Evolutionary Analysis of Human Vascular Endothelial Growth Factor, Angiopoietin, and Tyrosine Endothelial Kinase Involved in Angiogenesis and Immunity

Anton Dormer and Gregory Beck*

Department of Biology, University of Massachusetts at Boston, Boston, MA 02125-3393, USA

Edited by H. Michael; received 7 February 2005; revised and accepted 24 March 2005; published 30 April 2005

ABSTRACT: Human vascular endothelial growth factor (VEGF), angiopoietin (ANG) and tyrosine kinase with immunoglobulin and epidermal growth factor homology domains (TIE)-2 consist of a grouping of proteins that are involved in vascular homeostasis, vascular integrity and angiogenesis. There are nine proteins in the immediate VEGF family: VEGFA, VEGFB, VEGFC, VEGFD, VEGF-3, placental growth factor (PGF), VEGF receptor (VEGFR)-1, VEGFR-2, and VEGFR-1-related. They can be stimulated by cytokines to become involved in immune responses. By using in silico tools, we were able to identify several possible analogues or homologues of VEGF, ANG and TIE-2 in invertebrates. This is the first report to show that these proteins may be conserved through evolution. These proteins may have a role in vascular maintenance and immunity. In addition, since VEGF, ANG and TIE-2 have a role in mammalian immunity that is significantly influenced by cytokines, such as IL-1, this may indicate an interaction of the vascular system and the immune system over evolutionary time.

KEYWORDS: Immunology, angiogenesis, evolution, cytokine, protein sequence, human vascular endothelial growth factor angiopoietin, tyrosine endothelial kinase, vascular homeostasis, vascular integrity, invertebrate, endothelial growth factor

Abbreviations: VEGF: vascular endothelial growth factor; ANG: angiopoietin; TIE: tyrosine kinase with immunoglobulin and epidermal growth factor homology domains; VEGFR: VEGF receptor; FLT: fms-related tyrosine kinase; FN3: fibronectin type 3 domain; ANGPT: angiopoietin-like 1; TEK: tyrosine endothelial kinase.

INTRODUCTION

Vascular Endothelial Growth Factor (VEGF)

The VEGF family of proteins consists of nine different proteins. Vascular endothelial growth factor A (VEGFA) was first identified in 1989 as a vascular permeability factor that stimulated endothelial cell growth and was related to platelet derived growth factor (PDGF) [Keck et al., 1989; Ottaviani et al., 2001]. Other similar proteins were then quickly discovered. The protein VEGFB, was cloned and characterized
A. Dormer and G. Beck / Evolutionary Analysis of Human Vascular Endothelial Growth Factor Angiopoietin

In 1996, and found to behave similarly to VEGFA. It is active in angiogenesis and endothelial cell growth and can affect blood vessel permeability. This protein is structurally and functionally similar to VEGFD (VEGFC NCBI Locus Link). Vascular endothelial growth factor D has an additional function apart from angiogenesis, and endothelial cell growth as it is involved in lymph angiogenesis [Yokoyama et al., 2003]. The VEGF proteins exert multiple functions by binding to two receptors, VEGFR-1 and VEGFR-2. The angiopoietin (ANG) family is a group of proteins that are associated with VEGF that modulates and refines its activity.

**Angiopoietin (ANG)**

There are five ANG proteins: ANG-1, -2, -3, -4, and -5. Angiopoietin-1 is expressed by periendothelial cells [Satchell et al., 2001], stabilizes tumor neovasculature, and can inhibit tumor growth [Tian et al., 2002]. It also regulates angiogenesis, vascular permeability, and effects the expression of VEGFR-2 through the tyrosine kinase with immunoglobulin and epidermal growth factor homology domains (TIE)-2 receptor (TIE-2R; Stoeltzing et al., 2002; Christensen et al., 2002]. Loss of normal expression of ANG-1 can lead to disease abnormalities, such as, menorrhagia, and tumor angiogenesis promotion [Hewett et al., 2002; Shim et al., 2002].

Angiopoietin-2 is expressed in areas undergoing vascular remodeling and is involved in neovascularization. Pathological conditions change the expression of ANG-2 and that leads to an imbalance in the ratio between ANG-1 and ANG-2 [Paradis et al., 2003]. Angiopoietin-3 and -4 are newly discovered ANG proteins. Angiopoietin-3 seems to be expressed in the lung and cultured human umbilical vein endothelial cells. Angiopoietin-3 mRNA expression is different from that of ANG-2 indicating a possible regulatory pathway [Nishimura et al., 1999]. Angiopoietin-3 and -4 appear to be members of the angiopoietin family because they bind to the TIE-2R and not the TIE-1R. Angiopoietin-1 and -2 bind to the TIE-1R, but not the TIE-2R. Evidence shows that ANG-4 is an agonist for the TIE-2R (as is ANG-1), whereas ANG-2 and ANG-3 are antagonists [Valenzuela et al., 1999]. The overexpression of VEGF appears to promote growth of new vessels accompanied by plasma leakage, whereas overexpressing ANG-1 promotes the enlargement of existing vessels and a resistance to leakage [Thurston, 2002]. The other TIE family member the TIE-1R, seems to be important in endothelial vascular integrity [Sato et al., 1995].

The TIE-1R is a tyrosine kinase protein with Ig and epidermal growth factor homology domains. The VEGF protein can modulate the TIE-1R via its complex with the TIE-2R. It can therefore switch the TIE-1R:TIE-2R complex between two different forms in endothelial cells. This complex provides a mechanism whereby this initiator of vessel growth and remodeling can directly modulate receptors involved in vessel stabilization [Tsiamis et al., 2002]. In addition, these factors interact with interleukins, which suggest they also may have a role in mediating immunity.

**VEGF family members and interactions with interleukins**

Interleukin (IL)-1α and β are pro-inflammatory cytokines. Interleukin-1α stimulates VEGF secretion by human peripheral blood mononuclear cells in a dose-dependent manner. This is accomplished by its induction of angiogenesis by activating the VEGF-VEGFR-2 signaling pathway between inflammatory cells and the endothelial cells of blood vessels [Salven et al., 2002]. The IL-1β protein, along with VEGF, acting as endothelial cell mitogens and permeability factors in ovarian angiogenesis, has been the subject of increasing interest. Interleukin-1 may regulate ovarian VEGF activity [Levitas et al., 2000].

A similar association with tumor pathology is associated with the cytokine IL-6. Interleukin-6 increases angiogenic activity in human cervical cancer cells and this effect is specifically associated with up
regulation of VEGF. In regards to immunity, IL-6 can cause an increase of VEGF in viral retinopathy induced inflammation [Vinores et al., 2001]. Biologically active VEGF is stored in platelets. Platelets seem to prevent circulating VEGF from inducing the development of new blood vessels except at sites where coagulation takes place. The IL-6 protein may affect the amount of VEGF stored in the platelets. The interaction of IL-6 with angiogenic pathways in cancer could explain tumor growth associated with IL-6 administration. Tumor thrombocytosis and benign angiogenic diseases are associated with high IL-6 levels and poor patient outcome [Salgado et al., 1999].

Other interleukins interact with VEGF-family members as well. Interleukin-8 serum levels are increased in association with pathologies, such as, tissue injuries, ischemia, and tumor growth. Interleukin-8 promotes angiogenesis. The release of IL-8 may be initiated by the release of adenosine under mentioned pathological conditions. Adenosine also causes the release of IL-8 that leads to the release of VEGF and ANG-2 [Feoktistov et al., 2003]. Tumor necrosis factor (TNF) also plays a role in influencing the activity of ANG proteins [Kim et al., 2002; Scott et al., 2002]. The expression of IL-10 is correlated with clinical prognosis in lung cancer [Hatanaka et al., 2001]. In general, tumor-produced IL-10 promotes stromal vascularization through expression of ANG-1, ANG-2, and the TIE-2R [Hatanaka et al., 2001]. High levels of interferon gamma (IFNγ) released by IL-12-activated lymphocytes up-modulates ANG-2 genes, but down-modulates the expression of VEGF. These opposite effects of VEGF and ANG-2 expression fit in well with the inhibition of angiogenesis that is characterized by the anti-tumor activity of IL-12 [Cavallo et al., 2001].

The importance of the VEGF family of proteins in angiogenesis and immunity suggested to us the need to examine their possible evolutionary development. For example, VEGF has been found in Drosophila, yet to date, the associated proteins of ANG have yet to be found in any invertebrate [Cho et al., 2003]. Ottaviani [Ottaviani et al., 2001] has described PDGF molecules in invertebrates. We feel that we have identified several proteins that are possible analogs or homologs for VEGF, ANG and TIE. Our data strongly suggest an early evolutionary appearance of the VEGF family of proteins. In addition, it suggests that there could have been a concurrent development of the immune and vascular system over evolutionary time.

METHODS

BLAST

The National Center of Biotechnology information basic local alignment search tool (NCBI BLAST) was used to search the NCBI databases. In looking for VEGF family members, we used the BLAST tool that locates local alignments and therefore is able to detect relationships among sequences that share only isolated regions of similarity [Altschul et al., 1990; Altschul and Lipman, 1990]. Initially, all human VEGF family member sequences were used for a BLAST search in the NCBI human (h), mouse (m), rat (r), pufferfish (p; Fugu rubripes), mosquito (mq; Anopheles gambiae), fruit fly (d; Drosophila melanogaster), zebrafish (z; Danio rerio), and nematode (ce; Caenorhabditis elegans) databases. In our in silico experiments we felt that initial E-values from our BLAST results did not signify structural conservation over evolutionary time. Initial E-values may not have indicated that a recovered sequence was similar to our query. Therefore, each hVEGF-like family member sequence that was found in the initial BLAST analysis of the invertebrate and vertebrate genome databases was then used in a BLAST analysis back into the NCBI human genome database. This was done to make sure that the query sequence and the subject sequences were similar based on E-value. Here, we used protein sequence similarity in our search for invertebrate correlates of VEGF family members.
**A. Dormer and G. Beck / Evolutionary Analysis of Human Vascular Endothelial Growth Factor, Angiopoietin**

We utilized the European Bioinformatics Institute ClustalW program (www.ebi.ac.uk/clustalw). The ClustalW program was used to look for biologically meaningful sequence alignments of evolutionary distant sequences. This program was used to confirm whether the E-values obtained from our BLAST results would show significantly conserved regions. We used the default ClustalW parameters with an expect of 0.01; filter at default, descriptions at 100, and alignments at 100.

**Phylogenetic tree**

For phylogenetic tree analysis we employed the Gene Bee Phylogenetic Tree Programs (www.genebee.msu.su). The default mode was used with PHYLIP and unrooted parameters. No correction or manipulations were made in our tree generation [Chumakov et al., 1988; Ishmanov et al., 1988; Brodsky et al., 1992; Brodsky et al., 1995].

We estimated divergence time of the VEGF family of proteins using the molecular clock method. We performed a pairwise per-site divergence calculation with default parameter using a Dayhoff PAM substitution model [Dayhoff et al., 1978] to correct for multiple hits [Cutler, 2000a; Cutler, 2000b]. This was accomplished using the PHYLIP PROTDIST program to compute distance matrix from protein sequences as developed by Felsenstein [Felsenstein, 1989; Felsenstein, 1993].

**RESULTS**

To determine the evolutionary relationship of VEGF, ANG, and TIE molecules, all members within this family of proteins were examined (Table 1).

To examine the evolutionary relationship of the VEGF family an unrooted phylogenetic tree was first developed using Gene Bee (Fig. 1). These results show that the VEGF cluster of proteins is more divergent than the ANG cluster. This assumption is due to the bootstrap values that vary among the VEGF cluster (Fig. 1(a)). The VEGFR-1-related protein is a receptor that clustered with VEGFA. The VEGFR-1 and -2 sequences also clustered together. It was expected that the VEGF and ANG families would cluster together. What was surprising was the evolutionary distance between the ANG proteins in comparison to VEGF and VEGFR proteins (Fig. 1(a), (b)). Figure 1(b) shows the evolutionary distance within the hVEGF family. As shown, the distance is greatest between ANG-1 and VEGFA.

The next step was to choose representative proteins of the VEGF family. We chose to use, as the initial inquiry, VEGFA, VEGFR-1, ANG-1 and the TIE-2R. Although we identified conserved regions by BLAST analysis, we used the entire molecule as an inquiry tool in our subsequent BLAST searches (Table 2). The VEGFA molecule is 215 aa with a conserved region from aa 50 to aa 133. This region is similar to PDGF. In addition, VEGF forms homodimers and exists in four different isoforms. Overall, VEGF resembles PDGF, but its N-terminal is helical rather than extended; the cysteine knot motif is a common feature of this domain. Even though there is a conserved region in VEGF we decided to use the entire molecule in a BLAST analysis to eliminate other proteins that may have the PDGF domain, but not other unique features of the VEGFA molecule. Similarly, since VEGFR-1 has five-conserved Ig domains we decided to use the entire protein in a BLAST analysis instead of just a single conserved domain. (i.e., it has two-tyrosine kinase domain). We did observe however that BLAST analysis using one or both of the domains retrieved results specific for the domains and not proteins structurally related to VEGFR-1 (data not shown).
The ANG-1 protein, like VEGF, consists of only one major conserved domain. Its fibrinogen-related domains (FBG) consist of C-termini $\beta$ and $\gamma$ chains. This portion of ANG-1 is very similar to other fibrinogen proteins found in invertebrates, such as tenacin and Drosophila scabrous (ANG-1 isoform NCBI conserved domain search). The FBG in ANG-1 is approximately 215 aa (aa 281 to aa 496) and takes up approximately half the ANG-1 molecule. As with VEGFR-1, the ANG-1 BLAST analysis returned information on just its conserved domains and not any information on proteins related to ANG-1 (data not shown).

The primary receptor for ANG-1 is TIE-2. The TIE-2R is approximately 1,124 aa in length, has multiple domains like VEGFR, and has a calcium-binding epidermal growth factor (EGF)-like domain (present in a large number of membrane-bound and extracellular, mostly animal, proteins). There is one Ig and three FBG domains. Approximately 2% of all animal proteins contain FBG (e.g., membrane spanning cytokine receptors, growth hormone receptors, tyrosine phosphatase receptors, and adhesion molecules). Finally, the most important region of the protein is the tyrosine kinase catalytic domain. This domain plays an important role in the functional ability of the TIE-2R. The TIE-2R and VEGFR molecules are in the same protein family and relay their ligand signal with the use of their tyrosine kinase domain (TIE-2 NCBI Locus Link).

The BLAST analysis of VEGFA, VEGFR-1, ANG-1 and the TIE-2R led to the identification of protein homologs/orthologs in both vertebrates and invertebrates. The initial results were then used for BLAST analysis back to confirm that the sequences were related (data not shown). We used BLAST analysis of the VEGF family members in a non-vascular organism (i.e., yeast) and as expected the results indicated that there was no similarity to any yeast protein identified so far (data not shown).

The ClustalW results of the respective grouping of VEGF, VEGFR, ANG, and TIE family members show conserved protein sequence regions that may imply functional similarity (Fig. 2). The VEGF family has been, until now, only identified in human, mouse, rat and zebrafish. In the organisms we
Fig. 1. (a) Analysis of the hVEGF protein family. A cluster algorithm was used to create this phylogenetic tree with bootstrap values of the hVEGF protein family. (b) Analysis of the hVEGF protein family. Standard unrooted phylogenetic tree showing the evolutionary relationship of the hVEGF family of proteins.
Table 2

The identification of proteins of the VEGF family

<table>
<thead>
<tr>
<th>Organism*</th>
<th>VEGF family member</th>
<th>Mouse</th>
<th>Rat</th>
<th>Pufferfish</th>
<th>Zebrafish</th>
<th>Mosquito</th>
<th>Fly</th>
<th>Nematode</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td></td>
<td>VEGFA</td>
<td>VEGFA</td>
<td>00000051063</td>
<td>VEGF</td>
<td>00000017314</td>
<td>CG7103-PA</td>
<td>No similarity found</td>
</tr>
<tr>
<td></td>
<td>VEGFA</td>
<td>3e-78</td>
<td>e-118</td>
<td>4e-53</td>
<td>6e-61</td>
<td>1e-09</td>
<td>3e-09</td>
<td>No similarity found</td>
</tr>
<tr>
<td></td>
<td>VEGFR-1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>2e-63</td>
</tr>
<tr>
<td></td>
<td>ANG-1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>9e-16</td>
<td>9e-16</td>
</tr>
<tr>
<td></td>
<td>TIE-2</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1e-68</td>
<td>5e-63</td>
<td>1e-63</td>
</tr>
</tbody>
</table>

*Throughout the manuscript lower vertebrate and invertebrate sequences that have only been characterized by database searches are identified with an abbreviation suffix. Therefore, the Anopheles sequence that we identified as VEGFA-like is annotated as mqVEGFA-like and the Drosophila sequence that we identified as ANG-1-like is annotated as dANG-1-like. Similarly, mouse (m), rat (r), pufferfish (p; Fugu rubripes), mosquito (mq; Anopheles gambiae), fruit fly (d; Drosophila melanogaster), zebrafish (z; Danio rerio), and nematode (ce; Caenorhabditis elegans) are also so annotated.

investigated, the BLAST results identified structurally similar proteins we annotated as “like” proteins (we consider this to be the case since their biological functions have not yet been characterized). These proteins were identified and included in our ClustalW analysis.

The ClustalW analysis of VEGFA proteins (Fig. 2(a)) consisted of hVEGFA, mVEGFA, rVEGFA, zVEGFA, a pufferfish molecule (00000051063, pVEGFA-like), a mosquito molecule (00000017314, mqVEGFA-like), and a Drosophila molecule (CG31629-PA, dVEGFA-like). For VEGFR-1, the ClustalW (Fig. 2(b)) consisted of hVEGFR-1, mVEGFR-1, rVEGFR-1, zVEGFR-1, a pufferfish molecule (00000072413930, pVEGFR-1-like), a mosquito molecule (0000003182, mqVEGFR-1-like), a Drosophila molecule (CG10244, dVEGFR-1-like), and a C. elegans molecule (VEGFR-related 4; ceVEGFR-1-like).

The ClustalW analysis of the ANG-1 family (Fig. 2(c)) consisted of hANG-1, mANG-1, rNAG-1, zANG-1, a pufferfish molecule (00000069285620178, pANG-1-like), a mosquito molecule (0000003182, mqANG-1-like), a Drosophila molecule (CG30281-PA, dANG-1-like), and a C. elegans molecule (VEGFR-related 4; ceVEGFR-r4-like).

Finally, the last ClustalW analysis was for the TIE proteins (Fig. 2(d)). It consisted of hTIE-2, mTIE-2, rTIE-2, zTIE-2, a pufferfish molecule (00000061737609174, pTIE-1R-like), a mosquito molecule (0000003182, mqTIE-2R-like), a Drosophila molecule (CG10244-PA, dTIE-2R-like), and a C. elegans TIE-2R-like molecule (myoblast growth factor receptor – egg laying defective genes (egl), ceMGFR-like) (Fig. 2(d)).
A. Dormer and G. Beck / Evolutionary Analysis of Human Vascular Endothelial Growth Factor Angiopoietin

Fig. 2a. ClustalW analysis of VEGF proteins. The conserved region of hVEGFA is from amino acid 50 to 133. The PDGF/VEGF domain is a cell activator of mesenchymal origin. The VEGFA PDGF/VEGF domain differs from the PDGF PDGF/VEGF domain in that its N-terminal segment is helical rather than extended (VEGFA NCBI Conserved Domain Search). Our data shows that there is > 50% homology between the conserved region of the human, mouse, rat, zebrafish, fugu, mosquito, and Drosophila VEGFA and VEGFA-like molecules as indicated by consensus symbols. The structural homology among the sequences should imply functional homology. Residue coloring is according to standard physiochemical criteria (red: small hydrophobic; blue: acidic; magenta: basic; green: hydroxyl, amine and basic). Alignments display typical symbols denoting the degree of conservation observed in each column (* denotes identical nucleotides, : denotes conserved substitutions, and . denotes semi-conserved substitutions in the alignment).

DISCUSSION

This study presents an analysis of various angiogenesis-associated proteins. We applied a systemic approach in assessing the evolutionary relationship of proteins that may have been conserved in many species. The goal of this study was not to make a distinction among the protein we were looking for, but create a direction for the investigation of the VEGF family from a phylogenetic point of view. We believe that we have successfully accomplished that goal based on the overall similarity of the identified VEGF-like family members. Our findings are based on BLAST, ClustalW and phylogenetic tree analysis.
Fig. 2b. ClustalW analysis of VEGFR-1 proteins. The VEGFR-1 receptor has many conserved domains in vertebrates. The hVEGFR-1 protein contains two tyrosine kinase catalytic (TyrKc) domains, three immunoglobulin (Ig) cell adhesion molecule (cam) subfamily (Igcam) domains, one Ig-like domain, and one true Ig domain. Based on our ClustalW analysis only the two TyrKc domains (hVEGFR-1 amino acids 819–933 and 991–1157) were significantly conserved among both vertebrate and invertebrate VEGFR sequences queried.
Fig. 2c. ClustalW analysis of ANG-1 proteins. The conserved domain of hANG-1 is called the fibrinogen-related domains (FrEDs). This section of hANG-1 (amino acids 281–496) is similar to the C terminal globular domain of fibrinogen-like domain. The fibrinogen-like domain represents the most conserved region of the human angiopoietins. Recent studies indicate that it comprises the receptor-binding portion of angiopoietin [Valenzuela et al., 1999]. Our ClustalW results with representative ANG-1 and ANG-1-like proteins that are found in vertebrates and invertebrates appear to be consistent with those findings.
Fig. 2d. ClustalW analysis of TIE-2 proteins. Like VEGFR, TIE-2R has several functional domains. However, it is the tyrosine kinase catalytic (TyrKc) domain that is the conserved catalytic core of the TIE-2 protein. In hTIE-2 the TyrKc domain is found from amino acid 819 to 1095. The Fugu TIE-1-like protein seems to be incomplete although the upstream section of its TyrKc domain is conserved when compared to the other vertebrate TIE-2 proteins.
Our data identified possible VEGF-like molecules in pufferfish and mosquito. The pVEGFA-like molecule, which had the E-value of 4e-53 when compared to hVEGFA, could be considered a PDGF- and VEGF-related factor. It contained a conserved region that had a > 50% homology between sequences. This region is the PDGF conserved domain region that is located between aa 50 and aa 120 (Fig. 2). The same was true for A. gambiae although the homology with the PDGF region was lower (although, it was over the 25% threshold that is used to determine homology between proteins of different species). Therefore, we consider the mqVEGFA-like molecule, that had an E-value of 1e-09 when compared to the hVEGFA, to be a PDGF- and VEGF-related factor as well. This E-value is consistent with the E-values of dVEGFA-like molecules. For VEGF, the structural similarity indicates that all of these proteins have a PDGF and/or VEGF family homologous region (found between aa 50 to 132 of all our inquiries). The only surprise was that this protein could not be identified in the genome of C. elegans. Yet, other possible VEGF-like family members (ceVEGFR-1-like, ceANG-1-like, and ceTIE-2R-like) were found in this nematode (Table 2). The reason for this is not known at this time.

We have identified several pVEGFR-like molecules (Fugu sequences SINFRUP0000007241390, SINFRUP00000051063, SINFRUP00000060218, SINFRUP00000084136, and SINFRUP00000066122). All four had an E-value of zero when compared to hVEGFR-1. We believe that mqVEGFR-like, with an E-value of 2e-85; and 2 other Anopheles sequences (ENSAngP00000013218, with the E-value of 2e-69; and ENSAngP00000019144 with its E-value of 4e-64), can be considered PDGFR- and VEGFR-like though there is obviously less similarity.

When hVEGFR-1 was queried for BLAST analysis into the NCBI Drosophila genome database the results identified sequences that had lower E-values than Drosophila proteins that were previously designated as PDGFR- and VEGFR-like proteins (Celera direct submission to NCBI). The sequences we identified were Drosophila CG10244-PA (dVEGFR-1-like) with an E-value of 7e-64 and Drosophila CG3277-PB with an E-value of 5e-54 when compared to hVEGFR. The E-values here are lower than dPDGFR- and dVEGFR-like CG8222-PB (E-value 3e-53), dPDGFR- and dVEGFR-like CG8222-PC (E-value 3e-53), and dPDGFR- and dVEGFR-like CG8222-PB (E-value 8e-53). With the identification of C. elegans, VEGFR-4-like (E-value 2e-63) and VEGFR-3-like (E-value 2e-63) molecules (ceVEGFR-1-like; Table 2, Fig. 2(b)), a potential problem arises. As mentioned, according to our analysis there appears to be no hVEGF-like protein identified in the C. elegans database, yet there appears to be VEGFR-1-like proteins. Experiments should be done with these proteins to see if they are performing VEGFR type functions. These experiments could include: To challenge C. elegans with VEGF to see if there is an increase in activity and to challenge a mammalian model system with C. elegans VEGFR-like protein to see if it binds to vertebrate VEGF (Table 2).

The ceVEGFR-like proteins were quite similar in terms of their E-value to dVEGFR-like proteins. These E-values obtained reinforced the similarity found in the conserved region of VEGFR in general. Most importantly, the region of similarity is consistent among all of the VEGF like proteins we looked at. This domain region is designated as KOg0200 (the fibroblast/platelet-derived growth factor receptor). The domain consists of related receptor tyrosine kinases that may function in signal transduction mechanisms (VEGFR Conserved Domain Search) (Fig. 2(b)). Yet, a ligand for the ceVEGFR-like molecule was not found. This raises the question of the function of VEGFR-like proteins in C. elegans. Even though we had success in finding VEGF-like protein in invertebrates and other vertebrates, the data showing the existence and the initial characterization of an ANG-1 protein is important.

The analysis of hANG-1 identified a group of similar proteins. The E-value results reported here were significant. They include pANG-1-like, mqANG-1-like, dANG-1-like, ceFibrinogen β and γ chain-like molecules (Table 2, Fig. 3). By looking at the E-values, one can see the decreasing similarity in the
Fig. 3. Human VEGF family and homologues. Phylogenetic trees were generated using the Gene Bee Phylogenetic that constructed a rectangular tree with bootstrap value (a) and a branched unrooted tree (b).

conserved regions that confers activity of the ANG-1 protein. Again, we hypothesize that functional experiments using these molecules could reveal other VEGF- and VEGFR-like proteins in those species.

Among the dANG-1-like and mqANG-1-like proteins, the E-values are significantly higher than that of dVEGF-like and mqVEGF-like proteins. This finding suggests that the ANG protein family is less diverse than the VEGF family (Table 2). The conserved region that extends from aa 281 to aa 496 seems to indicate a strong structural similarity between hANG-1 and pANG-1-like and mqANG-1-like molecules (Fig. 2(c), Table 2). Such findings were also evident in the amino acid similarity of our identified TIE-2R proteins.

As mentioned, ANG-1 confers its activity via the TIE-2R. Again, based on E-values we identified TIE-2R-like proteins that could be the receptor to the ANG-1-like protein identified. These TIE-2R-like proteins include pTIE-2R-like, mqTIE-2R-like, dTIE-2R-like, and ceMGFR-like.

The hTIE-2R has a tyrosine kinase catalytic domain that is part of a large family of phosphotransferases. These proteins are involved in the signal transduction mechanism of many receptors. What makes the hTIE-2R unique is how it is integrated into the cell membrane and its specificity to ANG-1 and ANG-2 (TIE-2 NCBI Locus Link). Since similar conserved regions are found in pTIE-2-like, dTIE-2R-like, mqTIE-2R-like, and ceMGFR-like molecules they could all be considered TIE-2R-like proteins (Table 2 and Fig. 2).

We also looked at several members of the VEGF family of proteins to examine the development of these dynamic proteins over time. The unrooted phylogenetic trees presented are based on other publications that have found this technique accurate in determining the evolutionary relationship of proteins [Hodge and Cope, 2000]. First, we created trees of just the hVEGF family itself. As predicted, VEGF, VEGFR, ANG and TIE proteins all associated with each other. What was not predicted is how they aligned with each other. It appears that possible VEGF and ANG proteins may have evolved along two separate pathways. However, their respective receptors seem to have evolved from a common protein ancestor of VEGF and yet they are involved in mediating both VEGF and ANG biological activities. This led us to propose the concept of protein structural divergence and functional convergence. That is, two structurally diverging proteins can be brought to functional convergence by having a similarly evolved protein from a common evolutionary background (Fig. 1(a) and (b)). We then took representative VEGF-like members from our BLAST analysis and performed a phylogenetic tree analysis of the hVEGF family and similar proteins.

We define each class by the first node that is found in the center of the unrooted tree (Fig. 3). There are no very large branches; therefore, there has been a minimization of parsimony algorithms and unique similarity among similar molecules. We would thus call similar proteins that are found to be in the hVEGF group orthologs. However, though this is true of the vertebrate VEGF-like family members, it seems that the invertebrate sequences sequester on their own branch. Therefore, there were actually five branches: the hVEGF, VEGFR, ANG and TIE branches and the invertebrate VEGF-like family member branch. The exception to this was ceANG-1-like. We did not identify a VEGF-like protein in C. elegans (Table 2).

The VEGFA and PGF proteins were found to cluster together with a bootstrap value of 100. Therefore, VEGFA appears to be more similar to PGF than to VEGFB, VEGFC or VEGFD. The proteins mqVEGFA and rVEGFA are significantly related to both VEGFA and PGF suggesting a common precursor.

Both the vertebrate and invertebrate orthologues of hANG that we identified clustered together. A similar finding occurs with the TIE-2 cluster. We observed that the TIE-1R and the TIE-2R may have arisen from separate precursors. The rTIE-2R protein was the only vertebrate TIE-2 molecule that did not cluster with the others. Finally, both dTIE-2R-like and mqTIE-2R-like and dVEGFR-1-like and
mqVEGFR-like molecules showed a relationship that indicated they may have arisen from a similar precursor. This result is consistent with the results we found with the hTIE-2R and VEGFR-1 proteins. The ceVEGFR-3-like and ceVEGFR-4-like proteins are similar. However, their existence without a VEGF-like molecule, which we could not find, is unusual (Fig. 3) and may warrant further investigation.

We estimated divergence time using molecular clock methods for the VEGF, VEGFR, ANG and TIE protein isoforms [Cutler, 2000a; Cutler 2000b; Gonnet, 1992] (Fig. 4). Our calculations correlated with the recent finding of VEGF-like proteins in leeches [Tettamanti et al., 2003]. This is one of the earliest indications of a VEGF-like protein since we were not able to find a VEGF-like molecule in nematodes. Only a VEGFR-like molecule has been found in nematodes.

Nematodes are thought to have evolved around 970 Mya before the split between protostomes and deuterostomes which occurred approximately 500 Mya. As mentioned, a recent report identified a VEGF-like protein in the leech [Tettamanti et al., 2003]. Leeches are annelids and part of the protostome branch of animals. This finding is consistent with our estimation of the emergence of VEGFR and VEGF precursors at about 500 Mya (Fig. 4). We agree with the Cho hypothesis [Cho et al., 2003] that VEGF- and VEGFR-1-like and -2-like molecules have a role both in regulation of blood cells and the regulation of angiogenesis in invertebrates. Although, when looking for similar proteins in invertebrates, there needs to be analysis at both the protein and gene level along with biological activity to confirm structural and functional homology.

The goal of this study was to look at the possible pathways taken by a highly divergent and dynamic family of proteins that have a role in both angiogenesis and immunity. Again, the fact that both the
VEGFR and TIE molecules appear to be derived from a common ancestor that produced VEGF and not ANG is a unique and surprising development. Our studies using immuno-informatics appears to be a reliable method in trying to understand the complex evolution of the angiogenesis and immune system.

REFERENCES

transfers. Genome Biol. 4, R19.


