의학박사 학위논문

A novel design for drug pathways to contribute to personalized medicine

개인 맞춤 약물 제공에 기여하기 위한 새로운 약물 패스웨이 설계

2020 년 8 월

서울대학교 대학원

(협)의료정보학과 의료정보학전공

홍 주 영

A novel design for drug pathways to contribute to personalized medicine

지도 교수 김 주 한
이 논문을 의학박사 학위논문으로 제출함

2020 년 5 월

서울대학교 대학원
(협)의료정보학과 의료정보학전공
홍 주 영

홍주영의 의학박사 학위논문을 인준함
2020 년 7 월

위 원 장 최 진 욱 (인)

부위원장 김 주 한 (인)

위 원 장 인 진 (인)

위 원 강 혜 련 (인)

위 원 김 혜 리 (인)

Abstract

Introduction: Genetic variations in human enzymes or transporters cause changes in the drug concentration inside the human body, which result in individual differences in drug response. Therefore, maintaining the optimal drug concentration by adjusting drug doses or selecting alternatives is necessary to maximize drug efficacy and safety. Recently, Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines based on pharmacogenetics have been published. However, these guidelines have some problems with clinical applications or regulations or standardization related to system development. To this end, I propose a pharmacogenomics pathway: PG-path, which can predict changes in the plasma drug concentration inside the body.

Methods: The gene set that interacts with a specific drug is extracted from DrugBank, and the interaction types (enzyme, transporter, target) and action types (inhibition, induction, substrate) are obtained. The next step produces a frame and a background image to build the pharmacokinetic (PK) pathway and the pharmacodynamic (PD) pathway. Then, extracted elements are applied to the designed frame to visualize the interaction of drugs and genes. PathVisio is used to

standardize components and formats.

Results: The PG-path consists of a frame with organs in the human body and a background image with each organ figure. The interaction type and action type between drugs and genes were drawn with a standard symbol. The frame with the diagram was merged with the background image. We improved the understanding of the PG-path by hyperlinking the window containing information on genes and drugs to each node and popping it up. Genes were placed similar to inside the body to visualize the flow of the drug in the body. We applied the gene-wise variant burden (GVB) score which is the degree of accumulative damage of the gene to each gene, to make an individualized pathway.

Conclusions: The PG-path is designed to visualize the general flow of drugs in the human body. Applying the GVB score to each gene makes it possible to predict the change in plasma drug concentration. Since focusing on each drug, the PG-path can predict the effect of drug-gene-drug interactions on the drug response when multiple drugs are administered. Adding PK properties and clinical factors to the PG-path could improve the ability to predict drug response.

Keywords: pharmacogenomics, pharmacogenetics, pharmacokinetics, pharmacodynamics, pathway, personalized medicine, drug response, drug-gene-drug interaction

Student number: 2016-30616

Contents

Introduction 1

Personalized medicine based on pharmacogenetics 3

Prescription guideline 3

Implementation of the PGx-based prescription guideline system 6

Personalized medicine based on pharmacogenomics 9

The current pharmacogenomics-based pathways 9

Objective of this study 13

Materials and Methods 14

Materials 14

DrugBank 5.0.1 14

Sorting Intolerant From Tolerant (SIFT) 14

The 1000 Genomes Project 15

Gene-wise Variant Burden (GVB) score 15

Pathway development methods 16

Pathway development software 21

Results 22

Nodes and edges 22

Background frame and image 26

Convert and merge 28

Information on drugs and genes 31
Drawing a pathway with PathVisio 36
Usage of PG-path: Pathway analysis and visualization with GVB score 38
Discussion 42
Prediction of changes in plasma drug concentration by gene placement in pathways 42
PG-path: Single drug-centered pathway 46
Pathway analysis: GVB scoring method using variations obtained from DNA sequencing 50
Comparison of variant set between GVB scoring method and CPIC guideline 51
Conclusion 52
Reference 53
국문 초록 61

List of Figures

Figure 1. The factors affecting the drug response include physiological, environmental, and genetic factors..... 1

Figure 2. Method of maximizing the efficacy and minimizing the toxicity..... 3

Figure 3. Assigning CYP2C19 phenotypes based on genotypes in the CPIC guideline..... 4

Figure 4. To recommend antiplatelet therapy based on CYP2C19 status when prescribing clopidogrel for ACS/PCI patients. 5

Figure 5. One example of procedures implementing the CPIC guideline..... 8

Figure 6. PharmGKB PK/PD pathway of clopidogrel..... 11

Figure 7. SMPDB PK/PD pathway of clopidogrel..... 12

Figure 8. Procedure of pathway analysis and modeling. 19

Figure 9. To merge a diagram with a background frame and image. 20

Figure 10. The default pharmacokinetic (PK) pathway of clopidogrel. 29

Figure 11. The default pharmacodynamic (PD) pathway of clopidogrel. 30

Figure 12. A pop-up window showing clinical information on the drug.

.....32
Figure 13. A pop-up window showing information on the gene and its interaction with the administered drug.....33
Figure 14. A pop-up window showing information on the gene and its interaction with the administered drug.....34
Figure 15. A pop-up window showing PD pathway description.....35
Figure 16. The personalized pharmacokinetic (PK) pathways of clopidogrel.....40
Figure 17. The personalized pharmacodynamic (PD) pathways of clopidogrel.....41
Figure 18. The change in the plasma concentration of active ingredients depends on drug categories.....45
Figure 19. Drug-gene-drug interaction (DGDI) between Drug A and Drug B.....47
Figure 20. Drug-gene-drug interaction (DGDI) among Amiodarone, Fluoxetine, and Codeine.....49

List of Tables

Table 1. Interaction and action types on clopidogrel from DrugBank 5.0.1. 18

Table 2. The characteristics of a pharmacokinetic pathway..... 24

Table 3. The characteristics of a pharmacodynamic pathway. 25

Introduction

The drug responses vary from person to person even if patients administer the same drug at the same dose because age, gender, liver or kidney dysfunction, and drug interactions affect the drug response. Recently, many studies have shown that genetic factors have a significant effect on drug response [1] (Figure 1). The disciplines supporting this phenomenon are pharmacogenetics and pharmacogenomics (PGx). Pharmacogenetics is the discipline that seeks to enhance the therapeutic effect by finding a genetic variation affecting a drug response in a particular gene. While, PGx is a discipline that seeks to contribute to improving the therapeutic effect by detecting variants affecting the drug response at the genome level and revealing the relationship with the drug [2].

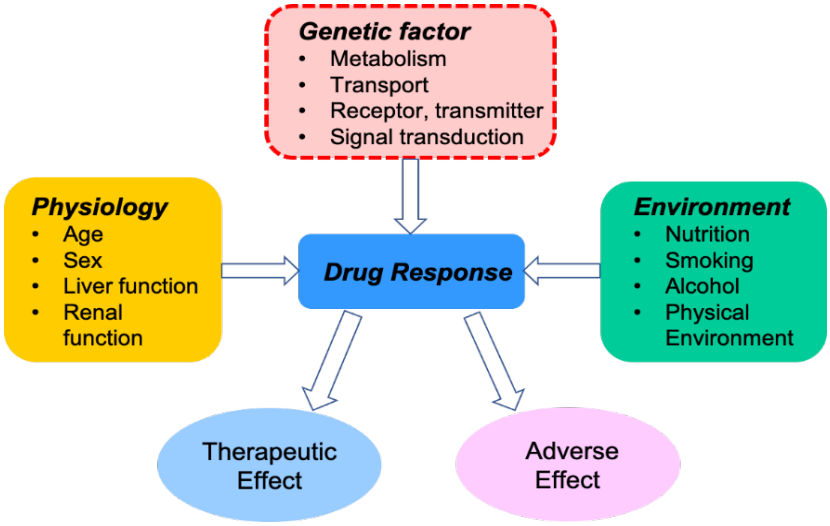


Figure 1. The factors affecting the drug response include physiological, environmental, and genetic factors.

When patients administer drugs, two stages of reactions occur in the body. Pharmacokinetics (PK) refers to the process of absorbing, metabolizing, distributing, and excreting a drug. This process transforms a drug into a suitable form and dosage to reduce the toxicity of the drug inside the human body. On the other hand, pharmacodynamics (PD) refers to a drug's action at a target cell to show its efficacy. Genetic variations in enzymes or transporters that affect the PK process of absorption, distribution, metabolism, and excretion (ADME) cause drug concentration changes. Therefore, taking into account the effect of genetic variation, it is possible to change the dose of the drug from outside the human body or to administer an alternative drug so that the range of the drug concentration in the plasma is within the therapeutic range. Thus, the optimized concentration of the drug reacts at the target, contributing to maximizing the effect of the drug and minimizing the drug's toxicity (Figure 2). This is the principle pursued by personalized medicine [3].

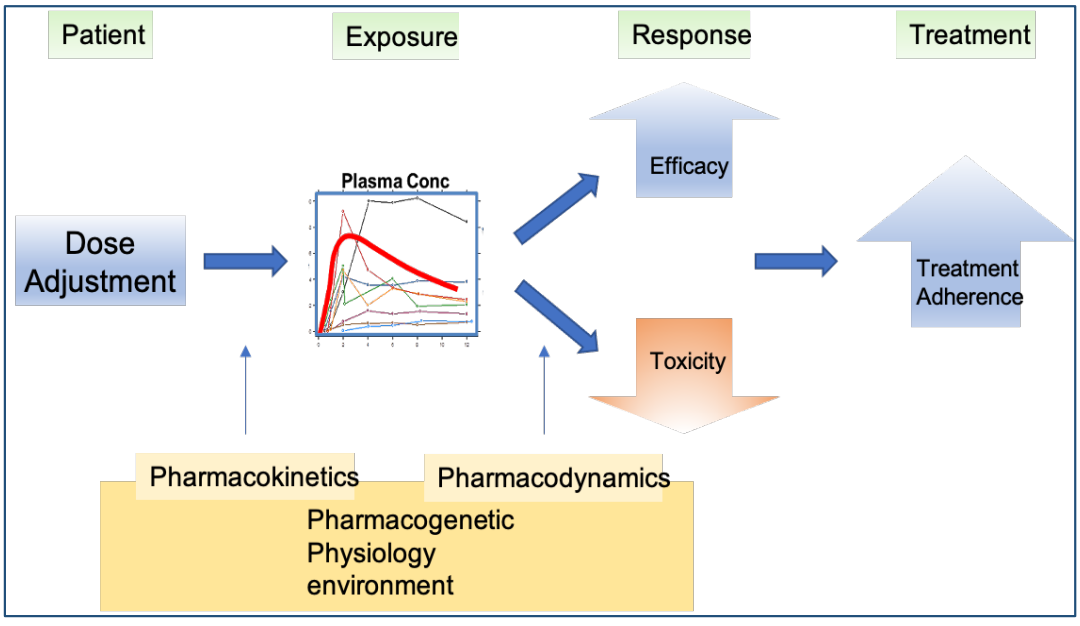


Figure 2. Method of maximizing the efficacy and minimizing the toxicity.[1]
Through dose adjustment, the drug concentration inside the body is optimized.

Personalized medicine based on pharmacogenetics

Prescription guideline

Both pharmacogenetics and PGx pursue personalized medicine, but their approaches are different. Many researchers have long tried to systematically apply the results of genetic testing obtained in the laboratory to clinical practice. As part of these studies, PGx-based prescription guidelines such as the Clinical Pharmacogenetics Implementation Consortium (CPIC) guidelines and Dutch Pharmacogenetics Working Group (DPWG) guidelines have begun to

[1] Translational PK/PD Modeling Facility-Quantitative clinical pharmacology.

be published [4]. It is the concept based on pharmacogenetics. These guidelines interpret variations found in a particular gene–drug pair, translate these to the corresponding phenotypes and present therapeutic recommendations for these phenotypes. These guidelines present recommendations such as altering drug doses or taking alternatives in cases of patients with significant genetic variants in enzymes or transporters that affect changes in protein function. By doing so, these guidelines try to maximize the efficacy and safety of the drug (Figures 3, 4).

Likely phenotype	Genotypes	Examples of diplotypes
Ultrarapid metabolizer: normal or increased activity (~5–30% of patients)	An individual carrying two increased activity alleles (*17) or one functional allele (*1) plus one increased-activity allele (*17)	*1/*17, *17/*17
Extensive metabolizer: homozygous wild-type or normal activity (~35–50% of patients)	An individual carrying two functional (*1) alleles	*1/*1
Intermediate metabolizer: heterozygote or intermediate activity (~18–45% of patients)	An individual carrying one functional allele (*1) plus one loss-of-function allele (*2–*8) or one loss-of-function allele (*2–*8) plus one increased-activity allele (*17)	*1/*2, *1/*3, *2/*17
Poor metabolizer: homozygous variant, mutant, low, or deficient activity (~2–15% of patients)	An individual carrying two loss-of-function alleles (*2–*8)	*2/*2, *2/*3, *3/*3

Figure 3. Assigning CYP2C19 phenotypes based on genotypes in the CPIC guideline.[2]

[2] SA Scott. et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C19 Genotype and Clopidogrel Therapy: 2013 Update, *Clinical pharmacology & Therapeutics,* 2013; 94(3).

Phenotype (genotype)	Implications for clopidogrel	Therapeutic recommendations	Classification of recommendations[a]
Ultrarapid metabolizer (UM) (*1/*17, *17/*17) and extensive metabolizer (EM) (*1/*1)	Normal (EM) or increased (UM) platelet inhibition; normal (EM) or decreased (UM) residual platelet aggregation[b]	Clopidogrel: label-recommended dosage and administration	Strong
Intermediate metabolizer (*1/*2, *1/*3, *2/*17)	Reduced platelet inhibition; increased residual platelet aggregation; increased risk for adverse cardiovascular events	Alternative antiplatelet therapy (if no contraindication), e.g., prasugrel, ticagrelor	Moderate
Poor metabolizer (*2/*2, *2/*3, *3/*3)	Significantly reduced platelet inhibition; increased residual platelet aggregation; increased risk for adverse cardiovascular events	Alternative antiplatelet therapy (if no contraindication), e.g., prasugrel, ticagrelor	Strong

Figure 4. To recommend antiplatelet therapy based on CYP2C19 status when prescribing clopidogrel for ACS/PCI patients.[3]

This method is ideal, but there are several problems. First, there is no standardized procedure from experimental results to prescription recommendations. Although multiple institutions or large hospitals have conducted many clinical studies to verify guidelines, since each organization has its processes, each study sometimes yields different results. Second, clinicians often do not believe in the necessity of gene–drug pairs, which results in a lack of confidence in whether the variations presented in the guidelines are reliable enough to be applied to clinical practice. Third, it is difficult to recommend the test to the patient because of the burden of the test cost and uncertain reimbursement. Moreover, if a therapeutic

[3] SA Scott. et al. Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C19 Genotype and Clopidogrel Therapy: 2013 Update, *Clinical pharmacology & Therapeutics,* 2013; 94(3).

recommendation does not help the patient's treatment, coping may be difficult [4].

Implementation of the PGx-based prescription guideline system

When developing a guideline system and implementing it in hospitals or institutions, there are some problems with standardization. First, the variant-extraction method that composes the panel for genotyping is not standardized. If the panel is composed of entirely actionable variants, testing with this panel will give the best results. However, it is desirable to extract appropriate variants in consideration of cost. However, since the extraction method is not standardized, sometimes, the variation in the patient is presented differently by the institution [5, 6]. Second, occasionally, different phenotypes are provided as a result because the standard for interpreting the variants varies from laboratory to laboratory [7, 8]. Researchers are trying to provide a standard for interpreting variants by delivering software such as PharmCAT or Stargazer. However, it is difficult to obtain the same results among studies because even software sometimes presents different results [9, 10]. Third, therapeutic recommendations are often different by guidelines because the range of sources for evidence-based guidelines is varied.

The CPIC guidelines even use case reports as evidence. On the other hand, the DPWG guidelines use only clinical trials or clinical studies with well-designed protocols, such as high or moderate rates, as sources. Therefore, several test results occasionally provide different recommendations for the same phenotype [11, 12] (Figure 5).

Currently, the CPIC guidelines have been published for only 19 genes; thus, the number of guidelines is small, and the velocity of generating them is very slow. Since the current guidelines refer to genome-wide association study (GWAS) research, the proposed variants are composed of common variants. Therefore, it is difficult to provide accurate prescription recommendations for patients with rare variants [13, 14]. Furthermore, since drug responses may vary by ethnicity, it is challenging to apply guidelines to a specific ethnicity often.

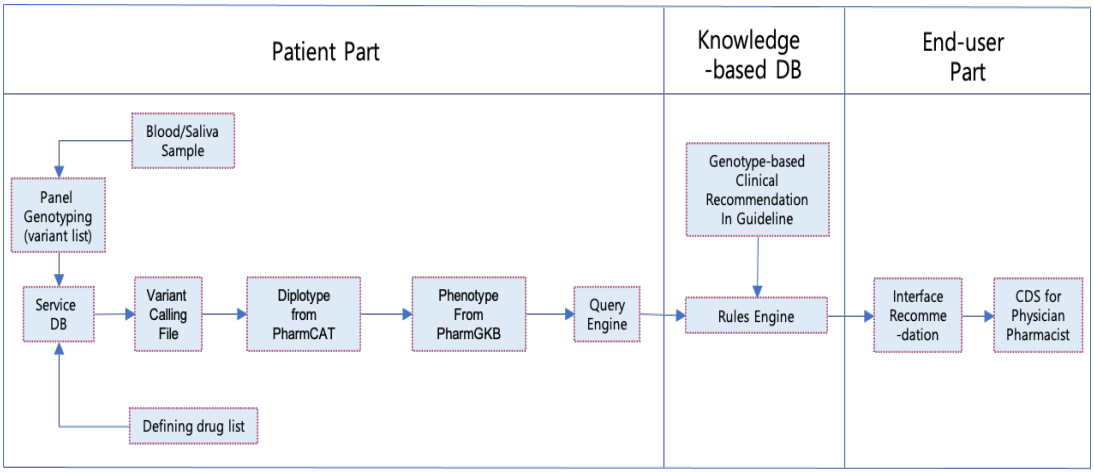


Figure 5. One example of procedures implementing the CPIC guideline

Although the guidelines can interpret the test results and apply them to clinical practice, each guideline must be evaluated with many clinical studies to be used as a complete alternative in actual clinical application. With recent advances in sequencing technology and reduced testing costs [15], expectations for whole-genome-level analysis are increasing. Although many genes interact with a drug when patients administer a medication, many studies have mainly adopted only one major drug-gene pair. Unless weakly interacting genes are considered, these studies can lead to incomplete results. In addition, attempts to find variations associated with drug responses at the whole genome level have increased usability by opening the likelihood of detecting rare or novel variants that cannot be found in the prescription guidelines.

New tools are needed to integrate PGx concepts that address drug responses and gene impacts with each drug. As a new tool supporting personalized medicine, I propose a PGx pathway in which the nature of the variation can be easily applied and modified while visualizing the entire genes interacting with the drug. PGx pathways could be as useful as pharmacogenetics-based prescription guidelines.

Personalized medicine based on pharmacogenomics

The current pharmacogenomics-based pathways

For many decades, pharmacologists have been producing pathway diagrams to explain the pharmacology of administered drugs. These diagrams have proven to be more robust in organizing, sharing, and discussing knowledge than other known methods [16].

A type of diagram that visualizes flows and actions of drugs inside the human body is a PGx-based pathway, which depicts a chain of reactions between an administered drug and genes in the human body [17]. Researchers utilize the PGx-based pathway to explain PK and PD. A PK pathway depicts a drug's ADME conducted by enzymes, transporters, or carriers at the systemic level. A PD

pathway describes the mechanism of action by which a drug affects its targets, such as components of signaling pathways or metabolic pathways, or the receptors themselves, at a cellular–molecular level [18, 19].

PGx studies analyze the gene variations involved in PK and PD and help understand how the variations alter protein function. Based on previous discoveries, PGx studies goal to maximize the efficacy and safety of drugs by adjusting the dosage or altering the drug itself [5, 20, 21]. A PGx–based pathway portrays the series of interactions between genes and a drug in terms of PK and PD. It clarifies how changes in the function of a protein, due to genomic variation, affect the ADME of a drug in the body and the drug action in a target cell [17].

As research on the interaction between drugs and genes has accelerated alongside advances on the Internet and systems pharmacology [22], many PGx–based pathways have been systematically built. Through multidisciplinary studies, many pathway resources have been developed for helping various goals, ranging from the identification of drug–related genes to the design of tools for drug discovery [23–25]. The drug–related pathways include metabolite signaling pathways and drug–metabolic or drug–

action pathways [26].

Pathway resources that make use of PK and PD in relation to drugs include the Pharmacogenomics Knowledgebase (PharmGKB) [27] and the Small Molecule Pathway Database (SMPDB) [26-28]. The PharmGKB pathway is a drug-centric diagram comprising PK/PD-related genes. The diagram is designed to present genes interacting with drugs and affecting drug metabolism and drug response. These are the means of concatenating separate data sets and act as snapshots of current knowledge. These pathways have been manually curated by experts and include novel or candidate genes, as well as common genes [29] (Figure 6).

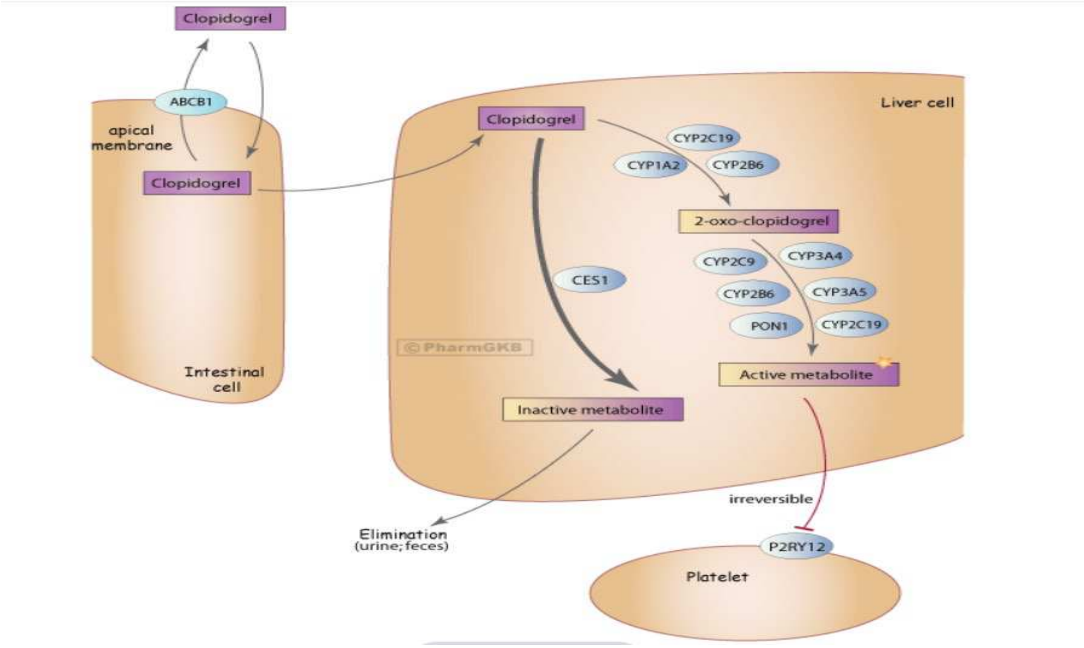


Figure 6. PharmGKB PK/PD pathway of clopidogrel.[4]

[4] https://www.pharmgkb.org/pathway/PA154424674

SMPDB is a set of diagrams drawn exclusively for small molecules. It is a database of metabolism-oriented pathways presenting human metabolic, drug metabolism, metabolic disease paths, etc. These pathway components are linked with detailed descriptions, and the elements and formats have been standardized. This resource is hyperlinked to massive PGx database resources. Each drug diagram forms a part of the overall metabolism process. Because it deals only with small molecules, except for bio drugs, both drugs and metabolites are represented by a chemical formula [28] (Figure 7).

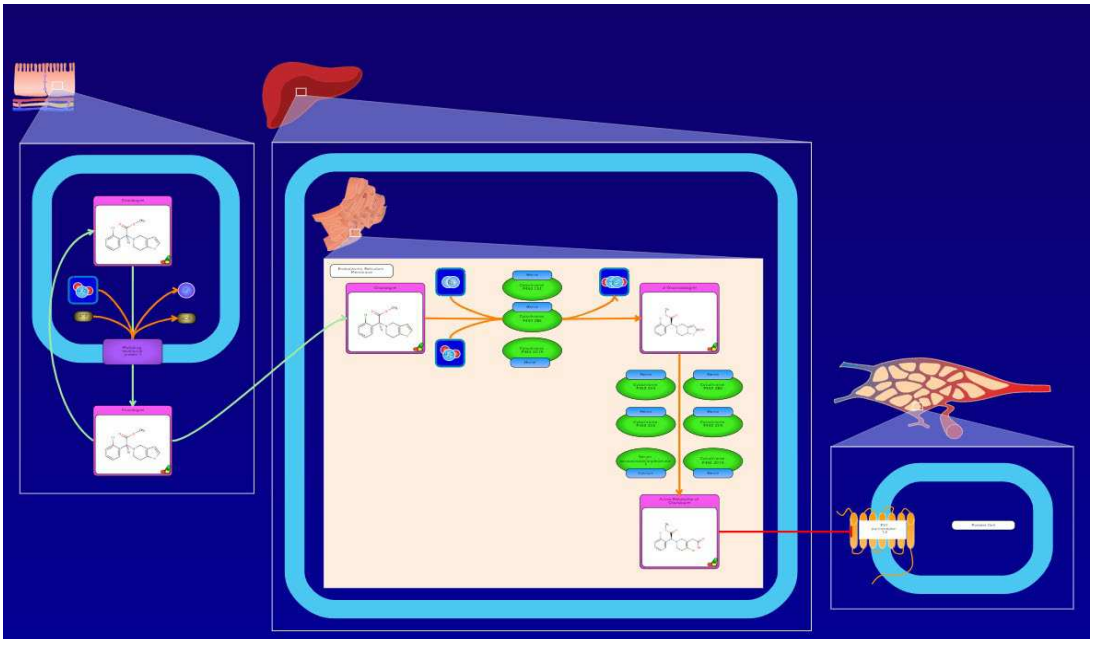


Figure 7. SMPDB PK/PD pathway of clopidogrel.[5]

[5] https://smpdb.ca/view/SMP0000610

Similar to these, several pathway resources have been standardized by applying internal norms to components and formats of pathways. However, many drug pathways remain unstandardized, and separately created pathways are also unstandardized in terms of interoperability. Most drug pathways are utilized to identify gene function or to search for novel genes. However, they are not suitable for supporting tailored drug treatments in clinical settings due to complexities.

.

Objective of this study

The objective of this study is to model the new knowledge-based PGx pathway: PG-path. The PG-path can display the interaction of entire genes with an administered drug and modify the attributes of variations for each gene. The pathway supplies an intuitive understanding of the PK and PD of a specific drug. Furthermore, the PG-path is designed to support tailored drug prescriptions by predicting the change in drug concentration in the plasma.

Materials and Methods

Materials

DrugBank 5.0.1

The components in the PG-path were extracted from DrugBank 5.0.1 [30]. The components extracted from DrugBank are composed of drug-gene pairs, drug interaction types (enzymes, transporters, carriers, and targets), and drug actions (inhibitors, inducers, and substrates). This information was curated from extensive primary literature sources by domain-specific experts and skilled biocurators related to DrugBank [30].

Sorting Intolerant From Tolerant (SIFT)

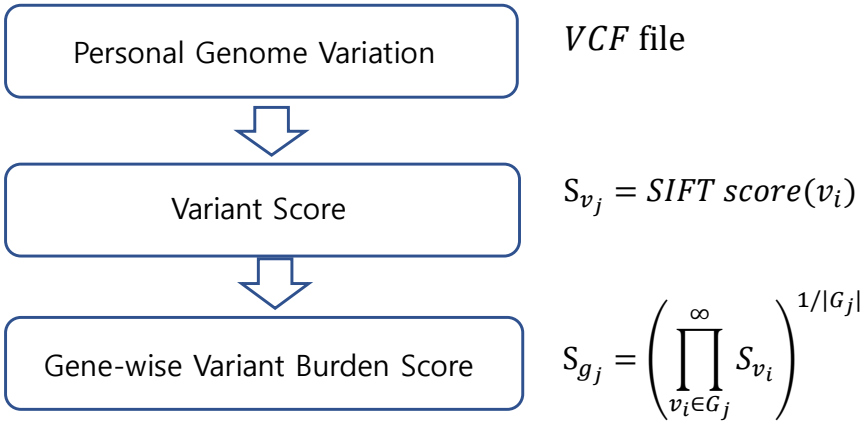
Using sequence homology, SIFT predicts whether an amino acid substitution will impact the protein function and therefore alter the phenotype [31-33]. SIFT is a standard tool for distinguishing missense variations. For amino acid substitutions, SIFT has been mainly used in human genetic research (e.g., cancer, Mendelian diseases, and infectious diseases) [34].

The 1000 Genomes Project

Established in 2007, The 1000 Genomes Project is a comprehensive, open-access database that includes the genetic variants of 2,504 individuals from 26 populations across the globe [35, 36]. The international research consortium has the purpose of sequencing the genomes of at least 1,000 volunteers from various populations worldwide to enhance current knowledge of genetic contributions to human health and disease [36].

Gene-wise Variant Burden (GVB) score

The GVB scores, which indicate the degree of protein function damage, are applied to the genes of a pathway as one method of utilizing the PG-path. The GVB scores use SIFT scores [32]. The GVB score, defined as the geometric mean of the SIFT scores for the set of variants in a gene, is applied to estimate the overall impact of all deleterious variants in the gene [37]. This method operates under the assumption that variants that potentially change protein function are not assured but may cause deleterious phenotypes [38]. A lower GVB score means that variants are likely to be more harmful to the function of the protein encoded by the gene [37, 38].



VCF file

$S_{v_j} = SIFT\ score(v_i)$

$$S_{g_j} = \left(\prod_{v_i \in G_j}^{\infty} S_{v_i} \right)^{1/|G_j|}$$

S_v: *nonsynanymous variant score*

S_g: $Gene - wise\ Variant\ Burden$

G_j: $A\ set\ of\ variants\ in\ gene\ j$

VCF: Variant call format

Pathway development methods

Each PK/PD pathway is developed by applying the following method (Figure 8B):

1. Extracting the drug–gene pair, drug interaction type, and drug action from DrugBank 5.0.1 for the selected drug (Table 1)
2. Creating a standard frame for the PK/PD pathway and producing a set of symbols (Figure 9A)
3. Drawing nodes and edges of the PK/PD pathway using PathVisio software (Figure 9B).
4. Saving the data as the Graphical Pathway Markup Language

(GPML) format and storing the information in a PK/PD GPML folder

5. Drawing the background image for the PK/PD pathway using Illustrator CS6 (Figure 9C)

   (PK: one background image for all drugs; PD: each drug has its own background image)

6. Saving the data in the Portable Network Graphics (PNG) format and storing the information in the PK/PD PNG folder
7. Converting the GPML file to the Scalable Vector Graphics (SVG) format using PathVisioRPC (PK/PD pathway)
8. Merging the SVG file with the background image PNG file (PK/PD pathway) (Figure 9D)
9. Linking to an SVG + Hypertext Markup Language (HTML) file with the gene and drug information windows and a PD description window
10. Generating the complete PK/PD pathway

Drugname	Uniprot_name	Gene	Interaction _type	Action
Clopidogrel	P2Y purinoceptor 12	P2RY12	target	antagonist
	Cytochrome P450 3A4	CYP3A4	enzyme	substrate
	Cytochrome P450 2B6	CYP2B6	enzyme	substrate, inhibitor
	Cytochrome P450 3A5	CYP3A5	enzyme	substrate
	Cytochrome P450 2C19	CYP2C19	enzyme	substrate
	Cytochrome P450 2C9	CYP2C9	enzyme	substrate, inhibitor
	Cytochrome P450 1A2	CYP1A2	enzyme	substrate
	Cytochrome P450 2C8	CYP2C8	enzyme	inhibitor
	Liver carboxylesterase 1	CES1	enzyme	substrate
	Multidrug resistance protein 1	ABCB1	transporter	substrate

Table 1. Interaction and action types on clopidogrel from DrugBank 5.0.1.[6]
Interaction type: the protein role, according to which the reaction between a drug and a gene happens; Action type: the type by which a drug acts on a protein or vice versa.

[6] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

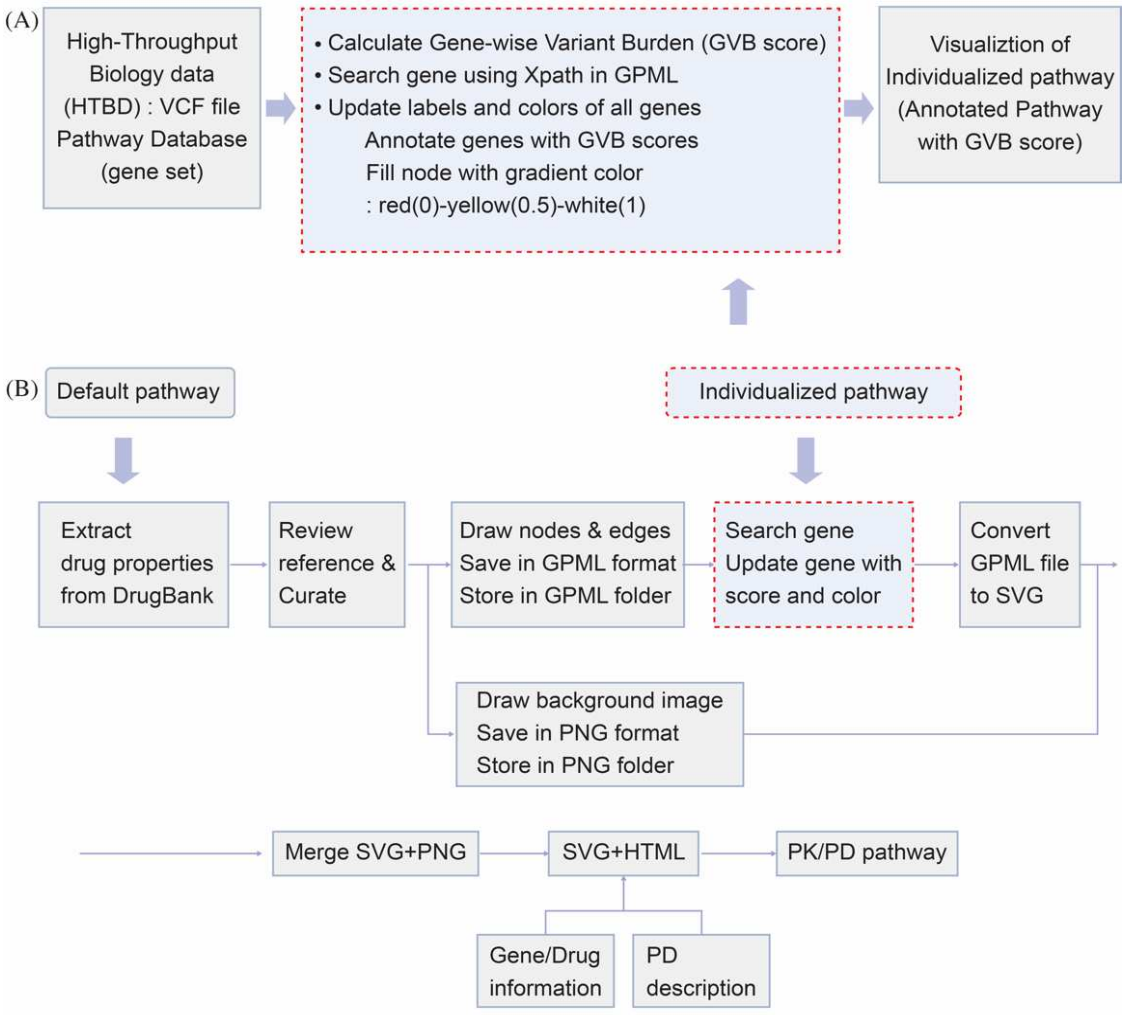


Figure 8. Procedure of pathway analysis and modeling.[7]

(A) Three phases of pathway analysis: (1) to input the high-throughput biological data (HTBD) variants from the VCF file; (2) to perform the algorithm-based analysis and GVB scoring; (3) to visualize the pathway and annotate with the GVB scores.

(B) Development procedure for the default pathway and personalized pathway.

VCF: Variant Call Format; GVB: Gene-wise Variant Burden; GPML: Graphical Pathway Markup Language; SVG: Scalable Vector Graphics; PNG: Portable Network Graphics; HTML: Hypertext Markup Language; PD: Pharmacodynamics.

---

[7] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

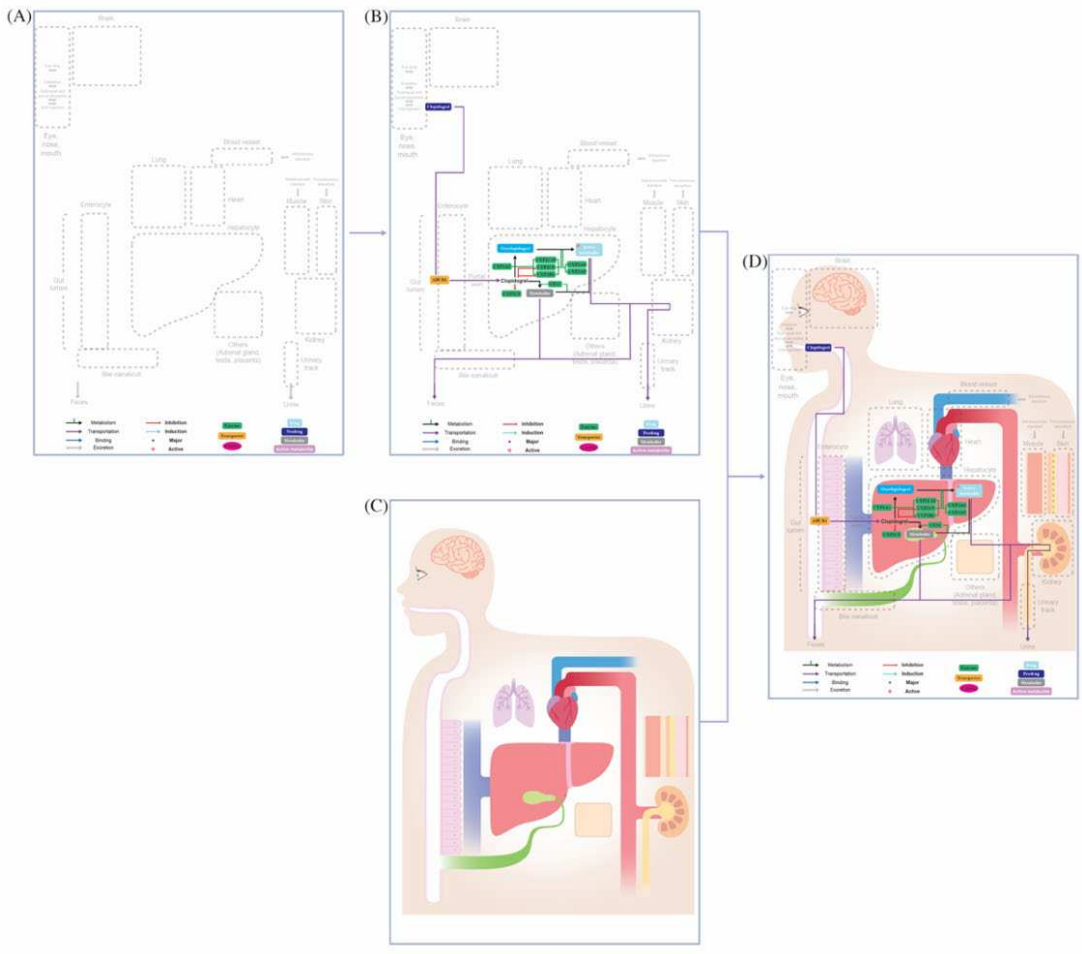


Figure 9. To merge a diagram with a background frame and image.[8]

(A) A standard frame for the PK pathway; (B) the nodes and edges of the PK pathway; (C) the background image for the PK pathway; (D) merging the SVG file with the background image PNG file (PK pathway).

[8] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

Pathway development software

Pathway diagrams were generated in the GPML format using PathVisio 3.2.4. The background images were created in the PNG format using Adobe Illustrator CS6. Each pathway diagram saved in GPML format was converted to SVG format using PathVisioRPC, an XML-RPC interface, after code modification.

Results

Nodes and edges

Although antibiotic, antifungal, and antiviral drugs are included in PG-paths, the PG path is a human-centered PGx pathway. The PG-path are produced by categorizing the PK pathway, which represents the drug's ADME in the human body at the systemic level, and the PD pathway, which depicts the action of the drug at the target cell at the cellular and molecular level.

Through each element's role in consideration, a series of symbols were produced to standardize each component of the pathway. Drugs are classified as prodrugs and active drugs. Prodrugs become activated when the liver enzymes metabolize them, and active drugs are already active in their present form. The same symbol with different colors distinguishes these two categories. Metabolites produced by the metabolism of a drug are also classified into inactive and active metabolites. Similarly, the two groups are classified by the same shape with different colors as a drug. Since the roles of genes encoding enzymes, transporters, and carriers are different from each other, their shapes and colors are differently granted. After standardizing each gene and drug, the edges

representing the actions between the drugs or genes were standardized. The actions are composed of metabolism, transportation, binding, and excretion. Each action's corresponding edge was chosen using PathVisio's molecular interaction map (MIM) tool [39]. Inhibitions and inductions, action types that refer to the actions of drugs on genes, were also symbolized as edges. One of the undefined symbols, the 'major' marker, was for enzymes that play a significant role in metabolism. Furthermore, the 'active' marker was for presenting the active metabolites produced while the prodrug was being metabolized (Table 2).

For the symbols in the PD pathway, the target, drug, and active metabolite were indicated with nodes, while the inhibition, activation, metabolism, binding, conversion, and action were represented with edges (Table 3).

Characteristic	Description
Drug	active drug, prodrug (activating enzyme)
Metabolite	inactive metabolite, active metabolite
Protein	enzyme (e.g., activating enzyme), transporter, carrier
Background anatomical organs	eye, nose, mouth, brain, blood–brain–barrier, lung, heart, muscle, skin, kidney, liver, adrenal gland, testis, intestines, and placenta
Transport structures	blood vessels (e.g., arteries and veins), bile ducts, and excretory tracts (e.g., urinary tract for urine, gut lumen for feces)
Methods of administration	eye drop, inhalation, sublingual and buccal absorption, oral ingestion, intravenous injection, intramuscular injection, and percutaneous absorption
Interaction types (from DrugBank)	metabolism, transportation, binding, excretion
Action types (from DrugBank)	inhibition, induction, substrate
Tissue site for ADME	expression levels of enzymes and transporters through the ProteinAtlas database

Table 2. The characteristics of a pharmacokinetic pathway.[9]

Methods of administration: the location where the drug is administered; Interaction type: the protein role, according to which the reaction between a drug and a gene happens; Action type: the type by which a drug acts on a protein or vice versa; ADME: Absorption, Distribution, Metabolism, Excretion; DrugBank version, 5.0.1.

[9] Hong JY, Kim JH. PG–path: Modeling and personalizing pharmacogenomics–based pathways. *PLoS One*. 2020;15(5): e0230950

Characteristic	Description
Components	active drugs, prodrug, active metabolites, genes (targets)
cellular-level component	cell components including the nucleus, endoplasmic reticulum, mitochondria, Golgi apparatus, lysosomes, peroxisomes, vesicles, cell membrane, ribosomes.
Action types (from DrugBank)	agonist, antagonist, activator, modulator, competitor, cofactor, ligand, stimulator, antibody, binder, potentiator, neutralizer, inhibitor, inducer, etc.
Interaction types (from DrugBank)	the role of the biological pathway of the gene at the molecular level (target)

Table 3. The characteristics of a pharmacodynamic pathway.[10]

Interaction type: the protein role, according to which the reaction between a drug and a gene happens; Action type: the type by which a drug acts on a protein or vice versa; DrugBank version, 5.0.1.

[10] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

Background frame and image

When for a PK pathway to be produced, a background frame modeled on the human body is made so that the proteins can be precisely superimposed on the human organs. Almost all kinds of standard background organs were selected to describe the ADME routes accurately. The background anatomical organs of a standard PK pathway frame consist of the eye, nose, mouth, brain, lung, heart, muscle, liver, intestines, kidney, adrenal gland, testis, placenta, and skin as well as other organs that are needed for a few drugs. Drugs are distributed through the arteries and veins and are excreted through the bile duct or urinary tract. The location and methods of drug administration must be observed closely as the drug concentration of the absorbed substances differs across routes of administration. The administration routes include eye drops, inhalation, oral ingestion, sublingual and buccal absorption, intravenous injection, intramuscular injection, and percutaneous absorption. The tissue sites at which metabolism or transport occurs depend on the expression levels of enzymes and transporters, which can be found by referring to the Protein Atlas databases (Figure 9A, 9B). To graphically describe the ADME of a drug in the PK pathway, a background image composed of standardized body organs was

drawn (Table 2, Figure 9C, 9D).

Each PD pathway is composed of drugs, active metabolites, and genes encoding targets. The PD diagram was produced with PathVisio to visualize the interactions between a drug and its targets. The standard biological mechanism inside a target cell was graphically represented as a background image at the molecular–cellular level. The cellular elements include the nucleus, endoplasmic reticulum, mitochondria, lysosomes, Golgi apparatus, peroxisomes, cell membranes, vesicles, ribosomes, etc. In this way, the effect of the drug can be interpreted intuitively through its mechanism of action on the target (Table 3).

Convert and merge

For a pathway diagram produced in the GPML format, the attributes of nodes and edges can be modified using XPath. The modified GPML file was converted into the SVG format using the PathVisioRPC interface. Each drug was retrieved using the Human Metabolome Database (HMDB) identifier, and each gene was retrieved using the HUGO Gene Nomenclature Committee (HGNC) identifier. This allows for the identification of nodes using XPath. PathVisioRPC is an XML-RPC interface for PathVisio. With this interface, PathVisio's functionalities could directly be applied to any desired analytical environment. The modified and transformed diagram and a background image PNG file were merged to generate complete PK and PD pathways in SVG + HTML format (Figure 9, 10, 11).

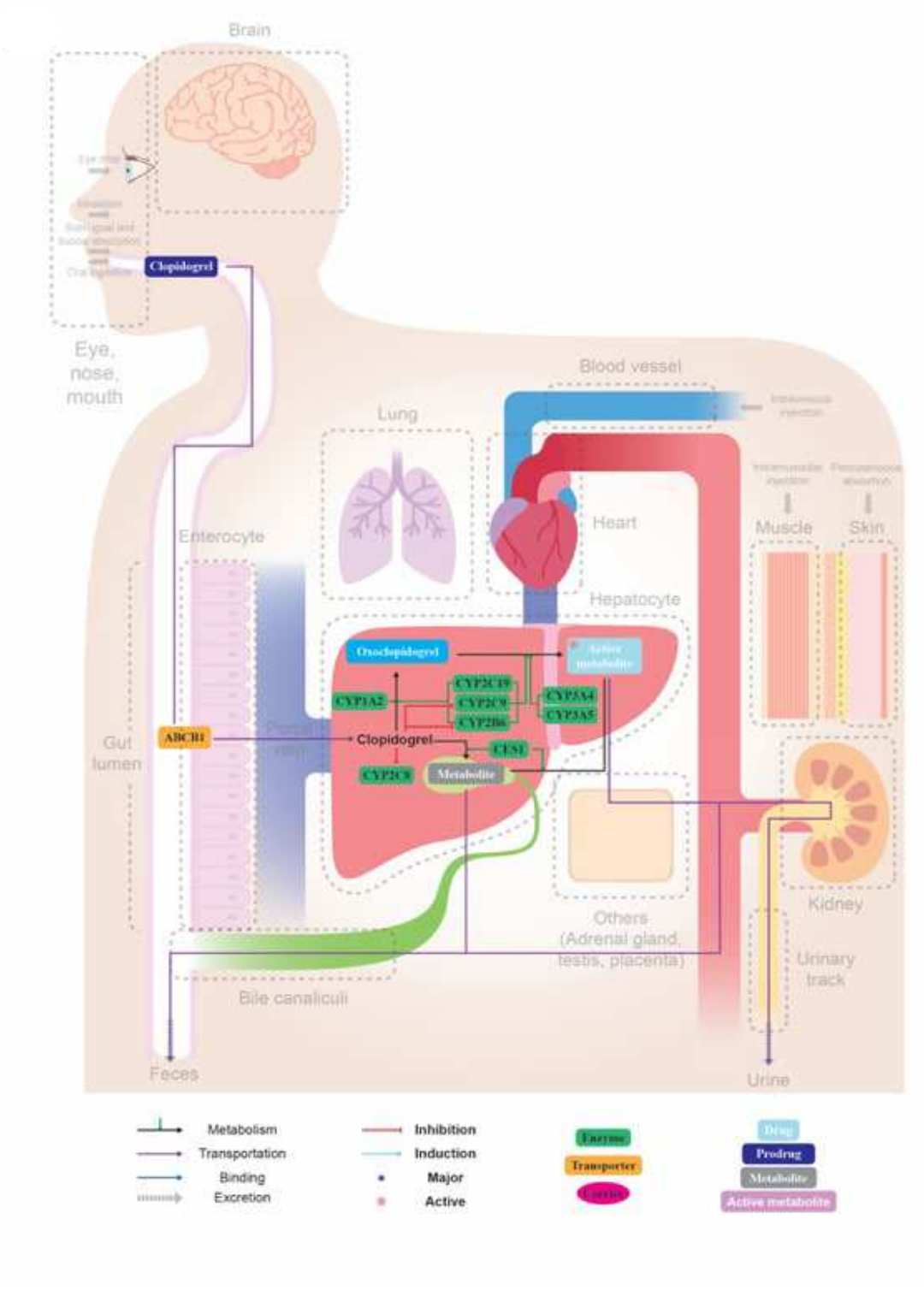


Figure 10. The default pharmacokinetic (PK) pathway of clopidogrel.[11]

[11] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

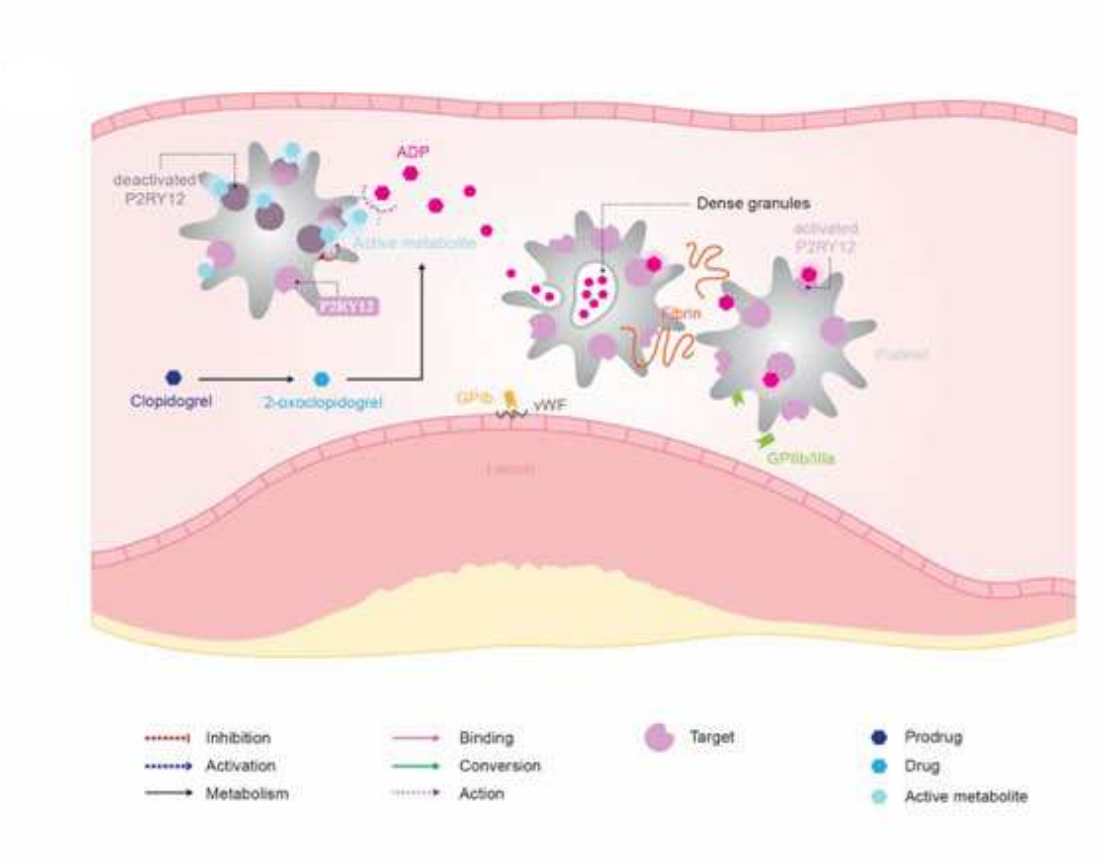


Figure 11. The default pharmacodynamic (PD) pathway of clopidogrel.[12]

[12] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

Information on drugs and genes

Drugs and genes (nodes) on the pathway were linked to a detailed information window, and the PD pathway was linked to a description window to explain the mechanism of action of a drug on a target. This knowledge-based database was manually curated with a wide range of references by experts. In a merged pathway, genes are at the state hyperlinked to the external HGNC database, and drugs are at the state hyperlinked to the external HMDB database. By the identifier of each node used to be linked to an external database, the pathways are hyperlinked to an internal window that includes detailed information on genes or drugs. When a drug symbol is clicked, a pop-up window shows the drug's information, such as its name, Anatomical Therapeutic Chemical (ATC) class, indications, drug interactions, adverse drug effects, and pharmacokinetic data (Figure 12). When a gene symbol is clicked, a pop-up window shows gene information, such as the gene symbol, gene name, interaction type with drug, chromosomal location, expression tissue, action type, and gene function description (Figure 13, 14). The PD pathway has a button to be linked to an internal PD description window to help users understand the mechanism of the action of the drug (Figure 15).

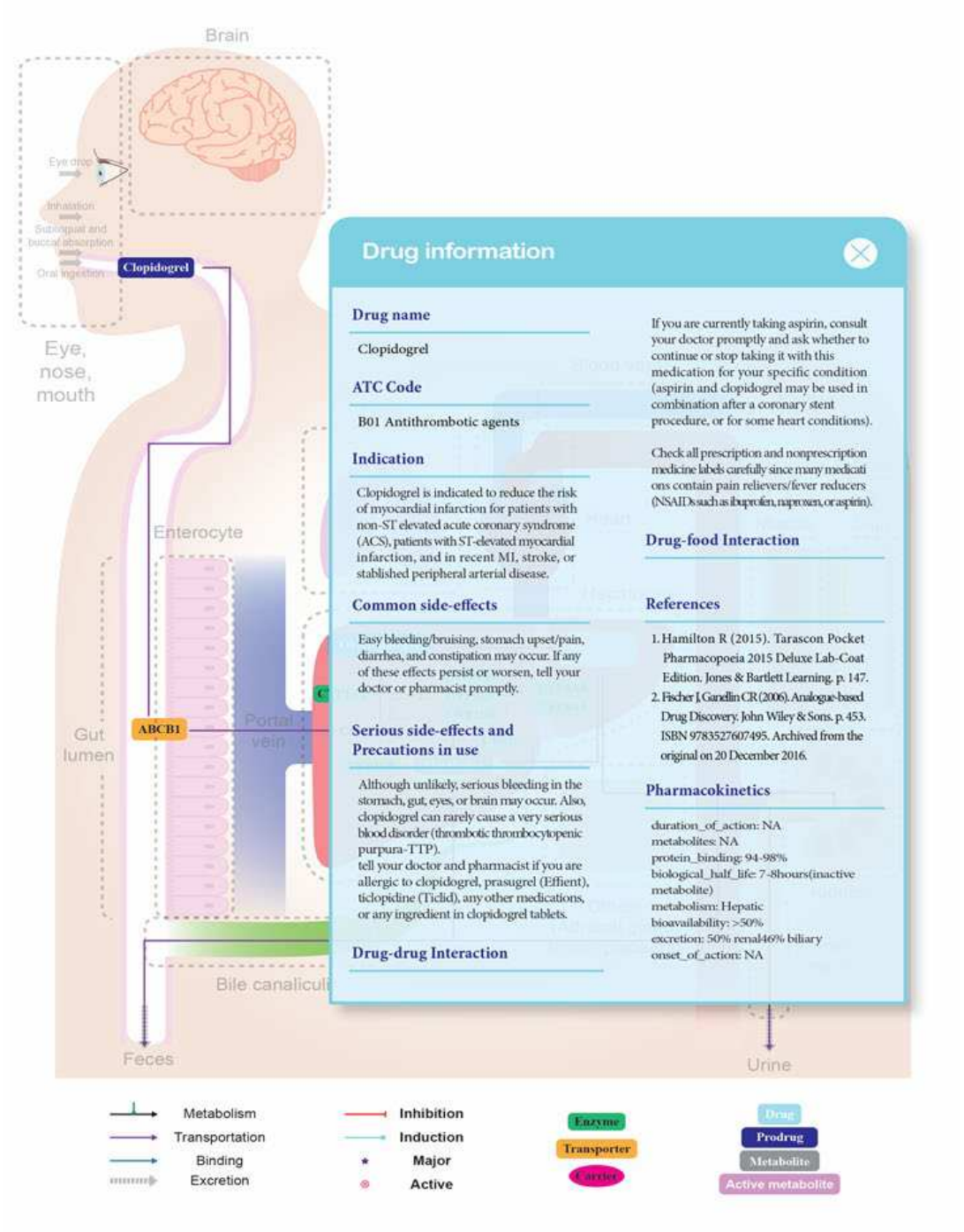


Figure 12. A pop-up window showing clinical information on the drug.[13]

[13] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

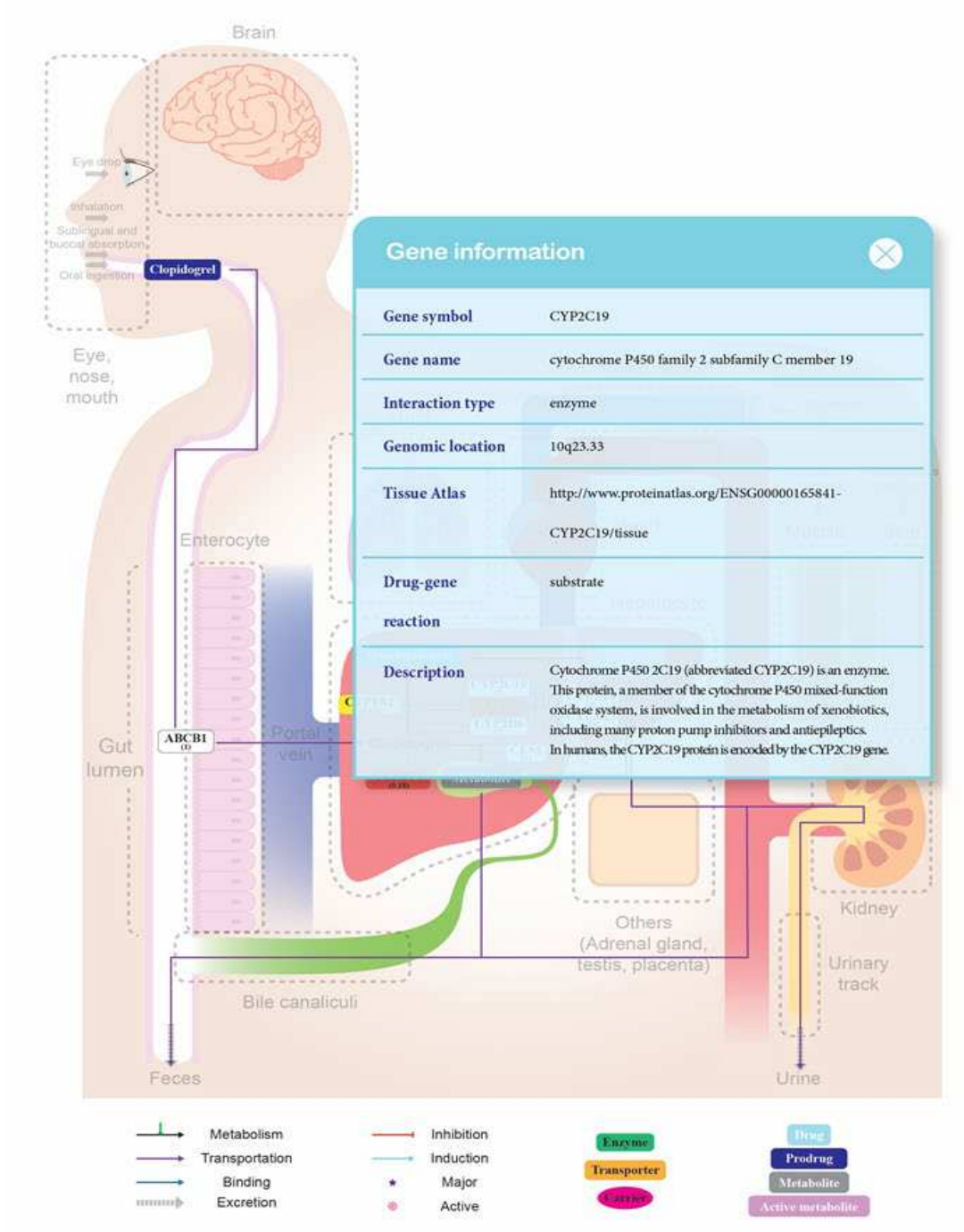


Figure 13. A pop-up window showing information on the gene and its interaction with the administered drug.[14]

[14] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

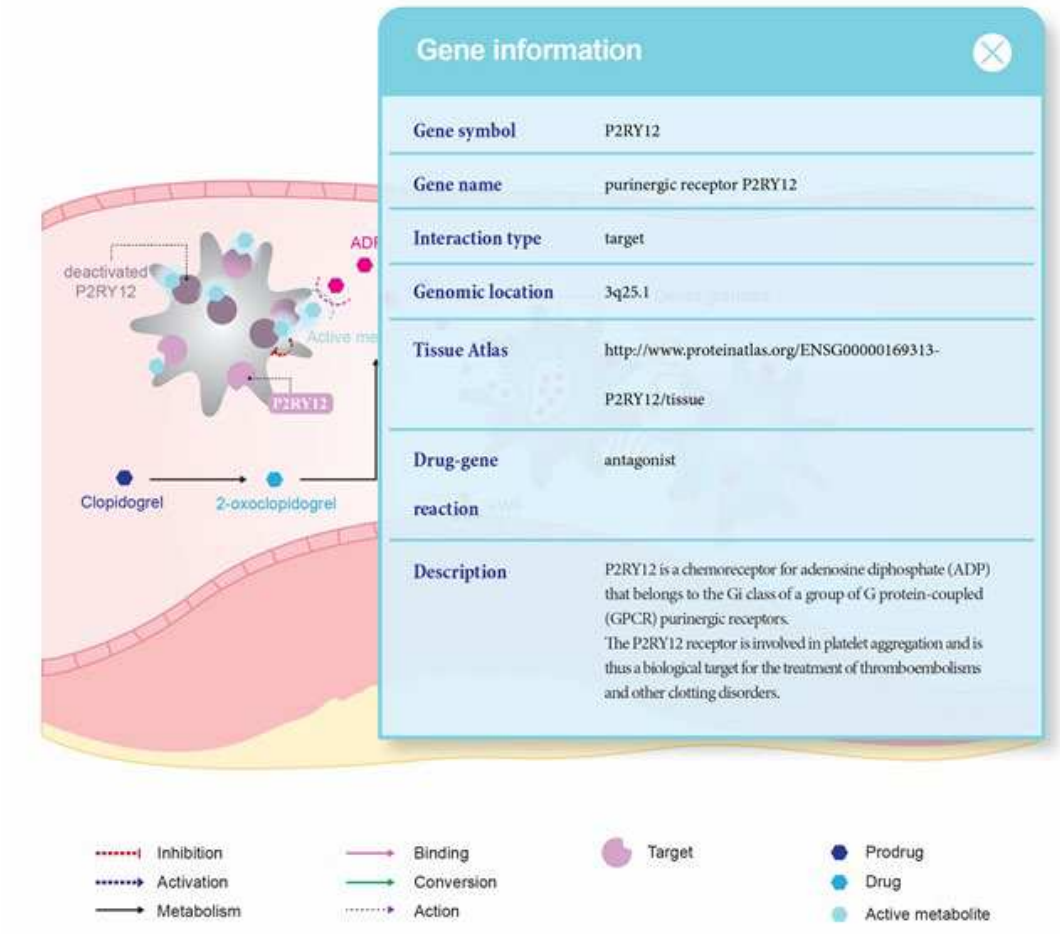


Figure 14. A pop-up window showing information on the gene and its interaction with the administered drug.[15]

[15] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

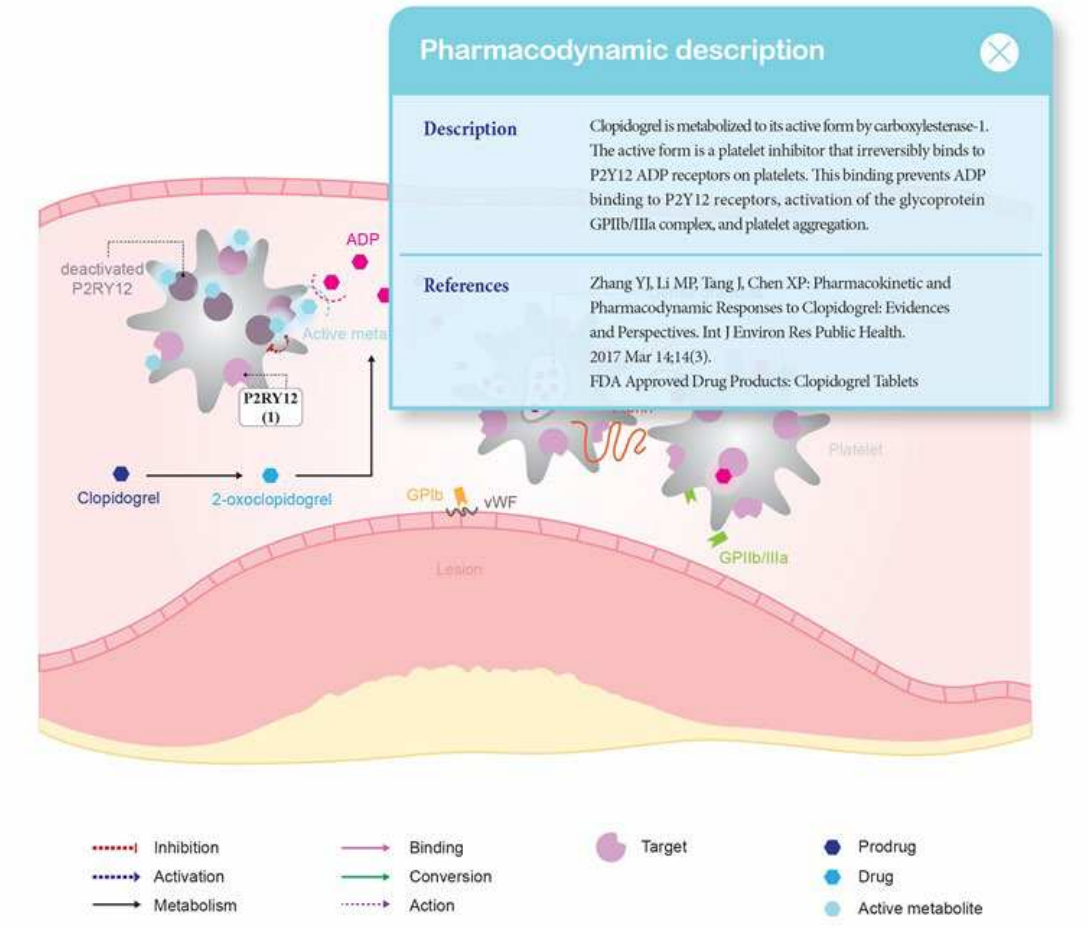


Figure 15. A pop-up window showing PD pathway description.[16]

[16] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

Drawing a pathway with PathVisio

Providing standard terminology and formatting in the order that the prewritten computer programs can transfer the information from the standard format to the local application format, and vice versa, is crucial. Even in PGx research, the importance of formatting is growing. The solution and adoption of the standard format will facilitate easier exchanges between different sources of data [17].

To supply the standard format, PG-paths using PathVisio, a free, open-source, and downloadable pathway software, were built. This software has flexible export formats. By editing the source code, the pathway, which is saved in the Biological Pathway Exchange (BioPAX) level3, GPML, and SVG formats, can be modified to numerous styles or for certain purposes [16]. When pathways are drawn with this software, the attributes (color, size, and labels) of the nodes (genes, drugs, and metabolites) are synchronized to a standard format. Additionally, the attributes (color and arrowheads) of the edges (metabolism, transportation, binding, and excretion) are also synchronized conceptually in agreement with the interaction and action types. Since PathVisio 3.2.4 contains an embedded MIM tool, the interaction and action types can be used consistently with an appropriate purpose [39]. Each pathway is stored in SVG and GPML

formats to modify the source code and relationship with other databases. Thus, the pathway can be utilized for computational analysis or visualization of genomic variation data.

.

Usage of PG-path: Pathway analysis and visualization with GVB score

Out of the existing gene scoring methods, the GVB score method was selected for application to algorithm-based pathway analysis. Since drug-related genes, unlike disease genes, can be presented as a variety of phenotypes in healthy individuals, the case sample was randomly selected from the public database, not considering the specific disease. The NA12878 sample was one of the 2504 samples stored in a public database called the 1000 Genome Project [40]. The GVB scoring method was applied to a set of variants extracted from the Variant Call Format (VCF) file of the NA12878 sample. The calculated GVB scores were annotated to genes on the default pathway to create a personalized pathway. Applying GVB scores to a pathway follows the following three phases: (1) to input the high-throughput biological data (HTBD), (2) to perform the algorithm-based analysis, and (3) to visualize the output data on the pathway [41] (Figure 8A). In the personalized PK and PD pathways, the GVB scores with a color gradient can be displayed on pathway genes. The pathway analysis step was processed between steps 7 and 8 of the default pathway-developing method (Figure 8). The personalized PK and PD pathways, which are generated by displaying the GVB score,

can predict the change in drug concentration in the plasma under the score distribution of the genes in the human body (Figure 16, 17).

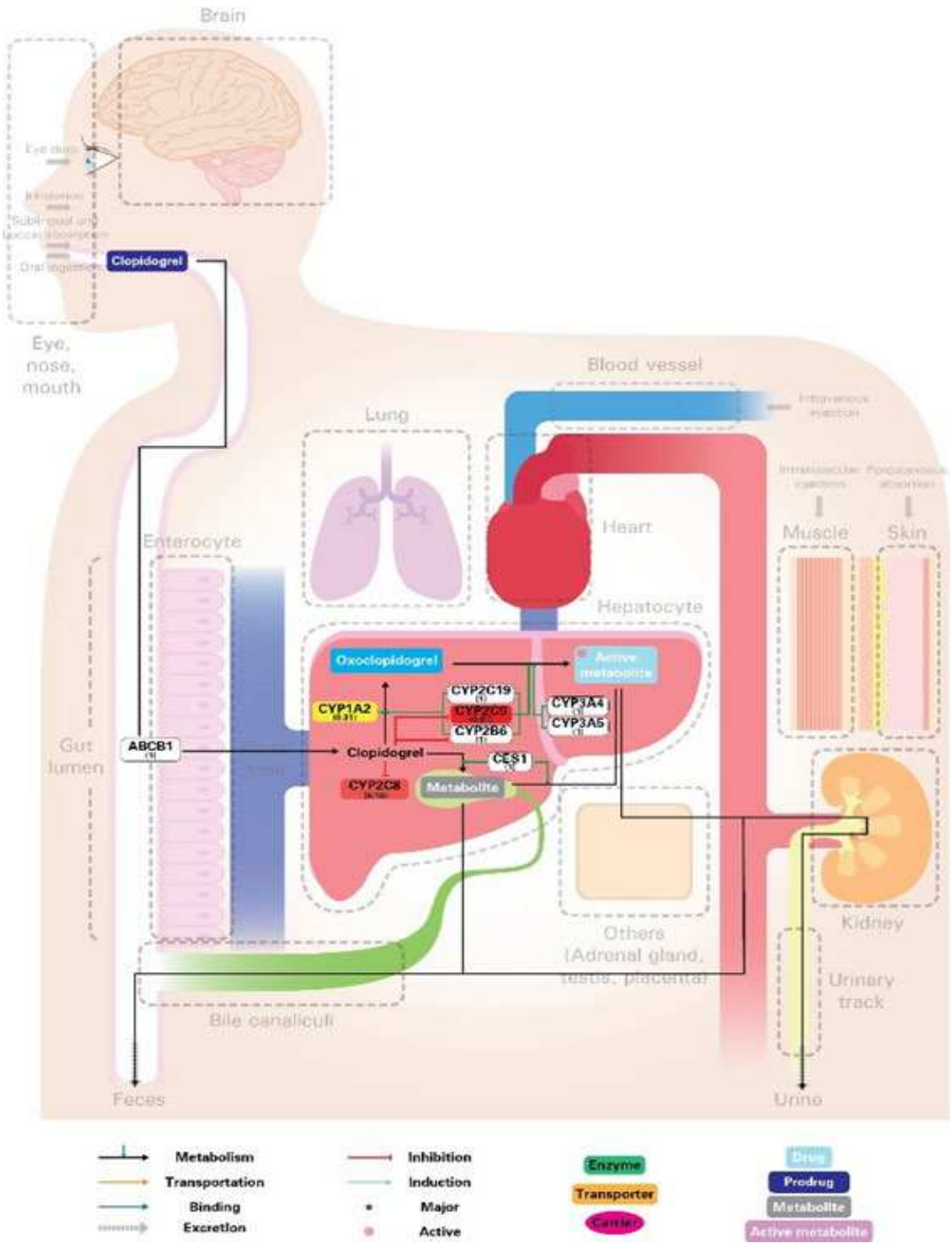


Figure 16. The personalized pharmacokinetic (PK) pathways of clopidogrel.[17]
The personalized PK pathway with GVB scores of the NA12878 sample from the 1000 Genomes Project, GVB: Gene-wise Variant Burden.

[17] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950

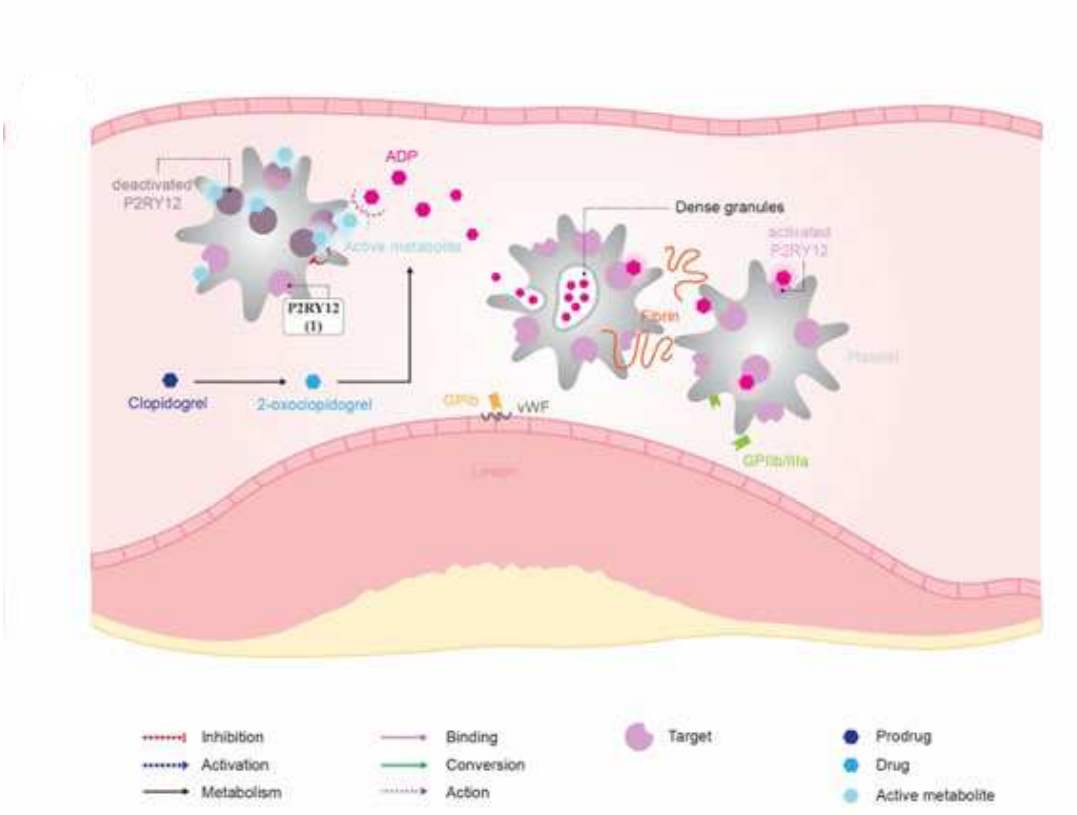


Figure 17. The personalized pharmacodynamic (PD) pathways of clopidogrel.[18] The personalized PD pathway with GVB scores of the NA12878 sample from the 1000 Genomes Project, GVB: Gene-wise Variant Burden.

[18] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950

Discussion

Prediction of changes in plasma drug concentration by gene placement in pathways

The PG-path differs from present-day pathways in that PG-paths indicate the location of the gene in the human body that interacts with the administered drug. To account for changes in plasma drug concentration, describing the gene location of the human body as accurately as possible is necessary, which will consequently make understanding systemic drug flow easier.

Assuming that the body is composed of one compartment and the elimination rate of the drug in the plasma remains constant for ADME of common oral drugs, the change in the concentration of active ingredients in plasma varies from person to person, but the patterns of changes are similar. From a pharmacokinetic perspective, a drug undergoes ADME and reabsorption once it enters the human body. Initially, the bioavailability of the absorption phase determines the quantity of oral drug absorbed by the gastrointestinal tract. Drugs transported into the liver through the portal vein experience the first-pass effect and are primarily metabolized and excreted as inactive metabolites. The remaining active drugs are distributed into

highly perfused tissues, such as the heart, liver, brain, brain, kidney, and red blood cells, or slowly infused tissues, such as the skin, fat, and muscle. After acting at the target site, the drug is metabolized in the liver and excreted as bile or urine. A few of the active drug returns to the plasma via reabsorption in the gastrointestinal or proximal tubules during excretion. However, some drugs present other pharmacokinetic behaviors. The plasma concentration of the active ingredient changes depending on whether the drug is an active drug or prodrug. Generally, active drugs that act at the target site are excreted after being metabolized in the liver. In contrast, prodrugs experience the first-pass effect, i.e., Metabolism, then become activated to exert a medicinal effect at the target site. From a pharmacodynamic perspective, many factors affect drug action. However, only the retention time and concentration of the active ingredient at the target site in connection with pharmacokinetics are considered. From a pharmacogenomics perspective, genetic variations in enzymes, transporters, and carriers have a crucial effect on the concentration of active ingredients in plasma with the usual drug reaction and elimination processes. The location of the impaired protein aids in detecting the problematic area in the ADME process and predicting changes in the drug concentration in plasma [42]

(Figure 18).

Consider that clopidogrel is a prodrug, and the case sample is NA12878 randomly extracted from the 1000Genome project. In the pathway of the NA12878 sample, CYP2C9 has a quite low GVB score among liver enzymes. During the first-pass effect, the GVB score of CYP2C9, the enzyme that metabolizes the prodrug into an active metabolite, is notably lower than other enzymes, reducing the formation of the active ingredient that generates the drug action. Although CYP2C9 is not a major enzyme of clopidogrel, deleterious variations in CYP2C9 can affect metabolic capacity by lowering protein activity. It is predicted that active metabolites of clopidogrel will be produced below the average-maximum concentration in the liver and that the standard drug concentration will not reach the maximum therapeutic level (Figure 16, 17).

Since the PG path specifies the location of the genes interacting with the administered drug, the change in PK properties can easily be predicted. Furthermore, the PG-path with properties of genetic variations makes it possible to predict drug efficacy and toxicity.

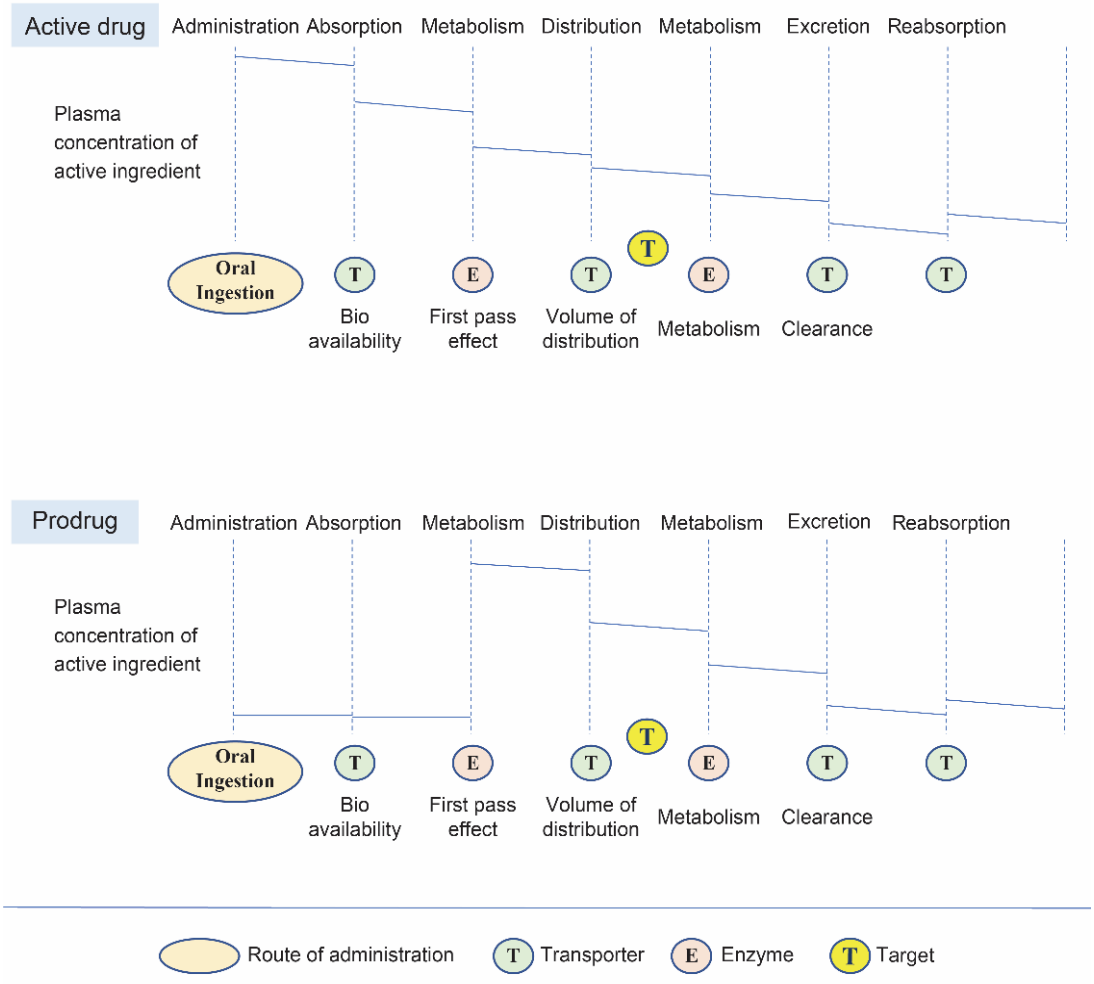


Figure 18. The change in the plasma concentration of active ingredients depends on drug categories.[19]

Assuming that the body is one compartment and the rate of drug elimination in the plasma is constant, when a drug is administered into the human body, it experiences absorption, metabolism, distribution, excretion (ADME), and resorption. In those stages, the plasma concentration of an active ingredient decreases step by step. However, when a prodrug is absorbed, after undergoing metabolism, the plasma concentration of an active ingredient shows, and the next steps occur the same as those of the active drug.

[19] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950

PG-path: Single drug-centered pathway

The PG-path allows visualization of just one drug at a time, considering PK and PD, which indicates that the PG-path does not depict the entire metabolic pathway or the whole signaling pathway. We must focus on the scope of analysis at the individual drug scale because individualized therapy only concentrates on taken drugs [41].

The PG-path can only focus on the genes interacting with a single drug, making the identification of which genes the drug induces or inhibits in the pathway easier. Furthermore, even two drugs taken together can concentrate on the two drug pathways separately, and the drug-drug interaction (DDI) between the two drugs can be considered from a genetic perspective, which permits drug-gene-drug interaction (DGDI) analysis. For example, assuming that Drug A inhibits gene E and Drug B is the substrate of gene E, the metabolism of Drug B slows down if Drug A and Drug B are taken together. In the case of taking multiple drugs, the PG-paths can help predict whether the drug effects will increase or decrease [43] (Figure 19, 20).

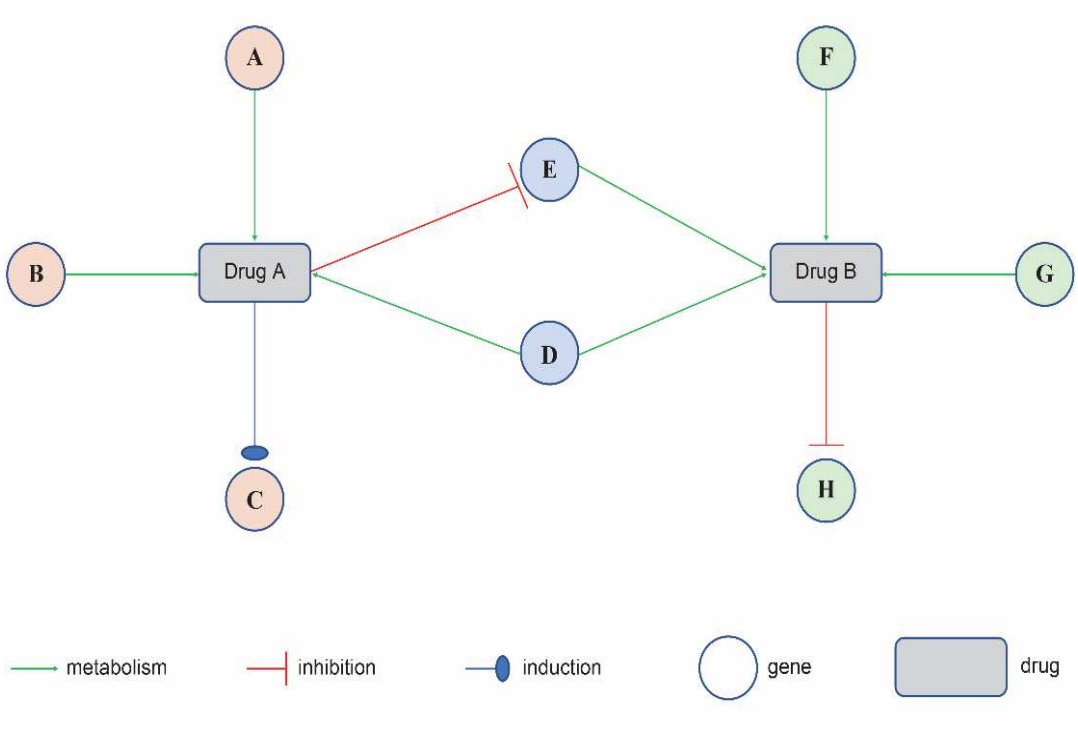


Figure 19. Drug-gene-drug interaction (DGDI) between Drug A and Drug B.[20]
If Drug A inhibits gene E and Drug B becomes the substrate of gene E, the metabolism of Drug B slows down when Drug A and Drug B are taken together.

[20] Hong JY, Kim JH. PG-path: Modeling and personalizing pharmacogenomics-based pathways. *PLoS One*. 2020;15(5): e0230950.

When three drugs, amiodarone, fluoxetine, and codeine, are administered, one of the genes, CYP2D6, is inhibited by fluoxetine and amiodarone. However, codeine is the substrate of CYP2D6, and it is a prodrug of morphine. Since amiodarone and fluoxetine inhibit CYP2D6, the metabolic ability of CYP2D6 when three drugs are administered is lower than that of CYP2D6 when only codeine is administered. Therefore, the prodrug codeine is converted to morphine at the level of lower metabolic capacity. The patients feel less analgesic effect as much as that at the standard states.

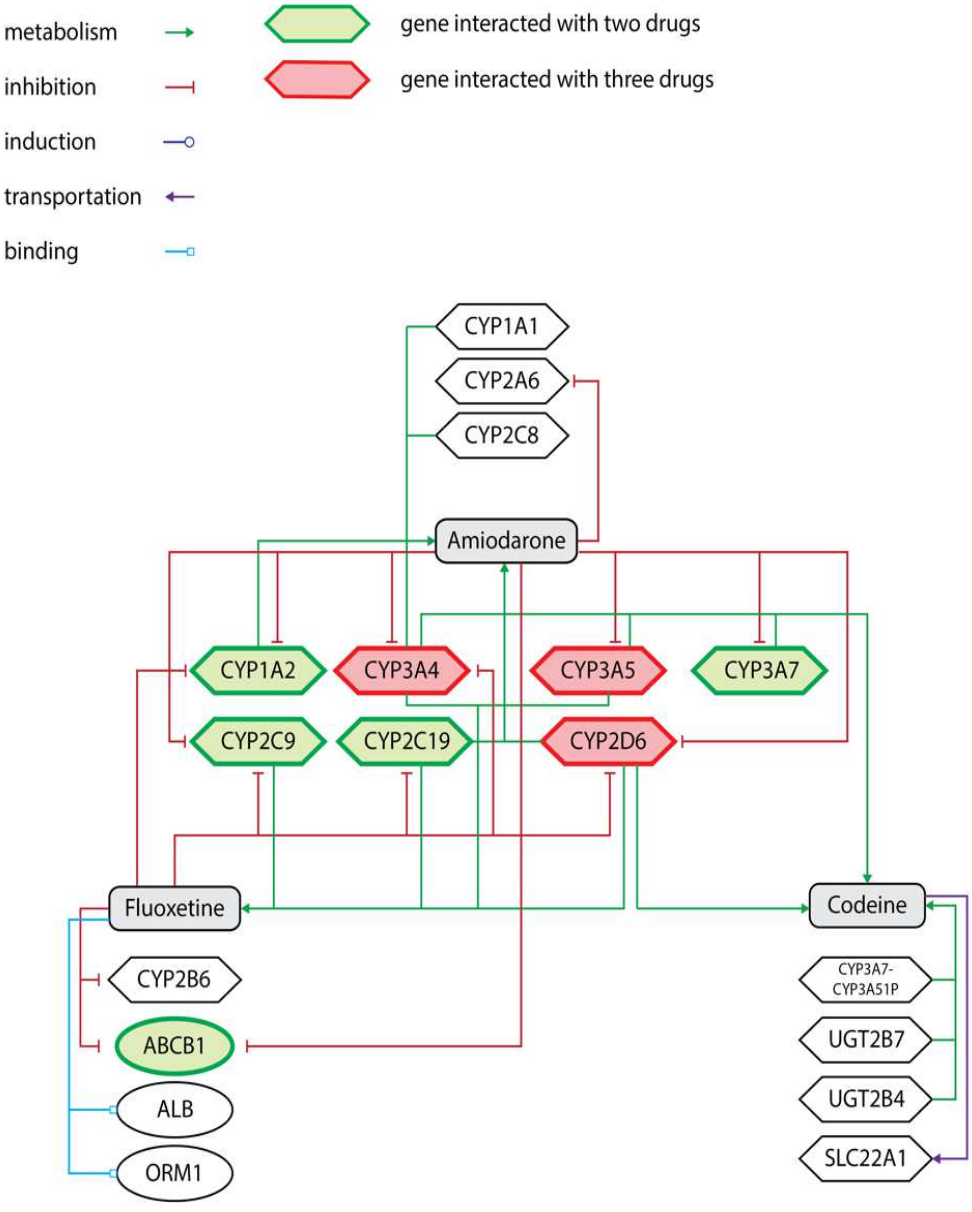


Figure 20. Drug–gene–drug interaction (DGDI) among Amiodarone, Fluoxetine, and Codeine.

Since fluoxetine and amiodarone inhibit CYP2D6 and codeine is the substrate of CYP2D6, the metabolism of codeine would be slowed when codeine is taken together with amiodarone and fluoxetine.

Pathway analysis: GVB scoring method using variations obtained from DNA sequencing

In contrast to preceding analyses using gene expression data resulting from microarrays or data from RNA sequencing [44], PG-path makes use of the GVB scoring method, aggregating scores of variants from the DNA sequencing data acquired through next-generation sequencing (NGS) [45]. Here, this algorithm-based analysis method applies the altered degrees of the activity of protein function encoded by the gene to the gene set that interacts with the drug. The change in drug concentration in plasma can be predicted by identifying the location of the impaired genes. Pathway analysis can convert massive variation data in the genome to changes in drug concentration in plasma, allowing variation information to have meaningful interpretations.

This method has a limitation in that it cannot assure the accuracy of the extracted gene set because there is no gold standard for drug-gene interactions. However, if further experimental or clinical studies are available, this analysis method may provide a more accurate result of analyses [41].

Comparison of variant set between GVB scoring method and CPIC guideline

The GVB scoring method makes use of the result of whole-genome sequencing (WGS) or whole-exome sequencing (WES) of which variant list is from a single cell in the human body. In contrast, CPIC guideline utilizes an array matrix which consists of common restricted variants in a specific pharmaco-gene.

Consequently, this PA can calculate the GVB score, including the rare and novel variants. It is different from CPIC guidelines, which indicate phenotypes and recommendations translated from common variants. It makes us address the wide range of interpretations based on the change of the protein function.

Conclusion

Drug responses vary from person to person, and genetic factors have a significant impact. For the efficient use of drugs, experts in pharmacogenomics have recently made prescription guidelines, such as the CPIC guidelines, based on variations in a specific drug–gene pair in the drug. It can recommend dose adjustment or alternative drugs. However, those guidelines have several problems with clinical applications or regulations. Moreover, since it addresses common variants, the effects of rare variants or novel variants cannot be considered.

To this end, I designed a PG–path that can display the entire genes interacting with the administered drug and modify the attribute of variation for each gene. Applying the GVB score to each gene can predict the change in plasma drug concentration. Since the PG–path is drawn individually for drugs when multiple drugs are administered, it can predict the effect of drug–gene–drug interactions on drug response. Adding PK properties and clinical factors to a PG–path that covers only genetic factors could improve the ability of the PG–path to predict drug response.

Reference

1. Joseph S, Nicolson TJ, Hammons G, Word B, Green-Knox B, Lyn-Cook B. Expression of drug transporters in human kidney: impact of sex, age, and ethnicity. BioMed Central. 2015;6(1).

2. Weinshilboum RM, Wang L. Pharmacogenetics and Pharmacogenomics: Development, Science, and Translation. Annual Review of Genomics and Human Genetics. 2006;7(1):223-45.

3. Mathur S, Sutton J. Personalized medicine could transform healthcare. Biomedical Reports. 2017;7(1):3-5.

4. Relling MV, Klein TE. CPIC: Clinical Pharmacogenetics Implementation Consortium of the Pharmacogenomics Research Network. Clinical Pharmacology & Therapeutics. 2011;89(3):464-7.

5. Pratt VM, Everts RE, Aggarwal P, Beyer BN, Broeckel U, Epstein-Baak R, et al. Characterization of 137 Genomic DNA Reference Materials for 28 Pharmacogenetic Genes. The Journal of Molecular Diagnostics. 2016;18(1):109-23.

6. Lauschke VM, Zhou Y, Ingelman-Sundberg M. Novel genetic and epigenetic factors of importance for inter-individual differences in drug disposition, response and toxicity. Pharmacology & Therapeutics. 2019;197:122-52.

7. Kalman L, Agúndez J, Appell ML, Black J, Bell G, Boukouvala S, et al. Pharmacogenetic allele nomenclature: International workgroup recommendations for test result reporting. Clinical Pharmacology & Therapeutics. 2016;99(2):172-85.

8. Relling MV, Evans WE. Pharmacogenomics in the clinic. Nature. 2015;526(7573):343-50.

9. Klein TE, Ritchie MD. PharmCAT: A Pharmacogenomics Clinical Annotation Tool. Clinical Pharmacology & Therapeutics. 2018;104(1):19-22.

10. Sangkuhl K, Whirl-Carrillo M, Whaley RM, Woon M, Lavertu A, Altman RB, et al. Pharmacogenomics Clinical Annotation Tool (Pharm CAT). Clinical Pharmacology & Therapeutics. 2019.

11. Caudle K, Klein T, Hoffman J, Muller D, Whirl-Carrillo M, Gong L, et al. Incorporation of Pharmacogenomics into Routine Clinical Practice: the Clinical Pharmacogenetics Implementation Consortium (CPIC) Guideline Development Process. Current Drug Metabolism. 2014;15(2):209-17.

12. Bank P, Caudle K, Swen J, Gammal R, Whirl-Carrillo M, Klein T, et al. Comparison of the Guidelines of the Clinical Pharmacogenetics Implementation Consortium and the Dutch Pharmacogenetics Working Group. Clinical Pharmacology &

Therapeutics. 2018;103(4):599-618.

13. Wouden CH, Van Rhenen MH, Jama WOM, Ingelman-Sundberg M, Lauschke VM, Konta L, et al. Development of the PG x-Passport: A Panel of Actionable Germline Genetic Variants for Pre-Emptive Pharmacogenetic Testing. Clinical Pharmacology & Therapeutics. 2019;106(4):866-73.

14. Van Der Wouden C, Cambon-Thomsen A, Cecchin E, Cheung K, Dávila-Fajardo C, Deneer V, et al. Implementing Pharmacogenomics in Europe: Design and Implementation Strategy of the Ubiquitous Pharmacogenomics Consortium. Clinical Pharmacology & Therapeutics. 2017;101(3):341-58.

15. Sato M, Hosoya S, Yoshikawa S, Ohki S, Kobayashi Y, Itou T, et al. A highly flexible and repeatable genotyping method for aquaculture studies based on target amplicon sequencing using next-generation sequencing technology. Scientific Reports. 2019;9(1).

16. Kutmon M, Van Iersel MP, Bohler A, Kelder T, Nunes N, Pico AR, et al. PathVisio 3: An Extendable Pathway Analysis Toolbox. PLOS Computational Biology. 2015;11(2):e1004085.

17. Thorn C, Whirl-Carrillo M, Klein T, Altman R. Pathway-based Approaches to Pharmacogenomics. Current Pharmacogenomics. 2007;5(1):79-86.

18. Sousa M, Pozniak A, Boffito M. Pharmacokinetics and pharmacodynamics of drug interactions involving rifampicin, rifabutin and antimalarial drugs. Journal of Antimicrobial Chemotherapy. 2008;62(5):872-8.

19. De Vries Schultink AHM, Zwart W, Linn SC, Beijnen JH, Huitema ADR. Effects of Pharmacogenetics on the Pharmacokinetics and Pharmacodynamics of Tamoxifen. Clinical Pharmacokinetics. 2015;54(8):797-810.

20. Franconi F, Campesi I. Pharmacogenomics, pharmacokinetics and pharmacodynamics: interaction with biological differences between men and women. British Journal of Pharmacology. 2014;171(3):580-94.

21. Drozda K, Pacanowski MA, Grimstein C, Zineh I. Pharmacogenetic Labeling of FDA-Approved Drugs. JACC: Basic to Translational Science. 2018;3(4):545-9.

22. Breitling R. What is systems biology? Frontiers in Physiology. 2010;1.

23. Rajasethupathy P, Vayttaden SJ, Bhalla US. Systems modeling: a pathway to drug discovery. Current Opinion in Chemical Biology. 2005;9(4):400-6.

24. Soukup T, Lamb B, Arora S, Darzi A, Sevdalis N, Green J.

Successful strategies in implementing a multidisciplinary team working in the care of patients with cancer: an overview and synthesis of the available literature. Journal of Multidisciplinary Healthcare. 2018;Volume 11:49-61.

25. Caraballo PJ, Hodge LS, Bielinski SJ, Stewart AK, Farrugia G, Schultz CG, et al. Multidisciplinary model to implement pharmacogenomics at the point of care. Genetics in Medicine. 2017;19(4):421-9.

26. Frolkis A, Knox C, Lim E, Jewison T, Law V, Hau DD, et al. SMPDB: The Small Molecule Pathway Database. Nucleic Acids Research. 2010;38(suppl_1):D480-D7.

27. Barbarino JM, Whirl-Carrillo M, Altman RB, Klein TE. PharmGKB: A worldwide resource for pharmacogenomic information. Wiley Interdisciplinary Reviews: Systems Biology and Medicine. 2018;10(4):e1417.

28. Jewison T, Su Y, Disfany FM, Liang Y, Knox C, Maciejewski A, et al. SMPDB 2.0: Big Improvements to the Small Molecule Pathway Database. Nucleic Acids Research. 2014;42(D1):D478-D84.

29. Thorn CF, Klein TE, Altman RB. PharmGKB: The Pharmacogenomics Knowledge Base. Humana Press; 2013. p. 311-

20.

30. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Research. 2018;46(D1):D1074-D82.

31. Ng PC. Accounting for Human Polymorphisms Predicted to Affect Protein Function. Genome Research. 2002;12(3):436-46.

32. Ng PC. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Research. 2003;31(13):3812-4.

33. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature Protocols. 2009;4(7):1073-81.

34. Sim N-L, Kumar P, Hu J, Henikoff S, Schneider G, Ng PC. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Research. 2012;40(W1):W452-W7.

35. Consortium TGP. A map of human genome variation from population-scale sequencing. Nature. 2010;467(7319):1061-73.

36. Devuyst O. The 1000 Genomes Project: Welcome to a New World. Peritoneal Dialysis International. 2015;35(7):676-7.

37. Seo H, Kwon EJ, You Y-A, Park Y, Min BJ, Yoo K, et al. Deleterious genetic variants in ciliopathy genes increase risk of ritodrine-induced cardiac and pulmonary side effects. BMC Medical

Genomics. 2018;11(1).

38. Park Y, Kim H, Choi JY, Yun S, Min B-J, Seo M-E, et al. Star Allele-Based Haplotyping versus Gene-Wise Variant Burden Scoring for Predicting 6-Mercaptopurine Intolerance in Pediatric Acute Lymphoblastic Leukemia Patients. Frontiers in Pharmacology. 2019;10.

39. Luna A, Sunshine ML, Van Iersel MP, Aladjem MI, Kohn KW. PathVisio-MIM: PathVisio plugin for creating and editing Molecular Interaction Maps (MIMs). Bioinformatics. 2011;27(15):2165-6.

40. Zheng-Bradley X, Flicek P. Applications of the 1000 Genomes Project resources. Briefings in Functional Genomics. 2016:elw027.

41. García-Campos MA, Espinal-Enríquez J, Hernández-Lemus E. Pathway Analysis: State of the Art. Frontiers in Physiology. 2015;6.

42. Russ B. Altman DAF, David B. Goldstein. Principles of Pharmacogenetics and Pharmacogenomics. Cambridge University Press: British Library; 2012.

43. Tannenbaum C, Sheehan NL. Understanding and preventing drug–drug and drug–gene interactions. Expert review of clinical pharmacology. 2014;7(4):533-44.

44. Ge SX, Son EW, Yao R. iDEP: an integrated web application for differential expression and pathway analysis of RNA-Seq data. BMC Bioinformatics. 2018;19(1).

45. Behjati S, Tarpey PS. What is next generation sequencing? Archives of disease in childhood - Education & practice edition. 2013;98(6):236-8.

국문 초록

서 론: 인체의 약물 분해 효소나 약물 수송체에서의 유전적 변이는 약물의 농도에 변화를 일으켜, 약물 반응의 개인차를 나타나게 한다. 따라서 외부에서 약물의 농도를 조절하거나 약물을 대체하여, 인체 내에서 최적의 약물 농도가 되도록 맞추어 주어야 한다. 이를 통해 약물의 효능과 안전성의 극대화를 추구할 수 있다. 현재의 해결 방안으로 약물 유전학에 근거한 CPIC 가이드라인이 출판되고 있다. 그러나 이 가이드라인들은 임상 적용과 규제, 또는 시스템 개발에서의 표준화 등에 문제점을 가지고 있다. 이에 약물과 유전자들 사이의 상호작용을 시각화하고 변이 데이터를 적용하여, 인체 내 약물 농도 변화를 예측할 수 있는, 약물 유전체학 패스웨이인 PG-path 를 제작해 보고자 한다.

방 법: DrugBank 에서 특정 약물과 유전자 셋을 추출하고, 그들 간의 상호작용 유형(효소, 수송체, 수용체)과 작용 유형(억제, 유도, 기질)을 추출한다. 전신 수준에서 흡수, 분포, 대사, 배설 과정을 다루는 약동학적 패스웨이와, 분자-세포 수준에서의 약물 작용을 다루는 약력학적 패스웨이로 나누어, 틀과 배경 그림을 제작한다. 그리고, DrugBank 에서 추출한 요소들을 제작된 틀에 적용하여, 약물과 유전자의 상호 작용을 시각화 한다. 성분 요소와 포맷의 표준화를 위해 공용 소프트웨어인 PathVisio 를 사용하였다.

결 과: PG-path 는 인체에 존재하는 장기를 표현한 하나의 틀과, 각 장기들의 배경 이미지로 구성된다. 약물과 유전자 사이의 상호작용 유형과 작용 유형을 표준 포맷에 맞게 틀 위에 그려 주고, 다이어그램이 그려진 틀과 배경 그림을 병합하였다. 유전자와 약물 정보를 담은 윈도우를 패스웨이와 연결하고 팝업 가능하게 함으로써, 패스웨이의 이해도를 높이고자 하였다. 각 유전자를 인체와 비슷한 위치에 배치하여, 약물의 인체 내에서 흐름을 시각화 하였다. 그리고, 각 유전자에 유전자의 누적 손상 정도를 나타내는 GVB 점수를 적용하여 개인별 약물 유전체학 기반의 패스웨이를 만들어 보았다.

결 론: PG-path 는 인체 내 약물의 일반적 흐름을 알 수 있게 제작되었다. GVB 점수를 각 유전자에 적용함으로써, 혈장 내 약물의 농도 변화를 예측할 수 있다. 또한, PG-path 는 약물 개별로 그려졌기 때문에, 여러 약물을 투약하였을 때, 약물-유전자-약물 간의 상호작용이 약물 반응에 미치는 영향을 예측할 수 있다. 유전 요소만을 다룬 PG-path 에 약동학적 요소나 임상적 요소 등이 추가되면, PG-path 의 약물 반응 예측 능력은 더 고도화될 수 있을 것이다.

주요어: 패스웨이, 약물 유전체학, 약물 유전학, 약물 반응, 맞춤약물, 약물-유전자-약물 상호작용, 약동학, 약력학

학 번: 2016-30616