GENES REGULATED BY CALORIC RESTRICTION HAVE UNIQUE ROLES WITHIN TRANSCRIPTIONAL NETWORKS

William R. Swindell
University of Michigan, Departments of Pathology and Geriatrics, Ann Arbor MI, 48109-2200

Abstract
Caloric restriction (CR) has received much interest as an intervention that delays age-related disease and increases lifespan. Whole-genome microarrays have been used to identify specific genes underlying these effects, and in mice, this has led to the identification of genes with expression responses to CR that are shared across multiple tissue types. Such CR-regulated genes represent strong candidates for future investigation, but have been understood only as a list, without regard to their broader role within transcriptional networks. In this study, co-expression and network properties of CR-regulated genes were investigated using data generated by more than 600 Affymetrix microarrays. This analysis identified groups of co-expressed genes and regulatory factors associated with the mammalian CR response, and uncovered surprising network properties of CR-regulated genes. Genes downregulated by CR were highly connected and located in dense network regions. In contrast, CR-upregulated genes were weakly connected and positioned in sparse network regions. Some network properties were mirrored by CR-regulated genes from invertebrate models, suggesting an evolutionary basis for the observed patterns. These findings contribute to a systems-level picture of how CR influences transcription within mammalian cells, and point towards a comprehensive understanding of CR in terms of its influence on biological networks.

Keywords
ageing; connectivity; dietary restriction; lifespan; longevity; microarray

1. Introduction
Caloric restriction (CR) has been found to increase longevity and is associated with a wide range of health benefits at late ages of the mammalian lifespan. CR diets represent a 30 – 40% reduction from ad libitum caloric intake, with a dietary composition carefully designed to avoid malnutrition and starvation effects. Several systemic effects of this diet that may contribute to longevity and reduction of age-related disease, such as elevated insulin sensitivity, reduced metabolic rate and decreased adiposity (Bertrand et al., 1980; Ballor, 1991; Gresl et al., 2001; Blanc et al., 2003). At the cellular level, some or all of these effects may lead to states favorable for longevity, and an important goal is to understand which genes, proteins and biochemical pathways are most critical in this regard. The simplest hypothesis is that CR diets promote similar cellular modifications in multiple mammalian tissues, such that extended lifespan is due to a small number of cellular processes that are not tissue-specific. In support of this notion, recent studies have documented gene expression responses to CR that occur in the same fashion across a variety of tissue types, several of which appear related to oxidative stress resistance or tumor suppression (Spindler and Dhabhi, 2007; Swindell, 2008). These
“common” responses to CR have been uncovered through the mining of multiple genome-wide microarray datasets. This data-driven approach to target selection is appealing because of its objectivity, which offsets the tendency of researchers to focus on certain well-known pathways that, from a genome-wide perspective, may or may not emerge as noteworthy candidates for detailed study. Currently, however, common gene expression responses to CR are understood only as a list of genes that are frequently upregulated or downregulated in mice provided CR diets (Swindell, 2008). This viewpoint provides limited insight into relevant biochemical pathways and does not illuminate the potentially unique roles of CR-regulated genes within a broader transcriptional network.

Network approaches to the analysis of DNA microarray data can reveal interesting properties of genes that are not discernable through standard differential expression analysis (Guelzim et al., 2002; Alon, 2003; Bray, 2003; Ge et al., 2003; Ferrarini et al., 2005). Gene expression is regulated by individual transcription factors, or combinations of such factors, that bind to DNA sequences to promote or inhibit expression of target genes. This basic molecular mechanism leads naturally toward a network view of transcription in which individual genes serve as nodes connected by edges to other genes (Lee et al., 2002). Transcriptional networks can be constructed from microarray data based upon co-expression analysis (Carter et al., 2004; Rice et al., 2005; Wolfe et al., 2005; Wang et al., 2006; Freeman et al., 2007; Huttenhower et al., 2007). Co-expression analysis connects genes within a network when the correlation between their expression profiles exceeds a certain threshold. This approach leads to networks in which connections may represent co-regulation of gene sets by the same transcription factor(s), or regulation of target genes by other genes that encode transcription factor proteins. In either case, connected genes are often elements of the same biological pathway (Genter et al., 2003; Zhang et al., 2004; Wolfe et al., 2005; Ulitsky and Shamir, 2007), and such genes may participate in similar biological processes (but see Allocco et al., 2004 and Yeung et al., 2004a). An interesting property of transcription networks, indeed most biological networks, is a “scale-free” topology, where few hub genes are highly connected to many other genes, with numerous weakly-connected genes located peripherally to highly-connected hubs (Guelzim et al., 2002; Siegal et al., 2007). The density of transcriptional networks, therefore, is distributed in non-random fashion, with certain compact regions corresponding to large gene groups with similar expression patterns. This architecture makes networks sensitive to disruption of highly connected hub genes, but has the advantage of imparting resistance to random perturbation (Albert et al., 2000).

The positioning of CR-regulated genes within transcriptional networks has not been previously investigated. One possibility is that CR-regulated genes are disproportionately associated with dense, highly connected regions of transcriptional networks. CR appears to be effective at increasing lifespan across a wide range of animal species, including yeast, worms, flies and mice, suggesting that mechanisms underlying the effects of CR have been conserved throughout evolution. These mechanisms may be linked to a conserved adaptive response that evolved to delay growth and conserve energy when resources are scarce (Holliday, 1989; Hart and Turturro, 1998; Kirkwood and Shanley, 2005), and genes underlying these processes may have key roles in transcriptional networks and survival in general. Previous studies have found that these properties, evolutionary conservation and essentiality, are characteristic of highly-connected hubs within transcriptional and protein networks (Jeong et al., 2001; Bergmann et al., 2003; Carlson et al., 2006). This observation is consistent with the view that transcriptional networks evolve by adaptively adding new connections, building denser webs around conserved genes over evolutionary time. There is some indication, for example, that evolution of cortical regions in the human brain has been accompanied by elevated density of transcriptional networks (Oldham et al., 2006). This evolutionary view suggests that CR, which is thought to involve conserved and essential pathways, will modulate expression of genes located in dense regions of transcriptional networks. This hypothesis has not been addressed.
by previous analyses, however, in part because an understanding of which genes have expression patterns consistently associated with CR diets in mammals has begun to emerge only recently (Spindler and Dhahbi, 2007; Swindell, 2008).

This study uses co-expression analysis to investigate network properties of genes regulated by CR across multiple tissue types in mice. Previously, Swindell (2008) pooled together data from independent microarray experiments, and identified 12 genes significantly downregulated by CR in five or more tissues, along with 16 genes significantly upregulated by CR in five or more tissues. In the present study, the co-expression and network properties of these genes are investigated using four independent datasets constructed from 674 Affymetrix microarrays. The results identify robust co-expression relationships among CR-regulated genes, which provide clues regarding key regulatory pathways activated or inhibited by CR. In addition, new and unexpected properties of CR-regulated genes are identified in the context of transcriptional networks. Genes downregulated by CR display unusually high connectivity, suggesting that such genes are disproportionately located in dense network regions. At the same time, however, genes upregulated by CR display unusually low connectivity, indicating that CR upregulates genes positioned in sparse, poorly connected network regions. Interestingly, network properties of CR-downregulated and CR-upregulated genes from mouse are partially shared by CR-regulated genes from the Drosophila and C. elegans model systems, and are strongly mirrored by C. elegans genes expressed specifically in nondauers and dauers, respectively. These observations both confirm and challenge intuitive notions regarding how CR influences transcriptional networks, and provide broad systems-level insight into the transcriptional properties of CR-regulated genes.

2. Materials and Methods

2.1 Mouse Expression Profiling Datasets

Genes regulated by CR across five or more tissue types were previously identified based upon comparative analysis of microarray datasets (Table 1) (Swindell, 2008). To investigate the co-expression and network properties of these genes, four independent expression profiling datasets were constructed (Tissue, Developmental, Treatment and Mutation series). Each dataset was based upon a set of CEL files obtained from the Gene Expression Omnibus database (Barrett et al., 2007). All CEL files were generated from the Affymetrix Mouse Genome 430 2.0 array platform, which is the most comprehensive oligonucleotide platform currently available, with approximately 45,000 transcripts based on sequences derived from the Genbank, dbEST and RefSeq databases. In total, 674 CEL files were used to construct the profiling datasets, with 115, 170, 204 and 185 files, respectively, used to generate the Tissue, Developmental, Treatment and Mutation series. CEL files assigned to each series were normalized as a group using robust multichip average (Irizarry et al., 2003). The assignment of CEL files to each series was performed deliberately to ensure that each dataset would contain 60 conditions for expression profiling (after averaging across replicate CEL files associated with the same experimental treatment). The 60 conditions included in each series represented a wide range of tissue and cell types, which follows a commonly adopted approach to transcriptional profiling that has been effective at identifying co-expression relationships in previous studies (e.g., Zhang et al., 2004; Schmid et al., 2005). At the same time, each dataset was also unique in the types of conditions emphasized, with emphasis placed on either maximal diversity in adult tissues (Tissue series), early developmental stages (Developmental series), experimental treatments applied to tissue and organ systems (Treatment series) or genetic mutations (Mutation series). Supplemental Data File 1 provides a full description of the 60 conditions in each dataset.
2.2 Preprocessing and Co-expression Analyses

A standard preprocessing procedure was applied to expression profiling datasets prior to evaluation of co-expression and network patterns. For each transcript, the average expression level across the 60 conditions was subtracted from gene expression measurements. This served to “center” expression values for each gene, such that highly expressed genes were comparable to genes with low expression levels. Next, for each of the 60 conditions, expression values were standardized using the average expression value and standard deviation across all transcripts. This ensured that each condition was weighted equally during clustering procedures and evaluation of co-expression.

Co-expression analyses were used to identify CR-regulated genes with correlated expression patterns. This was done by performing hierarchical cluster analysis on expression profiles associated with CR-regulated genes, and by identifying genes throughout the genome with expression profiles correlated with those of each CR-regulated gene. In both cases, distance between genes was determined by averaging absolute Pearson correlation coefficients from each of the four expression profiling datasets (Tissue, Developmental, Treatment and Mutation series). The use of multiple datasets to determine an “average distance” between expression profiles provides a reliable co-expression measure that is robust to spurious correlations driven by outliers or experimental noise (Lee et al., 2004).

2.3 Network Analyses of CR-regulated genes

The network connectivity properties of CR-regulated genes were investigated using three metrics (correlation order statistics, local connectivity index and correlation percentiles), which are similar to measures that have been used in previous studies (e.g., Oldham et al., 2006; Dong and Horvath, 2007). For a given transcript i, let \( r_i \) represent a vector containing \( N \) Pearson correlation coefficients, obtained by calculating correlations between the expression profile of transcript i and each of the other transcripts represented on an oligonucleotide array. In the case of the Affymetrix Mouse 430 2.0 array, for example, the value of \( N \) is approximately 45,000 (equal to the total number of probesets). The exact value of \( N \) varies slightly among transcripts on the same array, since for a given transcript \( i \), the vector \( r_i \) is here defined to exclude correlations generated from transcripts annotated with the same gene symbol as transcript \( i \). This ensures that the elements of the vector \( r_i \) are not biased for cases in which different transcripts on the same array represent the same gene.

Let \(|r_i|\) represent a vector obtained by taking the absolute value of the elements in \( r_i \), such that values within \(|r_i|\) range from 0 to 1. For a given transcript \( i \), the magnitude of values within \(|r_i|\) reflects the connectivity strength between transcript \( i \) and all other transcripts on the same array. In the framework developed by Dong and Horvath (2007), for example, the vector \(|r_i|\) is analogous to a single row of an adjacency matrix. The abundance of “large” values within \(|r_i|\) determines the number of connections to the node representing transcript \( i \) in a co-expression network, where “large” is determined by a thresholding rule. A network could be constructed, for example, by drawing connections between transcripts for which the correlation between expression profiles exceeds 0.80 or 0.90 (e.g., Freeman et al., 2007), but a wide range of other thresholding methods have been adopted as well (e.g., Carlson et al., 2006; Oldham et al., 2006). Since there are many options for selecting connectivity thresholds, methods in the present study do not require thresholding of connectivity strengths, but instead focus on numeric properties of the \(|r|\) vectors associated with CR-regulated genes. This approach avoids dichotomizing gene connectivity as either “present” or “absent”, and emphasizes continuity of network connection strengths. In addition, by using this approach, it is expected that results will have general implications, with minimal dependence on the particular method chosen to construct a co-expression network.
Network connections between a gene and its immediate neighbors are based on strong correlations and referred to as local connections (Dong and Horvath 2007). For a given transcript, the abundance and strength of these connections are determined by the largest values within the vector $|r_i|$. Such larger values are characterized by correlation order statistics, which provide an intuitive measure for evaluating the strength of the local connections for a given transcript. For a given transcript $i$, the correlation order statistic $r_{i,m}$ represents the $m$th largest value within the vector $|r_i|$. For example, the value $r_{i,1}$ represents the largest absolute correlation value between the $i$th transcript and any other transcript represented on the same array. The value of $r_{i,2}$ represents the second largest correlation, while $r_{i,3}$ represents the third largest correlation, and so on for the $k$th correlation order statistic. Transcripts positioned in dense network regions are expected to have large correlation order statistics in general, and for highly connected hubs, order statistics should decline slowly with increasing $k$.

It is useful to summarize the overall strength of local connections for a transcript based upon its $m$-largest correlation order statistics $r_{i(1)} \ldots r_{i(m)}$. For this purpose, the $m$th local connectivity index ($r_{L,m}$) is defined as the average of the first $m$ correlation order statistics $r_{i(1)} \ldots r_{i(m)}$ for a given transcript $i$. For example, the value of $r_{L,10}$ is obtained by averaging the ten largest correlations associated with transcript $i$, i.e., $r_{i(1)} \ldots r_{i(10)}$. This measure is exactly proportional to the node connectivity measure $k$ described by Yip and Horvath (2007), except here the average of the first $m$ correlation order statistics is used to focus on the largest absolute correlations associated with individual transcripts. This modified approach is sensible for genome-wide data, since it emphasizes stronger correlations, which are the most meaningful from a biological perspective, and most influential in determining the structure of co-expression networks.

The “weak” or non-local connectivity patterns of CR-regulated genes were investigated by calculating the correlation percentiles of $|r_i|$ vectors. For a given transcript $i$, the $(100 × p)$th correlation percentile $r_{i,p}$ refers to the $(100 × p)$th percentile of the values contained in $|r_i|$. For example, a 90th correlation percentile of 0.60 indicates that, among all correlations associated with a given transcript, 90% are less than 0.60. This metric is conceptually identical to correlation order statistics, except the percentile index $p$ is more convenient for characterizing patterns associated with weaker correlations. For example, if $N$ transcripts are included on a given array, the $k$th correlation order statistic $r_{i(k)}$ is equivalent to the $[100 \times (1 − k/N)]$th correlation percentile.

### 2.4 Invertebrate models of CR

Since fundamental mechanisms underlying the mammalian CR response are thought to be shared across taxonomic groups (Sinclair, 2005; Spindler and Dhahbi, 2007), network properties of genes regulated by CR in the *Drosophila* and *C. elegans* model systems were also investigated. For this purpose, two independent expression profiling datasets were generated for *Drosophila* (*Drosophila* Series A and B), as well as for *C. elegans* (*C. elegans* Series A and B). The two *Drosophila* datasets were generated from 186 and 173 CEL files, respectively, and were based upon the Affymetrix DrosGenome1 oligonucleotide array (with 14010 probesets). The two *C. elegans* datasets were generated from 180 and 184 CEL files, respectively, and were based upon the Affymetrix Celegens Genome array (with 22625 probesets). CEL files were downloaded from either Gene Expression Omnibus or ArrayExpress databases (Barrett et al., 2007; Parkinson et al., 2007), and arrays assigned to each dataset were normalized as a group using robust multichip average (Irizarry et al., 2003). Following normalization, replicate arrays corresponding to the same experimental treatment were averaged to obtain expression values for each of 60 conditions per dataset (approximately 3 replicates per condition on average). Conditions in each dataset represent whole-genome expression patterns associated with a variety of cell types, developmental
stages, mutations, as well as various pathogen or stress treatments (described in Supplemental Data File 1).

Genes regulated by CR in Drosophila were identified based upon microarray data from Pletcher et al. (2002). This study used Affymetrix roDROMEGa arrays to compare expression levels between flies maintained on control versus CR media (15% versus 5% yeast/glucose) at various points between 7 and 74 days of age. A total of 2188 transcripts were found to be significantly altered by CR early in the lifespan, where the effect of CR subsequently remained stable later in the lifespan (P < 0.05) (Pletcher et al., 2002). This was a large number of transcripts, so filtering was performed based upon the fold-change between CR and control treatments at ages for which data was available for both treatments (7, 28 and 47 days). Filtering led to the identification of 14 transcripts for which expression was downregulated by 40% or more, along with 18 transcripts for which expression was upregulated by 3-fold or greater. Based on these fold-change criteria, the total number of CR-regulated transcripts (32) identified was comparable to the number of mouse CR-regulated genes listed in Table 1 (28).

CR-regulated genes from the C. elegans model were identified from the microarray analysis of Szewczyk et al. (2006). This study used a custom array to evaluate gene expression changes between broods maintained on a CR (or axenic) growth medium versus a standard bacterial diet. The analysis of Szewczyk et al. (2006) identified 26 genes downregulated by CR, along with 22 genes upregulated by CR. Genes within these sets were matched with 26 and 32 transcripts, respectively, on the Affymetrix Celegans Genome array, and the network properties of these transcripts were examined.

In the nematode C. elegans, starvation at the L1 stage can induce the development of a growth-arrested dauer stage that is both stress-resistant and long-lived (Cassada and Russel, 1975; Riddle and Albert, 1997). Dauer-formation is not considered a model for the mammalian CR response, and mechanisms regulating dauer formation appear to be independent of those regulating response to standard C. elegans CR treatments (applied during the adult stage) (Lakowski and Hekimi 1998). However, since dauer formation represents a metabolic response to resource limitation that is accompanied by increased longevity, it was suspected that dauer-regulated genes could share some network properties with CR-regulated genes. Dauer-regulated genes were identified based upon consistencies from two previous microarray experiments (Jones et al., 2001; Wang and Kim, 2003). In particular, 28 dauer-specific genes were identified by intersecting the 358 dauer-enriched genes found by Jones et al. (2001) and 540 dauer-enriched genes found by Wang and Kim (2003). Likewise, 40 nondauer-specific genes were identified by intersecting the 533 nondauer-specific genes found by Jones et al. (2001) and 1984 genes induced during dauer recovery in the study of Wang and Kim (2003).

3. Results

3.1 Co-expression of CR-regulated genes

CR-regulated genes were clustered based upon expression patterns across four independent datasets (Figure 1 and Supplemental Data File 2). This revealed transcriptional relationships among CR-regulated genes, most of which were significantly stronger than expected by chance (Figure 1). In addition, for each CR-regulated gene, “nearest neighbors” within the genomewide transcriptional network were found by identifying other genes with similar expression patterns (Supplemental Data File 3). Based on these analyses, three main co-expression patterns were consistently discernable among CR-regulated genes. These patterns involved genes associated with immune response (H2-Aa, Cd74, Tgtp and Ifi27), collagen formation (Serphinh1, Col3a1 and Col1a1) and metallothionein production (Mt1 and Mt2).
H2-Aa and Cd74 exhibited nearly identical expression patterns across all four datasets (Figure 2). Both genes consistently clustered together when each dataset was evaluated individually (Supplemental Data File 2), as well as when datasets were combined to generate a single clustering solution (Figure 1). Tgtp and Ifi27, which also have immune response functions, always clustered in the same branch as H2-Aa and Cd74 (Supplemental Data File 2), and H2-Aa and Cd74 were neighbors of Tgtp and Ifi27 on a genome-wide basis (Supplemental Data File 3). Five overrepresented transcription factor binding sites were present within the promoter regions of H2-Aa, Cd74, Tgtp and Ifi27 (Table 1), which were associated with several trans-acting elements, including paired box gene 8 (Pax8), zinc finger protein 278 (Zfp278), zinc finger protein 98 (Zfp98), lymphoid enhancer binding factor 1 (Lef1) and the nuclear factor Y transcription factors (Nfya, Nfyb and Nfyc).

Serpinh1, Col3a1 and Col1a1 always clustered in the same branch, and with respect to three of four datasets, formed an exclusive branch (Figure 1 and Supplemental Data File 2). Four overrepresented transcription factor binding sites were present within the promoter regions of Serpinh1, Col3a1 and Col1a1 (Table 1). These binding sites were associated with numerous trans-acting elements, several of which are known to interact with collagen gene promoters, such as the myeloblastosis oncogene (Myb) (Nishina et al., 1989) and nuclear factor κB (NF-κB) (Rippe et al., 1999). Serpinh1, Col3a1 and Col1a1 were co-expressed with several other genes downregulated by CR in four of the tissue types considered by Swindell (2008), including Col5a2 (collagen, type V, alpha 2), Col1a2 (collagen, type I, alpha 2), Fbn1 (fibrillin 1) and Sparc (secreted acidic cysteine rich glycoprotein) (Supplemental Data File 3).

Strong co-expression of the metallothioneins Mt1 and Mt2 was unequivocally supported by all analyses (Figure 1 and Supplemental Data File 2). On a genome-wide level, Mt1 and Mt2 were reciprocal nearest neighbors, but only weakly associated with all other genes (Supplemental Data File 3). Twelve overrepresented transcription factor binding sites were present within the promoter regions of Mt1 and Mt2 (Table 1). These binding sites were associated with many trans-acting elements, but only few of these had previously been linked to Mt1 and Mt2 regulation, including signal transducer and activator of transcription factors 1 and 3 (Stat1 and Stat3) (Lee et al., 1999) and metal activated transcription factor 1 (Mtf1) (Heuchel et al., 1994).

There was limited evidence to suggest that any CR-regulated gene was co-expressed with genes that are often viewed as mediators of the mammalian CR response, including Sirt1, mTOR or Igf-I (Figure 1 and Supplemental Data File 2) (Sinclair, 2005). For example, Sirt1 clustered with Rbm3sim in the composite analysis (|r| ≈ 0.65 on average) (Figure 1), but closer inspection revealed inconsistency among independent datasets (Supplemental Data File 2). A more robust association was found between Ifitm3 and mTOR, which clustered in the same branch with respect to three of four datasets (Supplemental Data File 2) while forming an exclusive branch in the composite analysis (Figure 1). Interestingly, both Ifitm3 and mTOR have a role in cell growth and proliferation. Signaling through the mTOR pathway stimulates cell growth, while in contrast, over-expression of Ifitm3 has been found to inhibit proliferation in transfected HeLa and H1299 cells (Ropolo et al., 2004).

Promoter analysis suggested that, potentially, a small number of regulatory factors could modulate transcription of many CR-regulated genes. Among all 28 genes, three transcription factor binding sites were significantly overrepresented in the 500 bp region upstream of transcriptional start sites (CisView identifiers: TF_NFKB, TF_MYB, TF_SOX9) (P < 0.02) (Sharov et al., 2006). The overrepresentation of TF_NFKB, a binding site for nuclear factor κB (NF-κB), is especially noteworthy, since promoters of ten CR-downregulated genes contained this element (all CR-downregulated genes except Gng11 and Ifi27), as well as promoters of five CR-upregulated genes (Pnpla2, Pim3, Nfkbia, Ddit4, Sult1a1). The observed
upregulation of \textit{Nfkbia} under CR therefore provides an attractive regulatory mechanism, since \textit{Nfkbia} encodes I\(\kappa\)B\(\alpha\), an inhibitor of NF-\(\kappa\)B (Arenzana-Seisdedos et al., 1995).

### 3.2 Genes regulated by CR in mouse have unique network properties

Since the CR response appears to be evolutionarily conserved, it was hypothesized that CR-regulated genes would be highly connected elements within co-expression networks. Surprisingly, however, only CR-downregulated genes exhibited strong local connectivity, while in contrast, CR-upregulated genes exhibited weak local connectivity (Figure 3). This pattern was robust and emerged regardless of which of the four expression profiling datasets was used to quantify connectivity patterns (Supplemental Data File 4).

The network properties of CR-regulated genes were unusual with respect to the rest of the genome. For example, with respect to the Treatment series, the average \(r_{10}\) among the \(n = 12\) CR-downregulated genes was 0.906 (Figure 3). A simulation analysis was carried out in which 12 transcripts were randomly sampled from those represented on the Affymetrix Mouse 430 2.0 array, and for each of 10,000 samples, the average value of \(r_{10}\) was calculated. Only 0.07% of random samples yielded an average \(r_{10}\) value equal or larger than the 0.906 value associated with CR-downregulated genes. Likewise, with respect to the Treatment series, the average \(r_{10}\) value of the \(n = 16\) CR-upregulated genes was 0.797. Using the same simulation procedure with random samples of 16 transcripts, an equal or lower average \(r_{10}\) value was obtained in only 0.16% of simulation trials. These trends were consistent across each of the four datasets evaluated, and were statistically significant with respect to three of four datasets (Supplemental Data File 4).

Figure 3 demonstrates unique connectivity properties of CR-regulated genes based upon the largest 10 correlations associated with each gene (i.e., \(r_{10}\)), but similar conclusions are obtained based upon any order 1…50 of the local connectivity index (i.e., \(r_{1}…r_{50}\)). This is apparent from Figure 4, which plots average correlation order statistics \(r_{(1)}…r_{(50)}\) among the \(n = 12\) CR-downregulated and \(n = 16\) CR-upregulated transcripts, and provides a comparison to the average values of \(r_{(1)}…r_{(50)}\) obtained by randomly sampling \(n\) transcripts from the Affymetrix Mouse 430 2.0 array. For any value of \(k = 1…50\), the average \(k\)th correlation order statistic among CR-downregulated genes is significantly greater than expected by random sampling, while the average \(k\)th correlation order statistic among CR-upregulated genes is significantly less than expected by random sampling (Figure 4). These trends were consistent across all four datasets examined (Supplemental Data File 5). One exception was with respect to the Tissue series, for which correlation order statistics of \(k > 10\) were not, on average, high among CR-downregulated genes as compared to the rest of the genome (Supplemental Data File 5).

The contrasting properties of CR-downregulated and CR-upregulated genes were also apparent based upon other local connectivity metrics. The clustering coefficient, for example, provides a measure of “cliquishness”, or the tendency of a node’s neighbors to be connected with each other (Watts and Strogatz, 1998; Ravasz et al., 2002; Dong and Horvath, 2007). With respect to each of the four datasets examined, clustering coefficients associated with CR-downregulated genes were significantly greater than coefficients associated with CR-upregulated genes (\(P < 0.015\)) (Figure 5). Nearly identical results were obtained based upon the “line density” connectivity measure described by Dong and Horvath (2007) (data not shown). Overall, the network consequences of these local connectivity differences are illustrated by Figure 6, which shows that on a genome-wide scale, CR-downregulated genes are more highly connected than CR-upregulated genes.

The strong connectivity of CR-downregulated genes was limited to local network regions, and was not evident based upon analysis of correlation percentiles (Figure 7). In contrast, the weak connectivity of CR-upregulated genes did extend beyond local network regions. For any value
of $p = 0.99 \ldots 0.50$, the average correlation percentile $r_p$ among the $n = 16$ CR-upregulated genes was significantly less than expected based upon random sampling of $n$ transcripts from the Affymetrix Mouse 430 2.0 array (Figure 7). This pattern was consistent across each of the four expression profiling datasets analyzed (Supplemental Data File 6).

These results demonstrate that, as a group, genes regulated by CR across five or more tissue types in mice have unique network properties. It was expected that similar, albeit weaker, patterns would be found among genes regulated by CR across only four tissue types. The analysis of Swindell (2008) identified 79 such genes (31 CR-downregulated and 48 CR-upregulated). These genes include additional CR-downregulated heat shock proteins ($Hspa1a$, $Hspa1b$, $Hsp90b1$, $Hsp90aa1$, $Hspa8$) and collagen genes ($Col1a2$ and $Col5a2$), as well as a CR-upregulated forkhead box transcription factor ($Foxo3a$) (Supplemental Data File 7). Contrary to the expectation, the 30 CR-downregulated genes were not associated with unusually high local or non-local connectivity (Supplemental Data File 7). However, consistent with the expectation, the 48 CR-upregulated genes were associated with unusually low connectivity based upon all metrics evaluated. This pattern was consistent across each of the four expression profiling datasets (Supplemental Data File 7).

### 3.3 Other “aging genes” lack unique network properties

A number of “aging genes” have been identified for which mutations significantly increase mouse lifespan and delay the onset of age-related disease ($Prop1$, $Pit1$, $Ghr$, $Ghrhr$, $Irs1$, $Irs2$, $PappA$, $Clk1$, $Shc1$, $Igf1r$, $Kl$, $Adcy5$, $Surf1$, $Insr$, $Ucp2$, $Gpx4$) (Brown-Borg et al., 1996; Kopchick and Laron, 1999; Migliaccio et al., 1999; Flurkey et al., 2001; Bluher et al., 2003; Holzenberger et al., 2003; Kurosu et al., 2005; Liu et al., 2005; Conti et al., 2006; Conover and Bale, 2007; D’Agnello et al., 2007; Ran et al., 2007; Taguchi et al., 2007; Yan et al., 2007; Selman et al., 2008). The extent to which these genes interact with CR-regulated pathways is unclear (Miller et al., 2002; Tsuchiya et al., 2004; Bonkowski et al., 2006; Swindell, 2007). However, some evidence suggests that such aging genes are likely to be highly connected within biological networks (Ferrarini et al., 2005; Curtis et al., 2006). Surprisingly, however, while some genes were highly connected (e.g., $Insr$, $Ghrhr$, $Prop1$), and others were weakly connected ($Irs2$, $Surf1$, $Igf1r$) (Supplemental Data File 8), there was no evidence that, on average, aging genes were disproportionately associated with either dense or sparse network regions (Supplemental Data File 8). Genome-wide co-expression neighbors of all aging genes cited above are identified in Supplemental Data File 8.

### 3.4 Invertebrate models of CR

CR-regulated genes from invertebrates were also associated with unusual network properties that, in some cases, resembled those found for mouse (Figure 8). However, connectivity properties of CR-regulated genes were only partially consistent between mouse and invertebrate models (Figure 8).

In *Drosophila*, genes upregulated by CR were weakly connected based upon both expression profiling datasets examined (Figure 8B), and with respect to one expression profiling dataset (*Drosophila Series A*), CR-downregulated genes were strongly connected (Figure 8A). However, for CR-downregulated genes, a second expression profiling dataset (*Drosophila Series B*) revealed no evidence of unusually strong network connectivity (Supplemental Data File 9).

In *C. elegans*, CR-downregulated genes tended to be strongly connected, but this trend was significant only for one of two expression profiling datasets examined (Figure 8C and Supplemental Data File 9). For both datasets, however, CR-upregulated genes were strongly
connected (Figure 8D), which contrasts markedly with CR-upregulated genes in the mouse (and *Drosophila*) model systems.

Surprisingly, *C. elegans* dauer-regulated genes had network properties that corresponded very well those of CR-regulated genes from the mouse model system. Dauer-associated genes exhibited local connectivity metrics, that on average, were much stronger than expected within sets of randomly sampled transcripts (Figure 8E), while genes associated with the non-dauer state exhibited weak connectivity (Figure 8F). These conclusions were supported by separate analyses conducted with each of two expression profiling datasets (Supplemental Data File 9).

4. Discussion

The effects of caloric restriction (CR) can be studied from the standpoint of individual genes and proteins, as well as from a systems-level perspective that includes the collective interactions of many genes and proteins. These two approaches are much more likely to be complementary rather than conflicting, and ultimately, both will be necessary to completely understand the effects of CR on aging and lifespan. Recently, candidate genes associated with CR have emerged from microarray analyses (Spindler and Dhabl, 2007; Swindell, 2008), but have been understood only as a list of genes, which does not provide insight into connectivity properties and relationships between genes. This study analyzed CR-regulated genes from a network perspective, which revealed transcriptional relationships among such genes, clues related to regulatory mechanisms, and surprising connectivity properties. While genes regulated by CR are positioned throughout the transcriptional network, genes related to immunity, collagen and metallothioneins represent center points of the CR response in mice. Genes downregulated by CR exhibit a significant tendency to be well-connected and are thus associated with dense regions of the transcriptional network. The opposite tendency, however, was found among genes upregulated by CR, which are less connected and positioned in sparse network regions. This latter property may have an evolutionary basis, since CR-upregulated genes from the *Drosophila* model system are also weakly connected. Taken together, these data suggest that CR-regulated genes occupy unique positions within co-expression networks, and raise new questions that can be pursued using both vertebrate and invertebrate models of aging.

The network properties of CR-regulated genes may have implications regarding their evolutionary importance, as well as their involvement in central processes that underlie survival (Jeong et al., 2001; Carlson et al., 2006; Oldham et al., 2006). These implications, however, depend upon the interrelationships between genes, connectivity, network architecture and the phenotype, which are topics of ongoing investigation (Siegal et al., 2007). Given a scale-free topology, network theory suggests that highly connected genes will have an essential role in network operation and organism survival (Albert et al., 2000). This notion predicts that CR-downregulated genes will be involved in fundamental biological processes, which is consistent with the idea that CR influences processes related to growth and energy distribution (Holliday, 1989; Hart and Turturro, 1998; Kirkwood and Shanley, 2005). In contrast, however, network theory predicts that weakly connected CR-upregulated genes will be less essential for survival. One approach for evaluating this prediction is to examine phenotypes of knock-out mice that lack CR-upregulated genes, or to examine the essentiality of orthologues in the *S. cerevisiae*, *C. elegans* and *D. melanogaster* model systems (Ekman et al., 2006; Lee et al., 2008). An important consideration is whether connectivity properties documented in this study are reflected at other levels of the interactome, such as metabolic and protein-protein interaction networks. The Biomolecular Interaction Network Database (BIND), for example, now contains over 7,000 records corresponding to experimentally documented interactions between mouse proteins (Gilbert, 2005). Searches performed using this database did in fact identify more interaction records for protein products of CR-downregulated genes in comparison to CR-
upregulated genes (average of 4.25 (± 3.09) versus 3.20 (± 1.88) records per protein). Currently, however, this comparison is limited, since database records do not completely characterize all protein-protein interactions, and results were skewed towards heavily investigated proteins for which many interactions have been documented (e.g., H2-Aa, Hsp110 and Cdkn1a).

Inhibition of NF-κB signaling by CR may diminish activation of the innate immune system with increased age, which potentially, could contribute to the effects of CR on lifespan (Chung et al., 2006; Morgan et al., 2007). This study found that a motif associated with NF-κB was significantly overrepresented within promoter regions of CR-regulated genes, and was present in the promoters of 10 of 12 CR-downregulated genes. Moreover, the NF-κB inhibitor Nfkbia was among genes upregulated by CR in five or more tissue types (Swindell, 2008). The upregulation of Nfkbia by CR, in several tissue types (liver, heart, muscle, hypothalamus and lung), could therefore be an important mechanism by which CR inhibits the NF-κB pathway, with consequent effects on lifespan. An alternative possibility is that CR activates Sirt1, which then deacetylases components of the NF-κB complex (Yeung et al., 2004b; Ghosh et al., 2007; Salminen et al., 2008; but see Barger et al., 2008). Previous studies have indeed found that NF-κB signaling is inhibited by caloric restriction in both rodents and humans (Phillips and Leeuwenburgh, 2005; Dirks and Leeuwenburgh, 2006; Weiss et al., 2006; Ugochukwu and Figgers, 2007), and that inhibition of NF-κB blocks the accumulation of beta-amyloid plaques in Alzheimer’s mice (Paris et al., 2007), as well as neoplastic transformation of cells during the initial stages of cancer (Karin, 2006). Such effects result from blocking NF-κB activation, which is a characteristic feature of normal aging, as demonstrated by a recent genomic study (Adler et al., 2007). Adler et al. (2007) used microarray data to identify gene expression modules based upon shared motifs in the promoter regions of individual genes, and for each module, the correlation between module expression and age was evaluated in a wide variety of human and mouse tissues. Following this unbiased genomic approach, Adler et al. (2007) found that a module defined by an NF-κB motif was most strongly activated with increasing age, and further, showed experimentally that inhibition of NF-κB reversed age-associated gene expression patterns in murine skin.

Groups of CR-regulated genes that are co-expressed provide valuable points of reference for investigating factors mediating the cellular response to CR. For example, three co-expressed collagen genes are downregulated by CR (Col1a1, Col3a1 and Serpinh1), and in certain cell types, each gene is regulated by transforming growth factor-β (TGF-β) (Yamamura et al., 1998; Verrecchia et al., 2001). In mammals, it has been proposed that CR decreases TGF-β expression, and that this effect is due to lowered oxidative stress, along with reduced 4-hydroxy-2,3-nonenal (HNE) and activator protein 1 activity (AP-1) (Chiarotto et al., 2006). The generality of such an effect, however, is not clear, since Meyer et al. (2006) found that CR reduced TGF-β expression in human heart, but negative or opposite effects have been found with respect to rat liver and testis (Ando et al., 2002; Jung et al., 2004). In some tissues, therefore, declines in collagen gene expression under CR could involve other regulatory factors besides TGF-β. For instance, the Col1a1, Col3a1 and Serpinh1 promoters share four transcription factor binding sites, one of which is associated with NF-κB (see Table 1). In fact, NF-κB signaling has been found to influence collagen expression and accumulation (Novitskiy et al., 2004; Tang et al., 2004; Theiss et al., 2005; Nieto, 2007), and cross-talk between the NF-κB and TGF-β signaling pathways has been identified (Azuma et al., 1999; Hong et al., 2007). Describing how these factors influence collagen gene expression under CR will be an important task, since both TGF-β and NF-κB have a role in tumor development (de Caestecker et al., 2000; Sarkar and Li, 2008), and downregulation of Col1a1, Col3a1 and Serpinh1 by CR may contribute to increased lifespan by combating fibrosis in a variety of tissues (Wynn, 2008).
The CR-upregulated metallothionein genes are most often studied in relation to stress-response, but *Mt1* and *Mt2* appear to have broad physiological roles even in the absence of stress (Beattie et al., 1998; Waelput et al., 2000; Penkowa, 2006). It has been suggested, for example, that metallothioneins influence energy regulation, since mice lacking *Mt1* and *Mt2* have higher food intake, elevated plasma leptin concentrations and exhibit obesity following sexual maturation (Beattie et al., 1998). Additionally, like *Col1a1*, *Col3a1* and *Serpinh1*, the metallothioneins appear to affect collagen accumulation and development of fibrosis. Mice lacking *Mt1* and *Mt2* are in fact viewed as disease models for fibrosis of some organ systems (Jones, 2005), since these mice have increased susceptibility to liver fibrosis caused by carbon tetrachloride treatment (Elsheriff et al., 2004), and also have elevated collagen accumulation in the heart following alcohol consumption (Wang et al., 2005). The expression of *Mt1* and *Mt2* in the absence of stress is controlled by multiple factors (Haq et al., 2003), but in relation to CR, an interaction between *Mt* gene expression and leptin signaling is a noteworthy possibility, since leptin treatment of leptin-deficient *ob/ob* mice increases *Mt1* and *Mt2* expression in liver and decreases *Mt2* expression in the small intestine (Waelput et al., 2000).

The role of metallothioneins in response to CR may be evolutionarily conserved, since orthologues *mtl-1* and *mtl-2* are also upregulated by CR in *C. elegans* (Szewczyk et al., 2006).

CR decreases the expression of several genes related to immunity, most notably *Cd74* and *H2-Aa*, which reflects the complex and not fully understood influence of CR on immune function. Some evidence suggests that CR diets threaten energetically costly immunological memory mechanisms necessary for acquired immunity, leading to a diminished ability of CR-treated rodents to generate antibodies following secondary infection (Christadoss et al., 1984; Venkatraman et al., 1994; Fernandes et al., 1997; Dong et al., 2000; Martin et al., 2007). Since both *Cd74* and *H2-Aa* are involved in the antigen processing and presentation functioning of the adaptive immune system, decreased expression of these genes by CR could contribute to such an immune response compromise (Kontgen, 1993; Viville et al., 1993). This viewpoint contrasts, however, with studies showing positive effects of CR on immune function indicators, such as T-cell maintenance (Messaoudi et al., 2006), lymphocyte proliferation (Nikolich-Zugich and Messaoudi, 2005) and inducible natural killer cell activity (Pahlavani, 2005). All things considered, the most important investigations are those focused on survival and disease susceptibility of CR-treated animals following exposure to pathogens. In this regard, CR decreased survival of aged mice following primary exposure to influenza (Gardner, 2005), and a recent study showed that CR-treated mice were more susceptible to infection by an intestinal parasite (Kristan, 2007). A possible explanation of these results is that the low body weight of CR-treated mice fails to provide energetic resources to meet the metabolic costs of infection and immune response (Ritz and Gardner, 2006).

The application of genomic approaches provides an opportunity to characterize the effects of CR on the mammalian transcriptome in a comprehensive and unbiased manner, both at the level of individual genes as well as whole-genome co-expression networks. The present study has focused on genes commonly regulated by CR across multiple tissue types, with the expectation that such genes will be most important for mediating the lifespan increase associated with CR diets (Swindell, 2008). This approach is appealing because (i) it is the simplest of possibilities, (ii) lifespan necessarily depends upon continued functioning of many organ and tissue types with age, and (iii) evolutionary conservation of the CR response implicates fundamental mechanisms shared across mammalian cell types. Nevertheless, further data will be necessary to bear out the validity of this approach, and more generally, to distinguish effects of CR that increase lifespan from incidental effects of CR that do not influence survivorship. If genes commonly regulated by CR do underlie extended lifespan, as conjectured, then targeting key regulatory proteins associated with such genes will be an important CR-mimetic strategy. For instance, inhibition of NF-κB activity due to elevated...
Nfkbia expression across many tissue types appears to be a common response to CR. At least one proposed CR-mimetic compound, resveratrol, mimics this effect (Manna et al., 2000; Pendurthi et al., 2002; Iyori et al., 2008), and similar responses might also be attained by treatment with other NF-κB inhibitors now undergoing development (Gilmore and Herscovitch, 2006).

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.

**Acknowledgments**

This work was supported by NIA training grant AG000114 and the University of Michigan Department of Pathology. Two anonymous reviewers provided helpful comments on this manuscript. The author thanks laboratories for providing microarray data to the Gene Expression Omnibus and ArrayExpress databases.

**References**


Bonkowski MS, Rocha JS, Masternak MM, Al Regaiey KA, Bartke A. Targeted disruption of growth hormone receptor interferes with the beneficial actions of caloric restriction. Proc Natl Acad Sci USA 2006;103:7901–7905. [PubMed: 16682650]


Kristan DM. Chronic calorie restriction increases susceptibility of laboratory mice (Mus musculus) to a primary intestinal parasite infection. Aging Cell 2007;6:817–825. [PubMed: 17973970]


Mech Ageing Dev. Author manuscript; available in PMC 2009 May 29.


Figure 1.
Co-expression relationships among CR-regulated genes. The 28 CR-regulated genes were clustered with respect to expression pattern similarity. Labels of genes downregulated by CR are displayed in green, while labels of genes upregulated by CR are displayed in red. The horizontal axis indicates the average distance at which two clusters (or individual genes) were joined. A distance matrix was computed with respect to each of four independent datasets (Tissue, Developmental, Treatment and Mutation series), where distance between expression profiles was based upon absolute value of the Pearson correlation coefficient. A final distance matrix was generated by averaging distance values across the four datasets, and average linkage hierarchical clustering was then used to generate the tree shown in the figure. The dashed
vertical line provides indication of the lowest distance expected to arise by chance between any two genes (given that \( n = 31 \) genes are included in the tree). In particular, for 95% of 10,000 randomized simulation trials, with \( n = 31 \) genes sampled at random from the Affymetrix 430 2.0 array, the lowest distance was larger than the value represented by the vertical dashed line in Figure 1. Supplemental data file 2 shows clustering outcomes for the same genes based upon expression patterns in each of the four datasets individually.
Figure 2.
Expression patterns of H2-Aa and Cd74. Plots show expression patterns across the (A) Tissue, (B) Developmental, (C) Treatment and (D) Mutation series. In each plot, the vertical axis represents expression level, with the 60 conditions in each dataset arranged along the horizontal axis. The solid green line represents the expression pattern of H2-Aa while the dotted green line corresponds to that of Cd74. Expression patterns of each gene were centered to have the same average expression across the 60 conditions in each dataset. This average expression level is represented by the dashed horizontal line in each plot.
Figure 3.
Local connectivity index (Treatment series). CR-regulated genes are listed along the vertical axis, with green labels corresponding to CR-downregulated genes and red labels corresponding to CR-upregulated genes. The horizontal axis corresponds to the local connectivity index. This index was determined for each gene by first calculating correlations between that gene and all other transcripts on the Affymetrix Mouse Genome 430 2.0 array, and then averaging the largest ten of these correlations. A large index value is expected for genes located in dense regions of the transcriptional network. The green vertical line indicates the average local connectivity index among the $n = 12$ CR-downregulated genes, while the red vertical line indicates the average local connectivity index among the $n = 16$ CR-upregulated genes. Vertical black lines
indicate a conservative 95% confidence interval for the average index among \( n \) genes sampled at random (based upon simulation). The interval is conservative because it corresponds to \( \{ \min (L_1, L_2), \max (U_1, U_2) \} \), where random samples of size \( n = 12 \) yielded an average index in the range of \( \{ L_1, U_1 \} \) in 95% of trials, and random samples of size \( n = 16 \) yielded an average index in the range of \( \{ L_2, U_2 \} \) in 95% of trials. Results shown in the figure were generated from the Treatment series, but similar results were obtained based on the Tissue, Developmental and Mutation series (see Supplemental Data File 4).
Figure 4.
Local connectivity of CR-regulated genes (Treatment series). In each plot, the vertical axis represents the average $k$th correlation order statistic (among $n$ genes) and the horizontal axis corresponds to $k$. For a given gene, the $k$th correlation order statistic equals the $k$th largest correlation between that gene and all other transcripts represented on the Affymetrix Mouse Genome 430 2.0 array. Highly connected genes should have larger correlation order statistics for a given value of $k$, where the decay with increasing $k$ indicates the decline in connection strength to increasingly distant network neighbors. In part (A), green points represent the average $k$th correlation order statistic among the $n = 12$ CR-downregulated genes. In part (B), red points represent the average $k$th correlation order statistic among the $n = 16$ CR-upregulated genes. In both (A) and (B), dashed black lines outline a 95% confidence region for the average $k$th correlation order statistic among $n$ genes sampled at random (based upon simulation). In 95% of 10,000 simulation trials, random samples of $n$ genes from the Affymetrix Mouse 430 2.0 array yielded an average $k$th correlation order statistic within the dashed lines. Results shown were generated from the Treatment series, but similar results were obtained based on the Tissue, Developmental and Mutation series (see Supplemental Data File 5).
Figure 5.
Clustering coefficient. Green bars represent average clustering coefficients for CR-downregulated genes, while red bars represent average clustering coefficients for CR-upregulated genes. For each CR-regulated gene, the 20 genome-wide nearest neighbor transcripts were identified, and the clustering coefficient associated with each CR-regulated gene was calculated. The cluster coefficient ranges from 0 to 1 and measures the local density of expression networks surrounding a given node (transcript) (see Equation 5 from Dong and Horvath (2007)). Values near 0 indicate weak connectivity among neighbors at an individual node, while values near 1 indicate strong connectivity among neighbors at an individual node.
Nearly identical results were obtained using the “line density” measure of local connectivity (see Equation 2 from Dong and Horvath (2007)).
Figure 6.
Sub-graphs associated with CR-regulated genes (Treatment Series). In part (A), a subgraph consisting of CR-downregulated genes and their genome-wide neighbors is shown. Nodes representing CR-downregulated genes are colored green. In part (B), a sub-graph consisting of CR-upregulated genes and their genome-wide neighbors is shown. Nodes representing CR-upregulated genes are colored red. For both (A) and (B), connections between nodes are based upon a connectivity threshold of $|r|>0.90$. 
Figure 7.
Non-local connectivity of CR-regulated genes (Treatment series). In each plot, the vertical axis represents the average $p$th correlation percentile (among $n$ genes) and the horizontal axis corresponds to $p$. For a given gene, correlations with all other transcripts on the Affymetrix Mouse Genome 430 2.0 array were determined, and the $p$th correlation percentile is the correlation value that is larger than $p$ percent of all of these correlations. Genes that have strong “distant” connections (beyond immediately neighboring transcripts) should have large correlation percentiles, where the decay with increasing $p$ indicates the decline in connection strength to increasingly distant network neighbors. In part (A), green points represent the average $p$th correlation percentile among the $n = 12$ CR-downregulated genes. In part (B), red points represent the average $p$th correlation percentile among the $n = 16$ CR-upregulated genes. In both (A) and (B), dashed black lines outline a 95% confidence region for the average $p$th correlation percentile among $n$ genes sampled at random (based upon simulation). In 95% of 10,000 simulation trials, random samples of $n$ genes from the Affymetrix Mouse 430 2.0 array yielded an average $p$th correlation percentile within the dashed lines. Results shown were generated from the Treatment series, but similar results were obtained based on the Tissue, Developmental and Mutation series (see Supplemental Data File 6).
Figure 8.
Local connectivity of CR-regulated and dauer-regulated genes from invertebrate models. Parts (A) and (B) show mean correlation order statistics for Drosophila CR-downregulated and CR-upregulated genes, respectively. CR-regulated genes were chosen based upon microarray data from Pletcher et al. (2001). Parts (C) and (D) show mean correlation order statistics for CR-downregulated and CR-upregulated genes, respectively. CR-regulated genes were chosen based upon microarray data from Szewcyzk et al. (2006). Parts (E) and (F) show mean correlation order statistics for C. elegans nondauer-specific and dauer-specific, respectively. Dauer-regulated genes were chosen based upon microarray data from Jones et al. (2001) and Wang and Kim (2003). In each plot, the vertical axis represents the average $k$th correlation order statistic (among $n$ transcripts) and the horizontal axis corresponds to $k$. Dashed black lines outline a 95% confidence region for the average $k$th correlation order statistic among $n$ transcripts sampled at random (based upon simulation). Results shown are based upon the Drosophila and C. elegans series A datasets. Results based upon a second, independent dataset for each model (series B) are shown in Supplemental Data File 9.
## Table 1

Genes regulated by CR across five or more tissue types. Genes were identified based upon comparative analysis of thirteen microarray experiments that collectively evaluated the effects of CR in ten different mouse tissues (liver, heart, muscle, white adipose tissue, hypothalamus, hippocampus, colon, lung, kidney and cochlea) (Swindell, 2008). Each CR-regulated gene was identified based upon experimental data linked to a specific Affymetrix Mouse Genome 430 2.0 probeset. The identifier associated with that probeset is listed in the table below along with a gene symbol and description generated from the NetAffyx database in January 2008 (Liu et al., 2003). Groups of co-expressed genes are identified in footnotes, with significantly overrepresented transcription factor binding sites listed below the table. Overrepresented binding sites were identified based upon sequence analysis of the 500 bp upstream promoter region using the CisView database (Sharov et al., 2006). Descriptions of each binding site are available online (see: http://lgsun.grc.nia.nih.gov/cisview/).

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Affymetrix Identifier</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR-Downregulated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H2-AaA</td>
<td>1452431_s_at</td>
<td>Histocompatibility 2, class II antigen A, alpha</td>
</tr>
<tr>
<td>Cd74A</td>
<td>1425519_a_at</td>
<td>CD74 antigen (invariant polypeptide of major histocompatibility complex, class II antigen-associated). Receptor for macrophage migration inhibitory factor (MIF) (Leng et al., 2003).</td>
</tr>
<tr>
<td>TgtpA</td>
<td>1449009_at</td>
<td>T-cell specific GTPase</td>
</tr>
<tr>
<td>Ifi27A</td>
<td>1426278_at</td>
<td>Interferon, alpha-inducible protein 27</td>
</tr>
<tr>
<td>Col3alB</td>
<td>1427884_at</td>
<td>Procollagen, type III, alpha 1</td>
</tr>
<tr>
<td>Col1a1B</td>
<td>1423669_at</td>
<td>Procollagen, type I, alpha 1</td>
</tr>
<tr>
<td>Serpinh1B</td>
<td>1450843_a_at</td>
<td>Serine (or cysteine) peptidase inhibitor, clade H, member 1</td>
</tr>
<tr>
<td>Ccl9</td>
<td>1417936_at</td>
<td>Chemokine (C-C motif) ligand 9</td>
</tr>
<tr>
<td>Ifitm3</td>
<td>1423754_at</td>
<td>Interferon induced transmembrane protein 3</td>
</tr>
<tr>
<td>Gng11</td>
<td>1448942_at</td>
<td>Guanine nucleotide binding protein (G protein), gamma 11</td>
</tr>
<tr>
<td>C1qb</td>
<td>1417063_at</td>
<td>Complement component 1, q subcomponent, beta polypeptide</td>
</tr>
<tr>
<td>Hapl10</td>
<td>1425993_a_at</td>
<td>Heat shock protein 110</td>
</tr>
<tr>
<td>CR-Upregulated</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mt1C</td>
<td>1422557_s_at</td>
<td>Metallothionein 1</td>
</tr>
<tr>
<td>Mt2C</td>
<td>1428942_at</td>
<td>Metallothionein 2</td>
</tr>
<tr>
<td>Per1</td>
<td>1449851_at</td>
<td>Period homolog 1 (Drosophila)</td>
</tr>
<tr>
<td>Per2</td>
<td>1417602_at</td>
<td>Period homolog 2 (Drosophila)</td>
</tr>
<tr>
<td>Ppara</td>
<td>1449051_at</td>
<td>Peroxisome proliferator activated receptor alpha</td>
</tr>
<tr>
<td>Peci</td>
<td>1431012_a_at</td>
<td>Peroxisomal delta3, delta2-enoyl-Coenzyme A isomerase</td>
</tr>
<tr>
<td>Sub1a1</td>
<td>1427345_a_at</td>
<td>Sulfotransferase family 1A, phenol-prefering, member 1</td>
</tr>
<tr>
<td>MGI-1924575</td>
<td>1436316_at</td>
<td>RIKEN cDNA 9430029L20 gene</td>
</tr>
<tr>
<td>Daft4</td>
<td>1428306_at</td>
<td>DNA-damage-inducible transcript 4</td>
</tr>
<tr>
<td>Njkb1</td>
<td>1449731_s_at</td>
<td>Nuclear factor of kappa light chain gene enhancer in B-cells inhibitor, alpha. Inhibitor of NF-κB DNA binding (Arenzana-Seisdedos et al., 1995)</td>
</tr>
<tr>
<td>Fkbp5</td>
<td>1416125_at</td>
<td>FK506 binding protein 5</td>
</tr>
<tr>
<td>Roms3smD</td>
<td>1422660_at</td>
<td>Similar to RNA binding motif protein 3</td>
</tr>
<tr>
<td>Pim3</td>
<td>1451069_at</td>
<td>Proviral integration site 3</td>
</tr>
<tr>
<td>Tspan4</td>
<td>1448276_at</td>
<td>Tetraspanin 4</td>
</tr>
<tr>
<td>Cdhn1a</td>
<td>1421679_a_at</td>
<td>Cyclin-dependent kinase inhibitor 1A (P21)</td>
</tr>
</tbody>
</table>

 Mech Ageing Dev. Author manuscript; available in PMC 2009 May 29.
Gene Symbol | Affymetrix Identifier | Description
--- | --- | ---
Pnpla2 | 1428143_a_at | Patatin-like phospholipase domain containing 2

A. Immunity Group. Note that H2-Aa is an updated annotation and was previously H2-Ea in Swindell (2008). Promoter analysis identified five significantly overrepresented transcription factor binding sites that were present in the promoter regions of H2-Aa, Cd74, Tgtp and Ifit27 (ADD_PAX8, TF_NFY, TF_MAZR, TF_MZF, MIT_013LEF) (P < 0.05).

B. Collagen Group. For most analyses, since two transcripts were associated with Col3a1, expression values associated with both transcripts were averaged. Promoter analysis identified five significantly overrepresented transcription factor binding sites that were present in the promoter regions of genes in this module (TF_MAF, TF_MYB, TF_MEIS, TF_NFKB) (P < 0.004).

C. Metallothioneins. Promoter analysis identified twelve significantly overrepresented transcription factor binding sites that were present in the promoter regions of genes in this module (TF_MIF, TF_STAT, TF_ZIC, TF_HEN1, TF_HNF4, TF_SREBP, TF_OLF1, ADD_MTF1A, ADD_MTF1B, MIT_051TATA, TF_MYB, ADD_PAX8) (P < 0.04).

D. Rbm3sim is an unofficial gene symbol for LOC100043257. Updated from Swindell (2008).