International Journal of Applied Pharmaceutics

ISSN- 0975-7058 Vol 14, Issue 3, 2022

Original Article

DESIGN AND DEVELOPMENT OF A PHARMACOGENOMIC MODEL FOR BREAST CANCER TO STUDY THE VARIATION IN DRUG ACTION AND SIDE EFFECTS

HIMA VYSHNAVI A. M.[1], P. K. KRISHNAN NAMBOORI[2*]

[1.2*]**Computational Chemistry Group (CCG), Amrita Molecular Modeling and Synthesis Research Lab, Amrita School of Engineering, Coimbatore, Amrita Vishwa Vidyapeetham, India, 641112**
Email: n_krishnan@cb.amrita.edu

Received: 08 Feb 2022, Revised and Accepted: 16 Mar 2022

ABSTRACT

Objective: The proneness of disease, as well as drug action and side effects, vary from person to person. This may be due to individual variations in the genome. The individual variation demands the need to design a population-specific 'predictive, preventive, participatory and personalized (p4)' pharmacogenomics drug molecule. The present work aims at designing a pharmacogenomic model for breast cancer to explain the individual variation in the proneness of the diseases and susceptibility towards drug action.

Methods: The drug action and side effects of drugs were analyzed from clinical trial reports. The genes responsible for the drug action and the genes responsible for side effects have been identified and included in the variation analysis. The pharmacogenomic gene models have been designed by inducing population-specific genetic variations within the gene sequence. The 3D structures of the 'variation-specific' protein models have been generated by 'homology modelling.' These models have been used further for docking studies with the known drug molecules. The kinetic stability of the protein-ligand complexes obtained out of docking studies has been studied by the molecular dynamic simulation.

Results: By the interaction studies and the computational analysis using the 'population-specific protein models,' the drug molecule, *Capecitabine* showed the highest binding affinity (–6.30kcal/mol) with the African population, *Paclitaxel* was found to be the most interacting with the European population with a binding affinity of–9.5603 kcal/mol, and Lapatinib is found to be the most suitable ligand for the American population with a binding affinity of–6.90 kcal/mol. These observations agree with the clinical trial data found in the 'ClinTrial database'.

Conclusion: The designed models are found to be suitable for representing the respective population-specific target models. The interaction studies of known drug molecules with these population-specific target models correspond to the observations in the 'ClinTrial database.'

Keywords: Breast cancer, Drug action, Homology modelling, Pharmacogenomic model, Side effects and variation

DOI: https://dx.doi.org/10.22159/ijap.2022v14i3.44356. Journal homepage: https://innovareacademics.in/journals/index.php/ijap

INTRODUCTION

The recently introduced 'systems biology platforms' provide a new paradigm for integrating multiple omics approaches (e. g., genomics, transcriptomics, proteomics, and metabolomics) to create a more holistic understanding of basic biology to their adaptation, proneness, development, and progression towards disease [1-5]. The relevance of studying the molecular pathways of different diseases, especially degenerative diseases, has been introduced within the omics platforms [6]. The analysis was further extended to 'pharmacogenomics' to include the population-wise variation in drug action and proneness of diseases [7]. Pharmacogenomics cover multi-omics factors such as genomics, epigenomics, metagenomics and environmental genomics, meeting the requirements of individual variations [8, 9]. At present, the need for the design and development of 'pharmacogenomics and personalized drugs' for terminal diseases such as cancer and neurodegenerative diseases have been initiated [10]. As per the reports, cancer is the second leading cause of death next to cardiovascular diseases [11].

Breast cancer is reported as the most common type of cancer found among women [12]. Out of different drugs used for breast cancer, Capecitabine is found to be highly suitable for the South Asian population for controlling breast cancer. The project reports in 'the Clintrial database' identify 78.75% success and 18.33% adverse events while administering Capecitabine to Indian patients [13]. At the same time, for the USA population with the same drug, there were only 60.00% success and 40.00% adverse events [14]. Drug action and side effects vary among different ethnic groups. 'Identification of the genetic markers corresponding to these variations is a significant step in Pharmacogenomics [15]. The 'Single Nucleotide Variation (SNV)' has been identified as the most suitable variant for incorporating individual variation [16-18]. However, the significance of each SNV for any ethnic group depends upon its occurrence, leading to the need for a 'frequency-wise prioritization and classification' of the variants. The functional significance of these nucleotide variants corresponds to changes in the amino acids of their protein molecule [19]. The 3-dimensional structure of a target protein could be generated from its amino acid sequence through' homology modelling' [20]. Pre-dominant binding modes of the ligand molecule within the binding site of the target protein could be predicted through the molecular docking technique [21-23]. Among different platforms, AutoDock and Vina are two of the most widely used molecular docking tools [24]. The thermodynamic stability of the ligand within the target could be studied through molecular dynamic (MD) simulation. NAMD is one such molecular dynamic program designed to perform high-performance simulations of macromolecular systems [25].

In this manuscript, an illustration of the population-wise variation in drug response and side effects of common anti-breast cancer drugs such as Lapatinib, Capecitabine and Paclitaxel has been carried out.

MATERIALS AND METHODS

The clinical trial reports were studied from ClinicalTrials. gov [26]. The drug action and side effects among patients across the five populations, African, American, East Asian, South Asian, and European, have been studied. The drug targets have been identified from the 'DRUG BANK' [25]. Drug side effects have been studied through 'DRUG BANK', clinical trial reports, and 'Kyoto Encyclopaedia of Genes and Genomes (KEGG) pathway' [27, 28]. 'Database of SNP (Single Nucleotide Polymorphisms) (dbSNP)' was used for retrieving gene-specific variations. Population analysis of genetic variations (Single Nucleotide Variants-SNVs) has been carried out based on the 1000 genome population analysis considering the continental population classes, namely 'American (AMR), African (AFR), European (EUR), South Asian (SAS) and East Asian (EAS)' populations using SNPNexus [29]. The SNVs were further classified into two based on their frequency of occurrence

within the population. Mutated gene models were generated by inducing the variations into the wild gene sequence available in 'National Centre for Biotechnology Information (NCBI)' [30]. The mutated protein sequences were generated by incorporating the amino acid variants in the wild protein sequence present in Uniprot [31]. Homology modelling has been carried out using the 'SWISS Model' to generate 3D protein models [32]. The model structures were evaluated using primary analysis QMean, 'RMSD (root mean square deviation),' 'Ramachandran plot,' and 'ERRAT plot' using 'SAVES server' [33-35]. The model structures were further evaluated by the 'molecular dynamic (MD)' simulation using NAMD [25] tool. Screened models were then subjected to interaction analysis using Autodock Vina [36, 37]. The ligand's retention within the target's active site has been studied through MD simulation of the ligand-target complex [38]. The MD simulation was carried out using the ensemble where the 'number of molecules, volume, and temperature (NVT) ensemble' are fixed. The simulation temperature was set at 310K, and the simulation time was 20 ns.

RESULTS

Study of clinical trial reports

The pharmacogenomic projects related to breast cancer comprising variation in drug-likeness and common side effects and included in the Clintrial database have been taken for analysis. The ethnic groups included in these projects are the 'United States (American),' France and Spain (European), South Korea and China (East Asian), India and Sri Lanka (South Asian) and Egypt and Nigeria (African) populations.

The use of Lapatinib, Paclitaxel and Capecitabine among these populations has been included in these reports. Lapatinib is found to be suitable for the AMR population, followed by EAS and SAS population. Capecitabine is suitable for the SAS population, followed by AMR and EUR populations. These drugs impart various side effects such as gastrointestinal disorders, hepatic disorders, hemorrhage, and respiratory disorders. Most of these side effects result from the mutation of the genes SLC25a44, HTR3C, SERPINA14 and ABCA3 (table 2). It has been found that erbb2 and EGFR are targeted by Lapatinib, BCL2 is targeted by Paclitaxel and TYMS by Capecitabine.

Variation analysis

The computational techniques used in pharmacogenomics have been used to explain the above observations. The 'Single Nucleotide Variations (SNVs) are the major fingerprints of individual variation. Hence, the SNVs of the drug target genes and the genes corresponding to the side effects have been collected. These variants were further subjected to '1000 genome population analysis' with American (AMR), African (AFR), East Asian (EAS), European (EUR) and South Asian (SAS) populations. It has been observed that for the ERBB2 gene, there were 458 variants specific to AFR, 294 variants to AMR, 239 variants to EAS, 314 variants to EUR and 257 variants to the SAS population. EGFR (epidermal growth factor receptor) gene has eight variants among AFR, 2066 variants among AMR, 1564 variants among EAS, 1672 variants among EUR, 1783 variants among the SAS population. The gene BCL2 keeps 2809 variants among AFR, 1866 variants among AMR, 1599 variants among EAS, 1535 variants among EUR and 1640 variants among SAS. Similarly, TYMS contains 291 variants among AFR, 167 variants among AMR, 152 variants among EAS, 148 variants among EUR and 157 variants among the SAS population. Based on the frequency of occurrence, these variants were further classified into two classes, those with population frequency below 45% as class 1 and the rest (above 45%) as class 2.

Table 1: Evaluation of drug action models

Model	Template	Z-Score					RMSD (A)
		Qmean	C beta	All atom	Solvation	Torsion	
Capacitabine							
TYMS_AFR_C1	1HW3	-2.63	-5.14	1.16	-1.28	-1.37	2.14
TYMS_AFR_C2	1HW3	-2.83	-0.37	-2.67	-1.52	-2.17	2.37
TYMS_AMR_C1	1HW3	-1.69	-0.39	0.29	-2.35	-0.71	2.09
TYMS_EAS_C1	1HW3	-2.82	-0.31	-2.64	-1.54	-2.17	1.43
TYMS_EAS_C2	1HW3	0.21	0.34	-1.14	-0.98	0.89	1.33
TYMS_EUR_C1	1HW3	-0.70	-1.57	-1.10	-1.02	0.21	2.86
TYMS_EUR_C1	1HW3	-1.54	-2.25	0.87	-1.32	-0.78	1.29
TYMS_SAS_C1	1HW3	-6.81	-1.03	-4.02	-3.24	-5.00	2.33
TYMS_SAS_C2	1HW3	-0.84	-0.05	-0.49	-2.81	0.28	2.86
Lapatinib							
ERBB2_AFR-C1	3MZW	-3.26	-2.07	-1.21	-2.68	-1.65	2.33
ERBB2_AFR-C2	3MZW	-4.47	-2.29	-2.79	-2.11	-3.23	1.39
ERBB2_AMR-C1	3MZW	-3.63	-1.88	-0.87	-2.72	-2.15	3.29
ERBB2_AMR-C2	3MZW	-2.76	-2.06	-0.99	-2.30	-1.20	1.37
ERBB2_EAS-C1	3MZW	-2.78	-1.23	-0.91	-2.42	-1.51	2.02
ERBB2_EUR-C2	3MZW	-4.27	-3.36	-3.17	-2.55	-2.96	2.73
ERBB2_SAS-C1	3MZW	-2.35	-0.47	-0.97	-3.12	-1.06	2.49
ERBB2_SAS-C2	3MZW	-0.59	-0.63	-0.70	-0.92	-0.03	1.19
EGFR_AFR_C1	1XKK	-1.68	-2.92	-0.80	-0.54	-0.98	2.56
EGFR_AFR_C2	1XKK	-1.17	0.96	0.16	-1.24	-1.31	1.62
EGFR_AMR_C1	1XKK	-0.71	-1.13	-0.01	-0.54	-0.31	1.09
EGFR_AMR_C2	1XKK	-3.89	-2.11	-2.06	-1.10	-3.01	2.89
EGFR_EUR_C1	1XKK	-1.18	0.64	0.20	-1.08	-1.30	1.33
EGFR_EUR_C2	1XKK	-2.98	-2.53	-0.91	-2.03	-2.11	2.86
EGFR_SAS_C1	1XKK	-2.59	-2.56	-0.71	-2.56	-1.28	1.29
EGFR_SAS_C2	1XKK	-3.02	-0.33	-1.55	-2.41	-2.68	3.09
EGFR_EAS_C1	1XKK	-1.81	-2.14	-1.27	-1.58	-0.72	2.89
EGFR_EAS_C2	1XKK	-2.60	-2.07	-0.96	-1.34	-2.10	1.43
Paclitaxel							
BCL2_AFR_C1	2WL3	-2.41	-1.08	-2.51	-1.00	-2.28	2.06
BCL2_AFR_C2	2WL3	-1.74	-2.45	-2.14	0.28	-1.70	1.25
BCL2_AMR_C1	2WL3	-2.23	-3.19	-2.34	0.40	-2.26	1.01
BCL2_AMR_C2	2WL3	-2.19	-3.10	-2.25	0.17	-2.06	2.53
BCL2_EUR_C1	2WL3	-2.46	-2.99	-1.17	-1.19	-1.65	2.03
BCL2_EUR_C2	2WL3	-2.40	-3.02	-2.05	0.34	-2.55	1.34
BCL2_SAS_C1	2WL3	-3.12	-2.42	-0.91	-1.17	-2.80	2.09
BCL2_SAS_C2	2WL3	-2.61	-1.63	-1.46	-3.96	-0.81	2.09
BCL2_EAS_C1	2WL3	-2.51	-1.56	-1.25	-3.61	-0.90	2.49
BCL2_EAS_C2	2WL3	-2.72	-2.32	-1.28	-2.33	-1.61	1.15

Table 2: Evaluation of side effect models

Model	Template	Z-Score					RMSD (A)
		Qmean	CB	all atom	solvation	torsion	
Gastrointestinal disorder							
ABCA3_AFR_C1	AF-Q99758-F1	-0.68	-1.74	0.92	-0.49	-0.21	2.33
ABCA3_AFR_C2	AF-Q99758-F1	-3.58	-1.21	-3.16	-2.3	-2.38	1.39
ABCA3_AMR_C1	AF-Q99758-F1	-3.93	-2.58	-0.03	-0.88	-3.11	2.29
ABCA3_AMR_C2	AF-Q99758-F1	-325	-2.99	-1.52	-1.7	-2.43	2.36
ABCA3_EAS_C1	AF-Q99758-F1	-3.79	-0.94	-1.16	-2.11	-3.69	2.74
ABCA3_EAS_C2	AF-Q99758-F1	-2.45	-2.53	-1.48	-2	-1.41	1.79
ABCA3_EUR_C1	AF-Q99758-F1	-1.83	-2.33	-1.13	0.02	-1.63	1.40
ABCA3_EUR_C2	AF-Q99758-F1	-2.88	-1.68	-1.11	-2.89	-1.63	1.37
ABCA3_SAS_C1	AF-Q99758-F1	-2.92	-2.32	-1.4	-0.67	-1.954	2.02
ABCA3_SAS_C2	AF-Q99758-F1	-2.23	-2.99	-2.05	-1.47	-1.08	2.73
Respiratory disorder							
HTR3C_AFR_C2	AF-Q8WXA8-F1	-3.28	-0.79	-1.91	-6.24	-0.69	2.49
HTR3C_AMR_C1	AF-Q8WXA8-F1	-0.96	-0.44	1.21	-1.03	-0.77	1.19
HTR3C_AMR_C2	AF-Q8WXA8-F1	-3.32	-0.95	-0.72	-4.14	-1.45	7.95
HTR3C_EAS_C1	AF-Q8WXA8-F1	-2.93	-0.13	-0.87	-4.22	-1.22	2.90
HTR3C_EAS_C2	AF-Q8WXA8-F1	-2.39	-1.39	0.57	-1.22	-2.35	1.04
HTR3C_EUR_C1	AF-Q8WXA8-F1	-3.62	-0.98	-2.07	-5.18	-1.47	2.96
HTR3C_EUR_C2	AF-Q8WXA8-F1	-3.37	-1.62	0.9	-1.69	-2.92	2.48
HTR3C_SAS_C1	AF-Q8WXA8-F1	-3.71	-1.29	-1.73	-3.32	-2.14	2.49
HTR3C_SAS_C2	AF-Q8WXA8-F1	-2.53	-2.2	1.08	-1.05	-2.43	1.29
Hepatic disorder							
SERPINA1_AFR_C1	2QUG	-1.94	-0.66	-1.58	-2.51	-0.87	2.13
SERPINA1_AFR_C2	2QUG	-2.9	-2.19	-1.12	-2.46	-1.37	2.27
SERPINA1_AMR_C1	2QUG	-2.74	-0.04	-1.66	-2.8	-2.07	2.44
SERPINA1_AMR_C2	2QUG	-3.27	-2.07	1.17	-0.82	-2.66	3.14
SERPINA1_EAS_C1	2QUG	-2.26	-1.89	0.34	-0.61	-1.68	2.08
SERPINA1_EAS_C2	2QUG	-1.89	-1.35	1.1	-1.12	-1.6	1.71
SERPINA1_EUR_C2	2QUG	-2.2	-1.96	0.45	-0.47	-1.66	1.95
SERPINA1_SAS_C1	2QUG	-3.72	-3.28	-2.41	-1.08	-2.87	2.20
SERPINA1_SAS_C2	2QUG	-3.16	-3.73	1.16	-1.68	-1.66	2.13
Hemorrhage							
SLC25A44_AFR_C1	AF-Q96H78-F1	-2.31	-1.51	-2.01	-1.34	-1.8	2.55
SLC25A44_AFR_C2	AF-Q96H78-F1	-0.82	-1.72	-1.31	-1.33	0.29	2.58
SLC25A44_AMR_C1	AF-Q96H78-F1	-3.71	-2.75	-1.02	-1.67	-3.22	2.58
SLC25A44_AMR_C2	AF-Q96H78-F1	-4	-1.82	-1.94	-2.07	-2.68	1.39
SLC25A44_EAS_C1	AF-Q96H78-F1	-3.15	-1.79	-0.74	-1.35	-2.96	3.57
SLC25A44_EAS_C2	AF-Q96H78-F1	-3.83	-1.62	-3.37	-3.07	-2.23	1.22
SLC25A44_EUR_C1	AF-Q96H78-F1	-3.17	-2.99	-1.74	-2.48	-1.88	1.02
SLC25A44_EUR_C2	AF-Q96H78-F1	-2.17	-1.3	-1.58	-1.84	-1.43	1.72
SLC25A44_SAS_C1	AF-Q96H78-F1	-3.03	-2.69	-2.01	-2.07	-1.95	2.76
SLC25A44_SAS_C2	AF-Q96H78-F1	-3.44	-1.14	-1.88	-1.89	-2.37	1.47

Table 3: The binding free energy of ligand-target complexes

Population	Drug action (kcal/mol)	Respiratory disorders (kcal/mol)	Gastrointestinal disorder (kcal/mol)	Hemorrhage (kcal/mol)	Hepatic disorder (kcal/mol)
Lapatinib					
AFR C1	-5.20	-7.10	-3.20	-8.60	-6.60
AFR C2	-5.80	-9.80	-7.90	-10.40	-7.80
AMR C1	-10. 40	-6.70	-8.20	-7.80	-7.20
AMR C2	-11.10	-7.60	-9.80	-10.30	-8.700
EAS C1	-8.30	-7.30	-7.30	-7.70	-9.60
EAS C2	-8.60	-7.90	-8.00	-8.10	-7.60
EUR C1	-9.10	-7.90	-7.90	-7.70	-2.10
EUR C2	-8.60	-9.50	-9.50	-8.50	-9.90
SAS C1	-6.20	-8.80	-8.80	-7.30	-9.10
SAS C2	-6.50	-8.60	-8.60	-8.60	-8.70
Capecitabine					
AFR C1	-5.20	-7.10	-2.30	-8.60	-6.60
AFR C2	-5.80	-9.80	-7.90	-10.40	-7.80
AMR C1	-10. 40	-6.70	-8.20	-7.80	-7.20
AMR C2	-11.10	-7.60	-9.80	-10.30	-8.70
EAS C1	-8.30	-7.30	-7.30	-7.70	-9.60
EAS C2	-8.60	-7.90	-8.00	-8.10	-7.60
EUR C1	-9.10	-7.90	-7.90	-7.70	-2.36
EUR C2	-8.60	-9.50	-9.50	-8.50	-9.90
SAS C1	-6.20	-8.80	-8.80	-7.30	-9.10
SAS C2	-6.50	-8.60	-8.60	-8.60	-8.70

Population	Drug action (kcal/mol)	Respiratory disorders (kcal/mol)	Gastrointestinal disorder (kcal/mol)	Hemorrhage (kcal/mol)	Hepatic disorder (kcal/mol)
Paclitaxel					
AFR C1	-7.50	-9.10	-1.23	-9.10	-7.00
AFR C2	-7.80	-10.40	-9.20	-11.30	-8.20
AMR C1	-6.50	-7.40	-8.30	-8.10	-7.30
AMR C2	-7.30	-7.90	-12.80	-10.50	-9.70
EAS C1	-7.20	-7.60	-7.60	-8.00	-10.80
EAS C2	-8.30	-8.10	-8.50	-9.50	-7.90
EUR C1	-9.10	-8.30	-8.30	-8.00	-2.45
EUR C2	-10.02	-11.30	-11.30	-9.90	-10.40
SAS C1	-8.70	-10.10	-10.90	-11.10	-10.40
SAS C2	-9.10	-9.00	-9.40	-9.30	-9.50

Generating mutated sequence model

The variants belonging to different population classes have been included in the wild gene sequence to generate the mutated gene sequence. The amino acid variants corresponding to the 'SNVs' have been incorporated into the wild protein sequence. The 3D structures of the mutated protein molecules have been designed by 'homology modelling'.

Homology modelling

The 3D structure of proteins with an identity score>50 % with the mutated protein sequences has been identified and considered a template for the study. There were 50 templates for mutated BCL2 protein sequences, 15 templates for mutated TYMS protein sequences, 60 templates for mutated ERBB2 protein sequences and 30 templates for mutated EGFR protein sequences within drug action models. There were 12 templates for a respiratory disorder, 30 templates for gastrointestinal disorders, and 15 templates for hepatic disorder mutated protein sequences among side effect models. Models have been evaluated using the parameters GQME (Global Model Quality Estimation), QMean (the degree of nativeness), Cbeta, All-atom, Solvation, Torsion angle potential, RMSD, Ramachandran plot and ERRAT plot. The evaluation results of the drug action model are shown in table 1**,** and the side effect model is shown in table 2.

Interaction study

The binding site of the targets have been identified as; 3MZW-Ser783, Gln799, Thr798, Thr862, Lys753, Arg784, Asp863, Val734, Ala751, Ile752, Met774, Leu796, Val797, Leu852, Leu785, Leu800, Met801, Pro802, Tyr803, and Phe864; for 1xkk-Thr790, Gln791, Thr854, Lys45, Arg776, Asp855, Ala743, Ile744, Met766, Cys775, Leu777, Leu788, Ile789, Leu792, Met793, Leu844, Phe856, Leu858, Leu718, Val726, Cys797, Gly719, Ser720, Thr790, and Met766; 1HW3-Arg78, Val79, Phe80, Ile108, Leu221, Gly222, Phe225 and Met309; and for BCL2-TYR7, ASP8, ANN9, ARG10, GLU11, ILE12, VAL13, MET14, LYS15, TYR16, ILE17, HIS18, TYR19, LYS20, LEU21, SER22, GLN23, ARG24. The docking score values of the model target protein molecules with the drug molecules and the target genes corresponding to the selected side effects have been included in table 3. The interacting residues are shown in Supplementary materials 1 and 2. The binding affinity of the ligand-target complexes and their interaction are included in Supplementary material 3.

The ligand-target complex of Paclitaxel drug action and its side effects among SAS has been included in fig. 1. The ligand-target complexes of Lapatinib and Capecitabine for SAS have been included in Supplementary material 4.

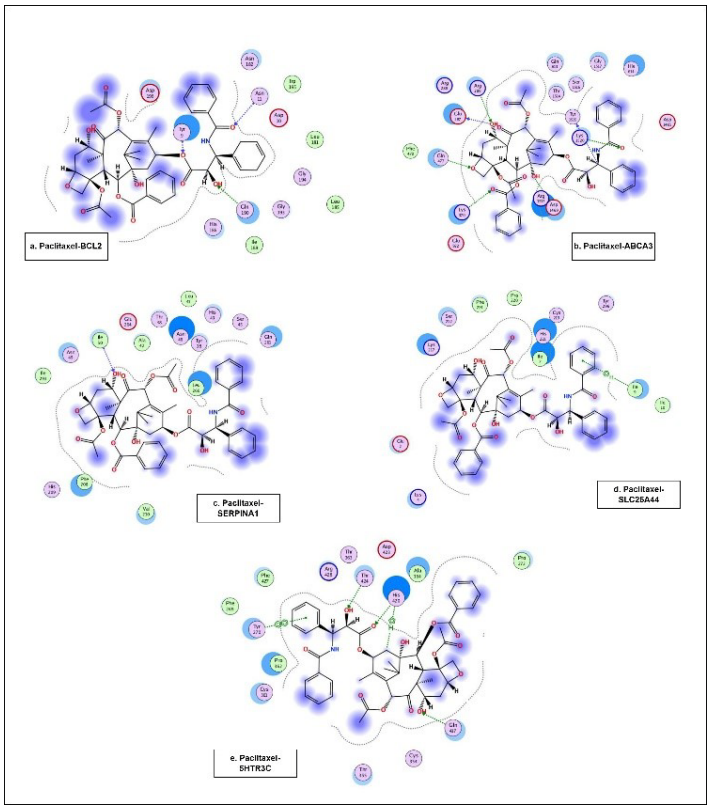

Fig. 1: Ligand-target complex of paclitaxel

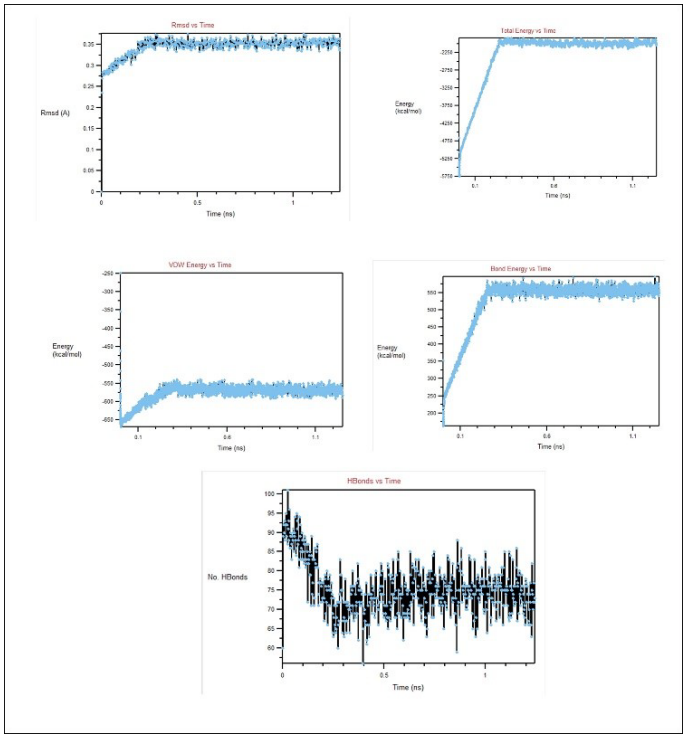

Fig. 2: The plots of MD simulation

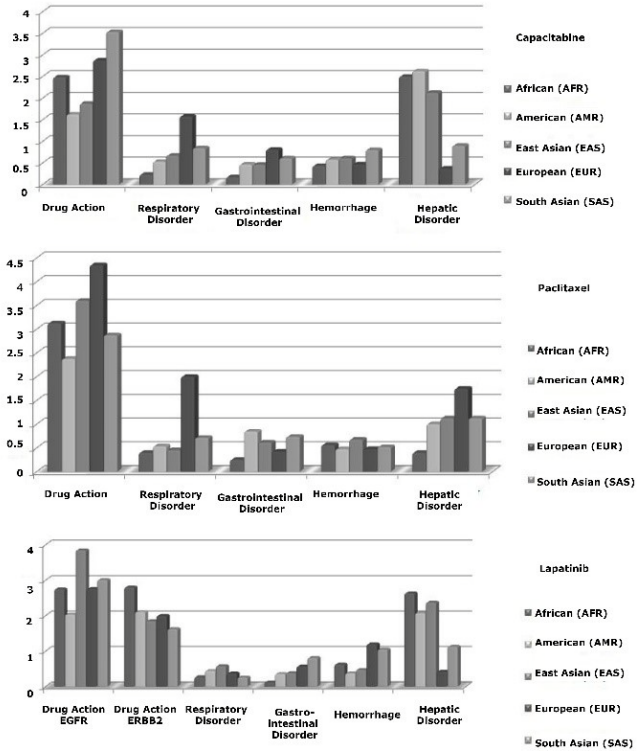

Fig. 3: The RMSD value of the ligand-target complex of different mutant gene models across the population class

Molecular dynamic simulation

The molecular dynamic simulations of the ligand-target complex have been carried out, and plots of MD simulation carried out for the Paclitaxel-SAS drug action model is shown in fig. 2. The RMSD plots have been generated in fig. 3.

The interacting residues after MD simulation are shown in Supplementary material 5.

DISCUSSION

The clinical trial study has been made based on the reports obtained in the clinical trial database with different ethnic groups. Based on the clinical trial analysis, the variation of drug action and the side effect of these drugs have been made. The anti-breast cancer drug molecules, Lapatinib, Capecitabine, and Paclitaxel, have been considered for the illustrative study based on the trials with reported results across the maximum number of populations. It has

been observed that there is the highest success rate, minimum death rate and minimal side effects for Lapatinib in the AMR population, followed by EAS and SAS populations. Even though the success rate of Capecitabine is high within SAS, there was a 2.91% death rate, and 18.33 % reported adverse events. The drug was successful among 60.00% of the patient group within AMR and with no death rate. But about40.00% of patients showed adverse events. This was followed by the EUR population with a success rate of 50.53%; 40.92% of patients were shown side effects and 0.54 % of death during the study.

Paclitaxel was found to be the most suitable drug for SAS, with 85.38 % of success, 14.61 % of adverse events and 0.00 % of death among the patient group followed by AFR. 60.80 % of patients showed positive drug action, 36.65 % developed side effects, and 2.54 % of them died during the trial period among the African population and could be considered least suitable for the EUR population. 44.68% of people responded positively to Paclitaxel, 25.53% of them had side effects, and 29.78 % of the patients died during the analysis. The reported targets of Capecitabine, Lapatinib and Paclitaxel have been selected for the' pharmacogenomic computational prediction strategy'. The gene TYSM has been identified as the targeted mutation for Capecitabine, and the respective protein 1HW3 has been considered as the target protein. Similarly, for Lapatinib, EGFR and ERBB2 have been identified as the mutations and 1XKK, and 3MZW have been included as the protein targets. For Paclitaxel, the identified mutation is BCL2, and the target protein is 2WL3. Regarding the side effects, gastrointestinal disorder (GID), respiratory disorder (RD), hepatic failure (HF) and haemorrhage have been selected for the study. Their respective mutations are ABCA3, HTR3C, SERPINA1, and SLC25A44 [39-42] and the corresponding protein targets are found to be AF-Q99758-F1, AF-Q8WXA8-F1, 2QUG, and AF-Q96H78-F1. The population-wise variants for the genes have been incorporated, and the respective 3D protein target models have been designed by homology modelling. The models with an alignment score greater than 50 %, RMSD of the template and the model<3A, Qmean value<-6.00, models with>80% of residues in most favoured regions, models>10 %<20 % of residues in additional allowed regions, models<10 % of residues in the generously allowed region and 0.00 % of residues in disallowed regions and models with<10 % of residues showing error have been screened. The evaluations parameters and the model properties have been included in table 1 and table 2.

The interaction study showed that Lapatinib interacted with the entire population model but was found to be more effective for ERBB2 models of AMR and EUR populations followed by EAS, SAS and AFR populations. Capecitabine showed good interaction with AFR and AMR followed by EAS and SAS populations and was least effective for the EUR population. Though Paclitaxel had a strong interaction with all populations, there was a slight variation in the binding affinity due to a difference in interaction type and bond length. The drug was found to be most interacting with EUR and SAS, followed by EAS, AFR and AMR populations.

Lapatinib had less interaction with the model for a respiratory disorder (RD) within the EUR population, whereas it showed good interaction with other population models. For the EUR population model, Lapatinib showed electrostatic interaction. Paclitaxel was found to be interacting with all the population models. The complex had 'electrostatic interaction,' 'hydrogen bonds interaction', as well as 'hydrophobic interactions with all the models.' While comparing the drug action with different population models, Paclitaxel had the least interaction for the EAS population. It had only a hydrogen bond interaction with the LEU1578 and S atom of the drug. The Capecitabine was found to be interacting with all the models. The drug is found to be the least interacting with AMR and EUR population models. Thus, it has been found that Lapatinib will have respiratory side effects for the European population due to the absence of sufficient hydrogen bond interaction within the ligand-target complex and the presence of weak interactions. Even though Paclitaxel and Capecitabine showed good interaction with all the population models, there are chances for the patients to develop respiratory disorders within the East Asian population for Paclitaxel, within the American and European population for Capecitabine.

All the three-drug molecules had good interaction with GID targets of African and American population models. In contrast, it was not found to be interacting with East Asian, South Asian, and European models, suggesting that patients from EAS, SAS and AFR populations develop gastrointestinal side effects on the administration of Capecitabine, Lapatinib, and Paclitaxel.

Lapatinib and Capecitabine showed good interaction with Hepatic Failure (HF) models across the population. Paclitaxel was not found to be interacting with models of EUR and AMR population, whereas it interacts well with other population models. This suggests that Lapatinib and Capecitabine do not have hepatic side effects. Though Paclitaxel does not have hepatic side effects among AFR, EAS and AMR populations, breast cancer patients from SAS and EUR populations subjected to Paclitaxel treatment show hepatic disorders as a side effect.

The ligand-target complexes were further evaluated for the retention of the ligand molecule within the active site of the target (kinetic stability) [43, 44]. The RMSD, bond energy, van der Waals energy, and total energy plots of the 'Paclitaxel-SAS drug action model complex' tend to converge after 1.5ns. The RMSD limit for all the experiments was found to be below 3.5 A. The hydrogen bonds tend to break after 5ns. It was found that RMSD of drug action models and Lapatinib converge between 0.2 to 0.3 ns. The RMSD of drug action models and Capecitabine are converged during 0.2 to 0.35 ns. Paclitaxel-drug action complexes converged at 0.30 to 0.45 ns. The RD-Lapatinib complex was found to converge at 0. 5 to 0.7 ns, the RD-Capecitabine complex started converging at 0.3 to 0.5 ns, and the RD-Paclitaxel complex converged at 0.4 to 0.55 ns. GID-Lapatinib complex was found to converge at 0.25 to 0.4 ns, GID-Capecitabine complex started converging at 0.20 to 0.25 ns, and GID-Paclitaxel complex converged at 0.30 to 0.40 ns. HF-Lapatinib complex was found to converge at 0. 5 to 0.7 ns, HF-Capecitabine complex started converging at 0.3 to 0.5 ns, and HF-Paclitaxel complex converged at 0.4 to 0.55 ns. When the interaction of the ligand-target complex was studied after the simulation, it was found that there were slight variations in hydrogen bonds whose bond length exceeded 2.8 A. Few interactions were lost, whereas there are complexes where a hydrogen bond interaction was added within 2.5A. Certain complexes retained their interaction throughout the simulation process.

Lapatinib is most suitable for the American population and least suitable for the South Asian population. There are chances for patients from the EUR population to develop gastrointestinal disorders when administered with Lapatinib. Capecitabine showed good interaction with the African population and was least useful for the European population. There are chances for the patients to develop respiratory disorders within American and European population's with gastrointestinal side effects for EAS, SAS and AFR when administered with Capecitabine. Paclitaxel was found to be most interacting with the European population. It was found that respiratory disorders within the East Asian population, EAS, SAS and AFR populations develop gastrointestinal side effects and SAS and hepatic disorders among SAS and EUR population.

CONCLUSION

According to the clinical trial reports of anti-breast cancer drugs, Lapatinib and Capecitabine are more suitable for the AMR population and least ideal for SAS populations, whereas Paclitaxel is most suitable for SAS and least suitable for the EUR population. The rate of side effects also varied among the populations. The importance of the pharmacogenomic model in drug action and the side effects of drugs have been illustrated in this manuscript. The designed models represented population-specific mutant models of the protein targets. The interaction studies of these target models with the drug molecules agree with the observations in the 'ClinTrial database.' The study emphasizes the requirement of the pharmacogenomic suitability of each drug before being suggested for any population. The requirement of the customized and population-specific design of drugs is highly recommended to meet the 'personalized, preventive, participatory and preventive (P4)' strategy of pharmacogenomics.

ABBREVIATIONS

AFR–African, AMR–American, BCL2-BCL2 apoptosis regulator, EAS-East Asian, EGFR-Epidermal Growth Factor Receptor, ERB2-Erb-B2

Receptor Tyrosine Kinase 2, EUR–European, FDA-Food and Drug Administration, GID-Gastrointestinal Disorder, GQME-Global Model Quality Estimation, HF-Hepatic Failure, KEGG-Kyoto Encyclopedia of Genes and Genomes, MD-Molecular Dynamics, NAMD-Nanoscale Molecular Dynamics, NCBI-National Centre for Biotechnology Information, NVT-Number Volume Temperature, RD-Respiratory Disorder, RMSD-Root Mean Square Deviation, SAS-South Asian, SNP-Single Nucleotide Polymorphism, SNV-Single Nucleotide Variation, TYMS-Thymidylate Synthetase

ACKNOWLEDGEMENT

The authors would like to acknowledge the 'Ministry of Electronics and Information Technology, Government of India' for meeting the research fellowship under the Visvesvaraya PhD Scheme for Electronics and IT.

The authors express their gratitude to the 'Biopharma solutions,' the industry partner of AMMAS research lab for their support and help in completing the project.

FUNDING

Nil

AUTHORS CONTRIBUTIONS

PK Krishnan Namboori: Study design and concept, Result in analysis and interpretation, Paper correction

Hima Vyshnavi A M: Data collection, carrying out the work, Result analysis and interpretation, Paper writing.

CONFLICT OF INTERESTS

The authors do not have any conflict of interest

REFERENCES

1. Subramanian I, Verma S, Kumar S, Jere A, Anamika K. Multi-omics data integration, interpretation, and its application. Bioinform Biol Insights. 2020 Jan;14:1177932219899051. doi: 10.1177/1177932219899051, PMID 32076369.
2. Pinu FR, Beale DJ, Paten AM, Kouremenos K, Swarup S, Schirra HJ, Wishart DD. Systems biology and multi-omics integration: viewpoints from the metabolomics research community. Metabolites. 2019 Apr 18;9(4):76. doi: 10.3390/metabo9040076, PMID 31003499.
3. Jamil IN, Remali J, Azizan KA, Nor Muhammad NA, Arita M, Goh HH, Aizat WM. Systematic multi-omics integration (MOI) approach in plant systems biology. Front Plant Sci. 2020;11:944. doi: 10.3389/fpls.2020.00944, PMID 32754171.
4. Hasin Y, Seldin M, Lusis A. Multi-omics approaches to disease. Genome Biol. 2017 May;18(1):83. doi: 10.1186/s13059-017-1215-1, PMID 28476144.
5. Haas R, Zelezniak A, Iacovacci J, Kamrad S, Townsend S, Ralser M. Designing and interpreting 'multi-omic' experiments that may change our understanding of biology. Curr Opin Syst Biol. 2017 Dec;6:37-45. doi: 10.1016/j.coisb.2017.08.009, PMID 32923746.
6. Misra BB, Langefeld CD, Olivier M, Cox LA. Integrated omics: tools, advances, and future approaches. J Mol Endocrinol. 2018;62(1):R21-45. doi: 10.1530/JME-18-0055, PMID 30006342.
7. Iyer PM, Karthikeyan S, Sanjay Kumar P, Krishnan Namboori PK. Comprehensive strategy for the design of precision drugs and identification of genetic signature behind proneness of the disease-a pharmacogenomic approach. Funct Integr Genomics. 2017 Jul;17(4):375-85. doi: 10.1007/s10142-017-0559-7, PMID 28470340.
8. AMH, CLA, OMD, Namboori PKK. Evaluation of colorectal cancer (CRC) epidemiology a pharmacogenomic approach. J Young Pharm. 2017 Jan;9(1):36-9. doi: 10.5530/jyp.2017.9.7.
9. Iyer PM, PSK, SK, Namboori PKK. "BRCA1" Responsiveness towards breast cancer-a population-wise pharmacogenomic analysis. Int J Pharm Pharm Sci. 2016 Jul;8(9):267-70. doi: 10.22159/ijpps.2016.v8i9.13457.
10. Vogenberg FR, Isaacson Barash C, Pursel M. Personalized medicine: part 1: evolution and development into theranostics. PT. 2010 Oct;35(10):560-76. PMID 21037908.
11. Nagai H, Kim YH. Cancer prevention from the perspective of global cancer burden patterns. J Thorac Dis. 2017 Mar;9(3):448-51. doi: 10.21037/jtd.2017.02.75, PMID 28449441.
12. Sharma GN, Dave R, Sanadya J, Sharma P, Sharma KK. Various types and management of breast cancer: an overview. J Adv Pharm Technol Res. 2010 Apr;1(2):109-26. PMID 22247839.
13. Pfizer. Phase III randomized, multi-center study of sunitinib malate (SU011248) or capecitabine in subjects with advanced breast cancer who failed both a taxane and an anthracycline chemotherapy regimen or failed with a taxane and for whom further anthracycline therapy is not indicated; 2012 Jun. Available from: https://clinicaltrials.gov/ct2/show/NCT00373113. [Last accessed on 25 Jan 2022]
14. Tracon Pharmaceuticals Inc. An open-label phase 1B dose-finding study of TRC105 in combination with capecitabine for progressive or recurrent metastatic breast cancer; 2019 Feb. Available from: https://clinicaltrials.gov/ct2/ show/ NCT01326481. [Last accessed on 25 Jan 2022]
15. Wang L. Pharmacogenomics: a systems approach. Wiley Interdiscip Rev Syst Biol Med. 2010 Jan–Feb;2(1):3-22. doi: 10.1002/wsbm.42, PMID 20836007.
16. Sanjay Kumar P, Karthikeyan S, Iyer PM, Krishnan Namboori PK. Prediction of epigenetic variations in Alzheimer's disease identification of ethnic variants through pharmacogenomic approach. Res J Pharm Biol Chem Sci. 2016 Nov;7:2742-5.
17. Katsonis P, Koire A, Wilson SJ, Hsu TK, Lua RC, Wilkins AD, Lichtarge O. Single nucleotide variations: biological impact and theoretical interpretation. Protein Sci. 2014 Dec;23(12):1650-66. doi: 10.1002/pro.2552, PMID 25234433.
18. Namboori PK, Vineeth KV, Rohith V, Hassan I, Sekhar L, Sekhar A, Nidheesh M. The ApoE gene of Alzheimer's disease (AD). Funct Integr Genomics. 2011 Dec;11(4):519-22. doi: 10.1007/s10142-011-0238-z, PMID 21769591.
19. Nishi H, Nakata J, Kinoshita K. Distribution of single-nucleotide variants on protein-protein interaction sites and its relationship with minor allele frequency. Protein Sci. 2016 Feb;25(2):316-21. doi: 10.1002/pro.2845, PMID 26580303.
20. Bordoli L, Kiefer F, Arnold K, Benkert P, Battey J, Schwede T. Protein structure homology modeling using SWISS-MODEL workspace. Nat Protoc. 2009;4(1):1-13. doi: 10.1038/nprot.2008.197, PMID 19131951.
21. Meng XY, Zhang HX, Mezei M, Cui M. Molecular docking: a powerful approach for structure-based drug discovery. Curr Comput Aided Drug Des. 2011 Jun;7(2):146-57. doi: 10.2174/157340911795677602, PMID 21534921.
22. Leyana PN, Manju PT, Vijayan M. In silico design of benoxazole bearing azetidinone derivatives as vegfr-2 agonist in cancer. Asian J Pharm Clin Res. 2021 Nov;14(11):112-5.
23. Manjunatha KS, Satyanarayan ND, Harishkumar S. Antimicrobial and *in silico* ADMET screening of novel (E)-N-(2-(1H-INDOL-3-YL-AMINO) vinyl)-3-(1-METHYL-1H-INDOL-3-YL)-3-phenylpropanamide derivatives. Int J Pharm Pharm Sci. 2016 Oct;8(10):251-6. doi: 10.22159/ijpps.2016v8i10.13957.
24. Vieira TF, Sousa SF. Comparing AutoDock and Vina in ligand/decoy discrimination for virtual screening. Appl Sci. 2019 Oct;9(21):4538. doi: 10.3390/app9214538.
25. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. Scalable molecular dynamics with NAMD. J Comput Chem. 2005 Dec;26(16):1781-802. doi: 10.1002/jcc.20289, PMID 16222654.
26. ClinicalTrials.gov [internet]. Bethesda: National Library of Medicine. US; 2000 Feb 29. Available from: https://clinicaltrials.gov. [Last accessed on 27 Jan 2022]
27. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, Sajed T, Johnson D, Li C, Sayeeda Z, Assempour N, Iynkkaran I, Liu Y, Maciejewski A, Gale N, Wilson A, Chin L, Cummings R, Le D, Pon A, Knox C, Wilson M. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res. 2018 Jan;46(D1):D1074-82. doi: 10.1093/nar/gkx1037, PMID 29126136.
28. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and

drugs. Nucleic Acids Res. 2017 Jan;45(D1):D353-61. doi: 10.1093/nar/gkw1092, PMID 27899662.

29. Oscanoa J, Sivapalan L, Gadaleta E, Dayem Ullah AZ, Lemoine NR, Chelala C. SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). Nucleic Acids Res. 2020 Jul;48(W1):W185-92. doi: 10.1093/nar/gkaa420, PMID 32496546.
30. Gene NCBI [internet]. (Us). Bethesda: National Library of Medicine. National Center for Biotechnology Information; 2004. Available from: https://www.ncbi.nlm.nih.gov/gene. [Last accessed on 27 Jan 2022]
31. UniProt Consortium. UniProt: the universal protein KnowledgeBase in 2021. Nucleic Acids Res. 2021 Jan;49(D1):D480-9. doi: 10.1093/nar/gkaa1100, PMID 33237286.
32. Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 2018 Jul;46(W1):W296-303. doi: 10.1093/nar/gky427, PMID 29788355.
33. Studer G, Rempfer C, Waterhouse AM, Gumienny R, Haas J, Schwede T. QMEAND is Co-distance constraints applied on model quality estimation. Bioinformatics. 2020 Mar 1;36(6):1765-71. doi: 10.1093/bioinformatics/btz828, PMID 31697312.
34. Chen YR, Peng SL, Tsay YW. Protein secondary structure prediction based on Ramachandran maps. In: Huang DS, Wunsch D, editors. C Levine D, S Jo KH. (eds) Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues. ICIC 2008. Lecture Notes in Computer Science. Berlin, Heidelberg. Springer; 2008;l 5226.
35. Khor BY, Tye GJ, Lim TS, Noordin R, Choong YS. The structure and dynamics of BmR1 protein from brugia malayi: *in silico* approaches. Int J Mol Sci. 2014 Jun;15(6):11082-99. doi: 10.3390/ijms150611082, PMID 24950179.
36. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010 Jan;31(2):455-61. doi: 10.1002/jcc.21334, PMID 19499576.
37. Agustini D, Vernadesly L Delviana, Theodorus. The potential of robusta coffee (Coffea Canephora) as a colorectal cancer therapy modality: an *in silico* study. Asian J Pharm Clin Res. 2021 Sep;14(10).
38. Yi F, Ji Z, Zhiguo C. Insights into the molecular mechanisms of protein-ligand interactions by molecular docking and molecular dynamics simulation: a case of oligopeptide binding protein. Comp Math Methods Med. 2018 Dec;2018:(35025140).
39. Gunn D, Garsed K, Lam C, Singh G, Lingaya M, Wahl V, Niesler B, Henry A, Hall IP, Whorwell P, Spiller R. Abnormalities of mucosal serotonin metabolism and 5-HT_3 receptor subunit 3C polymorphism in irritable bowel syndrome with diarrhoea predict responsiveness to ondansetron. Aliment Pharmacol Ther. 2019 Sep;50(5):538-46. doi: 10.1111/apt.15420, PMID 31342534.
40. Connolly B, Isaacs C, Cheng L, Asrani KH, Subramanian RR. SERPINA1 mRNA as a treatment for Alpha-1 antitrypsin deficiency. J Nucleic Acids. 2018 Jun;2018:8247935. doi: 10.1155/2018/8247935.
41. Beers MF, Mulugeta S. The biology of the ABCA3 lipid transporter in lung health and disease. Cell Tissue Res. 2017 Mar;367(3):481-93. doi: 10.1007/s00441-016-2554-z, PMID 28025703.
42. Carpenter AM, Singh IP, Gandhi CD, Prestigiacomo CJ. Genetic risk factors for spontaneous intracerebral haemorrhage. Nat Rev Neurol. 2016 Jan;12(1):40-9. doi: 10.1038/nrneurol.2015.226, PMID 26670299.
43. Sakano T, Mahamood MI, Yamashita T, Fujitani H. Molecular dynamics analysis to evaluate docking pose prediction. Biophys Physicobiol. 2016 Jul;13:181-94. doi: 10.2142/ biophysico.13.0_181, PMID 27924273.
44. Mateev E, Valkova I, Georgieva M, Zlatkov A. Through ensemble docking. Int J Pharm Pharm Sci. 2021 Aug;13(8):32-5.