Gene Regulation & Transcriptomics

Methodology for analyzing temporal patterns of differential coexpression using meta-analysis

Jesse Gillis[1] and Paul Pavlidis[1,*]

[1] Centre for High-Throughput Biology and Department of Psychiatry, University of British Columbia, Vancouver BC, 177 Michael Smith Laboratories 2185 East Mall, University of British Columbia,Vancouver, BC, V6T1Z4.

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Associate Editor: XXXXXXX

ABSTRACT

Motivation: Differential coexpression is a change in coexpression between genes that may reflect 'rewiring' of transcriptional networks. It has previously been hypothesized that such changes might be occurring over time in the lifespan of an organism (aging). While both differential expression and coexpression of genes have been previously studied in aging, differential coexpression has not. Generalizing differential coexpression analysis to many time points presents a methodological challenge. Here we introduce a method for analyzing changes in coexpression across ordered groups (e.g., over time) and test its usefulness.

Results: Our method is based on the use of the wavelet transform to efficiently represent changes in coexpression at multiple time scales. We used published microarray studies categorized by age to test the methodology. We validated the methodology by testing our ability to reconstruct Gene Ontology (GO) categories using our measure of differential coexpression and compared this result to using coexpression alone. Our method allows significant improvement in characterizing these groups of genes. In addition, we found that our method finds more significant changes in gene relationships compared to several other methods of expressing temporal relationships between genes, such as coexpression over time.

Supplementary data: http://www.chibi.ubc.ca/diffExAge

Contact: paul@bioinformatics.ubc.ca

1 INTRODUCTION

Differential coexpression is defined as a change in the correlation relationships between genes. It is a natural extension of the concept of 'guilt by association', in which functional relationships between genes are thought to be reflected in coexpression relationships (Eisen et al,, 1998, Lee et al. 2004). Differential coexpression posits that changes in coexpression can be biologically relevant, and occur with or without changes in gene expression levels (differential expression). We think of differential coexpression as potentially revealing 'rewiring' of gene network, reflecting dynamic changes in the regulatory relationships between genes which can then be 'read out' at the level of transcription. Because of the potential importance of network rewiring, differential coexpression could be useful for uncovering molecular mechanisms of normal processes such as development and aging as well as disease processes. A schematic outlining the features of differential coexpression is provided in Figure 1.

Differential coexpression has previously been studied primarily in the context of changes in coexpression between two groups (Watson, 2006, Choi et al., 2005, Kostka and Spang, 2004). However, no method to handle ordered groups, such as over age or time, has been proposed.

The current study was motivated by our interest in studying human aging. For our purposes, we take 'aging' to include both developmental and normal senescent changes. In searching for biomarkers for ageing it has been usual to look for differential expression over time (Zahn et al., 2007, Lee et al., 1999). The equivalent task in differential coexpression analysis would compare coexpression across time. Previous expression profiling studies have demonstrated that the expression patterns of age-regulated genes are indicators for a functional measure of ageing in humans (Zahn et al., 2006, Rodwell et al., 2004). Because a large array of functional changes occur over age, age-related change may be a rich resource for differential coexpression – many linked changes in functional relationships or rewiring of transcriptional networks.

The life-long and complex time course of aging means both that there are many potential natural divisions to group different ages together. One approach to analyzing changes would be to take a derivative of gene coexpression across time (over increasing age groups), thus providing the differential coexpression between each age group and the next. At the extreme one might consider comparing just two groups (e.g., 'old' and 'young'). However, the derivative or two-group comparison will fail to detect gradual changes which can only be characterized over the long term. Another possible approach to characterize multiple time points would be to compare every age group to every other, but this is highly redundant and ignores the temporal relationship between data points.

A good method for differential coexpression should have the following properties:

(1) It would characterize the change in coexpression at each time.

*To whom correspondence should be addressed.

(2) It would characterize the change in coexpression over the long term and the short term.

(3) It would form a basis set for the temporal data.

For two groups of data, it would reduce to conventional differential coexpression (i.e. a difference between gene correlations).

These three properties suggest we require a transformation incorporating both changes in scale and timing. Broadly, this defines a group of transformations of (typically) temporal data known as time-frequency transforms, the best-known of which is the spectrogram (Cohen, 1989), effectively a collection of bandpass filters plotted along the same axis. Thus, this is a characterization both of the change in time and scale of activity. In the context of aging, this would represent differential coexpression between both fine time points, old vs. very old, for example, and larger scales, post-adult vs. pre-adult, for example.

A more recent innovation which exploits the fact that large scale temporal processes are often themselves not changing rapidly is the wavelet transform. Wavelets are a popular technique to analyze temporal data in many fields and has previously been explored for use in temporal gene expression data (Song et al., 2007). The wavelet transform uses a window that varies in frequency resolution, so that at low frequencies, the frequency resolution is very good, and at high frequencies, the temporal resolution is very good (Daubechies, 1990). Typically the temporal data for this analysis consists of enough time points that wavelets are useful to summarize activity at multiple scales and times. In our case, the use of wavelets arises not just as a commonly useful way of representing data, but as a generalization of the definition of differential coexpression to multiple and continuous conditions (time).

Thus the wavelet transform of the correlation between two genes provides differential coexpression data over time at multiple scales. In order to reduce to conventional two-group differential coexpression, a wavelet transform of age-specific coexpression data should be seen as taking the difference between groups. This corresponds to the first proposed wavelet transform, the Haar wavelet transform (Haar, 1909).

In this study we use the Gemma database of publicly-available microarray studies (Hamer et al., submitted; http://www.chibi.ubc.ca/Gemma) for a meta-analysis of expression patterns, as a proof of principle. We show that functionally related genes tend to have similar differential coexpression values, indicating an expected common change in functional relationship over time. We further demonstrate that differential coexpression data allows similarly GO categorized genes to be grouped together with significantly greater efficacy than coexpression alone. Finally, we show that the Haar wavelet basis is more informative than using the ordered coexpression values basis or a discrete derivative basis (differences between successive age groupings). Because ageing is a process so strongly characterized by changes in function, these techniques developed to characterize ageing may also shed light on how ageing changes function and thus how dysfunction occurs.

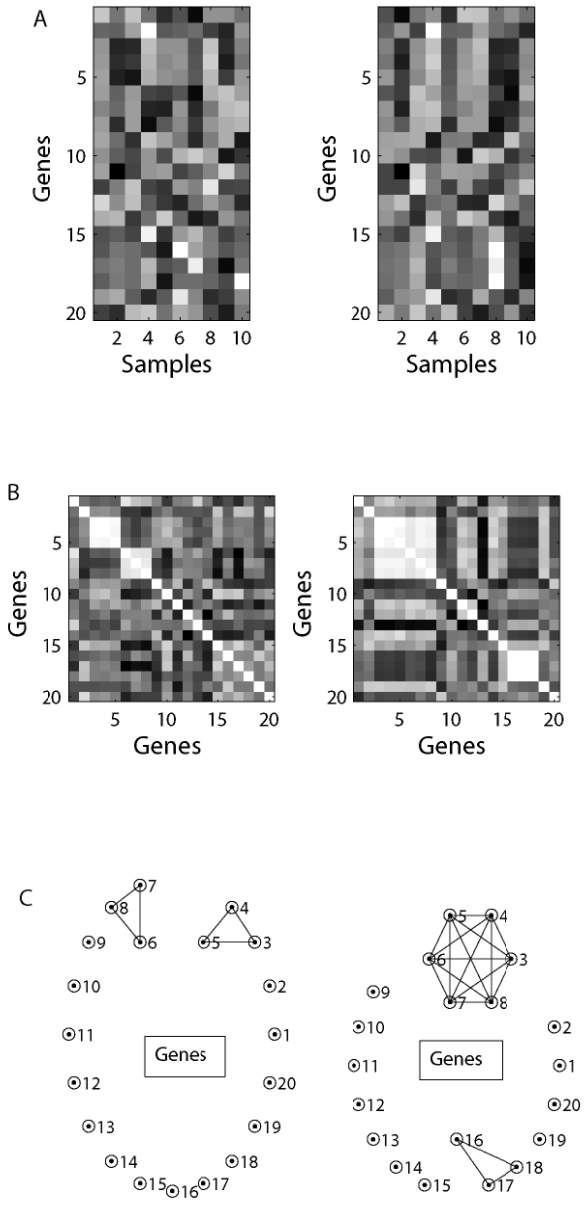

Figure 1. Schematic of differential coexpression. The left and right sides of the figure correspond to two hypothetical experimental conditions. A. Heatmap representation of expression levels of 20 genes in 10 samples per condition; lighter shades indicate higher relative expression. The correlations among some genes changes, e.g. genes 16-18. B. Correlation matrix heatmaps corresponding to the data in A. Light colors indicate higher correlations. The changes in the position and size of the 'blocks' of highly coexpressed genes changes between conditions. C. Coexpression networks generated by thresholding the correlations between each pair of genes, illustrating the concept of ‘rewiring’.

2 METHODS

2.1 Data grouping and normalization

Human microarray studies from Gemma's database were categorized by their subject's ages into the four groups; "prenatal", "child/young adult" (0-18 years), "adult" (19-54), and "older adult" (55+). Studies spanning more than one adult grouping were categorized by whether their ages ranged older or younger than adult. Thus a study overlapping adult and older adult, for example, 50 to 100 years old, would be categorized within the "older adult" category and similarly a study consisting of 10-20 would be classified only as "child/young adult". In this way, trends away from the mean age are captured. It is important to note that the studies used were not necessarily designed to study age effects, and include a variety of tissues. The selection procedure yielded 8 to 13 studies for each age group and 37 in all, encompassing 2803 individual microarrays (13 to 404 arrays per study). Pearson correlations between pairs of genes were calculated for each study. For genes annotated by Gemma as having multiple probes in a given study, the correlations for the probes were averaged. Correlation values between pairs of genes for each study were converted to standard deviations from 0 in the correlation distribution as determined by the theoretical standard deviation from 0 of the Fisher transformation z = 0.5 log ((1+r)/(1-r)) (Fisher, 1915), and averaged across common data groupings.

To allow the investigation of differential expression over age, we computed a relative rank-based measure of expression level for each gene. Each gene's expression level for each study was averaged across samples in each study, converted into a rank with the study, and then averaged within each age group.

2.2 Wavelet transform

For a time series, x(t), the continuous wavelet transform at scale a and time t, W(a, t) is given by:

$$W(a,t) = \frac{1}{\sqrt{a}} \int x(u) g^* \left(\frac{u-t}{a}\right) du \quad (1)$$

where g is the mother wavelet function and g* its complex conjugate, where g must satisfy certain constraints such as continuity, integrability, square integrability, and admissibility (wavelike). In our case, the wavelet is the step function, or Haar wavelet, g(x)=1 for 0>x>1/2, g(x)=-1 for 1/2>x>1, and g(x)=0 otherwise. The discrete wavelet transform is then an implementation of the transform over discrete time points.

Over our four age groups (prenatal, child/young adult, adult, older adult), the Haar basis consists of four values:

- The averaged correlation between genes across all four time points: (1/2, 1/2, 1/2, 1/2)
- The averaged correlation difference pre-adult and post-adult: (1/2, 1/2, -1/2, -1/2)
- The averaged correlation difference prenatal to child/young adult: (1/sqrt(2), -1/sqrt(2), 0, 0)
- The averaged correlation difference adult to older adult: (0, 0, 1/sqrt(2),-1/sqrt(2))

The four wavelet differential coexpression values are the dot product of these vectors with the age grouped expression data. The wavelet lifting scheme was used to calculate the coefficients (Kaplan, 2002; Jense and Cour-Harbo, 2001). Note that the first wavelet coefficient represents the average coexpression across all ages. The remaining three coefficients represent differential coexpression and no aspect of coexpression (since the coefficients are orthogonal). Henceforth "differential coexpression coefficients" excludes the first coefficient. The variance of temporal coefficient values across the experiments used was calculated.

We performed the Haar transform both on the averaged correlation between pairs of genes (as a time series across our age groups) as well as on the averaged ranks of individual genes. Differential expression may be a confound for differential coexpression so to ensure that the differential coexpression findings may not be explained by simpler underlying changes, genes exhibiting significant differential expression values at a given time and scale (top 5% rank change) were removed. If a gene exhibited differential expression for one coefficient, other values not exhibiting differential expression were retained. Following this procedure, 927 genes were characterized as exhibiting differential expression for each coefficient and removed from analysis for differential coexpression. This allows our analysis to focus on genes which are relatively stable in expression level, but which might exhibit changes in coexpression relationships with other genes.

2.3 Validation

All Gene Ontology (GO) groups of genes containing 25-30 members were chosen using the Gemma web services (see supplementary data). This size range was chosen for computational tractability of cross-validation, and to avoid using GO categories which overlap extensively. This generated 139 separate GO groups encompassing 2925 distinct genes out of a total 10764 covered by the GO categories.

We used these GO groups for validation of the differential coexpression results. By an analogy with the use of coexpression to predict gene function, we propose that each gene within a given GO set might have a characteristic differential coexpression relationship with genes inside the set and a characteristic relationship with genes outside the set. For each gene inside the set, gene B, the respective distributions of differential coexpression coefficients can be calculated. An arbitrary gene, gene j, outside the set has differential coexpression values with gene B that may also be calculated. We may then ask if gene j's relationship with gene B resembles gene B's relationship with other genes inside the set or with genes outside the set. As a control, the same calculation was performed using the coefficient representing coexpression.

A leave-one-out methodology was employed in the following fashion for genes A-Z:

(1) The Gene A of genes A-Z was left out.
(2) For the first differential coexpression coefficient, the distribution of differential coexpression coefficients between gene B and genes C-Z was calculated.
(3) For each gene, gene j, not in the set B-Z, the first differential coexpression coefficient was calculated between gene B and gene j.
(4) The distribution of first differential coexpression coefficients between gene B and all genes not in C-Z or j was constructed.
(5) The odds ratio for the distributions in steps 3 and 4 was calculated for the value between gene B and gene j.

(6) Steps 2-5 were repeated for genes C-Z and the geometric mean of the coefficients taken.

(7) Steps 2-6 were repeated for each differential coexpression coefficient and the geometric mean was taken.

(8) Steps 1-7 were repeated rotating through each gene in the set to generate receiver operator characteristic (ROC) curves.

An additional basis set of differential coexpression coefficients was generated as a comparison to the wavelet transform. The first coefficient remained the first value in time, while the remaining coefficients were calculated by taking a discrete derivative, or difference between successive pairs of values across time. This set will be referred to as the derivative set. Conceptually this overlaps with the time basis 1st coefficient and the wavelet 3rd, and 4th coefficients, but lacks any scale variation, such as that found in the wavelet second coefficient. Significant coefficients ($p<0.05$) were

calculated for the direct time series coefficients, the wavelet coefficients, and the derivative coefficients. The number of significant coefficients for each gene pair was calculated.

A schematic example of the wavelet transform method is shown in Figure 2.

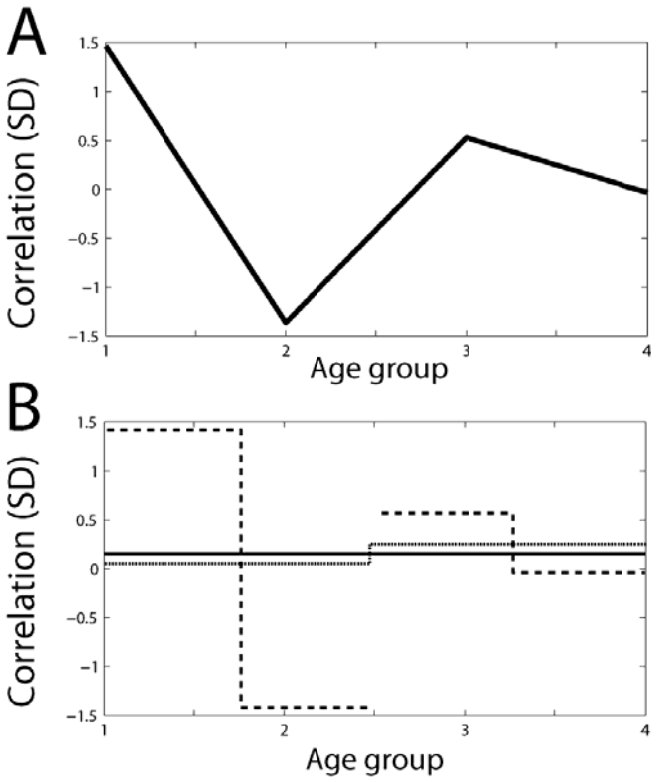

Figure 2. Method schematic. A. An examples of coexpression data for one pair of genes plotted across agre groups. B. The wavelet basis set is shown summing to generate the original data points from A. The solid line is the mean value, the dotted line shows the difference between the first half and second half, and the dashed line shows the difference between first and second groups then third and fourth groups. Each successively smaller scale is graphed on top of the sum of previous scales, summing to the original.

3 RESULTS

To test our approach, we analyzed differential coexpression across human lifespan in a corpus of 37 expression studies (2803 individual microarrays in total). This produces 4 symmetric matrices of 18534 by 18534 genes with potential wavelet coefficient values, consisting of coexpression or differential coexpression. Because this a large and complex data set, important trends may exist outside of examining only the most statistically significant cases. Our validation uses all the data; however, we have also constructed an adjacency matrix consisting of the most significant gene-pairs across all coefficients ($p<0.01$ and top 100 significance for at least one gene with respect to the other). This produced 1585917 significant gene-pair relationships, available as supplementary data. The focus of the remaining analysis presented in this paper is aimed at evaluating the proposed approach. First, we explored whether differential coexpression is relevant to gene function using an analysis of Gene Ontology categories. Second, we compared our Haar-wavelet transformation approach to one based on derivatives of coexpression changes. An analysis of the biological relevance of the age-related changes we found will be described elsewhere.

3.1 Gene function is reflected in differential coexpression patterns

Previous studies have found coexpression to be an indicator of functional relationships (Lee et al., 2004). Aging is a process characterized by a change in functional relationships (Zahn et al., 2007). Thus we hypothesized that aging may be a process in which differential coexpression is significant. Because ageing is a widespread process, we hypothesized that many functional categories may exhibit significant differential coexpression. GO sets are functionally categorized genes and thus they serve as a test set to observe whether differential coexpression is a useful tool to characterize gene function over time.

Results of leave-one-out validation using the GO sets (see Methods) are shown in Figure 3. Given a training set (a GO group minus one in-group gene, rotated through each in-group gene) classification of a testing set (all genes outside the GO group plus one in-group gene, rotated through each in-group gene) was performed. Note that genes exhibiting strong differential expression across age were removed to ensure that changes in coexpression are not better explained by differential expression of the same genes (Li et al., 2004). The area under the curve (AUC) for the ROC curve is 0.77 with a standard deviation 0.05 across the GO groups. The AUC of 0.77 represents the probability of correctly assigning a higher score to a random in-group gene over a random out-group gene, given the two. As a control we also attempted to classify using coexpression alone and obtained an AUC of 0.65

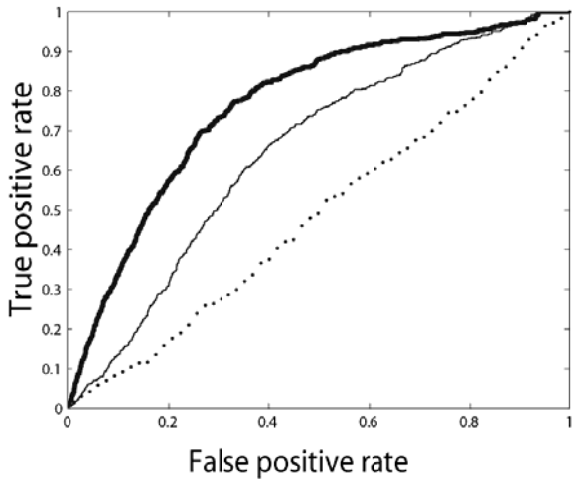

Figure 3. Functional categories of genes can be predicted by differential coexpression. The thick line curve shows the result of leave one out validation to generate an ROC curve for reconstructing GO categories containing 25-30 genes using their differential coexpression values (AUC 0.77). The thin line curve show the ROC curve if instead of differential coexpression values, coexpression values alone are used (AUC 0.65). The dotted line curve shows the result-ing ROC curve if in place of GO sets, random sets with 25-30 genes are used (AUC 0.49).

(Figure 3). Random sets of genes (of the same size) generated the expected identity line with an AUC of ~0.5 (Figure 3). Using coexpression in conjunction with differential coexpression does improve performance further, although slightly (AUC of 0.79, Figure 3), suggesting that coexpression is not captured entirely by differential coexpression.

3.2 Comparison to derivative basis set

The wavelet coefficients map naturally onto conventional definitions of differential coexpression, but there are other possible mappings. One such would be the discrete derivative, starting with the first time point's expression level and then subsequent coefficients reflecting changes from the previous value. It might also be reasonable to question whether changes in coexpression are helpful to consider at all, rather than simply independently observing each time point for coexpression. In that case, the natural basis set would be to use the time basis set (age groups 1-4) of coexpression data directly. Each pair of genes has 4 coefficients associated with it. There is little point to this if significance in one coefficient implies significance in another.

Figure 4 shows the fraction of significant coefficients sharing other significant coefficients for the same gene pair, and the number of common coefficients (with 1 indicating that the gene pair with that significant coefficient had only one such). The wavelet coefficient performs better, with 45% of its significant coefficients being unique in that gene pair, than either temporal coefficients with 35% such or derivative coefficients with 41% – more of the wavelet coefficient's significant values are independent across values.

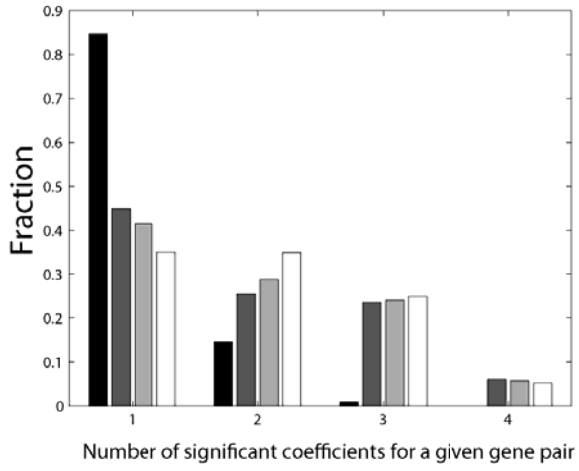

Figure 4. The percentage of significant coefficients in a basis set representing age for a gene pair in which other coefficients are also significant. The x axis shows the number of significant coefficients, e.g., the values at 1 indicate that the given percentage of coefficients are found in gene pairs with no other significant coefficients. Black shows the result of random data uncorrelated across coefficients; dark grey, the result of wavelet coefficients; light grey, the result of derivative coefficients; and white, the result of the age groups directly.

4 DISCUSSION

These results suggest that wavelet analysis of differential coexpression is a useful tool for capturing functional relationships between genes as they change over ordered sets of conditions. In addition to their good performance compared to a derivate approach, wavelets had an additional feature making them attractive for meta-analysis, in that individual studies with samples containing multiple 'ages' could be combined and analyzed at multiple scales. For example, if a study covered the age groups prenatal and child/young adult, it could be included in both, ensuring it contributed to only the first and second wavelet coefficients, as appropriate for the scale at which the study was performed.

In a sense our use of the wavelet transform is unconventional because we have only four time points, in contrast to more typical applications where temporal resolution is much higher. However, we think of the wavelet transformation as a useful basis set with a biological underpinning, and which has a convenient generalization. Thus the wavelet method could equally well be applied to studies specifically geared toward the analysis of data with different or finer temporal resolution, or to other ordered conditions.

4.1 Differential coexpression over time using wavelets

The inferior performance of the temporal coefficient method was not surprising, since the total coexpression effects are not an independent coefficient and dominate significance measures. That is, if a gene pair is highly coexpressed, even if it is also highly variable over age (yielding large coefficients in the other two bases), it is still quite likely to remain highly coexpressed at all ages. The inferior performance of the derivative basis suggests that longer scale dynamics are relevant to expression changes with age. This would seem consistent with the proposal of Barker et al. (1989) that fetal programming can play a significant role in determining factors affecting longevity (Gluckman et al., 2005).

The difference between derivative and wavelet bases may become more important with greater resolution age categories. The temporal basis set had an average standard deviation of 0.22 between experiments compared to the normalized 1.0 between mean gene pair values. Some part of this is likely due to overlapping age categories, and while this is large enough to substantially effect the ordering of many differential coexpression values near the mean, it is not large enough to alter extreme values. Because the temporal basis values were then transformed to constitute the derivative and wavelet bases, the significance values are stable with respect to experiment variability. This has bearing on the success of the GO validation since the distinction at each step was between potentially extreme or unusual (in-group) values and necessarily baseline (out-group) values.

One possible concern with our study is that we have mixed data sets from various tissues. The precise tissue for prenatal vs. adult, for example, may differ. Likewise there are study parameters (such as population type) that might vary with age in studies. It could then be argued that differential coexpression across age that we see is simply a proxy measure for differential coexpression between across tissue type. However, this is not consistent with the superior performance of the wavelet coefficients compared to the other basis. The wavelet coefficients average data at a larger scale, thus diminishing tissue effects. Were this effect dominant, treating

each age group independently (as attempted in the temporal basis) instead of as part of an ordered group would be most useful.

A final possibility in lieu of the suggested bases would be to determine a good basis post hoc, as is effectively done in any dimension reduction. This presents a number of problems of its own. First, it will vary from dataset to dataset and introduce new normalization difficulties. The variation from dataset to dataset would also reduce the general applicability of any results since any past findings would have to be reinterpreted in the context of whatever basis set is calculated for the new work. In addition, there would be no reason to see the new basis set as a form of differential coexpression precisely since it is unlikely that a single component would end up wholly representing coexpression (thereby removing it from the others, as was the case with the wavelet basis). More broadly, it is desirable that the method for producing the basis set be recursive in the sense that the addition of more age groupings should leave previous findings interpretable. Doubling the resolution (or duration) is naturally handled by the wavelet basis without altering the meaning of previous findings. This would not be so of a post hoc basis, which could change dramatically through the addition of new groupings, even for the same data. Finally, a principled transformation of the data can be geared to offer value as an interpretive tool, as opposed to a purely methodological tool.

4.2 Biological interpretation of the wavelet approach

Our data covers a range of age categories including both development and senescence. Two fundamental theories of senescence are Williams (1957) theory of antagonistic pleiotropy and Medawar's (1946) theory of mutation accumulation. These theories present an interesting interpretation in the context of our differential coexpression coefficients. Antagonistic pleiotropy posits a long-scale connection between early states and late states, in which a characteristic useful in youth is negative later (e.g., Rodier et al., 2007). A changing functional role over age in this way would be a good candidate to find differential coexpression, and, in particular, we would expect such differential coexpression to be present in the coefficient at the appropriate time-scale. That is, our second wavelet coefficient (representing long term differential coexpression) may be thought of as mapping onto youthful tradeoff theories of ageing. Mutation accumulation, on the other hand, maps more readily onto our last wavelet coefficient, representing (relatively) short term and late stage changes in coexpression. The third wavelet coefficient maps most readily onto developmental change, representing as it does, rapid and early changes in coexpression which we interpret at rapid and early changes in function for development. More specific mechanistic interpretations of function and dysfunction over age also map more readily onto a wavelet basis than others because they typically involve both a factor in time and scale. As previously mentioned, Barker's theory of fetal programming implies suggests a long term effect between early states and late states. One well studied mechanism reviewed by Maric (2007) involves fetal programming for high blood pressure. Because the wavelet coefficients can capture both the scale and timing of this event, they might serve to elucidate the unknown genetic causes for the well characterized physiological changes. With only four groupings by age, interpretation of this sort in our data must remain somewhat restrained, but finer resolution age groupings could make this a valuable characteristic of our method.

5 CONCLUSIONS

Our wavelet based methodology for determining age-related differential coexpression performs better than either a derivative based method, or using the age groups independently. The wavelet basis set also lends itself to ready interpretation in terms of both evolutionary and physiological mechanisms of ageing and can be seen as a natural generalization of two-category differential coexpression. The good performance across the multiple GO sets implies that age related differential coexpression may be a common process due to the degree to which ageing produces changes in function and functional relationships. Because our wavelet-based method for differential coexpression draws upon such a well established signal processing tool for temporal data, it offers a well characterized, efficient and convenient avenue for further study.

ACKNOWLEDGEMENTS

We thank Gavin Ha and Raymond Lim for technical support.

Funding: The work was supported by NIH grant GM076990, Michael Smith Foundation for Health Research career award and a CIHR New Investigator award.

REFERENCES

Barker, D.J., Osmond, C., Golding, J., Kuh, D. and Wadsworth, M.E. (1989) Growth in utero, blood pressure in childhood and adult life, and mortality from cardiovascular disease, *BMJ*, **298**, 564-567.

Becker, K.G., Owen, A.B. and Kim, S.K. (2006) Transcriptional profiling of aging in human muscle reveals a common aging signature, *PLoS Genet*, **2**, e115.

Choi, J.K., Yu, U., Kim, S. and Yoo, O.J. (2003) Combining multiple microarray studies and modeling interstudy variation, *Bioinformatics*, **19 Suppl 1**, i84-90.

Choi, J.K., Yu, U., Yoo, O.J. and Kim, S. (2005) Differential coexpression analysis using microarray data and its application to human cancer, *Bioinformatics*, **21**, 4348-4355.

Cohen, L. (1989) Time-frequency distributions-a review, *Proceedings of the IEEE*, **77**, 40.

Daubechies, I. (1990) The wavelet transform, time-frequency localization and signal analysis, *IEEE Transactions on Information Theory*, **36**, 44.

Eisen, M.B., Spellman, P.T., Brown, P.O., and Botstein, D. (1998) Cluster analysis and display of genome-wide expression patterns, *Proc Natl Acad Sci USA*, **95,** 14863—14868.

Gluckman, P.D., Hanson, M.A., Morton, S.M. and Pinal, C.S. (2005) Life-long echoes--a critical analysis of the developmental origins of adult disease model, *Biol Neonate*, **87**, 127-139.

Haar, A. (1909) Zur Theorie der orthogonalen Funktionensysteme, *Mathematische Annalen*, **69**, 40.

Jense, A. and Cour-Harbo, A.l. (2001) *Ripples in Mathematics: the Discrete Wavelet Transform*. Springer.

Kaplan, I. (2002) http://www.bearcave.com/software/java/wavelets/basiclift.html.

Kostka, D. and Spang, R. (2004) Finding Disease Specific Alterations in the Coexpression of Genes, *Bioinformatics*, **20**, 5.

Lee, C.K., Klopp, R.G., Weindruch, R. and Prolla, T.A. (1999) Gene expression profile of aging and its retardation by caloric restriction, *Science*, **285**, 1390-1393.

Lee, H.K., Hsu, A.K., Sajdak, J., Qin, J. and Pavlidis, P. (2004) Coexpression analysis of human genes across many microarray data sets, *Genome Res*, **14**, 1085-1094.

Maric, C. (2007) Mechanisms of fetal programming of adult hypertension: role of sex hormones, *Hypertension*, **50**, 605-606.

Medawar, P.B. (1946) Old age and natural death, *Mod.*, **1**, 27.

Rodier, F., Campisi, J. and Bhaumik, D. (2007) Two faces of p53: aging and tumor suppression, *Nucleic Acids Res*, **35**, 7475-7484.

Rodwell, G.E., Sonu, R., Zahn, J.M., Lund, J., Wilhelmy, J., Wang, L., Xiao, W., Mindrinos, M., Crane, E., Segal, E., Myers, B.D., Brooks, J.D., Davis, R.W., Higgins, J., Owen, A.B. and Kim, S.K. (2004) A transcriptional profile of aging in the human kidney, *PLoS Biol*, **2**, e427.

Song, J.Z., Duan, K.M., Ware, T. and Surette, M. (2007) The wavelet-based cluster analysis for temporal gene expression data, *EURASIP J Bioinform Syst Biol*, 39382.

Watson, M. (2006) CoXpress: differential co-expression in gene expression data, *BMC Bioinformatics*, **7**, 509.

Williams, G.C. (1957) Pleiotropy, natural selection, and the evolution of senescence, *Evolution*, **11**, 13.

Zahn, J.M., Poosala, S., Owen, A.B., Ingram, D.K., Lustig, A., Carter, A., Weeraratna, A.T., Taub, D.D., Gorospe, M., Mazan-Mamczarz, K., Lakatta, E.G., Boheler, K.R., Xu, X., Mattson, M.P., Falco, G., Ko, M.S., Schlessinger, D., Firman, J., Kummerfeld, S.K., Wood, W.H., 3rd, Zonderman, A.B., Kim, S.K. and Becker, K.G. (2007) AGEMAP: a gene expression database for aging in mice, *PLoS Genet*, **3**, e201.

Zahn, J.M., Sonu, R., Vogel, H., Crane, E., Mazan-Mamczarz, K., Rabkin, R., Davis, R.W., Becker, K.G., Owen, A.B. and Kim, S.K. (2006) Transcriptional profiling of aging in human muscle reveals a common aging signature, *PLoS Genet*, **2**, e115.