Identification of a Novel Mucin Gene HCG22 Associated With Steroid-Induced Ocular Hypertension

Shinwu Jeong,¹,² Nitin Patel,¹ Christopher K. Edlund,³ Jaana Hartiala,¹ Dennis J. Hazelett,⁴ Tatsuo Itakura,¹ Pei-Chang Wu,¹,⁵ Robert L. Avery,⁶ Janet L. Davis,⁷ Harry W. Flynn,⁷ Geeta Lalwani,⁶,⁷ Carmen A. Puliafito,²,⁷,⁸ Hussein Wafapoor,⁹ Minako Hijikata,¹⁰ Naoto Keicho,¹⁰ Xiaoyi Gao,¹¹ Pablo Argüeso,¹² Hooman Allayee,¹,³ Gerhard A. Coetzee,³,⁴,¹³ Mathew T. Pletcher,¹,¹⁴ David V. Conti,³ Stephen G. Schwartz,⁷ Alexander M. Eaton,⁹ and M. Elizabeth Fini¹,²,⁴,¹⁵

METHODS. A genome-wide association study was performed, followed by an independent candidate gene study, using a cohort enrolled from patients treated with off-label intravitreal triamcinolone, and handling change in IOP as a quantitative trait.

RESULTS. An intergenic quantitative trait locus (QTL) was identified at chromosome 6p21.33 near the 5′ end of HCG22 that attained the accepted statistical threshold for genome-level significance. The HCG22 transcript, encoding a novel mucin protein, was expressed in trabecular meshwork cells, and expression was stimulated by IL-1, and inhibited by triamcinolone acetate and TGF-β.

Bioinformatic analysis defined the QTL as an approximately 4 kilobase (kb) linkage disequilibrium block containing 10 common single nucleotide polymorphisms (SNPs). Four of these SNPs were identified in the National Center for Biotechnology Information (NCBI) GTEx eQTL browser as modifiers of HCG22 expression. Most are predicted to disrupt or improve motifs for transcription factor binding, the most relevant being disruption of the glucocorticoid receptor binding motif. A second QTL was identified within the predicted signal peptide of the HCG22 encoded protein that could affect its secretion. Translation, O-glycosylation, and secretion of the predicted HCG22 mucin gene product via two different mechanisms (transcription or secretion) is highly suggestive of a role in steroid-induced OH.

Keywords: ocular hypertension, glucocorticoid, GWAS, gene
Glucoma is a leading cause of irreversible blindness, presently affecting approximately 70 million people worldwide. The most common form is primary open-angle glaucoma (POAG), which accounts for approximately 90% of all cases. Ocular hypertension (OH) is the major risk factor for POAG and lowering IOP is the only effective treatment. However, many patients remain refractory to existing IOP-lowering interventions and eventually may become blind, underscoring the unmet medical need for novel approaches to control IOP.

Intraocular pressure is a product of the rate of aqueous humor production, resistance to outflow, and episcleral venous pressure. The aqueous outflow pathways in the angle of the eye are composed of a spongy tissue called the trabecular meshwork (TBM), which leads into Schlemm's canal. Much of the resistance to outflow resides within the TBM, putatively within 7 to 14 μm of the inner wall of Schlemm's canal in a region known as the juxtacanalicular or cribriform region. Primary open-angle glaucoma is caused by an increase in outflow resistance due to poorly understood defects in this region. Ocular hypertension in POAG has been associated with dysregulated intracellular signaling pathways controlled by IL-1/NF-κB, TGF-β superfamily, and Wnt.

The pathophysiology of OH leading to POAG shares many features with a secondary form of OH caused by treatment with glucocorticoids (GCs). GCs are a class of steroid hormones released in response to stress. As part of the natural feedback mechanism that turns down the inflammatory response, they are useful for treating a wide variety of diseases. Complicating this, however, is the considerable interindividual variability and tissue-specific GC sensitivity among individuals, which can cause a variety of systemic side effects. Treatment with steroids, such as dexamethasone or triamcinolone acetonide (TA), in the eye causes elevated IOP in predisposed individuals. It has been documented that approximately 40% of the normal population develops an IOP increase >6 mm Hg above baseline following topical administration of GCs, 4 times a day for 4 to 6 weeks. When GCs are administered intravitreally, IOP may increase by 30% or more in up to half of patients. These individuals are considered to be "steroid responders." Interestingly, almost all POAG patients are steroid responders; conversely, steroid responders who do not have POAG are at much higher risk of developing POAG compared to nonresponders. A better understanding of genetic and molecular mechanisms underlying individual variation in the response to steroids could shed light on our understanding of steroid-induced OH and OH leading to POAG. That being said, mechanisms of steroid-induced OH and OH leading to POAG also show distinct differences; therefore, studying the genetics of steroid-induced OH specifically could lead to discovery of genes involved in IOP regulation that could not have been revealed any other way.

The genome-wide association study (GWAS) represents an agnostic approach for prioritization of genes that are associated with a disease process, and makes it possible to discover novel genes that could not otherwise be identified. In the current study, we performed a pharma-GWAS for this purpose, using a cohort enrolled from patients treated with off-label intravitreal TA (IVTA) for various retinal indications, and followed up with an independent candidate gene study. We identified two independent quantitative trait loci (QTLs) at chromosomal locus 6p21.35 associated with the novel mucin gene HCG22.

Materials and Methods

Ethics Statement

This study was compliant with the Declaration of Helsinki, the Health Insurance Portability and Accountability Act (HIPAA), and the Institutional Review Boards of the University of Miami (Miami, FL, USA), Lee Memorial Health System (Cape Coral, FL, USA) and the University of Southern California (Los Angeles, CA, USA). Informed consent was obtained from all research subjects after explanation of the nature and possible consequences of the study. The Institutional Review Board of the University of Southern California reviewed the protocol for cell isolation from the discarded human corneal rims (which are coded specimens) and designated it as nonhuman subject research.

Study Cohorts

Research subjects for the “Florida-2” cohort served as the discovery group for the current study. Research subjects were enrolled at the Bascom Palmer Eye Institute, University of Miami (Miami, FL, USA; 13 subjects), the Retina Health Center (Fort Myers, FL, USA; 61 subjects), or the California Retina Associates (Santa Barbara, CA, USA; 5 subjects). Most subjects reported their race as “white.” Patients were offered entry into the study if they were to be treated with an injection of IVTA (approximately 4 mg in 0.1 cc) for various retinal diseases or previously had been treated with IVTA. All treatments were similar: approximately IVTA, 4 mg in 0.1 cc. Inclusion criteria were age 18 years or greater, at least two IOP measurements over six months post injection, and the ability to give appropriate informed consent. Exclusion criteria were prior pars plana vitrectomy in the study eye, previous trabeculectomy or glaucoma tube shunt surgery in the study eye, anterior segment neovascularization in the study eye, and uveitis other than cystoid macular edema status post cataract or scleral buckle surgery in the study eye.

The previously described “Florida-1” cohort, enrolled at the Bascom Palmer Eye Institute, University of Miami served as the replicate group for the current study. Inclusion/exclusion criteria were similar to the Florida-2 cohort, as described previously. Exit criteria described in the published study included subsequent intravitreal injection, any intervention intended to lower IOP or subsequent intraocular surgery; however, for the current study this was revised to allow subsequent IVTA or interventions intended to lower IOP, to parallel the Florida-2 cohort. Clinical data was updated by retrospective chart review.

Measurement of IOP

Intraocular pressure was recorded in the study eye before treatment with IVTA (baseline IOP). All IOP measurements were performed according to the treating physician’s standard practice, typically planation tonometry or Tono-Pen (Reichert Technologies, Buffalo, NY, USA). On each follow-up visit, IOP was recorded (follow-up IOP) for up to 1 year, or until the study eye met any of the predetermined exit criteria.

Change in IOP (ΔIOP) was defined as the difference between baseline IOP and the maximum IOP measurement for 6 months after each injection. This was used as the quantitative trait (QT) for each research subject. A positive ΔIOP indicated an elevation in IOP following IVTA. If both eyes were treated, the eye with the highest ΔIOP was selected. If a second IVTA injection was given during the 6-month follow-up, the last IOP used to determine the ΔIOP was the one the day of the subsequent IVTA injection. If both eyes were treated, the eye with the highest ΔIOP was selected.

For this study, a “steroid responder” was defined, according to Armaly and Becker, as an individual with normal IOP (12-22 mm Hg) who has a positive ΔIOP > 6 mm Hg above baseline IOP following injection of IVTA.

DNA Preparation

Peripheral venous blood was obtained from all research subjects. Genomic DNA was extracted from leukocytes using...
HCG22 and Steroid-Induced Ocular Hypertension

Genotyping and Quality Control

DNA samples from the 79 subjects enrolled into the Florida-2 cohort were genotyped using the Illumina HumanOmni2.5-Quad (Omni2.5) BeadChip at the Epigenomics Core facility of the University of Southern California using Illumina GenomeStudio software (v2011.1; Illumina, Inc., San Diego, CA, USA). Genotyping reproducibility was confirmed by including two duplicate samples. Genome-wide association study quality control (QC) was performed using PLINK (v1.07) software following Illumina protocols and published procedures at sample and single nucleotide polymorphism (SNP) levels. Of the 79 original subjects, 64 survived the QC procedure with all having a genotyping call rate > 95%. The 15 individuals who failed QC were dropped from further analyses. This included individuals with a substantial amount of African ancestry (n = 5), ambiguous sex (n = 2), unexpected duplicates (n = 2), and individuals with missing IOP phenotype (n = 6) in our genotyped samples.

Of the 2,449,708 SNPs successfully genotyped, a total of 1,545,328 SNPs were included in the final analysis after exclusion from either a low call rate < 90% (65,469), deviation from Hardy-Weinberg Equilibrium with a P value for HWE test < 0.001 (2058), or a minor allele frequency < 0.05 (849,444).

DNA samples from the 52 subjects enrolled into the Florida-1 cohort were previously genotyped using the GeneChip Human Mapping 500K Array Set from Affymetrix in the Genomics Core facility at The Scripps Research Institute (Jupiter, FL, USA).

Population Stratification

We used PLINK to remove outlier individuals and SNPs in high linkage disequilibrium (LD) or in regions of extended LD. Principal components of genetic ancestry were inferred using EIGENSOFT. The genomic control inflation factor was calculated and a quantile–quantile (Q–Q) probability plot was generated to visualize the distribution of the test statistics.

Statistical Analysis

Association analysis was conducted using linear regression, treating an inverse-normal transformation of ΔIOP as a continuous variable, with adjustment for age, sex, and principal components of genetic ancestry, and assuming an additive genetic effects model. Genotyped SNPs were analyzed using PLINK software. For a meta-analysis with the Florida-1 and Florida-2 cohorts, a fixed effects model was used, including a test for heterogeneity, as implemented in the meta R package on tests of the overall treatment effect in meta-analysis with normally distributed responses.

SNP Validation

TaqMan analysis was used to validate genotypes for specific SNPs of interest, or to determine genotypes for SNPs not on the microarrays. Allelic discrimination analysis was performed using SDS 2.3 software and the data were analyzed in TaqMan Genotype v1.1 (LTI).

Bioinformatics

HaploReg v2 software (update 2013.02.14) was used to search for proxy SNPs of the index SNPs identified in this study. ConSite was used to identify binding motifs for transcription factor NF-κB. Association of SNPs with gene expression was evaluated using the NCBI GTEx eQTL browser. ProtoNet software was used to identify proteins with sequence similarity to HCG22.

To assess the likely effect of SNPs on transcription factor response elements, we first identified matches to well-established positional weight matrices (PWM) from Factorbook and HOCOMOCO, and then determined whether the affected residue is critical within the motif. For consideration, the proposed effect had to be in a response element that scored 80% or better of the maximum PWM score, and the difference of frequency between the reference genome and alternate SNP value had to be greater than 60%. A full mathematical description of the algorithm has been reported previously.

Cultured Cells

Primary human corneal epithelial cells (purchased from Lifeline Cell Tech, Frederick, MD, USA) were grown in OcuLife Complete Medium Kit, provided by the supplier, and used for experiments at passage 4. The immortalized human trabecular meshwork cell line, TM-1, kindly provided by Donna Peters (University of Wisconsin, Madison, WI, USA), was maintained as described. Primary human TBM cells were prepared from corneal rims discarded at the time of surgical transplantation of donor corneas, as described. Primary TBM cells before passage 8 were used for experiments.

DNA Constructs and Transfections

Total RNA purified from BET1A, a human bronchial epithelial cell line, was used to synthesize cDNA, from which HCG22 cDNA (GenBank AB560771) was PCR-amplified and cloned into the pCR2.1-TOPO vector using the TOPO TA Cloning kit (Invitrogen, Carlsbad, CA, USA), using the oligonucleotide primers 5′-CTA GCA CCA TGT GGA TTC TCT TGA CAT GAT and 5′-CCA ACA AAG AGC ACA ACA GTG ACT G-3′. The DNA sequence of cloned cDNA was confirmed by DNA sequencing and was found to contain two amino acid substitutions (R to G and E to Q) in the signal peptide region (see Results). EcoR1 and Not I restriction enzyme sites in pcDNA3.1+ vector (Invitrogen) were used to clone and express recombinant HCG22 protein with a His tag at the C-terminus, using forward primer, 5′-GAT CCA ATT CAT GCC CGG CTA CGT CC-3′, and reverse primer, 5′-GCT ACG ATG CGG CCG CTC AGT GGT GGT GGT GGA GAA GGT GAG TGC CAC CAT GTT TA-3′. To express the HCG22-His tag cDNA, TM-1, and HEK293 cells were transfected with 2 μg of the vector plasmid using Lipofectamine LTX reagent (Invitrogen) according to the supplier’s protocol. A HEK293 cell line stably expressing HCG22-His was made by culturing the transfected cells in the presence of 800 μg/mL G-418 (a selective agent).

Expression Analysis

The TM-1 cells were grown in 6-well plates at 0.12 millions of cells/well and cultured for 4 days before serum starvation for 7 hours. Cells then were stimulated by adding IL-1β (10 ng/mL), TGF-β2 (10 ng/mL), or TA ((100 ng/mL; Sigma-Aldrich Corp., St. Louis, MO, USA) for 17 hours. Total RNA was purified using TRIzol (Ambion, Grand Island, NY, USA). The cDNA was synthesized using iScript cDNA Synthesis Kit (Bio-Rad, Hercules, CA, USA), and quantitative RT-PCR (qRT-PCR) was performed using iQ SYBR Green Supermix (Bio-Rad) under thermal cycling conditions of 30 seconds at 95°C, 10 seconds
at 60°C, and 10 seconds at 72°C. Relative transcript amounts were normalized by GAPDH expression in individual treatment samples. The following PCR primer pairs were used: GCGAAGCCTGGTGGGT and CAGGGACAGGCTGTTTCTCA for MUC21; ACCAACACACCTATGTCAGGC and TCCTGGCACAGTATGGCTG for MUC22; CGAGCGAAGGTACGGAATAGTA and GTCCTCAGTGGTCCTTCCATC for HCG22, which spans exon2-exon3 junction in the coding transcript that is not present in the noncoding HCG22 transcript; GAGTCAACGGATTTGGTCGT and GGCAACAATATCCACTTTACCAGAGT for GAPDH.

O-Glycosylation Test

The TM-1 and HEK293 cells were transiently transfected with HCG22-His vector for 1 day to observe the secretion of HCG22 protein. The HEK293 cells stably transformed with HCG22-His were grown until reaching 80% confluence in a 12-well plate. After washing the wells twice with serum-free Dulbecco’s modified Eagle’s medium (DMEM), cells were treated with benzyl-2-acetamido-2-deoxy-alpha-D-galactopyranoside (benzyl GalNAc; Sigma-Aldrich Corp.) in serum-free medium for 24 hours at a range of doses. The culture supernatant was collected and subjected to SDS-PAGE and Western blotting. The membrane was probed using anti-His antibody (Abcam, Cambridge, MA, USA) to detect the expressed recombinant protein.

RESULTS

Genome-Wide Association Study

Clinical data on the 64 white subjects from the Florida-2 cohort used for the present analysis are shown in Table 1. A Q–Q plot of SNP association \( P \) values (genomic control inflation factor corrected) showed that the observed \( P \) values did not differ from their expected values, except at the extreme tail (Supplementary Fig. S1). Thus, this result demonstrates good control of population stratification for our white subjects.

Linear regression analysis was performed to determine association between the quantitative trait and genotype. The results are represented by the Manhattan plot in Figure 1, where the genome-level significance threshold was set at \( P = 5 \times 10^{-8} \). Two SNPs, rs2523864 and rs2251830, on chromosome 6 exceeded this threshold and yielded \( P \) values of \( 3.92 \times 10^{-6} \) and \( 4.74 \times 10^{-6} \), respectively. These SNPs are in perfect LD (\( r^2 = 1 \)). A second peak on chromosome 9 reached borderline significance, with the top SNP having a \( P \) value of \( 1.56 \times 10^{-7} \). This SNP is located in an intragenic region upstream of the TLE4 gene, which encodes a transcription factor. Other regions of the genome with suggestive evidence of association included two SNPs in perfect LD (\( r^2 = 1 \)) on chromosome 10, located in an intron of the ADAR2 gene (an RNA editing enzyme), and a SNP on chromosome 20,

![Figure 1](https://example.com/figure1.png)

**FIGURE 1.** Manhattan plot of the GWAS for steroid-induced OH. The Manhattan plot depicts the genome-wide \( P \) values. The results for the total of 1,545,328 genotyped SNPs that passed the quality control criteria are plotted as \(-\log_{10}(P \text{ value})\) by genomic position. The red line indicates the genome-wide significance level of \( P = 5 \times 10^{-8} \).
The statistically significant peak is situated at chromosomal locus 6p21.33 (Fig. 2). This region lies within the 3.6 Mb major histocompatibility complex (MHC) class I locus in the region between the human leukocyte antigen genes HLA-A and HLA-B, approximately 300 kb telomeric of HLA-B. Early studies linked this region to psoriasis, an inflammatory skin disease.\textsuperscript{58-61} Single nucleotide polymorphisms in this region also have been associated with two inflammatory diseases of the eye. The first is Behçet’s disease, a rare immune-mediated systemic vasculitis that often presents with mucous membrane ulceration and ocular involvements that include a secondary form of glaucoma.\textsuperscript{62} The second is Stevens-Johnson syndrome, a hypersensitivity complex that affects the skin and mucous membranes, including the cornea of the eye.\textsuperscript{63} Most recently, SNPs in this region have been associated with dilated cardiomyopathy\textsuperscript{64} and diffuse panbronchiolitis (DPB), a rare, genetically inherited disease of the lungs that primarily affects Asians.\textsuperscript{65} A complete gene annotation was performed in connection with the DPB study, revealing two novel protein-coding genes \textit{PBMUCL1} and \textit{PBMUCL2}, that mapped to this region.\textsuperscript{65} Conceptual translation revealed that \textit{PBMUCL1} is a new member of the MUC family of mucin genes and has been given the HUGO designation of \textit{MUC22}. The \textit{PBMUCL2} (Genbank: AB560771) gene was found to overlap the gene \textit{HCG22}, previously associated only with a noncoding transcript (Genbank: AK094453.1). The HUGO Gene Nomenclature Committee (HGCN) subsequently recognized \textit{HCG22} as the gene for the coding transcript. The two lead SNPs we identified map to the intergenic region between \textit{MUC22} and \textit{HCG22} (Fig. 2).

**Replication**

To confirm the association signal at the chromosome 6p21.33, we used the Florida-1 cohort as an independent replication dataset. Clinical data on the 49 subjects surviving QC analysis are shown in Table 1. Since neither of the peak SNPs for the chromosome 6p21.33 locus was present on the Affymetrix chipset previously used for the Florida-1 cohort, we determined genotypes by Taqman analysis. As noted in the discussion of the Florida-2 cohort, rs2523864 and rs2251830 SNPs are in perfect LD ($r^2 = 1$). Therefore, we only analyzed rs2523864 for replication purposes. Linear regression analysis yielded an effect estimate in a similar direction as in the Florida-2 cohort, with a corresponding $P$ value of 0.18 for association of rs2523864 with IOP. Given the directionally consistent effect of rs2523864, we next carried out a meta-analysis that combined the association results in the Florida-1 and Florida-2 cohorts ($n = 113$). This analysis significantly strengthened the association of rs2523864 with IOP by one order of magnitude (combined $P = 5.34 \times 10^{-5}$), supporting the notion that this variant influences the response to GCs.

**HCG22 Expression Analysis**

Glucocorticoids bind the GC receptor (GCR), which acts as a transcription factor to stimulate expression of specific genes. The proximity of the peak SNPs on chromosome 6 to the transcriptional start site for \textit{HCG22} suggests that the molecular mechanism for association of these variants with IOP could be through regulation of \textit{HCG22} gene expression. If altered transcription of \textit{HCG22} is a causal factor in steroid-induced OH, we further reasoned that the gene would be expressed in the TBM. To test this hypothesis, we performed RT-PCR expression analysis with primers specific for the coding transcript of \textit{HCG22} using the immortalized TBM cell line, TM-1. For comparison, the adjacent genes for \textit{MUC22} and \textit{MUC21} also were included in this analysis (Fig. 3). The MUC-type mucins are known to be expressed at high levels in the ocular surface epithelia, but have not been connected previously with the aqueous outflow pathways of the eye. Thus, we compared expression in TM-1 cells to expression in primary HCE cells. The \textit{MUC21} mRNA expression was specific for HCE cells, but \textit{MUC22} and \textit{HCG22} were expressed in HCE and TM-1 cells. This represents a novel finding, since mucins have not previously been associated with cells of the aqueous outflow pathways. These results were confirmed using primary TBM cells (data not shown).

We next used RT-PCR to quantify \textit{HCG22} mRNA expression in TM-1 cells treated with GCs and with IL-1β and TGF-β, agents that are controlled by GCs and associated with OH.\textsuperscript{15,16,65} (Fig. 4). Treatment with IL-1β stimulated \textit{HCG22} expression by approximately 4-fold. Treatment with TA inhibited basal and IL-1β-stimulated expression. Transforming growth factor-β inhibited \textit{HCG22} expression to approximately 0.3-fold. These results also were confirmed using primary TBM cells (data not shown).

To determine any functional impact of the quantitative trait locus on \textit{HCG22} expression, we first used bioinformatics tools...
to search for proxy SNPs of the lead SNP, rs2523864. Using the 1000 Genomes CEU Phase 1 data implemented in HaploReg v2 software,96 we identified 10 SNPs spanning a region of approximately 4 kb that were in high LD with rs2523864 ($r^2 \geq 0.8$) (Table 2). This LD block is situated upstream of HCG22, located approximately 1.8 kb from the transcriptional start site for all transcripts.

The National Center for Biotechnology Information (NCBI) GTEx eQTL browser is a central resource that archives and displays results of a national research project for determining association between genetic variation and high-throughput molecular-level expression phenotypes, and this information can provide insight into the biological relevance of data from GWA studies.51 We used this resource to determine whether SNPs in the identified LD block were associated with HCG22 expression levels in the lymphoblastoid cell line used by NCBI. This analysis revealed highly significant eQTL $P$ values for four SNPs, two of which are our index SNPs (Table 2).

To assess the likely effect of the SNPs in the LD block on transcription factor response elements, we identified matches to well-established PWMs from Factorbook52 and HOCOMOCO,53 and then determined whether the affected residue is critical within the motif. We found that SNPs in our LD block alter binding motifs for a number of transcription factors (Supplementary Table S1). Those involved in developmental eye defects that cause ocular hypertension, NF-kB, TGF-β, or Wnt signaling (pathways associated with OH), and response to GCs are listed in Table 2.

The DLX2 and Pax6 cause Axenfeld-Rieger syndrome, a disorder that manifests as anterior segment dysgenesis and congenital glaucoma.66 The NF177 and AP269 genes control retinal development. The MITF gene is implicated in microphthalmia and Waardenburg syndrome, a retinal disorder.69 Genes SNAI1 and SNAI2 also are implicated in Waardenburg syndrome.59 Genes TCF4 and ZEB1 are implicated in Fuch’s endothelial dystrophy, a disease of the corneal endothelial layer.70,71 Genes TCF4 and ZEB1 also are involved in determining central corneal thickness, which is a modifying genetic factor in OH.72,73

Other altered motifs bind transcription factors are involved in intracellular signaling associated with OH. The SMAD motif mediates TGF-β signaling. The POS5FI and TCF4 motifs mediate Wnt signaling. Most striking, however, from the perspective of biological relevance, is predicted disruption of the binding motif for the GCR that mediates response to GCs.74,75

Response to IL-1 is mediated by transcription factor NF-kB, and this pathway has been associated with OH in POAG.15 Since no NF-kB or related transcription factor elements appeared in the transcription factor binding site search, we used the web-based bioinformatics tool Consite50 to search the DNA sequence of our approximately 4-kb LD block for transcription factor binding sites that could mediate response to NF-kB. We identified eight canonical NF-kB response elements.

**HCG22 Coding Analysis**

Previous analysis of the hypothetical protein product of the coding transcript of HCG22 (Genbank: BAJ41551.1; UniProtKB/Swiss-Prot: E2RYF7.1) suggested it as a non-MUC-type mucin.65 Following the N-terminal signal sequence are 15 tandem mucin repeats within a threonine-rich C-terminal region of 160 amino acids. The lack of a transmembrane domain suggests that HCG22 mucin protein is secreted. Glycosylation patterns were predicted by analysis of coding regions of the genes using web-based software, revealing numerous sites for O-linked glycosylation within the tandem repeats. The predicted molecular weight of the HCG22 protein backbone is approximately 26 kDa, but the actual size could be considerably larger due to the addition of sugar chains.

To validate the predictions based on amino acid sequence analysis, HEK293 cells transfected with the HCG22 expression construct were left untreated or treated with Benzyl-GalNAc, which selectively inhibits the elongation and sialylation of mucin-type O-glycan chains.76,77 Conditioned culture media was harvested for Western blotting after 24 hours. Figure 5 shows the results of Western blotting using an antibody to the His tag at the N-terminus, which revealed a band corresponding to 70 to 75 kDa protein. The molecular size of this protein decreased with increasing doses of Benzyl-GalNAc. These results are consistent with HCG22 translation into a protein product that is O-glycosylated and exported outside the cell.

Application of the web-based computer program ProtoNet52 reveals that the HCG22 protein product is most closely
related to peritrophin-A, a protein found in the insect digestive tract.

Second Quantitative Trait Locus

The cloned insert of the HCG22 expression construct was sequenced as part of a routine procedure to confirm its identity. Unexpectedly, we identified two nucleotide substitutions in the signal peptide corresponding to two SNPs also present in dbSNP (rs3873352 and rs2523855), each of which predicts an amino acid substitution (NH3-MP to G). The substitution due to rs3873352 would cause a change from a positive charge and hydropathy index of 4.5 to a neutral charge and hydropathy index of 0.4. Such a change is likely to affect secretion. The second substitution due to rs2523855 would result in a glutamic acid to glutamine substitution at 18th amino acid, a conservative change that is less likely to have functional consequences.

Discussion

The pathophysiology of OH caused by treatment with steroids shares many features with the much more common form of OH leading to POAG, but also has some distinct differences.19–25 Thus, steroid-induced OH provides a valuable model that could help shed light on our understanding of POAG, but also could provide novel insight into the regulation of IOP. Applying a pharmaco-GWAS and a candidate gene approach, we identified two independent QTLs associated with HCG22 at chromosomal locus 6p21.33. Results of follow-up studies indicate that one of these QTLs may affect transcription of HCG22, and one may affect secretion of HCG22 protein in the aqueous outflow pathways. Thus, the first QTL could affect levels of the noncoding and coding transcripts derived from HCG22, while both QTLs could alter levels of the encoded protein. The identification of two independent QTLs that are associated with IOP possibly through different molecular mechanisms (i.e., gene transcription and protein secretion), is highly suggestive that HCG22 has an important role in steroid-induced OH.

Genetic effects in pharmacologic response can be quite large in comparison with those involved in most complex diseases, and may involve fewer genes.78 In fact, several recent pharmacogenomic case-control studies of approximately 100 cases have detected polymorphisms of genome-wide significance.78 Our 2-stage GWAS comprised two cohorts, analyzed as discovery and replication groups, then combined by meta-analysis for a total of 113 subjects. Analysis of the discovery cohort dataset resulted in a P value of genome-wide significance and the overall P value increased in magnitude (P = 5.34 × 10^{-5}) when the cohorts were combined. However, these findings warrant further investigation, since GWA studies are inherently hypothesis-generating exercises, and function is the ultimate test of significance. We note that the genetic

![Benzyll-GalNAc](image)

**Figure 5.** Translation, O-glycosylation and secretion of HCG22 protein. The HEK293 cells expressing HCG22-His tag protein were cultured in the presence of benzyl-GalNAc at indicated doses. After 24 hours, the conditioned media were harvested. Secreted proteins were resolved on a denaturing 10% SDS/PAGE gel, blotted, and probed with anti-His antibody.
associations we identified are supported by initial functional follow up studies reported here, which serve as a starting point for more detailed characterization of HCG22 and its role in the development of OH in response to GCs.

The method used in this study allowed us to address the quantitative trait was conservative, and this conclusion partially explains why the Florida-2 results did not replicate in the Florida-1 cohort. The ΔIOP was calculated as the maximum difference between baseline and follow-up IOPs and the timing of the follow-up visits may not have corresponded with the highest IOP attained. Moreover, high responders are treated early with IOP-lowering medications, meaning these patients would never demonstrate their full potential. Our discovery and replication sets were obtained in two different physician practices, where steroid response may have been handled with different levels of aggressiveness. Each of these factors could lead to an underestimate of the ΔIOP in the Florida-1 cohort. Adding to this, the Florida-1 cohort was smaller than the Florida-2 cohort. However, the identified variants had directionally consistent effects in both cohorts and the combined meta-analysis increased the strength of the association by one order of magnitude, supporting the notion that variants at the HCG22 locus influence the response to GCs.

The two QTLs identified in this study are located at chromosomal locus 6p21.33 within the 3.6 Mb MHC class I locus, between the genes for HLA-A and HLA-B, a little more than 300 Kb telomeric of HLA-B. Interestingly, studies published more than 35 years ago, identified association between HLA-A, HLA-B, and POAG.79–81 and a study comparing cellular responsiveness to GCs and HLA classification for 25 POAG patients and 25 individuals who respond to GCs with an IOP > 31 mm Hg found an association with HLA-B12.82,83

The first QTL identified in this study mapped to an LD block of approximately 4.0 kb, which is situated in an intergenic area approximately 1.1 kb upstream of the transcriptional start site of the HCG22 gene. This suggested that the QTL might affect the regulation of HCG22 transcription by GCs. We found that TBM cells express HCG22, and expression is stimulated by IL-1 and inhibited by TGF-β. We also found that GCs inhibit constitutive and IL-1-stimulated expression of HCG22. Elevated IL-1/NF-κB and TGF-β signaling have been associated with OH leading to POAG, and GCs act in part by modifying IL-1 and TGF-β signaling. Thus, these results are consistent with a role in the disease process.

Gene expression phenotypes are heritable and can be mapped as eQTLs.75,76,84 The basis of eQTL analysis is that SNPs affecting with regards to potential mechanism was the GCR binding motif disruption that was identified. Functional analysis linking the DNA in our LD block to a transcriptional reporter gene can be used in future studies to determine how this and other candidate SNPs specifically affect GC-induced transcription in TBM cells. Since the SNPs in the identified LD block were perfectly correlated genetically, it is possible that more than one SNP contributes to the phenotype associated with this QTL.

The GCR gene NR3CI generates two transcription factor isoforms: GCRα and GCRβ.91 Isomorph GCRβ is a truncated form of GCRα that lacks the ligand-binding domain. It has a dominant-negative effect on GCRα-induced transactivation of promoters activated by GCs. However, GCRβ also has intrinsic, GCRα-independent transcriptional activity, regulating mRNA expression of a large number of genes negatively or positively.92,93 The GCRβ/GCRα ratio is reduced in a variety of diseases of GC sensitivity, including asthma, rheumatoid arthritis, and ulcerative colitis.75,76,79 The GCRβ/GCRα ratio also is reduced in glaucomatous TBM cells, and it was shown recently that this contributes to their GC sensitivity.90,93 It is conjectured that the relative reduction of GCRβ in relation to GCRα might be caused by elevated proinflammatory cytokines.94 This notion also is relevant to OH in POAG, which is associated with elevated IL-1α/NF-κB signaling.15

An important outcome of these studies is the implication that the HCG22 mucin protein may be a causative factor in steroid-induced OH. Previous analysis of the predicted amino acid sequence of the HCG22 coding transcript revealed features consistent with a small, secreted, O-glycosylated mucin-like protein.95 Here, we showed for the first time to our knowledge that this transcript can be translated, O-glycosylated, and exported outside the cell as an approximate-ly 70 to 75 kDa mucin protein. Our application of the web-based computer program ProtoNet reveals that HCG22 is most closely related to peritrophin-A, a mucin protein found in the peritrophic matrix of the insect digestive tract. Comprised of a network of glycoproteins, proteoglycans, and chitin fibrils, the peritrophic matrix functions as a molecular sieve to partially digest protein and carbohydrate, as a barrier to ingested pathogens, as a scaffold for proteases and glycosidases, and as a sink for toxic substances.96 As noted above, we found that GCs inhibit constitutive and IL-1-stimulated expression of HCG22. Thus, it is tempting to speculate that disrupted binding of the GCR in individuals with the minor genotype may prevent the GC-mediated inhibition of HCG22 expression and contribute to GC-induced OH. One idea is that too much HCG22 protein is detrimental to the trabecular meshwork cell under conditions of GC treatment or physiological stress. Further studies will be needed to test this hypothesis.

Sequence of the cDNA insert of the expression construct used for these studies revealed two SNPs within the putative signal peptide of HCG22, one that would cause a nonconservative amino acid change, and one that would cause a conservative change. Signal peptides are surprisingly complex and single amino acid substitutions can have major effects on their function in secretion.97 Statistical analysis revealed a significant association between the nonconservative SNP and steroid response. Future studies will be needed to quantify the SNP effects on secretion efficiency and/or intracellular distribution of HCG22 in cultured TBM cells.

While this manuscript was under review, sequences of three new HCG22 noncoding transcripts were submitted to Genbank (HG499131.1, HG499132.1, and HG499135.1). It has been reported that only approximately 20% of transcription progress across the human genome is associated with protein-coding genes.98 Because of recent advances in RNA sequencing and computational methods, long noncoding RNAs are emerging as key regulators of diverse cellular processes.99–105
The relative involvement in steroid-induced ocular hypertension of the HCG22 coding mRNA and the HCG22 noncoding mRNAs remains to be investigated in future studies.

Only a few other published studies have used genetics to prioritize genes associated with steroid-induced OH. Myocilin (MYOC), the first gene identified for hereditary forms of POAG,103 had already been defined as a highly induced protein in TBM cell cultures treated with GCs.102,103 However, subsequent analysis revealed no evidence for a link between steroid-induced OH and POAG-causing mutations.104,105 A candidate gene study using a cohort from trials of topically administered GCs following photorefractive keratectomy revealed no evidence for a link between steroid-induced OH in a small subgroup treated with prednisolone acetate.106 However, using the Florida-1 cohort described here, we found no statistically significant associations of response to IVTA with six candidate polymorphisms in NR3C1 previously associated with systemic GC sensitivity, including N363S.42 A similar study from another group on response to IVTA also found no significance.107 Thus, this result remains to be replicated by other researchers. To our knowledge, none of the currently ongoing GWA studies for POAG, OH, or related disorders has identified this locus. This suggests the possibility that it is specific for steroid-induced OH.

The past decade has seen the emergence of a first generation of effective pharmaceuticals for retinal diseases. Among these, GC drugs have shown much promise. For example, off-label use of TA, delivered by intravitreal injection (i.e., IVTA), has been used as primary or adjunctive treatment for a variety of retinal diseases, including choroidal neovascularization, macular edema secondary to diabetic retinopathy, and retinal vein occlusion. However, when GCs are administered intravitreally, IOP may increase by 30% or more in up to half of patients, usually within 3 months. The IOP elevation may be prolonged and intractable requiring surgical intervention. Identification of SNPs associated with response to steroids in the eye ultimately may provide a valuable predictive tool that could make GC drugs safer, by identifying patients who are steroid responders before treatment. Expansion of the current datasets and additional independent replication will be needed to determine whether the tag SNPs reported here might be used for this purpose.

Acknowledgments

The authors thank William Smiddy, MD (University of Miami, Miami, FL, USA) for patient referral to the Florida-1 cohort, Sabrina Gerzenstein (University of Miami, Miami, FL, USA) for technical assistance in preparation and banking of the Florida-1 cohort DNA samples, Donna Peters, MD (University of Wisconsin, Madison, WI, USA) for the gift of the TM-1 cell line, and Pragna Patel, MD (University of Southern California, Los Angeles, CA, USA) for helpful discussions on the genetic approaches.

Supported by National Institutes of Health (NIH; Bethesda, MD, USA) Grants R01-EY009828 (MEF), R01-EY022651 (XG), R01-CA136924 (GAC), U1L-RR031986 (pilot project subaward, MEF), P30-EY14801 (University of Miami), and P30-CA014089 (NIH/National Cancer Institute, US); Department of Defense Grants W81XWH09-1-0075 and W81XWH-13-1-0048 (University of Miami); a Schaffer Innovative Glaucoma Research Grant from the Glaucoma Research Foundation (available in the public domain at http://www.glaucoma.org; MEF); a Senior Scientific Investigator Award from Research to Prevent Blindness (available in the public domain at http://www.rpbusa.org; MEF); an unrestricted grant from Research to Prevent Blindness (available in the public domain at http://www.rpbusa.org/); a Schaffer Innovative Glaucoma Research Grant from the Glaucoma Research Foundation (available in the public domain at http://www.rpbusa.org/); an Investigator Award from Prevent Blindness Florida, Tampa, Florida, United States (available in the public domain at http://florida.preventblindness.org/; SGs); the Retina Health Center, Fort Myers, Florida, United States (available in the public domain at http://www.retinahalthcenter.com/); to support research staff involved in sample and data acquisition, AME); IC Labs, Fort Myers, Florida, United States (costs of the Illumina chip screening in the USC Epigenomics Center Core Facility, AME); Walter G. Ross Distinguished Chair in Ophthalmic Research at the University of Miami (MEF); and the Keck School of Medicine of USC (start-up funding for MEF).

Disclosure: S. Jeong, None; N. Patel, None; C.K. Edlund, None; J. Hartalia, None; D.J. Hazelett, None; T. Itakura, None; P.-C. Wu, None; R.L. Avery, IC Labs (F, I, S); P. J.L. Davis, None; H.W. Flynn, IC Labs (F, I, S), Retina Health Center (F, I, S); P. G. Lalwani, None; C.A. Puliafito, None; H. Wafapo, None; M. Hijikata, None; N. Keicho, None; X. Gao, None; P. Argueso, None; H. Allayee, None; G.A. Coetzee, None; M.T. Pletcher, None; D.V. Conti, None; S.G. Schwartz, None; A.M. Eaton, IC Labs (F, I, S), Retina Health Center (F, I, S); P. M.E. Fini, None

References


