

Identifying gene-protein interactions in response to environmental estrogenic chemicals using Bayesian networks: its implication in the etiology of pulmonary vascular lesions

Changwon Yoo^a, Quentin Felty^b

^a *Departments of Biostatistics and* ^b *Environmental & Occupational Health, Robert Stempel College of Public Health and Social Work, Florida International University, 11200 S.W. 8th Street, Miami, FL 33199; Email: Feltyq@fiu.edu*

Despite the extensive research in the field of vascular proliferative disease, the complex interaction of genes with environmental factors with respect to its impact on this disease is not yet fully understood. To date, results from the literature show no direct studies of vascular gene-environment interactions. Vascular lesions consist of actively proliferating endothelial cells that can lead to a tissue mass that blocks the artery and increases arterial pressure. Severe vascular lesions consequentially lead to right ventricular heart failure and death. This outcome is approximately twice more common in women than in men, but whether it may be attributed to the hormone estrogen is not clear. Based on the limited evidence, it appears that estrogenic environmental chemicals play a role in vascular lesion formation. In this study, Bayesian Networks was used to study which genes are influenced by proteins that we previously identified to be modulated by exposure to the environmental estrogen PCB153 and natural estrogen 17 β -estradiol. From the 96 spots obtained in our 2D-DIGE, we selected 13 spots for further protein identification analysis. The analyses identified 11 of the 13 spots of interest with high confidence protein score CI that was greater than 99%. The differentially expressed proteins regulated by PCB153 and estradiol were analyzed using IPA software to identify significant networks. The biological significance of TGFB1 network identified by IPA analysis coupled with the fact that some of the 11 identified proteins are involved in gene transcription lead us to use Bayesian networks to determine whether our identified proteins had any relation to our gene array data from similar estrogenic exposures. While our findings do not reveal any mechanisms involved in estrogen-induced vascular lesion formation, the identified relationships do provide a potential target pathway for future mechanistic studies.

Keywords: *Vascular lesions, environmental estrogens, polychlorinated biphenyls, pulmonary hypertension*

1. INTRODUCTION

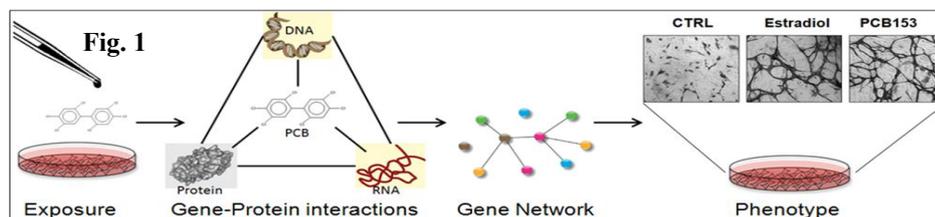
Adverse effects of estrogenic endocrine disruptors on the pulmonary vasculature:

Synthetic chemicals that mimic endogenous hormones can interfere with the normal development of both wildlife and humans. The presence of endocrine-disrupting chemicals (EDCs) within the environment raises concern as to whether the estrogenic EDCs exist at biologically significant levels. Although estrogen-induced gene expression has been well studied in vascular endothelial cells, there is a lack of knowledge about the gene networks affected by estrogenic EDCs such as polychlorinated biphenyls (PCBs). PCBs are a class of polychlorinated aromatic hydrocarbons composed of 209 discrete congeners. Due to their high lipophilicity and structural stability, PCBs are among the most extensively investigated persistent environmental pollutants that bioaccumulate in the food chain and are concentrated in fatty tissues (Kimbrough, 1995). Several epidemiological studies have shown a link between PCB exposure and increased risk of cardiovascular disease (Gustavsson and Hogstedt, 1997; Hay and Tarrel, 1997; Goncharov et al., 2008; Tokunaga and Kataoka, 2003; Sergeev and Carpenter, 2005). For example, serum PCB levels were reported to be significantly associated with the prevalence of hypertension in individuals not taking anti-hypertensive medication (Goncharov et al., 2011).

Pulmonary arterial hypertension (PAH) is roughly twice as common in women compared to men. Severe PAH is characterized by clustered proliferation of endothelial cells (ECs) in the lumina of small size pulmonary arteries resulting in the formation of complex vascular structures known as plexiform lesions. Although the etiology of PAH is not known, the possibility that unrestrained vascular endothelial cell proliferation is a fundamental problem in severe PAH cannot be ignored. Whether the formation of these vascular lesions is caused by estrogen or estrogenic EDCs is not known. Animal studies in models of pulmonary hypertension and scant clinical data suggest that estrogens are protective. Yet, the long-term use of estrogen therapy has been shown to increase the risk of cardiovascular disease in women (Rossouw, 2005). The dose of estrogens reportedly used in animal models and clinically may offer a potential explanation for the estrogen paradox. Clinically, estrogen is given at a “low dose” to minimize thrombotic risk and hormone-dependent malignancies. However, few *in vitro* and *in vivo* studies have studied the adverse effects of low dose estrogen exposure. Recently, ApoE mice exposed to a low dose estrogen treatment showed a significant increase in their atherosclerotic plaque score, whereas high-dose estrogen decreased aortic plaque burden (Freudenberger et al., 2010). It is biologically plausible that an estrogen lung disease connection exists. Blood vessels in the lung have been reported to increase and decrease in rhythm with a woman’s uterine lining changes (Farha et al., 2007). Furthermore, Farha et al. reported a higher level of lung blood vessel formation in ovariectomized female mice receiving estrogen compared with placebo. Angiogenic markers expressed in vascular lesions of severe pulmonary hypertension suggest a mechanism akin to dysregulated angiogenesis (Tuder et al., 2001). Since estrogen is a known mitogen of ECs that promotes vessel formation, exogenous estrogen exposure in the form of hormone therapy for contraception or management of menopause may support the angioproliferative endothelial phenotype. Considering the recent studies that question the previously reported cardio-protective effects of estrogens, there is a growing concern that estrogenic EDCs may increase the formation of pulmonary vascular lesions that manifest as severe PAH.

Using gene-protein interactions to predict toxic mechanisms of environmental EDCs:

There exists a need to construct disease-associated molecular networks to more accurately predict adverse health outcomes from exposure to environmental pollutants. As shown in **Fig. 1**, we observed an aggressive angiogenic phenotype in endothelial cells exposed to the endocrine disruptor PCB153. Using *in vitro* gene mRNA levels, we developed a molecular network to study the effects of estrogenic EDCs on vascular endothelial cells (Felty et al., 2010). Although our initial network was developed using genes that were directly affected by estrogen and PCB treatments, it ignored a more holistic approach by not including measurements such as xenoestrogen metabolites, proteins, post-translational modifications (i.e. phosphorylation, S-glutathionylation), or the interactions between proteins and DNA/RNA. Because physical interactions between proteins and DNA/RNA are likely to regulate the toxicological effects of estrogenic EDCs, relying solely on mRNA expression is not sufficient for estimating a gene network accurately. Therefore, in this paper we attempt to integrate data from both genomic and proteomic technologies with the aim of refining the previously estimated gene network. We also discuss the implications of this molecular network in understanding the pathogenesis of severe angioproliferative pulmonary arterial hypertension.



2. MATERIALS AND METHODS

Cell Culture & Treatment Conditions:

A telomerase-immortalized human microvascular endothelial cell line was used in this study (American Type Culture Collection, Manassas, VA). These endothelial cells have retained many of the phenotypic characteristics of primary endothelial cells including normal cell morphology and the capacity to form tubules in vitro (Venetsanakos et al., 2002). Cells were grown in Endothelial Cell Basal Medium-2 (EBM-2) supplemented with EGM 2-MV Single-Quots from Lonza (Walkersville, MD). Estrogen treatments were studied under the following culture conditions: 75–80% confluent cultures were washed and serum starved in phenol red-free medium for 3h. Thereafter, the cells were treated with either PCB126, PCB153, or 17 β -estradiol in phenol red-free medium supplemented with 10% charcoal/dextran-treated fetal bovine serum to avoid any estrogenic activity. Estrogen treatments were given under the following culture conditions: 75–80% confluent cultures were washed and serum starved in phenol red-free medium for 3h. The cells were treated with either PCB153 (100ng/ml) or 17 β -estradiol (1ng/ml) for another 24 hours. PCB blood levels have been reported to reach approximately 1000ng/ml (~3 μ M) in occupationally exposed individuals (Wassermann et al., 1979). Our previous studies also showed endothelial cell proliferation and vascular tube formation at physiological doses of 17 β -estradiol (Felty, 2006; Felty and Porther, 2008); therefore HMVEC were exposed to estradiol at 1ng/ml (~3.6nM). PCB congeners 2,2',4,4',5,5'-hexachlorobiphenyl (PCB153) and 3,3',4,4',5-pentachlorobiphenyl (PCB126) were purchased from AccuStandard (New Haven, CT). All other chemicals and reagents were purchased from Sigma (St. Louis, MO).

Proteomics:

Protein lysates were precipitated with 100% methanol, and each sample was resuspended in 2D cell lysis buffer. For each sample, 30 μ g of protein was mixed with 200 pmol CyDy2, Cy3 or Cy5 and kept in the dark on ice for 30 min. The labeled samples were then mixed together. 2X 2-D Sample buffer, destreak solution (GE Healthcare), and rehydration buffer were added. The samples were mixed well and spun down before loading onto the strip holder. After loading the labeled samples to pH 3-10 linear IPG strips (GE Healthcare), the Isoelectric Focusing (IEF) was run for 12 hours rehydration at 20°C, followed by 500 V for 1000 Vhr, 1000 V for 2000 Vhr, and 8000 V for 24000 Vhr. Upon finishing the IEF, the IPG strips were incubated in freshly made equilibration buffer-1 for 15 minutes with gentle shaking. Then the strips were rinsed in freshly made equilibration buffer-2 for 10 minutes with gentle shaking. IPG strips were rinsed in SDS-gel running buffer before transferring into 12% SDS-gels. The SDS-gels were run at 15°C until the dye front ran out of the gels. Gel images were scanned immediately following the SDS-PAGE using Typhoon TRIO (Amersham BioSciences) of Cy2, Cy3 and Cy5 labeled protein images. The scanned images were then analyzed by Image Quant software Version 6.0 (Amersham BioSciences), followed by in-gel analysis using DeCyder software Version 6.0 (Amersham BioSciences). The DeCyder spot detection algorithm calculated ratio (volume of a spot from the secondary image/volume of the corresponding spot from the primary image), and a threshold of 1.5 fold change was set.

The spots of interest were picked up by Ettan Spot Picker (Amersham BioSciences) based on the in-gel analysis and spot picking design by DeCyder software. The gel spots were digested in-gel with trypsin protease. The digested tryptic peptides were eluted from the Zip-tip with Matrix solution and spotted on the MALDI plate (model ABI 01-192-6-AB). MALDI-TOF (MS) and TOF/TOF (tandem MS/MS) were performed on an ABI 4700 mass spectrometer. MALDI-TOF mass spectra were acquired in reflection positive ion mode, averaging 4300 laser shots per spectrum. TOF/TOF tandem MS fragmentation spectra were acquired for each sample, averaging 4300 laser shots per fragmentation spectrum on each of the 10 most abundant ions present in each sample. Both of the resulting peptide mass and the associated fragmentation spectra were submitted to GPS Explorer version 3.5 equipped with MASCOT search engine (Matrix Science) to search the database of National Center for Biotechnology Information non-redundant (NCBI nr). The parameters were set at 800-4000 Da to create the “peak list.” Searches were performed without constraining protein molecular weight or isoelectric point, with variable carbamidomethylation of cysteine and oxidation of methionine residues, and with one missed cleavage allowed in the search parameters. Mass tolerance was set at 0.3 Da and 100 ppm. Candidates with either protein score confidence interval (C.I.) % or Ion C.I.% greater than 95 were considered significant.

Gene Array:

PCR array analysis of cDNA from RNA isolated from HMVEC was performed in triplicate using RT² Profiler PCR array for Human Angiogenesis from SABiosciences, in accordance with the manufacturer's instructions. PCR arrays were used to compare the relative expression of 84 genes for each array. Relative changes in gene expression between the treated and untreated HMVEC RNA were calculated by the $\Delta\Delta C_T$

method as previously described (Livak and Schmittgen, 2001). First, target genes in each array were normalized to the endogenous control or housekeeping genes: RPL13A, GAPDH, or ACTB. Next, each treatment group was normalized to the untreated group (vehicle control, 0.1% DMSO). Using the SABiosciences PCR Array Data Analysis Web Portal, fold changes in gene expression were analyzed between groups using the equation $2^{-\Delta\Delta CT}$, where $\Delta\Delta CT = [C_T^{\text{target gene}} - \text{Avg. } C_T^{\text{housekeeping genes}}]_{\text{treated group}} - [C_T^{\text{target gene}} - \text{Avg. } C_T^{\text{housekeeping genes}}]_{\text{control group}}$. Each value is a repeat of three independent experiments. ANOVA and Fisher's least significant difference test were used for multi-group comparisons, with a P value of less than 0.05 considered significant.

Bayesian Networks Analysis:

Given the emergence of datasets in medicine and biology with large number of variables, Bayesian Networks have been successful in developing efficient algorithms that scale up to datasets involving hundreds of variables in learning high quality predictive models from genomic data (Yoo and Cooper, 2004). A Bayesian Network is a directed acyclic graph in which each node represents a variable and each arc represents probabilistic influence. In Bayesian Networks, each arc is interpreted as a direct influence between a parent node (variable) and a child node, relative to the other nodes in the network.

Bayesian networks are also built based on the causal Markov conditions (Pearl, 1988). We can find out any conditional relationships among the variables in the Bayesian network (Charniak, 1991). This can be

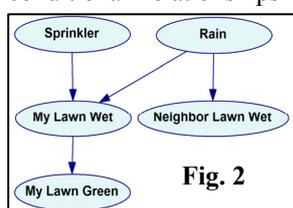
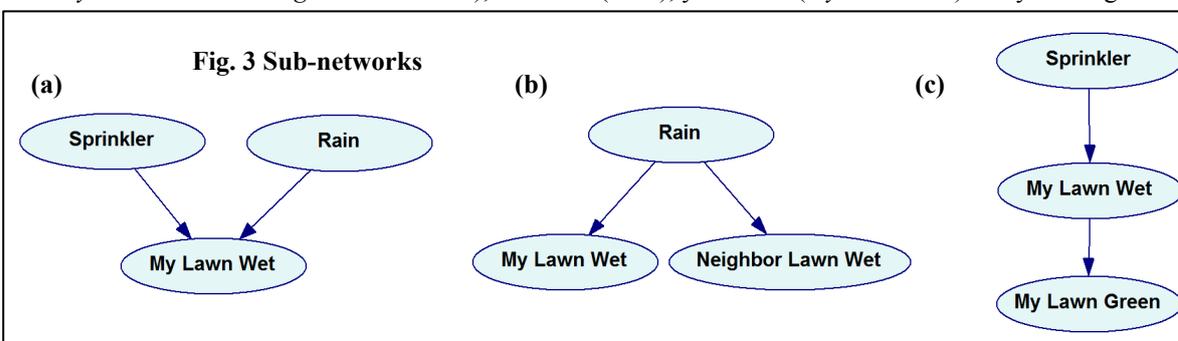


Fig. 2

understood with the following simple example of a Bayesian network (Fig. 2). In the above network, either it rains (Rain) or you turn on your sprinkler (Sprinkler) which can make your lawn wet (My Lawn Wet). Also if your lawn gets enough water, it gets green (My Lawn Green). Of course, your neighbor, who does not have a sprinkler, can get his lawn wet (Neighbor Lawn Wet) when it rains. In the network shown in Fig. 2, we can identify the following three sub-networks:

In Fig. 3 (a) (we call it **converging arcs** going into My Lawn Wet from Sprinkler and Rain), if you know your lawn is wet (My Lawn Wet) and you know it didn't rain (not Rain) then there is a high chance that your sprinkler was on (Sprinkler), in other words, **if converging arcs into a variable C from variables A and B, then A and B are associated given C**. Also note that in Fig. 3 (b) (we call it **diverging arcs** from Rain into My Lawn Wet and Neighbor Lawn Wet), if it rains (Rain), your lawn (My Lawn Wet) and your neighbor's



lawn gets wet (Neighbor Lawn Wet). If you know it rained (Rain), knowing your lawn is wet will not tell you much more about your neighbor's lawn being wet, in other words, **if diverging arcs from a variable C into variables A and B, then A and B are not associated given C**. In Fig. 3 (c) (we call it **serial arcs** from Sprinkler to My Lawn Wet to My Lawn Green), if sprinkler is on (Sprinkler), then my lawn gets wet (My Lawn Wet), and eventually, your lawn gets green (My Lawn Green). In this case, if you know your lawn is wet (My Lawn Wet), then knowing whether the sprinkler is on or off (Sprinkler) will not tell you much more about whether your lawn will get green (My Lawn Green), in other words, **if serial arcs from a variable A to a variable C to a variable B, then A and B is not associated given C**.

3. RESULTS & DISCUSSION

Proteomic Results

Spot number	Top Ranked Protein Name(Species)	Location*	Type(s)*	Accession No.	Protein MW	Protein PI	Pep.Count	Protein Score	Protein Score C. I. %	Total Ion Score	Total Ion C. I. %
6	filamin B [Homo sapiens]	Cytoplasm	other	gi 9664263	71134.8	6.25	10	88	99.963	26	0
21	Lamin B1 [Homo sapiens]	Nucleus	other	gi 15126742	66366.6	5.13	9	77	99.595	45	97.636
24	MX1 protein [Homo sapiens]	Nucleus	enzyme	gi 40225574	66810.8	5.9	9	80	99.756	23	0
30	lamin A/C transcript variant 1 [Homo sapiens]	Nucleus	other	gi 57014047	74025.6	6.44	33	443	100	150	100
34	far upstream element-binding protein [Homo sapiens]	Nucleus	transcription regulator	gi 17402900	67518.5	7.18	13	132	100	47	98.289
72	cathepsin B	Cytoplasm	peptidase	gi 741376	17143	5.44	7	327	100	259	100
79	Chain A. Human Electron Transfer Flavoprotein	Cytoplasm	transporter	gi 2781202	33075.4	6.95	11	188	100	94	100
85	vacuolar protein sorting 29 isoform 1 [Homo sapiens]	Cytoplasm	transporter	gi 7706441	20492.7	6.29	9	186	100	92	100
87	myosin regulatory light chain MRCL2 isoform A [Homo sapiens]	Cytoplasm	other	gi 15809016	19766.5	4.71	10	475	100	346	100
92	ISG15 ubiquitin-like modifier [Homo sapiens]	Cytoplasm	other	gi 14550514	17903.3	6.84	5	290	100	228	100
93	Chain A. Cyclophilin A Complexed With Dipeptide Gly-Pro	Cytoplasm	enzyme	gi 1633054	17869.8	7.82	10	340	100	225	100

A limitation of the 2D approach is that it will often miss small molecular weight and less abundant proteins. Therefore, spots of interest were not chosen entirely on dramatic changes by estrogen treatments. From the 96 spots, we selected 13 spots for further protein identification analysis. The analyses identified 11 of the 13 spots of interest with high confidence protein score C.I. that was greater than 99% (Table 1). We have applied Bayesian Networks learning method to further understand interactions among the proteins affected by the estrogenic EDC exposure. Bayesian Networks is a knowledge representation formalism that has been widely used in translational biomedical research for constructing statistical graphical models and for generating causal hypotheses from data (Yoo and Cooper, 2004). Since we wanted to identify interactions among proteins that respond to the treatments, we have categorized each proteins's expression level to low (if the expression level is one standard deviation lower than the mean expression level), no change, and high (if the expression level is one standard deviation higher than the mean expression level), based on the normalized expression level of all the genes in control and treatments. To learn Bayesian networks among the genes, we then used computer software called Bayesian network inference with Java objects (Banjo) (Hartemink, 2010). We have categorized each protein's fold change (compared to control) under PCB153 or Estradiol as very low, low, high, and very high. We also applied the Bayesian networks learning algorithm to learn probable interactions among the proteins. In the node labeled Experiment, it represents either control (no application of PCB153 or Estradiol) or application of PCB153 or Estradiol. Since we had a small sample size (n=3), we identified many equivalent networks that best fit the data. Thus, instead of searching for the

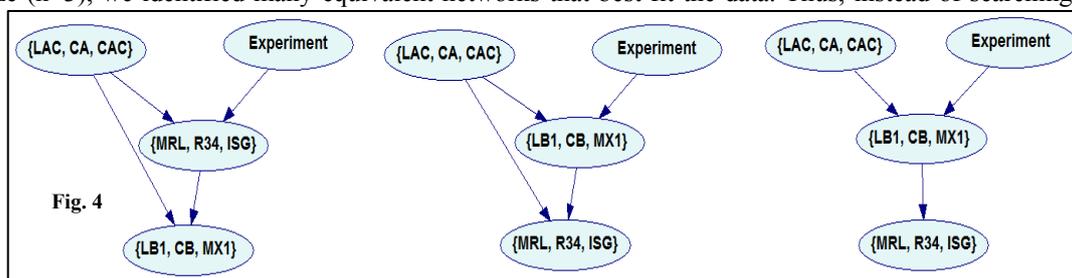


Fig. 4

Figure 4. Summary of Seven Equivalent Protein Bayesian network. Analyzing the seven equivalent networks, we have identified three groups of proteins that always showed tight association: {LAC, CA, CAC}, {LB1, CB, MX1} and {MRL, R34, ISG}. {LAC= lamin A/C transcript variant 1, CA= human electron transfer flavoprotein, CAC = cyclophilin A Chain}; {LB1= lamin B1, CB= cathepsin B , MX1= MX1 protein}; {MRL= myosin regulatory light chain MRCL2 isoform A, R34= RNA binding motif protein 34, ISG= ISG15 ubiquitin-like modifier}

highest scoring network, we have reviewed the equivalent networks for novel patterns. We have executed Banjo with ten independent analyses (each analysis ran for three hours, i.e., total of 30 hours) and identified seven equivalent networks. Analyzing the seven equivalent networks, we have identified three groups (sets) of proteins that always showed tight association: {LAC, CA, CAC}, {LR1, CB, MX1} and {MRL, R34, ISG}. We summarized the relationships among the three groups (sets) of proteins (Fig. 4). All three networks show that proteins LAC, CA, and CAC are not associated with the experiment, however, because of the **converging arcs** into {LR1, CB, MX1} (or {MRL, R34, ISG}) from Experiment and {LAC, CA, CAC}, we can infer that possible interactions between {LAC, CA, CAC} and {LR1, CB, MX1} (or {MRL, R34, ISG}) makes them associate with the experiment.

Integration of Gene-Protein data:

Bayesian network inference with Java objects (Banjo) software was used to analyze PCR array data of the estrogen treatments. Bayesian gene networks were reconstructed using the expression profile of all 84 genes from the RT² Profiler PCR Array for Human Angiogenesis. We then combined the gene-protein data to

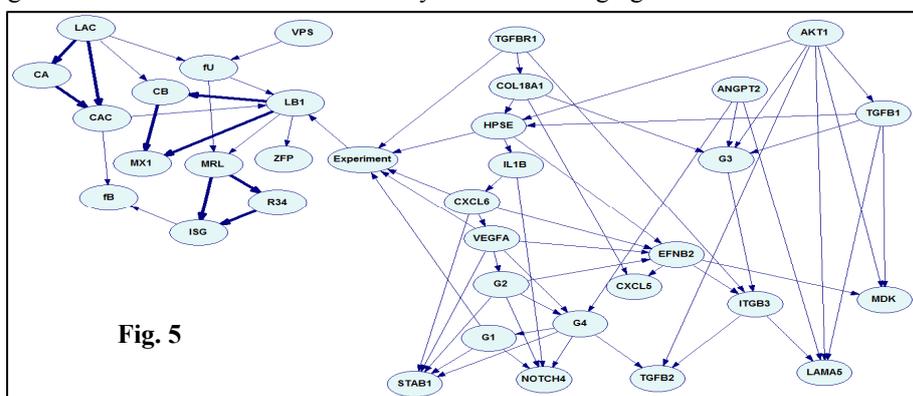
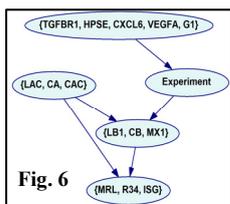


Fig. 5

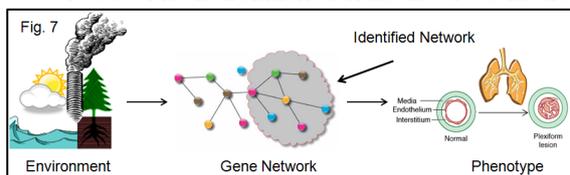
further construct a more complete network (Fig. 5). We have also summarized the interactions among the genes and proteins in Fig. 6. From Fig. 5, if we know the states of TGFB1, HPSE, CXCL6, VEGFA, and G1, then no other

Figure 5. Combined gene-protein Bayesian Network. Bold arcs represent consistent relationships from the seven independent Bayesian network analyses from the proteomics data. Groups of genes in the network are designated as G1= ANPEP, IL8, NRP2, PGF, TEK, TIMP2; G2=PLAU, HPRT1; G3=IL6, VEGFC, CXCL10; G4= FLT1, FGF2, PDGFA.



genes in the network will be associated with Experiment. Thus, in **Fig. 6**, we have only included TGFBR1, HPSE, CXCL6, VEGFA, and G1 into the summarized model. **Fig. 6**, however, indicates (because of the **converging arcs** into {LR1, CB, MX1} (or {MRL, R34, ISG}) from {TGFBR1, HPSE, CXCL6, VEGFA, G1} and {LAC, CA, CAC}) possible interactions among genes TGFBR1, HPSE, CXCL6, VEGFA, and G1 with proteins LAC, CA, CAC only through LB1, CB, MX1, MRL, R34, and/or ISG. For example, the gene VEGFA interacts with proteins lamin B1, cathepsin B, and MX1 protein when cells are exposed to estrogen and estrogenic EDCs. Therefore, if we know that vascular cells are exposed to estrogenic chemicals then we can predict that the gene VEGFA interacts with those proteins. Interestingly, the proteins lamin A/C, electron transfer protein, and cyclophilin A cannot directly interact with these genes, but do interact with these genes through proteins LB1, CB, MX1, MRL, R34, and/or ISG. From this combined gene-protein network we were able to identify a biologically relevant pathway involved in the process of angiogenesis that could explain the phenotype changes we observe in endothelial cells exposed to estrogenic EDCs. Matrix metalloproteases (MMPs) have been thought to be the most important proteases involved in angiogenesis. There was a recent report that vascular endothelial growth factor (VEGFA) did not increase MMPs in the process of angiogenesis, but instead the authors saw a striking increase in cathepsin B mRNA levels (Chang et al., 2009). Cathepsins are secreted to degrade matrix proteins and have significant roles in angiogenesis. This finding of VEGFA induction of cathepsin B is interesting because according to the authors of the study it was truly an unexpected result. Based on the study by Chang et al., we are encouraged that our molecular network of angiogenesis was able to identify this unique interaction between VEGFA and Cathepsin B. This may hold promise for using a molecular network to generate new hypotheses to explain the angiogenic phenotype.

Pulmonary hypertension is roughly twice as common in women compared to men. Severe PAH is characterized by clustered proliferation of endothelial cells. The mechanism of this pathology is not known but it has been postulated to be similar to dysregulated angiogenesis. Since estrogen is a known mitogen of endothelial cells, we postulated that long-term estrogen exposure from the environment may contribute to the formation of vascular lesions. Therefore, we generated a molecular network based on in vitro experiments (as described in **Fig. 1**) that will enable us to know not only genes but also the proteins affected by estrogenic EDCs. Limitations in our network come from the costs it would take to have characterized all 96 proteins identified in our proteomic experiments. Thus, we were constrained to only 11 proteins to include in this combined gene-protein network. Ideally, we would like to use a molecular network model to identify gene/proteins involved in the development of pulmonary vascular lesions. As shown in **Fig. 7**, the implications of this research is to create a complete molecular network so we can identify sets of molecules associated with environmental risks involved in the development of pulmonary vascular lesions. The results



from such a network will provide critical knowledge for developing protective measures against environmental estrogenic risk factors, and the development of novel preventive or therapeutic strategies for pulmonary hypertension related to these environmental risks.

REFERENCES

- Chang,S.H., Kanasaki,K., Gocheva,V., Blum,G., Harper,J., Moses,M.A., Shih,S.C., Nagy,J.A., Joyce,J., Bogyo,M., Kalluri,R., and Dvorak,H.F. (2009). VEGF-A induces angiogenesis by perturbing the cathepsin-cysteine protease inhibitor balance in venules, causing basement membrane degradation and mother vessel formation. *Cancer Res.* 69, 4537-4544.
- Charniak,E. (1991). Bayesian Networks Without Tears. *Ai Magazine* 12, 50-63.
- Farha,S., Asosingh,K., Laskowski,D., Licina,L., Sekiguchi,H., Losordo,D.W., Dweik,R.A., Wiedemann,H.P., and Erzurum,S.C. (2007). Pulmonary gas transfer related to markers of angiogenesis during the menstrual cycle. *J. Appl. Physiol* 103, 1789-1795.
- Felty,Q. (2006). Estrogen-induced DNA synthesis in vascular endothelial cells is mediated by ROS signaling. *BMC. Cardiovasc. Disord.* 6, 16.
- Felty,Q. and Porther,N. (2008). Estrogen-induced redox sensitive Id3 signaling controls the growth of vascular cells. *Atherosclerosis* 198, 12-21.

Yoo and Felty, Identifying gene-protein interactions in response to environmental estrogenic chemicals using Bayesian networks: its implication in the etiology of pulmonary vascular lesions

Felty,Q., Yoo,C., and Kennedy,A. (2010). Gene expression profile of endothelial cells exposed to estrogenic environmental compounds: implications to pulmonary vascular lesions. *Life Sci.* 86, 919-927.

Freudenberger,T., Oppermann,M., Heim,H.K., Mayer,P., Kojda,G., Schror,K., and Fischer,J.W. (2010). Proatherogenic effects of estradiol in a model of accelerated atherosclerosis in ovariectomized ApoE-deficient mice. *Basic Res. Cardiol.* 105, 479-486.

Goncharov,A., Haase,R.F., Santiago-Rivera,A., Morse,G., McCaffrey,R.J., Rej,R., and Carpenter,D.O. (2008). High serum PCBs are associated with elevation of serum lipids and cardiovascular disease in a Native American population. *Environ Res.* 106, 226-239.

Goncharov,A., Pavuk,M., Foushee,H.R., and Carpenter,D.O. (2011). Blood pressure in relation to concentrations of PCB congeners and chlorinated pesticides. *Environ. Health Perspect.* 119, 319-325.

Gustavsson,P. and Hogstedt,C. (1997). A cohort study of Swedish capacitor manufacturing workers exposed to polychlorinated biphenyls (PCBs). *Am J. Ind. Med.* 32, 234-239.

Hartemink, A. J. Banjo: structure learning of static and dynamic Bayesian networks. 2010.
Ref Type: Online Source

Hay,A. and Tarrel,J. (1997). Mortality of power workers exposed to phenoxy herbicides and polychlorinated biphenyls in waste transformer oil. *Ann. N. Y. Acad. Sci* 837, 138-156.

Kimbrough,R.D. (1995). Polychlorinated biphenyls (PCBs) and human health: an update. *Crit Rev. Toxicol.* 25, 133-163.

Livak,K.J. and Schmittgen,T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-408.

Pearl,J. (1988). Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. (San Mateo: Morgan Kaufmann).

Rossouw,J.E. (2005). Coronary heart disease in menopausal women: implications of primary and secondary prevention trials of hormones. *Maturitas* 51, 51-63.

Sergeev,A.V. and Carpenter,D.O. (2005). Hospitalization rates for coronary heart disease in relation to residence near areas contaminated with persistent organic pollutants and other pollutants. *Environ Health Perspect.* 113, 756-761.

Tokunaga,S. and Kataoka,K. (2003). [A longitudinal analysis on the association of serum lipids and lipoproteins concentrations with blood polychlorinated biphenyls level in chronic "Yusho" patients]. *Fukuoka Igaku Zasshi* 94, 110-117.

Tuder,R.M., Chacon,M., Alger,L., Wang,J., Taraseviciene-Stewart,L., Kasahara,Y., Cool,C.D., Bishop,A.E., Geraci,M., Semenza,G.L., Yacoub,M., Polak,J.M., and Voelkel,N.F. (2001). Expression of angiogenesis-related molecules in plexiform lesions in severe pulmonary hypertension: evidence for a process of disordered angiogenesis. *J. Pathol.* 195, 367-374.

Venetsanakos,E., Mirza,A., Fanton,C., Romanov,S.R., Ilsty,T., and McMahon,M. (2002). Induction of tubulogenesis in telomerase-immortalized human microvascular endothelial cells by glioblastoma cells. *Exp. Cell Res.* 273, 21-33.

Wassermann,M., Wassermann,D., Cucos,S., and Miller,H.J. (1979). World PCBs map: storage and effects in man and his biologic environment in the 1970s. *Ann. N. Y. Acad. Sci.* 320, 69-124.

Yoo,C. and Cooper,G.F. (2004). An evaluation of a system that recommends microarray experiments to perform to discover gene-regulation pathways. *Artif. Intell. Med.* 31, 169-182.