Journal of Clinical and Nursing Research

Review Article

BIO-BYWORD SCIENTIFIC PUBLISHING PTY LTD

Precision Medication of Tacrolimus based on the Genotype by Using Second-generation Sequencing

Yumeng Li*

Northeast Yucai Foreign Language School

Abstract: Organ transplantation has become a powerful strategy for the treatment of malignant diseases. Nevertheless, graft rejection is one of the main factors affecting graft survival after organ transplantation. Under this circumstance, the transplant-related mortality still keeps up. This invention includes the precise medication guidance of Tacrolimus (FK506) inapplicable population, against the side-effects of this drug. This invention, based on second-generation sequencing, has the advantages of relatively low cost and high sequencing throughput. During the design process, we collect the data of single nucleotide polymorphism (SNP) concerning the adverse drug reactions of Tacrolimus. Then we filter and summarize fifteen SNPs basing on importance degree (level >key enzyme>race). Thenceforth, after the process of analyzing the raw extract by operating BWA, Picard-tools, GATK, and Perl, we annotate SNPs by Annovar. Through this innovation, people can obtain further feedback on drugs that targets different genes in order to achieve the purpose of precision medication and minimizing the risks of misusing Tacrolimus.

Publication date: January, 2020

Publication online: 31 January 2020

Corresponding author: Yumeng Li, liaoquanneng@ivygate.cn

1 Background

Tacrolimus, a kind of immunosuppressive medicine, is used for transplantation rejection. Compared to the effects of other immunosuppressive drugs, its pharmacological effects are much more potent, and its rate of rejection is lower. Specifically, Tacrolimus is 100 times stronger than cyclosporine (CsA) in inhibiting T lymphocytes. However, its absolute bioavailability can be affected by many factors, such as age, food, disease. FKBP is one of these essential factors. FK506 binding protein, a highly conserved protein, is a vital binding partner of Tacrolimus. It functions as a molecular switch by directly binding to the substrate or changing the conformation of the target proteins. Tacrolimus binds to the FKBP12 after entering the body and inhibits the PPIase activity. The FK506-FKBP12 complex works as the competitive inhibitor, which allows the drug interacts with the calcineurin(CaN) which is an essential component for T-cell activation. As a result, the nuclear translocation and transcription of T-cell activation genes will be prevented.

Tacrolimus is always absorbed orally in the gastrointestinal tract and is metabolized in the liver, mainly via CYP3A. CYP3A5 is one of the predominant enzyme responsible for the metabolism of Tacrolimus. Genetic variations within CYP3A5 lead to changes to the activity of the CYP3A5 protein. As a result, concentrations of Tacrolimus are affected in the body. Moreover, individuals who have homozygous for the G allele have higher concentrations of Tacrolimus, which means that these people require lower doses of the Tacrolimus. Some researches have shown that genetic polymorphisms of genes can significantly influence pharmacokinetics and pharmacology of Tacrolimus.

The polymorphism in the CYP gene performs a vital role. For instance, CYP3A5 enzyme is an enzyme which belongs to cytochrome P450 family, and it is involved in the metabolism of the drug and the steroid hormones testosterone and progesterone and synthesis of cholesterol, steroids and other lipids. Recently, CYP3A5 have already been checked by the CPIC and the DPWG. They both illustrated that the demand for Tacrolimus needed to increase for patients who have CYP3A5. In the project, Illumina Miseq is used to

detect sequences in advantages of a short time, high flask, and low cost. The reads chosen are based on level, key enzyme, and race. The reads are highly supported by pieces of evidence and critical vital enzymes. Besides, we test not only CYP3A5 but also the other genes which may help us find the new variant. Moreover, the purpose of this project is to treat patients accurately not only by maximizing the drug effects but also by minimizing ADRs.

Due to the genetic polymorphism, different drugs may have a distinct influence on patients. Accurate classification of patients based on gene diagnosis and development of personalized medicine will not only result in reducing the risk of potential diseases but also can provide more efficient and specific plans. As a result, it will be more likely to maximum the curing effect and minimize the side- effects. In this project, we do the research and filter out 15 genes related to Tacrolimus, and by using the Illumina miseq. We can achieve the purpose of personalized medicine, which can realize the precision medication.

2 Methods

2.1 Panel design

First of all, we use the literature from the official website like PharmGKB, Pubmed, and Grugbank, and rank them based on the level. Then we collect 32 SNPs and arrange them in a file by listing chromosome where they located, start and end of GRCh37.p10, ref, alt, rs, level, gene, and biogeographic groups. In pharmGKB, it clearly illustrates the gene or polymorphisms, which is related to the Tacrolimus and the alleles and phenotype of each gene.

Secondly, we filter out 15 SNPs based on the level, key enzyme, and race. We first choose to use the gene with the level 1,2,3 because level 1 means that the locus is most highly supported by sufficient evidence and the reliability decreases in order. After that, the gene which contains the key enzyme should be included because key enzyme may affect the digestion of Tacrolimus. After that, we decide to do researches of one population, East Asian. Meanwhile, we base on the order (level>key enzyme>race). To be more specific, even though biogeographic groups of some genes are not East Asian, we still need to choose them because they have level 1,2 or key enzyme or both.

Finally, the information of each gene in the file will be compared with reference and will do further annotation by using the BWA, Picard-tool, GATk, and Perl.

2.2 Procedures:

(1) Quality evaluation

The quality of sequence outcome is undeniably paramount because it directly influences the quality of the project. In the study, we use Fastqc, which effectively test sequences with every four lines a group and remove joint sequences and low-quality sequences, to assure high-qualified sequences. We also use GATK to estimate library and site coverage depth.

(2) Mapping

Burrows-Wheeler Aligner (BWA) is used in this step for Genome alignment. In order to run script files, firstly, we create index files for reference sequences. Secondly, we sequence alignment and output alignment results - SAM (sequence alignment/map format) file. Thirdly, we describe the SAM file by commenting on header information and comparing the consequences. Finally, we acquire the elementary and more detailed SAM file.

(3) Sort and mark duplicate

This is the first step of preprocessing, and we use Picard-tools. We sorted Sam files, generating into bam files. Next, we classify the items in bam files corresponding to the same chromosome in Sam files from smallest to largest in coordinate order and mark the duplicate.

(4) Indel Realign

This is the second step of preprocessing, and GATK is used to mark areas which need to be rematched and compare each of them. Finally, we acquire sites where evidence suggests a hidden indel as well as indels seen in original alignments(in CIGARs)

(5) Base Revalibrator

This is the last step of preprocessing, also using GATK. Initially, we report quality score and establish models for quality value with the position within the read as well as the preceding and current nucleotide observed by the sequencing machine. Next, we calibrate the quality value based on the models and print the reads.

(6) Calling

In this step, haplotype caller (GATK) is used to detect and recognize SNV and Indel. It can determine if a region has the potential to be variable, construct a de Bruijn assembly of the region, provide the path that are potential haplotypes needed to be evaluated, calculate haplotype likelihoods, determine if there are

any variants on the most likely haplotypes, compute the allele frequency distribution to determine most likely allele count and emit a variant call if determined.

(7) Filtering

This step requires GATK to extract the SNPs from the call set, apply the filter to the them, extract the Indels from the call set, apply the filter to the Indel call set, chombine SNP and indel call set and get the passed call set(refilter). Finally, we achieve a set of concise and available data.

(8) Annotation

We use Annovar in this step to annotate single nucleotide variants (SNVs) and insertions/deletions, gene-based, region-based, filter-based variants or even custom annotations, only printing the worst consequence. After running the process, we get the vcf result file. Then we dispose the file by using Perl and get the necessary information.

3 Conclusions

Tacrolimus is a kind of immunosuppressive drug used for reducing transplantation rejection. With this invention, we can lower drug-induced risks when patients have immunological rejection after liver or kidney transplantation. We search and read the information to find SNPs related to the Tacrolimus by using some databases, such as PharmGKB, PubMed, and DrugBank. Then, the SNPs are filtered by the following criteria: first, SNPs should have enough evidence, which means the level of each chosen SNP should be smaller than 4; secondly, we give priority to the SNP whose gene is the key enzyme; lastly, SNPs based on East Asian are kept. Eventually, fifteen SNPs are chosen to make up a panel. We apply FastQC, GATK, BWA, Picard-tools, and Annovar to sequence the exome and compile a Perl script for the whole analysis process.

The race of some SNPs is mixed population, rather than East Asian. As a result, this research is less targeted to East Asian. The action we should do is to enhance this researcher's pertinence.This research is based on the Linux operation system. All processes of the exome analysis should be completed in the Linux environment. However, many laboratories cannot meet this requirement. Eventually, by using the designed Perl script, this invention can identify whether patients have SNPs, which may cause ADRs when patients take Tacrolimus. As a result, the ADRs of Tacrolimus can be reduced.