Age-dependent changes in mean and variance of gene expression across tissues in a twin cohort

Ana Viñuela[1*+], Andrew A Brown[2,3], Alfonso Buil[4,5,6], Pei-Chien Tsai[1], Matthew N Davies[1], Jordana T Bell[1], Emmanouil T Dermitzakis[4,5,6], Timothy D Spector[1], Kerrin S Small[1*]

1) Department of Twin Research and Genetic Epidemiology, King's College London, St Thomas' Campus, Westminster Bridge Road, SE1 7EH London, United Kingdom.

2) Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK 3) NORMENT, KG Jebsen Centre for Psychosis Research, Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway

4) Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland

5) Institute for Genetics and Genomics in Geneva (iGE3), University of Geneva, 1211 Geneva, Switzerland

6) Swiss Institute of Bioinformatics, 1211 Geneva, Switzerland

+ Current address: Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland

*** Corresponding authors**

Summary:

Gene expression changes with age have consequences for healthy aging and disease development. Here we investigate age-related changes in gene expression measured by RNA-seq of fat, skin, whole blood and derived lymphoblastoid cell lines (LCLs) expression from ~800 adult female twins. We see evidence of up to 60% shared transcriptional age-related effects across tissues on level of expression and 47% on splicing; amongst these we highlight effects on genes involved in diseases such as Alzheimer and cancer. We identify 137 genes with age-related changes in variance and 42 genes with age-related discordance between genetically identical individuals; implying the latter are driven by environmental effects. Investigation of methylation effects observed a widespread and stronger effect of age on methylation than expression however we did not find a strong relationship between changes in both expression and methylation. In contrast, we find 4 significant examples where the ageing process is modulated by genetic variants. In summary, we demonstrate that aging affects the splicing, level and variance of expression, and that these processes can be both environmentally and genetically influenced.

Introduction

Aging is a complex process, characterized by a progressive decline in an organism's biological function and phenotype characteristics, which leads to an increased chance of developing disease and ultimately the death of the organism (Valdes et al. 2013). Others have attempted to understand the aging process by identifying common denominators of aging in different organisms (López-Otín et al.). Many of these hallmarks, such as genome instability, epigenetic alterations, loss of proteostasis and telomere attrition, are accompanied by changes in gene expression. Identification of genes differentially expressed genes with age has proven useful in identifying pathways whose behavior is modified by age, as well as identifying biomarkers of aging and therapeutic targets (de Magalhaes et al., 2009; Glass et al., 2013; Rodwell et al., 2004). Expression studies into aging using animal models and whole organisms have discovered that the expression of up to 75% of genes can be associated to aging (Viñuela et al., 2010). These modifications can occur by acting on the level of expression of genes, on the splicing of the mRNA produced or on the genetic regulation of gene expression (Li et al., 2009; Viñuela et al., 2010). On the other hand, human studies have only recently managed to identify thousands of genes associated with age in multiple tissues (Glass et al., 2013; Peters et al., 2015; Yang et al., 2015), but are still far away from identifying the same scale of aging effects in expression or the same variety of changes. Reasons for this include a reduced power to see interactions due to the uncontrolled human environment and inbred nature of model organisms, and importantly the lack of sufficient human expression data using appropriate technologies and tissues.

In this study, we investigate changes in gene expression with age using RNA-seq measurements of fat, skin, whole blood and derived lymphoblastoid cell lines (LCLs) expression from ~800 monozygous (MZ) and dizygous (DZ) adult female twins (Supp. Table S1). We take a comprehensive approach that includes not only an analysis of the effect of age on the mean of gene expression and alternative splicing, but also analyses using gene expression variance and discordance between genetically identical MZ twin pairs. Although age-related changes in variance of gene expression have been identified in animal models, we believe this is one of the first human study succeeding in the identification of specific genes changing variance with age (Bahar et al., 2006; Lu et al., 2004; McCarroll et al., 2004; Somel et al., 2006). We show also how environmental exposures on MZ siblings change expression over time and how the aging process is a complicated interplay between genetic variance and environmental factors. To explore this in detail we also studied methylation changes with age in the same samples from fat tissue and genotype-by-age interaction on gene expression. Finally, in comparison with previous studies, we observe a greater degree of sharing of age expression effects across tissues, reflecting the large sample, benefits of the twin design and the more accurate quantification provided by RNA-seq.

Results

Effects of aging in gene expression levels

To investigate the wide range of changes in gene expression with age, we used RNA-seq data from 855 healthy individuals drawn from the TwinsUK cohort (Supp. Table S1) in four tissues: i) photo protected skin, ii) subcutaneous fat, iii) whole blood and iv) lymphoblastoid cell lines (LCLs). We consider a gene associated with age if at least one exon was associated with chronological age. We discovered that 36.6% of tested genes (5,631 of 15,353) had at least one exon where expression levels was significantly associated with age in at least one

tissue (adjusted *P* value < 0.05; Figure 1A, Supplementary File 1). This number is roughly double that we previously reported (18.3%, 3,019 genes) using exactly the same skin, fat and LCLs samples but measuring expression using microarrays (Glass et al., 2013)(Figure 1A, Supp. Figure S1). We also found that the total number of expressed genes increased as a function of age in fat tissue (adjusted *P* value = 0.00264) but not in the other tissues, suggesting that the number of genes expressed at any given moment of life may be also a function of age in some tissues. Application of Gene Set Enrichment Analysis (GSEA) to the differentially expressed genes showed significant enrichment in GO terms (FDR < 0.05) related to RNA processing, fat metabolism and oxidation reduction in skin; and cell adhesion, membrane structure and sodium channel complex structure in fat tissue (Supplementary files 2). In blood, there was no specific enrichment for GO terms, and in LCLs only 7 genes showed significant association with age, 3 of which were previously reported (Glass et al., 2013). In conclusion, we identify thousands of genes whose expression was associated to age.

To quantify the relative effect of age on gene expression, we estimated the proportion of variance of exon expression levels (removing technical confounders) explained by age and additive genetic effects (heritability). In exons associated with age, age explained only a small proportion of the variation in gene expression, with median values between 2.2% and 5.7% depending on tissue and with maximum values ranging from 12% to 27% (Supplementary File 3). Globally, the effect of age on expression was greatest in blood, then skin, fat, and finally LCLs had the least. In comparison, the proportion of variance explained by additive genetic effects on the same set of age-affected exons was greater than that explained by age in all tissues (median $h^2_{skin} = 0.12$, $h^2_{fat} = 0.22$, $h^2_{LCLs} = 0.20$, $h^2_{blood} = 0.23$). A small global effect of age in gene expression may explain the difficulties in identifying biomarkers of aging in gene expression and suggest a need of larger sample sizes adjusting for genetic effects for such studies.

Splicing is associated with aging

As well as changes in the average level of expression, age can also cause changes in mRNA maturation and splicing. To identify changes in splicing with age, we quantified splicing based on link reads between exons using Altrans (Ongen and Dermitzakis, 2015a). As before, we consider the splicing of a gene to be associated with age if at least one link was associated with chronological age. We found a total of 904 genes (6.3% of the 14,261 genes with more than one exon expressed) with at least one link differentially expressed with age in either fat or skin (adjusted *P* value < 0.05, Supp. Files 4, 5 and 6). For 51.8% of those genes in skin and 11.4% in fat, age was associated with level of expression as well. We did not see significant associations between age and splicing in LCLs and blood, probably due to the smaller age effect in LCL expression and the smaller sample size available for blood (N = 384). Among genes with age-related differential splicing we found *APOE* (Figure 1-D and Figure 5)(previously associated to extreme longevity, Alzheimer disease and cholesterol metabolism), *LMNA* (causal of progeria, an accelerated aging syndrome, Figure 6), *HTRA2* (Parkinson disease) and *AAP* (Alzheimer disease). In fat tissue, many thrombospondins and collagen genes had differentially expressed links, as well as genes such as *AKT1* and *AKT2* from the insulin-IGF1 pathway, which is known to play a central role in aging. Overall, we observed that some age-related changes in gene expression were associated to changes in

splicing, but that the effect of age in splicing was wicker than the observed in expression levels.

Variance and differences in gene expression between MZ twins is dependent of age

Age-dependent changes in the variance of gene expression (rather than mean expression levels) has been reported in different model organisms (Bahar et al., 2006; Lu et al., 2004; McCarroll et al., 2004), but previous studies in humans have not been conclusive (Somel et al., 2006). Changes in phenotypic variance with age can be due to different responses to environment, age-related damage accumulation leading to stochastic deregulation of gene expression or gene-age interactions where changes in relative genetic effects can increase heterogeneity across the population at a particular age (Paré et al., 2010). We looked for changes in variance with age and identified 137 genes where expression showed age-dependent variance in at least one tissue (adjusted P *value* < 0.05, Figure 1B, Supplementary File 7). Since changes in phenotypic variance have mainly been reported to increase with age, we were surprised to observe that for the majority of these genes we report a decrease in variance of expression. The biological functions associated to the genes with age-associated differential variance in skin included oxidation reduction, with affected genes such as *SOD2*, fatty acid metabolism with genes including *CPT1B*, *ELOV3* and *ELOV5* or cell cycle control like *p21*. In blood, enriched pathways included the VEGF signaling pathway with the *PIK3CD* and *PXN* genes (Figure 1B). Our analysis shows concrete examples of age-related changes in phenotypic variance affecting expression in humans and identified changes in variance with age as another process by which aging may be linked to disease.

Changes in variance with age occur either as a consequence of environmental exposures or as a result of changes in genetic regulation of gene expression. Since MZ twins are genetically identical (and the same age), differences in expression levels within twin-pairs must have an environmental cause, allowing us to learn whether changes in variance with age were indeed influenced by the environment experienced by the twins. Therefore, and exploiting the twin design, we calculated the difference in expression between MZ twins (Supp. Table S1). We have successfully used this strategy previously to classify genetic determinants of phenotypic variance in gene expression (Brown et al., 2014) and GxE interactions affecting allelic specific expression (Buil et al., 2015). Here, we identified 42 genes where difference (discordance) in expression between MZ co-twins changed with age in at least one tissue (Figure 1C and Supp. File 8). Of the 34 genes identified in skin, 11 also showed a decrease in variance with age. This indicates that the observed change in variance for those genes was environmentally, and not genetically determined. However, for the remaining genes, either changing environments which were concordant across MZ twins, or GxE interactions remain plausible explanations for the change in variance. In conclusion, changes in phenotypic variation with age can be attributed to different environmental exposures among the individuals and not only to a general decline in regulatory functions and increase in genome damage with age, as others have suggested (Bahar et al., 2006).

Age-related associations in expression are modulated by genetic variation

Changes in variance in expression with age could also be a result of gene-by-age interactions affecting expression (GxA), when the genetic regulation of expression changes with age (Kent et al., 2012; Viñuela et al., 2010; Yao et al., 2014). To discover these effects, we searched for SNPs whose effect on expression depends on the age of the individual. It is well

known that the power to discover such second order effects is much reduced compared to standard main effects; for this reason it is common to restrict the search space to those with known main effects, either genetic or on aging (Wheeler and Kim, 2011). We used the latter approach, and chose not to restrict testing SNPs with known genetic effects as the strongest main-effect eQTL usually lie within promoter regions and are not often environmentally influenced. Therefore, we tested *cis*-GxA regulatory interaction effects for the 12,830 exons which were either 1) differentially expressed with age; 2) variance changes with age and 3) discordant in expression between MZ co-twins with age in fat or skin. After multiple testing corrections, we identified one significant gxa-eQTL, affecting the expression of *CD82* among the genes differentially expressed with age in fat (Figure 2). We also detected three gxa-eQTL among the genes that were discordant for expression in skin for the *CNKSR1*, *ACO1* and *ACSS2* genes (Supp. Figure S3 and Supplementary Files 9, 10, and 11). Despite the inherent challenges in identifying interaction effects, we here identify four GxA effects on gene expression with a relatively modest sample size. Given the many examples of GxA interactions reported in model organisms, we expected further studies with larger samples sizes to identify more examples.

The effect of age on methylation in fat tissue

Methylation levels and discordance in methylation between MZ twins has been shown to increase globally with age in promotor regions (Bell et al., 2012; Hannum et al., 2013). Given our findings of genes with age-dependent changes in variance and discordance of expression, and the difficulties of identifying genetic effects responsible for those changes, we postulated that epigenetic drift could explain some of the age-related changes in expression. Therefore, we used 552 Infinium HumanMethylation450 BeadChip methylation profiles generated from the same fat biopsies as the RNAseq data to investigate the role of methylation in age-related changes in gene expression (Grundberg et al.). Methylation data was not available in the other tissues. Firstly, using a linear mixed model we identified 39,092 differentially methylated regions (DMRs) with age from the 370,731 array probes (10.54%)(Supplementary File 12). The proportion of DMRs associated with age was significantly larger (*P* value < 2.2e-16, X^2 test) compared to the proportion of differentially expressed exons with age in fat RNAseq data (N = 1,511, 1.4%), showing that globally the association of aging with methylation levels is larger than for gene expression. Of the 39,092 age-DMRs, 93.6% were hypermethylated with a median variance in methylation explained by age of 3.27% (Figure 3). In total, 3,555 genes have an age-DMR near their TSS (<200 bp), of which 444 were also differentially expressed with age, suggesting a possible influence of methylation in age-related expression. Secondly, we looked for associations between expression of all exons and methylation probes at less than 200 bp distance from their TSS. From 297,702 pairs of exon-methylation probes in the proximity of the TSS, we found 4,853 to be significantly correlated, 53% negatively and 46.91% positively correlated with expression (Supplementary File 13). From those 4,853 exon-probe pairs, 16.8% of exons and 15.3% of methylation probes were differentially expressed or methylated with age. In conclusion, we observed a widespread and stronger effect of age on methylation than expression and a lower number of significant associations between expression and methylation.

To investigate whether interactions involving methylation markers could be responsible for changes of variance with age in gene expression, we looked for interactions between

methylation and age (methylation*age) affecting gene expression. Since only 3 genes had a significant association between variance in gene expression in fat and age at an adjusted P value < 0.05, we chose to relax our threshold to an adjusted P value < 0.10 and test nine genes. We identified a Bonferonni significant methylation x age interaction effect on expression of *IRS1* at three methylation probes, the most significant at probe *cg19451698* (*P* value = 6.6e-05, Figure 2, Supplementary File 14). This significant interaction implies that the expression of the *IRS1* gene decreases with age in individuals with *cg19451698* hypomethylated. Such an effect was not present in individuals with high levels of methylation in the same region. Homologs of *IRS1* and other members of the insulin/IGF-1 pathway are known to regulate longevity in model organisms, a function that may be conserved in humans due to their involved in age-related diseases like diabetes. In summary, we detected interacting effects of methylation with age modulating genes expression. This indicates that more work is necessary to unravel these complex interactions and the relationship between changes in methylation and expression with age.

Age effects in expression are shared across tissues

Previous studies performed in multiple tissues identified a limited number of shared genes associated with age across tissues (Glass et al., 2013; Melé et al., 2015; Yang et al., 2015). Similarly, of the 5,631 genes (36.67%) affected by age in at least one tissues, we were only able to identify five genes significantly associated to age in all the three primary tissues (Figure 4). By pair-wise comparisons between tissues we found that 274 was the largest number of genes significantly associated between two tissues (Figure 4B). However, this degree of overlap of associated exons across tissues was significant (*P* value < 1e-216, Fishers test) indicating the presence of a common signature of aging across tissues. Furthermore, defining tissue-shared effects based on strict thresholds will underestimate the true sharing between tissues, particularly in blood which had reduced power to detect associations due to smaller sample size. Enrichment analysis, which can detect evidence of sharing which does not attain statistical significance by comparing the P value distributions across tissues, revealed shared age-related effects ranging from 21% to 60% (Figure 4), with skin and blood showing the least overlap while fat and skin showed the most. This is considerably greater than the degree of enrichment observed in microarrays, 27%-28% between fat and skin (Glass et al., 2013). For links associated with age reflecting changes in splicing, we observed similar levels of shared age-related effects between skin and fat and overall sharing ranging from 16% to 47%. Our results indicate that global biomarkers of aging with effects across multiple tissues are prevalent. It also supports the finding of multi-tissues studies like GTEx pilot studies showing that the inclusion of more tissues types will also probably allow the identification of more biomarkers, as they identified thousands of genes associated in expression with age in at least one of the 43 tissues tested with a smaller sample size (Melé et al., 2015).

Discussion

The association between aging and disease has been extensively demostrated by epidemiological and GWAS studies (Jeck et al., 2012; López-Otín et al.), but the association between specific genes with a disease in the context of the aging process remains elusive. Our analyses have identified thousands of genes affected by aging, some of which may explain the influence of aging in the onset and outcome of diseases. From the genes reported to show age effects in expression, two examples of genes linked to age-related

diseases illustrate the complicated effects of ageing that we observed in their expression. The first example is the *APOE* gene. Activity of *APOE* has been associated with Alzheimer and cardiovascular diseases, and genetic variants within the *TOMM40/APOE/APOC1* locus have been linked to longevity (Beekman et al., 2013). Our analysis showed that the expression of multiple exons and links of *APOE* change with age in skin tissue, producing different isoforms that can potentially induce changes in the activity of the gene (Figure 5). Furthermore, we previously reported an eQTL affecting the expression of *APOE* in skin and fat tissues (Buil et al., 2015). The skin eQTL (*rs439401*) has been implicated in triglycerides metabolism, Alzheimer and cardiovascular diseases (Chasman et al., 2009). Our gxa-eQTL analysis reported a nominally significant *P* value of 0.014. Given the strong association between expression and disease, such an effect could modulate age-related development and progression of the disease. The second example we choose to highlight here involves the *LMNA* gene (Figure 6), which causal of the Hutchinson-Gilford progeria syndrome. This syndrome is characterized by accelerated aging features as a consequence of the accumulation of a truncate progerin isoform of *LMNA.* The progerin transcript increases with age in normal cells (Rodriguez et al., 2009), with its protein known to accumulate in human skin in an age-dependent manner (McClintock et al., 2007). We reported changes in expression of exons (adjusted *P* values < 0.1) and links (adjusted *P* values < 0.05) between exons consistent with the production of different alternative isoforms in an age-dependent manner. Furthermore, we found an eQTL affecting the expression in skin, blood and LCLs tissues, the peak LCL eQTL (*rs915179*) has been previously linked to exceptional longevity in humans (Conneely et al., 2012; Sebastiani et al., 2012). The best gxa-eQTL with a *P* value = 2.01e-03 was nominally significant but did not pass multiple testing correction. These examples illustrate that studying the global effects of the aging process may lead to the identification of gene involved in age-related diseases.

An under-studied aspect of the relationship between aging and disease is the proposed link between aging and diseases as a consequence of a loss in regulatory capacities in aging organisms and manifested in an increase in phenotypic variance with age (Lu et al., 2004; McCarroll et al., 2004). Global changes in variance of gene expression have been reported to increase with age, but our analysis mainly identified individual genes with a decreased variance, contradicting the expectation of a stochastic increase of the phenotypic variance with age due to reduced regulatory capabilities. Three factors may induce changes in phenotypic mean or variance: genetic variation, environmental variation or an interaction between the two. Genes and pathways associated with longevity and age-related changes are often strongly regulated in older organisms with low levels of stochasticity and higher levels of heritability (Brown et al., 2015; McCarroll et al., 2004; Viñuela et al., 2010). Exons affected by age on the mean and variance of expression were highly heritable (Supp. Figure S4, Supp. file 3) suggesting, as previously reported, that age modulates genetic regulation of expression. We attempted to identify genetic and environmental factors involved in the changes of variance with age by testing for GxE interactions. We were able to identify a significant gxa-eQTL in fat tissue acting on the gene *CD82* (*rs10769002*). This gene is associated with tumor progression as it codes for a metastasis suppressor glycoprotein highly correlated with *p53* and its increase in expression has been associated with overall better survival to cancer (Gentles et al., 2015). In our analysis we observed that individuals homozygous for the reference allele increased gene expression with age compared to the alternative allele. Therefore, it is possible that the alternate allele in *rs10769002* may be a

risk factor for some types of cancer in older individuals. Three other examples were identified in skin tissue for genes also previously implicated in cancer and metabolism. In conclusion, we identify changes in phenotypic variance with age that would be explained by GxE and changes in regulation, suggesting that damage accumulation is not the only explanation to the observed change in phenotypic variance with age in many other phenotypes. Moreover, we show that the study of phenotypic variance with age in gene expression may identify new candidate genes relevant for age-related diseases.

We observed a widespread effect of age on methylation which is potentially stronger than the observed effect of age on gene expression, although both age effects were small compared to the relative influence of genetics. Our search for interactions that would explain changes in variance with age identified *IRS1* as a gene which expression changes as a consequence of an age-methylation interaction. The *IRS1* gene has been associated to T2D, an age-related disease and it has also been found to have T2D associated DMRs nearby (Nilsson et al., 2014). Our results suggest that although methylation changes are strong markers for the aging process their influence on expression changes with age may be only relevant for a small percentage of genes. In conclusion, our search to explain results on global changes in phenotypic variance with age indicates that increase or decrease in expression regulation jointly with the accumulation of environmental exposures may often be observed as multiple GxE interactions.

In summary, we have performed a large human transcriptomic study of aging in multiple tissues. We found that the shared effect of aging in humans across four tissues as well as the number of affected genes is larger than previously reported. We also report that the global effect of age in gene expression is small (median variance explained by age is between 2.2% and 5.74%). When compared to the large global effect of genetic factors on gene expression, the low age-related values may explain the difficulties in identifying biomarkers of aging in gene expression, and highlight the need of larger sample sizes that account for genetic variation. On the other hand, we observed a larger global effect of age in methylation levels, with age explaining up to 60% of the variance observed in methylation levels in some regions. However, we observed a low number of associations between expression and methylation, suggesting that the relationship between both phenotypes and age-related changes may be independent for most genes. Moreover, we have shown that age alters gene expression in multiple complex ways, including variance, mRNA maturation and genetic regulation. Many of these affected genes have been linked to age-related diseases, showing the need for future studies into the relationship between age-related changes in gene expression and its regulation, and age-related diseases. This is particularly relevant for genome wide association studies (GWAS) where eQTL are routinely used to identify target genes of genetic variants without accounting for the effects of age.

Experimental Procedures

Study design

The sample collection, and mRNA extraction has been described in detail in (Grundberg et al., 2012). In sort, 856 Caucasian female individuals (336 MZ and 520 DZ twins) from the TwinsUK Adult twin registry (Spector and Williams, 2006) were recruited with a ranged age from 39 to 85 years (mean 59 years). Samples were prepared for sequencing and processed as described in (Brown et al., 2014) and (Buil et al., 2015). The number of monozygotic (MZ), dizygotic (DZ) and unrelated individuals (individuals with no relatives in the dataset) included in the final analysis per tissue are described on Supp. Table S1

Exons and links quantification

The 49-bp sequenced paired-end reads were mapped to the GRCh37 reference genome (The International Human Genome Sequencing Consortium, (2001)) with BWA v0.5.9 (Li and Durbin, 2009). We use genes defined as protein coding in the GENCODE v10 annotation (Harrow et al., 2012), removing genes with more than 10% zero read count in each tissue. For the analysis presented in this paper, only exons from protein coding genes and LincRNAs from verified loci (level 1) and manually annotated (level 2) were investigated. We calculated the relative quantification of splicing events using Altrans (Ongen and Dermitzakis, 2015a). Read counts assigned to links and exons were scaled to 10 million reads.

Supp. Table S3 show the total number of exons and genes sequenced per tissue, as well as the total number of exons, genes used in the analysis here presented.

Genotying and imputation.

Genotyping of the TwinsUK dataset (N = ~6,000) was done with a combination of Illumina arrays as described in (Brown et al., 2014; Buil et al., 2015; Grundberg et al., 2012). Samples were imputed into the 1000 Genomes Phase 1 reference panel (data freeze, 10/11/2010) (2012) using IMPUTE2 (Howie et al., 2009) and filtered (MAF<0.01, IMPUTE info value < 0.8).

Splicing junction quantifications

We calculated the relative quantification of splicing events using Altrans (Ongen and Dermitzakis, 2015b). The method makes use of mate pairs mapped to different exons to count "links" between two exons based on the GENCODE v10 annotation for level 1 and 2 from protein coding genes and lincRNA. Exons that overlap were grouped into "exon groups" to identify unique portions of each exon from an exon group. The unique portions were used to assign reads to an exon. The quantitative metric produced by Altrans is the fraction of one link's coverage over the sum of overages of all the links that the primary exon produced. The values range from 0 to 1, representing the proportion of a give link among all the links produced by the primary exon. The metric is calculated in 5'-to-3' (forward) and 3'-to-5' (reverse) directions to capture splice acceptor and donor effects respectively. Supp. Table S4 show the total number of links identify per tissue, as well as the total number of links per gene detected.

Age effects on mean exon expression and links

Rank normalized reads per exon or links were used to assess the age effect on exon expression mean. A linear mixed model was fitted to examine age effect on gene expression in R (http://www.r-project.org/) with the lmer function in the lme4 package (Bates et al., 2011). Confounding factors in all models included fixed (primer insert size, GC content mean and batch (only for blood samples)) and random effects (primer index, date of sequencing, family relationship and zygosity). The *P* values to asses significance for age effect were calculated from the Chi-square distribution with 1 degree of freedom using likelihood ratio as the test statistic. A set of 100 permutations were used to adjust for multiple testing. Expression values were permuted while maintaining samples from twin pairs together. To correct for the number of exons per genes, which would allow genes with more exons to have more significant associations by chance than genes with fewer exons, we calculated adjusted *P* values in 16 groups, one per group of genes with similar number of exons. The adjusted *P* value values were calculated as the proportion of permuted statistics more significant, divided by 100. *P* values < 0.05 were considered significant. A gene was considered as significantly affected by age in its expression if at least one exon was significantly associated with it.

Tissue shared effects

For each pair of tissues comparison we extracted *P*values of exons in one tissue (e.g. skin) from significantly age associated exons in other tissue (e.g. fat). The *P*values distributions were used to assess the enrichment of age associated exons in other tissues. Analysis were performed in largeQvalue (Brown, 2014), an implementation of the R statistical software qvalue package (Dabney and Storey), for large datasets.

Number of genes expressed with age

Raw FPKM read counts were used to identify the number of genes expressed per individuals. A gene was considered expressed with FPKM read counts > 0.2. The numbers of expressed genes were rank normalized and used to assess the age effect on number of genes expressed. A linear mixed model with number of genes expressed per samples as response variable was used to assess the association between number of exons expressed and age. Confounding factors in all models included the same fixed and random effects as used before. *P* values were calculated from the Chi-square distribution as before.

Age effect on variance of gene expression

Residuals removing from technical covariates and family structure were used to assess the association for variance and age per tissue. Residuals were extracted from a linear mixed model fitted with the lmer function in the lme4 package (Bates et al., 2011) using R. Confounding factors in all models included fixed and random effects as detailed above. The residuals were fit on a loess function including age as response variable. Residuals from the loess regression were squared root to give a measure of the distance from the mean expression with age. A Spearman correlation test between this 'distance' and the age was used to asses evidences for an age effect on variance. Multiple testing corrections were performed as described for the expression association with age with 100 permutations.

Age effect on discordance of gene expression

Residuals removing only technical covariates were used to assess the change in discordance of gene expression with age per tissue from complete MZ pairs of twins (Supp. Table S1). Association with age was assessed by regressing the maximum expression of each twin pair on the expression of the sibling plus age to detect whether the relationship between maximum and minimum expression was conditional on age. Multiple testing was assess using 100 permutations and as described for the expression association.

Fat methylation analysis

Infinium HumanMethylation450 BeadChip (Illumina Inc, San Diego, CA) was used to measure DNA methylation. Details of experimental approaches have been previously described (Grundberg et al.). To correct the technical issues caused by the two Illumina probe types, the beta mixture quantile dilation (BMIQ) method was performed (Teschendorff et al., 2013). The methylation data is also background corrected. DNA methylation probes that mapped incorrectly or to multiple locations in the reference sequence were removed. Probes with >1% subjects with detection P-value > 0.05 were also removed. Subjects with more than 5% missing probes were also removed. All probes with non-missing values were included.

Differential methylation with age was investigated for probes around the 50,000 bp from the TSS of genes included in the age analysis, which give us a total of 370,731 probes tested from a total of 541,369 CpGs probes on the 450K array. A linear mixed model was fitted to examine age effect on gene expression as in previous analysis. Confounding factors in the models included fixed (beadchip, BS conversion efficiency and BS-treated DNA input) and random effects (family relationship and zygosity). Multiple testing was assessed using 100 permutations. Methylation expression association was tested using expression residuals after removing technical covariates and family structure using a linear model in R with. 100 permutations were used to correct for multiple testing.

Effect sizes and heritability analysis.

We calculated effect size of age in expression and methylation from the normalized data and as a proportion of variance attributed to age over the total variance in exon expression. We also calculated the variance attributed to additive genetic effects, common environment and unique environment. Variance components were calculated from a linear mixed model, as previously described in (Grundberg et al., 2012), and (Visscher et al., 2004) using all available complete twin pairs per tissue (Supp. Table S1). The model was fitted as described above.

Genotype-by-age and methylation-by-age interactions

Expression residuals removing from technical covariates and family structure were used to assess the association of exons and genetics variance interacting with age. To identify genotype-by-age interactions affecting gene expression we performed a linear regression of the residuals of each exon on the SNPs in a 1Mb window around the transcription start site for each gene, using a linear model in R. Only SNPs with MAF >= 0.05 were tested. We used 10 permutations to assess the significance of the interactions for exons with age-related effects, namely mean expression changes, variance changes and discordant effects. We

used a similar strategy as used by (Gerrits et al., 2009) and based on (Anderson and Braak, 2003). A linear model with main effects but without an interaction term was used to extract residuals for each exon-SNP association test. The residuals were permuted (10 times) and used in a linear association with a model for the interacting term (gxa). *P* values from this analysis were stored and used to adjusted *P* values correcting for the number of exons per genes, as described before.

Methylation-by-age interaction analysis used expression and methylation residuals after removal of technical covariates and accounting for family structure. A linear model was used to test the association between expression and methylation levels with age. Significant associations were considered those with a P value < 1.0e-4 (Bonferroni correction).

Code

Supp. File 15 contains code use for the analysis presented in this manuscript.

Ethics Statement

This project was approved by the ethics committee at St Thomas' Hospital London, where all the biopsies were carried out. Volunteers gave informed consent and signed an approved consent form prior to the biopsy procedure. Volunteers were supplied with an appropriate detailed information sheet regarding the research project and biopsy procedure by post prior to attending for the biopsy.

Data availability

RNAseq data was deposited in the European Genome-phenome Archive (EGA) under the accession EGAS00001000805. Methylation data was downloaded from ArrayExpress, accession number E-MTAB-1866.

Contributions

AV, AAB, AB and MND analyzed genotype and expression data as well as developed the methodology. AV, PCT and JTB analyzed methylation data. AV drafted the manuscript which contributions from all authors. AV, ETD, TDS and KSS designed the study. All authors read and approved the final manuscript.

Acknowledgments

The TwinsUK study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR)- funded BioResource, Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. SNP genotyping was performed by The Wellcome Trust Sanger Institute and National Eye Institute via NIH/CIDR. T.S. is an NIHR senior Investigator and is holder of an ERC Advanced Principal Investigator award. This work is also in part supported by the ERC (250157) and ESRC (ES/N000404/1).

AV, AAB, AB and MD were supported by the EU FP7 grant EuroBATS (No. 259749). AAB is also supported by a grant from the South-Eastern Norway Health Authority, (No. 2011060). KSS is supported by MRC grant MR/L01999X/1.

Some computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics.

List of Supplemental files:

Supplemental file 1: Summary statistics for exons association with age. Each sheet contains the pvalues for one tissue.

Supplemental files 2: Output from the functional analysis of differentially expressed genes with age using David. Only significant results (FDR < 0.05) are included. Each sheet contains the pvalues for one tissue.

Supplemental files 3: Results for variance decomposition per exon (variance explained by age and heritability), in fat, skin, whole blood and LCLs. Each sheet contains the pvalues for one tissue.

Supplemental files 4 and 5: Summary statistics for links (reads between exons) association with age. Tab delimited files, one per tissue and direction (reverse, forward): fat and skin. Each sheet contains the pvalues for one direcction.

Supplemental files 6: Output from the functional analysis of differentially expressed links with age using David. Only significant results (FDR < 0.05) are included. Each sheet contains the pvalues for one tissue (skin and fat).

Supplemental files 7: Summary statistics for variance of expression association with age. Each sheet contains the pvalues for one tissue: fat, skin, whole blood and LCLs.

Supplemental files 8: Summary statistics for discordance of expression between MZ twins and its association with age. Each sheet contains the pvalues for one tissue: fat, skin, whole blood and LCLs.

Supplemental files 9: Summary statistic for gxa-eQTL in differentially expressed exons in fat and skin. Only the best SNP-exon association per exon in included.

Supplemental files 10: Summary statistic for gxa-eQTL in exons with a significant change in variance with age in fat and skin. Only the best SNP-exon association per exon in included.

Supplemental files 11: Summary statistic for gxa-eQTL in exons discordant for expression between MZ twins in fat and skin. Only the best SNP-exon association per exon in included.

Supplemental files 12: Summary statistics for methylation association with age in fat tissue.

Supplemental files 13: Summary statistics for methylation association with expression.

Supplemental files 14: Summary statistics for methylation interaction with age.

Supplemental files 15: R code to perform the analysis discussed in the paper.

References

(2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

(2012). An integrated map of genetic variation from 1,092 human genomes. Nature *491*, 56-65.

Anderson, M., and Braak, C.T. (2003). Permutation tests for multi-factorial analysis of variance. Journal of Statistical Computation and Simulation *73*, 85-113.

Bahar, R., Hartmann, C.H., Rodriguez, K.A., Denny, A.D., Busuttil, R.A., Dolle, M.E.T., Calder, R.B., Chisholm, G.B., Pollock, B.H., Klein, C.A., et al. (2006). Increased cell-to-cell variation in gene expression in ageing mouse heart. Nature *441*, 1011-1014.

Bates, D., Maechler, M., and Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-41.

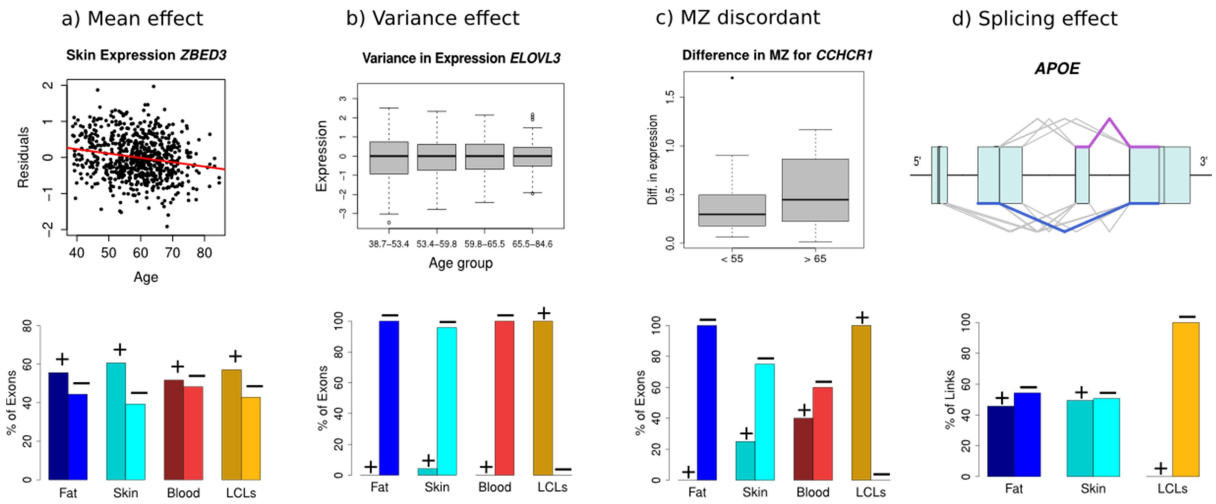
Beekman, M., Blanché, H., Perola, M., Hervonen, A., Bezrukov, V., Sikora, E., Flachsbart, F., Christiansen, L., De Craen, A.J.M., Kirkwood, T.B.L., et al. (2013). Genome-wide linkage analysis for human longevity: Genetics of Healthy Aging Study. Aging Cell *12*, 184-193.

Bell, J.T., Tsai, P.-C., Yang, T.-P., Pidsley, R., Nisbet, J., Glass, D., Mangino, M., Zhai, G., Zhang, F., Valdes, A., et al. (2012). Epigenome-Wide Scans Identify Differentially Methylated Regions for Age and Age-Related Phenotypes in a Healthy Ageing Population. PLoS Genet *8*, e1002629.
Brown, A., Ding, Z., Viñuela, A., Glass, D., Parts, L., Spector, T., Winn, J., and Durbin, R. (2015). Pathway Based Factor Analysis of Gene Expression Data Produces Highly Heritable Phenotypes that Associate with Age. G3: Genes|Genomes|Genetics.
Brown, A.A. (2014). largeQvalue: A program for calculating FDR estimates with large datasets.
Brown, A.A., Buil, A., Viñuela, A., Lappalainen, T., Zheng, H.-F., Richards, J.B., Small, K.S., Spector, T.D., Dermitzakis, E.T., and Durbin, R. (2014). Genetic interactions affecting human gene expression identified by variance association mapping.
Buil, A., Brown, A.A., Lappalainen, T., Vinuela, A., Davies, M.N., Zheng, H.-F., Richards, J.B., Glass, D., Small, K.S., Durbin, R., et al. (2015). Gene-gene and gene-environment interactions detected by transcriptome sequence analysis in twins. Nat Genet *47*, 88-91.
Chasman, D.I., Paré, G., Mora, S., Hopewell, J.C., Peloso, G., Clarke, R., Cupples, L.A., Hamsten, A., Kathiresan, S., Mälarstig, A., et al. (2009). Forty-Three Loci Associated with Plasma Lipoprotein Size, Concentration, and Cholesterol Content in Genome-Wide Analysis. PLoS Genet *5*, e1000730.
Conneely, K.N., Capell, B.C., Erdos, M.R., Sebastiani, P., Solovieff, N., Swift, A.J., Baldwin, C.T., Budagov, T., Barzilai, N., Atzmon, G., et al. (2012). Human longevity and common variations in the LMNA gene: a meta-analysis. Aging Cell *11*, 475-481.
Dabney, A., and Storey, J.D. qvalue: Q-value estimation for false discovery rate control.. R package version 1.40.0.
de Magalhaes, J.P., Curado, J., and Church, G.M. (2009). Meta-analysis of age-related gene expression profiles identifies common signatures of aging. Bioinformatics *25*, 875-881.
Gentles, A.J., Newman, A.M., Liu, C.L., Bratman, S.V., Feng, W., Kim, D., Nair, V.S., Xu, Y., Khuong, A., Hoang, C.D., et al. (2015). The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med *21*, 938-945.
Gerrits, A., Li, Y., Tesson, B.M., Bystrykh, L.V., Weersing, E., Ausema, A., Dontje, B., Wang, X., Breitling, R., Jansen, R.C., et al. (2009). Expression Quantitative Trait Loci Are Highly Sensitive to Cellular Differentiation State. PLoS Genet *5*, e1000692.
Glass, D., Vinuela, A., Davies, M., Ramasamy, A., Parts, L., Knowles, D., Brown, A., Hedman, A., Small, K., Buil, A., et al. (2013). Gene expression changes with age in skin, adipose tissue, blood and brain. Genome Biology *14*, R75.
Grundberg, E., Meduri, E., Sandling, Johanna K., Hedman, Åsa K., Keildson, S., Buil, A., Busche, S., Yuan, W., Nisbet, J., Sekowska, M., et al. Global Analysis of DNA Methylation Variation in Adipose Tissue from Twins Reveals Links to Disease-Associated Variants in Distal Regulatory Elements. The American Journal of Human Genetics *93*, 876-890.
Grundberg, E., Small, K.S., Hedman, A.K., Nica, A.C., Buil, A., Keildson, S., Bell, J.T., Yang, T.-P., Meduri, E., Barrett, A., et al. (2012). Mapping cis- and trans-regulatory effects across multiple tissues in twins. Nat Genet *44*, 1084-1089.
Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J.-B., Gao, Y., et al. (2013). Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. Molecular cell *49*, 359-367.
Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: The reference human genome annotation for The ENCODE Project. Genome Research *22*, 1760-1774.
Howie, B.N., Donnelly, P., and Marchini, J. (2009). A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. PLoS Genet *5*, e1000529.
http://www.r-project.org/ The R Project for Statistical Computing.
Jeck, W.R., Siebold, A.P., and Sharpless, N.E. (2012). Review: a meta-analysis of GWAS and age-associated diseases. Aging Cell *11*, 727-731.

Kent, J.W., Göring, H.H.H., Charlesworth, J.C., Drigalenko, E., Diego, V.P., Curran, J.E., Johnson, M.P., Dyer, T.D., Cole, S.A., Jowett, J.B.M., et al. (2012). Genotype × age interaction in human transcriptional ageing. Mechanisms of ageing and development *133*, 581-590.
Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics *25*, 1754-1760.
Li, Z., Wright, F.A., and Royland, J. (2009). Age-Dependent Variability in Gene Expression in Male Fischer 344 Rat Retina. Toxicological Sciences *107*, 281-292.
López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. The Hallmarks of Aging. Cell *153*, 1194-1217.
Lu, T., Pan, Y., Kao, S.-Y., Li, C., Kohane, I., Chan, J., and Yankner, B.A. (2004). Gene regulation and DNA damage in the ageing human brain. Nature *429*, 883-891.
McCarroll, S.A., Murphy, C.T., Zou, S., Pletcher, S.D., Chin, C.S., Jan, Y.N., Kenyon, C., Bargmann, C.I., and Li, H. (2004). Comparing genomic expression patterns across species identifies shared transcriptional profile in aging. Nat Genet *36*, 197 - 204.
McClintock, D., Ratner, D., Lokuge, M., Owens, D.M., Gordon, L.B., Collins, F.S., and Djabali, K. (2007). The Mutant Form of Lamin A that Causes Hutchinson-Gilford Progeria Is a Biomarker of Cellular Aging in Human Skin. PLoS ONE *2*, e1269.
Melé, M., Ferreira, P.G., Reverter, F., DeLuca, D.S., Monlong, J., Sammeth, M., Young, T.R., Goldmann, J.M., Pervouchine, D.D., Sullivan, T.J., et al. (2015). The human transcriptome across tissues and individuals. Science *348*, 660-665.
Nilsson, E., Jansson, P.A., Perfilyev, A., Volkov, P., Pedersen, M., Svensson, M.K., Poulsen, P., Ribel-Madsen, R., Pedersen, N.L., Almgren, P., et al. (2014). Altered DNA Methylation and Differential Expression of Genes Influencing Metabolism and Inflammation in Adipose Tissue From Subjects With Type 2 Diabetes. Diabetes *63*, 2962-2976.
Ongen, H., and Dermitzakis, Emmanouil T. (2015a). Alternative Splicing QTLs in European and African Populations. The American Journal of Human Genetics *97*, 567-575.
Ongen, H., and Dermitzakis, E.T. (2015b). Alternative splicing QTLs in European and African populations using Altrans, a novel method for splice junction quantification.
Paré, G., Cook, N.R., Ridker, P.M., and Chasman, D.I. (2010). On the Use of Variance per Genotype as a Tool to Identify Quantitative Trait Interaction Effects: A Report from the Women's Genome Health Study. PLoS Genet *6*, e1000981.
Peters, M.J., Joehanes, R., Pilling, L.C., Schurmann, C., Conneely, K.N., Powell, J., Reinmaa, E., Sutphin, G.L., Zhernakova, A., Schramm, K., et al. (2015). The transcriptional landscape of age in human peripheral blood. Nat Commun *6*.
Rodriguez, S., Coppede, F., Sagelius, H., and Eriksson, M. (2009). Increased expression of the Hutchinson-Gilford progeria syndrome truncated lamin A transcript during cell aging. Eur J Hum Genet *17*, 928-937.
Rodwell, G.E.J., Sonu, R., Zahn, J.M., Lund, J., Wilhelmy, J., Wang, L., Xiao, W., Mindrinos, M., Crane, E., Segal, E., et al. (2004). A Transcriptional Profile of Aging in the Human Kidney. PLoS Biol *2*, e427.
Sebastiani, P., Solovieff, N., DeWan, A.T., Walsh, K.M., Puca, A., Hartley, S.W., Melista, E., Andersen, S., Dworkis, D.A., Wilk, J.B., et al. (2012). Genetic Signatures of Exceptional Longevity in Humans. PLoS ONE *7*, e29848.
Somel, M., Khaitovich, P., Bahn, S., Pääbo, S., and Lachmann, M. (2006). Gene expression becomes heterogeneous with age. Current Biology *16*, R359-R360.
Spector, T.D., and Williams, F.M.K. (2006). The UK Adult Twin Registry (TwinsUK). Twin Research and Human Genetics *9*, 899-906.
Teschendorff, A.E., Marabita, F., Lechner, M., Bartlett, T., Tegner, J., Gomez-Cabrero, D., and Beck, S. (2013). A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics *29*, 189-196.
Viñuela, A., Snoek, L.B., Riksen, J.A.G., and Kammenga, J.E. (2010). Genome-wide gene expression regulation as a function of genotype and age in C. elegans. Genome Research *20*, 929-937.

Visscher, P.M., Benyamin, B., and White, I. (2004). The Use of Linear Mixed Models to Estimate Variance Components from Data on Twin Pairs by Maximum Likelihood. Twin Research (2000) *7*, 670-674.
Wheeler, H.E., and Kim, S.K. (2011). Genetics and genomics of human ageing. Philosophical Transactions of the Royal Society B: Biological Sciences *366*, 43-50.
Yang, J., Huang, T., Petralia, F., Long, Q., Zhang, B., Argmann, C., Zhao, Y., Mobbs, C.V., Schadt, E.E., Zhu, J., et al. (2015). Synchronized age-related gene expression changes across multiple tissues in human and the link to complex diseases. Scientific Reports *5*, 15145.
Yao, C., Joehanes, R., Johnson, A.D., Huan, T., Esko, T., Ying, S., Freedman, J.E., Murabito, J., Lunetta, K.L., Metspalu, A., et al. (2014). Sex- and age-interacting eQTLs in human complex diseases. Human Molecular Genetics *23*, 1947-1956.

Figures & Tables



	Fat	Skin	Blood	LCLs
Exons	1,511	11,695	688	7
Genes	905	4,307	480	7

	Fat	Skin	Blood	LCLs
Exons	3	250	21	3
Genes	3	112	19	3

	Fat	Skin	Blood	LCLs
Exons	1	40	5	2
Genes	1	34	5	2

	Fat	Skin	Blood	LCLs
Exons	928	1,696	688	1
Genes	310	679	451	1

Figure 1 | Effects of aging in gene expression: The effect of aging in gene expression is not limited to changes in mean expression values with age (a), but includes also changes in levels of phenotypic variance (b,c), and splicing (d). The top row graphs show real data examples for the effects of aging in expression investigated. The middle graphs show bar plots with the percentage of exons with positive (+) or negative (-) age effects in each analysis. And finally the bottom tables provide the number of exons and genes with significant association for each of the effects presented. All the real examples are from skin, the tissue with larger age effect in expression overall analyses. a) Effect of aging in mean gene expression, usually referred as differentially expression with age in exons. The example shows the residuals (after removing technical covariates) of the expression of the *ZBED3* gene decreasing with age. Skin is the tissues with a larger effect of age in expression and LCLs the smaller. b) The effect of aging in variance of gene expression is shown with the *ELOVL3* gene and a significant decrease of variance in expression with age. From the bar plot it is possible to appreciate that the majority of the significant exons had a decrease in variance with age. c) Differences in expression between monozygous (MZ) twins point out to environmental factors different among the siblings affecting gene expression, since MZ twins are genetically identical individuals with the same age. The example shows the difference in expression between MZ twins in the gene *CCHR1*. d) For the splicing analysis, only links (reads between two exons) were considered. The example shows the structure of the gene *APOE* with its exons (boxes) and lines connecting the exons representing reads spanning between two exons. The number of reads linking exons 3 and 4 (in purple) decreased in number with age, while reads linking exons 2 and 4 (blue) increased with age (Figure 5 for details). The model suggested that an isoform skipping the third exon (from the 5') may be more abundant in older individuals compare to an isoform that includes the third exon linked to the last exon.

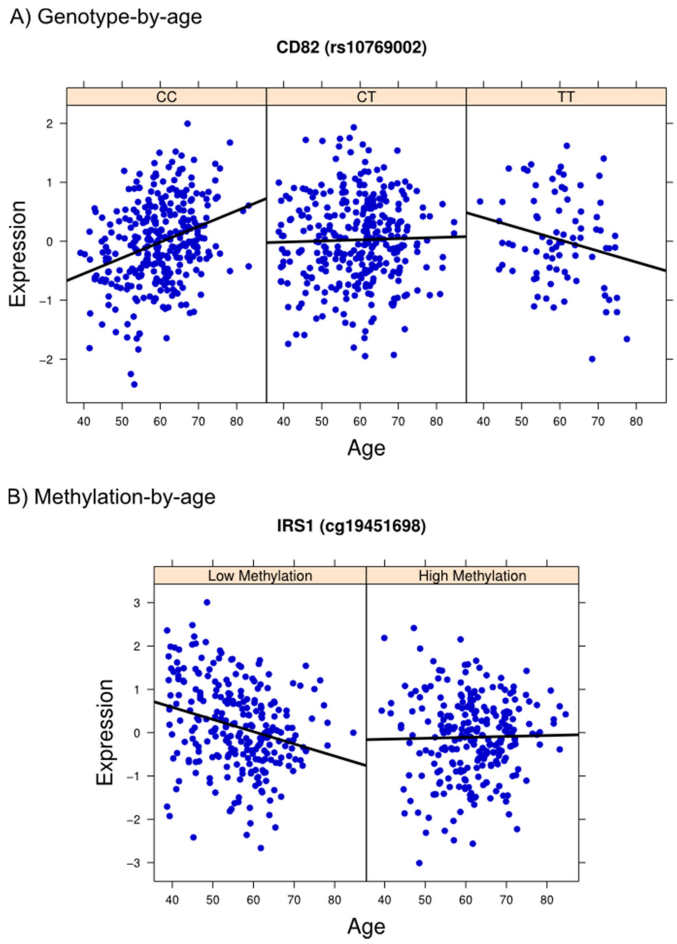


Figure 2 | Interacting effects of aging on gene expression: The two plots show the effects of genotypes (eQTL) and methylation on gene expression can be modulated by age. A) The graph shows a genotype-by-age expression quantitative trait locus (gxa-eQTL) in fat tissue affecting the expression of the *CD82* gene. The expression of the reported exon increased with age in homozygous individuals for the CC alleles in rs10769002. Homozygous individuals for the alternative allele (TT) showed a decreased in expression with age. B) The graph shows a methylation-by-age interaction affecting gene expression. The expression of the *IRS1* gene decreased with age in individuals with the methylation cg19451698 hypomethylated.

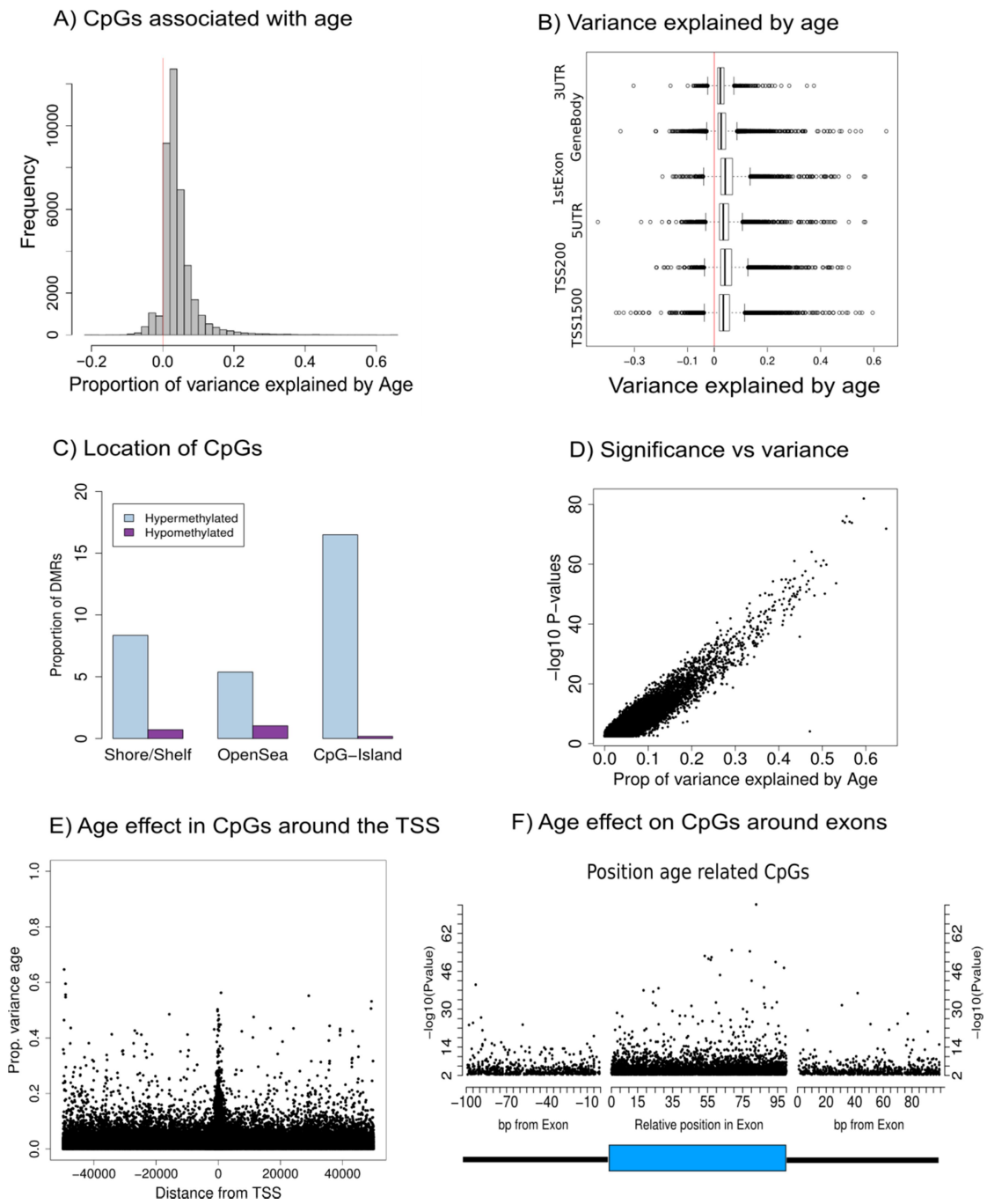


Figure 3 | A) CpG islands showed mainly an increased in methylation with age, independent form the genomic position (B and C). D) The estimates for the proportion of variance attributed to age in methylation show that up to 60% of the variance in methylation would be attributed to age. E) The position of CpG markers respect to the near TSS from genes (only CpGs at <50Mb from the TSS are shown) show a larger effect of age on methylated regions near the TSS. F) The left and right panels show age-associated CpGs positions at the near 0-100bp at the 3'and 5'of each exon. The central panel show the relative position of the CpGs associated with age within each exon (blue box). The CpGs show higher associated with age in the exon 5'region, probably due to the proximity to the TSS of the genes.

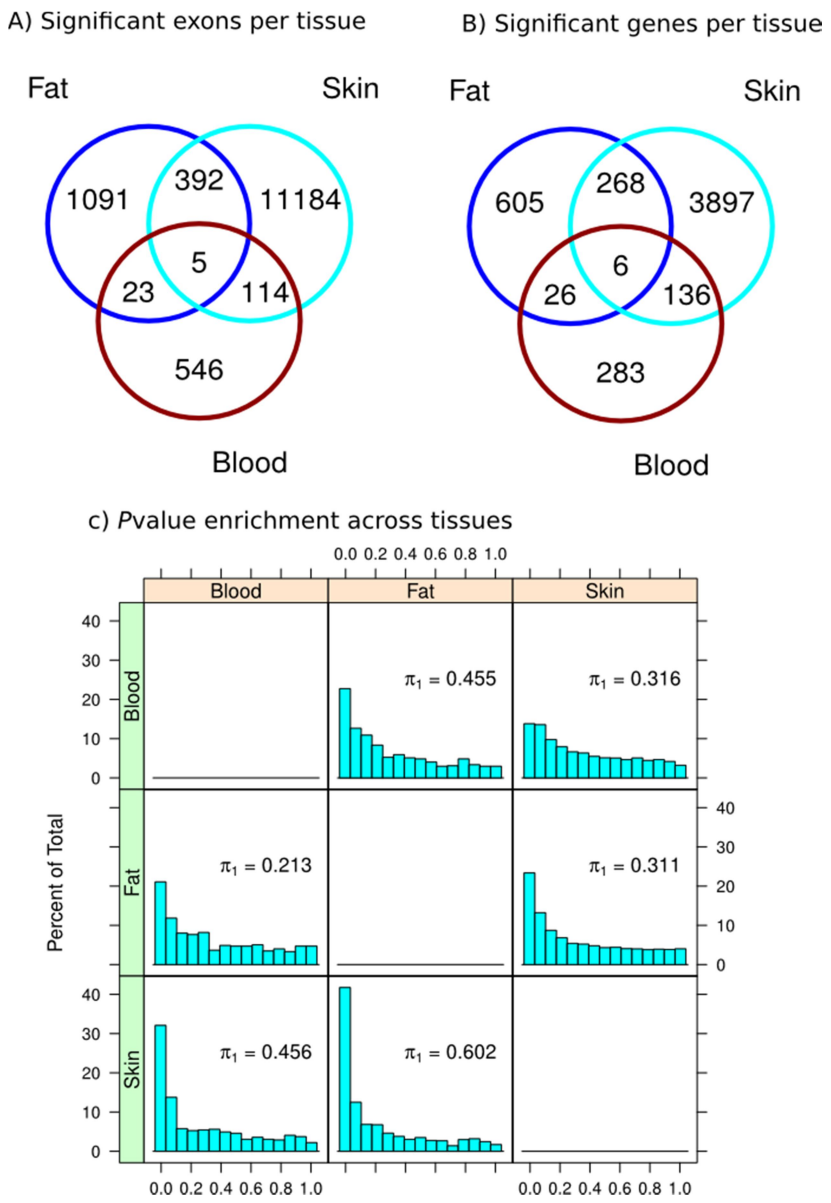


Figure 4 | Tissue shared and specific effects of aging in gene expression changes with age. The top venn diagrams show A) the number of exons (left) and B) number of genes (right) significantly associated with chronological age in fat, skin and whole blood. Five exons were commonly associated to age in the three tissues. LCLs were not included, as only 7 exons were significantly associated with age. C) The *P* values of significant exons associated with age in one tissue were extracted from the analysis in the other tissues for enrichment analysis (π1). The histograms show the *P* values for association between expression and age in one tissue (left, green color) if the exons were significantly associated exons in another tissue (top, orange color). As show in the graphs, age-related signals detected in fat shared an estimated 60.2% of the age effect signal skin tissue and 45.6% with blood.

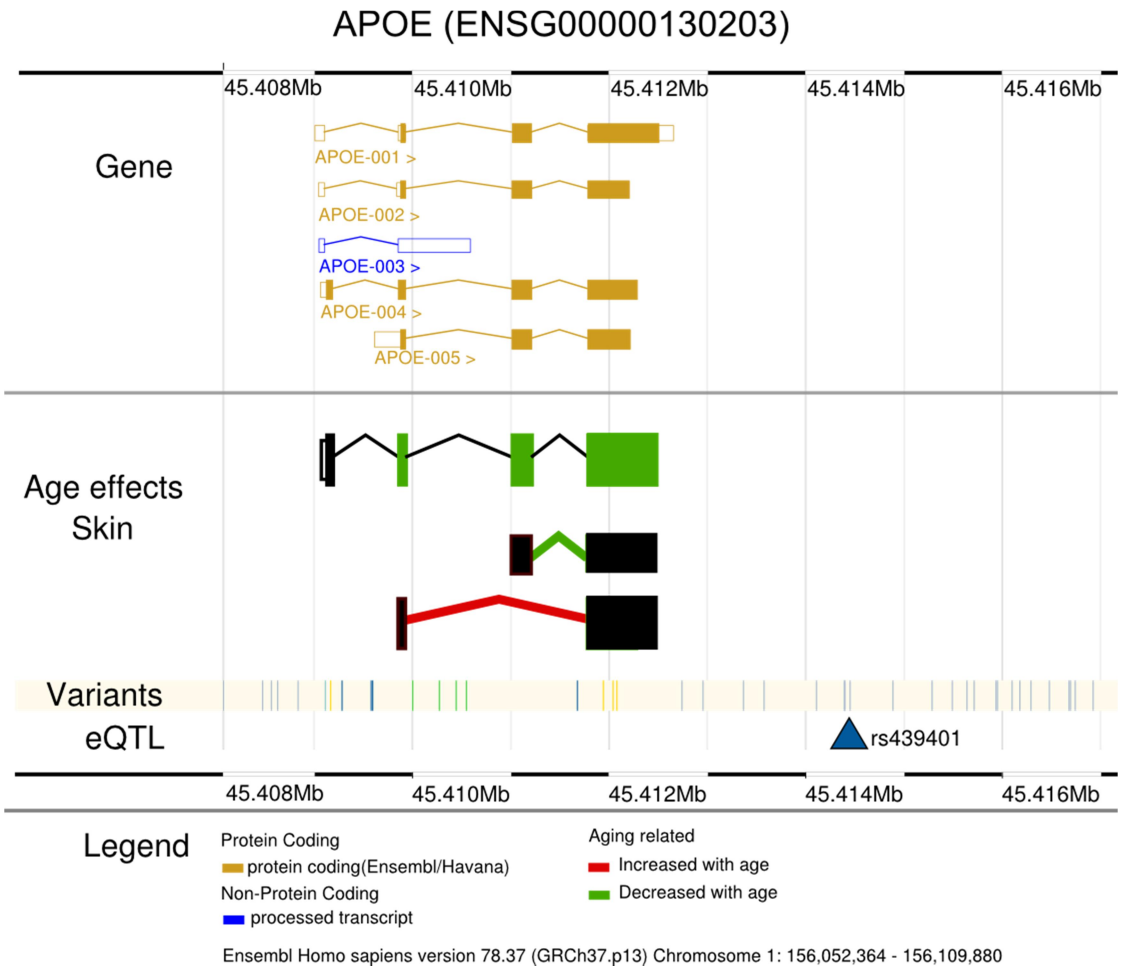


Figure 5 | Structure of the APOE gene. Mutations in this gene have been associated with alterations in fatty acid metabolism and cardiovascular diseases. Polymorphisms in and near the gene has been associated with Alzheimer and cardiovascular diseases. The gene produces multiple protein coding transcripts variants (yellow) and non-coding processed transcripts (blue). In the skin tissue, three exons and one link decreased their expression with age (green coloured exon and link between exons); and one link increase its expression (red coloured link). Furthermore, we detected one eQTL (rs439401) affecting the expression of the gene.

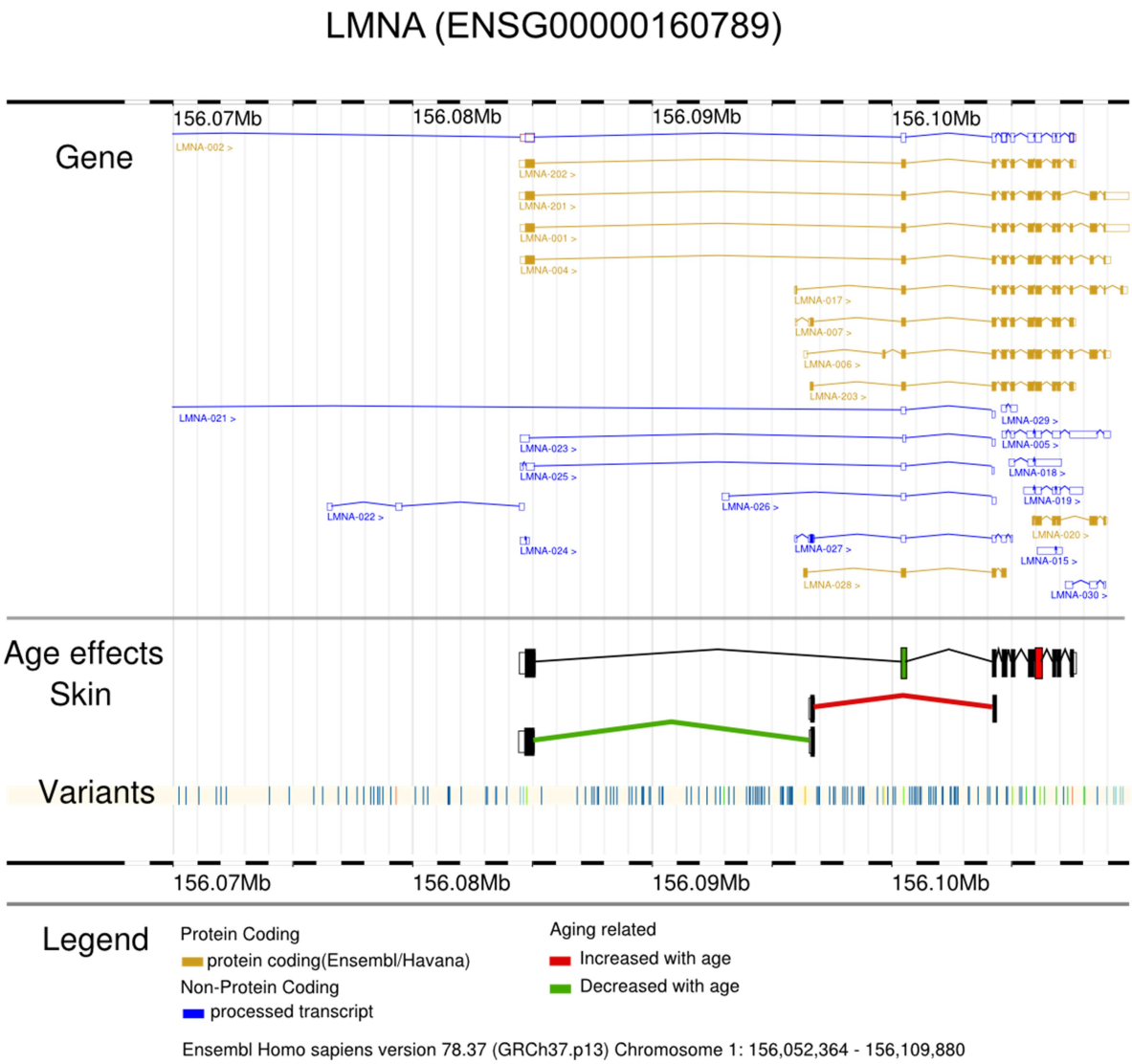


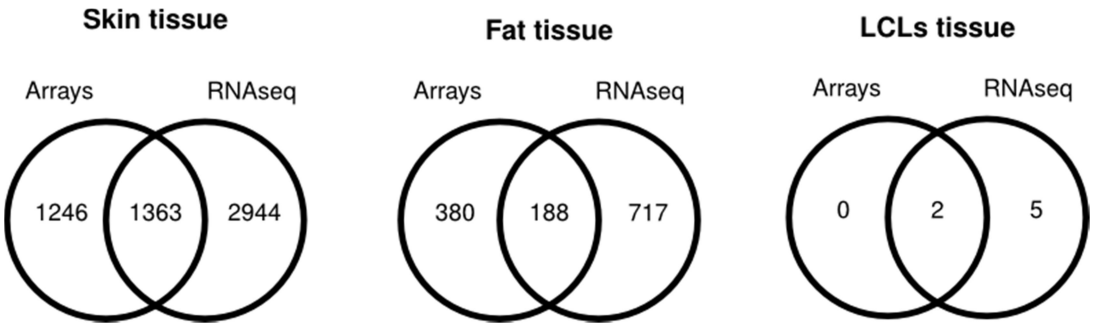
Figure 6 | Structure of the LMNA gene. Mutations in this gene has been associated with multiple diseases, including the Hutchinson-Gilford progeria syndrome, characterize by accelerated ageing features. The gene produces multiple protein coding transcripts (yellow) and non-coding processed transcripts (blue). In the skin tissue, two exons were affected in their expression by age by increasing expression (red coloured exon, corrected Pval < 0.1) and decrease expression with age (green coloured exon). Furthermore, two links were significantly associated with age in their expression (corrected Pval < 0.05). Our results suggested an increase in the production of isoforms using alternative 5'

Tables:

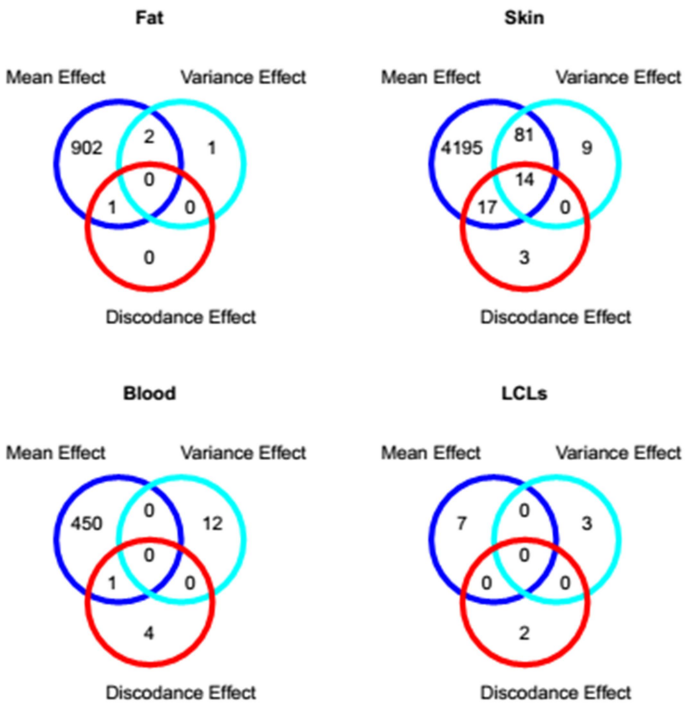
	Fat		Skin		Blood		LCLs	
	All	Age	All	Age	All	Age	All	Age
Age	0.0012	0.0287	0.0026	0.0224	0.0039	0.0542	0.0006	0.0366
Genetics	0.0809	0.2228	0.0856	0.1275	0.1301	0.2333	0.1089	0.2032
Commo Env.	0.0236	0.0573	0.0000	0.0162	0.0380	0.0341	0.0971	0.1312
Unique Env.	0.8550	0.6655	0.8566	0.7662	0.7403	0.6214	0.7320	0.6214
N. of exons	101,133	1,511	96,736	11,695	71,393	688	98,372	7

Table 1: Summary of median proportion of variance attributed to age, genetics, common environment and unique environment, for all exons (All) and for age-affected exons (Age). In general, age explained a small proportion of the variance attributed to gene expression. However, for exons affected by age in their expression, the genetic component (heritability) explained significantly higher proportion of the variance in expression compare to the rest of the genes in fat, skin and blood tissues (willconox test Pvalue < 2.1e-17).

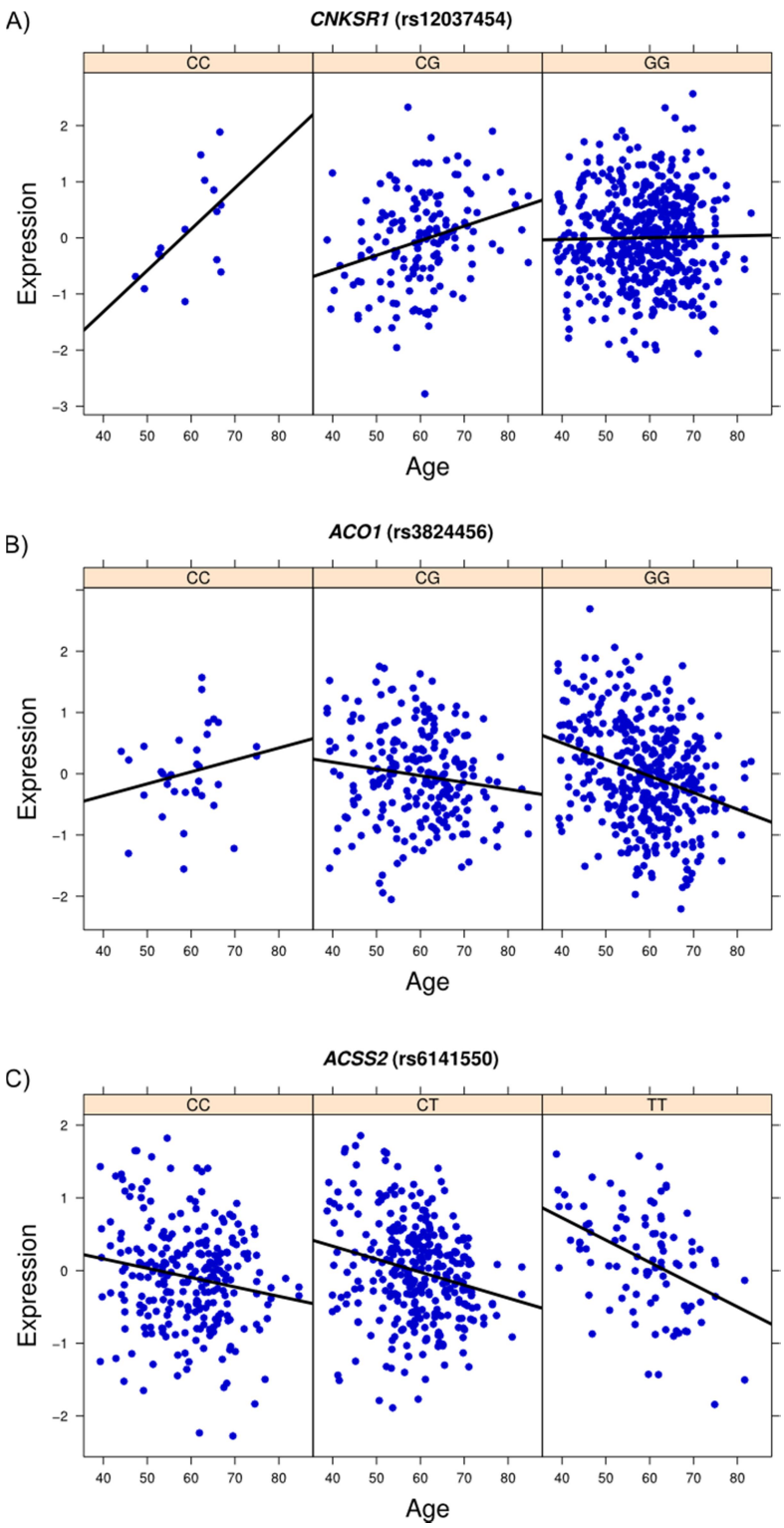
Supplemental Figures



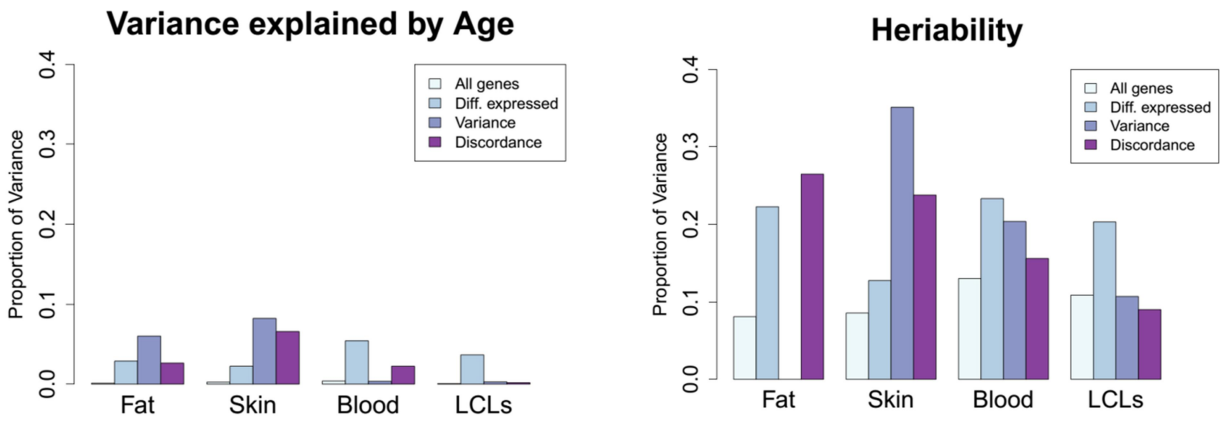
Supp. Figure S1 | Venn Diagrams per tissue comparing differentially expressed genes in array expriments (Glass et al, 2012) and gene swith at least an exon differentially expressed with age from RNA-seq data.



Supp. Figure S2 | Venn diagram showing the overlap in genes with at least one exon affected by changes in the mean (DE genes), variance and differences between MZ twins (discordance) in each tissue.



Supp. Figure S3 | Interacting effects of aging on gene expression: All the graphs show a genotype-by-age expression quantitative trait locus (gxa-eQTL) in skin tissue affecting the expression of three genes.



Supp. Figure S4 | Median proportion of variance explained by age (left) and genetics (right) in all tested genes (All genes), differentially expressed genes with age (diff. expressed), genes changing variance with age (Variance) and genes discordant in MZ twins with age (Discordant). In general, the amount of variance explained by age and heritability in genes significantly affected by age in different ways is larger than in the median of the whole genome. The exception applies to those groups of genes with very little number of genes, like discordance genes in fat with 1 gene. The complete variance decomposition analysis is shown in table S3.

Tables

Supp. Table S1 | Number of monozyigous (MZ), dizygous (DZ) and unrelated individuals (individuals with no relatives in the dataset) included in the final analysis per tissue are described on the following.

	Fat	Skin	LCLs	Whole Blood
Samples	766	716	814	384
MZ pairs	131	114	137	69
DZ pairs	187	173	217	91
Unrelated	130	142	106	64

	Fat				Skin				Blood				LCLs			
	All	**DE**	**Variance**	**Discordance**	**All**	**Age**	**Variance**	**Discordance**	**All**	**Age**	**Variance**	**Discordance**	**All**	**Age**	**Variance**	**Discordance**
Age	0.0012	0.0287	0.0601	0.0263	0.0026	0.0224	0.0822	0.0658	0.0039	0.0542	0.0035	0.0225	0.0006	0.0366	0.0030	0.0018
Genetics	0.0809	0.2228	8.9e-14	0.2649	0.0856	0.1275	0.3513	0.2380	0.1301	0.2333	0.2037	0.1560	0.1089	0.2032	0.1070	0.0901
Commo Env.	0.0236	0.0573	0.172	-0.1066	0.0000	0.0162	-0.1017	-0.066	0.0380	0.0341	0.0716	0.0293	0.0971	0.1312	0.1966	0.0599
Unique Env.	0.8550	0.6655	0.690	0.8094	0.8566	0.7662	0.6778	0.7512	0.7403	0.6214	0.6055	0.7179	0.7320	0.6214	0.6224	0.7644
N. of exons	101,133	1,511	3	1	96,736	11,695	239	40	71,393	688	13	5	98,372	7	3	2

Supp. Table S3 | Summary of mean proportion of variance attributed to age, genetics, common environment and unique environment, for exons affected by age in their variance (Variance) and exons discordant for expression with age (Discordant). The last row indicates the number of exons significant for each category (corrected *P*value < 0.05).