

Analysis of VEGF-A Regulated Gene Expression in Endothelial Cells to Identify Genes Linked to Angiogenesis

Corban G. Rivera^{1,2*}, Sofie Mellberg³, Lena Claesson-Welsh³, Joel S. Bader^{1,2}, Aleksander S. Popel¹

1 Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland, United States of America, **2** High-Throughput Biology Center, Johns Hopkins School of Medicine, Baltimore, Maryland, United States of America, **3** Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden

Abstract

Angiogenesis is important for many physiological processes, diseases, and also regenerative medicine. Therapies that inhibit the vascular endothelial growth factor (VEGF) pathway have been used in the clinic for cancer and macular degeneration. In cancer applications, these treatments suffer from a “tumor escape phenomenon” where alternative pathways are upregulated and angiogenesis continues. The redundancy of angiogenesis regulation indicates the need for additional studies and new drug targets. We aimed to (i) identify novel and missing angiogenesis annotations and (ii) verify their significance to angiogenesis. To achieve these goals, we integrated the human interactome with known angiogenesis-annotated proteins to identify a set of 202 angiogenesis-associated proteins. Across endothelial cell lines, we found that a significant fraction of these proteins had highly perturbed gene expression during angiogenesis. After treatment with VEGF-A, we found increasing expression of HIF-1 α , APP, HIV-1 tat interactive protein 2, and MEF2C, while endoglin, liprin β 1 and HIF-2 α had decreasing expression across three endothelial cell lines. The analysis showed differential regulation of HIF-1 α and HIF-2 α . The data also provided additional evidence for the role of endothelial cells in Alzheimer’s disease.

Citation: Rivera CG, Mellberg S, Claesson-Welsh L, Bader JS, Popel AS (2011) Analysis of VEGF-A Regulated Gene Expression in Endothelial Cells to Identify Genes Linked to Angiogenesis. *PLoS ONE* 6(9): e24887. doi:10.1371/journal.pone.0024887

Editor: Sumitra Deb, Virginia Commonwealth University, United States of America

Received: March 8, 2011; **Accepted:** August 23, 2011; **Published:** September 13, 2011

Copyright: © 2011 Rivera et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The work was supported by National Institutes of Health grants R01 HL101200 and R01 CA138264. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: cgrivera@jhu.edu

Introduction

Angiogenesis has been implicated in a wide spectrum of diseases including cancer, age-related macular degeneration, rheumatoid arthritis, diabetic nephropathy, pathologic obesity, and asthma. Compounds that inhibit angiogenesis represent potential therapeutics for many diseases. Judah Folkman performed pioneering research in the field of angiogenesis [1]; his work led to the identification of a number of proteins and polypeptides with anti-angiogenic activity [2]. Since then, many compounds have entered clinical trials as modulators of angiogenesis. Targeting the vascular endothelial growth factor (VEGF) pathway has been the leading anti-angiogenic strategy in the clinic [3]. Other anti-angiogenic targets include integrins [4] and angiopoietin 1 [5]. While these therapies suppress new blood vessel growth for a period of time, anti-angiogenic therapies suffer from the upregulation of compensating pathways that circumvent the inhibited pathway. Consequently, the field is in need of a comprehensive understanding of the proteins and pathways involved in angiogenesis.

Identifying angiogenesis-associated proteins is related to the problem of gene function prediction. Over the past decade, there have been many important studies on using network structure to functionally annotate gene products. These methods were reviewed in [6,7]. Early methods transferred gene functions from direct neighbors to annotate genes. Extensions to these methods allowed more distant annotations through shortest paths [8] and diffusion [9,10,11,12]. Some of these methods make use of diverse

machine learning techniques including SVMs [9] and Bayesian networks [10].

There has been recent interest in the area of applying bioinformatics approaches to study angiogenesis. These studies integrate gene expression data with molecular interactions with diverse goals. A recent study explored pathways associated with VEGF [11]. Another study validated their new method by recovering pathways known to be involved in angiogenesis [12]. A third study aimed to predict the impact of anti-angiogenic kinase inhibitors [13]. A previous study conducted in our lab identified crosstalk between angiogenesis modulating protein families [14]. These studies suggest momentum towards using bioinformatics to study angiogenesis. However none of these studies had the goal of finding new and missing angiogenesis gene annotations. We aimed to: (i) identify novel and missing angiogenesis annotations and (ii) verify their association with angiogenesis with statistical analysis and multiple gene expression datasets.

Results and Discussion

In Figure 1, we describe our computational strategy to identify and validate novel angiogenesis-associated proteins. To ensure that the first objective is satisfied, the basis for our search began with proteins known to be annotated with angiogenesis from the Gene Ontology database (GO). We integrated molecular interactions with the angiogenesis-annotated proteins to identify new highly associated proteins. We satisfied the second objective with

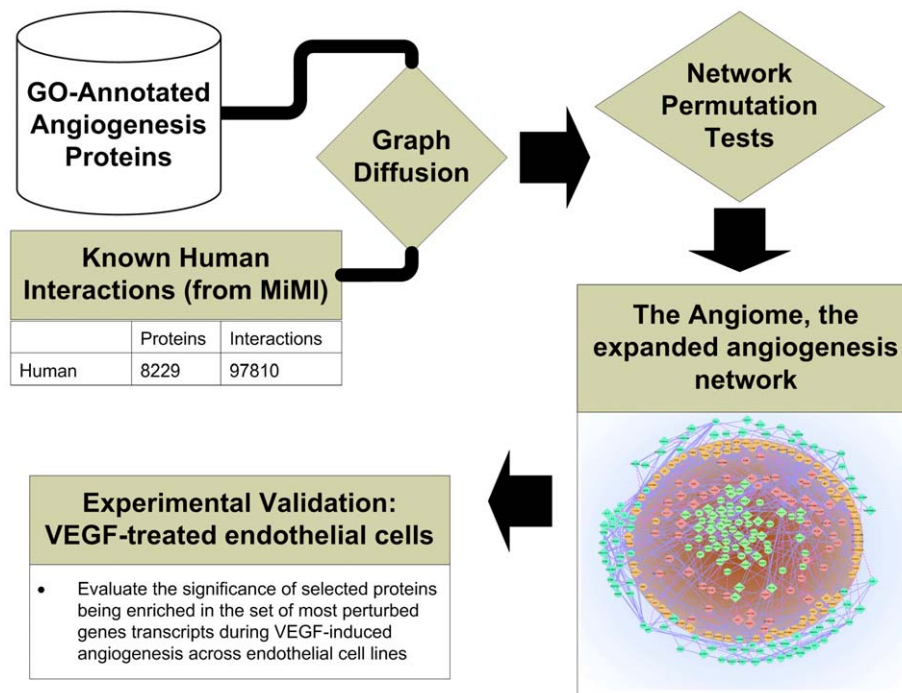


Figure 1. Computational strategy to identify and verify novel angiogenesis-associated proteins. We began by integrating known angiogenesis-annotated proteins (GO-annotated) with a set of known human interactions (from MiMI). We used a graph diffusion approach to identify new angiogenesis-associated proteins using a “guilt-by-association” approach. We verified the set of selected genes using two approaches. First, we evaluated the statistical significance of the selected genes by repeatedly permuting the human interactome (statistical hypothesis given in Methods). Secondly, we integrated a time series gene expression experimental dataset obtained during VEGF-induced angiogenesis to determine if the selected genes are enriched in the set of most perturbed transcripts. doi:10.1371/journal.pone.0024887.g001

statistical analysis of three published datasets (see Table 1) that showed the newly associated proteins frequently had perturbed gene expression during angiogenesis. We combined the angiogenesis annotated and associated sets of proteins to form the *Angiome*, the network of protein interactions associated with angiogenesis.

Angiogenesis annotation by graph diffusion

To determine the quality of graph diffusion annotations, we compared graph diffusion to several basic methods for transferring functional annotations including a first neighbors method and a second neighbours method. Graph diffusion (see Methods) works by ranking proteins by their interactions with annotated proteins over paths of all lengths. The accumulation of evidence from all path lengths generally leads to better performance. The first neighbors method ranks proteins by the number of direct

interactions with annotated proteins. The second neighbors method ranks proteins by the number of annotated proteins reachable by paths of length less than three. To evaluate the performance of these methods for angiogenesis annotations, we performed a leave-one-out cross-validation (LOOCV) procedure. In Figure 2, we show the performance of the diffusion kernel for angiogenesis annotation. The figure shows the receiver operator characteristic (ROC) curve and a precision recall curve for all methods. The performance of the methods can be summarized by the area under the ROC curve (AUC). Greater AUC indicates better performance. Graph diffusion achieves an AUC of 0.76, while the first and second neighbor methods achieved an AUC of 0.66 and 0.58 respectively. Based on this data, we proceeded to make angiogenesis annotations using graph diffusion.

Validation by network randomization

The highest ranking proteins by graph diffusion are topologically closer to the annotated proteins; however the method can be biased towards hubs and other highly connected proteins. To eliminate this bias, we tested the statistical significance of a protein achieving as high of an association score in randomized networks. For a given protein, we test the null hypothesis that the score is equal to the rank that could be achieved in a randomized network with identical degree distribution. We generated 300 randomized networks by repeated edge swapping. The probability of the null hypothesis is given by the fraction of randomized networks where the score of the protein exceeds the score from the real network. We reject the null hypothesis if the probability is less than 0.01. Using this procedure, we identified a set of 202 angiogenesis associated proteins.

Table 1. VEGF stimulated time series gene expression datasets analyzed.

Ref.	Source	Cells	Culture	Timepoints
[15]	GDS3567	HUVEC	Gelatine	0 m, 30 m, 1 h, 2.5 h
[16]	GDS2039	HMVEC	Matrigel	30 m, 2 h, 2 h, 4 h, 8 h 15 m, 1 h, 3 h, 6 h, 8 h, 12 h, 18 h
[17]	Mellberg et al.	TIME	Collagen I	24 h

Each dataset measured VEGF induced changes in endothelial cell gene expression over a time course. The datasets used different cell lines, cell culture, and time points for their studies.

doi:10.1371/journal.pone.0024887.t001

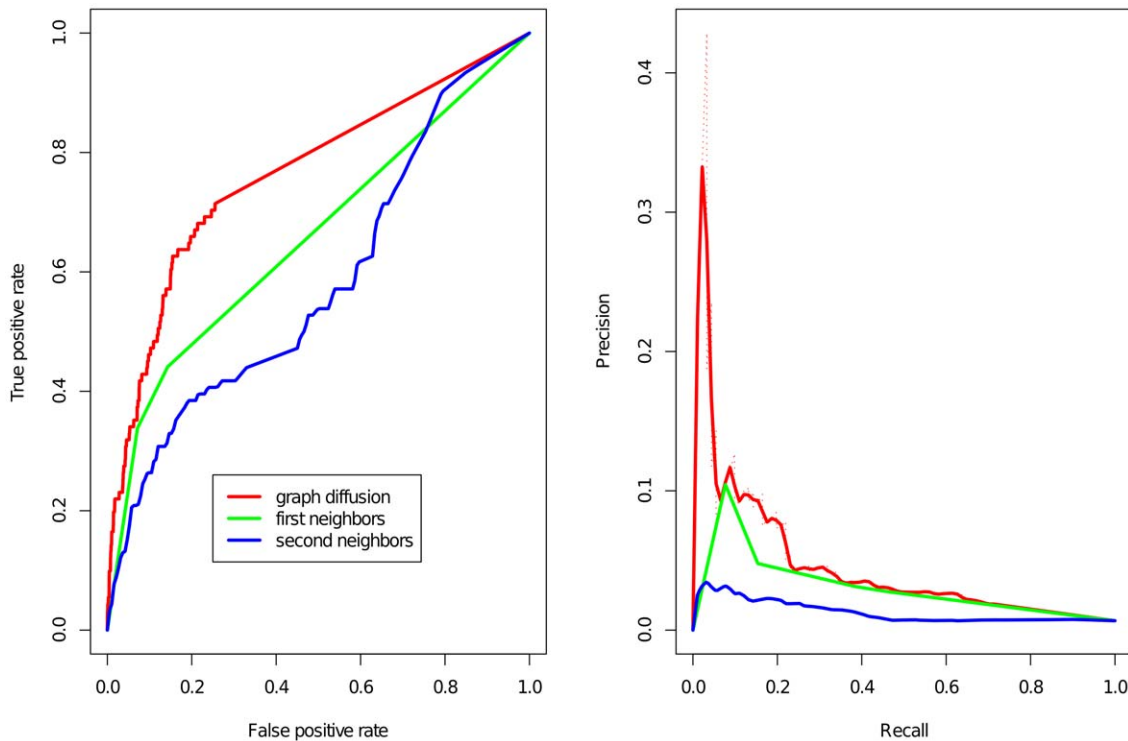


Figure 2. Graph diffusion for angiogenesis annotation. The receiver operator characteristic (ROC) curve and precision recall curve detailing the leave-one-out cross validation prediction accuracy of angiogenesis annotation. Using graph diffusion the area under the curve (AUC) was 0.76. Annotation by the first neighbors method achieved 0.66, and second neighbors method achieved 0.58. doi:10.1371/journal.pone.0024887.g002

The Angiome

The combined set included 293 angiogenesis annotated and associated proteins. The interactions (Table S1) and proteins (Table S2) included in this network are given as Supporting Information. In Figure 3, we visualized the combined network of angiogenesis-associated proteins (we call it the *Angiome*) by cellular compartment. Proteins were placed into the nucleus (green nodes), cytoplasm (yellow nodes), plasma membrane (orange nodes), or extracellular space (cyan nodes). Proteins that have been identified in multiple compartments were duplicated in the layout. Protein trafficking between compartments was indicated by dashed purple lines. Molecular interactions within a compartment were shown in grey. The interaction density of the figure illustrated the redundancy and complexity of the angiogenesis process.

Validation by transcriptional profiling

Using graph diffusion, we identified 202 proteins that are significantly connected to angiogenesis-annotated proteins. We needed additional validation that these proteins were functionally related to angiogenesis. We expected that many angiogenesis-related genes would have VEGF-regulated gene expression in time during angiogenesis. We analysed three separate time series gene expression studies of angiogenesis involving VEGF stimulated endothelial cells. To mitigate potential cell line bias, we chose studies performed on HUVEC (GDS3567) [15], HMVEC (GDS2039) [16], and TIME [17] cells (see Table 1).

To identify transcripts with high temporal correlation after VEGF stimulation, we ranked transcripts by the absolute value of the covariance between the transcript measurements and the measured time points. For each protein, the temporal correlation of the gene expression profile was given as Supporting Information

(Table S3). We tested the null hypothesis that the distribution of angiogenesis-associated proteins was uniform throughout the ranked list of transcripts. We determined the significance of the enrichment using the Kolmogorov-Smirnov test [18]. A significant p -value indicated that the angiogenesis-associated proteins were disproportionately ranked at either the head or tail of the ranked list of transcripts.

We first tested the set of angiogenesis-annotated genes (GO-annotated) to ensure that they were enriched in the ranked list of perturbed transcripts. As expected, we found that the set of angiogenesis-annotated (i.e. GO-annotated) proteins were significantly perturbed in HUVEC ($p = 0.04$) and TIME ($p = 0.041$) cell lines (see Table 2). We then tested the combined set of 293 proteins including the proteins found using graph diffusion and significant at the 0.01 level. We found that the combined set of proteins was more significantly perturbed than the set of GO-annotated proteins alone across all three endothelial cell lines (shown in Table 2).

Genes with VEGF-regulated expression

We identified seven genes that were consistently regulated by VEGF-A across all three cell lines (shown in Table 3). Only genes with reliable gene expression measurements in all three datasets were considered. Genes such as HIV-1 tat interactive protein 2 (HTATIP2), HIF-1 α , β -amyloid precursor protein (APP), and myocyte enhancer factor 2C (MEF2C) all had increasing gene expression across all three endothelial cell lines. Genes such as endoglin (ENG), liberin β 1 (PPFIBP1), and EPAS1 (HIF-2 α) had decreasing expression after VEGF treatment. We identified literature that supported the association of some of these proteins with angiogenesis. It was previously reported that HTATIP2 (also

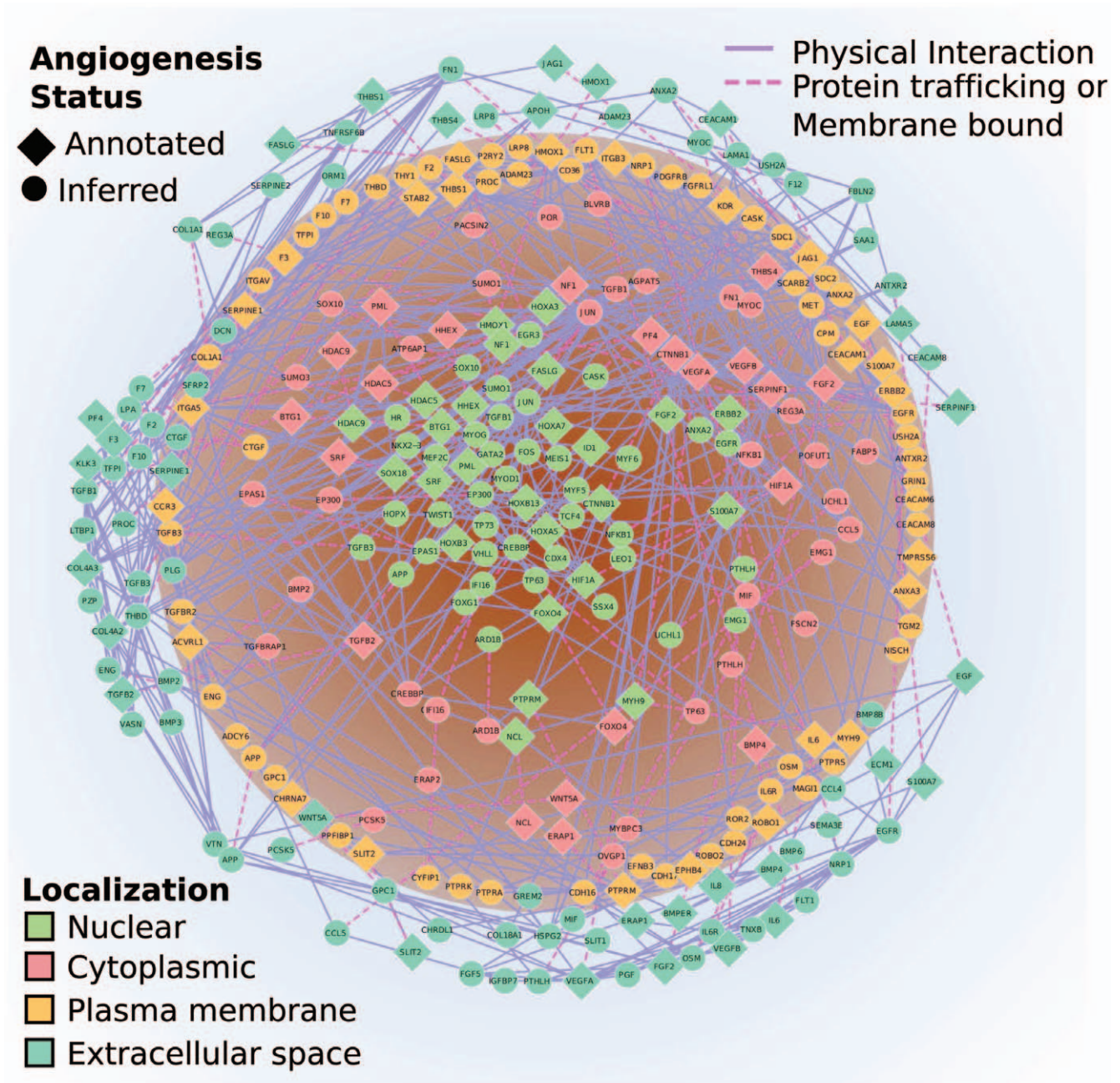


Figure 3. The Angiome. This figure represents all proteins currently annotated (diamond nodes) and topologically associated (circle nodes) with angiogenesis. The figure showed these proteins by their general cellular compartments: nucleus (green), cytoplasm (red), plasma membrane (orange), and extracellular (cyan). Protein trafficking between cellular components was shown as purple dashed lines. Protein-protein interactions within a compartment were shown as blue lines. The figure included the proteins in the single largest component of the network. A complete list of the proteins, interactions, and cellular component assignments could be found in the supplement.
doi:10.1371/journal.pone.0024887.g003

known as CC3) had antiangiogenic properties [19], and VEGF induced MEF2C expression [20]. We found that while endoglin had a known role in angiogenesis [21], the association between liperin β 1 and angiogenesis was not well established. Of these seven genes, only two had non-electronic angiogenesis annotations.

It was known that HIF-1 α was a transcription factor for VEGF [22]. Across three endothelial cell lines, we found a VEGF-regulated increase in HIF-1 α expression over time. It was also established that HIF2A was a hypoxia responsive transcription factor [23]. We found that EPAS1 (also known as HIF-2 α) had

decreasing expression after VEGF-treatment across three endothelial cell lines. The analysis showed differential regulation of HIF-1 α and HIF-2 α .

It was speculated that endothelial cells may have a role in Alzheimer's disease [24,25]. In this proposed mechanism, hypoxic brain cells produced VEGF and stimulated endothelial cells. It was also known that the disease resulted in the destruction of neurons by the accumulation of β -amyloid plaques (i.e. APP protein aggregates). After treatment with VEGF-A, we found increasing expression of APP across all three endothelial cell lines. These data

Table 2. Enrichment of angiogenesis-associated proteins in a ranked list of the most perturbed gene expression transcripts during angiogenesis.

Gene Set	HUVEC	HMVEC	TIME
GO angiogenesis annotated only	0.04	0.185	0.041
angiogenesis annotated and associated at $p < 0.01$	0.008	0.001	0.001
angiogenesis associated at $p < 0.01$ only	0.08	0.001	0.001
angiogenesis annotated and associated at $p < 0.05$	0.005	0.02	0.001
angiogenesis associated at $p < 0.05$ only	0.035	0.026	0.008

To evaluate the relevance of angiogenesis-associated proteins with angiogenesis, we evaluated the significance of gene sets in the ranked list of most perturbed gene expression profiles. We obtained the gene expression profiles from three studies performed on different endothelial cell lines. The table gave the p -value of the enrichment of the gene set at the head of the ranked list of perturbed gene transcript.
doi:10.1371/journal.pone.0024887.t002

supported the relation between endothelial cells and Alzheimer's disease.

Pathways linked to angiogenesis

Using the expanded network shown in Figure 3, we identified overrepresented pathways linked to angiogenesis (see Table 4). The pathways we report in Table 4 could not be identified using angiogenesis-annotated proteins alone. A pathway associated with acute myocardial infarction (AMI) was identified as angiogenesis-associated with the expansion of the network to type IV collagens (i.e. COL4A3 and COL4A2) and coagulation factors (e.g. F10, F2, and F7). It was known that angiogenesis was a natural mechanism to restore perfusion after AMI [26]. The coagulation factors were also associated with prothrombin activation. Fragments of prothrombin had been identified as regulators of angiogenesis [27]. Collagen IV proteins also contributed to the platelet β -amyloid precursor protein pathway (APP). While the molecular basis was unclear, angiogenesis was identified as part of a mechanism for Alzheimer's disease along with pathological secretion of β -amyloid proteins [24]. In this study, the expansion of the angiogenesis-associated network gave us increased statistical power to link pathways with angiogenesis.

Conclusions

Compounds that inhibit angiogenesis represent potential therapeutics for many diseases. Therapies that target the VEGF pathway have been used in the clinic. These treatments suffer from an "escape phenomenon" where alternative pathways are upregulated and angiogenesis continues. The redundancy of angiogenesis regulation indicates the need for additional studies and new drug targets.

To discover novel and missing angiogenesis annotations, we integrated the human interactome with known angiogenesis-annotated proteins to identify a set of 293 angiogenesis-related proteins. We verified these protein associations by repeated network randomization and analysis of relevant time series gene expression datasets. After treatment with VEGF-A, we found increasing expression of HIF-1 α , APP, HIV-1 tat interactive protein 2, and MEF2C, while endoglin, liprin β 1 and HIF-2 α had decreasing expression across three endothelial cell lines. The analysis showed differential regulation of HIF-1 α and HIF-2 α . The data also provided additional evidence for the role of endothelial cells in Alzheimer's disease.

Materials and Methods

The interaction dataset was taken from the Michigan Molecular Interaction database (MiMI) [28] (Feb 2009 version). The dataset was composed of 13,491 genes, proteins, and RNA connected by 126,763 physical interactions. The interactions included protein-protein, protein-DNA, protein-RNA, and RNA-RNA. As a result, the dataset captured diverse aspects of biomolecular interactions including protein complexation, transcriptional regulation, and RNA interference. The dataset consisted of interactions curated from reputable online databases such as Reactome [29], BIND, BioGrid [30], HPRD [31]. This network of physical interactions formed the basis for pathway expansion. Gene Ontology (GO) [32] annotations were used to determine the known set of angiogenesis-associated proteins (6/2010 version). MsigDB C2 was used for pathway annotations [33]. Preparation and computational preprocessing of the gene expression measurements was described in their manuscript.

Graph diffusion

We used graph diffusion to establish topological associations. Intuitively, the algorithm worked by initiating a series of fixed length random walks through the network originating at nodes of interest. The score of each node was given by the fraction of

Table 3. Summary of genes temporally regulated by VEGF across cell lines.

Symbol	Name	Annotated	HMVEC	HUVEC	TIME
HTATIP2	HIV-1 Tat interactive protein 2	yes	0.72	0.71	0.86
HIF1A	hypoxia inducible factor 1	yes	0.59	0.85	0.89
APP	amyloid beta (A4) precursor protein	no	0.85	0.81	0.85
MEF2C	myocyte enhancer factor 2C	no	0.99	0.82	0.63
ENG	endoglin	no	-0.60	-0.62	-0.74
PPFIBP1	liprin beta 1	no	-0.61	-0.89	-0.79
EPAS1	endothelial PAS domain protein 1	no	-0.81	-0.93	-0.53

We highlighted those proteins from Figure 3 that had consistently correlated gene expression in time after VEGF stimulation. HTATIP2, HIF-1 α , APP, and MEF2C had positively regulated gene expression, while ENG, PPFIBP1, and EPAS1 had negatively regulated gene expression. Correlation values are shown for HUVEC HMVEC and TIME endothelial cells. Only those genes with reliable measurements in all three datasets were considered.
doi:10.1371/journal.pone.0024887.t003

Table 4. Pathways significantly represented in the Angiome.

Pathway	Description	p-value
AMI	Acute myocardial infarction-associated pathway	2.00E-06
Extrinsic	Extrinsic prothrombin activation pathway	2.02E-04
Platelet APP	Platelet amyloid precursor protein pathway	3.50E-04
TGFB	TGF beta signaling pathway	3.00E-03
Nkt	Expression of chemokine receptors during T-cell polarization	3.80E-02
Keratinocyte	Keratinocyte differentiation	5.39E-02
Cardiac EGF	EGF receptor transactivation by GPCRs in cardiac hypertrophy	7.92E-02

These pathways could not be statistically identified from angiogenesis-annotated proteins alone. With the expansion of the angiogenesis-related network, these pathways were easily detected. The p -value was computed using Fisher's exact test followed by correction for the false discovery rate. doi:10.1371/journal.pone.0024887.t004

random walks that pass through the node. A parameter γ controlled the length of the random walks. The aggregation of evidence over multiple paths led to a more stable result from potentially unreliable data. The steady-state solution of this system associated each node with an association score. In this study, we set γ such that all nodes receive some non-zero association score. We made this functionality available through our webserver (sysbio.bme.jhu.edu).

For a weighted undirected graph $G(V,E)$ with vertex set V and edge set E , let \mathbf{A} be the symmetric adjacency matrix representing G . Let q_i be 1 if node i was in the query set or zero otherwise. Let $i \sim j$ be the set of nodes connected to node i . We expressed the time derivative \dot{s}_i of the diffusion kernel score $s_i(q)$ for node $i \in V$ as

$$\dot{s}_i = \sum_{i \sim j} \mathbf{A}_{ji} s_j - s_i \sum_{i \sim j} \mathbf{A}_{ij} - \gamma s_i + q_i \quad (1)$$

Let \mathbf{D} be the degree weighted diagonal matrix of \mathbf{A} . In matrix notation, equation (1) had the form in equation (2).

$$\dot{\mathbf{s}} = \mathbf{A}\mathbf{s} - \mathbf{D}\mathbf{s} - \gamma\mathbf{I}\mathbf{s} + \mathbf{q} \quad (2)$$

Our goal was to identify the values of \mathbf{s} at steady-state. We set $\dot{\mathbf{s}} = \mathbf{0}^T$ from equation (2) and solve for \mathbf{s} in equation (3).

$$\mathbf{s} = [\mathbf{D} - \mathbf{A} + \gamma\mathbf{I}]^{-1} \mathbf{q} \quad (3)$$

The values in \mathbf{s} gave the association of each node with the query nodes defined by \mathbf{q} .

Statistical significance by network randomization

We computed the statistical significance of protein associations by permutation testing. We tested the null hypothesis that the association score of a protein is equal to the score of the protein in randomized networks. The alternative hypothesis was that the score of a protein is greater than the score of the protein in randomized networks. To test these hypotheses, we generated 300 randomly edge swapped networks. The probability of the null hypothesis was given by the number of randomized networks where the score of a protein exceeded the score of the protein in the real network. Pseudocounts were added to avoid fitted probabilities of zero. We did not evaluate the statistical

significance of the angiogenesis-annotated proteins. Angiogenesis-annotated proteins were selected for the study and as such they are inherently biased.

Time series gene expression analysis

We reanalyzed three time series gene expression datasets to validate the angiogenesis-associated proteins and mitigate potential cell line bias. In Table 1, we presented the three time series datasets used in this study. These studies focused on early transcriptional changes during VEGF induced angiogenesis. The studies used different endothelial cell lines and measured different time points. Mellberg *et al.* [17] measured gene expression of VEGF stimulated telomerase-immortalized human microvascular endothelial (TIME) cells at 15 min, 1, 3, 6, 9, 12, 18, and 24 hours. Schweighofer *et al.* [15] measured gene expression in VEGF simulated HUVEC at 30 min, 1, 2.5, and 6 hours (GDS3567). Glesne *et al.* [16] measured gene expression in VEGF stimulated HMVEC at 30 min, 1, 2, 4, and 8 hours of tubulogenesis (GDS2039).

To identify temporal correlations, we computed the Pearson's correlation coefficient and covariance between each gene expression profile and the time course. Genes with fewer than 3 reliable gene expression measurements are excluded. We formed a ranked list of perturbed genes using the absolute value of the covariance. We tested the null hypothesis that angiogenesis-annotated genes are evenly distributed throughout the ranked list of genes. The alternative hypothesis was that the proteins were disproportionately ranked at either the head or tail of the list. We used the gene set enrichment analysis (GSEA) software to compute the p -value [34]. Pathway enrichment was computed using Fisher's exact test followed by correction for the false discovery rate (i.e. Benjamini-Hochberg [35]).

Supporting Information

Table S1 Molecular interactions and types included in Figure 3. The file listed the interactions included in Figure 3. Proteins trafficking between components and transmembrane proteins may be represented in multiple compartments. (XLS)

Table S2 Graph diffusion scores and associated p-values. The data included the individual genes included in Figure 3. The file indicated if the protein was previously annotated with angiogenesis, the graph diffusion score, and the permutation-based p -value. Pseudocounts may have been added to avoid fitted probabilities of zero. (XLS)

Table S3 VEGF-induced gene expression correlation measurements. For each protein, we list the temporal correlation of the gene expression profile during VEGF-induced angiogenesis. A positive (negative) correlation value indicated a trend towards increasing (decreasing) gene expression after VEGF treatment. Results were shown for the three VEGF-A induced gene expression datasets. (XLS)

References

- Folkman J (1971) Tumor angiogenesis: therapeutic implications. *N Engl J Med* 285: 1182–1186.
- Folkman J (2007) Angiogenesis: an organizing principle for drug discovery? *Nat Rev Drug Discov* 6: 273–286.
- Garber K (2002) Angiogenesis inhibitors suffer new setback. *Nature biotechnology* 20: 1067–1068.
- Rehn M, Veikkola T, Kukk-Valdre E, Nakamura H, Ilmonen M, et al. (2001) Interaction of endostatin with integrins implicated in angiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 98: 1024.
- Maisonpierre PC, Suri C, Jones PF, Bartunkova S, Wiegand SJ, et al. (1997) Angiopoietin-2, a natural antagonist for Tie2 that disrupts in vivo angiogenesis. *Science* 277: 55.
- Murali TM, Wu CJ, Kasif S (2006) The art of gene function prediction. *Nat Biotechnol* : - 24: 1474–1475; author reply 1475–1476.
- Pena-Castillo L, Tasan M, Myers CL, Lee H, Joshi T, et al. (2008) A critical assessment of Mus musculus gene function prediction using integrated genomic evidence. *Genome Biol* 9(Suppl 1): S2.
- Zhou X, Kao MC, Wong WH (2002) Transitive functional annotation by shortest-path analysis of gene expression data. *Proc Natl Acad Sci U S A* 99: 12783–12788.
- Lanckriet GR, De Bie T, Cristianini N, Jordan MI, Noble WS (2004) A statistical framework for genomic data fusion. *Bioinformatics* 20: 2626–2635.
- Troyanskaya OG, Dolinski K, Owen AB, Altman RB, Botstein D (2003) A Bayesian framework for combining heterogeneous data sources for gene function prediction (in *Saccharomyces cerevisiae*). *Proceedings of the National Academy of Sciences* 100: 8348.
- Huang Y, Li S (2010) Detection of characteristic sub pathway network for angiogenesis based on the comprehensive pathway network. *BMC Bioinformatics* 11: S32.
- Gu J, Chen Y, Li S, Li Y (2010) Identification of responsive gene modules by network-based gene clustering and extending: application to inflammation and angiogenesis. *BMC systems biology* 4: 47.
- Chen Y, Wei T, Yan L, Lawrence F, Qian HR, et al. (2008) Developing and applying a gene functional association network for anti-angiogenic kinase inhibitor activity assessment in an angiogenesis co-culture model. *BMC Genomics* 9: 264.
- Rivera CG, Bader JS, Popel AS (2011) Angiogenesis-associated crosstalk between collagens, CXC chemokines, and thrombospondin domain-containing proteins. *Ann Biomed Eng* 39: 2213–2222.
- Schweighofer B, Testori J, Sturtzel C, Sattler S, Mayer H, et al. (2009) The VEGF-induced transcriptional response comprises gene clusters at the crossroad of angiogenesis and inflammation. *Thromb Haemost* 102: 544–554.
- Glesne DA, Zhang W, Mandava S, Ursos L, Buell ME, et al. (2006) Subtractive transcriptomics: establishing polarity drives in vitro human endothelial morphogenesis. *Cancer Res* 66: 4030–4040.
- Mellberg S, Dimberg A, Bahram F, Hayashi M, Rennel E, et al. (2009) Transcriptional profiling reveals a critical role for tyrosine phosphatase VE-PTP in regulation of VEGFR2 activity and endothelial cell morphogenesis. *FASEB J* 23: 1490–1502.

Acknowledgments

The authors would like to thank Emmanouil D. Karagiannis for helpful discussions at the initial stage of the project. We also thank Amina A. Qutub for advice regarding HIF-1 α .

Author Contributions

Conceived and designed the experiments: CGR JSB ASP. Performed the experiments: CGR SM LC-W. Analyzed the data: CGR. Contributed reagents/materials/analysis tools: CGR SM LC-W. Wrote the paper: CGR SM LC-W JSB ASP.

- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102: 15545–15550.
- NicAmhlaioibh R, Shtivelman E (2001) Metastasis suppressor CC3 inhibits angiogenic properties of tumor cells in vitro. *Oncogene* 20: 270.
- Maiti D, Xu Z, Duh EJ (2008) Vascular endothelial growth factor induces MEF2C and MEF2-dependent activity in endothelial cells. *Investigative ophthalmology & visual science* 49: 3640.
- ten Dijke P, Goumans MJ, Pardali E (2008) Endoglin in angiogenesis and vascular diseases. *Angiogenesis* 11: 79–89.
- Talks KL, Turley H, Gatter KC, Maxwell PH, Pugh CW, et al. (2000) The expression and distribution of the hypoxia-inducible factors HIF-1 {alpha} and HIF-2 {alpha} in normal human tissues, cancers, and tumor-associated macrophages. *American Journal of Pathology* 157: 411.
- Koukourakis MI, Giatromanolaki A, Skarlatos J, Corti L, Blandamura S, et al. (2001) Hypoxia inducible factor (HIF-1a and HIF-2a) expression in early esophageal cancer and response to photodynamic therapy and radiotherapy. *Cancer research* 61: 1830.
- Vagnucci AH (2003) Alzheimer's disease and angiogenesis. *The Lancet* 361: 605–608.
- Grammas P, Moore P, Weigel PH (1999) Microvessels from Alzheimer's disease brains kill neurons in vitro. *Am J Pathol* 154: 337–342.
- Lee SH, Wolf PL, Escudero R, Deutsch R, Jamieson SW, et al. (2000) Early expression of angiogenesis factors in acute myocardial ischemia and infarction. *New England Journal of Medicine* 342: 626–633.
- Browder T, Folkman J, Pirie-Shepherd S (2000) The hemostatic system as a regulator of angiogenesis. *Journal of Biological Chemistry* 275: 1521.
- Jayapandian M, Chapman A, Tarcea VG, Yu C, Elkiss A, et al. (2007) Michigan Molecular Interactions (MiMI): putting the jigsaw puzzle together. *Nucleic Acids Res* 35: D566–571.
- Vastrik I, D'Eustachio P, Schmidt E, Gopinath G, Croft D, et al. (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biol* 8: R39.
- Breitkreutz BJ, Stark C, Reguly T, Boucher L, Breitkreutz A, et al. (2008) The BioGRID Interaction Database: 2008 update. *Nucleic Acids Res* 36: D637–640.
- Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, et al. (2009) Human Protein Reference Database–2009 update. *Nucleic Acids Res* 37: D767–772.
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, et al. (2000) Gene ontology: tool for the unification of biology. *The Gene Ontology Consortium. Nat Genet* 25: 25–29.
- Liberzon A, Subramanian A, Pinchback R, Thorvaldsdottir H, Tamayo P, et al. (2011) Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27: 1739–1740.
- Subramanian A, Kuehn H, Gould J, Tamayo P, Mesirov JP (2007) GSEA-P: a desktop application for Gene Set Enrichment Analysis. *Bioinformatics* 23: 3251–3253.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. pp 289–300.