

Prediction of *cis*-regulatory elements controlling genes differentially expressed by retinal and choroidal vascular endothelial cells

Dongseok Choi · Binoy Appukuttan · Sierra J. Binek ·
Stephen R. Planck · J. Timothy Stout ·
James T. Rosenbaum · Justine R. Smith

Received: 23 February 2008 / Accepted: 23 April 2008 / Published online: 26 June 2008
© The Author(s) 2008

Abstract Cultured endothelial cells of the human retina and choroid demonstrate distinct patterns of gene expression. We hypothesized that differential gene expression reflected differences in the interactions of transcription factors and respective *cis*-regulatory motifs(s) in these two endothelial cell subpopulations, recognizing that motifs often exist as modules. We tested this hypothesis in silico

by using TRANSFAC Professional and CisModule to identify *cis*-regulatory motifs and modules in genes that were differentially expressed by human retinal versus choroidal endothelial cells, as identified by analysis of a microarray data set. Motifs corresponding to eight transcription factors were significantly ($p < 0.05$) differentially abundant in genes that were relatively highly expressed in retinal (i.e., glucocorticoid receptor, high mobility group AT-hook 1, heat shock transcription factor 1, p53, vitamin D receptor) or choroidal (i.e., transcription factor E2F, Yin Yang 1, zinc finger 5) endothelial cells. Predicted *cis*-regulatory modules were quite different for these two groups of genes. Our findings raise the possibility of exploiting specific *cis*-regulatory motifs to target therapy at the ocular endothelial cells subtypes responsible for neovascular age-related macular degeneration or proliferative diabetic retinopathy.

Dongseok Choi and Binoy Appukuttan contributed equally to this work and share first authorship of the paper.

Electronic supplementary material The online version of this article (doi:10.1007/s12177-008-9007-1) contains supplementary material, which is available to authorized users.

D. Choi (✉)
Division of Biostatistics, Department of Public Health and Preventive Medicine, Oregon Health and Science University, Portland, OR 97239, USA
e-mail: choid@ohsu.edu

B. Appukuttan · S. J. Binek · S. R. Planck · J. T. Stout ·
J. T. Rosenbaum · J. R. Smith (✉)
Casey Eye Institute, Oregon Health and Science University, 3375 SW Terwilliger Boulevard, Portland, OR 97239, USA
e-mail: smithjus@ohsu.edu

S. R. Planck · J. T. Rosenbaum · J. R. Smith
Department of Cell and Developmental Biology, Oregon Health and Science University, Portland, OR 97239, USA

S. R. Planck · J. T. Rosenbaum
Department of Medicine, Oregon Health and Science University, Portland, OR 97239, USA

J. T. Stout
Department of Molecular Medicine, Oregon Health and Science University, Portland, OR 97239, USA

Keywords Endothelial cell · Retina · Choroid · *cis*-Regulatory motif · *cis*-Regulatory module

Introduction

Phenotypic differences exist between endothelial cells located within different extraocular vascular beds [1]. This phenomenon of “endothelial cell heterogeneity” [2] may also be observed in the human eye. Using oligonucleotide microarrays with probe sets for 8,746 transcripts, we demonstrated that approximately 10% of genes were significantly differentially expressed by donor-matched cultured human retinal and choroidal endothelial cells [3]. Consistent with these findings, we subsequently showed that protein expression differed between human retinal and choroidal endothelial cells; when retinal and choroidal endothelial protein extracts

were studied by two-dimensional difference gel electrophoresis, 25% of the 123 protein spots that were analyzed contained significantly different amounts of material derived from the retinal versus choroidal endothelial cells [4].

Differential gene expression by ocular endothelial subpopulations is likely controlled by multiple mechanisms including the interaction of transcription factors with their *cis*-regulatory motifs. Furthermore, it is expected that genes with similar expression patterns would have common motifs. This relationship is not simple, as gene expression is usually controlled by combinations of transcription factors acting via groups of *cis*-regulatory motifs, which are referred to as *cis*-regulatory modules [5]. Nonetheless, this provides a mechanism whereby activity of one or more transcription factors could result in cell type-specific expression of a subset of genes. A corollary to this statement is that existence of motifs or modules that are significantly more abundant in genes preferentially expressed in one cell type implicates activity of the respective transcription factor(s) in the regulation of gene expression of that cell specifically.

New *in silico* methods provide one avenue for comparing the regulation of gene expression between different cell types (or different states of the same cell). Biostatistics software programs that link to bioinformatics databases allow the investigator to identify and compute the abundance of *cis*-regulatory motifs or modules within the promoter region of any given gene. TRANSFAC [6], a database first conceived by Wingender [7] in 1988, is arguably the most comprehensive and widely used assembly of transcription factors and their respective *cis*-regulatory elements and regulated genes, with built-in search tools [8, 9]. CisModule is a recently published algorithm, best described as a “Bayesian module sampler,” that was developed for “discovery” of novel—or known—*cis*-regulatory modules and constituent motifs [10]. Although the literature does not contain published reports of the application of this computational approach to the investigation of human ocular endothelial cells, Waleev et al. [11] used a similar approach to study changes in gene regulation in human umbilical vein endothelial cells in response to stimulation with tumor necrosis factor alpha [12]. In this study, we performed analyses using TRANSFAC and CisModule to identify *cis*-regulatory elements and modules in genes that were differentially expressed by retinal versus choroidal endothelial cells.

Materials and methods

Gene expression microarray data

We analyzed a microarray data set that had been generated by comparing gene expression of resting retinal and choroidal

endothelial cells from eyes of three human cadaver donors using Affymetrix GeneChip Human Genome Focus Arrays (two technical replicates/cell type/donor; technical replicates were generated by preparing duplicate cultures of endothelial cells, from which RNA was extracted separately for hybridization to independent arrays). These data may be found in the Gene Expression Omnibus repository [13] (GSE7850: see retinal/choroidal endothelial cell_human_4/5/6_unstimulated_replicate1/2). Information about the experimental procedure and data normalization and processing were previously published [3]. Significant differences in gene expression between retinal and choroidal endothelial cells were determined using the significance analysis of microarrays [14] with the false discovery rate [15] set at 5%; transcripts identified by this method were further filtered to select for those that demonstrated a unidirectional and twofold or greater change across all three donors. This filtering of the data was performed to enrich the data set for genes that were clearly differentially expressed by retinal and choroidal endothelial cells, because genes that were not differentially expressed were not expected to influence the analysis. Alternative approaches would have been to study the entire group of expressed genes, including genes that were not differentially expressed, or a subset of these genes that had been chosen at random. However, manual determination of promoter sequences would have been impractical had we taken either of these approaches. We preferred to determine the promoter sequences manually (see “Collection of promoter sequences”), as this allowed us to define the length and position of the promoter sequence and, in addition, ensured that the transcription start site (TSS) was correctly identified.

Collection of promoter sequences

Human gene sequences were identified on the nucleotide website of the National Center for Biotechnology Information [16]. For each gene, nucleotide sequences, including 2,000 bp of the 5' flanking promoter region and 1,000 bp downstream of the TSS, were collected. If a particular gene encoded transcript variants with different TSSs, the defined 3,000 bp sequences for all variants were recorded. Sequences were stored as text files, with the positions of TSS, ATG start codon, exon(s), and intron(s) annotated.

Identification of *cis*-regulatory motifs

Before analysis of the promoter sequences, regions that contained interspersed repeats or were of low complexity were identified and masked using RepeatMasker [17]. The modified sequences were evaluated for the presence and abundance of known *cis*-regulatory motifs, using TRANSFAC Professional v11.4 (Biobase GmbH, Germany), with Match, to perform a matrix-based search. For any motif

identified in this way, relative abundance was compared for genes that were highly expressed in retinal versus choroidal endothelial cells by the Fisher's exact test, with a p value below 0.05 defining a significant difference in abundance.

Identification of *cis*-regulatory modules

We used CisModule [18] to predict potential *cis*-regulatory motifs and modules in the same modified promoter sequences that were studied in the previous analysis. CisModule is based on a hierarchical mixture model that recognizes the existence of the *cis*-regulatory module as a series of transcription factor binding sites within short genomic sequences and acting in concert to regulate gene expression; the algorithm, which uses Bayesian inference, involves simultaneous sampling of a promoter sequence for both putative modules and motifs [10]. Consequently, CisModule may be employed to identify both known and novel *cis*-regulatory elements and to predict *cis*-regulatory modules, by analysis of a group of co-regulated genes. We ran each set of promoter sequences separately through CisModule a total of three times with parameters set as follows: up to five motifs per module, motif length of 8 to 14 bp, and module length of up to 200 bp. The run with the highest posterior model odds for each cell subtype was selected for further analysis by TRANSFAC Professional, to identify *cis*-regulatory motifs.

Results

Promoter sequences

In the gene expression data set, 73 and 66 transcripts of the total of 8,746 transcripts represented on the array were present at significantly higher levels in retinal and choroidal endothelial cells, respectively. We retrieved genomic sequence encompassing 2,000 bp of the 5' flanking promoter region and 1,000 bp downstream of the TSS for 67 of the 73 genes that were highly expressed in retinal cells and 63 of the 66 genes that were highly expressed in choroidal cells. As we also retrieved the sequences for any genes encoding transcript isoforms with alternative TSSs, we collected, in total, 91 and 76 promoter sequences corresponding to gene transcripts that were present in relative abundance in retinal and choroidal endothelial cells, respectively. The differentially expressed genes are identified in Supplementary Tables 1 and 2.

Comparison of *cis*-regulatory motifs in genes differentially expressed by retinal and choroidal endothelial cells

The Match program of TRANSFAC Professional returned a total of 5,756 *cis*-regulatory motifs, corresponding to 159

different transcription factors, from the 91 retinal endothelial sequences, and a total of 4,973 *cis*-regulatory motifs corresponding to 158 different transcription factors, from the 76 choroidal endothelial sequences. To minimize spurious identification of false motifs, we filtered our search to identify only motifs with a "core match" of one; for any given potential binding site of 7 to 15 bp, an invariable "core" 5-bp sequence had to be present. In this fashion, we reduced the number to 2,879 motifs in retinal endothelial sequences and 2,597 motifs in choroidal endothelial sequences, corresponding to 146 and 147 transcription factors, respectively. A total of 135 transcription factors were identified in both retinal and choroidal endothelial sequences, while 23 transcription factors were identified in relation to either retinal or choroidal endothelial, but not both, sequences. Transcription factors and corresponding motif consensus core sequences (defined by International Union of Pure and Applied Chemistry [IUPAC] nucleotide code) [19] that were most abundant in retinal endothelial promoter sequences were (1) zinc finger 5 (ZF5), CGCGC ($n=140$); (2) avian myeloblastosis viral oncogene homolog (*v-Myb*), AACGG ($n=112$); (3) Hand1:E47, TCTGG ($n=91$); (4) epidermal growth factor receptor-specific transcription factor (ETF), GGCGG ($n=87$); and (5) transcription factor AP-2 (TFAP2), CMSGC ($n=72$). Transcription factors and motif consensus core sequences that were most abundant in choroidal endothelial promoter sequences were (1) ZF5, CGCGC ($n=179$); (2) ETF, GGCGG ($n=106$); (3) *v-Myb*, AACGG ($n=89$); (4) MYC-associated zinc finger protein (MAZ), GGGAG ($n=75$); (5) and Hand1:E47, TCTGG ($n=71$). Table 1 shows the distribution of binding sites by strand and by location. The transcription binding sites were located approximately 1:1, forward strand/reverse strand and approximately 3:2, 2,000 bp upstream/1,000 bp downstream of the TSS.

We used Fisher's exact test to compare the abundance of *cis*-regulatory motifs, corresponding to each of the transcription factors identified in the previous analysis, in the promoter regions of differentially expressed retinal and choroidal endothelial genes. Before testing, for genes having transcript variants with different TSSs, we collapsed all isoforms into one entry by taking the maximum frequencies of each motif identified for any isoform. For five transcription factors, motifs were significantly more abundant in retinal endothelial promoter sequences. Transcription factors and corresponding motif consensus core sequences (defined by IUPAC nucleotide code) [19] were (1) glucocorticoid receptor (GCCR; $p=0.015$), GTTCT; (2) high mobility group AT-hook 1 (HMG1Y; $p=0.022$), GGAAA; (3) heat shock transcription factor 1 (HSF1; $p=0.025$), AGAAY; (4) p53 ($p=0.025$), CATGY; and (5) vitamin D receptor (VDR; $p=0.024$), GGGTS and TGAMC. Motifs recognized by three transcription factors

Table 1 Distribution of *cis*-regulatory motifs in promoter sequences of genes that are differentially expressed by human retinal and choroidal endothelial cells

	Number of motifs with perfect core match	Number of corresponding transcription factors	Number of motifs on forward strand	Number of motifs on reverse strand	Number of motifs upstream of TSS	Number of motifs downstream of TSS	Number of motifs within specified region (in relation to TSS)					
							-2,000 to -1,500	-1,000 to -500	-500 to +1	+1 to +500	+500 to +1000	
							to	to	to	to	to	
Genes with ↑ expression in retinal EC	2,879	146	1,395	1,484	1,708	1,171	327	356	381	644	602	569
Percent of total			48%	52%	59%	41%	11%	12%	13%	22%	21%	20%
Density/sequence			15.3	16.3	18.8	12.9	3.6	3.9	4.2	7.1	6.6	6.3
Density/1000bp/sequence			5.1	5.4	9.4	12.9	7.2	7.8	8.4	14.2	13.2	12.5
Genes with ↑ expression in choroidal EC	2,597	147	1,311	1,286	1,593	1,004	346	322	436	489	498	506
Percent of total			50%	50%	61%	39%	13%	12%	17%	19%	19%	19%
Density/sequence			16.8	16.5	20.4	12.9	4.4	4.1	5.6	6.3	6.4	6.5
Density/1000bp/sequence			5.6	5.5	10.2	12.9	8.9	8.3	11.2	12.5	12.8	13.0
All genes	5,476	158	2,706	2,770	3,301	2,175	673	678	817	1,133	1,100	1,075
Percent of total			49%	50%	60%	39%	12%	12%	15%	21%	20%	19%
Density/sequence			16.0	16.4	19.5	12.9	4.0	4.0	4.8	6.7	6.5	6.4
Density/1000bp/sequence			5.3	5.5	9.8	12.9	8.0	8.0	9.7	13.4	13.0	12.7

EC endothelial cell, TSS transcription start site, bp base pairs

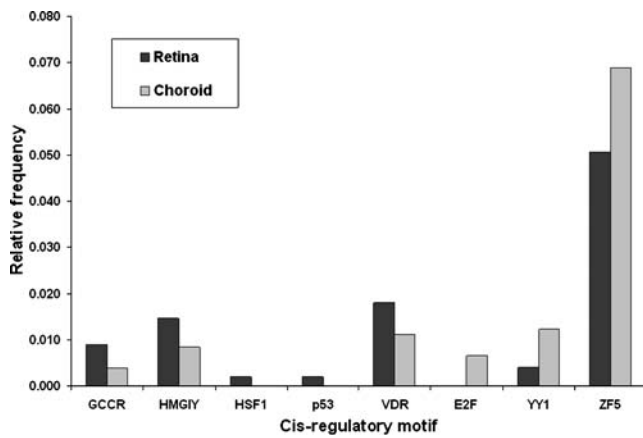


Fig. 1 Relative frequency of *cis*-regulatory motifs that were significantly over-represented within the promoters of genes differentially expressed by retinal versus choroidal endothelial cells. Frequency was determined with respect to total number of *cis*-regulatory motifs within all genes studied for each cell subtype. Motifs are identified according to the corresponding transcription factors (*GCCR* glucocorticoid receptor, *HMGIY* high mobility group AT-hook 1, *HSF1* heat shock transcription factor 1, *VDR* vitamin D receptor, *E2F* transcription factor E2F, *YY1* Yin Yang 1, *ZF5* zinc finger 5)

were significantly more abundant in choroidal endothelial promoter sequences. These transcription factors and motif consensus core sequences were (1) transcription factor E2F (E2F; $p < 0.001$), TGGCG and CGGCA; (2) Yin Yang 1 (YY1; $p = 0.001$), GCCAT; and (3) ZF5 ($p = 0.006$), CGCGC. In addition, the analysis identified motifs recognized by TFAP2 as being more abundant in retinal endothelial promoter sequences, and, simultaneously, it identified motifs recognized by TFAP2-alpha as being more abundant in choroidal endothelial promoter sequences; as TFAP2-alpha is a member of the TFAP2 family, we attributed this result to a database anomaly and did not consider it valid for further consideration. Supplementary Table 3 presents the total number of motifs in retinal and choroidal endothelial promoter sequences corresponding to the eight transcription factors identified as described above. Figure 1 illustrates graphically the relative frequencies of motifs recognized by these transcription factors.

Identification of putative *cis*-regulatory modules in genes differentially expressed by retinal and choroidal endothelial cells

We next employed CisModule to identify potential *cis*-regulatory modules, comprised of up to five *cis*-regulatory motifs, in the promoter regions of differentially expressed retinal and choroidal genes. The algorithm detected 220 representations of a putative module in retinal endothelial sequences and 441 representations of a putative module in choroidal endothelial sequences. The predicted *cis*-regulatory

modules for promoter sequences in genes highly expressed by retinal or choroidal endothelial cells were different. Table 2 shows the predicted modules for retinal and choroidal sequences and possible identities for respective constituent *cis*-regulatory motifs, obtained by submitting the motif sequences to TRANSFAC Professional. Nine of the ten predicted motifs were identified, and there was one novel motif. Among the 67 genes that were more highly expressed in retinal endothelial cells, 54 genes (81%) were identified as containing one or more of the predicted modules, while 63 of 63 genes (100%) that were more highly expressed in choroidal endothelial cells were similarly identified. Table 3 presents these data and indicates the number of motifs that composed the predicted retinal and choroidal endothelial modules. Supplementary Tables 4 and 5 give the position weight matrices of the predicted motifs with consensus sequences, and Fig. 2 shows the sequence logos. A sequence logo is a graphical representation of a sequence; at any given position, the height of each letter within a stack indicates the relative frequency of the corresponding nucleotide, and the total height of stack at any given position indicates how informative that position is within the binding site [20].

Discussion

Previously, we compared the relative abundance of gene transcripts between endothelial cells isolated from retina and choroid by microarray, to characterize vascular endothelial diversity within the eye [3]. In this study, we utilized a computational approach to identify *cis*-regulatory motifs and transcription factors that might account for the differences in gene expression we observed. Particular *cis*-regulatory motifs may cluster together within a promoter region as modules that control cell-specific gene expression on the binding of transcription factors [5]. Using the TRANSFAC database and the CisModule software, we have identified a number of *cis*-regulatory motifs and modules that appear to regulate gene expression in either retinal or choroidal endothelial cells. Identification of such promoter elements may have implications for the pathogenesis of diseases that specifically involve these endothelial cells, including proliferative diabetic retinopathy and neovascular age-related macular degeneration (AMD).

Mitochondrial dysfunction, overproduction of vascular endothelial growth factor (VEGF), and endothelial cell proliferation are key events in the pathogenesis of choroidal neovascularization in AMD [21, 22]. Interestingly, *cis*-regulatory motifs for E2F, YY1, and ZF5 are prevalent in promoters of genes that are preferentially expressed in choroidal endothelial cells. All three transcription factors could potentially influence directly or, via a signaling pathway, one or more of the aforementioned events. When

Table 2 Predicted *cis*-regulatory modules and constituent motifs, with possible identities of corresponding transcription factors as determined by TRANSFAC Professional, for genes that were highly expressed by human retinal or choroidal endothelial cells

	Motif or Module	Consensus sequence	Total number	Possible transcription factor
Genes with ↑ expression in retinal EC	Motif 1	GTCCCTGGGCTCTG	159	GATA4
	Motif 2	CCAGCGGCGCCAG	398	ZF5
	Motif 3	GGGTGGGGTGGGGG	105	SREBF1
	Motif 4	CAGGGCAGGAGG	249	Unidentified
	Motif 5	GGGGGGGGCAGGG	591	ZBTB7
	Module		221	
Genes with ↑ expression in choroidal EC	Motif 1	AAAAAAAAAATAAA	849	TBP
	Motif 2	AAAAAAAAAAAAAAAA	254	FALZ
	Motif 3	CCTCTCCCCCCCC	511	WT1
	Motif 4	TGTGTGTGTGTGTG	153	Pax-4
	Motif 5	CCCCGCGTCCCC	1,229	WT1
	Module		441	

If TRANSFAC Professional identified multiple transcription factors that could bind to a particular consensus sequence, the factor with core match and matrix match equal to or closest to one was selected for presentation.

EC endothelial cell, *GATA-4* GATA-binding factor 4, *ZF5* zinc finger 5, *SREBF1* sterol regulatory element-binding factor 1, *ZBTB7* zinc finger- and BTB domain-containing protein 7, *TBP* TATA box-binding protein, *FALZ* fetal Alzheimer antigen, *WT1* Wilms' tumor zinc finger protein, *Pax-4* paired box gene 4

neovascularization is induced in the E2F1 gene-deficient mouse, either by hind limb ischemia or tumor challenge, vascular proliferation is exacerbated, as a result of the production of excess VEGF; E2F1 interacts with p53 under hypoxic conditions to repress VEGF transcription [23]. In the mouse, ZF5 represses transcription of the *myc* gene [24], whose product regulates the activity of specific members of the E2F family of proteins to induce either cell proliferation or cell death [25]. Loss of cellular YY1 leads to a diminution of mitochondrial oxidative function via a pathway involving the mammalian target of rapamycin (mTOR) [26]. Overexpression of mTOR augments the activity of hypoxia-inducible factor (HIF)-1 alpha, and secretion of VEGF under conditions of low oxygen tension, thereby increasing the growth of new vessels [27].

Dysregulation of the transcription factors that interact with E2F, YY1, and ZF5, or polymorphisms within the *cis*-regulatory motifs themselves, might influence the onset or progression of neovascular AMD.

cis-Regulatory motifs that are more prevalent in the promoters of genes preferentially expressed in retinal endothelial cells (i.e., *GCCR*, *HMG1Y*, *HSF1*, *p53*, and *VDR*) interact with transcription factors that regulate cell proliferation, apoptosis, aging, and the stress response; all these processes are pertinent to neovascularization in diabetic retinopathy. Transcription factor, *HSF1*, is regulated in response to cellular stress and promotes the removal of accrued unfolded peptides from the endoplasmic reticulum [28]. This process may be perturbed in diabetes mellitus [29]. As proliferative diabetic retinopathy involves

Table 3 Number of predicted *cis*-regulatory modules and motifs in promoter sequences of genes that were differentially expressed by human retinal and choroidal endothelial cells

	Promoter sequence	Number (% of total)					
		Module	Motif 1	Motif 2	Motif 3	Motif 4	Motif 5
Genes with ↑ expression in retinal EC	Genes	53	42	32	52	52	54
	(67 total)	79%	63%	48%	78%	78%	81%
	Genes + isoforms ^a	69	57	46	62	64	68
	(91 total)	76%	63%	51%	68%	70%	75%
Genes with ↑ expression in choroidal EC	Genes	63	49	63	51	49	58
	(63 total)	100%	78%	100%	81%	78%	92%
	Genes + isoforms ^a	75	60	72	62	62	70
	(76 total)	99%	79%	95%	82%	82%	92%

EC endothelial cell

^a Promoter sequences corresponding to transcript isoforms with alternative TSSs were included in this analysis

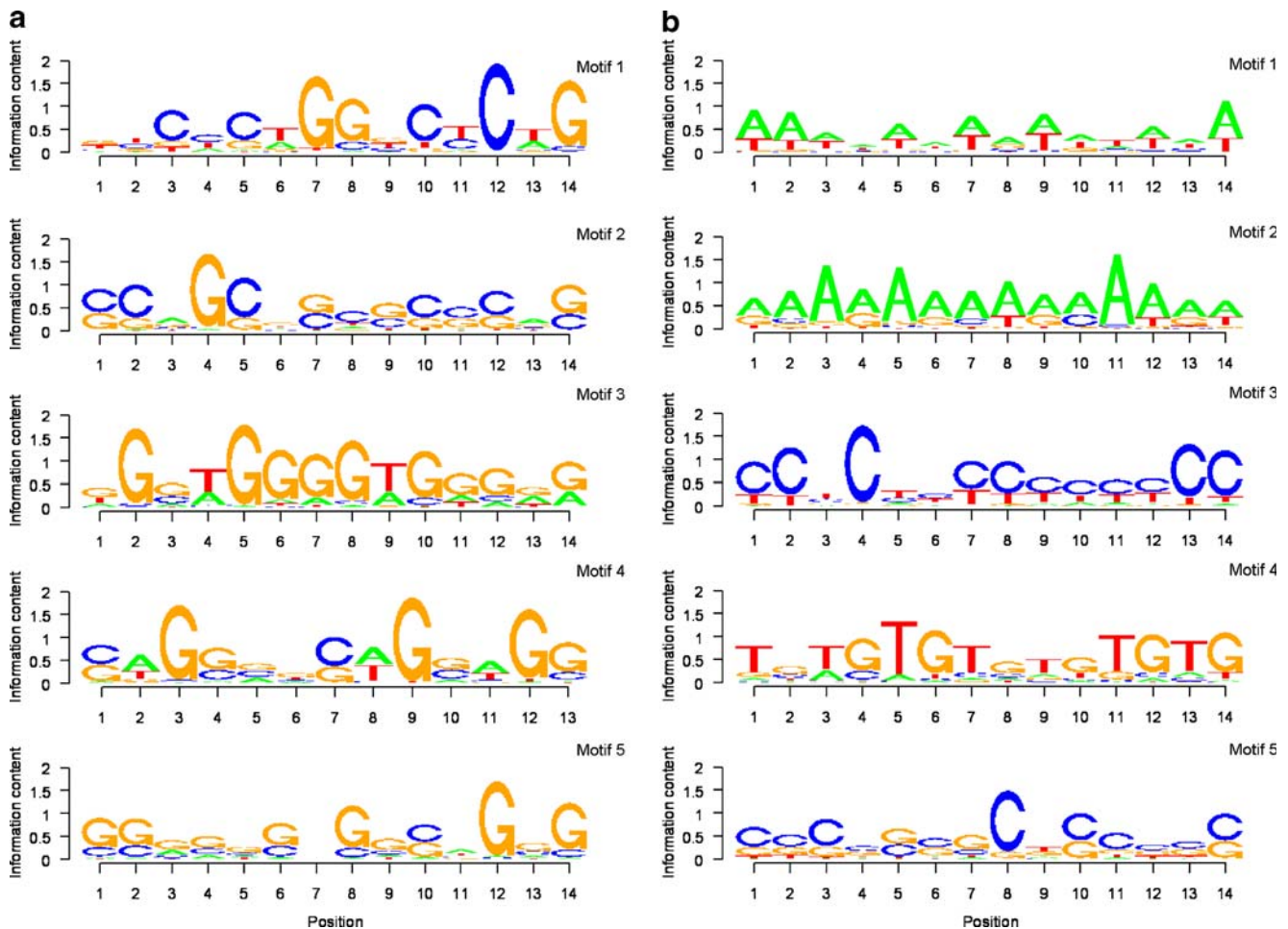


Fig. 2 Sequence logos for predicted *cis*-regulatory modules in **a** genes that were expressed at significantly higher levels in retinal endothelial cells and **b** genes that were expressed at significantly higher levels in choroidal endothelial cells

abnormal growth of retinal blood vessels, it is intriguing that a low level of HMGA1 leads to a decrease of cell-surface insulin receptors in patients with insulin resistance and type II diabetes [30]. Moreover, an overproduction of HMGA1 protein can lead to cell proliferation and inhibition of apoptosis mediated by p53 [31]. p53 has many activities, including induction of apoptosis in cells with damaged DNA, anti-oxidative effects with implications for the reduction of oxidative damage associated with aging, and anti-angiogenic effects via inhibition of HIF-1 alpha [32, 33]. The carboxyl region of VDR binds the active form of vitamin D, while the amino region binds to the *cis*-regulatory motif [34]. Activation of VDR by vitamin D prevents the proliferation of endothelial cells isolated from tumors [35], and, consistent with this observation, vitamin D treatment inhibits angiogenesis in a mouse model of retinoblastoma [36]. Interestingly, a polymorphism within the VDR gene, believed to alter the ability of the activity of the transcription factor, is associated with type I diabetes [37]. Like VDR, GCCR is a hormone-activated transcrip-

tion factor [38]. Activation of GCCR also suppresses neoplasia-associated angiogenesis *in vivo*, an effect that results from decreased levels of VEGF and interleukin-8 [39]. Interestingly, both VDR and GCCR are differentially expressed by different extraocular endothelial cell subpopulations [40, 41].

As with any *in silico* research, it will be important to confirm findings in *in vitro* and/or *in vivo* biological systems. For example, one might construct reporter vectors containing multiple copies of a *cis*-regulatory motif or a combination of these motifs for transfection of retinal and choroidal endothelial cells, to evaluate cell-specific interactions of transcription factor(s) and promoter element(s). It is also important to be aware that the output from a computational study of this nature depends very much on the selection of sequences for analysis. For example, we initially performed analysis of the same genes that were used for the present study, with Promoter Analysis and Interaction Network Tool (PAINT) v3.5 [35], an alternative program for identifying *cis*-regulatory motifs within a

genomic sequence. PAINT predicts the TSS position along the gene sequence and identifies *cis*-regulatory motifs upstream of this forecasted site [42]. In the pilot study with PAINT, we examined the 2,000-bp upstream region of the predicted TSS and obtained a different set of *cis*-regulatory motifs to those identified in this project (data not shown). In the current study, we analyzed both 2,000-bp downstream and 1,000-bp downstream of the predicted TSS, as sequences downstream of the TSS, including intronic sequences, may contain *cis*-regulatory elements [43]. In addition, we more accurately determined the position of the TSS by manually collecting promoter sequences from the literature. The same general principles apply to the prediction of *cis*-regulatory modules by CisModule. An apparent contradiction is that *cis*-regulatory motifs recognized by ZF5 were more abundant in genes highly expressed by choroidal endothelial cells and, yet, also present within a predicted retinal endothelial *cis*-regulatory module (Table 2). However, the result of the CisModule analysis implies that ZF5 may preferentially regulate gene expression in retinal endothelial cells only when acting in concert with other transcription factors. One novel putative motif was identified within the retinal *cis*-regulatory module (Table 2). If nuclear protein extract isolated from retinal endothelial cells specifically binds this novel motif, it would be of immediate interest to identify the corresponding transcription factor(s).

Identifying and characterizing retinal and choroidal endothelial cell-specific gene expression may have implications for the treatment of vascular diseases of the eye. Regarding the proliferating choroidal endothelial cells responsible for neovascular AMD, the logical goal of treatment is to inhibit division and/or induce apoptosis of these cells alone. The same reasoning applies to retinal endothelial cells and diabetic retinopathy. Delivery of a therapeutic transgene under the control of a promoter with *cis*-regulatory motifs that are preferentially activated in one endothelial population could potentially enhance gene expression exclusively in those cells, avoiding detrimental effects to neighboring non-diseased cells, including other endothelial populations. Retinal cell-specific promoters and enhancers have been characterized and used to achieve cell-specific expression within the retina [44, 45]. Endothelial cell-specific promoter elements isolated from the human VEGF-R1 gene effectively target transgene expression to the ocular vasculature when injected intravitreally [46], and enhancer elements that specifically limit expression to proliferating endothelial cells have been identified [47]. A promoter construct that combines specific retinal or choroidal endothelial *cis*-regulatory motifs and element(s) that initiate expression in proliferating cells could provide a safe and effective new treatment for proliferative diabetic retinopathy or neovascular AMD, respectively.

Acknowledgments This work was supported in part by grants from the National Institutes of Health (EY014909, EY010572), Research to Prevent Blindness (i.e., Career Development Award to JRS, Scholar Award to SRP, Robert E. McCormick Scholar Award to JTS, Senior Scholar Award to JTR, and unrestricted grant to Casey Eye Institute) and the Clayton Research Foundation.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Chi JT, Chang HY, Haraldsen G et al. Endothelial cell diversity revealed by global expression profiling. *Proc Natl Acad Sci USA*. 2003;100(19):10623–8.
- Aird WC. Phenotypic heterogeneity of the endothelium: I. Structure, function, and mechanisms. *Circ Res*. 2007;100(2):158–73.
- Smith JR, Choi D, Chipps TJ et al. Unique gene expression profiles of donor-matched human retinal and choroidal vascular endothelial cells. *Invest Ophthalmol Vis Sci*. 2007;48(6):2676–84.
- Zamora DO, Riviere M, Choi D et al. Proteomic profiling of human retinal and choroidal endothelial cells reveals molecular heterogeneity related to tissue of origin. *Mol Vis*. 2007;13:2058–65.
- Istrail S, De-Leon SB, Davidson EH. The regulatory genome and the computer. *Dev Biol*. 2007;310(2):187–95.
- BIOBASE. TRANSFAC databases. www.biobase.de/pages/index.php?id=transfacdatabases.
- Wingender E. Compilation of transcription regulating proteins. *Nucleic Acids Res*. 1988;16(5):1879–902.
- Knuppel R, Dietze P, Lehnberg W et al. TRANSFAC retrieval program: a network model database of eukaryotic transcription regulating sequences and proteins. *J Comput Biol*. 1994;1(3):191–8.
- Matys V, Kel-Margoulis OV, Fricke E et al. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. *Nucleic Acids Res*. 2006;34(Database issue):D108–10.
- Zhou Q, Wong WH. CisModule: de novo discovery of cis-regulatory modules by hierarchical mixture modeling. *Proc Natl Acad Sci U S A*. 2004;101(33):12114–9.
- Waleev T, Shtokalo D, Konovalova T et al. Composite Module Analyst: identification of transcription factor binding site combinations using genetic algorithm. *Nucleic Acids Res*. 2006;34(Web Server issue):W541–5.
- Viemann D, Goebeler M, Schmid S et al. Transcriptional profiling of IKK2/NF-kappa B- and p38 MAP kinase-dependent gene expression in TNF-alpha-stimulated primary human endothelial cells. *Blood*. 2004;103(9):3365–73.
- NCBI. GEO Datasets. www.ncbi.nlm.nih.gov/sites/entrez?db=gds.
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98(9):5116–21.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Stat Soc, Series B: Methodological*. 1995;57:289–300.
- NCBI. Human Genome. <http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide>.
- Institute for Systems Biology. RepeatMasker. www.repeatmasker.org.
- CisModule: A Bayesian module sampler by hierarchical mixture modeling. www.stat.ucla.edu/~zhou/CisModule.

19. Cornish-Bowden A. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Res.* 1985;13(9):3021–30.
20. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res.* 1990;18(20):6097–100.
21. Ambati J, Ambati BK, Yoo SH et al. Age-related macular degeneration: etiology, pathogenesis, and therapeutic strategies. *Surv Ophthalmol.* 2003;48(3):257–93.
22. Nowak JZ. Age-related macular degeneration (AMD): pathogenesis and therapy. *Pharmacol Rep.* 2006;58(3):353–63.
23. Qin G, Kishore R, Dolan CM et al. Cell cycle regulator E2F1 modulates angiogenesis via p53-dependent transcriptional control of VEGF. *Proc Natl Acad Sci U S A.* 2006;103(29):11015–20.
24. Sobek-Klocke I, Disque-Kochem C, Ronsiek M et al. The human gene ZFP161 on 18p11.21-pter encodes a putative c-myc repressor and is homologous to murine Zfp161 (Chr 17) and Zfp161-rs1 (X Chr). *Genomics.* 1997;43(2):156–64.
25. Leone G, Sears R, Huang E et al. Myc requires distinct E2F activities to induce S phase and apoptosis. *Mol Cell.* 2001;8(1):105–13.
26. Cunningham JT, Rodgers JT, Arlow DH et al. mTOR controls mitochondrial oxidative function through a YY1-PGC-1alpha transcriptional complex. *Nature.* 2007;450(7170):736–40.
27. Land SC, Tee AR. Hypoxia-inducible factor 1alpha is regulated by the mammalian target of rapamycin (mTOR) via an mTOR signaling motif. *J Biol Chem.* 2007;282(28):20534–43.
28. Westerheide SD, Morimoto RI. Heat shock response modulators as therapeutic tools for diseases of protein conformation. *J Biol Chem.* 2005;280(39):33097–100.
29. Haynes CM, Titus EA, Cooper AA. Degradation of misfolded proteins prevents ER-derived oxidative stress and cell death. *Mol Cell.* 2004;15(5):767–76.
30. Foti D, Chiefari E, Fedele M et al. Lack of the architectural factor HMGA1 causes insulin resistance and diabetes in humans and mice. *Nat Med.* 2005;11(7):765–73.
31. Pierantoni GM, Rinaldo C, Mottolise M et al. High-mobility group A1 inhibits p53 by cytoplasmic relocalization of its proapoptotic activator HIPK2. *J Clin Invest.* 2007;117(3):693–702.
32. Caelles C, Helmborg A, Karin M. p53-Dependent apoptosis in the absence of transcriptional activation of p53-target genes. *Nature.* 1994;370(6486):220–3.
33. Sano M, Minamino T, Toko H et al. p53-Induced inhibition of Hif-1 causes cardiac dysfunction during pressure overload. *Nature.* 2007;446(7134):444–8.
34. Hughes MR, Malloy PJ, Kieback DG et al. Point mutations in the human vitamin D receptor gene associated with hypocalcemic rickets. *Science.* 1988;242(4886):1702–5.
35. Bernardi RJ, Johnson CS, Modzelewski RA, Trump DL. Antiproliferative effects of 1alpha,25-dihydroxyvitamin D3 and vitamin D analogs on tumor-derived endothelial cells. *Endocrinology.* 2002;143(7):2508–14.
36. Shokravi MT, Marcus DM, Alroy J et al. Vitamin D inhibits angiogenesis in transgenic murine retinoblastoma. *Invest Ophthalmol Vis Sci.* 1995;36(1):83–7.
37. Motohashi Y, Yamada S, Yanagawa T et al. Vitamin D receptor gene polymorphism affects onset pattern of type 1 diabetes. *J Clin Endocrinol Metab.* 2003;88(7):3137–40.
38. Muller M, Renkawitz R. The glucocorticoid receptor. *Biochim Biophys Acta.* 1991;1088(2):171–82.
39. Yano A, Fujii Y, Iwai A et al. Glucocorticoids suppress tumor angiogenesis and in vivo growth of prostate cancer cells. *Clin Cancer Res.* 2006;12(10):3003–9.
40. Forster C, Waschke J, Burek M et al. Glucocorticoid effects on mouse microvascular endothelial barrier permeability are brain specific. *J Physiol.* 2006;573(Pt 2):413–25.
41. Nakagawa K, Sasaki Y, Kato S et al. 22-Oxa-1alpha,25-dihydroxyvitamin D3 inhibits metastasis and angiogenesis in lung cancer. *Carcinogenesis.* 2005;26(6):1044–54.
42. Vadigepalli R, Chakravarthula P, Zak DE et al. PAINT: a promoter analysis and interaction network generation tool for gene regulatory network identification. *Omics.* 2003;7(3):235–52.
43. Schroeder MD, Pearce M, Fak J et al. Transcriptional control in the segmentation gene network of *Drosophila*. *PLoS Biol.* 2004;2(9):E271.
44. Flannery JG, Zolotukhin S, Vaquero MI et al. Efficient photoreceptor-targeted gene expression in vivo by recombinant adeno-associated virus. *Proc Natl Acad Sci U S A.* 1997;94(13):6916–21.
45. Mancuso K, Hendrickson AE, Connor TB Jr. et al. Recombinant adeno-associated virus targets passenger gene expression to cones in primate retina. *J Opt Soc Am A Opt Image Sci Vis.* 2007;24(5):1411–6.
46. Luz-Madriral A, Clapp C, Aranda J, Vaca L. In vivo transcriptional targeting into the retinal vasculature using recombinant baculovirus carrying the human flt-1 promoter. *Virology.* 2007;4:88.
47. Shaw LC, Pan H, Afzal A et al. Proliferating endothelial cell-specific expression of IGF-I receptor ribozyme inhibits retinal neovascularization. *Gene Ther.* 2006;13(9):752–60.