COVER SHEET

TITLE: FCGR3A Variability: Copy Number, Single Nucleotide Polymorphisms, and Efficacy of Immunotherapy

AUTHOR'S NAME: Mikayla A. Gallenberger

MAJOR: Genetics

DEPARTMENT: Human Oncology

MENTOR: Dr. Paul Sondel

DEPARTMENT: Pediatrics/Human Oncology

MENTOR(2): Dr. Amy Erbe-Gurel

DEPARTMENT(2): Human Oncology

YEAR: 2016

(The following statement must be included if you want your paper included in the library's electronic repository.)

The author hereby grants to University of Wisconsin-Madison the permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part in any medium now known or hereafter created.

ABSTRACT

FCGR3A Variability: Copy Number, Single Nucleotide Polymorphisms, and Efficacy of Immunotherapy

In order to develop personalized immunotherapies for diseases like cancer, a key component is to understand the profiles of specific genes that may influence the response to treatment. For example, genes like the Fc gamma receptors (*FCGR*) are subject to genetic variations that alter the protein product, resulting in differential impact on the degree of response to immunotherapies. Some *FCGRs* are copy number variable (CNV), while others have single nucleotide polymorphisms (SNP); this combination adds complexity to the genotyping process. We adapted reference-query pyrosequencing (RQPS) to simultaneously genotype the SNP and CNV of *FCGR3A*. After further troubleshooting, we will utilize RQPS to genotype Renal Cell Carcinoma (RCC) clinical trial patients and delineate the relationship between their genotypes and responses to current immunotherapy treatments in order to help clinicians decide which drugs to prescribe to future patients.

Mikayla A. Gallenberger/Genetics

Author Signature

4/27/2016

Date

Dr. Paul Sondel/Pediatrics

Mentor Signature

RECEIVED

MAY - 2 2016

FCGR3A Variability: Copy Number, Single Nucleotide Polymorphisms, and Efficacy of Immunotherapy

Mikayla A. Gallenberger
Supported by Dr. Amy Erbe-Gurel, Wei Wang, and Dr. Paul Sondel

Abstract

In order to develop personalized immunotherapies for diseases like cancer, a key component is to understand the profiles of specific genes that may influence the response to treatment. For example, genes like the Fc gamma receptors (*FCGR*) are subject to genetic variations that alter the protein product, resulting in differential impact on the degree of response to immunotherapies. Some *FCGRs* are copy number variable (CNV), while others have single nucleotide polymorphisms (SNP); this combination adds complexity to the genotyping process. We adapted reference-query pyrosequencing (RQPS) to simultaneously genotype the SNP and CNV of *FCGR3A*. After further troubleshooting, we will utilize RQPS to genotype Renal Cell Carcinoma (RCC) clinical trial patients and delineate the relationship between their genotypes and responses to current immunotherapy treatments in order to help clinicians decide which drugs to prescribe to future patients.

Introduction

Classical patterns of genomic inheritance predict the presence of two copies of each gene per organism, one inherited from each parent. Yet, this is not true genome-wide. Genes can be absent, giving a copy number of zero, or they can be duplicated, with copy numbers typically ranging from two to six in mammalian organisms [1]. This phenomenon is called copy number variation (CNV), and can be caused by mutations, deletions, and insertions [2]. Not all genes have regions subject to CNV, but those that are of great interest for the study of their altered protein expression levels, effects of disease susceptibility, and response to treatments.

One example of clinical significance of CNV includes some of the genes belonging to the Fc Gamma Receptor (FCGR) family: *FCGR3A, FCGR2C,* and *FCGR3B* [2]. Copy number variation at the FCGR locus includes FCGR3A, FCGR2C and FCGR3B but not FCGR2A and FCGR2B [2]. The FCGRs are variably expressed on immune cells, including natural killer cells (NK cells), neutrophils, monocytes, and macrophages. The FCGRs are only expressed on cells of the innate immune system, but they provide a link to the adaptive immune system by physically

connecting the antibodies of the innate system to the effector molecules of the adaptive system [3]. Thus, their overwhelming influence on the entire immune system makes the optimization of the connection between the receptor and the antibody crucial. NK cells are thought to be the main immune cells that contribute to Antibody-Dependent Cell-mediated Cytotoxicity, and they express both *FCGR3A* and *FCGR2C* [2].

In addition to being subject to CNV, the *FCGRs* are also affected by single nucleotide polymorphisms (SNP). SNPs are nucleotide variations at a single position in a DNA sequence among individuals. The avidity and affinity of FCGRs for antibodies can vary based on their SNP genotypes and CNV, and this group of receptors is in a region particularly prone to CNV. *FCGR3A* is affected by CNV at amino acid position 158 in 4-16% of the population [4]. The SNP on *FCGR3A* at amino acid position 158 can alter the functional amino acid from a valine (V) to a phenylalanine (F), varying receptor affinity for the IgG antibody [5]. It is estimated that in Caucasian populations, the homozygous V/V genotype occurs about 20% of the time, while in Asian populations, the number is closer to 40% [4].

SNPs in certain *FCGRs* have been described to cause differential responses immunotherapy, including *FCGR2A* and *FCGR3A* [6-10]. An immunotherapeutic study showed patients with a homozygous V/V genotype at the SNP in position 158 in *FCGR3A* responded significantly better to Rituximab treatment of non-Hodgkin's Lymphoma, which is the expected result because this genotype promotes a higher affinity for the IgG antibodies [5].

Given the prevalence of CNV and SNPs within *FCGR3A* and its significance for human immune response, a method for genotyping the variations is imperative. Currently, most researchers use quantitative PCR with Taqman probes to determine the copy number and SNP type individually. However, this method does not allow for the simultaneous genotyping of both

SNP and CNV, and thus can lead to inaccuracies. Additionally, it is also susceptible to variability within samples, and thus often requires duplicates of each sample for both CNV and SNP assays.

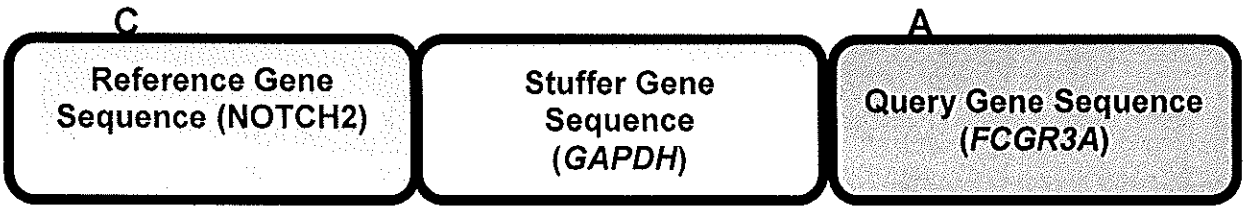
To combat the complexities of the standard Taqman PCR genotyping reactions, a new method was developed by Liu *et al.*: Reference Query Pyrosequencing (RQPS) [11,12]. RQPS is designed to produce the same results as the Taqman Real Time PCR test, but allows for simultaneous testing, thus increased accuracy of data interpretation. This new technique compares the known copy number and SNP location of a 'reference gene' to that of the unknown sequence being studied. This is accomplished by combining the reference and query genes in the same 'probe', a synthesized sequence of DNA that provides a baseline level of expression for each nucleotide, so that the patient sample can be compared to the probe. The probe is then later combined in a single well with genomic DNA, and a software program analyzes the copy number and SNP type of the test genome by comparing it to that of the known sequence. Given its decreased cost and increased efficiency, we are optimizing RQPS to simultaneously genotype the SNPs and CNV of *FCGR3A*, as a potential alternative.

Methods

This RQPS protocol for the FCGRs is adapted from the method published by Liu et al., 2010, to separately include the SNPs within the *FCGR3A* gene [11]. Further information on this method is also described in a paper published by the same authors [12]. We created a probe for *FCGR3A* through sequential PCR reactions, linking DNA sequences from three different genes, followed by cloning the linked DNA sequences into a plasmid. A graphic depiction of the probe is shown in Figure 1. The purpose of this is to provide a baseline level of expression to which the genomic sample can be compared. The generated probe is and mixed with genomic DNA. We used the mixed DNA in a pyrosequencing analysis to determine allelic quantifications for the

RQPS probe vs. genomic DNA of each of the sequences (query and reference) that were inserted into the RQPS probe. To validate the RQPS results, we separately assayed the SNPs and CNV of twenty-four samples from healthy donors by Taqman probes using Real Time PCR.

Figure 1. The probe sequence created for *FCGR3A* connects the reference gene *NOTCH2*, the stuffer sequence from *GAPDH*, and the query sequence from *FCGR3A* together. Also included within the sequence are 'artificial SNPs' within *NOTCH2* and *FCGR3A* of 'C' and 'A' nucleotides, respectively. The purpose of the artificial SNPs is to provide a reference to which the expression levels of the healthy donor sequence can be compared.



Probe Design for FCGR3A:

We designed two DNA sequences to be synthesized and used in the RQPS probe: A *NOTCH2* sequence with an artificial 'C' SNP and an *FCGR3A* sequence with an artificial 'A' SNP. The nucleotide at the *FCGR3A* SNP position is either T or G, which corresponds to amino acid position 158 F/V, so the genomic genotype can be compared to the probe via the pyrosequencing reaction. We chose the reference gene, *NOTCH2*, because it is not within a CNV variable region, and thus the copy number is always two [11,12]. By this method, the CNV and the SNP of the RQPS probe DNA can be compared to the genomic DNA in the pyrosequencing software for analysis, where the artificial 'G' nucleotide in *NOTCH2* will be compared to the genomic nucleotide 'A' which occurs in all genomic DNA at the same position. Similarly, for *FCGR3A*, an artificial 'C' nucleotide in the RQPS probe, will be placed at the nucleotide position of the true SNP (G/T) found in genomic DNA for *FCGR3A*, for use in analysis of the CNV of the SNP. The artificial SNPs allow for the probe DNA to later be distinguished from the genomic DNA.

More specifically, using the published DNA sequence for *NOTCH2,* we selected a 300 base pair (bp) region of the sequence. In the PyroMark software, we copied the sequence in and introduced 'G/A' SNP (where "A" is the nucleotide present, and "G" is the nucleotide which we will synthetically create in our probe). This sequence, with its artificial SNP included, was used to design both the forward and reverse primers (one of which will be biotinylated), and a sequencing primer. The biotinylated primer allows for isolation of our amplified *NOTCH2* sequence, which will be explained in more detail later. The designed sequence was then linked with a stuffer sequence of random nucleotides (about 30 nucleotides from the *GAPDH* gene sequence) to the modified the sequence of *FCGR3A* for the RQPS probe to include the artificial 'A' SNP. We used the NCBI's 'BLAST' database to ensure no homology or cross-reactivity between the sequences of *GAPDH, NOTCH2*, and the *FCGR3A* sequences. We ordered the probe sequence from the DNA technology company IDTDNA. We verified the probe length via gel electrophoresis.

Pyrosequencing Preparation:

To prepare for the pyrosequencing procedure, we combined the genomic DNA in molar ratios to the RQPS probe. The recommended ratios are 1:1, 2:1, and 3:1 (RQPS probe to DNA). Prior to performing the pyrosequencing reaction, we used a Real Time PCR reaction to ensure the fidelity of the ratios. The purpose of the ratios is to dilute the signal of the probe and to provide additional replicates built in within the same assay.

Using the mixed DNA (genomic DNA:probe) we perform PCR of the NOTCH2 sequence, and in a separate tube we perform PCR of the FCGR3A sequence. Following PCR using the forward primer and the biotinylated reverse primer, the biotinylated primer allows for replication of the biotin with the DNA strand, thus allowing for capturing of the PCR fragment

with streptavidin-coated beads. The isolated PCR fragment is then denatured, leaving behind a single stranded DNA, which is then annealed to the predesigned sequencing primer. The PyroMark Assay Design Program will predict the 'Analysis Sequence'; this gives us predicted histograms for what each genotype would look like.

Once the software and predicted sequences are set up, the PyroMark MD adds substrate and enzyme to each well of the 96 well plate that contains the PCR products annealed to sequencing primers. Based on the Analysis Sequence, the machine adds nucleotides in a specific order to each well. If the added nucleotides are incorporated in the DNA strand, together with enzymes in the sample mix, a light signal is generated. The intensity of the light signal corresponds to the number of nucleotides that incorporated into the DNA strand. The PyroMark software displays the results as histograms, which can be quantified as percentages of the peak height for each SNP within the software, in which the percentages compare the ratios of 'G' or 'T' nucleotides for *FCGR3A* in the genomic DNA to the 'C' nucleotides in the RQPS probe, and 'T' nucleotides for *NOTCH2* in the genomic DNA to the 'A' nucleotides in the RQPS probe. The percentage values can then be used to generate a scatter plot graph, with *FCGR3A* SNPs values along the y axis and *NOTCH2* values plotted along the x axis. By including a trend line, the scattered plots determine the copy number and SNP of the genomic DNA.

Calculation of Data Points

The results from the pyrosequencing machine are displayed as histograms, with the height of each peak representing the number of nucleotides added for each nucleotide. The software program interprets the ratio of the genomic DNA to the probe DNA as a value proportional to the copy number, from which the SNP type can be obtained.

Data Analysis

To analyze the data from the pyrosequencing histograms, we followed the analytical process suggested by Liu *et al.*, 2010 to determine the CNV and SNP. This involves calculating a value m_1, the ratio for the probe reference SNP in the 1:1 probe to genomic DNA mixture produced by the histogram, m_2, the corresponding ratio for the 2:1 mixture, as well as n_1, the ratio of the probe query SNP in the 1:1 ratio, and n_2, the corresponding value for the 2:1 mixture.

The data is then graphed into a scatter plot, where one data point comes from the value of ($n_1/[1-n_1]$, $m_1/[1-m_1]$), and the second data point comes from the value of $n_2/[1-n_2]$, $m_2/[1-m_2]$). After adding a trend line with the intercept set to zero, the slope of the line can be multiplied by the copy number of the reference gene, which in this case is *NOTCH2*, which has two copies.

Measurements of Success

The overall efficacy of the RQPS procedure is determined by the ability to interpret the results. We have a set of "expected" results, based on previously conducted Taqman assays for both SNP and CNV. We expect to be able to replicate the CNV results that were determined in the Taqman assay using the RQPS assay (i.e. if we determined sample A to have 2 copies via Taqman, we should find it has 2 copies of the allele via RQPS assay). For those samples that have predetermined accurate SNP genotyping, we expect to be able to replicate those with this assay. For those samples that have shown CNV via Taqman assays, we expect to improve the accuracy of the SNP readout using this RQPS assay.

Troubleshooting

During the development of these assays, a difficult step in preparing the reactions for *FCGR3A* was its high-sequence homology with *FCGR3B*. We worked with IDTDNA, to design RNase H Primers to specifically amplify *FCGR3A*. We have confirmed, through sensitive qPCR

assays, that we are in fact specifically amplifying *FCGR3A* in a *single* PCR assay. Using these same RNase H Primers for SNP analysis with Taqman probes, we have also analyzed the CNV for *FCGR3A*. This portion of our experimental design is in press, with myself included as a co-author, in Springer Protocols [13].

Results

After genotyping twenty-four individuals for the CNV and SNP of *FCGR3A* using RQPS, we have determined 0% to be CNV. Ten samples were F/F, eight samples were V/F, and 4 samples were V/V. Three samples were unable to be genotyped, and need to have new DNA isolated. The complete results are displayed in Table 1. Of the twenty-four genotyped, the results of seventeen were consistent with the Real Time PCR data, also shown in Table 1.

Table 1. RQPS and Real Time PCR results by sample number. Seventeen of the samples had consistent results across the two methods, while four did not.

Sample	RQPS	SNP PCR
KA 4	F/F	F/F
KA 5	V/F	V/F
KA 8	V/F	V/F
KA 12	F/F	F/F
KA 14	F/F	F/F
KA 18	V/F	V/F
KA 23	V/V	V/V
KA 26	V/V	V/V
KA 34	F/F	F/F
KA 37	V/V	V/F
KA 39	V/V	V/V
KA 40	?	F/F
KA 41	?	V/F
KA 42	F/F	F/F
KA 43	F/F	F/F
KA 44	?	F/F
KA 45	V/F	V/F
KA 47	V/F	F/F
KA 48	F/F	F/F
KA 51	V/F	F/F
MR	F/F	F/F
JS	V/F	V/F
KR	F/F	V/F
AH	V/F	V/F

Discussion and Future Directions

The adaptation of RQPS for genotyping of *FCGR3A* is not yet a viable option to simultaneously determine the genotype CNV and SNP. We have several explanations for these data, the most promising being that an imbalance in the ratio between genomic DNA and probe skews the results toward increased copy number. With further troubleshooting and trials, we expect to be able to resolve these issues and utilize RQPS in future projects.

If the ratio between genomic DNA and the synthesized probe is unequal in favor of the probe, we would expect to see a decrease in the dilution of the pyrosequencing signal such that more probe signal would be observed in the histograms, even at the 1:1 ratio. We did see this in several samples, and increasing the concentration of genomic DNA seemed to help. However, increasing concentration did not improve all of the samples, and so we do not believe this is the only problem with the assay. To combat this portion of the issue, we will order synthesized DNA sequences from IDT DNA (the same company that makes the probe) that represent the expected sequences for genotypes of V/V, V/F, and F/F for the *FCGR3A* SNP. We will dilute the sequences to 5 ng/μL, 10 ng/μL, and 15 ng/μL, and run the RQPS assay with these sequences as the 'genomic DNA'. This will help us determine a standard amount of DNA that should be used for each sample. It will also serve as a 'calibration' for the pyrosequencer to compare each result to, showing us the expected histograms for each possible genotype.

Once we have conducted further adaptation and can show that RQPS is a viable and cost-effective alternative to the laborious Real Time PCR method of simultaneous genotyping, the next step in this project will be to apply the adapted methodology to a population of Renal Cell Carcinoma (RCC) patients from the SELECT clinical trial. After genotyping all of the RCC patients for their SNP and CNV, we intend to perform flow cytometry on all of the samples to assess the protein expression levels of *FCGR3A*, and later, the other FCGRs. Then, we will determine whether the copy number influences the protein expression level, or if there are other factors. Finally, we will assess whether CNV and/or SNP of *FCGR3A* influences the response to the immunotherapeutic treatment given as part of the RCC SELECT Trial. Collectively, this information will shed light on whether some treatment courses are more significantly beneficial than others to individuals with specific genotypes.

References

[1] Duan J, Zhang J-G, Deng H-W, Wang Y-P (2013). "Comparative Studies of Copy Number Variation Detection Methods for Next-Generation Sequencing Technologies." *PLoS ONE* 8(3): e59128. doi:10.1371/journal.pone.0059128

[2] Bruhns P, et al (2008). "Specificity and affinity of human Fc gamma receptors and their polymorphic variants for human IgG subclasses." *Blood* 113: 3716-3725.

[3] Nimmerjahn F. and Ravetch J (2008). "Fc gamma receptors as regulators of immune responses." *Nature Reviews* 8:34-47.

[4] Cartron G, et al. (2011). "Interindividual variability of response to rituximab: from biological origins to individualized therapies." *Cancer Res* 17:19-30.

[5] Cartron, G., et al. (2002). "Therapeutic activity of humanized anti-CD2-monoclonal antibody and polymorphism in IgG Fc receptor Fc gamma RIIIa gene." *Blood* 99:754-758.

[6] Zhang, W., et al. (2007). "FCGR2A and FCGR3A polymorphisms associated with clinical outcome of epidermal growth factor receptor expressing metastatic colorectal cancer patients treated with single-agent cetuximab." *J Clin Oncol* 25(24): 3712-3718.

[7] Wu, J., et al. (1997). "A novel polymorphism of FcgammaRIIIa (CD16) alters receptor function and predisposes to autoimmune disease." *J Clin Invest* 100(5): 1059-1070.

[8] Delgado, D. C., et al. (2010). "Genotypes of NK cell KIR receptors, their ligands, and Fcgamma receptors in the response of neuroblastoma patients to Hu14.18-IL2 immunotherapy." *Cancer Res* 70(23): 9554-9561.

[9] Bakema, J. E. and M. van Egmond (2014). "Fc receptor-dependent mechanisms of monoclonal antibody therapy of cancer." Curr Top Microbiol Immunol 382: 373-392.

[10] Eckhardt, C. L., et al. (2014). "The Fc gamma receptor IIa R131H polymorphism is associated with inhibitor development in severe hemophilia A." *J Thromb Haemost* 12(8): 1294-1301.

[11] Liu, Z., et al. (2010). "Simple copy number determination with reference query pyrosequencing (RQPS)." *Cold Spring Harb Protoc.* 9:1-11.

[12] Liu, Z., et al. (2009). "Rapid identification of homologous recombinants and determination of gene copy number with reference/query pyrosequencing (RQPS)." *Genome Res* 19(11):2081-2089.

[13] Erbe, A.K., et al (2016). "Genotyping Single Nucleotide Polymorphisms (SNPs) and Copy Number Variability (CNV) of the FCGRs (FCGR3A and FCGR2C) expressed on NK cells." *Natural Killer Cells: Methods and Protocols.*