Age and Race related changes in Gene Expression during Human PBMC aging

Yang Hu[1,2]， Yudai Xu[1], Wen Lei[1], Jian Xiang[1], Lipeng Mao[1], Guobing Chen[1]*

1, Institute of Geriatric Immunology, School of Medicine, Jinan University, Guangzhou, China

2, Department of Neurology, the First Affiliated Hospital, Jinan University, Guangzhou, China

*Correspondence: Dr. Guobing Chen, Institute of Geriatric Immunology, School of Medicine, Jinan University, 601 Huangpu Avenue West, Guangzhou 510632, Guangdong Province, China. E-mails: guobingchen@jnu.edu.cn

Abstract

Human immune system functions over an entire lifetime, yet how and why the immune system becomes less effective with age are not well understood. Here, we characterize peripheral blood mononuclear cells from 172 healthy adults 21~90 years of age using RNA-seq and the weighted gene correlation network analysis (WGCNA). These data have revealed a meaningful of gene expression modules or representative biomarkers for human immune system aging in Aisan and White ancestry. Among them, several gene modules demonstrated a remarkably correlation with human immune aging progress. Besides, further analysis on these ageing related modules show age-related gene expression changes spike is around early-seventies. More importantly, we also focus on how race and ethnicity affect immune aging as race-specific effects on these gene expression changes have clinical applications for diagnosis and interpretation of immunosenescence. Thus, the top hub genes including NUDT7, CLPB, OXNAD1 and MLLT3 are identified from Asian and white aging related modules and further validated in humans PBMC at different ages. Finally, the impact of age and race on immune phenotypes we discuss may provide insights into future studies.

Introduction

Chronological age is a major risk factor for many common diseases including heart disease, cancer and stroke, three of the leading causes of death(He and Sharpless, 2017). Although chronological age is the most powerful risk factor for most chronic diseases, the underlying molecular mechanisms that lead to generalized disease susceptibility are largely unknown. It has been shown that aging alters the components of innate and adaptive immunity ranging from the expression of signaling molecules to the behavior of neutrophils, monocytes, lymphocyte, NK cells, etc(Cheung et al., 2018; Nikolich-Zugich, 2018). Interestingly, after the age of 60, the most consistent and dramatic alterations in adaptive immune system are the T cell compartment, including a decreased thymic output of naive T cells and an increased antigen-experienced memory cells(Moskowitz et al., 2017; Nikolich-Zugich, 2018).

Recent studies revealed that epigenetic variations and gene expression of purified immune cells, especially CD8+ T cells, changed significantly with aging, impacting the activity of important receptor molecules, signaling pathways, and transcription factors (TF)(Peters et al., 2015; Moskowitz et al., 2017). Besides, since cytokines were central to immune cell communication and effector activity, many researchers had investigated the contribution of changes in cytokine production to the age associated changes in immune response. Such as, T-cell proliferation and growth was induced by IL-2 but as T cells age, they lost their capacity to produce and respond to IL-2(Whisler et al., 1996). And the reduced IL-2 production ascribed age-related impairments in the activation of transcription factors AP-1 and NF-AT(Whisler et al., 1996). Thus, understanding the aging-associated immune gene expression changes is critical to explaining both the disease susceptibility and the different clinical course of diseases in the elderly.

Interestingly, analyses of human blood samples from different race and ethnicity uncovered significant aging-related changes in PBMC various population subsets(Noren Hooten et al., 2018). For example, a study on peripheral blood mononuclear cell subsets described not only the presence of benign ethnic neutropenia among African Americans but further described a higher proportion of CD19+ cells and a lower proportion of CD3+ cells than in Whites (Tollerud et al., 1989). Moreover, the proportions of PBMCs' subpopulation in Asian cohorts were also different. Choong and colleagues observed that there were differences in cell counts for T, NK, and CD4+ cells as well as in the CD4/CD8 ratio among healthy Malaysians, Chinese, and Indians across the life span (18–71 years) (Choong et al., 1995). Besides, Indians were significantly different from Malays and Chinese. Indians had higher T cells, higher CD4 cells, higher CD4/CD8 ratio, and lower NK cells(Choong et al., 1995). Chinese donors had lower B-cell levels than Malays and Indians(Choong et al., 1995). Despite the importance of age and race in shaping immune cell numbers and functions, it is not known whether Asian and White immune systems go through similar gene expression changes throughout their lifespan, and to what extent these aging-associated changes are shared between White and Asian.

To study this, we collected and profiled PBMCs of healthy adults by carefully matching the ages of Asian and White donors. Computational pipelines for WGCNA analysis, differentially expressed genes (DEGs) and functional enrichments analyses revealed immune system aging signatures in Asian and White. These findings uncovered in which ways aging differentially affects the immune systems between Asian and White populations and discovered a common genetic variant that greatly impacts normal PBMC aging.

Key words: Immune aging, race and ethnicity, PBMC, WGCNA

Results

Profiling PBMCs of healthy adults. We recruited 19 community dwelling healthy volunteers (10 women, 9 men) whose ages span 21–93 years old (Supplementary Table S1): 9 young (ages 21–30: 4 men, 5 women), and 10 older subjects (74+: 5 men, 5 women) in Guangdong area from China. No significant differences were detected between sexes in their frailty scores or age distributions. Male and female samples for each assay were comparable in terms of age; old women were slightly older than men (~85.2 vs. ~81.4; t-test p-value = 0.41) (Supplementary Table S1). Then, PBMCs were profiled using RNA-seq (10 women, 9 men; Supplementary Table S7). Prior to the construction of hierarchical clusters, 29,367 genes were selected after normalization of raw gene counts, excluding the genes with no expression in all samples. Moreover, a total of 153 normal healthy human subjects, whose ages span 20–90 years old and come from different races (Figure S1, Supplementary Table S2): 65 young (ages 21–40: 24 men, 41 women), 40 middle-aged (ages 41–64: 18 men, 22 women), and 48 older subjects (65+: 20 men, 28 women) were downloaded from The 10,000 Immunomes Project (10KIP, http://10kimmunomes.org/). Then, their PBMCs gene expression array data were processed for further WGCNA clustering analysis.

Aging and race have influenced the transcriptomic changes over human adult lifespan. To identify major sources of variation in transcriptomic data, we conducted WGCNA analysis using expressed genes (n = 19,089) from the 10KIP. First, by using the soft thresholding power (β=6) in this algorithm, the co-expression network could satisfy the approximate scale-free topology criterion with R^2 > 0.80 while maintaining a high mean connectivity with enough information (Figure S2A). Then, we merged the modules of eigengenes with a correlation coefficient of over 0.75 (Figure S2B), and the size of the co-expression modules ranged from 68 to 3,891 genes (Figure1A, Supplementary Table S3). After merging the highly correlated modules, the co-expression genes were clustered into 22 modules with similar gene expression patterns and labeled with different colors (Figure 1B). To further quantify co-expression similarity of entire modules, we calculated their eigengenes adjacency on their correlation of the entire modules. Each module showed independent validation to each other, and the progressively more saturated blue and red colors indicated the high co-expression interconnectedness (Figure1B). Genes within the corresponding modules indicated higher correlation than the genes between different modules from the network heatmap plot of selected genes.

In order to figure out which factor mainly affected the variation in PBMC transcriptomes, the correlation coefficients between modules and the trait of age, gender and race were calculated respectively. Three of these traits were slightly correlated with one or more modules, and the trait-module relationships with $P \leqslant 0.01$ were further analyzed (Figure 1C, Supplementary Table S4). Notably, among them, darkolivegreen module were negatively correlated with age (Pearson r=0.45, *p*=2.02e-08; Figure 1C), while Green module positively correlated with age (Pearson r=0.39, *p*=8.30e-07; Figure

1C). Furthermore, a significant correlation was also detected between sex and race in terms of their transcriptomic aging signatures respectively. Then, to study whether PBMC transcriptomic changes were acquired gradually over lifespan or more rapidly at certain ages, we detected age brackets during which abrupt changes take place, referred to as breakpoints. Within each aging related modules, we compared transcriptomic profiles observed at ages. Finally, these analyses revealed two periods in adult lifespan during which rapid changes occur: (i) a timepoint in earlyforties, and (ii) a later timepoint after 70 years of age (Figure 1D, E, Supplementary Table S5).

Besides, racial/ethnic differences in PBMC aging among adult lifespan were also important because of the profound effect on health. To identify major sources of variation in transcriptomic data, we conducted the principal component analyses (PCA) using expressed genes (n = 19,089) from high-quality samples (153 microarray data, Supplementary file1 Table S15). The first principal component (PC1) captured 14.9% of the variation in 153 microarray data and associated to age groups (Figure 1F, Supplementary Table S6). We could see PC1 differences between Asian and White samples were more significant (Figure 1G), and this also took place in Aisan and Black or Africa American(Figure 1G). Together, these results suggested that aging and race had both influenced the variation in PBMC transcriptomes, where it was unclear to what extent these aging-associated changes were shared in different races, such as Asian and White.

Aging-related changes in PBMC transcriptomes in Asian and its aging-related modules detection. To determine transcriptome profiles over PBMC of aging, 19 community dwelling healthy volunteers (10 women, 9 men) whose ages spanned 21–93 years old were collected for bulk RNA-seq at various stages. Principal component analysis (PCA) revealed that young and old samples were divided into two parts, and women changed more largely than men, especially in the old women (Figure 2A). To further identify the related genes of PBMC aging, weighted gene coexpression network analyses (WGCNA) was conducted using FPKM of 29,367genes (FPKM >1 of all sequenced points, Figure 2B) and the trait of age and sex. Genes with the same expression pattern were clustered into the same module to generate a cluster dendrogram (Figure 2B). The sample dendrogram and trait heatmap were visualized to understand the relationship between the corresponding gene expression data and biological traits (Figure 2C). 40 modules were obtained, of which the four module (cyan, darkturquoise, orange, brown) showed significant correlations with age, with the pearson correlation coefficients of absolute R $\geqq$0.70 (P < 0.01) (Figure 2C, Supplementary File 1, Tables S9). Further, a consensus clustering also confirmed the four main group were clearly separated by the 19 aging samples from young to old (Figure 2D), respectively. Similarly, the four interesting modules based on ME expression profile and 19 samples with extract age trait from young to old were also displayed in Figure 2E. The module eigengene E in Y-value was defined as the first principal component of a given module and it could be considered a representative of the gene expression profiles in a module,

as shown in figure S3 (Supplementary Table S8). These results suggested these four gene modules were highly associated with chronological age in Asian, especially for the brown and darkturquoise module.

Novel and known age-associated genes and pathways associated with PBMC aging in Asian. WGCNA analysis defined that the ME was the first principal component of a given module and could be considered as a representative of the module's gene expression profile. Based on ME expression profile of the four interesting modules, the expression of cyan, darkturquoise and orange modules were downregulated, while brown module showed the opposite results (Figure 3A). To further explore the biological functions of the brown and darkturquoise modules, we performed GO term enrichment analysis, as well as pathway ontology analyses by using clusterProfiler R package(Yu et al., 2012). Top GO biological processes and KEGG pathway in each module was shown in Figure 3B, C. For the brown module, the top two enriched terms in GO ontology were "Cellular amino acid metabolic process" (FDR =5.74E-04) or " Negative regulation of neuron apoptotic process" (FDR =8.43E-04). For the KEGG pathway analysis, the top enriched terms were "Herpes simplex virus 1 infection" (FDR = 9.19E-09) and "Valine, leucine and isoleucine degradation" (FDR = 1.33E-03). For darkturquoise module genes, the top enriched terms in the "GO databases were Protein-DNA complex subunit organization" (FDR = 7.68E-07) and " ncRNA processing" (FDR = 1.33E-06). Moreover, genes in darkolivegreen module were found to be significantly enriched in protein export and lysine degradation signaling pathway. The complete annotation for these two module was provided in Supplementary File 1, Tables S10. These findings together with previous research, which found persistent virus infections and metabolic disregulation were closely related with immune aging(Brunner et al., 2011; Hamrick and Stranahan, 2020), implied that the above signaling pathways virus infection and dysregulated metabolic process played an important role in aging. To identify key genes associated with chronological age, we performed a more detailed analysis of the brown and darkturquoise modules. First, based on the cut-off criteria of $|logFC| \geq 1$ and adj. $P < 0.05$, a total of 924 DEGs in the 19 Chinese PBMC transcriptomic data were found to be dysregulated in old individuals by the limma package. Then, as shown in veen diagram Figure S4, 278 and 19 co-expression genes were found in the brown module and darkturquoise module in 19 Chinese PBMC DEG dataset, respectively. In total, the overlapping genes were extracted for validating the genes of co-expression modules. Then, the PPI network among the overlapped genes was established by using the STRING database. The hub genes selected from the PPI network using the MCC algorithm of CytoHubba plugin were shown in (Figure 3D,E). According to the MCC sores, the top six highest-scored genes, including probable ATP-dependent RNA helicase (DDX27), Signal recognition particle subunit (SRP68) , E3 ubiquitin-protein ligase (RNF25), Transmembrane protein 131-like (TMEM131L), UDP-N-acetylglucosamine--peptide N-acetylglucosaminyltransferase 110 kDa subunit (OGT), Early endosome antigen 1 (EEA1) , exhibiting the highest connections with other genes were identified for further investigation (Figure 3D,E). Strikingly, the mRNA

abundance of these hub genes were significantly associated with chronological age, as shown in figure (Figure 3F, G). Surprisingly, according to reports in the literature, Nesrine Maharzi, et.al demonstrated TMEM131L could regulate immature single-positive thymocyte proliferation arrest by acting through mixed Wnt-dependent and -independent mechanisms(Maharzi et al., 2013). Reports demonstrated O-GlcNAc transferase (OGT) level was decreased in multiple aged tissues and reminded us that dysregulation of OGT related O-GlcNAc formation might play an important role in the development of age-related diseases (Fulop et al., 2008). Researcher also reported the abundance of EEA1 proteins was altered in the brains of aged mice(Ve et al., 2020). Moreover, SRP68 has been reported its association with cellular senescence, while the ubiquitination-related genes RNF25 is not clear in immune aging. These data support the notion that TMEM131L, OGT, EEA1, DDX27, SRP68 and RNF25 play an important role during PBMC aging, which may function as the novel candidate biomarkers for Chinese individuals.

Aging-related changes in PBMC transcriptomes in White and its hub gene detection. Similarly, to figure out the aging-related gene modules in PBMC transcriptomes in White individuals, we performed weighted gene co-expression network analysis (WGCNA) to reconstruct miroarray data in 113 White individuals, including 48 young (<40 years), 25 middle aged (40–65 years), and 41 normal aged (65–90 years). Then, a total of 16,376 genes from these transcriptomic data were used for this computation (Supplementary Table S11). 20 major gene modules (brown and turquoise, each containing ≧296 genes) were identified. Then, we plotted the heatmap of module-trait relationships to evaluate the association between each module and the trait of age and sex. The results of the module-trait relationships were presented in Figure 4A (Supplementary Table S12), revealing that the brown module and turquoise module were found to have the highest association with chronological age (brown module: r = 0.52, p = 2.45e-09; turquoise module: r = -0.47, p = 2.05e−07). More interestingly, we found these two aging related modules revealed there were two periods in the human lifespan during which the immune system underwent abrupt changes: (i) a timepoint in early thirties, and (ii) a later timepoint after 65 years of age (Figure 4B, C). To gain further insight into the potential functions of these two co-expression modules, gene enrichment analysis was performed by the *clusterProfiler* package. After screening of GO and KEGG enrichment analysis, we observed several enriched gene sets shown in Figure 4D, E. The biological process (BP) of brown and turquoise modules were mainly enriched in hormone transport, and postsynaptic specialization respectively (Figure 4D, Supplementary table S13). Moreover, on the KEGG pathway enrichment analysis, the genes of brown module was mainly categorized into long-term depression and gap junction, while the turquoise module was mainly enriched in phototransduction and hedgehog signaling pathway (Figure 4E). In the following, we focused on the core genes of the brown and turquoise modules. By using the differential expression analysis, we identified 1185 genes differentially expressed with chronological age in White. The venn diagram for the DEG, brown and turquoise modules

were listed as shown in Figure 4F. Similarly, the 50 overlapping genes in brown module and 177 overlapping genes in turquoise module were then to processed to detect the hub genes by using the STRING database, repectively. Then, the hub genes selected from the PPI network using the MCC algorithm of CytoHubba plugin were shown in Figure 4G. In the following, we focused on the core genes of the blue and turquoise module. The top two hub genes (Adenylate Cyclase 4, ADCY4; Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit alpha isoform, PIK3CA) in turquoise module were significantly down-regulated in PBMCs of old adults (Figure 4H), whereas immunoglobulin superfamily DCC subclass member 2 (NEO1) from brown module showed the opposite result in the white cohorts (Figure 4H). From the aging atlas website (https://bigd.big.ac.cn/aging/age_related_genes), ATP Pyrophosphate-Lyase 4 (ADCY4) and Serine/Threonine Protein Kinase (PIK3CA) have both involved in Longevity regulating pathway. As reported, ADCY4 catalyzes the formation of the signaling molecule cAMP in response to G-protein signaling(Ludwig and Seuwen, 2002), and PIK3CA participates in cellular signaling in response to various growth factors, which also involved in the activation of AKT1 upon stimulation by receptor tyrosine kinases ligands such as EGF, insulin, IGF1, VEGFA and PDGF(Yamaguchi et al., 2011). Besides, neogenin-1 (NEO1) has been reported associated with the long-term HSCs (LT-HSCs) expand during age(Gulati et al., 2019). Taken together, these data also revealed that ADCY4, PIK3CA, and NEO1 involved in aging importantly, which might serve as the novel candidate biomarkers in White individuals during the PBMC aging.

Shared transcriptomic signatures of aging between White and Asian. As racial/ethnic differences in age-expectations(Menkin et al., 2017), our first aim was to test whether PBMC aging differed across racial/ethnic groups. As the brown module from Asian and the turquoise module from White were both negatively correlated with chronological age, we compared these two modules to find out the common expressed genes by venn diagram. So as shown in Figure 5A, 95 genes in Asian and White significantly overlapped, despite thousands of race-specific gene associated with aging corresponding to 2623 and 1688 genes in Asian and White. Functional annotation of the 95 shared genes using R package clusterprofier revealed that these genes were highly enriched in the GO biological process of hindbrain development and coenzymeA metabolic process, as well as in the KEGG pathway of TGF-beta signaling significantly (Figure 5B,C). To uncover potential regulators of common transcriptomic changes in Asian and White, we identified hub genes by using the STRING database. According to the MCC sores from CytoHubba plugin, the top highest-scored genes, including peroxisomal coenzyme A diphosphatase (NUDT7) and caseinolytic peptidase B protein homolog (CLPB), were selected as the hub genes (Figure 5D). Meanwhile, by using the genes between the DEG lists and co-expression modules, the venn analysis of differentially expressed genes (DEGs) with the age-related modules in White and Asian revealed two aging-specific gene markers (Figure 5E). And then, as

shown in figure 5F, two overlapping genes (OXNAD1 and MLLT3) were both downregulated in the old adults in Asian and White. Combined analysis of DEGs and aging-related modules in Asian/White uncovered shared changes in pathway enrichment and hub genes. These analysis further highlighted the stark differences between races regarding aging of transcriptomic.

Validated shared genes involved in PBMC aging. After the 4 hub genes (NUDT7, CLPB, OXNAD1, MLLT3) common shared in Asian and White, we verified the expression levels of the hub genes among the individuals using the RNA-seq data and qPCR assay. As shown in Figure 6A, B, all of the 4 hub genes were found to be significantly downregulated in old individuals compared with the youth in Asian and White. Interestingly, they were all down-regulated in women during their lifespan in both White and Asian, as shown in Figure 6C. To further investigated whether these 4 hub genes expressed differentially across the progressive stages of PBMC aging, we measured these four hub genes mRNA levels (NUDT7, CLPB, OXNAD1 and MLLT3) in extracts of PBMC from 7 young adult (ages 21 – 30), and 5 aged health adults (ages 74+). Similarly, the mRNA level of NUDT7, CLPB, OXNAD1 and MLLT3 were both remarkably down-regulated in the aging individuals' PBMC, as verified by quantitative real time RT-PCR (qRT-PCR) (Figure 6D). The data in vivo above indicates a rather close relationship between hub genes and normal PBMC aging in Asian and White.

Methods

Human subjects. All studies were conducted following approval by the Ethics Committee of Jinan University (Approval#:KY-2020-027). Following informed consent, blood samples were obtained from 31 healthy volunteers residing in the Guangzhou, China region recruited by the Guangzhou First People's Hospital. For older adults 65 years and older, recruitment criteria were selected to identify individuals who are experiencing "usual healthy" aging and are thus representative of the average or typical normal health status of the local population within the corresponding age groups. Selecting this type of cohort is in keeping with the 2019 NIH Policy on Inclusion Across the Lifespan (NOT-98-024)(Kuchel, 2019), increasing the generalizability of our studies and the likelihood that these findings can be translated to the general population(Robertson et al., 2009). Subjects were carefully screened in order to exclude potentially confounding diseases and medications, as well as frailty. Individuals who reported chronic or recent (i.e., within two weeks) infections were also excluded. Subjects were deemed ineligible if they reported a history of diseases such as congestive heart failure, ischemic heart disease, myocarditis, congenital abnormalities, Paget's disease, kidney disease, diabetes requiring insulin, chronic obstructive lung disease, emphysema, and asthma. Subjects were also excluded if undergoing active cancer treatment, prednisone above 10 mg day, other immunosuppressive drugs, any medications for rheumatoid arthritis other than

NSAIDs or if they had received antibiotics in the previous 6 months. Finally, smoking history data are not typically collected in these studies—including 19 Chinese individuals—since smoking is a rare habit among older adults.

Ethics. The study was conducted following approval by the Ethics Committee of the first affiliated Hospital of Jinan University (Approval#:KY-2020-027). All study participants provided written informed consent at baseline using institutional review board approved forms. Individual-level human transcriptomic data (RNA-seq) are shared in the Sequence Read Archive (SRA) database (SRA: PRJNA703752). Our study complies with Tier 1 characteristics for " Biospecimen reporting for improved study quality " (BRISQ) guidelines.

RNA-seq library generation and processing. Total RNA was isolated from PBMCs using the TRIzol (Invitrogen, United States) following manufacturer's protocols. During RNA isolation, DNase I treatment was additionally performed using the RNase-free DNase set (Qiagen). RNA quality was checked using an Agilent 2100 Bioanalyzer instrument, together with the 2100 Expert software and Bioanalyzer RNA 6000 pico assay (Agilent Technologies). RNA quality was reported as a score from 1 to 10, samples falling below threshold of 8.0 being omitted from the study. cDNA libraries were prepared using either the TruSeq Stranded Total RNA LT Sample Prep Kit with Ribo-Zero Gold (Illumina) or KAPA Stranded mRNA-Seq Library Prep kit (KAPA Biosytems) according to the manufacturer's instructions using 100 ng or 500 ng of total RNA. Final libraries were analyzed on a Bioanalyzer DNA 1000 chip (Agilent Technologies). Paired-end sequencing (2 × 100 bp) of stranded total RNA libraries was carried out in either Illumina NextSeq500 using v2 sequencing reagents or the HiSeq2500 using SBS v3 sequencing reagents. Quality control (QC) of the raw sequencing data was performed using the FASTQC tool, which computes read quality using summary of per-base quality defined using the probability of an incorrect base call. According to our quality criteria, reads with more than 30% of their nucleotides with a Phred score under 30 are removed, whereas samples with more than 20% of such low-quality reads are dropped from analyses. Benchmarking is also applied on RNA-seq data using the same benchmark parameters as ATAC-seq, which resulted in 304 benchmark genes, none of the RNA-seq samples were dropped due to poor quality. Reads from samples that pass the quality criteria were quality-trimmed and filtered using trimmomatic. High-quality reads were then used to estimate transcript abundance using RSEM51. Finally, to minimize the interference of non-messenger RNA in our data, estimate read counts were renormalized to include only protein-coding genes. Supplementary TableS7 summarizes the FKPM values of PBMC from 19 Chinese individuals.

Microarray data obtaining. the microarray-based expression from 10 KIP provided by Lu et al.(Zalocusky et al., 2018) , was downloaded from the 10,000 immunomes project (10KIP, http://10kimmunomes.org/). This dataset contained quantile normalized genome-wide expression profiles of 153 adult human PBMC samples from all ages,

including samples from 65 young (ages 21–40: 24 men, 41 women), 40 middle-aged (ages 41–64: 18 men, 22 women), and 48 older subjects (65+: 20 men, 28 women) and containing three races including 19 Asian, 113 White and 21 Black or Africa American.

Identification of Key Co-expression Modules Using WGCNA

Co-expression networks facilitate methods on network-based gene screening that can be used to identify candidate biomarkers and therapeutic targets. In our study, the gene expression data profiles of microarray data and RNA-seq profile were constructed to gene co-expression networks using the WGCNA package in R respectively. WGCNA was used to explore the modules of highly correlated genes among samples for relating modules to external sample traits(Langfelder and Horvath, 2008). To build a scale-free network, optimum soft powers β were selected using the function pickSoftThreshold. Next, the adjacency matrix was created by the following formula: aij = |Sij|β (aij: adjacency matrix between gene i and gene j, Sij: similarity matrix which is done by Pearson correlation of all gene pairs, β: softpower value), and was transformed into a topological overlap matrix (TOM) as well as the corresponding dissimilarity (1-TOM). Afterwards, a hierarchical clustering dendrogram of the 1-TOM matrix was constructed to classify the similar gene expressions into different gene co-expression modules. To further identify functional modules in a co-expression network, the module-trait associations between modules, and clinical trait information were calculated according to the previous study(Hu et al., 2018). Therefore, modules with high correlation coefficient were considered candidates relevant to clinical traits, and were selected for subsequent analysis. A more detailed description of the WGCNA method was reported in our previous study(Hu et al., 2018).

Differential Expression Analysis and Interaction With the Modules of Interest

The R package limma (linear models for microarray data) provides an integrated solution for differential expression analyses on RNA-Sequencing and microarray data(Ritchie et al., 2015). In order to find the differentially expressed genes (DEGs) between Asian and White, limma was applied in the Asian RNA-seq and White dataset, respectively, to screen out DEGs. The p-value was adjusted by the Benjamini–Hochberg method to control for the false discovery Rate (FDR)(Benjamini and Hochberg, 1995). Genes with the cut-off criteria of $|logFC| \geq 0.50$ and adj. $P < 0.05$ were regarded as DEGs. The DEGs of the Asian and White dataset were visualized as a volcano plot by using the R package ggplot2 (Wickham H, 2016). Subsequently, the overlapping genes between DEGs and co-expression genes that were extracted from the co-expression network were used to identify potential prognostic genes, which were presented as a Venn diagram using the R package VennDiagram(Chen and Boutros, 2011).

Functional Annotation for the Modules of Interest

For genes in each module, Gene Ontology (GO) and KEGG pathway enrichment analysis were conducted to analyze the biological functions of modules. Significantly enriched GO

terms and pathways in genes in a module comparing to the background were defined by hypergeometric test and with a threshold of false discovery rate (FDR) less than 0.05. The clusterProfiler package offers a gene classification method, namely groupGO, to classify genes based on their projection at a specific level of the GO corpus, and provides functions, *enrichGO* and *enrichKEGG*, to calculate enrichment test for GO terms and KEGG pathways based on hypergeometric distribution(Yu et al., 2012). Thus, we input the interesting modules into the clusterProfiler by comparing them to the annotated gene sets libraries, with a cut-off criterion of adjusted $p < 0.05$. GO annotation that contains the three sub-ontologies—biological process (BP), cellular component (CC), and molecular function (MF)—can identify the biological properties of genes and gene sets for all organisms (Yu et al., 2012).

Construction of PPI and Screening of Hub Genes

In our study, we used the STRING (Search Tool for the Retrieval of Interacting Genes) online tool, which is designed for predicting protein–protein interactions (PPI), to construct a PPI network of selected genes(Szklarczyk et al., 2019). Using the STRING database, genes with a score ≥ 0.4 were chosen to build a network model visualized by Cytoscape (v3.7.2)(Shannon et al., 2003). In a co-expression network, Maximal Clique Centrality (MCC) algorithm was reported to be the most effective method of finding hub nodes(Chin et al., 2014). The MCC of each node was calculated by CytoHubba, a plugin in Cytoscape(Chin et al., 2014). In this study, the genes with the top 10 MCC values were considered as hub genes.

Verification of the Hub Genes

In order to confirm the reliability of the hub genes, we verified the expression patterns of the hub genes from healthy individuals including 7 young (ages: 23-30) and 5 old (ages: ≧ 74). The expression level of each hub gene between young and old individuals was plotted as a violin graph. Total RNA from PBMCs was extracted by TRIzol (Invitrogen, United States). Synthesis of cDNA was performed by using 2 μg of total RNA with PrimeScriptTM Reverse Transcriptase (Takara) according to the manufacturer's instructions. Specific primers used for qPCR were listed in the supplementary table S14. The gel image was acquired in the Gel Doc 1000 system and analyzed using the Quantity One software (Bio-Rad Laboratories, Hercules, CA, United States). ACTB was chosen as the endogenous control and cycle dependence was carried out to ensure that the PCR products fell within the linear range. Quantitative real-time PCR was performed using the SYBR® Premix Ex Taq Kit (Takara) in a 7900 Real Time PCR System (Applied Biosystems, United States) for at least three independent experiments. The relative quantification expression of each gene was normalized to ACTB, and calculated using the $2^{-\Delta\Delta\Delta CT}$ method.

Statistical Analysis

All experiments were performed for at least three independent times, and the data were expressed as the mean ± standard deviation (SD). All statistical analysis was performed using GraphPad Prism 6 Software (GraphPad Software, San Diego, CA, United States). Comparison between two groups was conducted by using Student's t-test. P-values less than 0.05 were considered as statistically significant.

Discussion

Age-associated changes in gene expression levels point towards altered activity in defined age-related molecular pathways that may play vital roles in the mechanisms of increased susceptibility to ageing diseases. In contrast to earlier studies of human age-related molecular differences(van den Akker et al., 2014; Peters et al., 2015), we detected 177 adult individuals from all ages among Asian, White and Africa and American ancestry. In our study, a total of four significant gene modules with the same expression trends were identified using integrated bioinformatic analysis in Asian and White populations. As suggested in functional annotation analysis by the *clusterProfiler* package, these module genes were mainly enriched in amino acid metabolic and differentiation, which are basic processes in ageing mechanisms including dysregulation of herpes simplex virus 1 infection, Valine, leucine and isoleucine degradation, long-term depression, gap junction, and hedgehog signaling pathway. Furthermore, according to MCC scores from the *CytoHubba plugin* in Cytoscape, the top chronological age related genes were screened out (namely TMEM131L, OGT, EEA1, DDX27, SRP68 and RNF25 in Asian; ADCY4, PIK3CA and NEO1 in White). According to reports in the literature, all of these genes are more or less closely associated with aging. Consistent with these reports, the expression of these genes were also found be significantly regulated among young and old individuals in our study, supporting these genes plays a causal role in human PBMC ageing. More importantly, the breakpoint analyses uncovered that although aging related transcriptomic changes accumulate gradually throughout adult life, there are two periods in the human lifespan during which the immune system undergoes abrupt changes. The two breakpoints (30 and 65-70 years old) were similar among races during the whole lifespan. The differences in the timing of age-related changes can be helpful in clinical decisions regarding when to start interventions/therapies.

Despite well-characterized race differences in immune responses, disease susceptibility, and lifespan, it is unclear whether aging differentially affects peripheral blood cells of European and Asian ancestry. To fill this gap, we generated RNA-seq data in PBMCs from 19 age-matched Chinese healthy adults in Guangdong area and downloaded microarray data of 153 individuals' PBMCs from 10 KIP (http://10kimmunomes.org/) including European, Asian and the other ancestry groups (19 Asian, 113 White and 21 Black or Africa American). Weighted gene correlation network analysis (WGCNA) is an integrated bioinformatic analysis, which could provide a

comprehensive characterization of the transcriptomic changes for disease's functional interpretation and led to new insights into the molecular aspects of clinical-pathological factors(Langfelder and Horvath, 2008). So by using this integrated bioinformatic analysis, we discovered a genomic signature of aging that is shared between race including (1) 95 age-associated genes in 132 individuals of European and Asian ancestry, (2) four hub genes (NUDT7, CLPB, OXNAD1 and MLLT3) all decreased in old ages. According to reports in the literature about these four hub genes, NUDT7, acts as a coenzyme A (CoA) diphosphatase, which mediates the cleavage of CoA. NUDT7 functions as a house-keeping enzyme by eliminating potentially toxic nucleotide metabolites, such as oxidized CoA from β-oxidation in the peroxisome, as well as nucleotide diphosphate derivatives, including NAD+, NADH, and ADP-ribose(Song et al., 2018). Furthermore downregulation of NUDT7 in mice accelerated senescence(Cho et al., 2003), and suppressed expression level was observed in the liver of starved mice(Bauer et al., 2004). Interestingly, OXNAD1 also known as oxidoreductase NAD-binding domain-containing protein, has been reported differentially expressed with chronological age(Peters et al., 2015). And according to the uniprot anotation for CLPB, it may function as a regulatory ATPase and be related to secretion/protein trafficking process, involves in mitochondrial-mediated antiviral innate immunity, and activates RIG-I-mediated signal transduction and production of IFNB1 and proinflammatory cytokine IL6(Yoshinaka et al., 2019). Moreover, the hub gene of MLLT3 is a component of the superelongation complex and co-operates with DOT1L, which di/trimethylates H3K79 to promote transcription(Steger et al., 2008; Li et al., 2014). Recently, Vincenzo Calvanese, et.al, found MLLT3 could govern human haematopoietic stem-cell self-renewal and engraftment(Calvanese et al., 2019). From above, NUDT7 and OXNAD1 both have an important role in cellular metabolism and aging, which was consistent with our finding of PBMC aging analysis, while the role of CLPB and MLLT3 in immune aging or senecence is unclear. Thus, by using co-expression networks, we identified new genes that are likely important in PBMC aging in European and Asian ancestry, opening new avenues of enquiry for future studies.

Changes in bulk PBMCs in different races were shown using aging-specific regulatory modules and hub genes by WGCNA analysis. Although this approach was effective in annotating the aging signatures, it is prone to biases in the differences of data quality and formats. Future studies are needed to describe these race differences at single-cell resolution and in sorted cells and to establish their functional implications. Moreover, future studies are needed to study important molecules identified here (NUDT7, CLPB, OXNAD1 and MLLT3) as aging specific biomarkers of immune system aging. Besides, we had much smaller sample sizes for both PBMCs in European, Asian and the other ancestry groups, we used a nominal P-value threshold (p<0.05) in these specific sub-analyses. Larger sample sizes will ultimately be needed to fully understand the transferability of the aging-transcriptome signatures. Taken together, these findings indicate that aging plays a critical role in human immune system aging and should be

taken into consideration while searching for molecular targets and time frames for interventions/therapies to target aging and age-related diseases.

Reference

Bauer, M., Hamm, A.C., Bonaus, M., Jacob, A., Jaekel, J., Schorle, H., et al. (2004). Starvation response in mouse liver shows strong correlation with life-span-prolonging processes. *Physiol Genomics* 17(2), 230-244. doi: 10.1152/physiolgenomics.00203.2003.

Benjamini, Y., and Hochberg, Y. (1995). Controlling The False Discovery Rate - A Practical And Powerful Approach To Multiple Testing. *J. Royal Statist. Soc., Series B* 57, 289-300. doi: 10.2307/2346101.

Brunner, S., Herndler-Brandstetter, D., Weinberger, B., and Grubeck-Loebenstein, B. (2011). Persistent viral infections and immune aging. *Ageing Res Rev* 10(3), 362-369. doi: 10.1016/j.arr.2010.08.003.

Calvanese, V., Nguyen, A.T., Bolan, T.J., Vavilina, A., Su, T., Lee, L.K., et al. (2019). MLLT3 governs human haematopoietic stem-cell self-renewal and engraftment. *Nature* 576(7786), 281-286. doi: 10.1038/s41586-019-1790-2.

Chen, H., and Boutros, P.C. (2011). VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* 12, 35. doi: 10.1186/1471-2105-12-35.

Cheung, P., Vallania, F., Warsinske, H.C., Donato, M., Schaffert, S., Chang, S.E., et al. (2018). Single-Cell Chromatin Modification Profiling Reveals Increased Epigenetic Variations with Aging. *Cell* 173(6), 1385-1397 e1314. doi: 10.1016/j.cell.2018.03.079.

Chin, C.H., Chen, S.H., Wu, H.H., Ho, C.W., Ko, M.T., and Lin, C.Y. (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 8 Suppl 4, S11. doi: 10.1186/1752-0509-8-S4-S11.

Cho, Y.M., Bae, S.H., Choi, B.K., Cho, S.Y., Song, C.W., Yoo, J.K., et al. (2003). Differential expression of the liver proteome in senescence accelerated mice. *Proteomics* 3(10), 1883-1894. doi: 10.1002/pmic.200300562.

Choong, M.L., Ton, S.H., and Cheong, S.K. (1995). Influence of race, age and sex on the lymphocyte subsets in peripheral blood of healthy Malaysian adults. *Ann Clin Biochem* 32 (Pt 6), 532-539. doi: 10.1177/000456329503200603.

Fulop, N., Feng, W., Xing, D., He, K., Not, L.G., Brocks, C.A., et al. (2008). Aging leads to increased levels of protein O-linked N-acetylglucosamine in heart, aorta, brain and skeletal muscle in Brown-Norway rats. *Biogerontology* 9(3), 139. doi: 10.1007/s10522-007-9123-5.

Gulati, G.S., Zukowska, M., Noh, J.J., Zhang, A., Wesche, D.J., Sinha, R., et al. (2019). Neogenin-1 distinguishes between myeloid-biased and balanced Hoxb5 (+) mouse long-term hematopoietic stem cells. *Proc Natl Acad Sci U S A* 116(50), 25115-25125. doi: 10.1073/pnas.1911024116.

Hamrick, M.W., and Stranahan, A.M. (2020). Metabolic regulation of aging and age-related disease. *Ageing Res Rev* 64, 101175. doi: 10.1016/j.arr.2020.101175.

He, S., and Sharpless, N.E. (2017). Senescence in Health and Disease. *Cell* 169(6), 1000-1011. doi: 10.1016/j.cell.2017.05.015.

Hu, Y., Pan, J., Xin, Y., Mi, X., Wang, J., Gao, Q., et al. (2018). Gene Expression Analysis

Reveals Novel Gene Signatures Between Young and Old Adults in Human Prefrontal Cortex. *Front Aging Neurosci* 10**,** 259. doi: 10.3389/fnagi.2018.00259.

Kuchel, G.A. (2019). Inclusion of Older Adults in Research: Ensuring Relevance, Feasibility, and Rigor. *J Am Geriatr Soc* 67(2)**,** 203-204. doi: 10.1111/jgs.15802.

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9**,** 559. doi: 10.1186/1471-2105-9-559.

Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., et al. (2014). AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell* 159(3)**,** 558-571. doi: 10.1016/j.cell.2014.09.049.

Ludwig, M.G., and Seuwen, K. (2002). Characterization of the human adenylyl cyclase gene family: cDNA, gene structure, and tissue distribution of the nine isoforms. *J Recept Signal Transduct Res* 22(1-4)**,** 79-110. doi: 10.1081/rrs-120014589.

Maharzi, N., Parietti, V., Nelson, E., Denti, S., Robledo-Sarmiento, M., Setterblad, N., et al. (2013). Identification of TMEM131L as a novel regulator of thymocyte proliferation in humans. *J Immunol* 190(12)**,** 6187-6197. doi: 10.4049/jimmunol.1300400.

Menkin, J.A., Guan, S.A., Araiza, D., Reyes, C.E., Trejo, L., Choi, S.E., et al. (2017). Racial/Ethnic Differences in Expectations Regarding Aging Among Older Adults. *Gerontologist* 57(suppl_2)**,** S138-S148. doi: 10.1093/geront/gnx078.

Moskowitz, D.M., Zhang, D.W., Hu, B., Le Saux, S., Yanes, R.E., Ye, Z., et al. (2017). Epigenomics of human CD8 T cell differentiation and aging. *Sci Immunol* 2(8). doi: 10.1126/sciimmunol.aag0192.

Nikolich-Zugich, J. (2018). The twilight of immunity: emerging concepts in aging of the immune system. *Nat Immunol* 19(1)**,** 10-19. doi: 10.1038/s41590-017-0006-x.

Noren Hooten, N., Longo, D., and Evans, M. (2018). "Age-and Race-Related Changes in Subpopulations of Peripheral Blood Lymphocytes in Humans."), 1-30.

Peters, M.J., Joehanes, R., Pilling, L.C., Schurmann, C., Conneely, K.N., Powell, J., et al. (2015). The transcriptional landscape of age in human peripheral blood. *Nat Commun* 6**,** 8570. doi: 10.1038/ncomms9570.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., et al. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7)**,** e47. doi: 10.1093/nar/gkv007.

Robertson, D. & Williams, G. H. (2009). Clinical and Translational Science: Principles of Human Research. (Academic Press, 2009).

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11)**,** 2498-2504. doi: 10.1101/gr.1239303.

Song, J., Baek, I.J., Chun, C.H., and Jin, E.J. (2018). Dysregulation of the NUDT7-PGAM1 axis is responsible for chondrocyte death during osteoarthritis pathogenesis. *Nat Commun* 9(1)**,** 3427. doi: 10.1038/s41467-018-05787-0.

Steger, D.J., Lefterova, M.I., Ying, L., Stonestrom, A.J., Schupp, M., Zhuo, D., et al. (2008). DOT1L/KMT4 recruitment and H3K79 methylation are ubiquitously coupled with gene transcription in mammalian cells. *Mol Cell Biol* 28(8)**,** 2825-2839. doi: 10.1128/MCB.02076-07.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., et al. (2019).

STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47(D1), D607-D613. doi: 10.1093/nar/gky1131.

Tollerud, D.J., Clark, J.W., Brown, L.M., Neuland, C.Y., Pankiw-Trost, L.K., Blattner, W.A., et al. (1989). The influence of age, race, and gender on peripheral blood mononuclear-cell subsets in healthy nonsmokers. *J Clin Immunol* 9(3), 214-222. doi: 10.1007/BF00916817.

van den Akker, E.B., Passtoors, W.M., Jansen, R., van Zwet, E.W., Goeman, J.J., Hulsman, M., et al. (2014). Meta-analysis on blood transcriptomic studies identifies consistently coexpressed protein-protein interaction modules as robust markers of human aging. *Aging Cell* 13(2), 216-225. doi: 10.1111/acel.12160.

Van, P., Jiang, W., Gottardo, R., and Finak, G. (2018). ggCyto: next generation open-source visualization software for cytometry. *Bioinformatics* 34(22), 3951-3953. doi: 10.1093/bioinformatics/bty441.

Ve, H., Cabana, V.C., Gouspillou, G., and Lussier, M.P. (2020). Quantitative Immunoblotting Analyses Reveal that the Abundance of Actin, Tubulin, Synaptophysin and EEA1 Proteins is Altered in the Brains of Aged Mice. *Neuroscience* 442, 100-113. doi: 10.1016/j.neuroscience.2020.06.044.

Whisler, R.L., Beiqing, L., and Chen, M. (1996). Age-related decreases in IL-2 production by human T cells are associated with impaired activation of nuclear transcriptional factors AP-1 and NF-AT. *Cell Immunol* 169(2), 185-195. doi: 10.1006/cimm.1996.0109.

Wickham H (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Yamaguchi, H., Yoshida, S., Muroi, E., Yoshida, N., Kawamura, M., Kouchi, Z., et al. (2011). Phosphoinositide 3-kinase signaling pathway mediated by p110alpha regulates invadopodia formation. *J Cell Biol* 193(7), 1275-1288. doi: 10.1083/jcb.201009126.

Yoshinaka, T., Kosako, H., Yoshizumi, T., Furukawa, R., Hirano, Y., Kuge, O., et al. (2019). Structural Basis of Mitochondrial Scaffolds by Prohibitin Complexes: Insight into a Role of the Coiled-Coil Region. *iScience* 19, 1065-1078. doi: 10.1016/j.isci.2019.08.056.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16(5), 284-287. doi: 10.1089/omi.2011.0118.

Zalocusky, K.A., Kan, M.J., Hu, Z., Dunn, P., Thomson, E., Wiser, J., et al. (2018). The 10,000 Immunomes Project: Building a Resource for Human Immunology. *Cell Rep* 25(2), 513-522 e513. doi: 10.1016/j.celrep.2018.09.021.

Data Availability Statement

All data relevant is contained within the article. The sequence data presented in the study has been deposited at the Sequence Read Archive (SRA) database, which is hosted by the NCBI, under accession number (SRA: PRJNA703752).

Ethics Statement

The studies involving human participants were reviewed and approved by the local ethics committee of the First Affiliated Hospital of Jinan University. The patients/participants provided their written informed consent to participate in this study.

Author Contributions

The concept of the study was planned by GC and YH. Experiments were conducted, analyzed, and interpreted by YH, LM, YX, and WL. Sample preparation for mRNA, and sequencing were done by YX, JX and WL. YH drafted the manuscript. GC edited the manuscript and provided advice. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from the National Key Research and Development Program of China (No. 2018YFC2002003) and the China Post-doctoral Science Foundation (No.55350399).

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgments

We would like to thank all donors as well as the Guangzhou First People's Hospital for provision of the samples. We thank the Beijing Novogene Biotechnology Co., Ltd. for assistance with RNA-sequencing as well as Lipeng Mao and Jian Xiang for excellent technical assistance.

Figure legend

Figure 1. Age and race have influenced the transcriptome changes over human adult lifespan. (A) Cluster dendrogram. Each color represents one specific co-expression module. In the colored rows below the dendrogram, the two colored rows represent the original modules and merged modules; (B) Eigengene adjacency heatmap of different modules; (C) Heatmap of the correlation between trait (age and sex) and module eigengenes. Each column corresponds to a module eigengene, and each row corresponds to a trait. Each cell contains the corresponding correlation. The table is color-coded by correlation according to the color legend. P-value < 0.05 represents statistical significance;(D-E) The characteristic gene expression changes during PBMC aging. The left-hand Y-axis represents the eigengene expression of each module, and the

right-hand Y-axis represnts the trend line for each individuals; (F-G) Principal component1 scores (PC1) were calculated for each individual from RNA-seq (left) principal component analyses (PCA) results. PC1 scores from RNA-seq data were differentially expressed among different races.

Figure 2. The characteristic gene expression of PBMC aging in Asian. (A) principal component analyses (PCA) for 19 Chinese PBMC RNA-seq data. Young and old individuals were largely separated according to the principal component1 scores (PC1). (B) Cluster dendrogram. Each color represents one specific co-expression module. In the colored rows below the dendrogram, the two colored rows represent the original modules and merged modules; (C) Supervised hierarchical cluster of each row correspond to a module eigengene (n = 40), column to a trait. Each cell contained the corresponding correlation. High correlations was colored in orange, low correlation in blue. (D) Hierarchical cluster analysis of four interested modules, based on the module-trait's correlation and p value (absolute $r > 0.5$, $P < 10^{-2}$), three modules (cyan, darkturquoise, orange, brown) showed relatively lower expression in young adults and high expression in the aged population. Conversely, the brown modules showed the opposite result. Each circle represented an individual. (E) The histograms described the eigengene expression of the four age-related module from young to old.

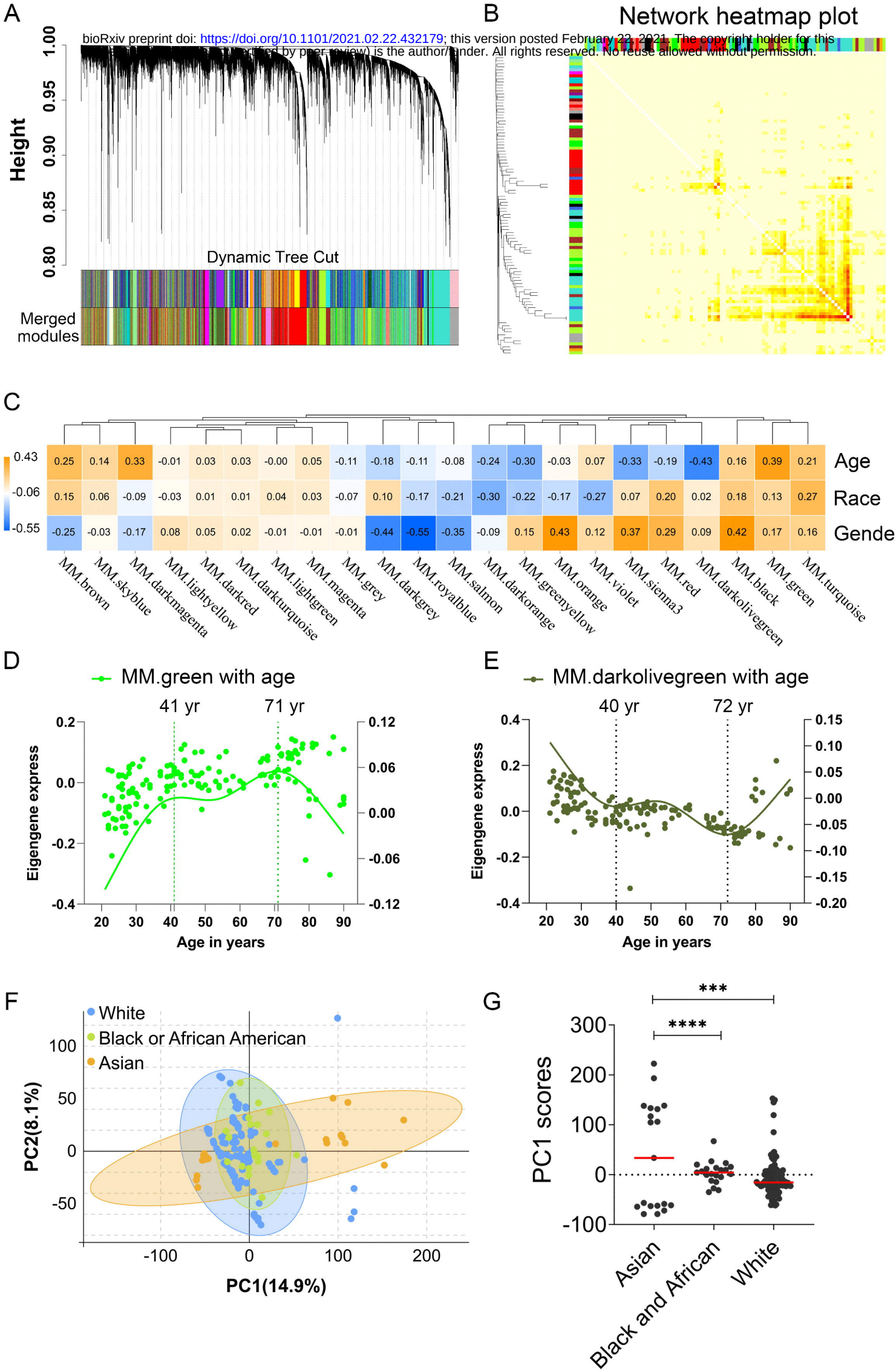
Figure 3. Novel and known age-associated genes and pathways associated with PBMC aging in Asian. (A) The transcriptomic expression of age related modules changed siginificantly among young and old individuals. Based on ME expression profile of the four interesting modules, the expression of cyan, darkturquoise and orange modules were downregulated, while brown module showed the opposite results. (B-C) GO functional annotation and KEGG pathway enrichment analysis for the brown and darkturquoise modules.(B) Top 10 enriched biological process; (C) Top 10 enriched kegg pathway. The color represents the adjusted p-values. (D-E) Hub gene detection for the brown and darkturquoise modules. PPI network of the brown and darkturquoise modules based on the STRING database. And each node represents a protein-coding gene and the size of each node is mapped to its degree. (F, G)The verification of the hub genes. The top of three genes in brown and darkturquoise modules were selected and its mRNA abundance of these hub genes were detected in young and old individuals.

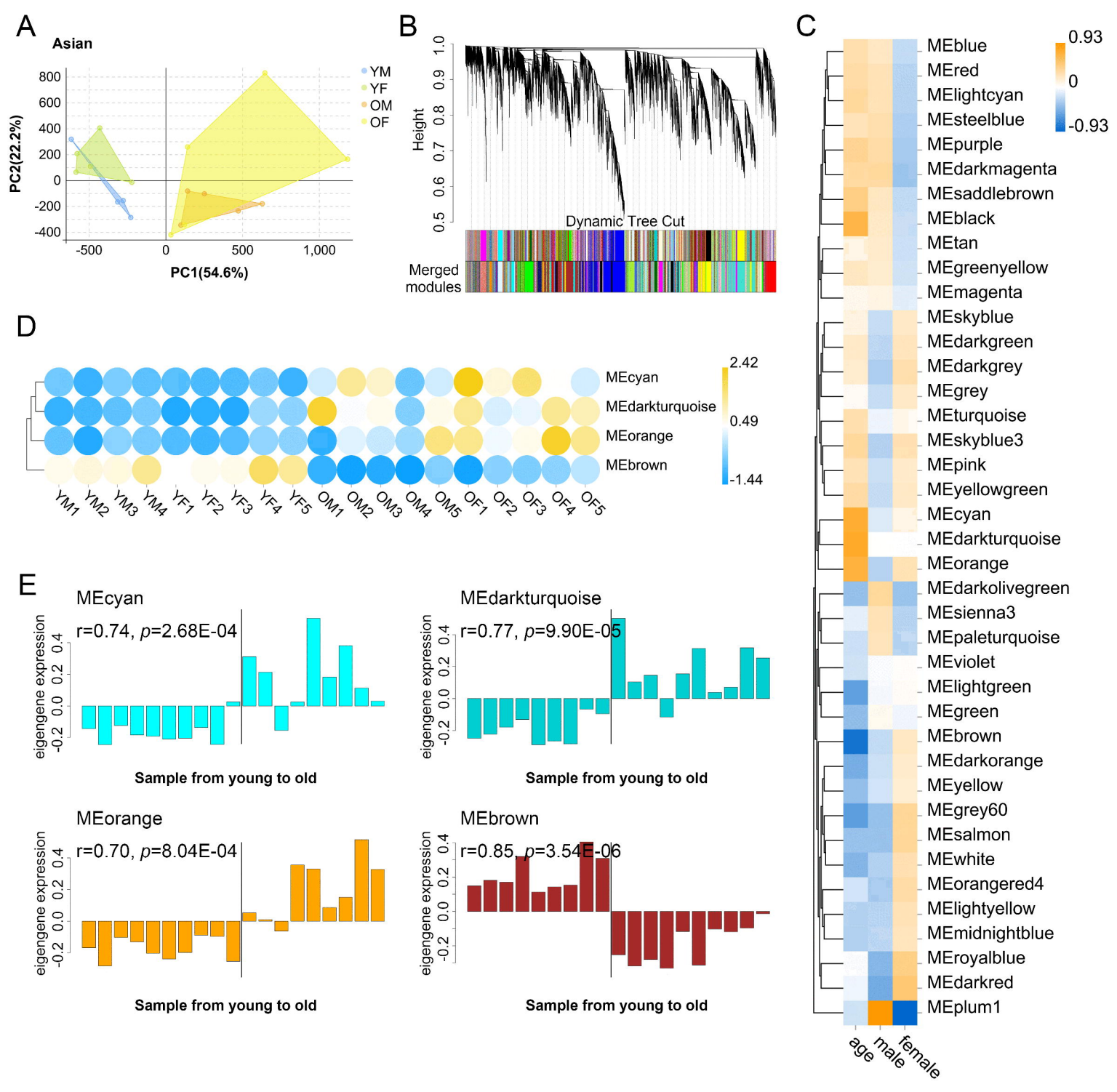
Figure 4. Aging-related changes in PBMC transcriptomes in White and its hub gene detection. (A) Module-trait relationships. Each row corresponds to a color module and column corresponds to a clinical trait (age and sex). Each cell contains the corresponding correlation and P-value. (B) The histograms described the eigengene expression of the four age-related module (brown and turquoise) from young to old. (C) Eigengene expression of the age-related modules (brown and turquoise). The transcriptomic expression of age related modules changed siginificantly among young and old individuals. (D-E) Gene Ontology (GO) and KEGG enrichment analysis for the genes in the brown and turquoise modules. The top 10 of the GO enriched biological process and

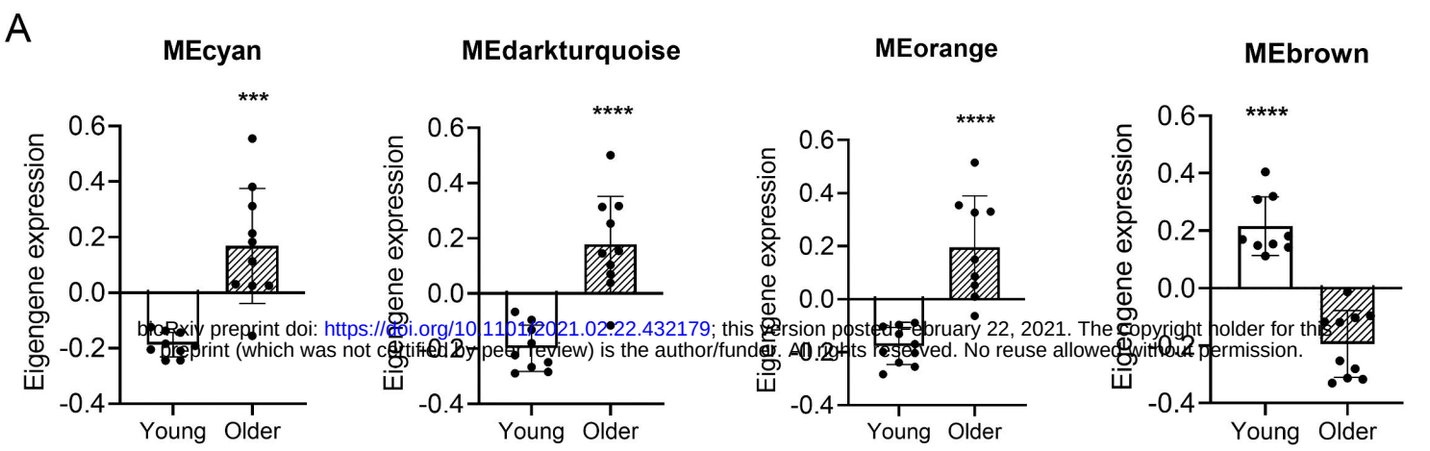
enriched kegg pathway were shown. The color represents the adjusted p-values. The color represents the adjusted p-values, and the size of the bars represents the gene number. (F)The venn diagram of genes among DEG lists and co-expression module. In total, 50 and 177 overlapping genes were listed in the intersection of DEG lists and two co-expression modules. (G) Hub genes detection for the brown and turquoise modules. (H) The verification of the hub genes. The top of genes in brown and turquoise modules were selected and its mRNA abundance of these hub genes were detected in young and old individuals.

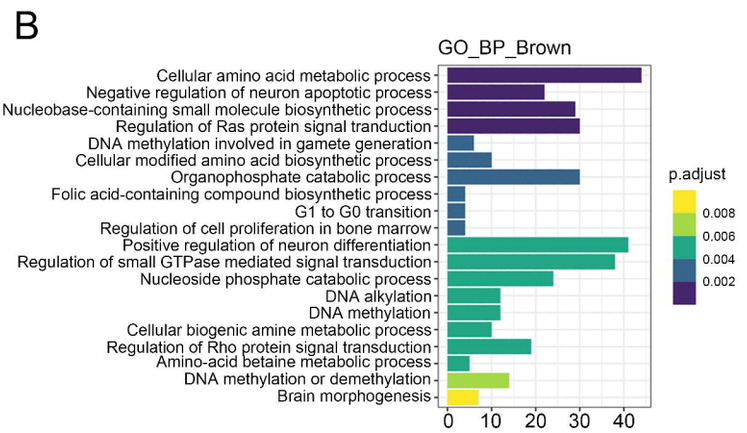
Figure 5. Shared transcriptomic signatures of aging between White and Asian. (A)The venn diagram of genes among turquoise module in White and brown module in Asian. Despite thousands of race-specific gene associated with aging corresponding to 2623 and 1688 genes in Asian and White, 95 genes in Asian and White significantly overlapped. (B-C) GO and KEGG enrichment analysis for the 95 common shared genes. (D) PPI network for the 95 genes. (E) The veen analysis of differentially expressed genes (DEGs) with the age-related modules in White and Asian revealed two aging-specific gene markers. (F) The two overlapping genes (OXNAD1 and MLLT3) were both downregulated, as shown in the volcano analysis for the DEG genes in Asian and White dataset.

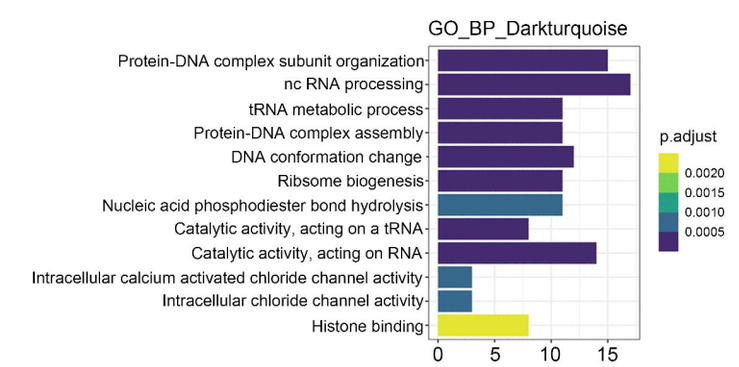
Figure 6. Validation of expression levels of the four common hub genes involved in PBMC aging(A, B) Gene expression value of the hub genes among young and old samples in Asian and White. Student's t-test was used for statistical analysis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$, $****p < 0.0001$. (C) Gene expression value of hub genes among samples of man and woman during their lifespan. (D) Quantification of the four hub genes was confirmed and presented by the qPCR assay. The P-value was calculated by the student's t-test.

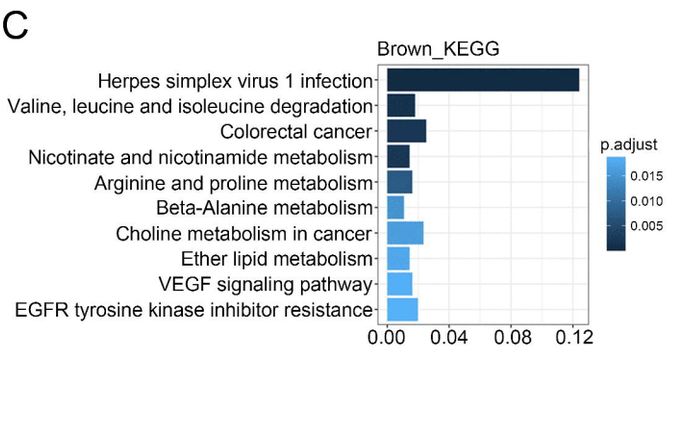


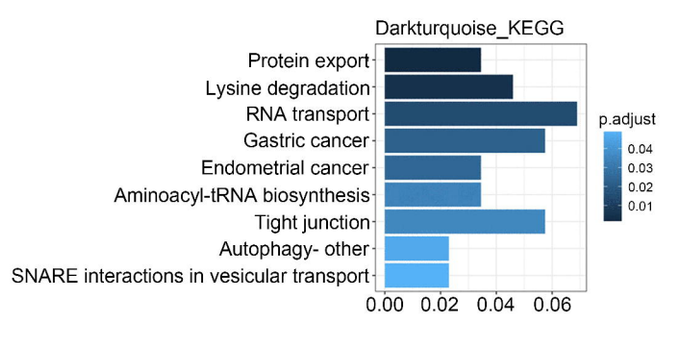

A
Asian
YM
YF
OM
OF
PC2(22.2%)
PC1(54.6%)
B
Height
Dynamic Tree Cut
Merged modules
C
MEblue
MEred
MElightcyan
MEsteelblue
MEpurple
MEdarkmagenta
MEsaddlebrown
MEblack
MEtan
MEgreenyellow
MEmagenta
MEskyblue
MEdarkgreen
MEdarkgrey
MEgrey
MEturquoise
MEskyblue3
MEpink
MEyellowgreen
MEcyan
MEdarkturquoise
MEorange
MEdarkolivegreen
MEsienna3
MEpaleturquoise
MEviolet
MElightgreen
MEgreen
MEbrown
MEdarkorange
MEyellow
MEgrey60
MEsalmon
MEwhite
MEorangered4
MElightyellow
MEmidnightblue
MEroyalblue
MEdarkred
MEplum1
0.93
0
-0.93
age
male
female
D
MEcyan
MEdarkturquoise
MEorange
MEbrown
2.42
0.49
-1.44
YM1
YM2
YM3
YM4
YF1
YF2
YF3
YF4
YF5
OM1
OM2
OM3
OM4
OM5
OF1
OF2
OF3
OF4
OF5
E
MEcyan
r=0.74, p=2.68E-04
MEdarkturquoise
r=0.77, p=9.90E-05
MEorange
r=0.70, p=8.04E-04
MEbrown
r=0.85, p=3.54E-06
eigengene expression
Sample from young to old

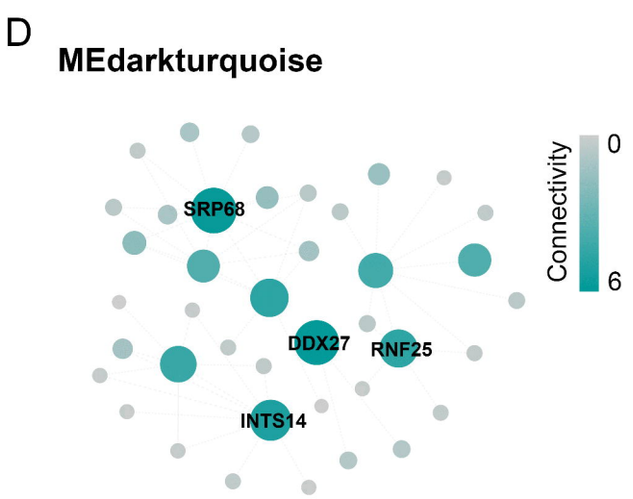


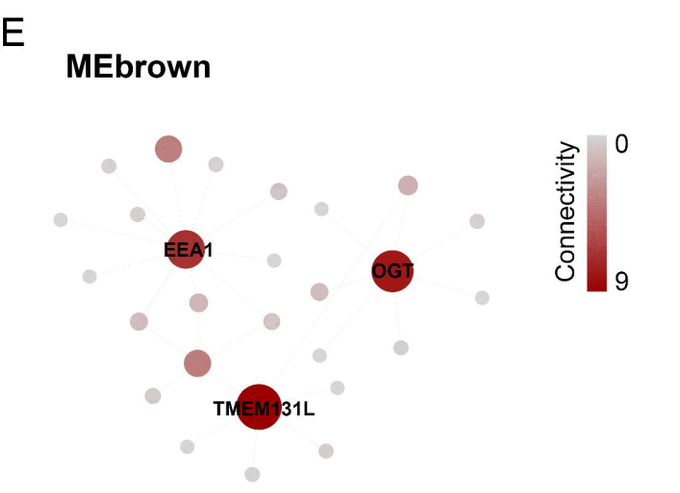


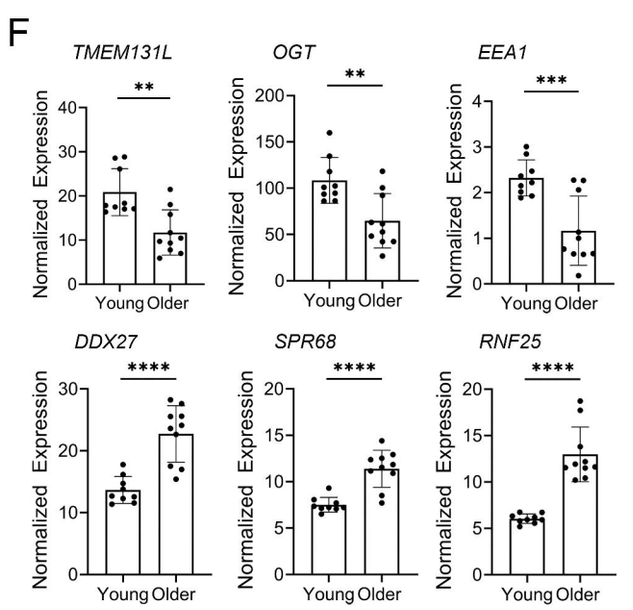


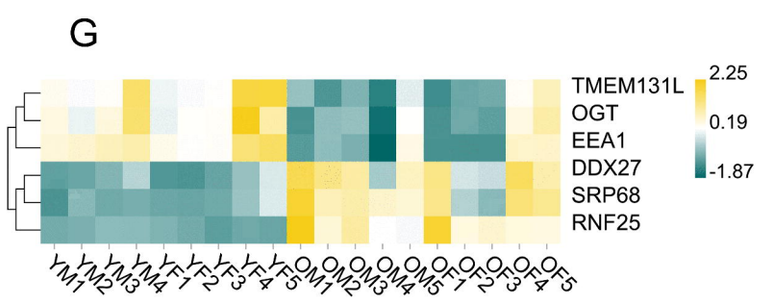


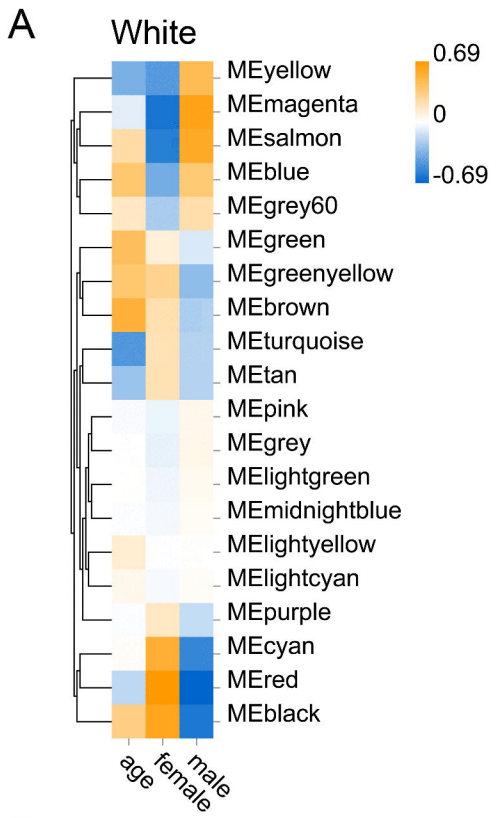


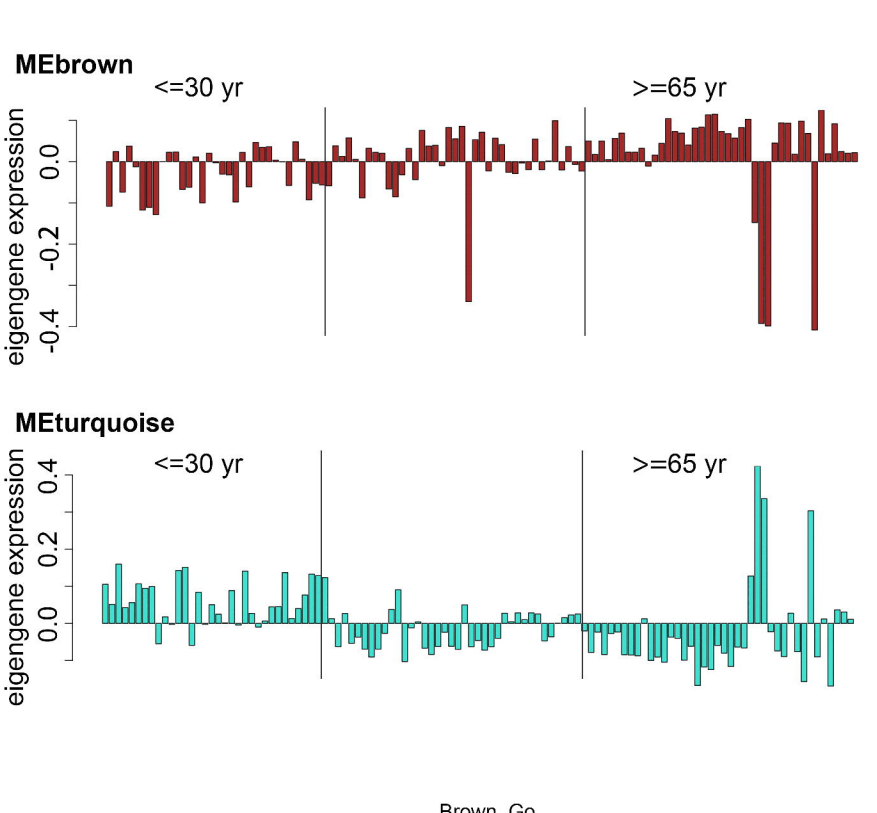


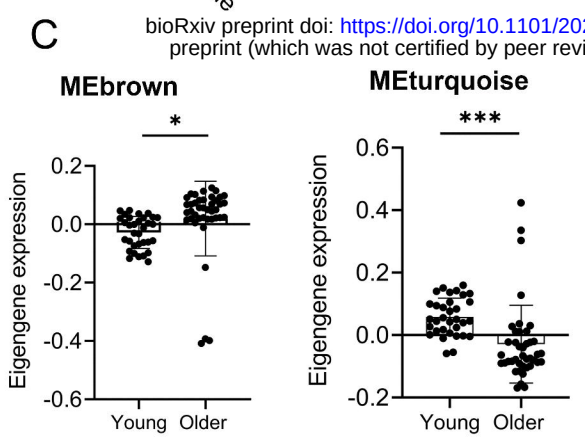


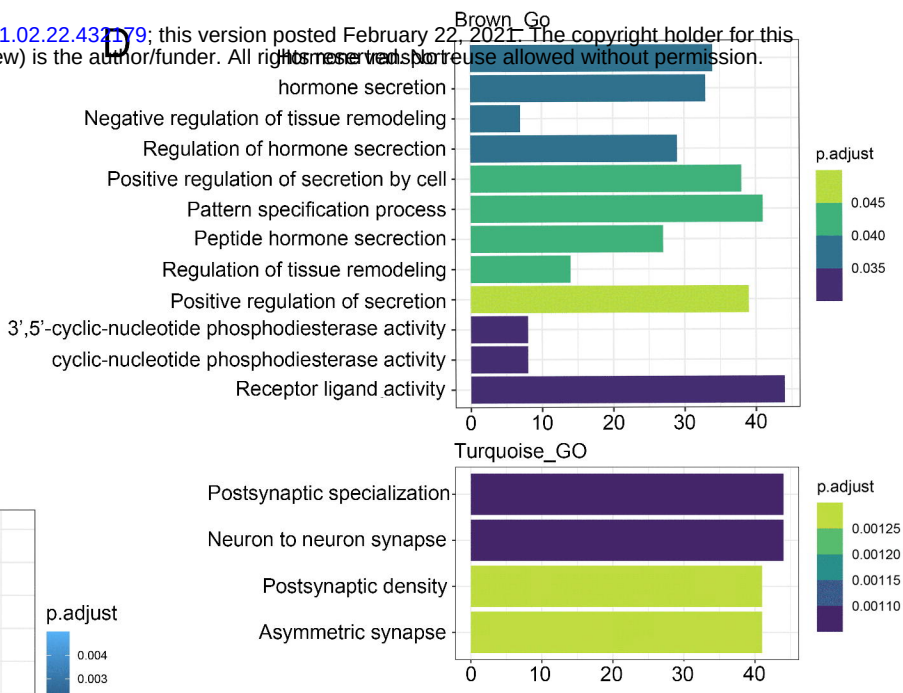


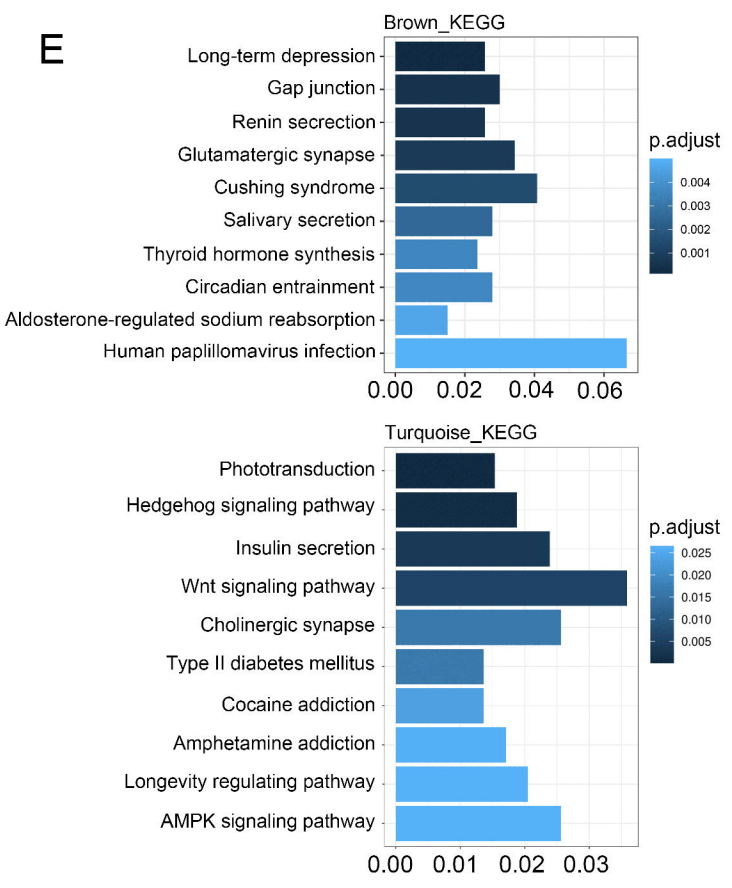


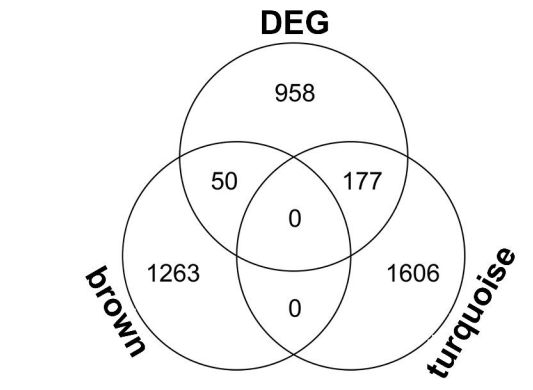


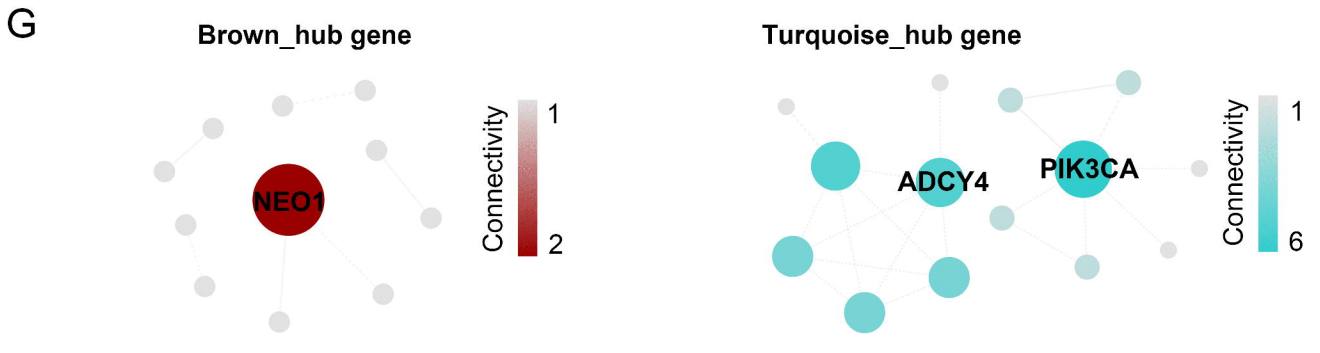


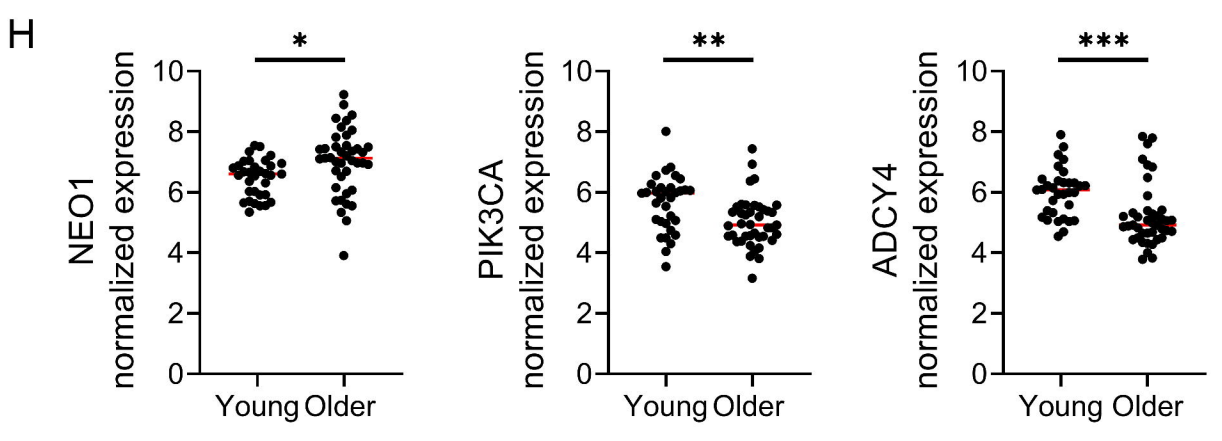


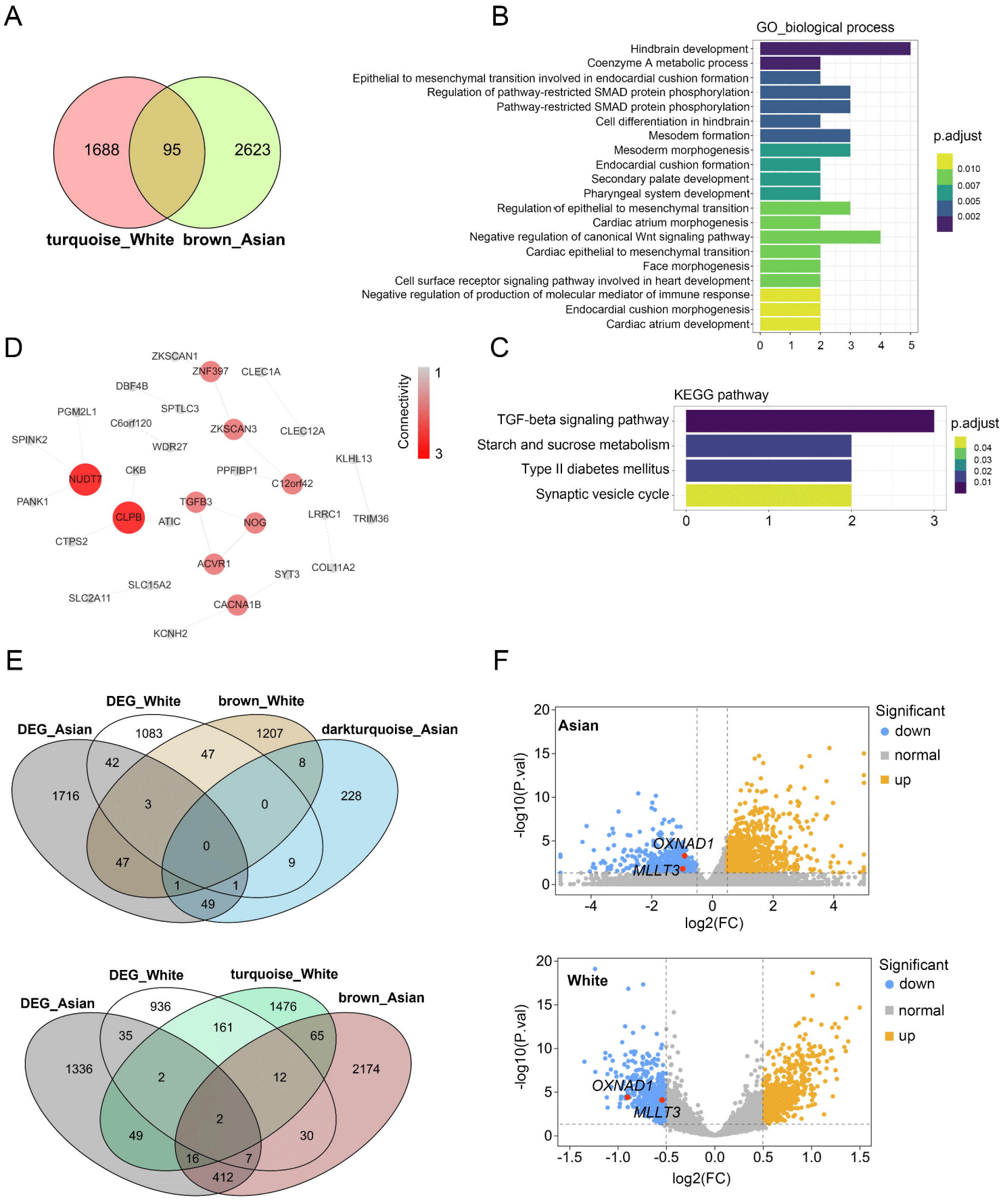


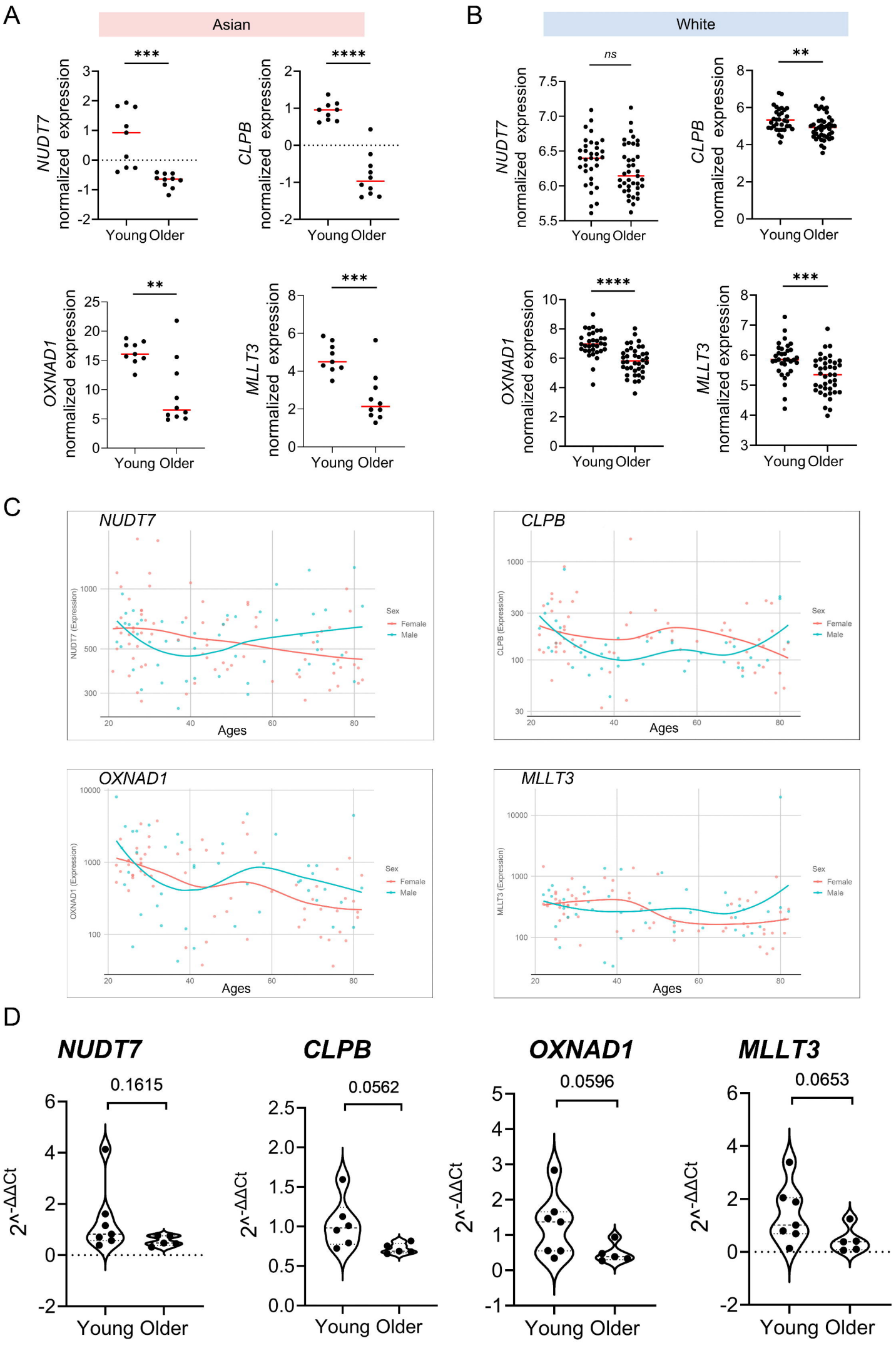









A
1688
95
2623
turquoise_White
brown_Asian
B
GO_biological process
Hindbrain development
Coenzyme A metabolic process
Epithelial to mesenchymal transition involved in endocardial cushion formation
Regulation of pathway-restricted SMAD protein phosphorylation
Pathway-restricted SMAD protein phosphorylation
Cell differentiation in hindbrain
Mesodem formation
Mesoderm morphogenesis
Endocardial cushion formation
Secondary palate development
Pharyngeal system development
Regulation of epithelial to mesenchymal transition
Cardiac atrium morphogenesis
Negative regulation of canonical Wnt signaling pathway
Cardiac epithelial to mesenchymal transition
Face morphogenesis
Cell surface receptor signaling pathway involved in heart development
Negative regulation of production of molecular mediator of immune response
Endocardial cushion morphogenesis
Cardiac atrium development
0 1 2 3 4 5
p.adjust
0.010
0.007
0.005
0.002
D
ZKSCAN1
ZNF397
CLEC1A
DBF4B
PGM2L1
SPTLC3
C6orf120
ZKSCAN3
CLEC12A
SPINK2
WDR27
KLHL13
NUDT7
CKB
PPFIBP1
C12orf42
PANK1
TGFB3
CLPB
ATIC
NOG
LRRC1
TRIM36
CTPS2
ACVR1
COL11A2
SYT3
SLC15A2
SLC2A11
CACNA1B
KCNH2
Connectivity
1
3
C
KEGG pathway
TGF-beta signaling pathway
Starch and sucrose metabolism
Type II diabetes mellitus
Synaptic vesicle cycle
0 1 2 3
p.adjust
0.04
0.03
0.02
0.01
E
DEG_White
brown_White
DEG_Asian
darkturquoise_Asian
1083
1207
42
47
8
1716
3
0
228
0
47
9
1
1
49
DEG_White
turquoise_White
DEG_Asian
brown_Asian
936
1476
35
161
65
1336
2
12
2174
2
49
30
16
7
412
F
Asian
Significant
down
normal
up
-log10(P.val)
log2(FC)
OXNAD1
MLLT3
White
Significant
down
normal
up
-log10(P.val)
log2(FC)
OXNAD1
MLLT3


A
Asian
***
****
**
***
NUDT7 normalized expression
CLPB normalized expression
OXNAD1 normalized expression
MLLT3 normalized expression
Young Older
B
White
ns
**
****
***
C
NUDT7
CLPB
OXNAD1
MLLT3
NUDT7 (Expression)
CLPB (Expression)
OXNAD1 (Expression)
MLLT3 (Expression)
Sex
Female
Male
Ages
D
NUDT7
CLPB
OXNAD1
MLLT3
0.1615
0.0562
0.0596
0.0653
2^-ΔΔCt
Young Older