Inter-tissue convergence of gene expression during ageing suggests age-related loss of tissue and cellular identity

Hamit Izgi[1], DingDing Han[2,+], Ulas Isildak[1], Shuyun Huang[2,], Ece Kocabiyik[1], Philipp Khaitovich[3*], Mehmet Somel[1*], Handan Melike Dönertaş[4,5*]

[1] Department of Biological Sciences, Middle East Technical University, Ankara, Turkey

[2] CAS Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

[3] Center for Neurobiology and Brain Restoration, Skolkovo Institute of Science and Technology, Moscow, Russia

[4] European Molecular Biology Laboratory, European Bioinformatics Institute EMBL-EBI, Wellcome Trust Genome Campus, Cambridge, UK

[5] Leibniz Institute on Aging - Fritz Lipmann Institute (FLI), Beutenbergstraße 11, 07745, Jena, Germany

[+] present address: Department of Clinical Laboratory, Shanghai Children's Hospital, Shanghai Jiaotong University, Shanghai, China

*Correspondence: melike.donertas@leibniz-fli.de (H.M.D), msomel@metu.edu.tr (M.S.), P.Khaitovich@skoltech.ru (P.K)

Abstract (124/150 words)

Developmental trajectories of gene expression may reverse in their direction during ageing, a phenomenon previously linked to cellular identity loss. Our analysis of cerebral cortex, lung, liver and muscle transcriptomes of 16 mice, covering development and ageing intervals, revealed widespread but tissue-specific ageing-associated expression reversals. Cumulatively, these reversals create a unique phenomenon: mammalian tissue transcriptomes diverge from each other during postnatal development, but during ageing, they tend to converge towards similar expression levels, a process we term Divergence followed by Convergence, or DiCo. We found that DiCo was most prevalent among tissue-specific genes and associated with loss of tissue identity, which is confirmed using data from independent mouse and human datasets. Further, using publicly available single-cell transcriptome data, we showed that DiCo could be driven both by alterations in tissue cell type composition and also by cell-autonomous expression changes within particular cell types.

Keywords

Ageing, development, transcriptome, mouse, reversal

Introduction

Development and ageing in multicellular organisms are highly intertwined processes. On the one hand, certain ageing-related phenotypes, such as presbyopia and osteoporosis (Luegmayr et al. 2004) are believed to represent the continuation of developmental processes into adulthood (Blagosklonny 2006; de Magalhães and Church 2005)). Such cases of "runaway development" or higher than optimal function during ageing (recognized as the hyperfunction theory of ageing (Gems and Partridge 2013)), may arise due to declined natural selection pressure failing to optimise expression regulation after sexual reproduction starts (Fisher, 1930; Medawar, 1953; Williams, 1957). Indeed, recent experimental studies in *C. elegans* show that senescence phenotypes promoted by insulin-IGF-1 signalling pathways support the hyperfunction theory (Lind et al. 2019; Ezcurra et al. 2018)). On the other hand, molecular studies have also reported a reversal of the ageing transcriptome towards pre-adult levels in various contexts, including primate brain regions (Somel et al. 2010; Dönertaş et al. 2017; Colantuoni et al. 2011), and mouse liver and kidney (Anisimova et al. 2020). Studying the functional consequences of this reversal pattern in the ageing human brain, we previously interpreted it as an indication of loss of cellular identity in neurons, possibly exacerbated by a reduction in the relative frequencies of neurons (Dönertaş et al. 2017). Such changes, in turn, could be caused by the accumulation of stochastic damage at the genetic, epigenetic, and proteomic levels over an adult lifetime, causing deregulation of gene expression networks.

Several major questions remain. First, the prevalence of reversal phenotypes across tissues is unclear, as most research has been conducted in the brain (Somel et al. 2010; Dönertaş et al. 2017). A second question pertains to the similarity of reversal-exhibiting genes and pathways across tissues. Ageing-related expression changes are partly shared among organs (Zahn et al. 2007), and reversal trends are also shared across different regions of the primate brain (Dönertaş et al. 2017). Distinct tissues might hence show parallel reversal patterns. Alternatively, as mammalian tissues diverge from each other during development in their transcriptome profiles (Cardoso-Moreira et al. 2019), one may hypothesise that during ageing, tissues converge back toward similar transcriptome profiles. Such a putative late-age convergence phenomenon would be consistent with the notion of ageing-related cellular identity loss (Yang et al. 2019; Dönertaş et al. 2017). A final question concerns the mechanism behind the observed reversal trends at the bulk tissue level. Specifically, the contribution of cell type composition

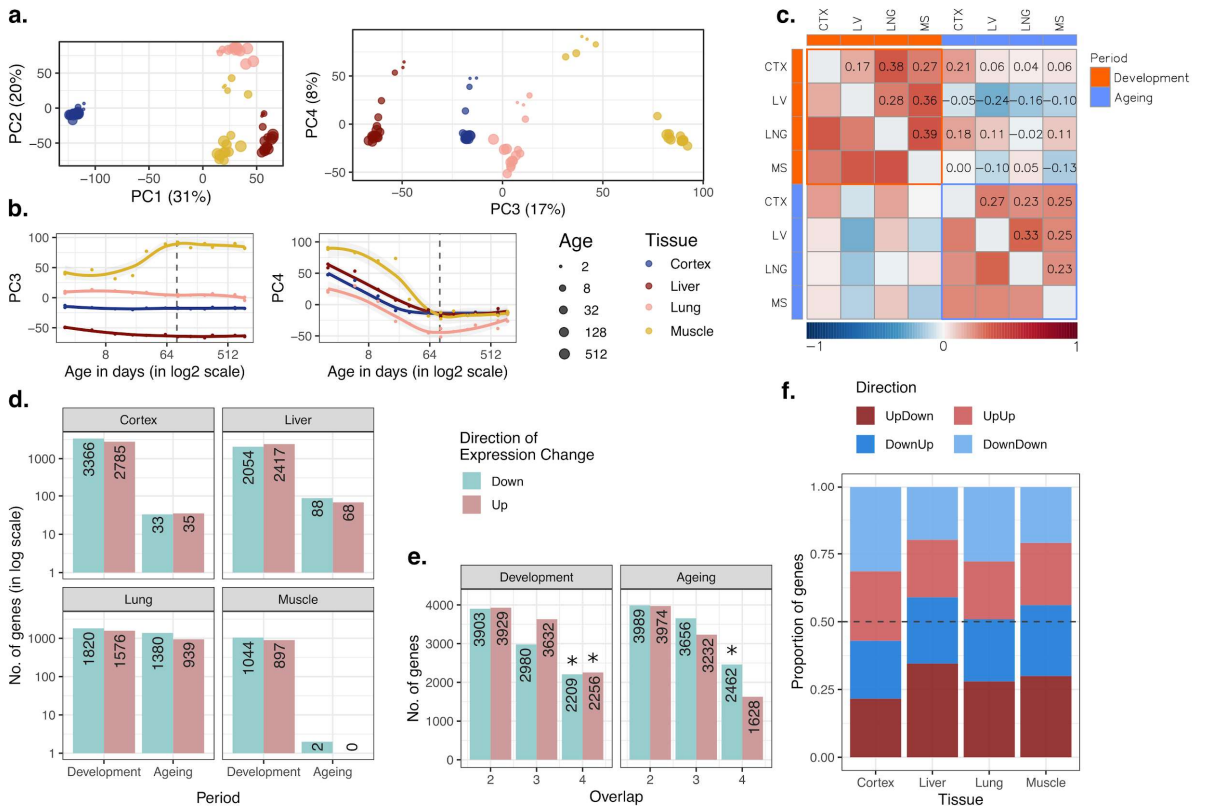
and cell-autonomous changes to the reversals at the tissue level remains unexplored.

Documenting the reversal phenomenon is critical to better understand the proximate mechanisms of mammalian ageing, and its ultimate mechanisms, such as the stochastic disruption versus continued expression of developmental genes. However, such work has been limited by the scarcity of studies that include both development and ageing periods of the same organism and across different tissues. This work presents an age-series analysis of bulk transcriptome profiles of mice, including samples of four tissues across postnatal development and ageing periods covering the whole postnatal lifespan. Using this dataset, we study the prevalence, mechanisms, and functional consequences of the reversal phenomenon in different mouse tissues. We further test the related hypothesis of tissue convergence during ageing and investigate the contribution of cell type composition and cell-autonomous changes.

Results

We generated bulk RNA-seq data from 63 samples covering the cerebral cortex (which we refer to as cortex), liver, lung, and skeletal muscle (which we refer to as muscle) of 16 male C57BL/6J mice, aged between 2 to 904 days of postnatal age (Methods). As mice reach sexual maturity by around two months (Tacutu et al. 2018), we treated samples from individuals aged between 2 and 61 days (n=7) as the development series, and those aged between 93 and 904 days (which roughly correspond to 80-year-old humans (Flurkey, M. Currer, and Harrison 2007)) (n=9) as the ageing series **(Figure 1-figure supplement 1)**. The final dataset contained n=15,063 protein-coding genes expressed in at least 25% of the 63 samples (one 904 days old mouse lacked cortex data).

Figure 1. Data summary and age-related expression patterns



a) *Principal components analysis (PCA) of expression levels of 15,063 protein-coding genes across four tissues of 16 mice. Values in parentheses show the variation explained by each component.* ***b)*** *Age trajectories of PC3 (left) and PC4 (right). Spearman's correlation coefficients between PC4 and age in each tissue in development range between 0.88 and 0.99 (See* ***Figure 1-source data*** *for all tests). The dashed vertical line indicates 90 days of age, separating development and ageing periods. Age distribution of samples are given in* ***Figure 1-figure supplement 1****.* ***c)*** *Similarity between the age-related gene expression changes (Spearman's correlation coefficient between expression and age without a significance cutoff) across tissues in development and ageing. Similarities were calculated using Spearman's correlation coefficient between expression-age correlations across tissues. CTX: cortex, LV: liver, LNG: lung, MS: muscle.* ***d)*** *The number of significant age-related genes in each tissue (FDR corrected p-value<0.1).* ***e)*** *Shared age-related genes among tissues identified without using a significance cutoff. The x-axis shows the number of tissues among which age-related genes are shared. Significant overlaps are indicated with an asterisk (*) **(Figure 1-figure supplement 4)**.* ***f)*** *The proportion of age-related expression change trends (no significance cutoff was used) in each tissue across the lifetime. UpDown: up-regulation in development and down-regulation in the ageing; DownUp: down-regulation in development and up-regulation in the ageing; UpUp: up-regulation in development and up-regulation in the ageing; DownDown: down-regulation in development and down-regulation in ageing. We confirmed the robustness of the results using VST normalisation in* ***Figure 1-figure supplement 10.***

Tissues diverge during postnatal development. Consistent with earlier work (Brawand et al. 2011; Cardoso-Moreira et al. 2019), we found that variation in gene expression is largely explained by tissue differences, such that the first three **p**rincipal **c**omponents (PCs) separate samples according to tissue (ANOVA $p<10^{-20}$ for PC1-3, **Figure 1-source data**), with the cortex most distant from the others **(Figure 1a)**. Meanwhile, PC4, which explains 8% of the total variance, displayed a shared age-effect across tissues in development (Spearman's correlation coefficient ρ=[-0.88, -0.99], nominal $p<0.01$ for each test; **Figure 1b**). Also, after the tissue effect was removed by standardisation, **p**rincipal **c**omponents **a**nalysis (PCA) showed a strong influence of age on the first two PCs, which explains 31% of the variance in total **(Figure 1-figure supplement 2)**. We further observed higher similarity among tissues at the juvenile stage compared to the young-adult stage. In other words, distances between tissues increased with age (change in mean Euclidean distance among tissues with age during development in PC1-PC4 space ρ_{dev}=0.99, $p_{dev}=1.5x10^{-5}$, **Figure 1-source data**), which resonates with previous reports of inter-tissue transcriptome divergence during development (Cardoso-Moreira et al. 2019). This divergence pattern was also observed when PCA was performed with developmental samples only (days 2 to 61: change in mean Euclidean distance among tissues in PC1-PC4 space; ρ=0.95, p=0.0008; **Figure 1-figure supplement 3a-b**).

Tissues involve common gene expression changes with age. We next characterised age-related changes in gene expression shared across tissues by i) studying overall trends at the whole transcriptome level and testing their consistency using permutation tests, and ii) studying statistically significant changes at the single gene level. First, we investigated similarities in overall trends of gene expression changes with age using the Spearman's correlation coefficient (ρ) between expression levels and age, for each gene, in each tissue, separately for the developmental and ageing periods (Methods; tissue-specific age-related gene expression changes and functional enrichment test results are available as **Supplementary File 1**). We then examined transcriptome-wide similarities across tissues during development and ageing by comparing these gene-wise expression-age correlation coefficients **(Figure 1c)**. Considering the whole transcriptome without a significance cutoff, we found a weak correlation of age-related expression changes in tissue pairs, both during development (ρ=[0.17, 0.39], permutation test $p<0.05$ for all the pairs, **Figure 1-source data**), and ageing (ρ=[0.23, 0.33], permutation test $p<0.05$ in 4/6 pairs, **Figure 1-source data**). We then tested whether developmental

patterns among tissues may be shared more than ageing-associated patterns, but we did not find significant difference between inter-tissue similarities within the development and those within ageing (Wilcoxon signed-rank test, p=0.31). Moreover, the number of genes with the same direction of change (without applying a significance cutoff) across four tissues was consistently more than expected by chance (permutation test p<0.05), except for genes upregulated in ageing (**Figures 1e, Figure 1-figure supplement 4**). This attests to overall similarities across tissues both during postnatal development and during ageing, albeit of modest magnitudes. We obtained similar results using another normalisation approach, **v**ariance **s**tabilising **t**ransformation or VST from the DESeq2 package (Love et al. 2014), and confirmed that the observed patterns are not affected by the choice of normalisation method (**Figure 1-figure supplement 10-11)**.

In the second approach, we focused on genes showing a significant age-related expression change, identified separately during development or during ageing (using Spearman's correlation coefficient and **f**alse-**d**iscovery **r**ate (FDR) corrected p-value<0.1, **Figure 1d**). We found that the developmental period was accompanied by a large number of significant changes (n=[1,941, 6,151], 13-41% across tissues), with the most manifest changes detected in the cortex. The genes displaying significant developmental changes across all four tissues also showed significant overlap (**Figure 1-figure supplement 5a, Figure 1-figure supplement 6**; permutation test: p_{shared_up}=0.027, p_{shared_down}<0.001). Using the **G**ene **O**ntology (GO), we found that shared developmentally up-regulated genes were enriched in functions such as hormone signalling pathways and lipid metabolism (FDR-corrected p-value<0.1). Meanwhile, shared developmentally down-regulated genes were enriched in functions such as cell cycle and cell division (FDR-corrected p-value<0.1; **Supplementary File 2**). Contrary to widespread expression change during development (13-41%), the proportion of genes undergoing significant expression change during ageing was between 0.013-15% **(Figure 1d)**. This contrast between postnatal development and ageing was also observed in previous work on the primate brain (Somel et al. 2010; Işıldak et al. 2020). In terms of the number of genes with a significant ageing-related change, the most substantial effect we found was in the lung (n=2,319), while close to no genes showed a statistically significant change in the muscle (n=2), a tissue previously noted for displaying a weak ageing transcriptome signature across multiple datasets (Turan et al. 2019). Not unexpectedly, we found no common significant ageing-related genes across tissues **(Figure 1-figure supplement 5a)**.

Considering the similarity between the ageing and development datasets **(Figure 1c)** and the similar sample sizes in development (n=7) and ageing periods (n=9), the lack of overlap in significant genes in ageing might be due to low signal-to-noise ratios in the ageing transcriptome, as ageing-related changes are subtler compared to those in development **(Figure 1-figure supplement 5b)**.

Gene expression reversal is a common phenomenon in multiple tissues. We then turned to investigate the prevalence of the reversal phenomenon (*i.e.* an opposite direction of change during development and ageing) across the four tissues. We first compared the trends of age-related expression changes between development and ageing periods in the same tissue, without a significance cutoff, to assess transcriptome-wide reversal patterns **(Figure 1c)**. This revealed weak negative correlation trends in liver and muscle (though not in the lung and cortex), *i.e.* genes up- or down-regulated during development tended to be down- or up-regulated during ageing, respectively. These reversal trends were comparable when the analysis was repeated with the genes showing relatively high levels of age-related expression change ($|\rho|>0.6$ in both periods; **Figure 1-figure supplement 7**). We further studied the reversal phenomenon by classifying each gene expressed per tissue (n=15,063) into those showing up- or down-regulation during development and during ageing. Here, again, we did not use a statistical significance cutoff and summarised trends of continuous change versus reversal in each tissue. This approach follows Dönertaş et al. (2017) and focuses on global trends instead of single genes. In line with the above results, as well as earlier observations in the brain, kidney, and liver (Dönertaş et al. 2017; Anisimova et al. 2020), we found that ~50% (43-58%) of expressed genes showed reversal trends (**Figure 1f**), although these proportions were not significantly more than randomly expected in permutation tests (**Figure 1-figure supplement 8**, Methods). Overall, we conclude that although the reversal pattern is not ubiquitous, the expression trajectories of the genes do not necessarily continue linearly into the ageing period.

Pathways related to development, metabolism and inflammation are associated with the reversal pattern. We then asked whether genes displaying reversal patterns in each tissue may be enriched in functional categories. Our earlier study focusing on different brain regions had revealed that up-down genes, *i.e.* genes showing developmental up-regulation followed by down-regulation during ageing, were enriched in tissue-specific pathways, such as neuronal functions (Dönertaş et al. 2017).

Analysing up-down genes compared to all genes up-regulated during development, we also found significant enrichment (FDR corrected p-value<0.1) in functions such as “synaptic signaling” in the cortex, as well as “tube development” and “tissue morphogenesis” in the lung, “protein catabolic process” in the liver and “cellular respiration” pathways in the muscle **(Supplementary File 3)**. Meanwhile, down-up genes (down-regulation during development followed by up-regulation during ageing) showed significant enrichment in functions such as “wound healing”, and “peptide metabolic process” in the cortex, “translation” and “nucleotide metabolic process” in the lung, “inflammatory response” in the liver and ‘leukocyte activation’ in the muscle **(Supplementary File 3)**.

Genes showing a reversal pattern are not shared among tissues. As tissues displayed modest positive correlations in their development- or ageing-related expression change trends **(Figures 1c, Figure 1-figure supplement 7)**, and as we had previously observed that distinct brain regions show similarities in their reversal patterns (*i.e.* the same genes showing the same reversal type), different tissues might also be expected to show similarities in their reversal patterns. Interestingly, we found no overlap between gene sets with the reversal pattern (up-down or down-up genes) across tissues, relative to random expectation (permutation test, $p_{up\text{-}down}$=0.08, $p_{down\text{-}up}$=0.53; **Figure 1-figure supplement 9**). Such a lack of overlap might be explained if genes showing reversal patterns in each tissue tend to be tissue-specific. It would also be consistent with the notion that reversals involve loss of cellular identities gained in development, during which tissue transcriptomes appear to diverge from each other (**Figures 1a, Figure 1-figure supplement 3**) (Cardoso-Moreira et al. 2019). This result led us to ask whether, in accordance with the reversal phenomenon, inter-tissue transcriptome divergence may be followed by increasing inter-tissue similarity, or convergence, during ageing.

Inter-tissue divergence during development and convergence during ageing. We studied the inter-tissue divergence/convergence question using two approaches. In the first, we analysed how transcriptome-wide expression variation among tissues changes with age regardless of their age-related expression patterns in any particular tissue. To do this, for each individual, we calculated the **c**oefficient **o**f **v**ariation (CoV) across the four tissues for each commonly expressed gene (n=15,063), which represents a measure of expression variation among tissues. Then, we assessed how such inter-tissue variation changes over the lifetime, by calculating the Spearman’s correlation coefficient between

CoV and age, separately for development and ageing periods (correlation values for all genes are given in **Figure 2-source data**).

Using the CoV values calculated across all 15,063 genes (excluding one 904 days-old individual for which we lacked the cortex data), we observed a significant mean CoV increase in development (Spearman's correlation coefficient ρ=0.77, two-sided p=0.041), confirming that tissues diverge as development progresses **(Figure 2a)**. Interestingly, during ageing, we observed a decrease in mean CoV with age, albeit not significant (ρ=-0.50, p=0.204, **Figure 2a**), suggesting that tissues may tend to converge during ageing. This was also supported by the PCA analysis in which we observed a trend of ageing-associated decrease in mean Euclidean distance among tissues (using PC1-PC4 space with quantile normalised data: ρ=-0.87, p=0.0026; with VST normalised data ρ=-0.58, p=0.102, **Figure 1-source data**). We obtained the same divergence-convergence pattern by calculating the median CoV values for each individual instead of the mean (**Figure 2-figure supplement 1**). Figure 2b exemplifies this pattern of increasing and then decreasing CoV through lifetime for the gene displaying the strongest such signal.

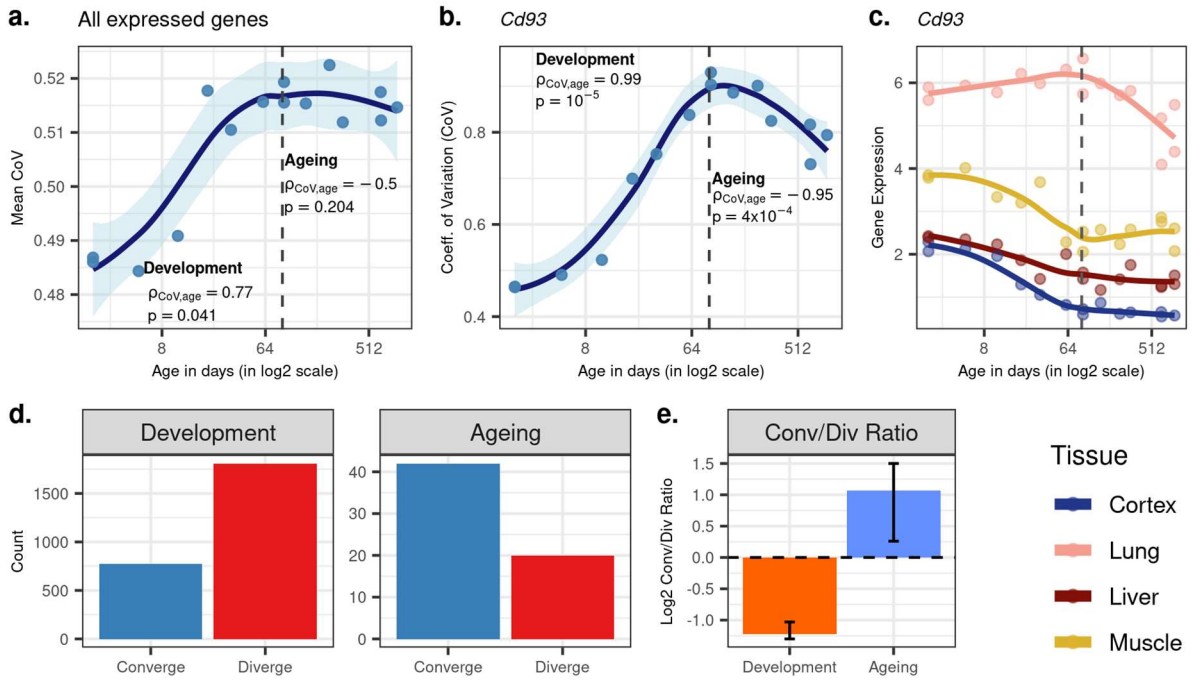
We identified n=9,058 genes showing divergent trends among tissues in development based on their CoV change with age (without using a significance cutoff per gene). Among these, n=4,802 showed convergent trends in ageing, which we refer to as **di**vergent-**co**nvergent (DiCo) genes. We next studied the transition points between divergence and convergence by clustering genes showing the DiCo pattern (n=4,802) based on their CoV values **(Figure 2-figure supplement 2)**. Notably, Cluster 1, which shows a slightly delayed divergence starting after 8-days and peaks around 3-months, was associated with metabolic and respiration-related processes (FDR-corrected p-value<0.1), and Cluster 5, which shows a relatively delayed convergence after 4 months, was enriched in categories related to vascular development (FDR-corrected p-value<0.1) **(Supplementary File 4)**. To assess the contribution of different tissues to the DiCo pattern, we further clustered DiCo-displaying genes (n=4,802) based on their expression levels **(Figure 2-figure supplement 3)**. Not surprisingly, the clusters with relatively higher expression levels of a tissue (*e.g.* muscle in Cluster 9) were enriched in functional categories (FDR-corrected p-value<0.1) related to that tissue (*e.g.* muscle cell development) **(Supplementary File 5)**.

We then studied DiCo at the single-gene level. We tested each gene for a significant CoV change in their expression levels (*i.e.* divergence or convergence) in development and ageing (Spearman's correlation test with FDR corrected p-value<0.1). We found that the ratio of divergent and convergent genes differed significantly between development (70% divergence among 2,581 significant genes) and ageing (68% convergence among 62 significant genes) (**Figure 2d-e**). The same pattern was also observed without using significance cutoff (**Figure 2-figure supplement 4**). We also confirmed that this pattern is also observed with VST-normalised data (Methods), and is thus not affected by the data preprocessing approach (**Figure 2-figure supplement 14**).

To our knowledge, inter-tissue convergence during ageing is a novel phenomenon. We first considered the possibility that convergence during ageing could be explained by heteroscedasticity which could arise due to increased inter-individual variability in gene expression during ageing (Somel et al. 2006). To test this hypothesis, we compared expression-age heteroscedasticity levels between two gene sets; 1) genes with the DiCo pattern, 2) genes showing divergent patterns throughout lifetime (DiDi, n=4,182) for each tissue, separately (Methods). We did not observe any significant difference in heteroscedasticity between DiCo and DiDi genes in any of the tissues (two-sided KS test, p>0.05 in all tissues, **Figure 2-figure supplement 15**), which suggests that heteroscedasticity due to increased inter-individual variability probably does not drive the observed age-related convergence during ageing. Visual inspection of gene expression clusters also suggested that the DiCo pattern is not particularly associated with non-linear changes in gene expression with age (**Figure 1-figure supplement 12-15**).

In order to further verify the DiCo pattern, we used a second approach to test it in our mouse dataset. For each individual, we calculated correlations between pairs of tissues across their gene expression profiles. Under the DiCo pattern, we would expect pairwise correlations to decrease during development and increase during ageing. Among all pairwise comparisons, we observed a strong negative correlation during development (ρ=[-0.61, -0.9], nominal p<0.05 in 5 out of 6 tests), while during ageing, 4 out of 6 comparisons showed a moderate positive correlation (ρ=[0.16, 0.69], nominal p<0.05 in 1 out of 6 comparisons, **Figure 2-figure supplement 5**). Calculating the mean of pairwise correlations among tissues for each individual, we observed the same DiCo pattern (nominal p<0.05 for both periods**,** **Figure 2-figure supplement 6**).

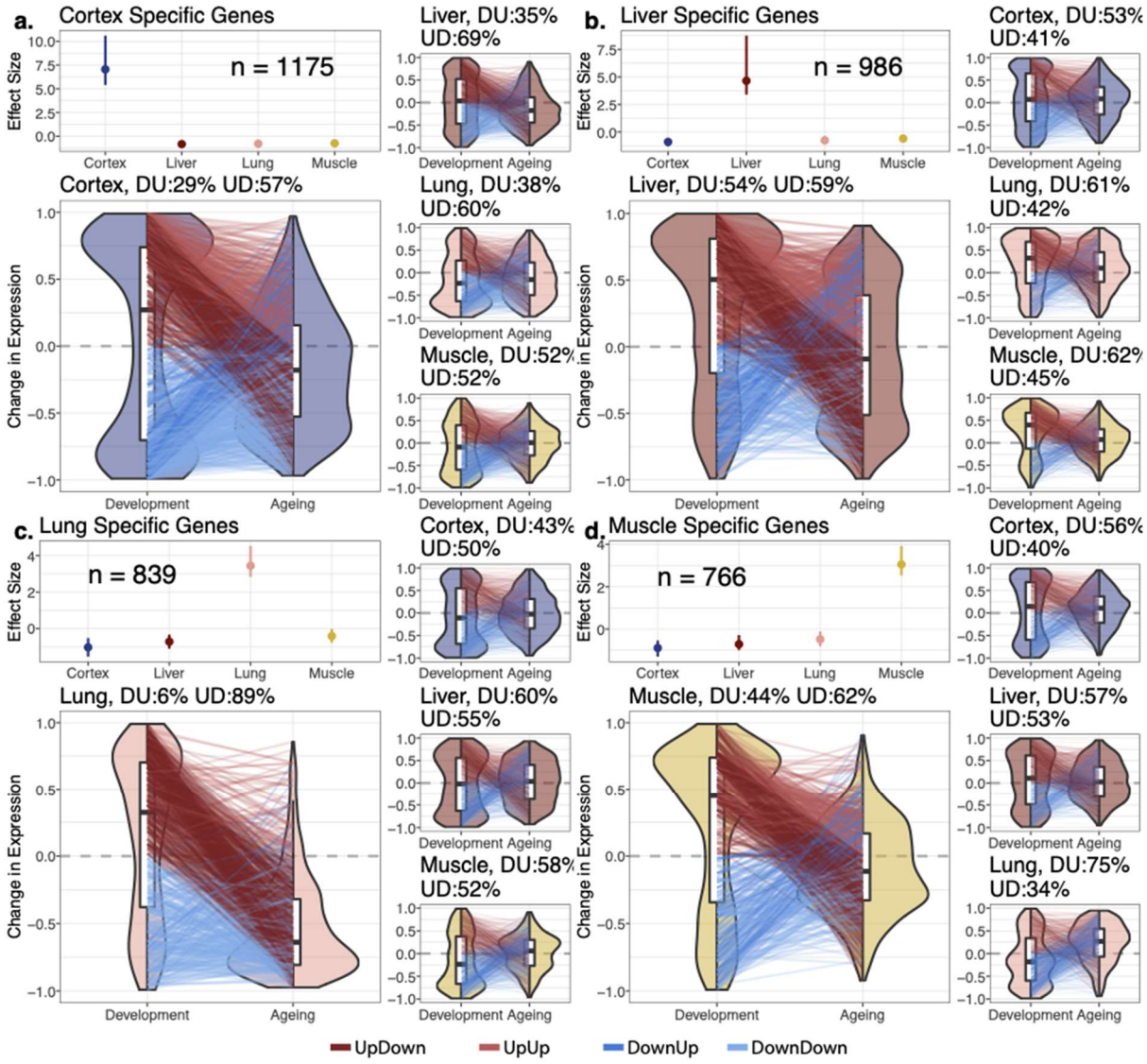
Figure 2. Age-related change in gene expression variation among tissues estimated with CoV



a) *Transcriptome-wide mean CoV trajectory with age. Each point represents the mean CoV value of all protein-coding genes (15,063) for each mouse (n=15) except the one that lacks expression data in the cortex.* ***b)*** *Age effect on CoV value of the Cd93 gene which has the highest rank for the DiCo pattern, in four tissues (Methods). CoV increases during development and decreases during ageing, indicating expression levels show DiCo patterns among tissues.* ***c)*** *Expression trajectories of the gene Cd93 in four tissues.* ***d)*** *The number of significant CoV changes with age (FDR corrected p-value<0.1) during development (left, $n_{conv.}$=772, $n_{div.}$=1,809) and ageing (right, $n_{conv.}$=42, $n_{div.}$=20). Converge: genes showing a negative correlation (ρ) between CoV and age; Diverge: genes showing a positive correlation between CoV and age.* ***e)*** *Log2 ratio of convergent/divergent genes in development and in ageing. The graph represents only genes showing significant CoV changes (FDR corrected p-value<0.1, given in panel d). Error bars represent the range of log2 ratios calculated from leave-one-out samples using the jackknife procedure (Methods, values are given in* ***Figure 2-source data****).*

The divergence-convergence (DiCo) pattern indicates loss of tissue-specificity during ageing. Potential explanations of the DiCo pattern involve two scenarios consistent with the age-related loss of identity: i) decreased expression of tissue-specific genes in their native tissues, or ii) non-specific expression of tissue-specific genes in other tissues. To test these predictions, we first identified tissue-specific gene sets based on relatively high expression of that gene in a particular tissue (cortex: 1,175, lung: 839, liver: 986, muscle: 766 genes). We noted that tissue-specific genes show clear up-down reversal patterns, being mostly up-regulated during development, and down-regulated during ageing (**Figure 3**, 57-89%). The up-down reversal pattern was particularly strong among tissue-specific genes for the three of four tissues tested (OR = [1.65, 6.52], $p<0.05$ for each tissue except in liver: OR=0.87, $p=0.09$, **Figure 3-source data**). Tissue-specific genes were also enriched among DiCo genes (**Figure 3-source data**, OR=1.56, Fisher's exact test $p<10^{-16}$).

Figure 3. Reversal patterns among tissue-specific genes



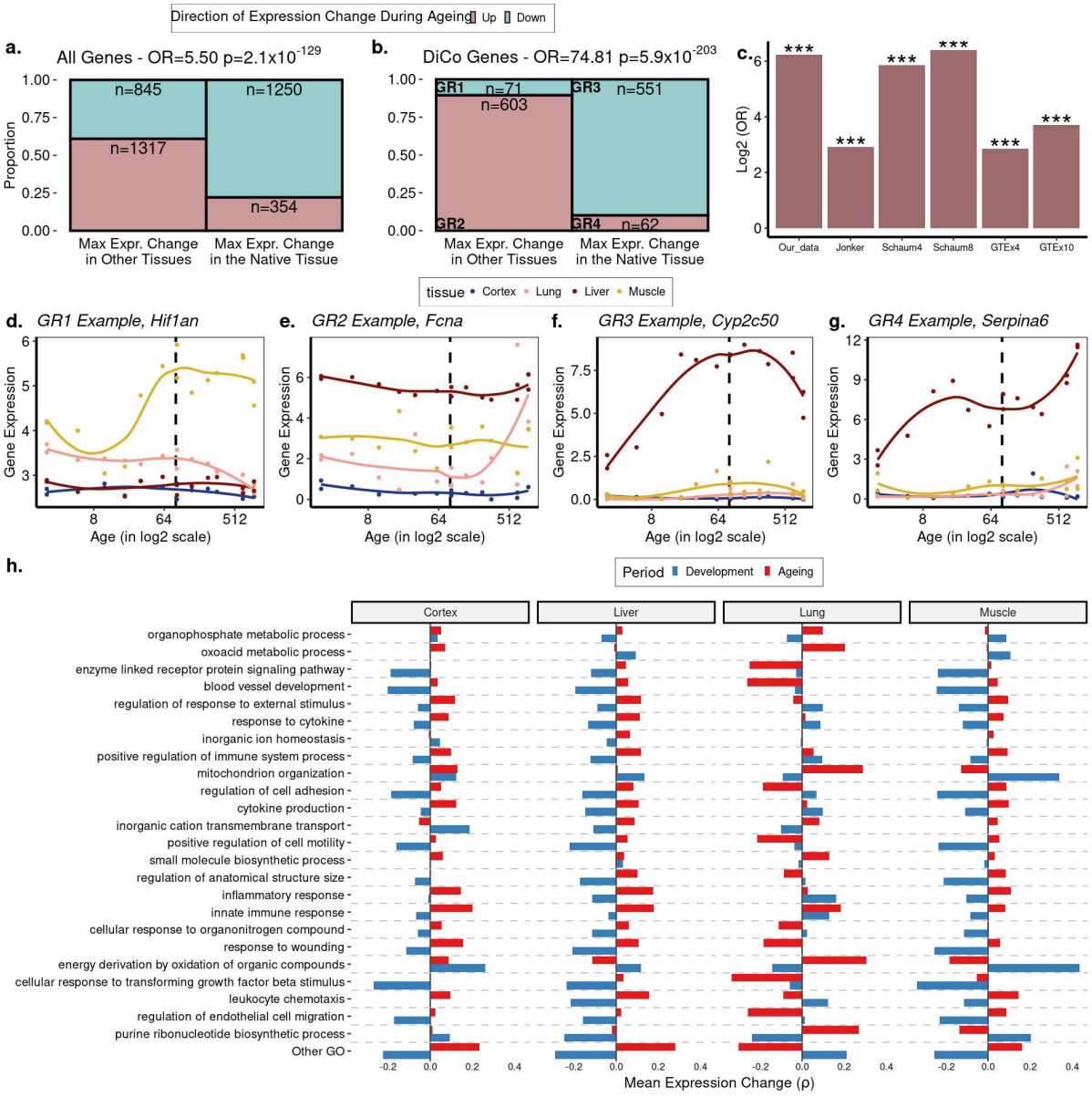
Age-related expression changes of the tissue-specific genes. In each panel ***a-d****, the upper left subpanels show effect size (ES) calculated with the Cohen's D formula, using expression levels of each gene among tissues (Methods). The IQR (line range) and median (point) effect size for each tissue is shown. The number of tissue-specific genes is indicated inside each subpanel. The lower left subpanels show violin plots of the distribution of age-related expression change values (Methods) among tissue-specific genes, in development and in ageing. Each quadrant represents the plots for each tissue-specific gene group. The red and blue lines connect gene expression changes for the same genes in development and ageing. DU: percentage of down-up reversal genes among down-regulated, tissue-specific genes in development. UD: percentage of up-down reversal genes among up-regulated, tissue-specific genes in development. Tissue-specific genes are enriched among UD reversal genes except in the liver (Fisher's exact test;* $OR_{cortex}=1.65$, $OR_{lung}=6.52$, $OR_{liver}=0.87$, $OR_{muscle}=1.26$, $p<0.05$ for each test except in liver).

We then tested our initial prediction that the DiCo pattern is related to tissue-specific genes losing their expression in their native tissue and/or gaining expression in non-native tissues during ageing. We first tested this hypothesis by considering all tissue-specific genes. We found a positive odds ratio between loss of expression in native tissue and gain in other tissues during ageing (OR = 5.50, Fisher's exact test $p=2.1x10^{-129}$, **Figure 4a**). The same analysis conducted with only the DiCo genes yielded a much stronger association (OR=74.81, Fisher's exact test $p=5.9x10^{-203}$, **Figure 4b**). This suggests that loss of tissue-specific expression is observed across the transcriptome, with a particularly strong association among DiCo genes. Figure 4c-f exemplifies the expression trajectories of genes chosen from each group defined in Figure 4b.

We then asked whether genes displaying the DiCo pattern may be related to specific functional pathways or share specific regulators. Using GO, we searched for functional enrichment among convergent genes during ageing, using developmentally divergent genes as the background (Methods). We found enrichment for 184 GO **B**iological **P**rocess (BP) categories for the DiCo pattern (**K**olmogorov-**S**mirnov (KS) Test, FDR-corrected p-value<0.1, **Figure 4-source data**) and summarised enriched categories by clustering them based on the number of genes they share. We then studied the trends of gene expression changes with age (without a significance cutoff) in each representative category for each tissue (Methods) (**Figure 4h**; we provide detailed clustering for the categories in 'Other GO' (**Figure 4-figure supplement 1**))**.** On average, energy metabolism, mitochondria and tissue function-related categories, as well as immune response-related categories, exhibit DiCo type expression changes over time and across tissues, where temporal changes in different tissues occur in opposite directions**.** Notably, for the majority of representative GO categories, the lung had the most distinct expression patterns in both periods (**Figure 4h, Figure 4-figure supplement 1**).

Contrary to the functional enrichment results, we did not find any specific regulators (miRNA or transcription factors) associated with DiCo using the same background as above (at 235 tests for miRNA and 158 tests for TF, FDR corrected p-value>0.1 for both tests) (Methods), which suggests that DiCo pattern may not be driven by a limited number of specific regulators, but may instead be a transcriptome-wide phenomenon.

Figure 4. The loss of tissue-specific expression during ageing and functional enrichment of DiCo genes



a) *Mosaic plot showing the association between maximal expression change in native vs. non-native tissues (x-axis) vs. down- (cyan) or up- (pink) regulation during ageing across all tissue-specific genes (n=3,766). The highly significant odds ratio indicates that genes native to a tissue tend to be down-regulated during ageing in that native tissue, if they show maximal expression change during ageing in that tissue. Conversely, if they show maximal expression change during ageing in non-native tissue, those genes are up-regulated during ageing. Consequently, tissue-specific expression patterns established during development will tend to be lost during ageing.* ***b)*** *The same as (a) but using only the tissue-specific genes that show the DiCo pattern (n=1,287).* ***c)*** *Summary of the association tests for 'direction of maximal expression change in native vs. non-native tissues' across all datasets analysed. The y-axis shows log_2 transformed Odds Ratio (OR) for each dataset (x-axis) - Schaum4: using the same four*

*tissues as our dataset. Schaum8: using eight tissues. GTEx4: using the same four tissues as our dataset. GTEx10: using ten tissues. ***: FDR-corrected p-value<10^{-87}. P-values are given in* ***Table 1****. The 4 groups are annotated as GR1-4 and gene expression changes for each group in our dataset is exemplified in* ***d-g. h)*** *Trends of expression change with age of genes (x-axis) in categories enriched in DiCo (GSEA). Enriched categories (n=184) are summarised into representatives (y-axis) using hierarchical clustering and Jaccard similarities (Methods). Categories are ordered by the number of genes they contain from highest (bottom, n = 290) to lowest (top, n = 26). The most distant cluster with low within-cluster similarity in the hierarchical clustering (Other GO) was clustered separately and given in* ***Figure 4-figure supplement 1****.*

Additional mouse and human datasets confirm the association between loss of tissue-specificity and inter-tissue convergence during ageing. We investigated inter-tissue convergence during ageing in three additional datasets where multiple tissue samples were available for the same individuals (**Table 2**). We conducted the analysis using a subset of the same four tissues in our dataset and also larger sets when additional samples were available. Age-related expression changes showed small to moderate correlations among all datasets analysed, with our dataset being most similar to the mouse dataset from Jonker et al., while the GTEx human dataset was the most distinct (**Figure 4-figure supplement 2a**).

First, using the Jonker et al. dataset (Jonker et al. 2013) comprising 5 tissues (**Table 2**), we observed transcriptome-wide convergence during ageing with a significant decline in mean Euclidean distance between PCs (ρ = -0.57, p = 0.014, **Figure 2-figure supplement 7a-c**) and a strong decrease in mean CoV during ageing (ρ = -0.48, p = 0.044, **Figure 2-figure supplement 7d**). Moreover, we found that 7/10 tissue pairs showed increased pairwise tissue correlations during ageing, although none of them was significant after multiple testing correction (**Figure 2-figure supplement 7f**). Sixty-six percent of the genes with a significant change in CoV were convergent, comparable to our dataset showing 68% convergence among significant changes. We also tested the association between the loss of identity and convergence pattern by repeating the same analysis as in Figure 4b with the Jonker et al. dataset, using only the convergent genes in ageing as we lack developmental period. We again found strong association, consistent with convergent genes losing expression in their native tissue and gaining in other tissues during ageing (OR=7.52, p<10^{-16}, **Figure 4c**). The results are summarised in **Table 1**.

Next, we used another mouse dataset by Schaum et al. (Schaum et al. 2020) (**Table 2**). Repeating the analysis on the same 4 tissues and also a larger set of 8 tissues, we did not find support for transcriptome-wide convergence (**Table 1**, **Figure 2-figure supplement 17, 19**). In the 4-tissue comparison 4/6 tissue-pairs, and in the 8-tissue comparison only 16/28 tissue-pairs showed positive correlations, supporting the inter-tissue convergence during ageing (**Figure 2-figure supplement 18c, 20c**). Interestingly, 75% of the negative correlations involved muscle and subcutaneous fat. Convergence ratios among genes showing significant change in CoV (FDR corrected p-value<0.1) were marginally above 50%. Although we did not observe widespread convergence during ageing in this dataset, we still detected strong associations between convergence in ageing and tissue specificity ($OR_{4\text{-tissue}}$=1.33, p = $1.08\text{x}10^{-8}$) and identity loss ($OR_{4\text{-tissue}}$=58.3 p < 10^{-16}; $OR_{8\text{-tissue}}$=84.2 p < 10^{-16}) (**Figure 4c**).

Lastly, we used the GTEx dataset to investigate inter-tissue convergence during ageing in humans. Calculating the change in mean Euclidean distance based on PCA and mean CoV values, we found a non-significant tendency towards convergence across the whole transcriptome in the same 4 tissues and a larger set of 10 tissues (**Table 1, Figure 2-figure supplement 8, 10**). We also performed the 4-tissue comparison with female and male individuals separately and observed relatively strong inter-tissue convergence among ageing females (ρ_{female}= -0.58, p_{female}=0.059) but less in males (ρ_{male}= -0.052, p_{male}=0.77) which lack individuals at the youngest and oldest age groups (**Figure 2-figure supplement 16**). Moreover, 5/6 and 29/45 tissue-pairs showed increased correlation with age in 4-tissue and 10-tissue comparisons, consistent with inter-tissue convergence during ageing (**Figure 2-figure supplement 9, 11**). Notably, 8 of 16 negative correlations in the 10-tissue comparison involved the skin tissue (**Figure 2-figure supplement 11c**). We also studied significant changes in CoV per gene, but found no significant gene in the 4-tissue comparison and only 3 genes in the 10-tissue comparison, all of which were convergent. Finally, we tested the association between the loss of expression in native tissue and gain in other tissues during ageing among convergent genes, confirming the association with the tissue identity (**Figure 4c, Table 1**).

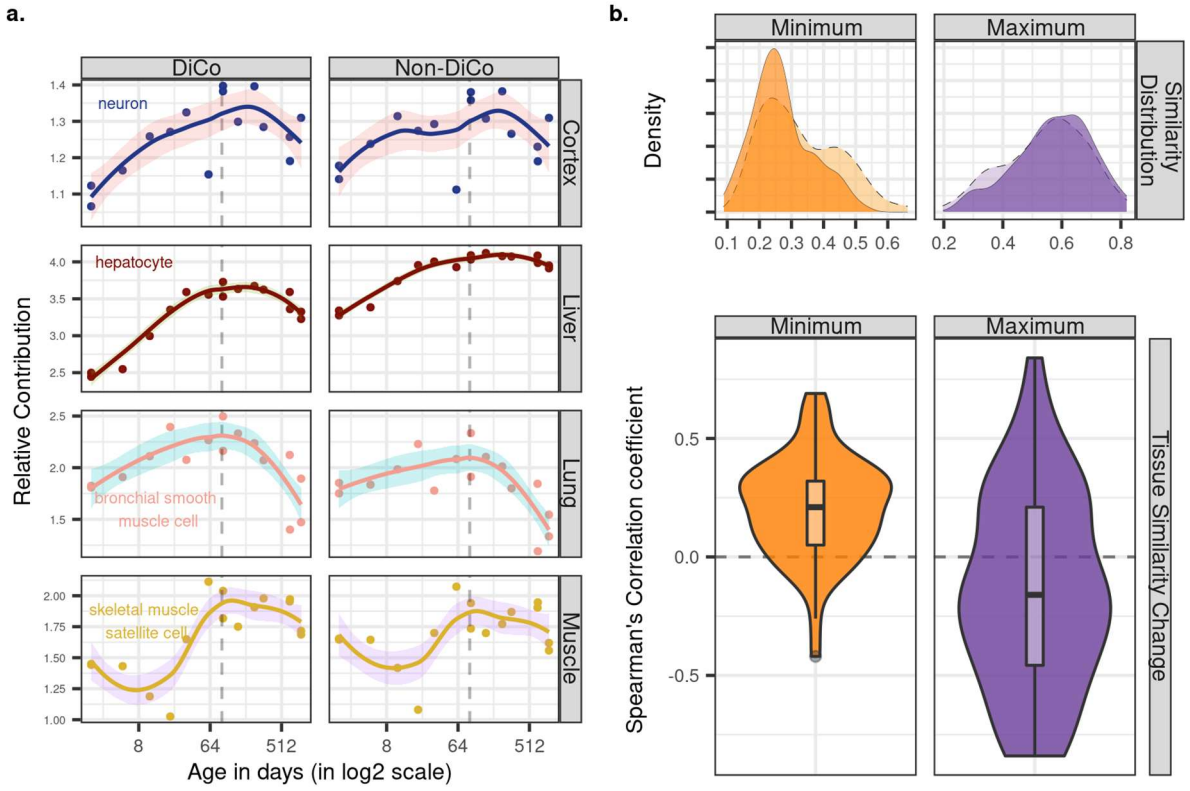
Overall, analysis of these three additional datasets indicates that inter-tissue convergence during ageing is commonly, but not always, observed at the transcriptome-wide level in mice and in humans.

458 Notably, the transcriptome-wide trend was weak in the Jonker et al. and GTEx datasets and not evident
459 in the Schaum et al. dataset. The association between the loss of identity and convergence, on the
460 other hand, was strong across all datasets (**Table 1**).
461
462 We further asked whether convergent gene sets identified in different datasets overlap. Eleven of 15
463 comparisons were significant, but the effect sizes were small **(Figure 4-figure supplement 2b).** We
464 reason that the low overlap across datasets might reflect that transcriptome-wide convergence was
465 weak and that we lack the developmental samples for the external datasets, *i.e.* we can only compare
466 convergence during ageing but not the DiCo pattern. Noteworthy, only 62% of convergent genes in
467 ageing are divergent during development in our dataset, and low overlap between convergence does
468 not rule out overlap across DiCo genes.
469
470 These results suggest that inter-tissue convergence in ageing may be a weak but widespread
471 phenomenon and associated with the loss of tissue identity. Overall, while mouse and human tissues
472 display divergence in development (**Figures 1a, 2a**, (Cardoso-Moreira et al. 2019)), this appears to be
473 followed by a trend towards inter-tissue convergence in ageing **(Figures 2a, Figure 2-figure**
474 **supplement 1-20)**, and could be linked to loss of tissue identity.
475
476 **Table 1**: Result summary of the all datasets analysed. First column shows the names of datasets analysed.
477 Numbers in parentheses show the sample sizes. 'Among all genes' column refers to the analyses performed using
478 all genes relevant to those analyses (subcolumns) without a significance cutoff. 'Within significant CoV changes':
479 genes show significant CoV change with age with FDR corrected p-value<0.1. In the 'DiCo vs Tissue specificity
480 (Di- as background)' column, divergent genes in development (Di-) were chosen as background. 'Co vs expression
481 change in native tissue association (Fig 4b)' column refers to the analysis performed in Figure 4b for each dataset
482 and the results were presented in Figure 4c. The association tests were performed among convergent genes in
483 ageing except in our dataset which was performed with DiCo genes. Significant test results were indicated with
484 bold fonts. Green cells show the results that support convergence or tissue-specific expression loss in ageing
485 whether as a significant result or as a trend. Unsupportive test results and inapplicable tests were coloured grey
486 and light yellow, respectively. rho: Spearman's correlation coefficient. OR: Odds ratio. * FDR corrected p-value<0.1.
487

	Among all genes						Within significant CoV changes
	PCA Change in Euclidean distance	Mean CoV change	Median CoV change	Pairwise tissue correlations	DiCo vs tissue specificity (Di- as background)	Co vs expression change in native tissue association (Fig 4b)	Co vs. Di proportions
Izgi2021	**rho=-0.87, p=0.0026**	rho=-0.5, p=0.2	rho=-0.48, p=0.23	4/6 positive, none significant*	**OR=1.56, p=1.3x10^{-18}**	**OR=74.81 p=5.9x10^{-203} (among 1287 DiCo genes)**	68% convergence (among 62 significant genes*)
Jonker2013 5 tissues, 2 different than ours (n=18)	**rho=-0.57, p=0.014**	**rho=-0.48, p=0.044**	rho=-0.03, p=0.91	7/10 positive, none significant*	Di- background missing	**OR=7.52, p=6.5x10^{-109} (among 2967 convergent genes)**	66% convergence (among 1735 significant genes*)
Schaum2020 Same 4 tissues (n=37)	rho=0.13, p=0.46	rho=0.25, p=0.14	rho=0.13, p=0.43	4/6 positive, **2 significant***	**OR=1.33, p=1.07x10^{-8}**	**OR=58.03, p=1.5x10^{-197} (among 2124 convergent genes)**	53% convergence (among 319 significant genes*)
Schaum2020 8 tissues (n=26)	rho=0.1, p=0.62	rho=0.16, p=0.43	rho=0.04, p=0.86	16/28 positive, **5 significant***	Di- background missing	**OR=84.2, p=9.7x10^{-96} (among 2380 convergent genes)**	54% convergence (among 244 significant genes*)
GTEx Same 4 tissues	rho=-0.23, p=0.12	rho=-0.12, p=0.42	rho=-0.18, p=0.23	5/6 positive, none significant*	Di- background missing	**OR=7.21, p=7x10^{-87} (among 2407 convergent genes)**	(no significant CoV changes)
GTEx 10 tissues	rho=-0.26, p=0.13	rho=-0.14 p=0.44	rho=-0.3, p=0.08	29/45 positive, none significant*	Di- background missing	**OR=13.01, p=5.7x10^{-114} (among 2195 convergent genes)**	(all 3 significant genes were convergent)

Changes in cellular composition and cell-autonomous expression can both explain the divergence-convergence pattern. Ageing-related transcriptome changes observed using bulk tissue samples may be explained by temporal changes in cell type proportions within tissues, by cell-autonomous expression changes, or both. To explore whether the observed inter-tissue DiCo patterns may be attributed to changes in cell type proportions, we used published data from a mouse single-cell RNA-sequencing experiment (Tabula Muris Consortium 2020). For each of the four tissues in our original experiment, we collected cell type-specific expression profiles from 3-month-old young adult mice in the Tabula Muris Senis dataset. We deconvoluted bulk tissue expression profiles in our mouse dataset using the corresponding tissue's cell type-specific expression profiles by regression analysis (Methods), and studied the relative contributions of each cell type to tissue transcriptomes and how these change with age. The analysis was performed with three gene sets; all genes (n=[12,492, 12,849]), DiCo (n=[4,007, 4,106]) and non-DiCo genes (n=[8,485, 8,743]). Studying these deconvolution patterns, we observed a weak but consistent trend involving the most common cell types in different tissues. For instance, analysing DiCo genes in the liver and lung, we found that the most common cell type's contribution (hepatocyte in the liver, and bronchial smooth muscle cell in the lung) tends to increase during development (Spearman's correlation coefficient ρ_{liver}=0.95, ρ_{lung}=0.81, nominal p<0.05). This contribution then decreases during ageing (ρ_{liver}=-0.77, ρ_{lung}=-0.86, nominal p<0.05) (**Figure 5a, Figure 5-figure supplement 1**). This pattern was also observed in muscle and cortex, albeit not significantly (**Figure 5a, Figure 5-figure supplement 1**). These changes most likely reflect shifts in cellular composition, some of which were demonstrated directly in mice using *in situ* RNA staining (Tabula Muris Consortium 2020). Repeating the analysis with non-DiCo genes resulted in highly similar patterns considering the most common cell types in tissues, except in muscle ageing in which the age-related decrease was significantly higher with DiCo genes than the non-DiCo genes (permutation test with re-sampling all genes, $p_{skeletal\text{-}muscle\text{-}satellite\text{-}cell}$=0.04) **(Figure 5a, Figure 5-figure supplement 1, Figure 5-figure supplement 2-5)**. These results indicate that the observed cellular composition changes may partly explain DiCo, although the influence of composition changes is not exclusive to genes displaying the DiCo pattern.

Figure 5. Contribution of tissue composition and cell-autonomous changes to the DiCo pattern



a) *Deconvolution analysis of our mouse dataset with the 3-month-old scRNA-seq data (Tabula Muris Senis) using DiCo* (n=[4,007, 4,106]) *and non-DiCo* (n=[8,485, 8,743]) *genes. Only the cell types with the highest relative contributions to each tissue bulk transcriptome are shown (cell type names are given within each plot). Contributions of all cell types to bulk tissue transcriptomes are shown in* ***Figure 5-figure supplement 1****.* ***b)*** *Distribution of correlations for minimally (left) and maximally (right) correlated cell type pairs among tissues (n=54 pairs). For each cell type of a given tissue, one minimally (or maximally) correlated cell type is chosen from other tissues among the 3-month age group of the Tabula Muris Senis dataset (density plots with solid line edges). Dashed lines show the correlation distributions in 24-months age of minimally or maximally correlated cell type pairs identified in the 3-months age group. Bottom panel shows age-related expression similarity (*ρ*) changes of minimally (left) and maximally (right) correlated cell type pairs. The correlation between age and tissue similarity (expression correlations) were calculated for each pair of cell types identified in the 3-months age group. All pairwise cell type correlations and their age-related changes are given in* ***Figure 5-source data.***

Next, we investigated the possible role of cell-autonomous changes in the DiCo pattern. Cell-autonomous changes could contribute to inter-tissue convergence during ageing in two ways. First, expression profiles of similar cell types shared across different tissues, such as immune cells, might converge with age. Another possible scenario, consistent with the notion of age-related cellular identity loss, is that the expression profiles of unrelated cell types, such as tissue-specific cell types in different tissues converge with age. To test these scenarios, we first ordered the pairwise correlations between cell types in different tissues at 3 months age group to determine the most similar and dissimilar cell types across tissues (Methods). Then, we studied how these similarities (*i.e.* pairwise correlations) change with age **(Figure 5b)**. Intriguingly, we found that pairs of similar cell types (*i.e.* those with the highest correlations) among tissues tend to become less similar with age (36/54 [67%] of pairwise comparisons, **Figure 5-source data**). On the contrary, the most distinct cell types (*i.e.* those with the lowest correlations) among tissues become more similar with age (45/54 [83%], **Figure 5-source data**). Repeating the analysis considering DiCo genes only yielded a similar trend (30/54 [56%] decrease in correlation among the most similar cell types, permutation test with re-sampling non-DiCo genes, $p>0.1$; and 47/54 [87%] increase in correlation among the most distinct cell types, permutation test, $p>0.1$). These trends are consistent with age-related cellular identity loss, and they suggest that cell-autonomous changes may also contribute to inter-tissue convergence during ageing, although further data and analyses would be needed to fully establish their validity.

Finally, we tested the possibility of intra-tissue convergence of cell types in the Tabula Muris Senis dataset, by calculating expression variation among cell types using the CoV measure for each individual. However, we did not observe a consistent trend of increasing similarity among cell types within tissues from 3m- to 24m-old mice **(Figure 5-figure supplement 6)**.

Discussion

Our findings confirm a number of ageing-associated phenomena identified earlier, while also revealing new patterns. First, we report parallel age-related expression changes among the four tissues studied, during development, as well as in ageing. The inter-tissue correlation distributions were modest and also comparable between development and ageing (**Figure 1c**). This last point may appear surprising at first glance, given the stochastic nature of ageing relative to development (Bahar et al. 2006;

Martinez-Jimenez et al. 2017; Angelidis et al. 2019; Somel et al. 2006; Feser et al. 2010; Kim, Villeponteau, and Jazwinski 1996; Enge et al. 2017), and also given earlier observations that developmental expression changes tend to be evolutionarily conserved, while ageing-related changes much less so (Zahn et al. 2007; Somel et al. 2010). At the same time, when we consider that tissues diverge during development, and also that ageing is characterised by parallel expression changes among tissues related to damage response, inflammation, and reduced energy metabolism (Zahn et al. 2007; Yang et al. 2015), similar magnitudes of correlations during development and ageing may be expected.

Second, we verify the generality of the reversal pattern, *i.e*. up-down or down-up expression change patterns across the lifetime, among distinct mouse tissues that include both highly mitotic (lung and liver) and less mitotic ones (skeletal muscle and cortex). Consistent with earlier observations in fewer tissues (Anisimova et al. 2020; Dönertaş et al. 2017), we find that about half the expressed genes display reversal in all cases studied. Importantly, expression reversal is not ubiquitous across all genes and our findings do not necessarily contradict the hyperfunction theory. Instead, we suggest that reversal is a common phenomenon that influences a notable fraction of the transcriptome and is a likely contributor to mammalian ageing.

Two observations here are notable. One is that reversal-displaying genes, especially those displaying the up-down pattern in each tissue, can be associated with tissue-specialisation-related pathways (*e.g.* morphogenesis) and tissue-specific functions (*e.g.* synaptic activity). The second observation is the lack of significant overlap among reversal genes among tissues. We thus hypothesised that reversals might be reflecting tissue specialisation during development (hence lack of overlap among tissues), and loss of specialisation during ageing. These processes could manifest themselves as inter-tissue divergence and convergence patterns over lifetime. We indeed observed that the up-down reversal pattern is enriched in tissue-specific genes, except in the liver. Studying inter-tissue similarity across mouse lifespan, we further found that the four tissues' transcriptomes diverged during postnatal development, and we further detected a trend towards inter-tissue convergence during ageing. We then further investigated this phenomenon through different approaches: i) by studying overall trends using PCA, ii) by analysing transcriptome-wide trends of inter-tissue CoV without considering gene-wise significance

cutoffs, iii) by focusing on genes with significant age-related changes in inter-tissue CoV, iv) by studying age-related changes in pairwise tissue correlations, and v) by analysing different cell-types using scRNA-seq data, and vi) by repeating the same analysis using independent mouse and human ageing datasets. The patterns we found were mostly consistent with inter-tissue convergence, but the majority of transcriptome-wide results were associated with low effect sizes, and some were not statistically significant. Importantly, all significant results suggested convergence during ageing. We therefore conclude that (1) developmental inter-tissue divergence does not continue into ageing; (2) convergence during ageing may be common although possibly not ubiquitous.

The weakness of the inter-tissue convergence signal per dataset and the limited overlap between convergent gene sets among datasets could have multiple reasons. These include the low signal-to-noise ratios characterising ageing-related expression patterns, the lack of old age individuals in our mouse dataset (>3-year-old mice) and the GTEx dataset (>90-year-old humans), limited overlap of tissues between our mouse dataset (cortex, liver, lung and muscle) and the Jonker et al. dataset (cortex, liver, lung, spleen, kidney), as well as differences in ageing patterns between species or between sexes. Further research involving larger sample sizes and diverse species are needed to confirm the generalisability of the observations.

Finally, we report a number of interesting observations on DiCo. We determine that tissue-specific genes tend to be down-regulated in the tissues that they belong to during ageing, while non-tissue-specific genes are up-regulated, which was confirmed by all external datasets (**Figure 4c**). Second, using deconvolution, we infer that cell types most common in a tissue (*e.g.* hepatocytes in the liver) tend to increase in frequency during development, but then decrease in frequency during ageing, as also shown recently using immunohistochemistry in a number of mouse tissues (Tabula Muris Consortium 2020). Accordingly, the DiCo phenomenon may at least partly be explained by shifts in cellular composition. This is intriguing as both highly mitotic and low mitotic tissues share this trend, indicating that an explanation based on stem cell exhaustion may not be applicable here. Third, we find increased expression similarity between distinct cell types in different tissues during ageing, but decreased similarity between similar cell types. Cell-autonomous expression changes, therefore, likely also contribute to the divergence-convergence phenomenon. We note that higher expression variability

among cells at old age (Hernando-Herraez et al. 2019; Enge et al. 2017) could also lead to inter-tissue convergence during ageing. A fourth interesting observation was the absence of significant enrichment for specific transcription factor or microRNA targets among DiCo genes. This result may not be surprising if inter-tissue convergence is mostly driven by stochastic damage accumulation, such as loss of epigenetic marks. It is also possible that instead of specific regulators, their interaction and cooperativity are associated with the DiCo. Future experimental studies could test both mechanistic aspects and functional link to tissue specificity.

We also note two major limitations of our study. One is related to the fact that our dataset represents bulk tissue samples, which may suffer from infiltration of foreign cell-types into tissues. Indeed, one of the external datasets, Schaum et al., included samples from perfused mice (Schaum et al. 2020) and we did not find support for the transcriptome-wide convergence during ageing, even though the association between tissue identity loss and convergence was also evident. The scRNA-seq dataset we analysed further suggested that DiCo is associated with tissue-specific genes and not immune- or blood-related categories, but we still cannot rule out possible infiltration artefacts that may affect our results. A second limitation is related to ageing being highly sex-dimorphic in mammals (Yuan et al. 2012; Sampathkumar et al. 2020). Hence, in-depth analysis of sex-specificity of the DiCo pattern could be relevant. Our mouse dataset included only male mice, while that of Jonker et al. was female-only. The fact that both revealed DiCo patterns suggest DiCo is not particular to one sex, but there could still exist sex-specific effects. In fact, when we analysed DiCo among human male and female individuals in the GTEx dataset separately, we observed slightly stronger inter-tissue convergence among ageing females than in males, although the GTEx male samples has also a drastically narrower age range (**Figure 2-figure supplement 16**). Accordingly, the prevalence of DiCo among humans and sexes waits to be determined.

Despite the open questions that remain, our results consistently support a model where ageing mammals suffer from loss of specialisation at the tissue level, and possibly also at the cellular level, which are observed as expression reversals and the newly discovered divergence-convergence phenomenon we report here.

Materials and Methods

Sample Collection

We collected bulk tissue samples from 16 male C57BL/6J mice. The samples were snap frozen in liquid nitrogen and stored at -80C. No perfusion was applied. The mice were of different ages covering the whole lifespan of *Mus musculus,* comprising both postnatal development and ageing periods. The samples included four different tissues; cerebral cortex, liver, lung and skeletal muscle. One 904 days-old mouse had no cortex tissue sample, and was thus excluded from the analysis. As a result, we generated 63 RNA-seq libraries in total.

Separation of development and ageing periods:

In order to compare gene expression changes during postnatal development and ageing we studied the samples before sexual maturation (covering 2 to 61 days of age, n=7) as the postnatal development period, and samples covering 93 to 904 days (n=9 in all tissues except in cortex where we had n=8) as the ageing period.

RNA-Seq Library Preparation

RNA sequencing was performed as previously described (Liu et al. 2016) with slight modifications. Briefly, total RNA was extracted using the Trizol reagent (Invitrogen) from frozen tissue samples. For sequencing library construction, we randomised all samples to avoid batch effects, and used the TruSeq RNA Sample Preparation Kit (Illumina) according to the manufacturer's instruction. Libraries were then sequenced on the Illumina HiSeq 4000 system in three lanes within one flow-cell, using the 150-bp paired-end module.

RNA-Seq Data Preprocessing

The quality assessment of the raw RNA-seq data was performed using FastQC v.0.11.5 (Andrews 2010). Adapters were removed using Trimmomatic v.0.36 (Bolger, Lohse, and Usadel 2014). The low-quality reads were filtered using the parameters: "PE ILLUMINACLIP: TruSeq3-PE-2.fa:2:30:1:0:8:true, SLIDINGWINDOW:4:15, MINLEN:25". The remaining high-quality reads were aligned to the mouse reference genome GRCm38 using STAR-2.5.3 (Dobin et al. 2013) with parameters: "--sjdbOverhang 99 --outSAMattrIHstart 0 --outSAMstrandfield intronMotif --sjdbGTFfile GRCm38.gtf". The percentage

of uniquely mapped reads in libraries ranged from 80 to 93%. We used cufflinks v.2.2.1 (Trapnell et al. 2010) to generate read counts for uniquely aligned reads (samtools "-q 255" filter) and calculated expression levels as fragment per kilobase million (FPKM). In total, we quantified expression levels for 51,608 genes in the GRCm38.gtf GTF file. We identified 50 duplicated genes with 1> FPKM value assigned, and the sum of their FPKM values were used.

All the remaining analysis was performed in R v.4.1. We restricted the whole analysis to only protein-coding genes obtained by the 'biotype' feature of the biomaRt library v.2.48.2 (Durinck et al. 2009). We also excluded genes which were not detected (zero FPKM) in 25% or more of the samples (at least 15 of 63), resulting in 15,063 protein-coding genes in total. As FPKM normalisation does not effectively account for cross-library variability, we additionally performed two normalisation approaches:

(a) Quantile normalisation: using all the samples together (n=63, regardless of their age or tissue), FPKM values were log2 transformed (after adding 1) and quantile normalised with 'normalize.quantiles' function from 'preprocessCore' library v.1.54 (Bolstad 2020). This approach equalises the distributions of different libraries. The assumption is that any large-scale differences in expression level distributions reflect technical factors.

(b) Variance stabilising transformation (VST): To assess the robustness of quantile normalisation on downstream analysis, we additionally implemented this approach, which ensures homoscedasticity, i.e. variances of expression levels are independent of the mean (Anders and Huber 2010). Uniquely aligned reads obtained from the STAR alignment were used to calculate read counts by HTSeq v.0.13.5 (Anders, Pyl, and Huber 2014) with parameters: "--format=bam --order=pos --stranded=no --type=exon --mode=union --nonunique=none". Read counts were then imported into R using the 'DESeqDataSetFromHTSeqCount' function in DESeq2 v.1.32.0 package (Love, Huber, and Anders 2014). The same filtration steps were applied as above, resulting in 14,973 protein-coding genes in total. Normalisation was performed with the 'vst' function and 'blinded=T' option in the DESeq2 package. The VST-normalised expression matrix was used to reproduce Figure 1 and Figure 2 results which are given in Figure 1-figure supplement 10, 11 and Figure 2-figure supplement 14.

Principal component analysis:

We studied the main sources of variation in the whole dataset using principal component analysis (PCA) on the scaled expression matrix with 'prcomp' function in the R base. The first four components, PC1 to PC4, explained 31%, 20%, 17% and 8% of the total variance. We observed a clear separation of tissues in PC1 and PC2 and a strong age effect in PC4. To statistically confirm tissue differences, we performed ANOVA on individual PC scores with tissue as explanatory variable; this was run on each of the first four PCs (PC1-PC4), separately. The magnitude of the age effect on PCA analysis was measured with Spearman's correlation test between individual age and each individual's PC score, separately in each tissue. PCA was also repeated for development and ageing periods, separately (**Figure 1-figure supplement 3**). We further calculated Euclidean distance in pairwise manner among tissues of each individual in PC1-4 space constructed in three different ways: (a) using all the samples together, (b) using only the developmental samples, (c) using only the ageing samples. Then, we tested the effect of age on mean Euclidean distance among tissues using the Spearman's correlation test. To study only the age effect on PC scores without the tissue effect we performed the following; (i) we removed the tissue-specific effects from the data by scaling the expression levels of each gene to mean=0 and sd=1 in each tissue separately, and (ii) we combined the four scaled expression matrices, (iii) we conducted PCA on the combined dataset (**Figure 1-figure supplement 2**).

Age-related gene expression change

To identify genes showing age-related expression change in each tissue, we used Spearman's correlation coefficient between individual age and expression level, separately for development and ageing periods. To capture potential non-linear but monotonic changes in expression, we chose the non-parametric two-sided Spearman's correlation test for both periods. We have used two-sided tests for all statistical tests throughout the article except the permutation tests. Significance of age-related genes was assessed with the **f**alse-**d**iscovery-**r**ate (FDR corrected p-value<0.1 cutoff, calculated with the Benjamini-Hochberg (BH) procedure (Benjamini and Hochberg 1995)) using the 'p.adjust' function in the R base library. Throughout the article, BH procedure with 0.1 cutoff was used for multiple test corrections of all statistical tests.

Functional associations:

We tested the functional associations of age-related gene expression change in separate tissues for each period (development and ageing) separately, employing the gene set over-representation analysis (GORA) procedure with Gene Ontology (GO) (Ashburner et al. 2000) Biological Process (BP) categories using the 'topGO' package v.2.44 (Alexa and Rahnenfuhrer 2019). We applied the 'classical' algorithm and performed Fisher's exact test on categories that satisfy the criteria of a minimum 10 and maximum 500 number of genes. We used the whole set of expressed genes (n=15,063) as the background. P-values were corrected for multiple testing using the BH procedure. Categories with FDR corrected p-value<0.1 were considered as significant.

Correlation between age-related gene expression changes in different tissues

We calculated Spearman's correlation coefficients between age-related gene expression change ρ_{gene} values (i.e. correlation between gene expression levels and age) calculated per gene in each tissue pair (**Figure 1c**). In order to test the statistical significance of the correlations, we used a permutation scheme as the expression levels across tissues are not independent but belong to the same mice. In order to account for the dependence, the individual ages were permuted in each round, but the permuted values were kept constant across tissues (similar to permutation tests applied in (Dönertaş et al. 2017; Işıldak et al. 2020; Dönertaş et al. 2018)). Specifically, we performed 1000 permutation rounds. In each round, we randomised the individual ages using the 'sample' function in R, while keeping the permuted age labels constant for individuals across tissues. We calculated the age-related gene expression changes with permuted ages in development and ageing datasets separately, thus simulating the null distribution with no age effect in each period. We then calculated the Spearman's correlation coefficient between the age-related expression levels from the permutations across tissues and assigned the p-value by calculating the proportion of permuted calculations with a more extreme correlation. All permutation tests in the article were performed as one-sided tests. The estimated false-positive-proportion (eFPP; proportion of false positives among all true non-significant results (true negatives+false positives)) was calculated as the median value of expected values divided by the observed value (**Figure 1-source data**).

Shared gene expression changes across tissues

We summarised the number of shared age-related genes among tissues for up- and down-regulated genes separately, using FDR corrected p-value<0.1 (**Figure 1-figure supplement 5**). The development and ageing datasets were tested separately. For each gene, we counted the number of tissues with the same direction of expression change with age. We calculated this overlap statistic among tissues (a) using genes with FDR-corrected p-value<0.1, and (b) with all genes without using any significance cutoff (**Figure 1e, Figure 1-figure supplement 4**).

Permutation test:

We again used a permutation scheme to assess the significance of shared age-related genes to account for the dependence among tissues. We tested the significance of shared up- and down-regulated genes, selected with or without an FDR cutoff, in development and in ageing periods separately. We used the age-related expression change values (ρ'_{gene}) calculated by permuting individual ages, 1000 times. To test the significance of the overlap of significantly up- or down- regulated genes (FDR corrected p-value<0.1) among tissues, we used the following procedure: (i) For each permutation round, we ranked the ρ'_{gene} values for each tissue in each period separately. (ii) We chose the highest N_u (to test the up-regulation), or lowest N_d (to test the down-regulation) number of genes, where N_u and N_d are the number of significantly up- or down- regulated genes, respectively, in a given tissue (FDR corrected p-value<0.1). (iii) For each permutation round, we calculated the number of overlaps across tissues using the chosen gene sets, i.e. the number of tissues with the same direction of expression change with age for those genes. Doing this for 1000 permutation results yielded a null distribution representing the expected overlaps if there were no age effect. (iv) We calculated the p-value as the proportion of 1000 permutations where the number of overlaps was higher than the observed value. The estimated false-positive-proportion (eFPP) was calculated as the median number of overlaps in permutations divided by the observed value.

Likewise, to test the significance of the overlap of shared up- and down-regulated genes selected without FDR cutoff, we used the same permutation scheme explained above, but this time using all the age-related expression changes created using permutations (ρ'_{gene}), without applying a significance cutoff for any tissue, and calculating the overlap across tissues in the same way.

Functional Associations:

We tested the functional associations of shared expression change trends among tissues in each period, separately, following the GORA procedure using the same criteria and algorithms explained in the previous section. To test shared up-regulated (n=45) or down-regulated genes (n=138) in development, we chose all significant age-related genes across tissues (n=10,305) in the development period as background. Since we could not identify any shared ageing-related genes across tissues (**Figure 1-figure supplement 5**), we did not perform a functional test for the ageing period.

Analysis of gene expression reversals

We compared the direction of gene expression change during development and during ageing to identify reversal genes in each tissue, separately. Genes showing up-regulation (positive correlation with age) in development and down-regulation (negative correlation with age) in ageing were assigned as up-down (UD) reversal genes, while the genes with the opposite trend (down-regulation in development and up-regulation in ageing) were assigned as down-up (DU) reversal genes. Without using any significance level for expression-age correlation values, we calculated the proportion of genes showing reversal by keeping the expression change direction in development the same, *i.e.* UD%=UD/(UU+UD) and DU%=DU/(DD+DU).

Permutation test:

To test the significance of reversal proportions, we kept the developmental changes constant and randomly permuted the individual ages only in the ageing period (as described earlier). Among developmental up-regulated genes, we calculated the UD% in each permutation, simulating a null distribution for UD reversal. We applied the same principle for the DU genes. Thus, we created a null distribution with the expected reversal ratios and tested the significance of observed values for each tissue separately (**Figure 1-figure supplement 8**).

Functional associations:

We used the GORA procedure as described earlier to test functional associations of reversal genes in each tissue but kept the developmental changes constant in the background. More specifically, we

tested the functional enrichment of UD reversal genes against UU genes, and DU genes against DD genes. We thereby specifically test the functions associated with the reversal pattern, but not development-associated functions.

Overlap of reversal genes - permutation test:

We tested the significance of overlap using the same permutation scheme described above. Specifically, among developmental up- (or down-) regulated genes shared among tissues, we constructed null distributions by calculating the ratio of UD vs UD+UU (or DU vs DU+DD) genes shared among tissues, identified in 1000 random permutations of individual ages only in the ageing period. (**Figure 1-figure supplement 9**). The number of shared up-regulated genes was n_{up}=2,255 (one gene excluded since it has constant expression in one tissue in ageing period), and the number of shared down-regulated genes was n_{down}=2,209.

Tissue convergence and divergence calculations using coefficient of variation (CoV)

For each individual mouse, for each gene (n=15,063), we calculated the inter-tissue coefficient of variation (CoV) estimate using normalised expression levels from the four tissues, dividing the standard deviation by the mean. We studied inter-tissue expression-variation change with age in development and ageing periods separately, using two approaches: (a) using the change in mean or median CoV across genes, and (b) studying significant CoV patterns at the single gene level.

Mean/median CoV across all genes:

We assessed transcriptome-wide variation among the tissues of each individual mouse by calculating the mean (or median) CoV of genes and then performing the Spearman's correlation test between mean-CoV (or median-CoV) and individual age.

CoV at the single gene level:

In the second approach, we tested the correlation between the CoV value of a gene and individual age for each commonly expressed gene using the Spearman's correlation test. P-values were corrected for multiple testing, using the 'BH' procedure. We used FDR corrected p-value<0.1 as cutoff. The genes showing positive correlation between CoV and age were called "divergent", and the ones showing

875 negative correlation were called “convergent” **(Figure 2b)**. Genes that display a divergent pattern during
876 development and convergent pattern in ageing (without using a significance level) were called
877 divergent-convergent (DiCo) genes (n=4,802).
878
879 *Permutation Test:*
880 To test the significance of DiCo genes (n=4,802), we kept the developmental divergent genes constant
881 (n=9,058, without a significance cutoff) and randomly permuted the individual ages only in the ageing
882 period (as described earlier). Among developmental divergent genes, we calculated the DiCo% for each
883 permutation, simulating a null distribution for the DiCo pattern (**Figure 2-figure supplement 12**).
884
885 *Clustering of DiCo genes:*
886 We used the k-means algorithm to cluster DiCo genes according to their CoV or expression changes
887 with age, separately (**Figure 2-figure supplement 2-3**). To find the optimum number of clusters for
888 both procedures, we applied gap statistics using the ‘clusGap’ function in the ‘cluster’ package v.2.1.2
889 with 500 simulations (Tibshirani, Walther, and Hastie 2001). We used the ‘kmeans’ function in base R
890 with ‘iter.max=20’ and ‘nstart=50’ parameters to cluster CoV values or expression levels which were
891 standardised to mean=1 and sd=0 across genes.
892
893 *Effect of gene expression trajectories on DiCo:*
894 To identify potential non-monotonic expression changes with age that could not be detected with the
895 Spearman’s correlation coefficient, we clustered all expressed genes (n=15,063) in each tissue,
896 separately, using the k-means algorithm following the same steps explained above (**Figure 1-figure**
897 **supplement 12-15**). The list of genes belonging to each cluster is given in **Figure 2-source data**. Then,
898 for each cluster, separately in each tissue, we performed a Fisher’s exact test to assess if a particular
899 cluster pattern is enriched or depleted in DiCo genes relative to all other expressed genes (the
900 background).
901
902 *Functional association analysis:*
903 To test the functional associations of the genes showing the DiCo pattern among tissues, we performed
904 GSEA using GO BPs. We retrieved developmental divergent genes (with $\rho_{CoV\text{-}age}>0$, n=9,058) and

multiplied these $\rho_{CoV\text{-}age}$ values with the ones calculated in the ageing period. Therefore, the genes with a negative value represent a DiCo pattern, while the ones with a positive value represent a divergent-divergent (DiDi) pattern. We then ranked the genes according to the calculated product values and sought enrichment for the upper and lower tail of the distribution using the Kolmogorov-Smirnov (KS) test implemented in the 'clusterProfiler' package v.4.0.0 (Yu et al. 2012). The 'gseGO' function was used with parameters: "nPerm=1000, minGSSize=10, maxGSSize=500 and pValueCutoff=1". Therefore, the enriched categories for the genes in the lower tail of the distribution would represent DiCo enrichment. Categories with FDR corrected p-value<0.1 were considered as significant.

We summarised DiCo enriched categories into representative ones following (Dönertaş et al. 2021) and used hierarchical clustering on gene similarities among categories. The tree was cut into 25 clusters. For each cluster, we chose as representative the category that has the highest mean Jaccard similarity to the other categories in the same cluster. Then, we calculated the mean age-expression correlation across all the genes in each representative category, in each tissue and in each period. As the unrelated categories, those with the low within cluster similarity, were grouped into one cluster, we denoted them 'Other GO', and performed the same clustering steps to further summarise them (**Figure 4-figure supplement 1**).

We further sought functional enrichment among DiCo genes that were clustered with the k-means algorithm for both CoV and expression clusters, separately (**Figure 2-figure supplement 2-3**). Genes in each cluster were tested among all DiCo genes using the same GORA procedure as described before.

Jackknife to test the Di/Co ratio between dev and ageing:

We tested the significance of divergent/convergent gene ratios using a jackknife resampling procedure in development and in ageing periods, separately. Leaving out an individual in each iteration, we re-calculated the number of significant divergent and convergent genes and their ratios. As we could not obtain any gene with significant CoV changes when the youngest adults were left-out due to the decreased power, standard error and confidence interval calculation was not possible. Instead, we report the range of pseudovalues. We note that the range of ratios in leave-out samples do not contain

the value 1 either in the development (0.41-0.49) or in the ageing (1.20-2.83) period (**Figure 2e**).

Pairwise tissue divergence-convergence test

In order to further verify the inter-tissue divergent-convergent pattern that we observed between development and ageing periods, we used a different approach based on expression correlations among tissues. We calculated pairwise Spearman's correlation coefficients among tissues of the same individual mouse, using all commonly expressed genes among the tissues (n=15,063). For each tissue pair, we tested the correlation between age and inter-tissue expression-correlations using the Spearman's correlation test in development and in ageing periods, separately. In addition, we calculated the mean (or median) of all six pairwise tissue correlations for each individual mouse, and tested the correlation between age and average inter-tissue expression-correlations using the Spearman's correlation test (**Figure 2-figure supplement 6**).

Determination of tissue-specific genes

To identify which tissue(s) contribute to the reversal pattern, we assigned each gene to a tissue to identify tissue-specific expression patterns. First, we calculated an effect size (ES) between the expression of a gene in a tissue versus other three tissues using the development samples only, and repeated this procedure for all tissues. Hence, we obtained ES for each commonly expressed gene in each tissue. ES was calculated using the 'Cohen's d' formula defined as the difference between the two means divided by the pooled standard deviation. We then assigned each gene to a tissue in which the gene has the highest ES. Finally, we retrieved only the fourth quartile (>Q3) of genes assigned to a tissue to define tissue-specific expression. Using this approach, we identified 3,766 tissue-specific genes in total (cortex: 1,175, lung: 839, liver: 986, muscle: 766 genes).

Enrichment test with the direction of age-related change:

We tested the association between tissue-specificity and age-related expression change during ageing using Fisher's exact test. Specifically, we constructed a contingency table with two categorical variables; the first variable defines the direction (either positive or negative) of maximum expression change during ageing identified in a tissue-specific gene, which is determined by the slope of the regression between log2 age and expression. The second variable defines whether this maximum expression change

identified in a tissue-specific gene occurs in its native tissue or not (either yes or no). Hence, a positive odds ratio (OR) suggests that (a) either the expression of genes decrease the most in their native tissue, and/or (b) the expression of genes increase the most in a non-native tissue during ageing.

Enrichment of tissue-specific genes in DiCo genes:

We tested the association between tissue-specificity [being either tissue-specific (n=3,766) or not (n=11,297)] and the DiCo pattern [either showing DiCo (n=4,802) or not (n=10,261)] using the Fisher's exact test, calculating the enrichment of tissue-specific genes within DiCo genes.

Additional publicly available bulk tissue transcriptome datasets

Jonker:

We downloaded the raw data from the GEO database with GSE34378 accession number (Jonker et al. 2013) and followed the same analysis pipeline described above using all the samples from 5 tissues ("Brain - Cortex", "Lung", "Liver", "Kidney", "Spleen") of 18 female mice comprising 90 samples in total. This dataset represents the ageing period of the mouse, ranging from 90 to 900 days. Using the oligo package v.1.56.0 (Carvalho and Irizarry 2010), we retrieved the expression matrices and performed "rma" normalisation followed by removing the probesets that were annotated to more than one gene. We confined the analysis to only the protein-coding genes expressed in at least 25% of all samples. The resulting 17,661 genes were log2 transformed (after adding 1) and quantile normalised using the preprocessCore library (Bolstad 2020) across all samples. Downstream analysis was the same as described above.

Schaum:

We downloaded the raw count matrix from the GEO database with GSE132040 accession number (Schaum et al. 2020) and performed the same filtrating steps as described above. We discarded the samples that have less than 4 million reads which was the cutoff used in the article. We restricted the analysis to only protein-coding genes expressed in at least 25% of the samples that have expression in 4 tissues ("Brain", "Lung", "Liver", "Muscle"). One individual was removed from the analysis due to being an outlier in PCA analysis after visual inspection (mouse ID: '3m7', PCA plots before and after outlier removal are present in our github repository). Final dataset contained 16,806 protein-coding

genes from 37 mice that range from 3 to 27 months of age covering the ageing period. There were 11 female mice ranging from 3 to 21 months of age and 26 male mice ranging from 3 to 27 months of age. We performed the same normalisation method and downstream analyses described above. We extended the analysis to 8 tissues ("Brain", "Heart", "Kidney", "Liver", "Lung", "Muscle", "Spleen", "Subcutaneous Fat") which were chosen based on the highest number of individuals that have the same tissue samples and that cover the whole ageing period (3 to 27 months). For the fat tissue, "Subcutaneous Fat" was chosen as representative tissue which has the highest number of samples among all minor fat tissues. After performing the same preprocessing steps explained above, the final dataset contained 17,619 genes from 26 mice. Downstream analysis was the same as above.

GTEx:

We downloaded the processed GTEx v8 dataset (GTEx Consortium et al. 2017) from the data portal and repeated the analysis in human tissues. We first confirmed our results in the same 4 tissues ("Brain - Cortex", "Lung", "Liver", "Muscle - Skeletal") and then expanded the analysis to 10 tissues ("Adipose - Subcutaneous", "Artery - Tibial", "Brain - Cerebellum", "Lung", "Muscle - Skeletal", "Nerve - Tibial", "Pituitary", "Skin - Sun Exposed (Lower leg)", "Thyroid", "Whole Blood"). In order to choose which tissues to analyse, we first choose the minor tissues with the highest number of samples for each major tissue, which prevents the representation of the same tissue multiple times. We then performed hierarchical clustering of tissues based on the presence of samples from the same individuals (**Figure 2-figure supplement 13**) and cut the tree into 3 clusters based on visual inspection. We selected the cluster with the highest number of overlapping individuals to analyse. The same procedure was followed for both 4- and 10-tissue analyses. In particular, we restricted the analysis to the individuals with samples in all tissues analysed and with a death circumstance of 1 (violent and fast deaths due to an accident) and 2 (fast death of natural causes) on the Hardy Scale (n =47 for 4 tissue, n=35 for 10 tissue). We removed duplicated genes from the analysis. Similar to our analysis with the mice data, we used only the protein-coding genes that are expressed in at least 25% of all samples, totalling 16,197 for 4 tissues and 16,305 for 10 tissues. The TPM values obtained from the GTEx data portal were log2 transformed (after adding 1), and quantile normalised using the preprocessCore library (Bolstad 2020) in R. Downstream analysis was the same as other datasets. To study the sex-specific convergence patterns, we repeated the same analysis separating female (n=11) and male (n=36) individuals.

Table 2. Dataset characteristics summarising species, tissues, number of individuals, age range, sex, and platform used for measuring gene expression values.

Dataset	Species	Tissues	N	Age range	Sex	Method
Izgi et al. 4 tissues	Mice	Brain, lung, liver, muscle	8	3 to 30 months	Male	RNAseq
Jonker et al. 5 tissues	Mice	Brain, lung, liver, kidney, spleen	18	3 to 30 months	Female	Microarray
Schaum et al. 4 tissues	Mice	Brain, lung, liver, muscle	37	3 to 27 months	Male (n=26) Female (n=11)	RNAseq
Schaum et al. 8 tissues	Mice	Brain, lung, liver, muscle, subcutaneous fat, kidney, heart, spleen	26	3 to 27 months	Male (n=20) Female (n=6)	RNAseq
GTEx 4 tissues	Humans	Brain, lung, liver, muscle	47	20 to 75 years	Male (n=36) Female (n=11)	RNAseq
GTEx 10 tissues	Humans	Adipose, tibial artery, cerebellum, lung, skeletal muscle, tibial nerve, pituitary, sun-exposed skin, thyroid, and whole blood	35	20 to 75 years	Male (n=27) Female (n=8)	RNAseq

Comparison of datasets

We compared the age-related expression change patterns across tissues of all datasets analysed using Spearman's correlation coefficient. We used the 'pheatmap' function from pheatmap package v1.0.12 (Raivo 2019) using hierarchical clustering (**Figure 4-figure supplement 2a**).

We performed Fisher's exact test to test the enrichment of convergent genes among datasets during ageing. We used only the convergent genes in ageing in our dataset (n=7,748) for comparison. For GTEx and Schaum et al. datasets, we performed enrichment for the same four tissues as our dataset and also for the larger sets, indicated as GTEx10 and Schaum8, respectively (**Figure 4-figure supplement 2b**).

Regulatory analysis

We used MiRTarBase (downloaded in 03/08/2021) (Hsu et al. 2010, 2014) and TRANSFAC (downloaded in 03/08/2021) (Matys et al. 2003, 2006) resources from the Ma'ayan lab database (Rouillard et al. 2016) for miRNA and transcription factor binding site (TFBS) enrichment analyses, respectively. As the database contains target information only for human HGNC IDs, we first converted

1045 those IDs to human Ensembl IDs and then to mouse Ensembl IDs only for the one-to-one ortholog
1046 genes, using 'getBM' and 'getLDS' functions from the biomaRt package. In total, we analysed 235
1047 miRNAs associated with 5,458 target genes and 158 TFs associated with 7,427 target genes. We
1048 conducted the overrepresentation analysis in the same way as for the DiCo functional enrichment
1049 analysis: specifically, we tested the targets of each regulator for enrichment in -Co genes (convergent
1050 genes in ageing) among Di- genes (divergent genes in development) used as background to keep
1051 developmental patterns fixed. We restricted the analysis for miRNA and TFs that have at least 5 target
1052 genes. After multiple testing correction with the BH procedure, we found no enrichment among either
1053 of the regulator types. Enrichment results are given in **Figure 4-source data**.
1054
1055 **Heteroscedasticity tests on the DiCo pattern**
1056 To test the hypothesis that the convergence pattern observed in the ageing period could be explained
1057 by the increased noise with age, thus regression towards the mean, we performed two distinct
1058 heteroscedasticity tests to compare DiCo genes against the lifelong-divergent genes (DiDi). In the first,
1059 we followed the method used to measure heteroscedasticity in Isildak et al. (2020) and Kedlian et al.
1060 (2019). We first fit a linear model between log2 transformed age and expression level, for each gene in
1061 each tissue (Kedlian et al. 2019; Işıldak et al. 2020; Somel et al. 2006). This represents the variability
1062 of error along the explanatory variable, age. Then, we calculated Spearman's correlation coefficient
1063 between the absolute residual values and age, which can be used as an estimate of heterogeneity
1064 change with age. We compared the heterogeneity change values of DiCo and DiDi genes using a two-
1065 sided KS test in each tissue. In the second approach, we used the 'ncvTest' function from the 'car'
1066 package v.3.0.11 (Fox and Weisberg 2018) which is a chi-squared test for heteroscedasticity estimated
1067 using a linear model. Again, we compared the heteroscedasticity measures of DiCo and DiDi genes
1068 using a two-sided KS test in each tissue.
1069
1070 **Single-cell RNAseq**
1071 *Preprocessing:*
1072 We used the Tabula Muris Senis dataset (Tabula Muris Consortium 2020) for scRNA-seq analysis as
1073 it is the only dataset to our knowledge that includes time-series samples covering old age, and the
1074 tissues present in our dataset. Seurat-processed FACS data of the tissues lung, liver, skeletal muscle

and non-myeloid brain were downloaded from the figshare database (Pisco 2020). The Seurat package v.4.0.0 (Stuart et al. 2019) was used to retrieve the expression matrix of the cells that are annotated to cell types in the original article. Each tissue contains samples from three time points: 90 (3m), 540 (18m) and 720 (24m) days-old mice, totalling 14 samples each in lung, liver and brain, and 9 samples in liver. We excluded cell types with less than 15 cells among all samples, and excluded genes if the expression level is 0 for all cells at a given age. This resulted in a median number of 99-382 cells assigned to cell types, 6-24 cell types and 16,951-22,122 genes across tissues. Using 3-month-old mice, we calculated cell type-specific expressions in each tissue. Specifically, we first calculated the mean expression levels among cells of an individual mouse for each cell type, and then calculated the mean among individuals to obtain an average expression value for each cell type. Uniprot gene symbols were converted to Ensembl gene IDs using the "biomaRt" R package (Durinck et al. 2009).

Deconvolution:

We used cell type-specific expression profiles of 3-month-old mice to estimate relative contributions of cell types to the transcriptome profiles of tissues in our mouse dataset. For a given tissue in our mouse dataset, we used single cell expression profiles of that tissue from the Tabula Muris Senis dataset. We used a linear regression-based deconvolution method for each tissue using three genesets: all genes (n=[12,492, 12,849]), DiCo genes (n=[4,007, 4,106]) and non-DiCo genes (n=[8,485, 8,743]). Regression coefficients were used as relative contributions of cell types according to the following linear model:

$$Yi = a + b_{j1}*X_{i1} + b_{j2}*X_{i2} + \ldots + b_{jn}*X_{in},$$

where i represents the tissue,

Y_i is the expression level of a sample in a tissue,

$b_{j1\ldots jn}$ represent the relative contributions of the n cell types in a tissue,

$X_{i1\ldots in}$ represent the expression levels of the n cell types in a tissue.

We then tested the effect of age on cell type contributions ($b_{j1,}\ldots b_{jn}$) using the Spearman's correlation test in development and in ageing.

Cell type similarities and their change during ageing:

To investigate the contribution of cell autonomous changes to inter-tissue convergence in ageing, we calculated pairwise cell type expression correlations among tissues and studied how these correlations change with age. Based on pairwise correlations in the 3-months age group, we identified the maximally and minimally correlated cell type pairs among tissues. Specifically, for each cell type in a given tissue, we chose the minimally correlated cell type in each of the other three tissues. For example, for each of the 10 cell types in the liver, we chose the minimally correlated cell type among the 15 cortex cell types, the minimally correlated cell type among the 24 lung cell types, and the minimally correlated cell type among the 6 muscle cell types. We repeated this procedure for all cell types in all four tissues, resulting in 54 cell type pairs. Then, we calculated Spearman's correlation coefficients between age and minimally correlated cell type pairs identified in the 3-months age group. Likewise, we repeated the same analysis for the maximally correlated cell type pairs among tissues.

Permutation tests:

To test whether DiCo genes are significantly more associated with cell type proportion changes than non-DiCo genes, we performed a permutation test based on a re-sampling procedure. For each tissue, we took random samples among all genes (n=[12,492, 12,849]) with size *N,* where *N* is the number of DiCo genes in that tissue, and repeated the deconvolution analysis as explained above. By calculating cell type proportion changes with age for each random sample repeated 1000 times, we created the null distribution for each cell type. Then, we calculated the p-values as the number of random samples having the same or higher cell type proportion change values divided by the observed value (cell type proportion changes with DiCo genes).

We applied a similar permutation scheme as explained above to test cell type similarity change differences between DiCo and non-DiCo genes. For each random sample of non-DiCo genes with size *N*, we calculated the pairwise correlations among cell types of tissues and identified maximally and minimally correlated cell types in the 3-months age group. Then, we calculated age-related changes of those correlations using Spearman's correlation coefficient to construct the null distribution.

Analysis of within-tissue convergence of cell types:

Analogous to inter-tissue convergence analysis, we also studied intra-tissue convergence of cell types in scRNA-seq data by calculating CoV among cell types within a tissue for each individual of ages 3m, 18m and 24m, separately. We filtered the data to obtain cell types present in at least 2 individual mice in every time point for each tissue which yielded 4, 7, 20 and 6 cell types in brain, liver, lung and muscle, respectively. We then tested the mean CoV (or CoV per gene) change with age using Spearman's correlation test.

Ethics statement

Post-mortem samples were obtained from 16 C57BL/6J mice aged between 2 days and 904 days. All mouse experiments were overseen by the Institutional Animal Welfare Officer of the Max Planck Institute for Evolutionary Anthropology (MPI-EVA). They were performed according to the German Animal Welfare Legislation, ("Tierschutzgesetz") and registered with the Federal State Authority Landesdirektion Sachsen (No. 24-9162. 11-01 (T62/08)). The mice were sacrificed for reasons independent of this study, their tissues were harvested and frozen immediately, and stored at -80°C.

Competing Interests

The authors report no competing interests.

Data Availability

Raw and processed RNA-seq data have been deposited in GEO with the accession number GSE167665. All summary statistics and analysis output are also provided as supplementary tables.

Code Availability

All the code used to perform analyses is available in GitHub: https://github.com/hmtzg/geneexp_mouse

Author Contributions

H.I., H.M.D, M.S. and P.K. conceived and conceptualised the study. H.I, H.M.D, and M.S. design the models to analyse the data. U.I. and H.M.D validated the results. H.I. analysed the data with

contributions by U.I., H.M.D, and E.K. D.D.H conducted the experiments and produced the raw data. S.H. helped D.D.H. with experimental procedures. H.I., U,I., H.M.D. curated and visualised the data. H.I., H.M.D., and M.S. wrote the manuscript. M.S and P.K. acquired funding for the experiments. H.M.D., M.S., and P.K. supervised and administered the study. All authors read, revised and approved the final version of this manuscript.

Acknowledgement

We thank Wolfgang Enard and Wulf Hevers for help with the mouse experiments and sharing samples, Nurcan Tuncbag, Nihal Terzi Çizmecioğlu, and the whole METU CompEvo team for helpful comments and fruitful discussions, and Zeliha Gözde Turan and Melih Yıldız for the critical reading of the manuscript and their suggestions.

Funding Statement

This work was supported by EMBL (H.M.D), the Scientific and Technological Research Council of Turkey (TÜBİTAK 2232, M.S.), the Science Academy (of Turkey) BAGEP Award (M.S.), and a METU Internal Grant (BAP, M.S.).

References

Alexa, Adrian, and Jorg Rahnenfuhrer. 2019. “topGO: Enrichment Analysis for Gene Ontology.”

Anders, Simon, and Wolfgang Huber. 2010. “Differential Expression Analysis for Sequence Count Data.” *Genome Biology* 11 (10): R106.

Anders, Simon, Paul Theodor Pyl, and Wolfgang Huber. 2014. “HTSeq—a Python Framework to Work with High-Throughput Sequencing Data.” *Bioinformatics* 31 (2): 166–69.

Andrews, Simon. 2010. “FastQC: A Quality Control Tool for High Throughput Sequence Data.” 2010. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Angelidis, Ilias, Lukas M. Simon, Isis E. Fernandez, Maximilian Strunz, Christoph H. Mayr, Flavia R. Greiffo, George Tsitsiridis, et al. 2019. “An Atlas of the Aging Lung Mapped by Single Cell Transcriptomics and Deep Tissue Proteomics.” *Nature Communications* 10 (1): 963.

Anisimova, Aleksandra S., Mark B. Meerson, Maxim V. Gerashchenko, Ivan V. Kulakovskiy, Sergey

1192 E. Dmitriev, and Vadim N. Gladyshev. 2020. “Multifaceted Deregulation of Gene Expression and
1193 Protein Synthesis with Age.” *Proceedings of the National Academy of Sciences of the United*
1194 *States of America*, June. https://doi.org/10.1073/pnas.2001788117.
1195 Ashburner, M., C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, et al. 2000.
1196 “Gene Ontology: Tool for the Unification of Biology. The Gene Ontology Consortium.” *Nature*
1197 *Genetics* 25 (1): 25–29.
1198 Bahar, Rumana, Claudia H. Hartmann, Karl A. Rodriguez, Ashley D. Denny, Rita A. Busuttil, Martijn
1199 E. T. Dollé, R. Brent Calder, et al. 2006. “Increased Cell-to-Cell Variation in Gene Expression in
1200 Ageing Mouse Heart.” *Nature* 441 (7096): 1011–14.
1201 Benjamini, Yoav, and Yosef Hochberg. 1995. “Controlling the False Discovery Rate: A Practical and
1202 Powerful Approach to Multiple Testing.” *Journal of the Royal Statistical Society. Series B,*
1203 *Statistical Methodology* 57 (1): 289–300.
1204 Blagosklonny, Mikhail V. 2006. “Aging and Immortality: Quasi-Programmed Senescence and Its
1205 Pharmacologic Inhibition.” *Cell Cycle* 5 (18): 2087–2102.
1206 Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for
1207 Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20.
1208 Bolstad, Ben. 2020. “preprocessCore: A Collection of Pre-Processing Functions.”
1209 https://github.com/bmbolstad/preprocessCore.
1210 Brawand, David, Magali Soumillon, Anamaria Necsulea, Philippe Julien, Gábor Csárdi, Patrick
1211 Harrigan, Manuela Weier, et al. 2011. “The Evolution of Gene Expression Levels in Mammalian
1212 Organs.” *Nature* 478 (7369): 343–48.
1213 Cardoso-Moreira, Margarida, Jean Halbert, Delphine Valloton, Britta Velten, Chunyan Chen, Yi Shao,
1214 Angélica Liechti, et al. 2019. “Gene Expression across Mammalian Organ Development.” *Nature*
1215 571 (7766): 505–9.
1216 Carvalho, Benilton S., and Rafael A. Irizarry. 2010. “A Framework for Oligonucleotide Microarray
1217 Preprocessing.” *Bioinformatics* 26 (19): 2363–67.
1218 Colantuoni, Carlo, Barbara K. Lipska, Tianzhang Ye, Thomas M. Hyde, Ran Tao, Jeffrey T. Leek,
1219 Elizabeth A. Colantuoni, et al. 2011. “Temporal Dynamics and Genetic Control of Transcription in
1220 the Human Prefrontal Cortex.” *Nature* 478 (7370): 519–23.
1221 Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha,

1222 Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. “STAR: Ultrafast Universal RNA-
1223 Seq Aligner.” *Bioinformatics* 29 (1): 15–21.
1224 Dönertaş, Handan Melike, Daniel K. Fabian, Matías Fuentealba Valenzuela, Linda Partridge, and
1225 Janet M. Thornton. 2021. “Common Genetic Associations between Age-Related Diseases.”
1226 *Nature Aging* 1 (4): 400–412.
1227 Dönertaş, Handan Melike, Matías Fuentealba Valenzuela, Linda Partridge, and Janet M. Thornton.
1228 2018. “Gene Expression-Based Drug Repurposing to Target Aging.” *Aging Cell* 17 (5): e12819.
1229 Dönertaş, Handan Melike, Hamit İzgi, Altuğ Kamacıoğlu, Zhisong He, Philipp Khaitovich, and Mehmet
1230 Somel. 2017. “Gene Expression Reversal toward Pre-Adult Levels in the Aging Human Brain and
1231 Age-Related Loss of Cellular Identity.” *Scientific Reports* 7 (1): 5894.
1232 Durinck, Steffen, Paul T. Spellman, Ewan Birney, and Wolfgang Huber. 2009. “Mapping Identifiers for
1233 the Integration of Genomic Datasets with the R/Bioconductor Package biomaRt.” *Nature*
1234 *Protocols* 4 (8): 1184–91.
1235 Enge, Martin, H. Efsun Arda, Marco Mignardi, John Beausang, Rita Bottino, Seung K. Kim, and
1236 Stephen R. Quake. 2017. “Single-Cell Analysis of Human Pancreas Reveals Transcriptional
1237 Signatures of Aging and Somatic Mutation Patterns.” *Cell* 171 (2): 321–30.e14.
1238 Ezcurra, Marina, Alexandre Benedetto, Thanet Sornda, Ann F. Gilliat, Catherine Au, Qifeng Zhang,
1239 Sophie van Schelt, et al. 2018. “C. Elegans Eats Its Own Intestine to Make Yolk Leading to
1240 Multiple Senescent Pathologies.” *Current Biology: CB* 28 (20): 3352.
1241 Feser, Jason, David Truong, Chandrima Das, Joshua J. Carson, Jeffrey Kieft, Troy Harkness, and
1242 Jessica K. Tyler. 2010. “Elevated Histone Expression Promotes Life Span Extension.” *Molecular*
1243 *Cell* 39 (5): 724–35.
1244 Fisher, R. A. (1930). *The genetical theory of natural selection. 272.*
1245 https://doi.org/10.5962/bhl.title.27468
1246 Flurkey, Kevin, Joanne M. Currer, and D. E. Harrison. 2007. “Chapter 20 - Mouse Models in Aging
1247 Research.” In *The Mouse in Biomedical Research (Second Edition)*, edited by James G. Fox,
1248 Muriel T. Davisson, Fred W. Quimby, Stephen W. Barthold, Christian E. Newcomer, and Abigail
1249 L. Smith, 637–72. Burlington: Academic Press.
1250 Fox, John, and Sanford Weisberg. 2018. *An R Companion to Applied Regression*. SAGE
1251 Publications.

1252 Gems, David, and Linda Partridge. 2013. “Genetics of Longevity in Model Organisms: Debates and
1253 Paradigm Shifts.” *Annual Review of Physiology* 75: 621–44.
1254 GTEx Consortium, Laboratory, Data Analysis &Coordinating Center (LDACC)—Analysis Working
1255 Group, Statistical Methods groups—Analysis Working Group, Enhancing GTEx (eGTEx) groups,
1256 NIH Common Fund, NIH/NCI, NIH/NHGRI, et al. 2017. “Genetic Effects on Gene Expression
1257 across Human Tissues.” *Nature* 550 (7675): 204–13.
1258 Hernando-Herraez, Irene, Brendan Evano, Thomas Stubbs, Pierre-Henri Commere, Marc Jan
1259 Bonder, Stephen Clark, Simon Andrews, Shahragim Tajbakhsh, and Wolf Reik. 2019. “Ageing
1260 Affects DNA Methylation Drift and Transcriptional Cell-to-Cell Variability in Mouse Muscle Stem
1261 Cells.” *Nature Communications* 10 (1): 4361.
1262 Hsu, Sheng-Da, Feng-Mao Lin, Wei-Yun Wu, Chao Liang, Wei-Chih Huang, Wen-Ling Chan, Wen-
1263 Ting Tsai, et al. 2010. “miRTarBase: A Database Curates Experimentally Validated microRNA–
1264 target Interactions.” *Nucleic Acids Research* 39 (suppl_1): D163–69.
1265 Hsu, Sheng-Da, Yu-Ting Tseng, Sirjana Shrestha, Yu-Ling Lin, Anas Khaleel, Chih-Hung Chou,
1266 Chao-Fang Chu, et al. 2014. “miRTarBase Update 2014: An Information Resource for
1267 Experimentally Validated miRNA-Target Interactions.” *Nucleic Acids Research* 42 (Database
1268 issue): D78–85.
1269 Işıldak, Ulaş, Mehmet Somel, Janet M. Thornton, and Handan Melike Dönertaş. 2020. “Temporal
1270 Changes in the Gene Expression Heterogeneity during Brain Development and Aging.” *Scientific
1271 Reports* 10 (1): 4080.
1272 Jonker, Martijs J., Joost Pm Melis, Raoul V. Kuiper, Tessa V. van der Hoeven, P. F. K. Wackers, Joke
1273 Robinson, Gijsbertus Tj van der Horst, et al. 2013. “Life Spanning Murine Gene Expression
1274 Profiles in Relation to Chronological and Pathological Aging in Multiple Organs.” *Aging Cell* 12
1275 (5): 901–9.
1276 Kedlian, Veronika R., Handan Melike Donertas, and Janet M. Thornton. 2019. “The Widespread
1277 Increase in Inter-Individual Variability of Gene Expression in the Human Brain with Age.” *Aging*
1278 11 (8): 2253–80.
1279 Kim, S., B. Villeponteau, and S. M. Jazwinski. 1996. “Effect of Replicative Age on Transcriptional
1280 Silencing near Telomeres in Saccharomyces Cerevisiae.” *Biochemical and Biophysical Research
1281 Communications* 219 (2): 370–76.

1282 Lind, Martin I., Sanjana Ravindran, Zuzana Sekajova, Hanne Carlsson, Andrea Hinas, and Alexei A.
1283 Maklakov. 2019. “Experimentally Reduced insulin/IGF-1 Signaling in Adulthood Extends Lifespan
1284 of Parents and Improves Darwinian Fitness of Their Offspring.” *Evolution Letters* 3 (2): 207–16.
1285 Liu, Xiling, Dingding Han, Mehmet Somel, Xi Jiang, Haiyang Hu, Patricia Guijarro, Ning Zhang, et al.
1286 2016. “Disruption of an Evolutionarily Novel Synaptic Expression Pattern in Autism.” *PLoS*
1287 *Biology* 14 (9): e1002558.
1288 Love, Michael I., Wolfgang Huber, and Simon Anders. 2014. “Moderated Estimation of Fold Change
1289 and Dispersion for RNA-Seq Data with DESeq2.” *Genome Biology* 15 (12): 550.
1290 Luegmayr, E., H. Glantschnig, G. A. Wesolowski, M. A. Gentile, J. E. Fisher, G. A. Rodan, and A. A.
1291 Reszka. 2004. “Osteoclast Formation, Survival and Morphology Are Highly Dependent on
1292 Exogenous Cholesterol/lipoproteins.” *Cell Death and Differentiation* 11 Suppl 1 (July): S108–18.
1293 Magalhães, João Pedro de, and George M. Church. 2005. “Genomes Optimize Reproduction: Aging
1294 as a Consequence of the Developmental Program.” *Physiology* 20 (August): 252–59.
1295 Martinez-Jimenez, Celia Pilar, Nils Eling, Hung-Chang Chen, Catalina A. Vallejos, Aleksandra A.
1296 Kolodziejczyk, Frances Connor, Lovorka Stojic, et al. 2017. “Aging Increases Cell-to-Cell
1297 Transcriptional Variability upon Immune Stimulation.” *Science* 355 (6332): 1433–36.
1298 Matys, V., E. Fricke, R. Geffers, E. Gößling, M. Haubrock, R. Hehl, K. Hornischer, et al. 2003.
1299 “TRANSFAC ® : Transcriptional Regulation, from Patterns to Profiles.” *Nucleic Acids Research*
1300 31 (1): 374–78.
1301 Matys, V., O. V. Kel-Margoulis, E. Fricke, I. Liebich, S. Land, A. Barre-Dirrie, I. Reuter, et al. 2006.
1302 “TRANSFAC® and Its Module TRANSCompel®: Transcriptional Gene Regulation in
1303 Eukaryotes.” *Nucleic Acids Research* 34 (suppl_1): D108–10.
1304 Medawar, P. B. (1953). Unsolved problem of biology. *The Medical Journal of Australia, 1*(24), 854–
1305 855.
1306 Pisco, Angela. 2020. “Tabula Muris Senis Data Objects.”
1307 https://doi.org/10.6084/m9.figshare.12654728.v1.
1308 Raivo, Kolde. 2019. “Pheatmap: Pretty Heatmaps.” *R Package Version* 1 (8).
1309 Rouillard, Andrew D., Gregory W. Gundersen, Nicolas F. Fernandez, Zichen Wang, Caroline D.
1310 Monteiro, Michael G. McDermott, and Avi Ma’ayan. 2016. “The Harmonizome: A Collection of
1311 Processed Datasets Gathered to Serve and Mine Knowledge about Genes and Proteins.”

Database: The Journal of Biological Databases and Curation 2016 (July). https://doi.org/10.1093/database/baw100.

Sampathkumar, Nirmal K., Juan I. Bravo, Yilin Chen, Prakroothi S. Danthi, Erin K. Donahue, Rochelle W. Lai, Ryan Lu, Lewis T. Randall, Nika Vinson, and Bérénice A. Benayoun. 2020. "Widespread Sex Dimorphism in Aging and Age-Related Diseases." *Human Genetics* 139 (3): 333–56.

Schaum, Nicholas, Benoit Lehallier, Oliver Hahn, Róbert Pálovics, Shayan Hosseinzadeh, Song E. Lee, Rene Sit, et al. 2020. "Ageing Hallmarks Exhibit Organ-Specific Temporal Signatures." *Nature* 583 (7817): 596–602.

Somel, Mehmet, Song Guo, Ning Fu, Zheng Yan, Hai Yang Hu, Ying Xu, Yuan Yuan, et al. 2010. "MicroRNA, mRNA, and Protein Expression Link Development and Aging in Human and Macaque Brain." *Genome Research* 20 (9): 1207–18.

Somel, Mehmet, Philipp Khaitovich, Sabine Bahn, Svante Pääbo, and Michael Lachmann. 2006. "Gene Expression Becomes Heterogeneous with Age." *Current Biology: CB* 16 (10): R359–60.

Stuart, Tim, Andrew Butler, Paul Hoffman, Christoph Hafemeister, Efthymia Papalexi, William M. Mauck 3rd, Yuhan Hao, Marlon Stoeckius, Peter Smibert, and Rahul Satija. 2019. "Comprehensive Integration of Single-Cell Data." *Cell* 177 (7): 1888–1902.e21.

Tabula Muris Consortium. 2020. "A Single-Cell Transcriptomic Atlas Characterizes Ageing Tissues in the Mouse." *Nature* 583 (7817): 590–95.

Tacutu, Robi, Daniel Thornton, Emily Johnson, Arie Budovsky, Diogo Barardo, Thomas Craig, Eugene Diana, et al. 2018. "Human Ageing Genomic Resources: New and Updated Databases." *Nucleic Acids Research* 46 (D1): D1083–90.

Tibshirani, Robert, Guenther Walther, and Trevor Hastie. 2001. "Estimating the Number of Clusters in a Data Set via the Gap Statistic." *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 63 (2): 411–23.

Trapnell, Cole, Brian A. Williams, Geo Pertea, Ali Mortazavi, Gordon Kwan, Marijke J. van Baren, Steven L. Salzberg, Barbara J. Wold, and Lior Pachter. 2010. "Transcript Assembly and Quantification by RNA-Seq Reveals Unannotated Transcripts and Isoform Switching during Cell Differentiation." *Nature Biotechnology* 28 (5): 511–15.

Turan, Zeliha Gözde, Poorya Parvizi, Handan Melike Dönertaş, Jenny Tung, Philipp Khaitovich, and Mehmet Somel. 2019. "Molecular Footprint of Medawar's Mutation Accumulation Process in

1342 Mammalian Aging." *Aging Cell* 348 (May): e12965.

1343 Williams, G. C. (1957). Pleiotropy, Natural Selection, and the Evolution of Senescence. *Evolution;*

1344 *International Journal of Organic Evolution, 11*(4), 398–411.

1345 Yang, Jae-Hyun, Patrick T. Griffin, Daniel L. Vera, John K. Apostolides, Motoshi Hayano, Margarita V.

1346 Meer, Elias L. Salfati, et al. 2019. "Erosion of the Epigenetic Landscape and Loss of Cellular

1347 Identity as a Cause of Aging in Mammals." *Cold Spring Harbor Laboratory*.

1348 https://doi.org/10.1101/808642.

1349 Yang, Jialiang, Tao Huang, Francesca Petralia, Quan Long, Bin Zhang, Carmen Argmann, Yong

1350 Zhao, et al. 2015. "Synchronized Age-Related Gene Expression Changes across Multiple

1351 Tissues in Human and the Link to Complex Diseases." *Scientific Reports* 5 (October): 15145.

1352 Yuan, Yuan, Yi-Ping Phoebe Chen, Jerome Boyd-Kirkup, Philipp Khaitovich, and Mehmet Somel.

1353 2012. "Accelerated Aging-Related Transcriptome Changes in the Female Prefrontal Cortex."

1354 *Aging Cell* 11 (5): 894–901.

1355 Yu, Guangchuang, Li-Gen Wang, Yanyan Han, and Qing-Yu He. 2012. "clusterProfiler: An R Package

1356 for Comparing Biological Themes among Gene Clusters." *Omics: A Journal of Integrative Biology*

1357 16 (5): 284–87.

1358 Zahn, Jacob M., Suresh Poosala, Art B. Owen, Donald K. Ingram, Ana Lustig, Arnell Carter, Ashani T.

1359 Weeraratna, et al. 2007. "AGEMAP: A Gene Expression Database for Aging in Mice." *PLoS Genetics*

1360 3 (11): e201.