



**HAL**  
open science

**For consideration as an Article in the Discoveries section  
of MBE Aging at evolutionary crossroads: longitudinal  
gene co-expression network analyses of proximal and  
ultimate causes of aging in bats**

Bernard Guillaume, Teulière Jérôme, Lopez Philippe, Corel Eduardo,  
Lapointe François-Joseph, Bapteste Eric, Eduardo Corel

► **To cite this version:**

Bernard Guillaume, Teulière Jérôme, Lopez Philippe, Corel Eduardo, Lapointe François-Joseph, et al..  
For consideration as an Article in the Discoveries section of MBE Aging at evolutionary crossroads:  
longitudinal gene co-expression network analyses of proximal and ultimate causes of aging in bats.  
Molecular Biology and Evolution, 2021, 10.1093/molbev/msab302/6400255 . hal-03391118

**HAL Id: hal-03391118**

**<https://hal.sorbonne-universite.fr/hal-03391118>**

Submitted on 21 Oct 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

For consideration as an Article in the Discoveries section of MBE

**Aging at evolutionary crossroads: longitudinal gene co-expression network analyses of proximal and ultimate causes of aging in bats**

Bernard Guillaume<sup>1</sup>, Teulière Jérôme<sup>1</sup>, Lopez Philippe<sup>1</sup>, Corel Eduardo<sup>1</sup>, Lapointe François-Joseph<sup>2</sup>, Bapteste Eric<sup>1,\*</sup>

<sup>1</sup>Institut de Systématique, Evolution, Biodiversité (ISYEB), Sorbonne Université, CNRS, Museum National d'Histoire Naturelle, EPHE, Université des Antilles, 75005 Paris, France.

<sup>2</sup>Département de sciences biologiques, Complexe des sciences, 1375 avenue Thérèse-Lavoie-Roux, Université de Montréal, Montréal (Québec) H2V 0B3, Canada.

\*Corresponding author: Bapteste Eric, 9 Quai Saint-Bernard, Sorbonne Université, Equipe AIRE, Bat A 4<sup>ème</sup> étage, Pièce 427, 75005 Paris, France. Tel : +(33) 1 44 27 21 64, E-mail: eric.bapteste@mnhn.fr

## Abstract

How, when and why do organisms, their tissues and their cells age remain challenging issues, although researchers have identified multiple mechanistic causes of aging, and three major evolutionary theories have been developed to unravel the ultimate causes of organismal aging. A central hypothesis of these theories is that the strength of natural selection decreases with age. However, empirical evidence on when, why and how organisms age is phylogenetically limited, especially in natural populations. Here, we developed generic comparisons of gene co-expression networks that quantify and dissect the heterogeneity of gene co-expression in conspecific individuals from different age-classes to provide topological evidence about some mechanical and fundamental causes of organismal aging. We applied this approach to investigate the complexity of some proximal and ultimate causes of aging phenotypes in a natural population of the greater mouse-eared bat *Myotis myotis*, a remarkably long-lived species given its body size and metabolic rate, with available longitudinal blood transcriptomes. *Myotis* gene co-expression networks become increasingly fragmented with age, suggesting an erosion of the strength of natural selection and a general dysregulation of gene co-expression in aging bats. However, selective pressures remain sufficiently strong to allow successive emergence of homogeneous age-specific gene co-expression patterns, for at least seven years. Thus, older individuals from long-lived species appear to sit at an evolutionary crossroad: as they age, they experience both a decrease in the strength of natural selection and a targeted selection for very specific biological processes, further inviting to refine a central hypothesis in evolutionary aging theories.

## Introduction

Aging studies confront us with major, pressing questions: how, when and why do organisms, their tissues, their cells and their molecules age? Decades of scientific work have identified a diversity of plausible mechanistic causes of aging (Harman 1956; Finch 1990; Vijg 2007; Kenyon 2010). In addition to such proximal causes of aging (Mayr 1961), during the past seventy years, three compatible, mutually non-exclusive evolutionary theories have sought to unravel the ultimate causes of aging in a diversity of organisms (Kirkwood 2005; Johnson et al. 2019). These evolutionary theories rest on a common argument: all organisms inevitably die from extrinsic sources, be it by accident, predation, disease outbreak or due to exceptionally

harsh conditions (Reichard 2017). Since no individual is immortal, irrespective of the existence of aging, as progressively fewer individuals within populations escape these extrinsic sources of mortality and are thus able to experience the effects of natural selection, from an evolutionary viewpoint the later periods of any organisms' lives are predicted to be less important than their early periods (Reichard 2017). Thus, the Mutation Accumulation (MA) theory of aging (Medawar 1952) predicts that late-expressed deleterious mutations will accumulate in organismal genomes, when the strength of natural selection has diminished so much that it cannot purge the older members of the populations from their 'rotten' late-expressed genes or further postpone these genes expressions in the lifetime of organisms. Hence, senescence, the gradual decrease in biological functions, occurs when such detrimental genes contribute to the decline of organisms late in their lives. The Antagonistic Pleiotropy (AP) theory of aging (Williams 1957) also fundamentally explains senescence by the decreasing strength of natural selection acting on chronologically aging individuals. As external death is inevitable, the selection for genes that benefit organisms should be more active early in their lives, even when these genes with early positive fitness effects, known as antagonistic pleiotropic genes, impair the survival of their bearers later in life. Finally, the Disposable Soma theory (DS) (Kirkwood 1977; Kirkwood and Holliday 1979) stresses that, because resources are limited, most organisms will be more successful if they invest their finite energy into their germline and into their reproduction rather than into the maintenance of their soma (Reichard 2017). Thus, according to DS, as selection will tend to favor investment in early reproduction over long-term somatic maintenance, evolutionary trade-offs get established in populations and natural selection is expected to weaken with age. This drives senescence, when investing energy in mechanisms that fight against somatic aging becomes, from an evolutionary perspective, exceedingly costly (Kirkwood and Rose 1991; Johnson et al. 2019). Of note, in contrast with MA, which should primarily explain senescence in preserved populations (e.g., human beings benefiting from medical progresses, or animals living in protected niches outside their natural environment (Medawar 1952; Nussey et al. 2013), AP and DS can also be mobilized to explain senescence beyond such protected or medicalized populations (Nussey et al. 2013). This difference in explanatory scope matters, because senescence is commonly detected in nature in birds, mammals, other vertebrates and insects, whose free-living populations present age-related changes in adulthood, including decrease in survival probability, decline in reproductive performance traits or functional alterations of other physiological or behavioural traits (Nussey et al. 2013).

Reassuringly for the classic evolutionary explanations of aging, analyses of aging phenotypes in model species returned observations compatible with the central hypothesis that the strength of natural selection decreases with age. Longitudinal studies conducted in various species generally report an increase in variation across traits as a function of age (Mendenhall et al. 2021), which ultimately suggests that natural selection exerts a weaker constraint on late than on early phenotypes. This increasing variation across traits with age has been proposed to result from increased gene expression changes within aging organisms. For example, chromosomal instability might produce major disturbances in gene expression with age in mammals (Dollé et al. 2000; Dollé and Vijg 2002; Vijg 2007) . Likewise, it has also been proposed that cumulative somatic mutations, disrupting the transcriptional networks that regulate cell structure and function, might contribute to the aging process (Bahar et al. 2006).

Supporting these proximal explanations of aging, (Bahar et al. 2006) reported a significant increase in transcriptional noise of nuclear genes in single postmitotic cardiomyocytes collected from fresh heart tissue samples in old male mice with respect to young male mice. This observation was considered to result from an accumulation of somatic DNA damage during aging, with detrimental effects on normal cell functioning. Similarly, (Martinez-Jimenez et al. 2017) reported that cell-to-cell transcriptional variability upon immune stimulation increases with aging in memory CD4<sup>+</sup> T cells in young and old mice, and in two divergent species. Likewise, single-cell transcriptome analysis of human pancreatic cells showed that islet endocrine cells from older donors display significantly increased levels of transcriptional noise with respect to that of younger donors (Enge et al. 2017). Moreover, single cell transcriptomics analyses across 30 cell types of young and old mice suggested that aging leads to increased transcriptional noise in most cell types, suggestive of a deregulated epigenetic control (Angelidis et al. 2019). Altogether, these results were interpreted to support an association between animal aging on the one hand, and an increased transcriptional dysregulation caused by a stochastic gradual accumulation of both epigenetic and genetic errors, on the other hand.

Yet, (Warren et al. 2007) nuances these simple correlations. The analysis of gene expression noise for six different mRNA transcripts in four hematopoietic cell types from phenotypically equivalent cells in young and old mice revealed a trend toward higher transcript levels in cells isolated from old animals, but no significant increase in transcriptional heterogeneity with age (Warren et al. 2007). These authors concluded that large-scale regulatory destabilization is not a universal hallmark of aging, but rather that transcriptional

dysregulation may be primarily a relevant aging mechanism in nonrenewing tissues. Similarly, Kimmel et al. (Kimmel et al. 2019) analyzed single-cell transcriptomic from three tissues in young and old mice and reported both increased and decreased variations in gene expression with age, depending on the genes and on the cell types investigated. Overall, longitudinal single cell transcriptomic analyses presently support a subtle consensual view; that is, that increased cell-to-cell variation in transcription appears as a hallmark of aging for some genes, in some tissues (Mendenhall et al. 2021).

This nuanced conclusion raises novel questions about the ultimate causes of aging. Increased gene expression heterogeneity within older organisms is surely compatible with a relaxation of the selective pressures exerted on gene expression and with a random accumulation of damage within individuals with age (Tarkhov et al. 2019). However, if some somatic tissues and cells can resist a stochastic deregulation of their gene expressions with age (e.g., as they genuinely maintain homogeneous gene expression profiles), more elaborated models on the strength of selective pressures acting on aging organisms and their eventual targets in older individuals may be warranted. Furthermore, there seems to be a theoretical tension between explanations of aging that rely on a general weakening of natural selection with age, and explanations that hold that natural selection actively shapes some trade-offs and some longevity programs, operating until very late in organismal life. On the one hand, selection would then appear too weak to sustain the fitness of older organisms, whereas, on the other hand, selection would be sufficient to enforce regular expression of late biological processes throughout their lives. Such seemingly different expectations on the influence of selective pressures suggest that the causes of aging are probably more complex than an increasing evolutionary neglect with time. Therefore, scientists need to know more about whether, when and how a decrease in the strength of natural selection contributes to aging (Cohen et al. 2020), especially in natural populations.

To this end, we introduce a generic network approach that quantifies and compares the heterogeneity of gene co-expression profiles in conspecific individuals from different age-classes (e.g., young, old, very old). Our approach focuses on gene co-expression, because aging involves multiple interacting biological pathways (Huang et al. 2019), which can be tracked using gene co-expression networks (GCN). Comparisons of GCNs constructed for successive age-classes within a population effectively provide a systemic view on the changes affecting gene co-expression over time for a broad range of biological processes. Unlike past applications of GCNs that mostly focused on identifying co-expression modules associated with aging

(Southworth et al. 2009; Baumgart et al. 2016; Huang et al. 2019), our novel approach exploits the whole topology of GCNs to determine i) how aging occurs (e.g., what sets of co-expressed genes are affected by aging or are age-specific); ii) when aging occurs (e.g., which age-classes present altered co-expression patterns associated with aging); and iii) what ultimate causes may drive topological changes detected between GCNs representing different age-classes (e.g. whether GCNs are increasingly fragmented with age, a topological dynamics suggestive of a weakening natural selection). These three goals can be simultaneously achieved. How organisms age can be addressed because GCNs are commonly used to discover the molecular bases, mechanisms and regulations of complex processes (Fuente 2010; Meng et al. 2018; Chowdhury et al. 2020), using module-centered approaches, hub gene finding, guilt by association analysis or by identifying genes with strongly altered connectivity under different conditions (Fuente 2010; Chowdhury et al. 2020), in our case under different age-classes. When organisms age is deducible because GCNs can help in understanding disease progression, by identifying genes that show similar co-expression across different stages of a disease (Chowdhury et al. 2020). Consistently, by discovering local and global topological similarities between networks from different age-classes (e.g., by detecting common subnetworks between young and intermediate bats, between intermediate and old bats, etc.), it can be assessed in which age-classes differences in co-expression associated with aging occurs. Finally, interpreting the loss of network connectivity associated with age as an effect of relaxed selective constraints offers a way to address why organisms age.

We applied this network approach on longitudinal blood transcriptomes from a natural population of greater mouse-eared bats (*Myotis myotis*), a suitable species to investigate the complexity of proximal and ultimate causes of aging phenotypes, due to their remarkable long lifespans, given their small body size and high metabolic rate (Austad 2010; Munshi-South and Wilkinson 2010; Wilkinson and Adams 2019; Gorbunova et al. 2020). Whereas some authors have recently postulated that longer lifespan could be correlated with a smaller litter size in bats compared to other mammals (Garbino et al. 2021), the unusual longevity of bats is traditionally explained by genetic and epigenetic adaptations that counteract aging until late in their lives. Namely, (Wilkinson et al. 2021) compared DNA methylation (DNAm) profiles across 26 bat species to identify epigenetic changes associated with age and longevity. This study showed that DNAm accurately predicts chronological age, and that longevity was negatively associated with the rate of DNAm change at age-associated sites. (Huang et al. 2020) conducted a comparative genomic and transcriptomic study on long-lived versus short-lived bats.

Expression analysis of 2,086 genes with significant differences indicated that long-lived bats exhibited a transcriptomic profile of enhanced DNA repair and autophagy pathways. These results were further confirmed by a phylogenomic longitudinal analysis across eutherian mammals, which uncovered 10 autophagy-associated genes under selective pressure in bat lineages (Kacprzyk et al. 2021). Thus, according to the evolutionary theories of aging, aging bats should display a general dysregulation of their individual transcriptomes with age; yet the unusual longevity of bats is conversely traditionally explained by adaptations that counteract aging until late in their lives. Consistently, we observed that bats GCNs become significantly more fragmented with age, which testifies of an increased heterogeneity in gene co-expression with age, compatible with a decrease of the strength of natural selection acting on older bats. However, GCN analyses also unraveled a succession of age-specific biological processes, unraveled by age-specific rewiring of co-expression networks, and the persistence of core functional processes. Altogether, these observations suggest that, despite an overall weakening regulation of gene co-expression with time, aging bats are still under sufficient selection to experience a transition towards a new, age-specific biology, superimposed onto persistent central biological programs.

## Results

### Introducing an original network-based approach to study aging

We constructed GCNs from longitudinal transcriptomes of blood, sampled in natural populations of greater mouse-eared bats (Huang et al. 2019). These networks were inferred by age-class, by applying Weighted Gene Co-expression Network Analysis (WGCNA) on 16 blood samples for each age-class, with a double stringency condition to consider gene co-expression as significant (e.g.,  $p$ -value  $< 0.05$  and Pearson coefficient  $> 0.8$ ) (see M&M). In these networks, nodes correspond to protein-coding genes, connected by weighted edges that represent the strength of their co-expression. Thus, each of our GCNs features the global set of significant gene co-expression, homogeneously shared by individual adult bats from a given age-class.



GCNs have already been used to unravel a broad range of subtle age-related transcriptional variations across tissues and across individuals within species (Somel et al. 2010; Baumgart et al. 2016; Pacifico et al. 2018; Huang et al. 2019), with a traditional focus on the heuristic detection of network modules (densely connected sets of nodes), associated with aging (Southworth et al. 2009; Ferreira et al. 2021). Here, we developed the use of GCNs for aging studies in another direction by exploiting different aspects of their architectures via comparisons of the general and local topological properties of GCNs constructed for different age-classes. Our goal was to characterize some under-exploited data, informative about both the proximal and ultimate causes of aging. For this, we quantified the general and local connectedness within each network (see M&M) and assessed how these metrics change with age. In particular, when GCNs produced from a similar number of samples between age-classes become increasingly disconnected with age, this topological change indicates that gene co-expression has become increasingly heterogeneous with age, a pattern compatible with a general decrease of the strength of natural selection on aging organisms. We hypothesize that this is because younger individuals have their gene co-expression more constrained by natural selection and are less exposed to the impact of individual random damage accumulation, and therefore tend to express more similar sets of genes within their age-class. By contrast, the gene co-expression of older individuals was predicted to be more affected by chance events (due to their different experiences and the decrease of natural selection strength).

### **GCN topologies are compatible with a decreasing strength of natural selection with age**

We computed simple exact statistics on GCNs constructed from the same number of samples from three age-classes, corresponding to young (1-2 years old), intermediate (3-4 years old) and older (6-7 years old) adult bats, respectively, at three increasingly high stringency thresholds of correlation values (see M&M, Fig. SI-1). These metrics show that the general and local topological properties of the GCNs produced for different age-classes vary markedly, regardless of stringency thresholds. Analyses of the distributions of clustering coefficient values (Fig. SI-2, Tables SI-3, SI-4) and of shortest path lengths (Fig. SI-2, Tables SI-3, SI-4) show that GCNs from different age-classes present significantly different profiles of local connectedness and of general connectedness, first investigated with the minimal distances between any pairs of nodes. The clustering coefficient values significantly decrease with age, meaning that local networks from older bats appear less connected than local networks from

younger bats. Moreover, the lengths of the shortest paths significantly increase with age, suggesting that nodes in networks are increasingly distant with age. However, the distributions of clustering coefficient values and that of shortest path lengths from young, intermediate, and older bats are too entangled to be amenable to any simple biological interpretation, meaning that these two indices are not the simplest ones to interpret for this dataset.

By contrast, other network metrics obviously show that heterogeneity in gene co-expression increases with age. First, GCNs from older bats are smaller (comprising less nodes and edges) than GCNs from young and intermediate bats (Table 1), even though the numbers of transcripts are comparable among individuals and age-classes. This reduction in network size implies that, in terms of absolute numbers, there is less homogeneous gene co-expression detected in older bats than in younger bats. More importantly, in addition to getting smaller, GCNs become increasingly fragmented with age (Table 1). Accordingly, the numbers of disconnected subgraphs (i.e., connected components in Table 1) increase in GCNs associated with increasingly older bats (Fig. 1A). This increased disconnectedness is noticeable even for GCNs with comparable numbers of nodes (e.g., for stringency thresholds of p-value  $< 0.05$  and Pearson coefficient  $> 0.9$ ), so it is not trivially caused by the reduced size of the GCN from older bats. These first series of observations indicates an increased heterogeneity in gene co-expression with age.

Statistical comparisons of the normalized degrees of nodes in GCNs from increasingly old age-classes support a similar conclusion (Fig. 1): nodes of increasingly older bats have significantly fewer direct neighbors with age (Wilcoxon rank-sum tests, p-value  $< 2.10^{-16}$  for all comparisons between pairs of GCNs from different age-classes built at identical stringency). In Figure 1, this trend is illustrated by the progressively fading red coloring: the GCN from younger bats contains a higher proportion of high degree nodes than the GCN of intermediate bats, which itself presents more highly connected nodes than the much more pale colored GCN of older bats.

Likewise, analyses of the distributions of closeness values unravel a significant decrease in the closeness of nodes belonging to the older age-class (Fig. 2). These distributions are significantly different between age-classes (Wilcoxon rank-sum tests, p-value  $< 2.10^{-16}$  for all comparisons between pairs of GCNs from different age-classes built at identical stringency), and increasingly shift to lower values of closeness with age. These measures indicate that nodes are further away from one another in GCNs from increasingly older bats, consistent with a

general fragmentation of the gene co-expression networks. In figure 3, this is illustrated by the progressively fading red coloring with age: the GCN from younger bats contains a higher proportion of nodes with higher closeness values than the GCN of intermediate bats, which itself presents more central nodes than the much more pale colored GCN of older bats.

Overall, our topological analyses report an increasing local and general fragmentation of GCNs with age, which strongly suggests that gene co-expression becomes increasingly heterogeneous across individuals from the same age-class. Moreover, heterogeneity in gene co-expression is also noticeable when considering the numbers and proportions of genes, expressed by individuals within an age-class, that are not integrated into the GCN of that age-class. Such genes do not display any significant co-expression patterns, so they cannot be connected to other genes, which stresses the uniqueness of their expression. These ‘singletons’ represent 10% (for the 1-2 and 3-4 years old bats) to 14.1% (for the 6-7 years old bats) of expressed genes with more than 10 total read counts (i.e., for the 1-2 years old bats: 1232 singletons in 12240 expressed genes; for the 3-4 years old bats: 1224 singletons in 12160 expressed genes; for the 6-7 years old bats: 1681 singletons in 11859 expressed genes).

Importantly, such a pattern is compatible with a general decrease of the strength of natural selection with age, leading to the presence of different transcriptional noises in distinct individuals from the same age-class, possibly as a result of distinct random damage accumulations in distinct aging organisms. Accordingly, any increasingly old organism would present increasingly stochastically perturbed genetic interactions, which reduces the likelihood that all or most individuals from the same old age-class would express the same sets of genes, with the same intensity. This interindividual variation in gene co-expression results in increasingly fragmented GCNs, compatible with the central hypothesis that natural selection generally exerts weaker constraints on the biological processes implemented as organisms age. Moreover, the increase in heterogeneity in gene co-expression may be due to individual-specific experience. By definition, over time, different individuals are increasingly exposed to different situations, so they are expected to express increasingly different genes. In this case, an increased heterogeneity in gene co-expression would also reflect the passage of time: it would correlate with aging, but not necessarily be a cause of aging.

Under all these interpretations, the partial homogeneity of gene co-expression observed between age-classes and within age-classes, especially for the older bats, requires some specific

explanation, because it appears increasingly difficult for aging adult bats to maintain stable, cohesive gene co-expression.

### **Gene co-expression resists the decreasing strength of natural selection with age**

Comparisons of GCNs from different age-classes unraveled another remarkable topological property of these networks. A small proportion of the genetic co-expression realized in bats (hereafter called ‘core’ and corresponding to white nodes and edges in Fig. 4) are conserved in the GCNs across age-classes (for at least seven years), reporting homogeneous persistent gene co-expression across age-classes. For example, at a stringency threshold of  $p$ -value  $< 0.05$  and Pearson coefficient  $> 0.8$  (see M&M), identical connections between identical nodes are observed in GCNs from younger, intermediate, and older bats. This core network represents 18.1% of the nodes and 1.1% of the edges of the GCN of younger bats, 18.2% of the nodes and 4% of the edges of the GCN of intermediate bats, and 19.4% of the nodes and 6.4% of the edges of the GCN of older bats (SI-5). The existence of a small core gene co-expression network, corresponding to a modest yet larger proportion of age-specific GCNs with age, implies that a part of the GCNs, actually their most central part in terms of closeness centrality (Fig. 5), can resist the general fragmentation of networks with age, hence the general dysregulation.

There are several possible interpretations for such a persistent, albeit limited in size, network pattern. First, core nodes and edges may correspond to more robust biological processes, that, for systemic reasons, better resist the transcriptomic stochastic dysregulation induced by evolutionary neglect in older individuals. For example, their regulation is maybe more redundant, and robust to failure. Accordingly, core co-expressed genes may belong to biological processes that keep on running in organisms in spite of their aging. In theory, such biological processes may even have detrimental fitness effects, as in the theoretical case of quasi-programs (Blagosklonny 2006), which, despite deleterious effects are not turned-off due to the weakening of natural selection. In the long run, some of the core gene co-expressions may then be involved in aging, or they may correspond to basic, vital processes, whose failure would be lethal. Second, the stability of core nodes and edges in the face of a decreasing natural selection fragmenting the rest of the GCN may reflect the ability of renewing cells within blood to resist, at least partly, a general dysregulation in older bats, consistently with the work by Warren et al. (Warren et al. 2007). This resistance to dysregulation with age may be expected, for example, for transcripts from hematopoietic cells, which escape the worst effects of genomic

instability, as such damaged cells are rigorously purged from the population (Warren et al. 2007), or for transcripts of granulocyte and naïve B and T cell populations in the peripheral blood, which are continuously renewed from the stem cell pool in the bone marrow (Warren et al. 2007). The turnover rate of granulocytes is especially rapid –perhaps less than a day (Galbraith et al. 1965)– meaning that these cells are in some respect younger than many other components of the organism. Yet, our longitudinal analysis also shows that, even within a tissue composed of renewing cells like blood, there are obvious signs that GCNs become generally more fragmented with age, hence signs of a general dysregulation of gene co-expression in older organisms. Third, within renewing cells, and within blood, the decreasing strength of natural selection with age may unevenly affect different genetic processes and different genes, some of them being submitted to more targeted selective pressures than the rest of the processes and genes of the genomes.

A functional analysis of these core nodes does not allow to distinguish among these alternative explanations, but it shows that core gene co-expression encompasses a functionally biased set of genes with respect to the general proteome in *Myotis* bats (SI-6). Core co-expressed genes are involved in a variety of enriched biological processes, associated with renewing blood cell biology, but also with ongoing viral and bacterial infection. At any rate, GCNs of bats seems at an evolutionary crossroads, with some but not all of their gene co-expression becoming increasingly variable with age.

The labeling of GCNs nodes and edges by the age of their hosts (1-2 years, 3-4 years, or 6-7 years, or a combination of these ages, see M&M) unravels another remarkable pattern in GCNs from *Myotis* bats (Fig. 4). In addition to conserved, core gene co-expression, each GCN contains significantly peripheral nodes (Wilcoxon rank-sum test, corrected p-value <0.05) involved in age-specific gene co-expressions (Fig. 5). That is that, for this dataset, younger, intermediate, and older bats appear to homogeneously co-express distinct sets of genes. Remarkably, the proportion of age-specific co-expression edges within a GCN significantly decreases across age-classes, at all stringency thresholds tested (Unilateral Z-test, p-value <0.05, SI-5), consistently with the hypothesis that natural selection decreases with age. For example, the analyses performed at a minimum Pearson threshold of > 0.8 show that 90.7% of the edges in the GCN of younger bats, 75.1% of the edges in the GCN of intermediate bats, and 59.6% of the edges in the GCN of older bats are detected only in the corresponding age-classes. Thus, proportionally and in absolute numbers of edges, increasingly older bats show less

homogeneous age-specific gene co-expression, but still they harbor age-specific gene co-expression.

To further demonstrate that aging correlated with noticeable topological changes in age-specific subnetworks, we used the labels associated with the different age-classes to decompose the GCNs constructed from young, intermediate, and older bats into subnetworks only containing core nodes and edges ('core-only' subnetworks), or only containing nodes connected by age-specific gene co-expression edges ('younger-only', 'intermediate-only' and 'older-only' subnetworks). Topological analyses of these subnetworks also show that, at all stringency thresholds, age-specific subnetworks are significantly more fragmented (e.g. composed of significantly more connected components) than the core subnetwork from a given GCN.

We compared the sizes of the connected components (i.e., number of edges) associated exclusively with younger, intermediate, and older bats, respectively. At each stringency threshold, we observed that age-specific connected components from older bats are significantly smaller, both in absolute and normalized numbers of edges (Wilcoxon rank-sum test, corrected by the Bonferroni method, SI-7), than age-specific connected components from younger age-classes (see M&M). Assuming that the number of edges within a connected component of an age-specific subnetwork reflects the number of realized molecular interactions supported by the co-expressed genes, then age-specific gene co-expressions would likely support shorter biological processes (in numbers of involved interactions) with age. Therefore, the ability to implement novel, age-specific complex biological processes or to maintain complex biological processes would be reduced in older bats with respect to younger bats. This conclusion would also be compatible with a decrease of selective pressures with age.

To confirm this interpretation, we computed statistics based on motifs (networks patterns that are significantly enriched with respect to null networks of similar sizes and degree distributions than each of our GCNs (Milo et al. 2002)). The existence of motifs in networks is often (but not always (Stone et al. 2019)) considered as topological evidence for selection on a network architecture (Shen-Orr et al. 2002). Specifically, we quantified the ratio of triangles over all motifs of size  $n=3$  nodes in each GCN. This metrics is normalized, can be compared between networks and reflects the connectivity within the network. By definition, this ratio equals 1 for a maximally connected GCN. Under the common assumption that the detection of motifs suggests the impact of some selection on network topology, our motif analyses also showed that bats GCNs are under selection. First, GCNs from aging bats are more densely

connected than expected by chance (hence under selection) (TABLE SI-8). Second, the strength of this selection is greater in the networks from younger bats (i.e., the ratio of connected motifs significantly decreases with age in GCNs (TABLE SI-9)).

Yet, although various lines of evidence suggest that natural selection could be decreasing, selective pressures on gene co-expression would still be sufficiently strong to constrain the implementation of multiple homogeneous age-specific biological processes after six or seven years of life.

We complemented these topological analyses by a functional analysis. For all tested networks, regardless of the stringency level, genes involved in core processes are significantly enriched in the functional Eukaryotic orthologous groups (KOG) category T (Signal transduction mechanisms) with respect to genes involved in age-specific co-expression (SI-10). By contrast, non-core genes were depleted in the functional categories T and Z (Cytoskeleton) for all age-classes at all thresholds of Pearson correlations (SI-11). Moreover, for each functional KOG category, finer functional ontology enrichment analyses (see M&M) show that age-specific processes are enriched in similar GO functions in the different age-classes with respect to the proteome (Fig. SI-12), suggesting that GCNs are rewired with age but encode similar functions through different genetic associations. Moreover, functional analysis of expressed genes absent of GCNs, or singletons, show that they are enriched in mitosis-, apoptosis- and stress response-related terms, suggesting that interindividual expression variation is at least in part associated to differences in white blood cells rates of proliferation and death, and experienced external stresses, a likely consequence of the variable external stimuli encountered by individual wild bats in their natural environment (Fig. SI-13A).

Overall, while their general and local gene co-expression patterns differ with age, in terms of enriched biological functions, aging bats do not seem to rely on a fundamentally novel biology during the courses of their lives. Rather, despite an age-specific rewiring of the GCNs, older bats seem to realize the same functions (or a subset of the functions) also achieved by younger and/or intermediate bats. These results suggest that as they age, adult bats maintain a largely similar biology, implemented however through different, age-specific gene co-expression patterns. This observation suggests either that gene regulation for achieving the same enriched functions is age-specific in bats, and thus that some gene regulatory networks change with age. Or, it suggests that our dataset subsampled the whole sets of gene co-expression truly involved to realize these enriched functional categories, with the result that

different sets of gene co-expression supporting these functions were reported in the networks for different age-classes. Future enhanced transcriptomic sampling will allow to distinguish between these two, mutually non-exclusive hypotheses. Yet, remarkably, because even older bats continue to display the same enriched biological processes realized by younger bats, some selection seems still active that preserves the decline of these particular functions in aging organisms, in spite of an increased heterogeneity of gene co-expression in aging bats.

### **Network analyses of longitudinal blood transcriptomes from mice**

To learn whether our findings on bats aging were specific to this species, we compared them with observations gathered from a model aging species, the short-lived mouse *M. musculus*.

Although our differential network-based analysis of aging bats blood transcriptomes from three age-classes is not directly comparable with the module-centered differential gene co-expression analyses of 13 tissues realized for a smaller sample of mice from two adult age-classes (16-month-old and 24-month-old mice) by (Southworth et al. 2009). These two studies are the closest one can compare. They suggest commonalities and differences in bats and mice aging at the level of their gene co-expression. While the networks from aging mice do not become significantly smaller in terms of number of edges, they present a significant loss of connectivity within modules with age (Southworth et al. 2009). This pattern dynamics suggests that, as for bats, some aspects of gene co-expression would become more heterogeneous in mice with age. However, whereas some inflammation pathways become constitutively induced in the older mice through the high activity of NF- $\kappa$ B, we did not detect any sign of age-specific activation of NF- $\kappa$ B target genes in aging bats. Rather, we noticed that the functional category kog T, corresponding to the Toll like receptor cascades, was even depleted in old bats (in the least stringent of our GCNs (Pearson  $> 0.8$ )), a seemingly opposite trend in aging bats and in aging mice.

In general, regarding the functions affected during mice and bats aging, systematic comparisons of the functions found enriched by Metascape in our analysis of bat age-specific subnetworks with the functions from modules significantly correlated with age in mice according to (Southworth et al. 2009) suggest some similarities (e.g., a decrease in ribonucleoprotein complex biogenesis and assembly (KOG A), DNA replication (KOG L),



microtubule cytoskeleton, ...) but also some differences (in the ER to Golgi vesicle-mediated transport, in the response to DNA damage stimulus, ...).

To further test this conclusion, we reanalyzed publicly available blood transcriptomes from mice using the same methodology than for bats (see M&M). Specifically, we analyzed GCNs for two age-classes of mice (2-3 months old and 4-9 months old, respectively) with similar numbers of individuals than bats age-classes. Importantly, wild bats and lab mice have very different lifestyles and lifespans. While the mice we studied were chronologically aging, their two age-classes remained composed of biologically young individuals (Flurkey et al. 2007). Thus, although the mice dataset can be used to track effects of aging in the same tissue than for bats, one has to bear in mind that the two age-classes we could analyze for mice span over a time period where lab mice are known to be healthy, without reported reduction in their fertility, hence physiologically healthy aging adults (Flurkey et al. 2007). Therefore, these two age-classes from mice have *a priori* no reason to match with the lifecycle stages captured by the age-classes for bats, themselves composed of aging adults, which may well be physiologically more aged than mice from the other dataset. Consequently, GCNs from these aging bats and mice likely differ, and comparative interpretations can only be speculative and careful.

Still, some differences between aging bats and aging mice GCNs are worth mentioning to highlight the originality of the findings made for aging bats. First, despite these species harboring close genome sizes (typical of mammals (Kapusta et al. 2017) with a genome size of 2148.63 MB and a median protein count of 61156 for *Myotis* vs a genome size of 2689.66 MB and a median protein count of 61156 for *Mus musculus*), GCNs from mice were much larger than GCNs from bats, for all age-classes, at all considered thresholds, both in terms of nodes and edge numbers (SI-Table 14). Thus, wild bats always displayed a much greater heterogeneity in co-expression than lab mice. Second, while GCNs from bats displayed marked topological differences and appeared increasingly fragmented with aging, GCNs from mice up to 9 months old did not display topological evidence of fragmentation (Table SI-15). Rather, GCNs topologies from our two age-classes of mice were much more similar than that of aging bats. Core nodes represented 52.8 % (Pearson correlation > 0.8; 40.3% at  $P_c > 0.85$ ; 33.1% at  $P_c > 0.9$ ) of the younger mice GCN and 51.6 % (Pearson correlation > 0.8; 38.5% at  $P_c > 0.85$ ; 28.6% at  $P_c > 0.9$ ) of the older mice GCN. Likewise, conserved co-expression edges (core edges) represented 25.6 % (Pearson correlation > 0.8; 22 % at  $P_c > 0.85$ ; 17.4 % at  $P_c > 0.9$ ) of the younger mice GCN and 23.7 % (Pearson correlation > 0.8; 21.4% at  $P_c > 0.85$ ; 18.2% at  $P_c$

> 0.9) of the older mice GCN. Age-specific gene co-expression in mice were largely peripheric, forming notably small-sized connected components in their GCNs (Fig. SI-16), which are responsible for slight statistical differences in the network metrics for 2-3 months old mice and for 4-9 months old mice (Fig. SI-17). Therefore, while mice aged of 2-3 months implement common as well as some different gene co-expression with respect to mice aged of 4-9 months, no general process of functional decline that would result in detectable increased gene co-expression heterogeneity was identified in aging mice for this dataset. Consistently, the proportion of singletons was stable and represented ~8% of expressed genes with more than 10 total read counts for 2-3 and 4-9 months old mice age-classes (for 2-3 months old mice: 1646 singletons in 20803 expressed genes; for the 4-9 months old mice: 1650 singletons in 21242 expressed genes). Thus, the relative topological stability without evidence of fragmentation for mice GCNs between 2 to 9 months of life, compared to the topological changes observed for bats GCNs between 1 to 6+ years of life, unravels a greater heterogeneity in gene co-expression during aging in wild bats than in lab mice, which may be related to differences in lifestyle, lifespan and aging rates throughout the lifecycle of these two species, and the fact that mice are inbred to be genetically homogeneous.

Finally, in terms of functions, genes involved in core co-expression in mice, on the one hand and in bats in the other hand, were rather similar. They only consistently differed at all Pearson correlation thresholds for the KOGs I (Lipid transport and metabolism), O (Posttranslational modification), U (Intracellular trafficking, secretion, and vesicular transport), R (General function prediction only) and S (Function unknown). The functions enriched in the KOG V (Defense mechanisms) do not indicate a generalized inflammation associated with aging in these young mice (Fig. SI-18). We also analyzed the functions associated with age-specific genes in bats and mice (Fig. SI-19). At the level of KOG distributions, the two species only consistently differed for the KOG T (Signal transduction mechanisms) at all Pearson correlation thresholds.

By contrast, bats and mice differed at the level of singletons-associated functions, i.e., genes expressed by individual mice or bats whose expression was not correlated with that of other genes within a given age-class, corresponding to the most heterogenous population of genes one can associate with aging (Fig. SI-13B). Interestingly, mice did not display an enrichment in functions associated with cell proliferation, cell death and stress responses, but rather in functions associated with organismal development and hormonal control. Such a result may be explained by the controlled, standardized artificial laboratory environment in which

mice were raised, which may lead to more interindividual coexpression regarding blood transcriptional responses to external stimuli than for bats living in nature. Interindividual variation between aging mice would rather appear to originate from differences in growth mechanisms, possibly affected by differential epigenetic regulation (Lathe 2004; Voelkl et al. 2020). Overall, these results suggest that our findings regarding topological changes in GCNs from blood transcriptomes of aging bats reflect aspects specific of the aging of bats in their environment, which were not recovered for the same tissue in mice aging in the lab.

## Discussion

We tested the hypothesis that variation of overall gene co-expression would increase with age in bats, using network-based differential gene co-expression analyses that provide a holistic view of how gene co-expression relationships change during aging. Taking advantage of the tools and concepts of network science to quantify the general and local topological properties of gene co-expression networks, we used exact network-metrics to circumvent debates associated with modules definitions (Kleinberg 2002; Fortunato 2010; Habib and Paul 2010; de Montgolfier et al. 2011), rather than performing mainstream module-centered analyses of GCNs (Gaiteri et al. 2014; Meng et al. 2018; Chowdhury et al. 2020), typically conducted using WGCNA. Although module-centered analyses are important, our original approach based on exact metrics reported an increase in variation of overall gene co-expression, supported by a general pattern of network fragmentation, which could not have been trivially predicted through differential gene expression analyses. Indeed, differential expression is not readily correlated with differential co-expression: differential expression can be positively, negatively correlated or even uncorrelated with differential co-expression (Oldham et al. 2006; Fuente 2010; Lui et al. 2015; Zamora-Fuentes et al. 2020). Moreover, we observed that, while a core subnetwork seems to be robust during bat aging, other subnetworks appear age-specific. Thus, on the one hand, GCN comparisons identified which interactions are conserved during aging. This characterized some essential housekeeping processes that provide a basic life support to bats. On the other hand, GCN comparisons detected age-specific co-expression with similar associated biological functions, providing evidence of a re-wiring in the architecture of the bat interactome with age. The core subnetwork and the rewiring of GCN with age would also not have been predictable from differential gene expression studies alone (Ideker and Krogan 2012; Anglani et al. 2014; Gaiteri et al. 2014; Mähler et al. 2017; Meng et al. 2018; Zamora-Fuentes et al. 2020). Furthermore, we interpreted the size reduction and the loss of connectivity during aging in bats as the result of a weaker constraint exerted by natural selection later in life. This

interpretation is consistent with studies on gene co-expression connectivity and selection that suggest the existence of selective forces in the overall design of at least some genetic pathways to maintain a highly connected class of genes, associating high connectivity in GCNs with purifying and/or stabilizing selection (Stuart 2003; Oldham et al. 2006; Mähler et al. 2017). This interpretation was also reinforced by motif analysis (Milo et al. 2002; Shen-Orr et al. 2002), showing that the ratio of triangular motifs over all size 3 motifs significantly decrease in aging bats, a metrics compatible both with the existence of selection in GCNs of aging bats, and with the decrease of that selection with age.

Ancestral bats species were already long-lived (Wilkinson and Adams 2019) and specific adaptations have further enhanced the longevity of extant species, such as the Brandt's bat (*Myotis brandtii*), which can live up to 41 years, or the greater mouse-eared bat (*M. myotis*), which can live up to 37 years (Foley et al. 2018). While classic theories of aging predict that the strength of natural selection should decrease in older bats, and should consequently lead to a general dysregulation of individual transcriptomes with age, conversely the unusual longevity of *M. myotis* is traditionally explained by adaptations that counteract aging until late in their lives. Accordingly, longitudinal transcriptomes studies of free-living populations of *Myotis* bats (Huang et al. 2016) have characterized some genetic bases for such adaptations. Previous studies have shown that DNA repair and DNA damage signaling pathways are maintained throughout *M. myotis* lifespan, consistent with low levels of cancer (Huang et al. 2016), which indicates robust strategies for genome maintenance. Moreover, bats are also able to tolerate active DNA transposons, possibly through evolved dampened mechanisms of cytoplasmic DNA sensing. Bat species are also remarkably resistant to oxidative damage, thanks to efficient mitochondria, which produce less H<sub>2</sub>O<sub>2</sub> per unit oxygen consumed (Brunet-Rossinni 2004). Consistently, *M. myotis* presents low rates of age-independent oxidative lesions, which suggests that this species evolved efficient repair or removal of damaged mitochondria (Jebb et al. 2018). Overall, *M. myotis* does not show the same transcriptomic changes with age as commonly observed in humans and other mammal species, but rather exhibit a unique, age-related gene expression pattern associated with DNA repair, autophagy, immunity and tumor suppression, beneficial for the maintenance of cellular functions into old age and therefore suspected to drive their extended healthspans (Huang et al. 2019).

Our results are consistent with traditional empirical findings on gene expression and with evolutionary theories of aging. It is likely that gene co-expression, affected by random damage accumulation, becomes more stochastic with age within individual bats. This individual

stochasticity of gene co-expression results in heterogeneous gene co-expression within an age-class. Consequently, at the same age, individual bats co-express different sets of genes, and, for this reason, GCNs from older age-classes are more fragmented (i.e., more disconnected and of smaller sizes) than GCNs from younger age-classes, even though, based on the longevity record for *M. myotis*, the older age classes are still rather young for this species. Such a topological change is therefore compatible with a decrease of the strength of natural selection acting on older bats. We found that this decrease of the strength of natural selection seems to affect age-specific gene co-expressions, because the age-specific subnetworks present in the GCNs of aging bats were, in particular, markedly reduced in size with age. However, despite this signal of a weakening regulation of the biological processes with time, aspects of the GCN of older bats suggest that these organisms are still under sufficient selection to experience a transition towards a new, age-specific biology and to maintain a fraction of central biological programs that were already running in young adults.

A possible interpretation of these contrasted results (a fading yet present selection) is that older bats simultaneously experience: a decrease of the strength of natural selection and the robust resistance to this form of evolutionary neglect by a fraction of homogeneous gene co-expressions. If this is correct: after seven years, an individual bat is not simply progressively succumbing to wear and tear (and possibly failing to turn-off some of its ongoing programs) as natural selection recedes, this bat also seems to keep on adapting, even if less effectively, some of its biological processes.

This claim is compatible either with the existence of successive stages of homogeneous age-specific gene co-expressions, or with a biased subsampling of the bats population in the different age-classes. In the first case, whereas stochastic gene co-expression increases with age, some selection would then remain active as individual bats age. In the second (non mutually exclusive) case, older bats would constitute a non-representative subset of the younger bats population from which they arose. Namely, subpopulations of older bats would be enriched in individuals which are able to implement a particular gene co-expression, and to keep on performing critical biological functions that affect survival, because different aging individuals are constrained to rely on shared genetic co-expression, which support the same sets of biological functions. While within each individual aging appears largely stochastic, the existence of constraints imposing critical gene co-expression and critical functional implementation to survive, would explain that at the population level, aging is not completely random. As a result, increasingly older individuals would be genuinely different on many

accounts, and in particular would display heterogeneous gene co-expression (which results in smaller, more disconnected GCNs for increasingly old age-class and a higher proportion of singletons), but at the same time, these increasingly older bats would share some critical features without which they simply would not survive. In other words, there would be many ways for bats to age, but less ways to survive to aging.

Under both scenarios, older bats from long-lived species seem to experience a remarkable evolutionary crossroad: as they age, their long lives make them both experience a decrease in the strength of natural selection, yet their belonging to long-lived species makes them benefit from inherited adaptations that partly resist the raising neglect of natural selection with age, through a targeted selection for very specific biological processes or a biased selection for surviving older individuals, relying on similar gene co-expression to perform convergent critical biological functions.

Importantly, our approach is generic, and could be applied to any longitudinal transcriptomic dataset, from single-cells (Mendenhall et al. 2021) to tissues, to analyze how, when and why these cells, tissues, or their host organisms are aging. It would be fascinating to use this approach on long-lived individuals from long-lived species, as such organisms have plenty of opportunities not only to experiment aging but also to evolve adaptations counteracting aging and enhancing their longevity. A succession of distinct longevity and trade-off programs may well be running, under some form of selection, in the older individuals of long-lived species. Thus, interacting molecular components contributing to such processes may possibly be captured through comparisons of age-associated GCNs as shared late age-specific gene co-expression patterns. Beyond bats, it would be fascinating to use comparisons of longitudinal GCNs to analyze the causes of aging for many other exciting species (e.g., theoretically immortal/extremely long-lived species (Johnson et al. 2019) or humans, whose average lifespan has considerably expanded in the past century. This would allow one to determine whether gene co-expression variation increases with age in these species, compatible with a weakening of natural selection, or instead, maybe, to observe that their GCNs do not show many signs of such an increased heterogeneity, but possibly rather some plateau in the fragmentation of networks associated with increasingly old individuals, as their general topological properties and connectedness become stable past a given age.

## Conclusion

We presented an approach that can simultaneously show how organisms age (i.e., what gene co-expressions change over time, hence what potential biological processes are associated with aging), when organisms age (i.e., after which age, GCNs of older individuals present significant topological differences with respect to GCNs of younger individuals) and also offers topological clues on why they age (assuming that an increased fragmentation, i.e. a disconnectedness and size reduction of networks, and a depletion in triangular motifs, is compatible with a weakening of natural selection). Investigating gene co-expression in a long-lived bat species (*M. myotis*), we observed a general dysregulation of gene co-expression with age, compatible with a progressive weakening of natural selection, which is already detectable in younger bats (i.e., 3-4 years old). However, we also detected a succession of age-specific shared gene co-expression, indicating that the decrease in strength of natural selection was not detrimental to the evolution of adaptations, even reduced, as organisms age and are increasingly neglected by natural selection. Our results appear compatible with classic theories of aging predicting that natural selection on older organisms decreases in natural populations, but they also indicate that theoretical refinements are needed to account for the persistence of early biological processes and for the emergence of (likely noisy) biological processes later in life in a context where natural selection is fading, but remains active, in some cells and on some targets.

## Materials and methods

### Data collection and sampling of bats

We collected the gene expression counts of 12,263 protein-coding genes from 88 bats (*M. myotis*) from the original data published in (Huang et al. 2019), most kindly provided by these authors. The dataset is composed of individuals ranging from 0 to 7+ year old. Bats were aged as follows as described in (Huang et al. 2019). *M. myotis* individuals have been marked with unique transponders since 2010, and the initial ages of bats when first captured were determined by examining the epiphyseal cartilage in their finger bones. Individuals were recorded as juvenile (0 years old) if the epiphyseal plates in their finger bones were open; otherwise, they were recorded as adult (1+ years old, true age unknown). In our study, we focused on adult bats only (1+ year old) and removed individuals marked as 5 and 5+ due to a lack of samples to

allow robust network construction for this particular age-class, using WGCNA (Langfelder and Horvath 2008), which recommends a minimum of 15 samples per age-class to infer gene co-expression networks. Accordingly, blood samples were pooled into three age-classes: 1-2, 3-4 and 6+. We then randomly selected 16 samples for each category for a total of 48 samples. This protocol, resulting in younger, intermediate and older bats with comparable numbers of individuals, reduces sampling size biases between age-classes, while providing enough signal to reduce the risk of missing data.

### **Data collection and sampling of mice**

Publicly available whole peripheral blood Illumina transcriptomic datasets containing RNAseq data from wild-type control, C57BL/6 mice of documented age were collected from the Gene Expression Omnibus (GEO) repository (<https://www.ncbi.nlm.nih.gov/geo/>) to build a longitudinal mouse blood supertranscriptome for two age-classes corresponding to ages 2-3 and 4-9 months old (SI-20). After controlling for batch effects using the function `removeBatchEffect()` from the `limma` package, matrices of raw counts were generated after mapping reads from the selected datasets to the most recent reference genome assembly (GRCm39/mm39) using STAR (default parameters, genome indexes were recomputed for each class of read lengths, by setting `sjdbOverhang` to read length minus one, as advised), to calculate correlation coefficients and perform GCN analysis. Then, we performed similar analyses on these longitudinal blood transcriptomes from mice than we had conducted for bats.

### **Gene co-expression network inference for bats**

To infer the GCN of each age-class, we used the R packages DEseq2 (Love et al. 2014) and WGCNA (Langfelder and Horvath 2008). We first normalized the raw counts of the 12,263 protein-coding genes using the variance stabilizing transformation (VST, from the DEseq2 package), and used the normalized count matrix as input for WGCNA. We used the function `corAndPvalue()` to compute the Pearson correlations and associated Student p-values. In the GCN, nodes represent protein-coding genes and edges represent the Pearson correlation between two genes. For each age-class, we produced three GCNs using three minimum stringency thresholds: 0.8, 0.85, and 0.9 (to be able to detect any potential threshold effects). With such high thresholds, only strong correlations are preserved in the GCNs, thus reducing the size of the networks while removing uninformative signal.



## **Metrics of networks fragmentation**

For bats and mice, we compared the topology of the GCN for each age-class using standard network topological properties and seven different network metrics: 1) number of nodes, 2) number of edges, 3) node degrees, 4) node closeness, 5) local clustering coefficients, 6) the number of distinct connected components (i.e., disconnected subsets of connected nodes) within the network, and 7) the shortest paths between pairs of nodes. More precisely, the node degree is the number of edges connecting a node to its neighbors (SI-1, panel A). The normalized degree of a node is the degree of this node divided by the number of nodes of the component it belongs to minus 1, therefore ranging from 0 to 1. The shortest path (SI-1, panel B) is the shortest distance (i.e., the minimum number of edges) between any pair of nodes. The closeness centrality (formula in SI-1, panel C) measures how central a node is within a connected component (i.e., how close it is to the other nodes). Since our GCNs were composed of multiple connected components, the closeness of a node is only computed over the component it belongs to, without loss of generality. The clustering coefficient (formula in SI-1, panel D) measures how close from being a clique (a complete graph) are the neighbors of a node (i.e., the connectivity of a node's neighborhood). These metrics were used to quantify different aspects of the fragmentation within each network and between networks. Thus, networks were said to be increasingly fragmented with age, when their constitutive nodes present less direct neighbors (node degree decreases), when their nodes become less central (node closeness decreases), when the neighbors of each of their nodes become less interconnected (local clustering coefficient decreases), when the shortest distances between pairs of their nodes increase, and when these networks present an increasing number of connected components with respect to networks associated with younger age-classes. Moreover, in addition to these metrics of disconnectedness, we also consider that networks were increasingly fragmented when they were smaller (i.e., presenting less nodes and less edges at increasing age-classes). Therefore, what we used to define fragmentation were general properties of network disconnectedness and network sizes; and the term of fragmentation does not therefore necessarily imply that fragmented networks from older age-classes are built from subsets of nodes and of edges from networks from younger age-classes. To compute these topological properties and metrics, we used an in-house Java program based on the Java library Grph (Hogie and et al. 2014).

## **Statistical analyses on metrics distribution**

The R package `ggplot2` was used to generate the distributions of network metrics for the GCNs of each age-class, at three stringency thresholds, for bats and mice, respectively. Statistical significance of the differences between these metric distributions, at each stringency threshold, was assessed using the non-parametric Wilcoxon rank-sum test and corrected by the Bonferroni method for multiple tests.

### **Statistical analyses on motifs**

For bats or mice GCNs (i.e., for each age-class at the three stringency thresholds), we used the R package `igraph` to enumerate all possible motifs of size three (Csardi and Nepusz 2006) and computed the ratio (`ratioR`) of triangle motifs divided by the sum of the three other types of disconnected motifs. Using the same package, we generated for each GCN 999 random networks with the same number of nodes and the same degree distribution as in the original GCN. Using the distribution of motifs in random GCNs, two distinct statistical comparisons among age-classes were performed. First, we tested the hypothesis that the `ratioR` of the original GCN is significantly larger than the `ratioR` of random GCNs. Second, we computed, at each stringency threshold, three values to compare `ratioR` at different age-classes (i.e.,  $x_{12} = (\text{ratioR}_{\text{young GCN}} - \text{ratioR}_{\text{intermediate GCN}})$ ;  $x_{13} = (\text{ratioR}_{\text{young GCN}} - \text{ratioR}_{\text{old GCN}})$ ;  $x_{23} = (\text{ratioR}_{\text{intermediate GCN}} - \text{ratioR}_{\text{old GCN}})$ ) and tested the hypothesis that these values ( $x_{12}$ ,  $x_{13}$  and  $x_{23}$ ) in the original GCNs are significantly larger than the corresponding values computed from random GCNs.

### **Core subnetworks and age-specific subnetworks**

For bats and for mice, we used an in-house Java script to achieve pairwise comparisons of GCNs, generated at a given stringency threshold. We define as ‘core nodes’ and ‘core edges’, the nodes and edges shared by the GCNs from the younger, intermediate and older age-classes. Nodes exclusively associated with a GCN from a single age-class were labeled as specific for that age-class. Nodes shared by a pair of networks were labeled as present in these two age-classes. In addition, we implemented specific labels for edges shared by young and intermediate bats, and for edges shared by intermediate and older bats. The absolute sizes of all connected components of all subnetworks composed of core-only or age-specific-only nodes and edges (obtained from GCN built for each age-class at all stringency thresholds) were measured in number of edges. Moreover, these absolute sizes were normalized by dividing the number of edges of each connected component of a subnetwork of interest (e.g., all connected components exclusively found in older bats) by the total number of edges in the GCN. This normalization

takes into account the fact that GCNs from aging bats have a reduced size and a reduced connectedness with respect to GCNs from bats from younger age-classes, and therefore may display smaller age-specific connected components for this reason. The absolute and normalized size distributions of age-specific connected components, resulting from these computations, were respectively compared across age-classes using a Wilcoxon rank-sum test corrected by the Bonferroni correction to assess whether some age-classes were comprised of significantly smaller age-specific connected components.

### **Functional enrichment analysis**

For bats and mice, co-expressed genes involved in core interactions, co-expressed genes not involved in core interactions and co-expressed genes associated with a given age-specific interactions, were respectively binned by age-class and functionally annotated using KOG (Tatusov et al. 2003). This approach returned a general functional profile that was represented as a bar plot presenting the proportion of age-specific and core co-expressed genes for distinct KOG categories. Distributions of these KOG categories were compared using pairwise Z test between subgroups to identify statistical differences in the amount of co-expressed genes involved in interactions at different ages or with robust co-expression in all age-classes. The functional ontology of these classes of co-expressed genes were analyzed with greater accuracy by performing functional enrichment analyses to identify enriched functions of age-specific and core co-expressed genes, within each KOG category when a large number of genes (> 3000) were analyzed, or in bulk for bat core genes, using Metascape (Zhou et al. 2019). Moreover, genes expressed at more than 10 total read counts but absent from gene coexpression networks, were defined as ‘singletons’, and ‘non-core co-expressed genes’ were defined as the sum of co-expressed genes not involved in core interactions and co-expressed genes associated with a given age-specific interactions. Both singletons and ‘non-core co-expressed genes’ were functionally analyzed similarly as above.

### **Acknowledgments and funding sources**

This work was supported by the overheads of ERC project, FP7/2007-2013 Grant Agreement # 615274, which supported GB and JT salaries.

## Bibliography

- Angelidis I, Simon LM, Fernandez IE, Strunz M, Mayr CH, Greiffo FR, Tsitsiridis G, Ansari M, Graf E, Strom T-M, et al. 2019. An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nat Commun* 10:963.
- Anglani R, Creanza TM, Liuzzi VC, Piepoli A, Panza A, Andriulli A, Ancona N. 2014. Loss of Connectivity in Cancer Co-Expression Networks. *PLOS ONE* 9:e87075.
- Austad SN. 2010. Cats, “Rats,” and Bats: The Comparative Biology of Aging in the 21st Century. *Integrative and Comparative Biology* 50:783–792.
- Bahar R, Hartmann CH, Rodriguez KA, Denny AD, Busuttill RA, Dollé MET, Calder RB, Chisholm GB, Pollock BH, Klein CA, et al. 2006. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* 441:1011–1014.
- Baumgart M, Priebe S, Groth M, Hartmann N, Menzel U, Pandolfini L, Koch P, Felder M, Ristow M, Englert C, et al. 2016. Longitudinal RNA-Seq Analysis of Vertebrate Aging Identifies Mitochondrial Complex I as a Small-Molecule-Sensitive Modifier of Lifespan. *Cell Systems* 2:122–132.
- Blagosklonny MV. 2006. Aging and Immortality: Quasi-Programmed Senescence and Its Pharmacologic Inhibition. *Cell Cycle* 5:2087–2102.
- Brunet-Rossinni AK. 2004. Reduced free-radical production and extreme longevity in the little brown bat (*Myotis lucifugus*) versus two non-flying mammals. *Mech Ageing Dev* 125:11–20.
- Chowdhury HA, Bhattacharyya DK, Kalita JK. 2020. (Differential) Co-Expression Analysis of Gene Expression: A Survey of Best Practices. *IEEE/ACM Trans Comput Biol Bioinform* 17:1154–1173.
- Cohen AA, Kennedy BK, Anglas U, Bronikowski AM, Deelen J, Dufour F, Ferbeyre G, Ferrucci L, Franceschi C, Frasca D, et al. 2020. Lack of consensus on an aging biology paradigm? A global survey reveals an agreement to disagree, and the need for an interdisciplinary framework. *Mech Ageing Dev* 191:111316.
- Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJournal, Complex Systems*:1695.
- Dollé MET, Snyder WK, Gossen JA, Lohman PHM, Vijg J. 2000. Distinct spectra of somatic mutations accumulated with age in mouse heart and small intestine. *PNAS* 97:8403–8408.
- Dollé MET, Vijg J. 2002. Genome Dynamics in Aging Mice. *Genome Res* 12:1732–1738.
- Enge M, Arda HE, Mignardi M, Beausang J, Bottino R, Kim SK, Quake SR. 2017. Single-Cell Analysis of Human Pancreas Reveals Transcriptional Signatures of Aging and Somatic Mutation Patterns. *Cell* 171:321-330.e14.
- Ferreira M, Francisco S, Soares AR, Nobre A, Pinheiro M, Reis A, Neto S, Rodrigues AJ, Sousa N, Moura G, et al. 2021. Integration of segmented regression analysis with weighted gene correlation network analysis identifies genes whose expression is remodeled throughout physiological aging in mouse tissues. *Aging (Albany NY)* 13:18150–18190.
- Finch CE. 1990. Longevity, senescence, and the genome. Chicago, Ill.: Univ. of Chicago Press

- Flurkey K, M. Curren J, Harrison DE. 2007. Chapter 20 - Mouse Models in Aging Research. In: Fox JG, Davisson MT, Quimby FW, Barthold SW, Newcomer CE, Smith AL, editors. *The Mouse in Biomedical Research (Second Edition)*. Burlington: Academic Press. p. 637–672. Available from: <https://www.sciencedirect.com/science/article/pii/B9780123694546500741>
- Foley NM, Hughes GM, Huang Z, Clarke M, Jebb D, Whelan CV, Petit EJ, Touzalin F, Farcy O, Jones G, et al. 2018. Growing old, yet staying young: The role of telomeres in bats' exceptional longevity. *Sci Adv* 4:eao0926.
- Fortunato S. 2010. Community detection in graphs. *Physics Reports*:75–174.
- Fuente A de la. 2010. From 'differential expression' to 'differential networking' – identification of dysfunctional regulatory networks in diseases. *Trends in Genetics* 26:326–333.
- Gaiteri C, Ding Y, French B, Tseng GC, Sibille E. 2014. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders: Beyond modules and hubs. *Genes, Brain and Behavior* 13:13–24.
- Galbraith PR, Valberg LS, Brown M. 1965. PATTERNS OF GRANULOCYTE KINETICS IN HEALTH, INFECTION AND IN CARCINOMA. *Blood* 25:683–692.
- Garbino GST, Feijó A, Beltrão-Mendes R, Da Rocha PA. 2021. Evolution of litter size in bats and its influence on longevity and roosting ecology. *Biological Journal of the Linnean Society* 132:676–684.
- Gorbunova V, Seluanov A, Kennedy BK. 2020. The World Goes Bats: Living Longer and Tolerating Viruses. *Cell Metab* 32:31–43.
- Habib M, Paul C. 2010. A survey of the algorithmic aspects of modular decomposition. *Comput. Sci. Rev.*:41–59.
- Harman D. 1956. Aging: a theory based on free radical and radiation chemistry. *J Gerontol* 11:298–300.
- Hogje L, et al. 2014. Grph, the high performance graph library for Java (Version 1.4.17). *The Grph Development Team*. [Internet]. Available from: <http://grph.inria.fr>.
- Huang Z, Jebb D, Teeling EC. 2016. Blood miRNomes and transcriptomes reveal novel longevity mechanisms in the long-lived bat, *Myotis myotis*. *BMC Genomics* 17:906.
- Huang Z, Whelan CV, Dechmann D, Teeling EC. 2020. Genetic variation between long-lived versus short-lived bats illuminates the molecular signatures of longevity. *Aging (Albany NY)* 12:15962–15977.
- Huang Z, Whelan CV, Foley NM, Jebb D, Touzalin F, Petit EJ, Puechmaile SJ, Teeling EC. 2019. Longitudinal comparative transcriptomics reveals unique mechanisms underlying extended healthspan in bats. *Nat Ecol Evol* 3:1110–1120.
- Ideker T, Krogan NJ. 2012. Differential network biology. *Mol Syst Biol* 8:565.
- Jebb D, Foley NM, Whelan CV, Touzalin F, Puechmaile SJ, Teeling EC. 2018. Population level mitogenomics of long-lived bats reveals dynamic heteroplasmy and challenges the Free Radical Theory of Ageing. *Sci Rep* 8:13634.

- Johnson AA, Shokhirev MN, Shoshitaishvili B. 2019. Revamping the evolutionary theories of aging. *Ageing Res Rev* 55:100947.
- Kacprzyk J, Locatelli AG, Hughes GM, Huang Z, Clarke M, Gorbunova V, Sacchi C, Stewart GS, Teeling EC. 2021. Evolution of mammalian longevity: age-related increase in autophagy in bats compared to other mammals. *Aging (Albany NY)* 13:7998–8025.
- Kapusta A, Suh A, Feschotte C. 2017. Dynamics of genome size evolution in birds and mammals. *Proc Natl Acad Sci U S A* 114:E1460–E1469.
- Kenyon CJ. 2010. The genetics of ageing. *Nature* 464:504–512.
- Kimmel JC, Penland L, Rubinstein ND, Hendrickson DG, Kelley DR, Rosenthal AZ. 2019. Murine single-cell RNA-seq reveals cell-identity- and tissue-specific trajectories of aging. *Genome Res* 29:2088–2103.
- Kirkwood T, Holliday R. 1979. The evolution of ageing and longevity. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 205:531–546.
- Kirkwood TB. 1977. Evolution of ageing. *Nature* 270:301–304.
- Kirkwood TB, Rose MR. 1991. Evolution of senescence: late survival sacrificed for reproduction. *Philos Trans R Soc Lond B Biol Sci* 332:15–24.
- Kirkwood TBL. 2005. Understanding the odd science of aging. *Cell* 120:437–447.
- Kleinberg JM. 2002. An Impossibility Theorem for Clustering. In: *Advances in Neural Information Processing Systems 15*. Suzanna Becker and Sebastian Thrun and Klaus Obermayer (Eds.). Vancouver, British Columbia, Canada: MIT Press. p. 446–453.
- Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.
- Lathe R. 2004. The individuality of mice. *Genes Brain Behav* 3:317–327.
- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15:550.
- Lui TW, Tsui NB, Chan LW, Wong CS, Siu PM, Yung BY. 2015. DECODE: an integrated differential co-expression and differential expression analysis of gene expression data. *BMC Bioinformatics* 16:182.
- Mähler N, Wang J, Terebieniec BK, Ingvarsson PK, Street NR, Hvidsten TR. 2017. Gene co-expression network connectivity is an important determinant of selective constraint. *PLoS Genet* 13:e1006402.
- Martinez-Jimenez CP, Eling N, Chen H-C, Vallejos CA, Kolodziejczyk AA, Connor F, Stojic L, Rayner TF, Stubbington MJT, Teichmann SA, et al. 2017. Aging increases cell-to-cell transcriptional variability upon immune stimulation. *Science* 355:1433–1436.
- Mayr E. 1961. Cause and effect in biology. *Science* 134:1501–1506.
- Medawar PB. 1952. An unsolved problem of biology. College

- Mendenhall AR, Martin GM, Kaeberlein M, Anderson RM. 2021. Cell-to-cell variation in gene expression and the aging process. *Geroscience*.
- Meng S, Liu G, Su L, Sun L, Wu D, Wang L, Zheng Z. 2018. Functional clusters analysis and research based on differential coexpression networks. *Biotechnology & Biotechnological Equipment* 32:171–182.
- Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D, Alon U. 2002. Network motifs: simple building blocks of complex networks. *Science* 298:824–827.
- de Montgolfier F, Soto M, Viennot L. 2011. Asymptotic Modularity of Some Graph Classes. In: Lecture Notes in Computer Science. Takao Asano and Shin-Ichi Nakano and Yoshio Okamoto and Osamu Watanabe (Eds.). Algorithms and Computation - 22nd International Symposium. Yokohama, Japan: Springer. p. 435–444.
- Munshi-South J, Wilkinson GS. 2010. Bats and birds: Exceptional longevity despite high metabolic rates. *Ageing Res Rev* 9:12–19.
- Nussey DH, Froy H, Lemaître J-F, Gaillard J-M, Austad SN. 2013. Senescence in natural populations of animals: Widespread evidence and its implications for bio-gerontology. *Ageing Research Reviews* 12:214–225.
- Oldham MC, Horvath S, Geschwind DH. 2006. Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *PNAS* 103:17973–17978.
- Pacifico R, MacMullen CM, Walkinshaw E, Zhang X, Davis RL. 2018. Brain transcriptome changes in the aging *Drosophila melanogaster* accompany olfactory memory performance deficits. Skoulakis EMC, editor. *PLoS ONE* 13:e0209405.
- Reichard M. 2017. Evolutionary perspectives on ageing. *Seminars in Cell & Developmental Biology* 70:99–107.
- Shen-Orr SS, Milo R, Mangan S, Alon U. 2002. Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat Genet* 31:64–68.
- Somel M, Guo S, Fu N, Yan Z, Hu HY, Xu Y, Yuan Y, Ning Z, Hu Y, Menzel C, et al. 2010. MicroRNA, mRNA, and protein expression link development and aging in human and macaque brain. *Genome Research* 20:1207–1218.
- Southworth LK, Owen AB, Kim SK. 2009. Aging Mice Show a Decreasing Correlation of Gene Expression within Genetic Modules. Copenhaver GP, editor. *PLoS Genet* 5:e1000776.
- Stone L, Simberloff D, Artzy-Randrup Y. 2019. Network motifs and their origins. *PLoS Comput Biol* 15:e1006749.
- Stuart JM. 2003. A Gene-Coexpression Network for Global Discovery of Conserved Genetic Modules. *Science* 302:249–255.
- Tarkhov AE, Alla R, Ayyadevara S, Pyatnitskiy M, Menshikov LI, Shmookler Reis RJ, Fedichev PO. 2019. A universal transcriptomic signature of age reveals the temporal scaling of *Caenorhabditis elegans* aging trajectories. *Sci Rep* 9:7368.

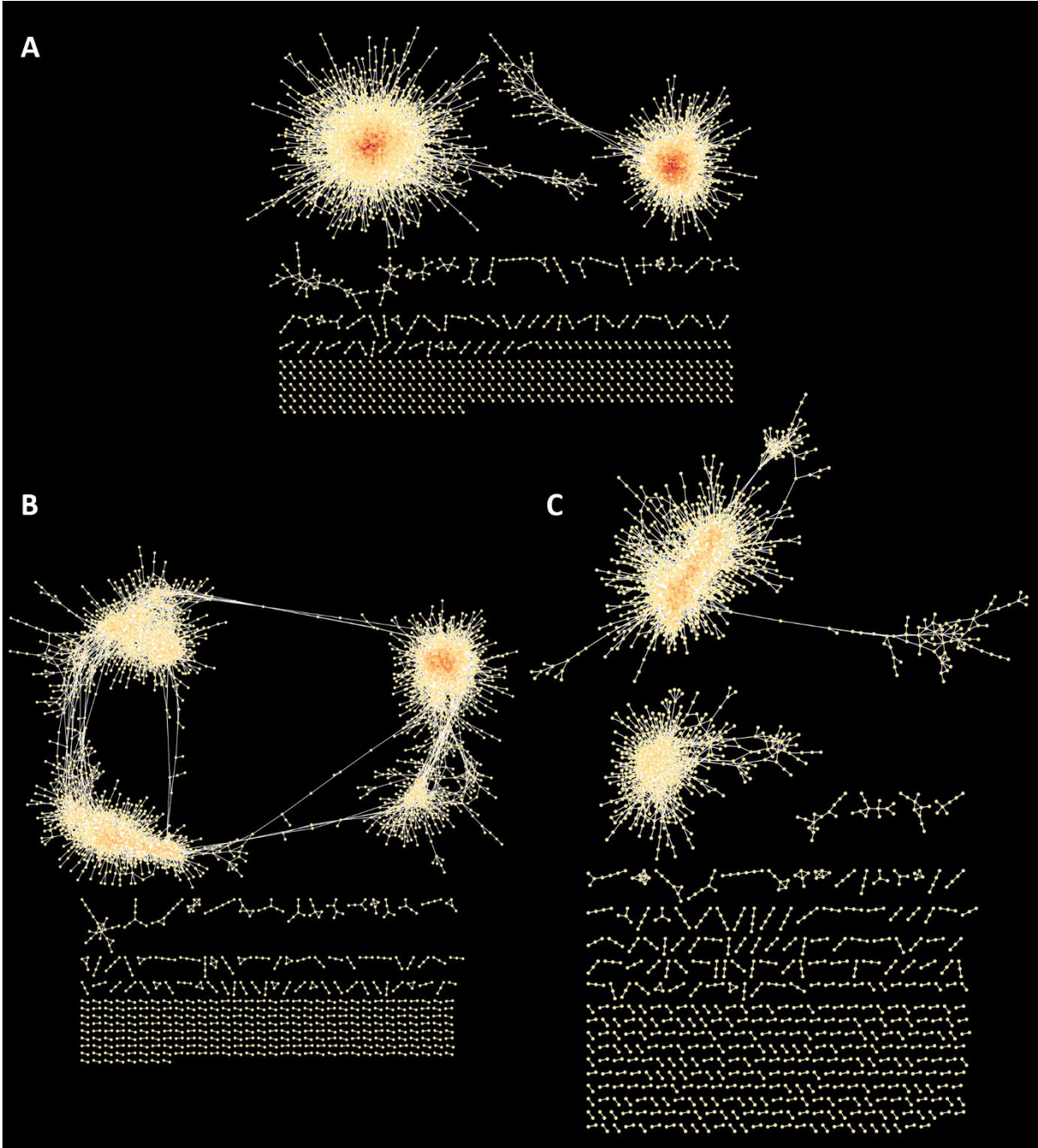
- Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, Krylov DM, Mazumder R, Mekhedov SL, Nikolskaya AN, et al. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* 4:41.
- Vijg J. 2007. *Aging of the Genome*. Oxford, UK: Oxford University Press.
- Voelkl B, Altman NS, Forsman A, Forstmeier W, Gurevitch J, Jaric I, Karp NA, Kas MJ, Schielzeth H, Van de Castele T, et al. 2020. Reproducibility of animal research in light of biological variation. *Nat Rev Neurosci* 21:384–393.
- Warren LA, Rossi DJ, Schiebinger GR, Weissman IL, Kim SK, Quake SR. 2007. Transcriptional instability is not a universal attribute of aging. *Aging Cell* 6:775–782.
- Wilkinson GS, Adams DM. 2019. Recurrent evolution of extreme longevity in bats. *Biol. Lett.* 15:20180860.
- Wilkinson GS, Adams DM, Haghani A, Lu AT, Zoller J, Breeze CE, Arnold BD, Ball HC, Carter GG, Cooper LN, et al. 2021. DNA methylation predicts age and provides insight into exceptional longevity of bats. *Nat Commun* 12:1615.
- Williams GC. 1957. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11:398–411.
- Zamora-Fuentes JM, Hernández-Lemus E, Espinal-Enríquez J. 2020. Gene Expression and Co-expression Networks Are Strongly Altered Through Stages in Clear Cell Renal Carcinoma. *Front Genet* 11:578679.
- Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. 2019. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* 10:1523.

## Figure legends

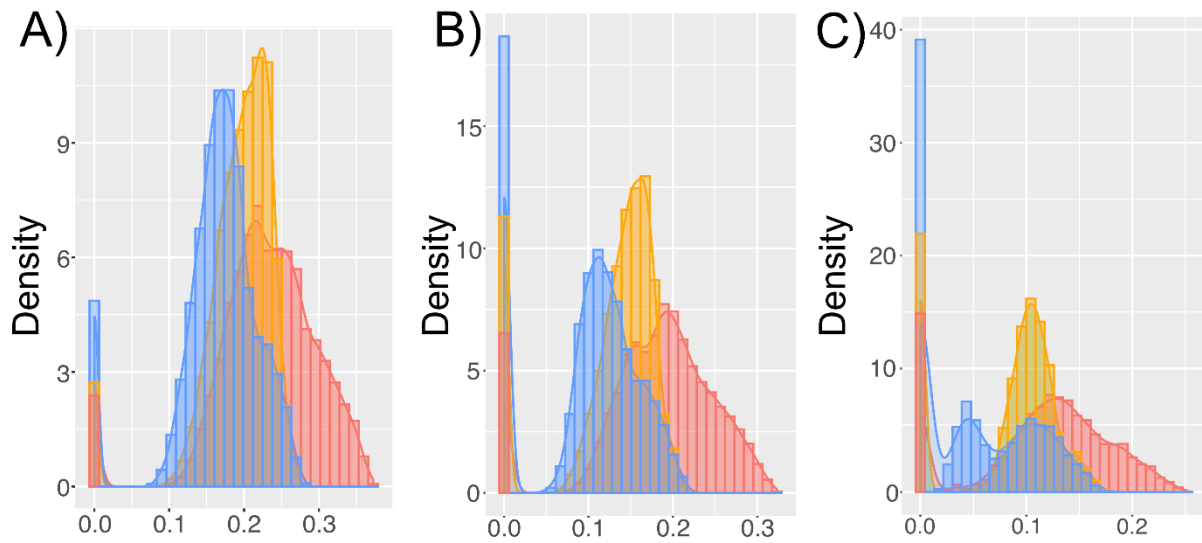
### Figure 1. GCNs for three age-classes with normalized degree values

All these GCNs are built for p-value < 0.05 and Pearson coefficient > 0.9. Panel A represents the GCN for the younger bats. Panel B represents the GCN for the intermediate bats. Panel C represents the GCN for the older bats. The node degree (the number of direct edges connecting a node to others) is normalized by dividing the degree of each node by the total number of nodes - 1 within each network. Nodes are colored based on their normalized degree value (from white, lowest degree to red highest degree), using the same scale for all networks.





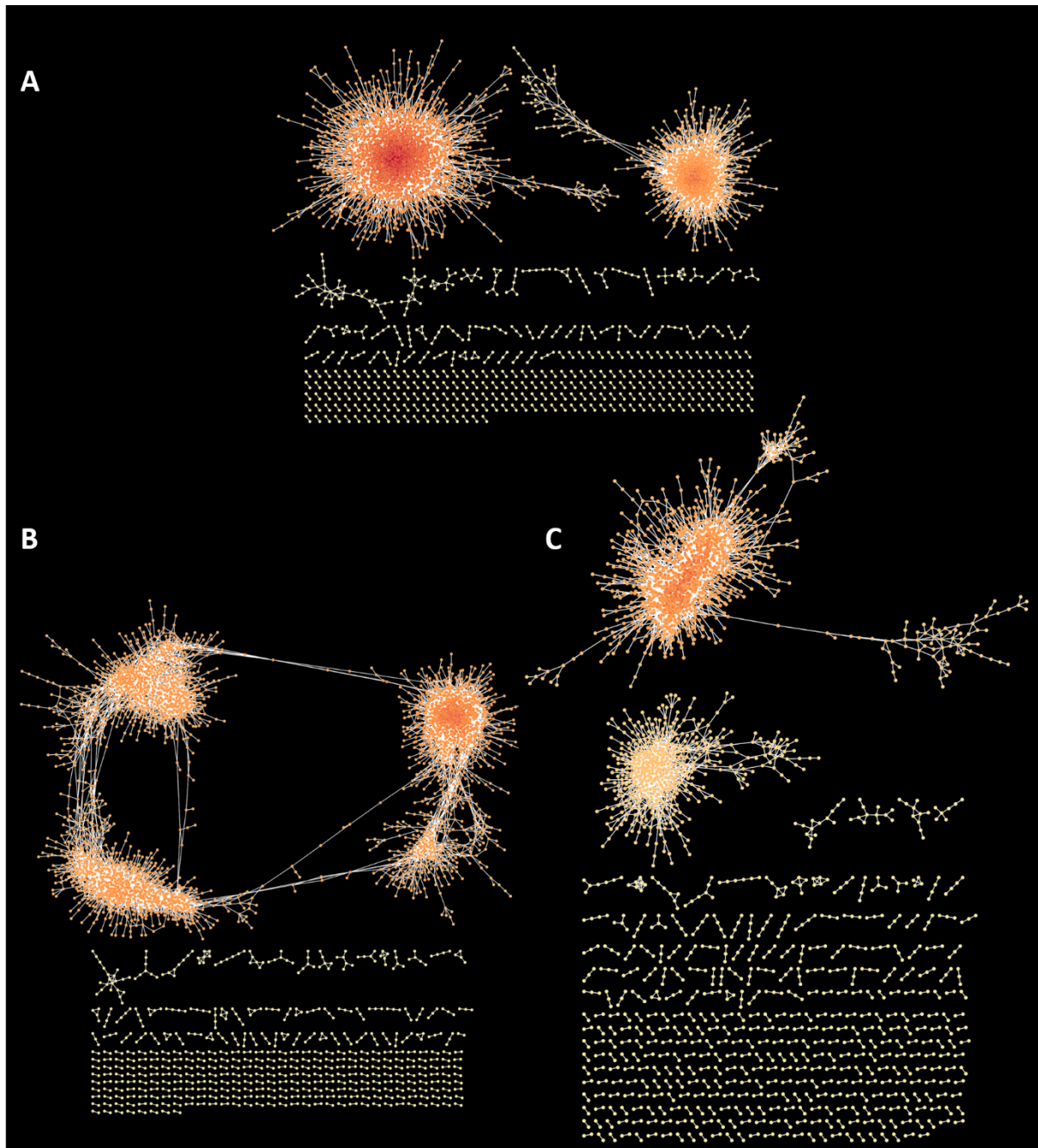
**Figure 2. Distribution of closeness values for GCN from different age-classes, at different stringency levels**



X-axis values correspond to closeness values. Y-axis values correspond to the kernel density estimation of the histogram computed by the R package ggplot2. Panel A shows the histogram of frequencies for nodes in GCNs built for p-value  $< 0.05$  and Pearson coefficient  $> 0.8$ . Panel B shows the histogram of frequencies for nodes in GCNs built for p-value  $< 0.05$  and Pearson coefficient  $> 0.85$ . Panel C shows the histogram of frequencies for nodes in GCNs built for p-value  $< 0.05$  and Pearson coefficient  $> 0.9$ . Blue bars indicate values for older bats, orange bars indicate values for intermediate bats, red bars indicate values for younger bats.

**Figure 3. GCNs for three age-classes with closeness values**

All these GCNs are built for p-value  $< 0.05$  and Pearson coefficient  $> 0.9$ . Panel A represents the GCN for the younger bats. Panel B represents the GCN for the intermediate bats. Panel C represents the GCN for the older bats. The closeness value (a measure of a node centrality, and distance to any other nodes in the network) is bounded between 0 and 1. Nodes are colored based on their closeness value (from white, lowest closeness value, to red highest degree), using the same scale for all networks.

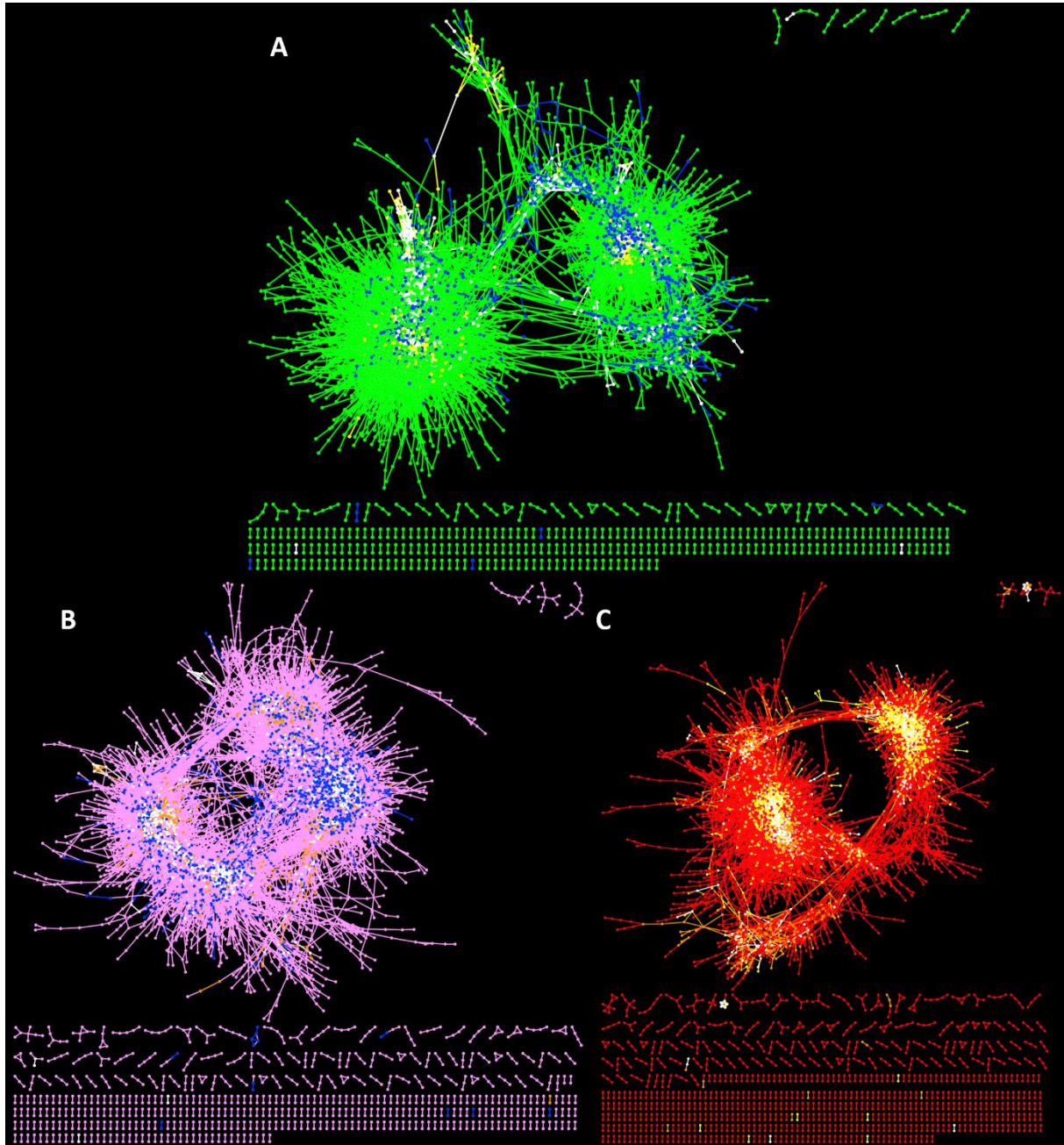


**Fig. 4 Core and age-specific interactions within GCNs from younger, intermediate, and older bats**

All these GCNs are built for  $p$ -value  $< 0.05$  and Pearson coefficient  $> 0.8$ . Panel A represents the GCN for the younger bats. Panel B represents the GCN for the intermediate bats. Panel C represents the GCN for the older bats. Edge and node colors indicate with which age-class the co-expression of two genes is associated: white for core, co-expressed genes present in all age-classes, green for co-expressed genes exclusively associated with younger bats, pink for co-expressed genes exclusively associated with intermediate bats, blue for co-expressed genes



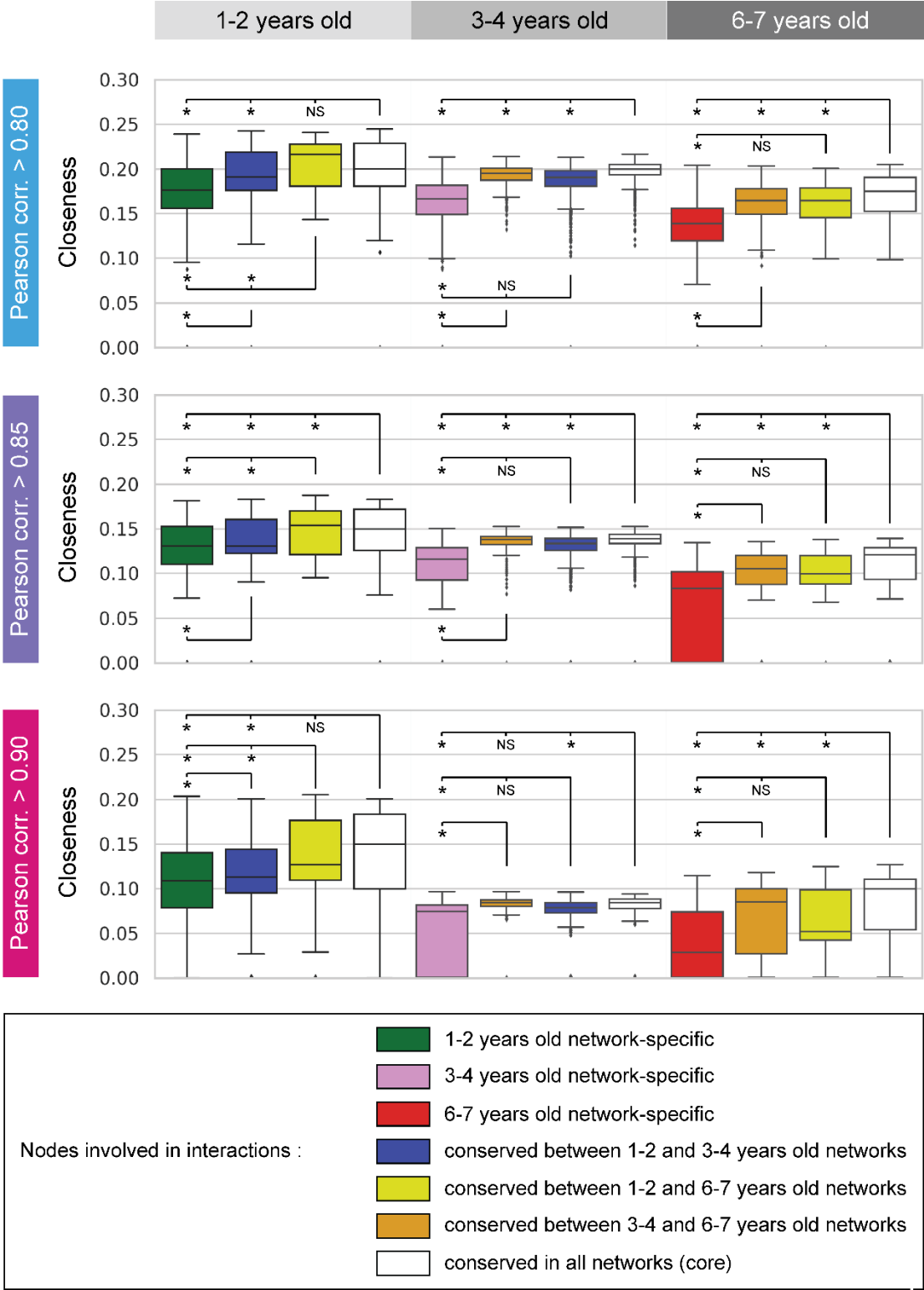
associated with both younger and intermediate bats, red for co-expressed genes exclusively associated with older bats, orange for co-expressed genes associated both with intermediate and older bats.



**Figure 5. Closeness distribution in age-specific subnetworks in bats**

For each Pearson correlation threshold and each age-class, the distribution of the closeness centrality metric is displayed for each age-specific network in box plots, colored as in Figure 3 and as indicated in the legend. Distributions were compared using the unilateral Mann-Whitney U-test to determine whether subnetworks indicated by the longest bracket branch had

significantly higher values of closeness. \* : closeness was significantly higher than the indicated distribution ( $P < 0.05$ ); NS : closeness was not significantly higher.



**Table 1. Summary of the analyzed GCN topological properties**

Correlation threshold	Age-classes	Number of nodes	Number of edges	Number of Connected componentss
0.80	1-2 y old	11013	932579	142
	3-4 y old	10956	255767	164
	6+ y old	10233	157058	264
0.85	1-2 y old	8730	377443	288
	3-4 y old	8006	95315	417
	6+ y old	6801	50402	592
0.9	1-2 y old	5404	92685	291
	3-4 y old	4074	21850	335
	6+ y old	2770	9438	397