the potential impact of global *DPYD* variation on anti-cancer treatments. This understanding may guide global healthcare systems to ensure medication safety by implementing personalized medicine to improve therapeutic safety and efficacy for every individual using FP drugs in the world including SAS.

In this research the frequencies of most commonly prevalent *DPYD* gene variants in SAS such as *DPYD*2A (rs3918290), DPYD*9(rs1801265), DPYD*5, rs2297595, DPYD*6, rs17376848, rs56038477, DPYD*4(rs1801158), rs67376798* and *rs75017182* (White et al., 2021) (Maekawa et al., 2007) (Hariprakash et al., 2018) were compared with world populations such as AFR, AMI, AMR, ASJ, EAS, FIN and NFE. The allele frequencies were obtained from the Genome Aggregation Database in the PharmGKB database. The PharmGKB is a centralized location for Pharmacogenomic data used by medical professionals, which contains genetic data from various sources managed by the National Institute of General Medical Sciences of the National Institutes of Health and the Pharmacogenomics Research Network.

Methodology

The most common *DPYD* gene variants involved in FP toxicity among SAS and their reference SNP cluster IDs were identified after a thorough literature review (White et al., 2021) (Naushad et al., 2021) (Hariprakash et al., 2018). Allele frequencies, sample sizes, the gene's wild-type number and the mutation numbers of *rs3918290(DPYD*2A), rs1801265 (DPYD*9A), rs1801159 (DPYD*5), rs2297595, rs1801160 (DPYD*6), rs17376848, rs56038477, rs1801158(DPYD*4), rs67376798* in SAS, AFR, AMR, EAS, FIN, NFE, ASJ and AMI were retrieved from the Genome Aggregation Database in PharmGKB Database. Allele frequencies of SAS and other world populations (positive and negative SNP values) were compared using χ^2 test of independence to determine an association between the gene variant frequencies of SAS with other world populations. 70 χ^2 analyses were performed. The significance of each finding was calculated using the *p*-value. $p < 0.05$ was considered significant. The null hypothesis of this study suggested that the distribution of *DPYD* gene variants in SAS is not significantly different from other world populations, whereas the alternative hypothesis suggested that the distribution of *DPYD* gene variants in SAS is significantly different from other world populations.

Results

Table 1 consists of extracted allele frequencies of *DPYD* gene variants in different populations (SAS, AFR, AMI, AMR, ASJ, EAS, FIN, and NFE) from the Genome Aggregation Database in the PharmGKB Database. Table 2 shows the *p*-values of the world population compared to SAS. ***DPYD*5, rs17376848, DPYD*6, rs56038477***: Obtained *p*-values less than 0.05 for all populations compared to SAS.

rs67376798: Obtained *p*-values higher than 0.05 for all populations compared to SAS. ***DPYD*2A***: AFR, ASJ, FIN, AMI and NFE populations, have shown *p*-values less than 0.05 when compared with SAS. AMR and AMI have shown *p*-values greater than 0.05.

DPYD*9A: Demonstrated *p*-values less than 0.05 when compared SAS with AFR, AMR, ASJ, EAS, FIN, and NFE populations and a *p*-value greater than 0.05 was obtained for the AMI population compared to SAS.

***rs2297595*:** Demonstrated *p*-values less than 0.05 for AFR, AMI, ASJ, EAS, FIN and NFE when compared with SAS. A *p*-value greater than 0.05 was demonstrated when the AMR population was compared with SAS.

Table 1. Extracted allele frequencies of DPYD gene variants in different populations from the PharmGKB Database (Genome Aggregation Database)

	Allele frequency							
Gene Variants	SAS	AFR	AMI	AMR	ASJ	EAS	FIN	NFE
*rs3918290/DPYD*2A*	0.27%	0.06%	0.00%	0.17%	0.69%	0.00%	2.43%	0.27%
*rs1801265/DPYD*9A*	24.76%	40.32%	22.04%	24.74%	11.07%	7.22%	28.23%	22.54%
*rs1801159/DPYD*5*	9.61%	15.82%	13.30%	23.69%	19.43%	25.52%	28.23%	22.52%
rs2297595	7.34%	3.36%	12.20%	5.66%	8.42%	1.61%	17.57%	10.11%
*rs1801160/DPYD*6*	9.17%	2.45%	1.75%	4.60%	10.68%	1.54%	2.21%	4.52%
rs17376848	3.91%	2.16%	12.61%	7.29%	1.30%	12.37%	4.83%	4.52%
rs56038477	1.61%	0.30%	2.85%	0.73%	0.55%	0.06%	1.42%	2.14%
*rs1801158/DPYD*4*	0.70%	0.40%	0.33%	1.53%	3.31%	0.02%	1.52%	1.96%
rs67376798	0.06%	0.12%	0.00%	0.17%	0.06%	0.00%	0.01%	0.60%
rs75017182	1.62%	0.25%	2.86%	0.75%	0.55%	0.06%	0.01%	0.60%

Table 2. p values of chi-square analysis of DPYD gene variants

	p-values in comparison with SAS						
Gene Variants	AFR	AMI	AMR	ASJ	EAS	FIN	NFE
*rs3918290/DPYD*2A*	<0.001	0.1168	0.1742	0.0044	0.0002	<0.001	0.0257
*rs1801265/DPYD*9A*	<0.001	0.0789	<0.001	<0.001	<0.001	<0.001	0.0004
*rs1801159/DPYD*5*	<0.001	0.0007	<0.001	<0.001	<0.001	<0.001	<0.001
rs2297595	<0.001	<0.001	0.1034	<0.001	<0.001	<0.001	<0.001
*rs1801160/DPYD*6*	<0.001	<0.001	<0.001	0.0224	<0.001	<0.001	<0.001
rs17376848	<0.001	<0.001	<0.001	<0.001	<0.001	0.0113	0.0485
rs56038477	<0.001	0.0101	<0.001	<0.001	<0.001	0.0013	0.0139
*rs1801158/DPYD*4*	0.3154	0.1938	0.0333	<0.001	<0.001	<0.001	0.1103
rs67376798	0.7219	0.4516	0.3963	0.9342	0.2123	0.059	0.2121
rs75017182	<0.001	0.0107	0.0079	<0.001	<0.001	0.3344	0.5331

DPYD*4 (rs1801158): AMR, ASJ, EAS and FIN have demonstrated *p*-values less than 0.05 when compared with SAS while AFR, NFE and AMI individuals obtained greater *p*-values (>0.05). ***rs75017182:*** *p*-values less than 0.05 were obtained for AFR, AMR, ASJ and EAS compared to SAS. A *p*-value greater than 0.05 were obtained for NFE, and FIN compared to SAS individuals.

Discussion

This study has examined the distribution of common *DPYD* gene variants in SAS compared to other world populations. We hypothesized that the distribution of *DPYD* variants in SAS populations would differ from that of other groups worldwide. According to the findings, the frequencies of *DPYD*5, rs17376848, DPYD*6, and rs56038477* gene variants in SAS are completely different from other world populations (Table 2). However, the frequency of *rs67376798* in SAS shows a similar pattern with all other world populations (Table 3). Based on the study's findings, it was observed that the *p*-values of *DPYD*2A* for AFR, ASJ, FIN, AMI and NFE populations, compared with SAS, were less than 0.05. This implies a noteworthy difference in the distribution of the *DPYD*2A* gene variant between SAS and the above-mentioned populations. While, AMR and AMI have shown *p*-values greater than 0.05, which implies that there is no significant difference between the *DPYD*2A* allele distribution in AMR and AMI compared to SAS. Moreover, a significant difference was found between the SAS population and AFR, AMR, ASJ, EAS, FIN, and NFE populations for *DPYD*9A* (See Table 2) suggesting a notable variation in the frequency of the *DPYD*9A* gene variant between SAS individuals and the populations mentioned above. However, there was no significant difference in the distribution of *DPYD*9A* in the AMI population compared to SAS (Table 2). Furthermore, the study found that the frequency of the *rs2297595* gene variant in SAS differs significantly from populations like AFR, AMI, ASJ, EAS, FIN, and NFE (This difference was established by obtaining a *p*-value of less than 0.05) indicating a notable variation in the frequency of the *rs2297595*

gene variant, whereas, no significant difference was found in between SAS with AMR population for the *rs2297595*. A significant difference was found in the frequency of the *DPYD*4 (rs1801158)* gene variant between SAS individuals and in AMR, ASJ, EAS and FIN individuals. However, no significant difference was found in the distribution of *DPYD*4 (rs1801158*) in AFR, NFE and AMI individuals compared to SAS.

Furthermore, according to the analysis, the distribution of allelic frequency of the *rs75017182* gene variant across populations such as AFR, AMR, ASJ and EAS individuals was found to be significantly distinct (*p*-values for all these populations were calculated to be less than 0.05). Nonetheless, no significant difference was observed in the distribution of *rs75017182* in NFE, FIN compared to SAS individuals.

In summary, considering the gene variants and their frequencies distributed in SAS would likely establish a relationship with other world populations. From the perspective of pharmacogenomics, if the *p*-value is lower than 0.05, the p-value is statistically significant. In other words, a significant difference can be found in the distribution of *DPYD* gene variants in SAS with other populations. Therefore, the number of hypersensitive reaction cases that might occur upon the administration of FP among cancer patients in SAS would vary compared to other world populations. Vice versa, if the *p*-value is higher than 0.05, the *p*-value is not statistically significant. In other words, no significant difference can be found in the distribution of gene variants in other populations compared to SAS. Therefore, the number of hypersensitive reaction cases that might occur upon the administration of FP in cancer patients would probably be similar.

However, factors such as environmental effects, other gene variants that affect gene expression and disease heterogeneity should be further considered to determine the likelihood of hypersensitivity occurrence between two populations. Gender, age, kidney functionality and body composition are linked to FP toxicity within populations (Knikman et al., 2021). Therefore, further studies should be to study the association of the above-mentioned factors with the distribution of *DPYD* gene variants across world populations. Moreover, some studies have demonstrated that the prevalence of functional variations may vary significantly owing to ethnicity. Gene variations might be due to natural selection, gene flow, mutations, genetic drift, and environmental effects. According to Farinango et al., in 2022, populations of European descent have a lower prevalence of *DPYD* variations (around 3-5% with partial insufficiency and 0.02% with total deficiency). SAS, AFRs and some Middle Eastern ethnicities have a greater incidence of *DPYD* variations (Farinango et al., 2022).

In contrast to other populations worldwide, the prevalence of variations in the *DPYD* gene among SAS provides fascinating insights into genetic diversity and potential clinical implications. This study has shown that there are both similarities and differences in the frequency of *DPYD* variants across the world population compared to SAS. Past literature also has shown that populations exhibit distinct patterns of *DPYD* gene variations compared to other populations. Moreover, in Caucasian populations, the most prevalent *DPYD* polymorphism, such as *DPYD*2A*, is less prevalent than in other ethnic groups and has a similar prevalence in EAS populations as Caucasians (Farinango et al., 2022). For instance, *DPYD*2A (rs3918290)* allele frequency is very low or zero in SAS and Japanese populations (0.05% and 0%) however it is slightly higher in American-Caucasian (2.5%) and European-Caucasians (1.5%) (Farinango et al., 2022). Other variations with potential functional significance, such as *DPYD*4A*, may have higher frequencies in EAS populations.

Furthermore, the understanding of common *DPYD* gene variations in SAS compared to other populations globally is important in the era of personalized medicine. Despite regional differences, some *DPYD* gene variants may exhibit consistent frequencies across diverse populations. Therefore, the “one fits

all" theory should be eliminated when prescribing chemotherapy drugs to patients who might have mutations in their genes. Although clinicians' current understandings of personalized medicine and pharmacogenomics may not apply to all ethnicities due to incomplete genotyping of relevant variants in Laboratories worldwide. For instance, non-western countries focus only on specific *DPYD* variants, such as the *DPYD*2A* allele, which is the most widely detected and tested mutation in commercial genotyping platforms. Therefore, *DPYD* sequencing as a screening method for identifying patients at considerable risk of toxicity must be developed to personalize the FP chemotherapeutic medicines. This would lead to reduce the side effects and the burden due to FP toxicity among cancer patients which would save time and money. Moreover, primary research exploring the distinctions of *DPYD* gene variants in diverse populations is essential for advancing precision medicine initiatives and optimizing therapeutic approaches based on genetic diversity and individual patient profiles. The major drawback of this study is the use of secondary data from the PharmGKB database. The secondary data accuracy level may depend on the data collection processes and the quality of the primary research. In contrast, by identifying differences and similarities in the frequency of *DPYD* variants, researchers can design specific drugs that target specific ethnicities.

Conclusion

This study thoroughly examines the genetic diversity in the DPYD gene in several ethnicities compared to SAS. The results indicate notable disparities in the occurrence rates of medically significant DPYD gene variations and genetic patterns across SAS compared to those of European, African, and EAS descent. A significant difference was found in DPYD*5, rs17376848, DPYD*6(rs1801160), and rs56038477 gene variants between SAS individuals and other populations. rs67376798 in SAS shows a probable similar pattern with other world populations. Moreover, this study impacts tailoring chemotherapy dosages and reducing the risk of toxicity in SAS patients, considering their particular DPYD genotypes. In brief, this study emphasizes the significance of having pharmacogenomics data tailored to different populations to guide personalized medicine approaches which would highlight the potential advantages of DPYD genotyping before treatment, especially in SAS communities where clinically relevant genetic variations and haplotypes are prevalent.

References

Amirfallah, A., Kocal, G. C., Unal, O. U., Ellidokuz, H., Oztop, I., & Basbinar, Y. (2018). *DPYD*, TYMS and MTHFR genes polymorphism frequencies in a series of Turkish colorectal cancer patients. *Journal of Personalized Medicine*, *8*(4). https://doi.org/10.3390/jpm8040045

Amstutz, U., Froehlich, T. K., & Largiadr, C. R. (2011). Dihydropyrimidine dehydrogenase gene as a major predictor of severe 5-fluorouracil toxicity. In *Pharmacogenomics* (Vol. 12, Issue 9, pp. 1321–1336). https://doi.org/10.2217/pgs.11.72

Farinango, C., Gallardo-Cóndor, J., Freire-Paspuel, B., Flores-Espinoza, R., Jaramillo-Koupermann, G., López-Cortés, A., Burgos, G., Tejera, E., & Cabrera-Andrade, A. (2022). Genetic Variations of the *DPYD* Gene and Its Relationship with Ancestry Proportions in Different Ecuadorian Trihybrid Populations. *Journal of Personalized Medicine*, *12*(6). https://doi.org/10.3390/jpm12060950

Hariprakash, J. M., Vellarikkal, S. K., Keechilat, P., Verma, A., Jayarajan, R., Dixit, V., Ravi, R., Senthivel, V., Kumar, A., Sehgal, P., Aradhana, A., Sonakar, A. K., Ambawat, S., Giri, A. K., Philip, A., Sivadas, A., Faruq, M.,Bharadwaj, D., Sivasubbu, S., & Scaria, V. (2018). Pharmacogenetic landscape of *DPYD* variants in South Asian populations by integration of genome-scale data. *Pharmacogenomics*, *19*(3), 227–241. https://doi.org/10.2217/pgs-2017-0101

Henricks, L. M., Lunenburg, C. A. T. C., de Man, F. M., Meulendijks, D., Frederix, G. W. J., Kienhuis, E., Creemers, G. J., Baars, A., Dezentjé, V. O., Imholz, A. L. T., Jeurissen, F. J. F., Portielje, J. E. A., Jansen, R. L. H., Hamberg, P., ten Tije, A. J., Droogendijk, H. J., Koopman, M., Nieboer, P., van de Poel, M. H. W., … Schellens, J. H. M.(2018). *DPYD* genotype-guided dose individualisation of fluoropyrimidine therapy in patients with cancer: a prospective safety analysis. *The Lancet Oncology*, *19*(11). https://doi.org/10.1016/S1470-2045(18)30686-7

Hishinuma, E., Narita, Y., Saito, S., Maekawa, M., Akai, F., Nakanishi, Y., Yasuda, J., Nagasaki, M., Yamamoto, M., Yamaguchi, H., Mano, N., Hirasawa, N., & Hiratsuka, M. (2018). Functional characterization of 21 allelic variants of dihydropyrimidine dehydrogenase identified in 1070 Japanese individuals. Drug Metabolism and Disposition, 46(8). https://doi.org/10.1124/dmd.118.081737

Hishinuma, E., Rico, E. G., & Hiratsuka, M. (2020). In vitro assessment of fluoropyrimidine-metabolizing enzymes: Dihydropyrimidine dehydrogenase, dihydropyrimidinase, and β-ureidopropionase. In *Journal of ClinicalMedicine* (Vol. 9, Issue 8). https://doi.org/10.3390/jcm9082342

Knikman, J. E., Gelderblom, H., Beijnen, J. H., Cats, A., Guchelaar, H. J., & Henricks, L. M. (2021). Individualized Dosing of Fluoropyrimidine-Based Chemotherapy to Prevent Severe Fluoropyrimidine-Related Toxicity: What Are the Options? In *Clinical Pharmacology and Therapeutics* (Vol. 109, Issue 3, pp. 591–604). Nature Publishing Group. https://doi.org/10.1002/cpt.2069

Lamont, E. B., & Schilsky, R. L. (1999). The oral fluoropyrimidines in cancer chemotherapy. In *Clinical Cancer Research* (Vol. 5, Issue 9).

Lunenburg, C. A. T. C., van der Wouden, C. H., Nijenhuis, M., Crommentuijn-van Rhenen, M. H., de Boer-Veger, N. J., Buunk, A. M., Houwink, E. J. F., Mulder, H., Rongen, G. A., van Schaik, R. H. N., van der Weide, J., Wilffert, B., Deneer, V. H. M., Swen, J. J., & Guchelaar, H. J. (2020). Dutch Pharmacogenetics Working Group (DPWG) guideline for the gene–drug interaction of *DPYD* and fluoropyrimidines. *European Journal of Human Genetics*, *28*(4). https://doi.org/10.1038/s41431-019-0540-0

Maekawa, K., Saeki, M., Saito, Y., Ozawa, S., Kurose, K., Kaniwa, N., Kawamoto, M., Kamatani, N., Kato, K., Hamaguchi, T., Yamada, Y., Shirao, K., Shimada, Y., Muto, M., Doi, T., Ohtsu, A., Yoshida, T., Matsumura, Y.,Saijo, N., & Sawada, J. I. (2007). Genetic variations and haplotype structures of the *DPYD* gene encoding dihydropyrimidine dehydrogenase in Japanese and their ethnic differences. *Journal of Human Genetics*, *52*(10), 804–819. https://doi.org/10.1007/s10038-007-0186-6

Naushad, S. M., Hussain, T., Alrokayan, S. A., & Kutala, V. K. (2021). Pharmacogenetic profiling of dihydropyrimidine dehydrogenase (*DPYD*) variants in the Indian population. *Journal of Gene Medicine*, *23*(1). https://doi.org/10.1002/jgm.3289

White, C., Scott, R. J., Paul, C., Ziolkowski, A., Mossman, D., & Ackland, S. (2021). Ethnic diversity of DPD activity and the DPYD gene: Review of the literature. *Pharmacogenomics and Personalized Medicine,* 14, 157–169. https://doi.org/10.2147/PGPM.S337147