Integrative and Comparative Biology

Integrative and Comparative Biology, volume 54, number 3, pp. 452–462
doi:10.1093/icb/icu048

Society for Integrative and Comparative Biology

THEME AND VARIATIONS: HETEROTHERMY IN MAMMALS

Comparative Genomics of Mammalian Hibernators Using Gene Networks

José Luis Villanueva-Cañas,* Sheena L. Faherty,† Anne D. Yoder† and M. Mar Albà[1,*,‡]

*Evolutionary Genomics Group, Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Universitat Pompeu Fabra (UPF), Barcelona 08003, Spain; †Department of Biology, Duke University, Durham, NC 27708, USA; ‡Catalan Institution for Research and Advanced Studies (ICREA), Barcelona, Spain

[1]E-mail: malba@imim.es

Synopsis In recent years, the study of the molecular processes involved in mammalian hibernation has shifted from investigating a few carefully selected candidate genes to large-scale analysis of differential gene expression. The availability of high-throughput data provides an unprecedented opportunity to ask whether phylogenetically distant species show similar mechanisms of genetic control, and how these relate to particular genes and pathways involved in the hibernation phenotype. In order to address these questions, we compare 11 datasets of differentially expressed (DE) genes from two ground squirrel species, one bat species, and the American black bear, as well as a list of genes extracted from the literature that previously have been correlated with the drastic physiological changes associated with hibernation. We identify several genes that are DE in different species, indicating either ancestral adaptations or evolutionary convergence. When we use a network approach to expand the original datasets of DE genes to large gene networks using available interactome data, a higher agreement between datasets is achieved. This indicates that the same key pathways are important for activating and maintaining the hibernation phenotype. Functional-term-enrichment analysis identifies several important metabolic and mitochondrial processes that are critical for hibernation, such as fatty acid beta-oxidation and mitochondrial transport. We do not detect any enrichment of positive selection signatures in the coding sequences of genes from the networks of hibernation-associated genes, supporting the hypothesis that the genetic processes shaping the hibernation phenotype are driven primarily by changes in gene regulation.

Introduction

Hibernation is an intricate physiological response that some mammalian species employ and thereby evade energetic demands posed by their environment. It is defined as a seasonal period of heterothermy characterized by periods of torpor (i.e., a controlled reduction in metabolic rate and body temperature) and metabolically active periods of rewarming known as interbout arousals (IBAs). Hibernation functions to reduce the expenditure of energy during the cold winter months when resources of favored food and water are scarce, and the demand for generating metabolic heat is high (Boyer and Barnes 1999; Carey et al. 2003; Geiser 2004).

Hibernators, as a group, encompass a vast amount of phylogenetic diversity (Carey et al. 2003). Within the Class Mammalia, they are found in all three of the deepest mammalian lineages: Monotremata, Marsupialia, and Eutheria (placental mammals). Within the Eutheria, they are found in different groups, including the main divisions Afrotheria and Laurasiatheria. Mammalian species that hibernate often are closely related to species that do not display hibernating behaviors. One possibility is that the common ancestor of mammals was a non-hibernator and that this trait has been gained independently due to similar pressures acting upon disparate lineages. The literature is replete with examples of adaptive phenotypic convergence acting at the morphological and physiological level, including reductions in anatomical features (e.g., tails, eyes, and ears) shared among subterranean mammals (Nevo 1979) or echolocation in bats and dolphins (Liu et al. 2010). Another possibility is that the common ancestor of Mammalia exhibited seasonal heterothermy and the capability to hibernate has subsequently been lost multiple times along several lineages. Regardless of the evolutionary origin of the hibernation phenotype,

Advanced Access publication May 31, 2014

its widespread distribution suggests that the genes and pathways modulating this phenotype may be similar among different mammals.

A fundamental area of inquiry in the study of hibernation physiology relates to defining the principal molecular mechanisms involved in hibernation. Comparative analysis of genomic data from several hibernating and non-hibernating species can provide important clues to specific genetic controls associated with the hibernation phenotype. By measuring changes in level of gene expression in hibernating versus non-hibernating conditions we can identify genes that are likely to be involved in the hibernation process. Srere et al. (1992) were the first to document differential gene expression to be a major player in the molecular regulation of hibernation. This study elegantly demonstrated that concentrations both of protein and mRNA expression of the molecule α2-macroglobulin increased during bouts of torpor, relative to the active state, in Richardson's and Columbian ground squirrels (*Spermophilus richardsonii* and *Spermophilus columbianus*, respectively). The authors hypothesized that the hibernation phenotype is manifest via a small number of regulatory changes in existing mammalian genes as opposed to the de novo creation of hibernation-specific biochemical functions. Currently, there is a large body of literature describing genes that are differentially expressed (DE) during hibernation in different species (Srere et al. 1992; Wilson et al. 1992; Bauer et al. 2001; Buck et al. 2002; Eddy and Storey 2004; Eddy et al. 2005, 2006; Williams et al. 2005; Yan et al. 2008; Fedorov et al. 2009).

Since the groundbreaking study of Srere et al. (1992), much subsequent work has illuminated the means by which differential gene expression is involved in the switch from a summer-active phenotype to a winter-hibernation phenotype. With the advent of high-throughput genome-wide approaches, such as microarrays and RNA-seq, it has become possible to investigate multiple genes simultaneously. Large-scale transcriptomic investigations have been conducted in some of the most well-studied hibernating species, including squirrels (*Spermophilus* spp.) (Williams et al. 2005; Fedorov et al. 2011; Hampton et al. 2011; Schwartz et al. 2013; Xu et al. 2013), American black bears (*Ursus americanus*) (Fedorov et al. 2011), and little brown bats (*Myotis lucifugus*) (Seim et al. 2013). Shotgun proteomics has also been employed to gain insights into the hibernation process in Arctic ground squirrels (Shao et al. 2010). Typically, these studies explore how genes are selectively expressed during hibernation, by sampling a variety of tissues (e.g., brain, white adipose tissue [WAT], brown adipose tissue [BAT], skeletal muscle, heart, and liver) at different physiological states of the hibernation circannual cycle (i.e., active, torpor, and IBA).

In addition to changes in levels of gene expression, some of the genes in hibernation-related pathways may have experienced adaptive substitutions in hibernating species, for example to increase the encoded protein's baseline activity when temperature or pH conditions vary. One example is the Leptin gene in bats. It has been proposed that this enzyme has experienced positive selection in heterothermic bats, thereby increasing its lipolytic activity (Yuan et al. 2011). Studies of the black bear suggest that, in general, hibernation-related genes are slowly evolving (Zhao et al. 2010). Clearly, more studies are required to evaluate the impact of positive selection on the evolution of these genes in different species.

Through the recent explosion of cost-effective means to sequence whole genomes, we now have the capability to perform comparative genomics analyses to reveal how changes at the sequence or expression level drive the manifestation of adaptive phenotypes. The number of hibernation-related gene candidates has increased in recent years by high-throughput investigations aimed at identifying hibernation DE genes, but no attempt has been made yet to compare and integrate the data from the different studies. We attempt to address this deficit here. Protein–protein, pathway, and genetic interaction information have been successfully used in a wide number of studies which, in turn, have revealed new candidate proteins correlated with diverse human diseases (Vidal et al. 2011). Utilizing this novel approach, we built interaction networks from the gene expression experiments to obtain not only a visualization of the biological processes occurring during hibernation, but also to gain knowledge of how genes may be connected at the amino-acid level. We find that, whereas only a few DE genes are shared in experiments performed on different hibernating species, the number of shared genes increases when we focus on interaction networks and biological processes. We also explore the impact of positive selection in the coding sequences of hibernation-related genes by tests based on maximum-likelihood. We do not detect a significant increase in the number of positively selected genes in hibernating species when compared with non-hibernators, suggesting that species-specific changes in amino acids have not played a predominant role.

Materials and methods

Orthologous gene dataset

Protein-coding genes from 10 different mammalian species, including three seasonally heterothermic species (*Ictidomys tridecemlineatus/Spermophilus tridecemlineatus*, *M. lucifugus*, and *Microcebus murinus*) and seven additional non-hibernating species (*Bos taurus*, *Cavia porcellus*, *Homo sapiens*, *Mus musculus*, *Otolemur garnettii*, *Oryctolagus cuniculus*, and *Pteropus vampyrus*), were downloaded from Ensembl (Flicek et al. 2012). The species were selected on the basis of having a well-annotated genome, and a broad phylogenetic representation. Using orthological data from Ensembl release 70 (January 2013), we obtained 8233 gene families containing one-to-one orthologs from all the species. The list of genes is available in the Supplementary Material.

Multiple sequence alignments

We built multiple sequence alignments for each mammalian gene family. In order to minimize the number of false positives, care was taken to select the best possible method and strategy for alignment. For that purpose, we used PRANK+F (Löytynoja and Goldman 2008) with the -codon option. For the input tree, we generated a neighbor-joining tree produced from a distance matrix with information about divergence time from the Timetree website (Hedges et al. 2006) and Agnarsson et al. (2011) for the bat distances and using fossil information from the closest groups (Agnarsson et al. 2011). We also employed a novel algorithm for selecting similar isoforms developed by our group to improve the quality of the alignments and decrease the number of false positives in positive selection analysis (Villanueva-Cañas et al. 2013). The protein alignments were converted to nucleotide sequence alignments with pal2nal (Suyama et al. 2006).

Positive selection test

We performed branch-specific tests of positive selection using the branch-site test (Zhang et al. 2005), implemented in the Phylogenetic Analysis by Maximum Likelihood (PAML) software package (Yang 2007). This test compares the null model in which codon-based dN/dS for all branches can only be less than or equal to 1, with the alternative model in which the labeled foreground branch may include codons evolving at $dN/dS > 1$. The likelihood ratio test was used to compare the two models. It was calculated as $2 \times (L1 - L0)$, where L1 is the maximum-likelihood value of the alternative hypothesis and L0 the maximum-likelihood value of the null hypothesis. A chi-square distribution with one degree of freedom was used to calculate the *P*-value and a multiple testing correction was also applied afterward (Benjamini and Hochberg 1995). We tested for positive selection in all terminal branches of the tree relating to the 10 species included in the orthologous gene dataset (see above).

Hibernation-related gene datasets

We obtained lists of hibernation DE genes from five previously published studies comprising 11 sets of sampled tissues subjected to high-throughput experiments (Table 2). These datasets included genes that were either significantly up-regulated or down-regulated during hibernation when compared with the active state. Each study was based on one of the following species: arctic ground squirrel (*Spermophilus parryii*) (Yan et al. 2008), 13-lined ground squirrel (*S. tridecemlineatus*) (Hampton et al. 2011; Schwartz et al. 2013), and American black bear (*U. americanus*) (Fedorov et al. 2011). Seven different tissues were interrogated in total, of which liver, skeletal muscle, hypothalamus, and heart were analyzed in more than one study, and thus were amenable for cross-comparison between the aforementioned species. Two studies were based on microarray platforms (Yan et al. 2008; Fedorov et al. 2011) and two utilized next-generation transcriptome sequencing (RNA-seq) (Hampton et al. 2011; Schwartz et al. 2013). The lists of genes were cross-validated using HUGO Gene Nomenclature Committee identifiers.

We also gathered a manually curated hibernation-related set of 64 proteins that had been previously described in the literature to be important for the hibernation response (Carey et al. 2003; Andrews 2007; Melvin and Andrews 2009). They corresponded to 91 Ensembl gene identifiers, after accounting for genes encoding different protein subunits and gene duplicates. All datasets are available from the Supplementary Material.

Networks construction and analysis

We constructed gene networks using the aforementioned sets of hibernation-related gene datasets. We obtained gene interaction and pathway connections using the Gene Multiple Association Network Integration Algorithm (GeneMANIA) plugin (Montojo et al. 2010) for Cytoscape 2.8 (Smoot et al. 2011). The GeneMANIA algorithm is divided into two parts; the first part is an algorithm, based on linear regression, which integrates different

genomic or proteomic sources of data and calculates a single, composite, functional association network. The second part consists of a label propagation algorithm that predicts gene function given this composite network. We used the latest release of the human interactome (version from August 2, 2012). The networks we obtained are available in the Supplementary Material. We obtained Gene Ontology (GO) terms from all genes using Cytoscape. We subsequently identified significantly overrepresented GO terms ($P<0.0001$) in each network.

We also created two additional networks that combined genes from different sources. The first one was constructed using a manually curated list of hibernation-related genes extracted from the literature (see the "Hibernation-related gene datasets" section). The second one combined hibernation DE genes from three tissues (WAT, skeletal muscle, and heart) from RNA-seq experiments performed using 13-lined ground squirrels (*S. tridecemlineatus*) (Hampton et al. 2011).

We calculated the clustering coefficient for each gene and network. In undirected networks, such as the ones we created, the clustering coefficient C_n of a node n is defined as $C_n=2e_n/(k_n(k_n-1))$, where k_n is the number of neighbors of n and e_n is the number of connected pairs between all neighbors of n (Watts and Strogatz 1998; Barabási and Oltvai 2004). The clustering coefficient is a ratio N/M, where N is the number of edges between the neighbors of n and M is the maximum number of edges that could possibly exist between the neighbors of n. Therefore, C_n of a node is always a number between 0 and 1. The network C_n is the average of the clustering coefficients for all nodes in the network. Ten additional networks were created using randomly selected protein-coding genes in order to calculate the random average clustering coefficient and explore its attributes.

To compare the genes from different networks we built a contingency table with the data and tested for significance using a one-tailed Fisher test. The table contained the proportion of shared genes in the two datasets with respect to the largest one and we specifically tested for an increase in the proportion of shared genes after expansion of the network.

We extracted the GO terms associated with the genes in the different networks using Cytoscape and determined the level of GO term-enrichment agreement between different networks. Due to the hierarchical nature of the GO terms, there was high redundancy in GO term annotation. We kept a GO term representative for each GO term group, generating a non-redundant list of terms characteristic of the genes that were DE in each tissue (Table 4).

Results and discussion

Distribution of mammalian hibernating species with completely sequenced genomes

Through the efforts of initiatives like the Mammalian Genome Project (Lindblad-Toh et al. 2011) and the 10 K genome Project (Genome 10 K Community of Scientists 2009), coupled with the cost decrease of next-generation sequencing and the improvements in both sequence-quality and software, characterizing new genomes has dramatically increased in the past few years. Most of the newly sequenced vertebrate species are members of the Class Mammalia, and a growing list of them are seasonally heterothermic; that is, capable of hibernation or torpor (Carey et al. 2003; Gummer 2005). Ensembl contains gene catalogs for four seasonally heterothermic species with low-coverage ~2× assemblies, the European hedgehog (*Erinaceus europaeus*), lesser hedgehog tenrec (*Echinops telfairi*), gray mouse lemur (*M. murinus*), and Ord's kangaroo rat (*Dipodomys ordii*), and two species with higher coverage ~7× assemblies, the 13-lined ground squirrel (*I. tridecemlineatus*) and the little brown bat (*M. lucifugus*). In addition, the genomes of several additional bat species that hibernate, David's Myotis (*Myotis davidii*), Brandt's bat (*Myotis brandtii*), Parnell's mustached bat (*Pteronotus parnellii*), and greater horseshoe bat (*Rhinolophus ferrumequinum*), have recently been subjected to shotgun sequencing (Parker et al. 2013; Seim et al. 2013; Zhang et al. 2013). Genomic sequencing data for the polar bear have also been made available (Li et al. 2011).

Figure 1 depicts a tree with the phylogenetic position of seasonally heterothermic species with a completely sequenced genome, including representative species from very distinct groups. These include the Afrotheria (lesser hedgehog tenrec), Insectivora (European hedgehog), Chiroptera (little brown bat, David's Myotis, Brandt's bat), Rodentia (13-lined ground squirrel and Ord's kangaroo rat), and Strepsirrhini (gray mouse lemur). In addition, transcriptome sequencing data (RNA-seq) for hibernating animals is available for 13-lined ground squirrels (Hampton et al. 2011; Schwartz et al. 2013) and Brandt's bats (Seim et al. 2013). Exploiting this genomic data can provide key insights about the evolution of hibernation in modern mammals by identifying shared, as well as divergent, molecular patterns.

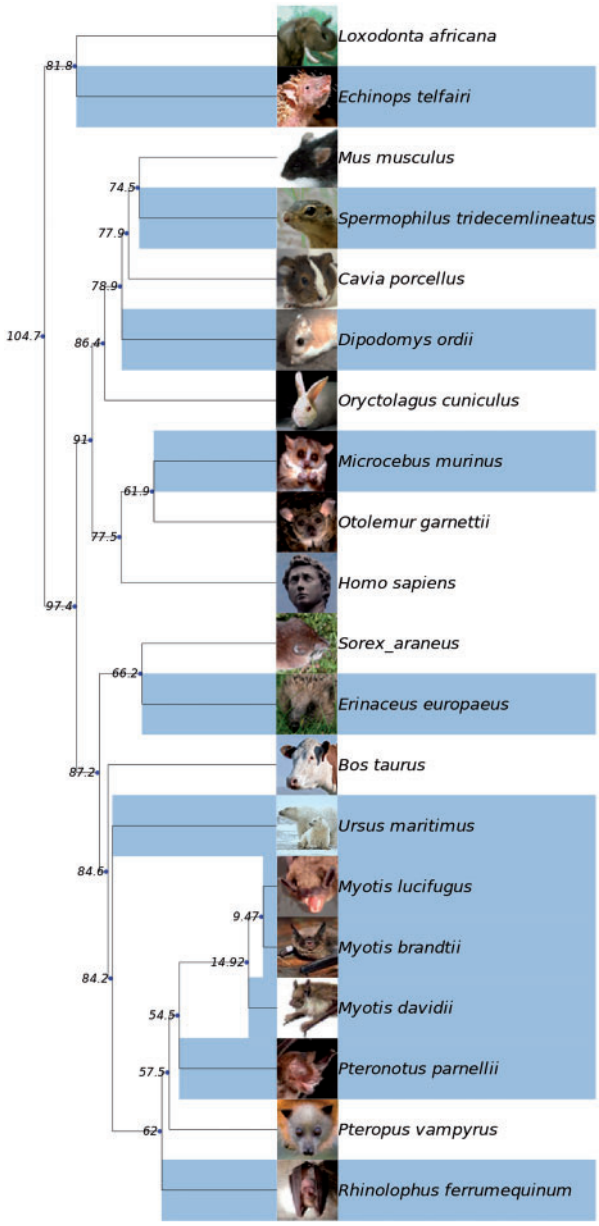

Fig. 1 (a) Phylogenetic tree showing the position of heterothermic species with a sequenced genome. Nodes show time of last common ancestor in millions of years. Legend: human (*Homo sapiens*), mouse (*Mus musculus*), cow (*Bos taurus*), guinea pig (*Cavia porcellus*), 13-lined ground squirrels (*Spermophilus tridecemlineatus*), gray mouse lemur (*Microcebus murinus*), bushbaby (*Otolemur garnettii*), rabbit (*Oryctolagus cuniculus*), megabat (*Pteropus vampyrus*), little brown bat (*Myotis lucifugus*), common shrew (*Sorex araneus*), elephant (*Loxodonta africana*), European hedgehog (*Erinaceus europaeus*), Ord's kangaroo rat (*Dipodomys ordii*), lesser hedgehog tenrec (*Echinops telfairi*), polar bear (*Ursus maritimus*), Brand's bat (*Myotis brandtii*), greater horseshoe bat (*Rhinolophus ferrumequinum*), Parnell's mustached bat (*Pteronotus parnellii*), David's myotis (*Myotis davidii*).

Adaptive substitutions in hibernation-related genes

To investigate whether hibernation-related genes are enriched in positive selection signatures, we selected species with high genome coverage and well-defined gene orthological relationships from Ensembl (Flicek et al. 2012). The set of species included three seasonally heterothermic species, little brown bat, 13-lined ground squirrel, and gray mouse lemur, as well as representative mammals that experience neither torpor nor hibernation, as controls (Table 1). We built protein multiple-sequence alignments for 8233 gene families comprising one to one orthologous genes from each of the species. Subsequently, we tested for positive selection in species-specific branches, using the branch-site test in the PAML package (Yang 2007).

In the complete dataset of gene families, the percentage of species-specific, positively selected genes after correction for false discovery was between 0.5% and 5% ($P<0.05$) (Table 1). This variation may in part be attributed to the different quality of the genomes considered. We found that species with higher-coverage genomes, such as human, mouse, and cow, usually showed a smaller fraction of positively selected genes when compared with species with lower-coverage genomes.

We next examined positive selection signatures in two sets of hibernation-related genes. The first set was derived from a list of manually curated hibernation-related genes, initially containing 64 different proteins, but later expanded to 203 by the inclusion of interacting partners in the gene network (see the "Materials and Methods" section). In this set, the number of positively selected genes was small and there were no significant differences between seasonally heterothermic species and non-hibernating species. The second set, using hibernation-related genes from the 13-lined ground squirrel, was considerably larger and included 748 genes. These genes were obtained in two steps: first, sets of DE genes were obtained from studies using RNA-seq (Hampton et al. 2011; Schwartz et al. 2013). Second, the original lists were expanded to include the interacting partners. In ground squirrels, 42 genes showed signatures of positive selection. This corresponds to 13.64% of the positively selected genes in the total dataset, a percentage in line with non-hibernating species (Table 1). No significant increase was detected either for the little brown bat or the gray mouse lemur. In addition, all positively selected candidate genes that were shared by several seasonally heterothermic species were also candidates for positive selection in at least one of the other species, indicating that these genes may have

Table 1 Number of positively selected (PS) genes in different mammals

Species	Total PS (8233)	PS curated dataset (204)	PS Hampton (748)	(PS Hampton/ total PS) ×100
Bos taurus	181	1	11	6.08
Cavia porcellus	300	4	30	10
Homo sapiens	45	3	4	8.89
Microcebus murinus	137	4	6	4.38
Mus musculus	70	1	4	5.71
Myotis lucifugus	418	2	26	6.22
Oryctolagus cuniculus	381	3	21	5.51
Otolemur garnettii	326	3	21	6.44
Pteropus vampyrus	60	0	8	13.33
Spermophilus tridicemlineatus	308	3	20	6.49

Notes: The results are based on the branch-site test; hibernating species are highlighted in gray. The number of genes tested is in parenthesis. PS curated: number of positive selection gene candidates in the manually curated list. PS Hampton: number of positive selection gene candidates in the list of genes derived from the Hampton et al. study (Table 2, All*).

Table 2 Description of hibernation-related gene datasets

Reference	Species	Technique	Tissue	DE genes	sampled genes	%DE genes	MC	Expanded
Hampton et al. (2011) Supplementary Table S2	*Spermophilus tridecemlineatus*	RNA-seq	WAT	133	14,351	0.92	5	290
			Skeletal muscle	93	12,496	0.74	3	331
			Heart	154	13,637	1.13	9	352
			All*	381	—	—		748
Fedorov et al. (2011) Supplementary Table S1	*Ursus americanus*	Microarray	Liver	319	5092	6.26	1	506
			Heart	245	7359	3.33	0	439
Yan et al (2008) Supplementary Table S2	*Spermophilus parryii*	Illumina 96-sample array matrix	BAT	51	1407	3.62	3	250
			Liver	120	1407	8.53	3	313
			Hypothalamus	29	1407	2.06	3	229
			Heart	51	1407	3.62	4	246
Schwartz et al. (2013) Supplementary Table S1	*Spermophilus tridecemlineatus*	RNA-seq	Hypothalamus	1063	—	—	7	1562
Seim et al. (2013). Supplementary data 4	*Myotis brandtii*	RNA-seq	Liver	138	—	—	2	264

Notes: DE: differentially expressed genes. MC: the number of genes also present in the manually curated dataset. Expanded: number of genes after expansion of the network. *Considering genes expressed in any tissue.

experienced positive selection; however, these selection forces may not be linked to the hibernation phenotype.

Comparison of DE genes from diverse studies

During hibernation some genes experience changes in their expression levels when compared with the active state. High-throughput experiments such as microarrays and RNA-seq data can provide a comprehensive view of the shifts in gene expression that occur in different tissues and at different times. We gathered 12 datasets of hibernation-related DE genes from the literature, corresponding to four different species: arctic ground squirrel (*S. parryii*) (Yan et al. 2008), 13-lined ground squirrel (*S. tridecemlineatus*) (Hampton et al. 2011; Schwartz et al. 2013), American black bear (*U. americanus*) (Fedorov et al. 2011), and Brandt's bat (*M. brandtii*) (Seim et al. 2013) (Table 2). We selected these studies because the experimental setting was easily comparable and they represented a broad spectrum of hibernating species.

Given that several tissues were analyzed in more than one experiment, we were able to do a cross-comparison from the perspectives both of candidate genes and of expanded networks. Our first goal was to compare the different datasets to determine whether the DE genes identified were similar across the different studies. We found that a small fraction of the genes (<9%) were shared between any two

Table 3 Shared genes between DE gene datasets

Tissue	Datasets	Shared genes	Shared genes networks	*P*-value
Heart	Yan_2_96 & Fedorov	4 (1.63)	127 (28.93)	0.0001
	Yan_2_96 & Hampton	7 (4.52)	43 (12.21)	0.004
	Fedorov & Hampton	8 (3.27)	36 (8.2)	0.0071
Liver	Fedorov & Yan_2_96	27 (8.46)	55 (10.87)	0.1574
	Fedorov & Seim	7 (2.19)	106 (20.94)	0.0001
	Yan_2_96 & Seim	4 (2.89)	17 (5.43)	0.1727
Hypothalamus	Schwartz & Yan_2_96	1 (0.094)	42 (2.69)	0.0001

Notes: See Table 1 for more details on the dataset. *P*-value was calculated using a one-tailed Fisher test comparing proportions of shared genes before and after expansion of the gene network. The percentage of shared genes is in parenthesis.

studies (Table 3). Considering the specific experimental design employed in each study, this can be considered a lower bound of the number of common genes in two species. We found remarkable examples of genes showing significantly increased or decreased tissue-specific expression in several species during hibernation relative to the active state. Two genes were identified in the hearts of the two ground squirrel species, as well as in the American black bear: *ACADVL* and *GAPDH*. The first gene, *ACADVL*, encodes the enzyme that catalyzes the first step of the mitochondrial fatty-acid beta-oxidation pathway. Fatty-acid beta-oxidation becomes increasingly important during bouts of torpor associated with hibernation. During the long hibernation season, fat-storing hibernators, such as ground squirrel species and American black bears, cease feeding and instead rely on stored lipids as their primary source of fuel (Bert et al. 1999). Therefore, it is not unsurprising that *ACADVL* shows increased levels of expression during hibernation relative to the active state in the species investigated. *GAPDH* encodes glyceraldehyde 3-phosphate dehydrogenase, which catalyzes the conversion of glyceraldehyde 3-phosphate to D-glycerate 1,3-bisphosphate, as part of the breakdown of glucose. Additionally, Soukri et al. found that *GAPDH* was significantly down-regulated in hibernating jerboas, compared with control animals that were euthermic (Soukri et al. 1996). In light of the fuel switch to lipids that occurs as an animal is engaged in hibernation, this result might be expected. As the breakdown of glucose is suppressed during hibernation, this contributes to profound reductions in the metabolic rate that an animal experiences while engaged in torpor. Another gene, α2-macroglobulin (*α2M*), was found to be up-regulated during hibernation in livers from arctic ground squirrels and American black bears. These results are consistent with the study of Srere et al. (1992), which found similar expression profiles in mRNA levels that were isolated from the livers of Richardson's and Columbian ground squirrels during hibernation, relative to the active state. *α2M*, a broad-spectrum protease inhibitor, is hypothesized to function in reducing blood-clotting times (Srere et al. 1992). This is especially important for survival during hibernation as circulation is nearly arrested due to a reduction in heart rate, thereby increasing the potential risk for fatal blood clots to form during venous stasis.

We also compared DE genes from different experiments with a manually curated list of 64 genes that have been correlated with hibernation in gene-focused studies (see the "Materials and Methods" section). Some DE genes were found in the manually curated list (Table 2). Examples of such genes are *GAPDH*, found in all datasets on hearts, and *ACSL1* detected in both samples of squirrel hearts. *ACSL1* converts free long-chain fatty acids into fatty acyl-CoA esters, and thereby plays a key role in the biosynthesis of lipids and in the degradation of fatty acids. We also noted that the DE sets included many genes that have not been previously documented as hibernation-related and which are interesting candidates for further exploration.

Hibernation-related networks

It is well established that genes do not work in isolation but instead function in an orchestrated way as part of gene networks. Networks have been used in a wide variety of disciplines in biology to capture different kinds of relations, such as trophic relationships in ecological fields, or protein or gene interactions in molecular biology and genetics. In networks, nodes represent entities (e.g., genes, proteins, and species) and the interactions are represented as edges. There are a high number of studies that use protein–protein interactions to unravel new candidates for targets in diseases (Vidal et al. 2011). Therefore, our approach represents a novel way to

utilize a network-level approach to better elucidate the molecular mechanisms regulating complex adaptive phenotypes, such as hibernation.

To develop a better understanding of the molecular processes guiding the complex physiological changes associated with hibernation, we expanded each set of DE genes to include closely related genes in the gene network. These are defined as genes from the same molecular pathway or that are linked by physical or genetic interactions (see the "Materials and Methods" section). Additionally, we created a network considering all tissues together for the RNA-seq experiments in 13-lined ground squirrels performed by Hampton et al. (2011). This network comprised 748 genes and a snapshot of it is displayed in Fig. 2. The beta-oxidation cluster was one of the most compact groupings in this network, stressing the importance of this metabolic pathway in the hibernation phenotype. Another important cluster of genes corresponded with muscular development, which supports the hypothesis that preceding hibernation, ground squirrels undergo a period of muscle-building that will be used in winter as a fuel source in addition to WAT (Boonstra et al. 2011).

The networks showed high overall connectivity and highlighted fundamental molecular processes in hibernation. For example, from the expanded network created from the gene sets of Hampton and colleagues, only 88 out of the 748 genes (12%) remained unconnected, suggesting that most of the DE genes are playing a role in the same biological processes. The clustering coefficients for the DE sets, which provide information on how well connected the neighborhood of the node is (see the "Material and Methods" section for more details), also support this notion. The clustering coefficient for the Hampton dataset, and the manually curated gene set, was 0.338 and 0.372, respectively. This is significantly higher than the clustering coefficients calculated from 10 similar sized networks created from randomly selected protein-coding genes (average clustering coefficient 0.0972, *t*-test $P = 0.00302$).

Remarkably, after expansion of the network, we found a higher proportion of genes in common

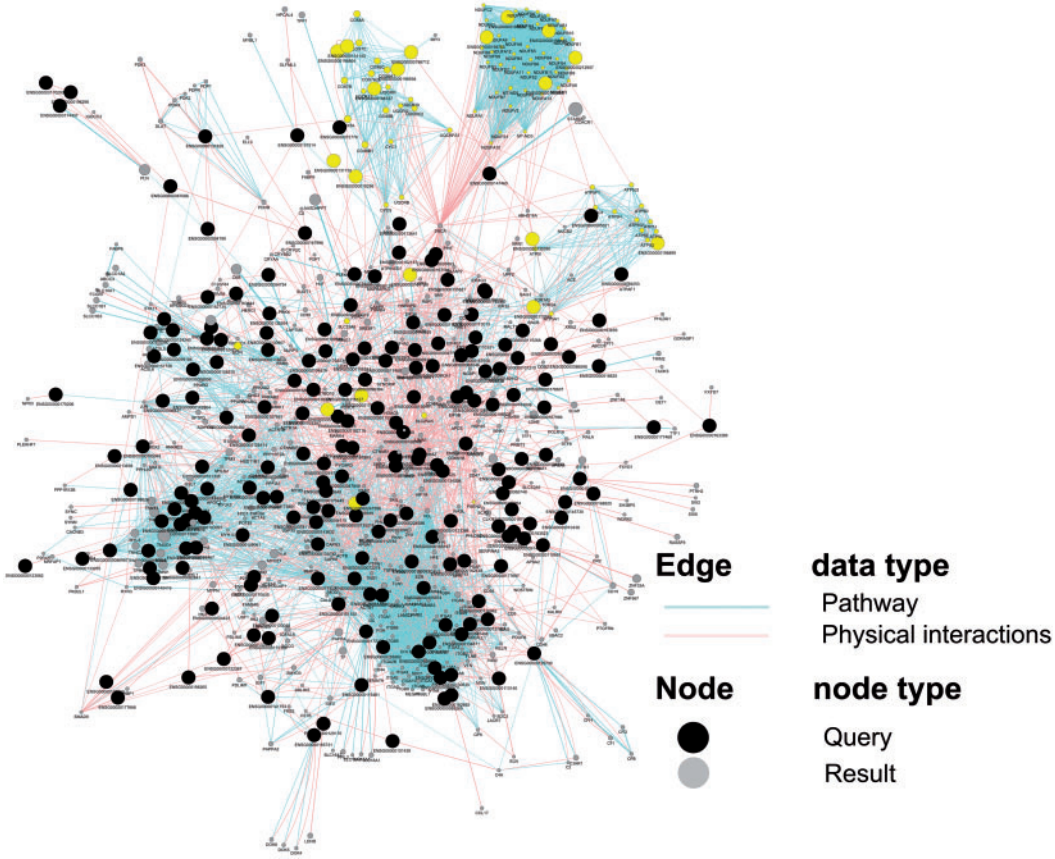

Fig. 2 Expanded network from the data of Hampton et al. (2011) in organic layout. Edges show type of interaction. Beta-oxidation process is highlighted in yellow. A navigable, high-resolution Cytoscape version of the network can be downloaded as Supplementary Material.

between DE datasets than when using the individual genes (Table 3). Regardless of the technology used, additional shared genes were identified after expanding the network. For example, fatty acid synthase (*FASN*), a molecule known to play a key role in lengthening the fatty acyl chain, was not present at first in any DE datasets but appeared in most datasets after expansion of the gene network.

The network's expansion also increased the fraction of genes shared between the DE datasets and the manually curated list. Before expansion the network generated using DE genes from different tissues contained 17 genes in common with the manually curated list (out of 64; 26.5%) and after expansion of both gene sets this number increased to 72 (out of 204; 35.5%).

Functionally related gene classes

We were interested in obtaining the primary molecular functions behind the agreement between networks. To that end, we obtained GO terms for each group of genes in the networks and identified those that were significantly over-represented. We employed GO terms for the corresponding human orthologous genes, which are better annotated than for other species. Table 4 provides a list of non-redundant GO terms that were significant ($P<10^{-4}$) in the gene networks for all the experiments from a given tissue.

Many of the shared GO terms are related to metabolic processes. The most frequently occurring GO terms were related to the regulation, mobilization, and oxidation of lipids, stressing the fact that lipid metabolism is a central physiological process that occurs during hibernation. Fat-storing hibernators engage in a period of aggressive deposition of fat in the period leading up to hibernation and subsequently switch from a carbohydrate-based to a lipid-based metabolism throughout the winter. Oxidation of fatty acid releases more than twice the energy per gram than does oxidation of carbohydrate; therefore, it appears that hibernating species have evolved the most efficient utilization of energy. In skeletal muscle, metabolism of pyruvate was overrepresented. Pyruvate acts as an intermediary in glycolysis and gluconeogenesis, or can also be converted into fatty acids through acetyl-CoA. In liver, gluconeogenesis and mitochondrial matrix, which contains enzymes involved in the tricarboxylic acid cycle, were also overrepresented. In BAT, we found overrepresented GO terms related to the generation of fatty acids and to aerobic respiration. In the hypothalamus, we found an enrichment of terms related to the cell cycle. This organ contains the suprachiasmatic nucleus (or nuclei) which controls the primary circadian clock in mammals (Albrecht and Eichele 2003). However, the DE genes were not significantly enriched in GO terms related to the circadian clock.

Table 4 Functional enrichment of hibernation-related genes

Tissue	GO term ID	Description
Heart	GO:0006818	Hydrogen transport
	GO:0015980	Energy derivation by oxidation of organic compounds
	GO:0034654	Nucleobase-containing compound biosynthetic process
	GO:0006839	Mitochondrial transport
	GO:0015672	Monovalent inorganic cation transport
	GO:0042776	Mitochondrial ATP synthesis coupled proton transport
Hypothalamus		
	GO:0000075	Cell cycle checkpoint
	GO:0000082	G1/S transition of mitotic cell cycle
	GO:0000216	M/G1 transition of mitotic cell cycle
	GO:0007346	Regulation of mitotic cell cycle
Liver		
	GO:0006956	Complement activation

Notes: GO terms are significantly overrepresented ($P<10^{-4}$) in all expanded gene datasets from a given tissue.

Conclusions

We have compared the results from different high-throughput, gene-expression studies in relation to the hibernation phenotype. These studies used different hibernating species and employed different methodology to measure changes in levels of gene expression during the circannual cycle by sampling tissues at multiple times throughout the year. Despite the heterogeneous nature of the data, we have succeeded in reconstructing gene networks using information from the interactome, revealing that the underlying molecular mechanisms are similar in the different species included in our analysis. Notably, genes in key pathways regulated during hibernation do not show increased signatures of positive selection in hibernating species when compared with non-hibernating ones, suggesting that gene regulation is the primary process driving the hibernation phenotype.

Next-generation sequencing technologies are becoming commonplace in investigations of how changes at the genetic level play a role in

evolutionary dynamics. The cost-effectiveness of the new techniques and the standardization of bioinformatic tools undoubtedly will transform how we address the study of complex physiological processes. Future studies using high-throughput transcriptome sequencing in a larger number of hibernating species will undeniably provide a more complete picture of how hibernation has evolved in modern day mammals and how differential gene expression functions in giving rise to complex phenotypes, such as hibernation.

Acknowledgments

The authors thank the members of the Evolutionary Genomics Group (evolutionarygenomics.imim.es, UPF-IMIM, Barcelona) for thoughtful discussions.

Funding

The Ministerio de Economía y Competitividad (FPI BES-2010-038494 to J.L.V.-C., Plan Nacional BIO2009-08160 and BFU2012-36820) and Fundació ICREA (to M.M.A.).

Supplementary data

Supplementary Data are available at evolutionarygenomics.imim.es (Publications, Datasets). This includes an excel supplementary file with the gene lists as well as gene network files to be visualized with Cytoscape 3, using the GeneMania plugin and human interactome data.

References

Agnarsson I, Zambrana-Torrelio CM, Flores-Saldana NP, May-Collado LJ. 2011. A time-calibrated species-level phylogeny of bats (Chiroptera, Mammalia). PLoS Curr 3:RRN1212.

Albrecht U, Eichele G. 2003. The mammalian circadian clock. Curr Opin Genet Dev 13:271–7.

Andrews MT. 2007. Advances in molecular biology of hibernation in mammals. BioEssays 29:431–40.

Barabási A-L, Oltvai ZN. 2004. Network biology: understanding the cell's functional organization. Nat Rev Genet 5:101–13.

Bauer VW, Squire TL, Lowe ME, Andrews MT, Nelson CJ, Otis JP, Carey HV, Yan J, Burman A, Nichols C, et al. 2001. Expression of a chimeric retroviral-lipase mRNA confers enhanced lipolysis in a hibernating mammal expression of a chimeric retroviral-lipase mRNA confers enhanced lipolysis in a hibernating mammal. Am J Physiol Regul Integr Comp Physiol 281:R1186–92.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 57:289–300.

Boonstra R, Bradley AJ, Delehanty B. 2011. Preparing for hibernation in ground squirrels: adrenal androgen production in summer linked to environmental severity in winter. Funct Ecol 25:1348–59.

Boyer BB, Barnes BBM. 1999. Molecular and metabolic aspects of mammalian hibernation. Bioscience 49:713–24.

Buck MJ, Squire TL, Andrews MT. 2002. Coordinate expression of the PDK4 gene: a means of regulating fuel selection in a hibernating mammal. Physiol Genomics 8:5–13.

Carey HV, Andrews MT, Martin SL. 2003. Mammalian hibernation: cellular and molecular responses to depressed metabolism and low temperature. Physiol Rev 83:1153–81.

Eddy SF, McNally JD, Storey KB. 2005. Up-regulation of a thioredoxin peroxidase-like protein, proliferation-associated gene, in hibernating bats. Arch Biochem Biophys 435:103–11.

Eddy SF, Morin PJR, Storey KB. 2006. Differential expression of selected mitochondrial genes in hibernating little brown bats, *Myotis lucifugus*. J Exp Zool A Comp Exp Biol 305:620–30.

Eddy SF, Storey KB. 2004. Up-regulation of fatty acid-binding proteins during hibernation in the little brown bat, *Myotis lucifugus*. Biochim Biophys Acta 1676:63–70.

Fedorov VB, Goropashnaya AV, Tøien O, Stewart NC, Chang C, Wang H, Yan J, Showe LC, Showe MK, Barnes BM. 2011. Modulation of gene expression in heart and liver of hibernating black bears (*Ursus americanus*). BMC Genomics 12:171.

Fedorov VB, Goropashnaya AV, Tøien O, Stewart NC, Gracey AY, Chang C, Qin S, Pertea G, Quackenbush J, Showe LC, et al. 2009. Elevated expression of protein biosynthesis genes in liver and muscle of hibernating black bears (*Ursus americanus*). Physiol Genomics 37:108–18.

Flicek P, Amode MR, Barrell D, Beal K, Brent S, Carvalho-Silva D, Clapham P, Coates G, Fairley S, Fitzgerald S, et al. 2012. Ensembl 2012. Nucleic Acids Res 40:D84–90.

Geiser F. 2004. Metabolic rate and body temperature reduction during hibernation and daily torpor. Annu Rev Physiol 66:239–74.

Genome 10K Community of Scientists. 2009. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. J Hered 100:659–74.

Gummer DL. 2005. Geographic variation in torpor patterns: the northernmost prairie dogs and kangaroo rats. Canada: University of Saskatchewan.

Hampton M, Melvin RG, Kendall AH, Kirkpatrick BR, Peterson N, Andrews MT. 2011. Deep sequencing the transcriptome reveals seasonal adaptive mechanisms in a hibernating mammal. PLoS One 6:e27021.

Hedges SB, Dudley J, Kumar S. 2006. TimeTree: a public knowledge-base of divergence times among organisms. Bioinformatics (Oxford, England) 22:2971–2.

Li B, Zhang G, Willersleve E, Wang J. 2011. Genomic data for the polar bear (*Ursus maritimus*). GigaScience. gigadb.org/dataset/100008.

Lindblad-Toh K, Garber M, Zuk O, Lin MF, Parker BJ, Washietl S, Kheradpour P, Ernst J, Jordan G, Mauceli E, et al. 2011. A high-resolution map of human evolutionary constraint using 29 mammals. Nature 478:476–82.

Liu Y, Cotton JA, Shen B, Han X, Rossiter SJ, Zhang S. 2010. Convergent sequence evolution between echolocating bats and dolphins. Curr Biol 20:R53–4.

Löytynoja A, Goldman N. 2008. Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. Science 320:1632–5.

Melvin RG, Andrews MT. 2009. Torpor induction in mammals: recent discoveries fueling new ideas. Trends Endocrinol Metab 20:490–8.

Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, Morris Q, Bader GD. 2010. GeneMANIA Cytoscape plugin: fast gene function predictions on the desktop. Bioinformatics 26:2927–8.

Nevo E. 1979. Adaptive convergence and divergence of subterranean mammals. Annu Rev Ecol Syst 10:269–308.

Parker J, Tsagkogeorga G, Cotton JA, Liu Y, Provero P, Stupka E, Rossiter SJ. 2013. Genome-wide signatures of convergent evolution in echolocating mammals. Nature 502:228–31.

Schwartz C, Hampton M, Andrews MT. 2013. Seasonal and regional differences in gene expression in the brain of a hibernating mammal. PLoS One 8:e58427.

Seim I, Fang X, Xiong Z, Lobanov AV, Huang Z, Ma S, Feng Y, Turanov AA, Zhu Y, Lenz TL, et al. 2013. Genome analysis reveals insights into physiology and longevity of the Brandt's bat *Myotis brandtii*. Nat Commun 4:2212.

Shao C, Liu Y, Ruan H, Li Y, Wang H, Kohl F, Goropashnaya AV, Fedorov VB, Zeng R, Barnes BM, et al. 2010. Shotgun proteomics analysis of hibernating arctic ground squirrels. Mol Cell Proteomics 9:313–26.

Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. 2011. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27:431–2.

Soukri A, Valverde F, Hafid N, Elkebbaj MS, Serrano A. 1996. Occurrence of a differential expression of the glyceraldehyde-3-phosphate dehydrogenase gene in muscle and liver from euthermic and induced hibernating jerboa (*Jaculus orientalis*). Gene 181:139–45.

Srere HK, Wang LC, Martin SL. 1992. Central role for differential gene expression in mammalian hibernation. Proc Natl Acad Sci USA 89:7119–23.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res 34:W609–12.

Vidal M, Cusick ME, Barabási A-L. 2011. Interactome networks and human disease. Cell 144:986–98.

Villanueva-Cañas JL, Laurie S, Albà MM. 2013. Improving genome-wide scans of positive selection by using protein isoforms of similar length. Genome Biol Evol 5:457–67.

Watts DJ, Strogatz SH. 1998. Collective dynamics of ''small-world'' networks. Nature 393:440–2.

Williams DR, Epperson LE, Li W, Hughes MA, Taylor R, Rogers J, Martin SL, Cossins AR, Gracey AY. 2005. Seasonally hibernating phenotype assessed through transcript screening. Physiol Genomics 24:13–22.

Wilson BE, Deeb S, Florant GL. 1992. Seasonal changes in hormone-sensitive and lipoprotein lipase mRNA concentrations in marmot white adipose tissue. Am J Physiol 262:R177–81.

Xu Y, Shao C, Fedorov VB, Goropashnaya AV, Barnes BM, Yan J. 2013. Molecular signatures of mammalian hibernation: comparisons with alternative phenotypes. BMC Genomics 14:567.

Yan J, Barnes BM, Kohl F, Marr TG. 2008. Modulation of gene expression in hibernating arctic ground squirrels. Physiol Genomics 32:170–81.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol 24:1586–91.

Yuan L, Zhao X, Lin B, Rossiter SJ, He L, Zuo X, He G, Jones G, Geiser F, Zhang S. 2011. Adaptive evolution of Leptin in heterothermic bats. PLoS One 6:e27189.

Zhang G, Cowled C, Shi Z, Huang Z, Bishop-Lilly KA, Fang X, Wynne JW, Xiong Z, Baker ML, Zhao W, et al. 2013. Comparative analysis of bat genomes provides insight into the evolution of flight and immunity. Science 339:456–60.

Zhang J, Nielsen R, Yang Z. 2005. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. Mol Biol Evol 22:2472–9.

Zhao S, Shao C, Goropashnaya AV, Stewart NC, Xu Y, Tøien Ø, Barnes BM, Fedorov VB, Yan J. 2010. Genomic analysis of expressed sequence tags in American black bear *Ursus americanus*. BMC Genomics 11:201.