The Microanatomy of Human Skin in Aging

Kyu Sang Han[1,2], Inbal Sander[3], Jacqueline Kumer[4], Eric Resnick[5], Clare Booth[6], Bartholomew Starich[1,2], Jeremy Walston[7], Ashley Kiemen[8], Sashank Reddy[2,9], Corrine Joshu[10], Joel Sunshine[3], Denis Wirtz[1,2,11], Pei-Hsun Wu[1,2]

[1] Department of Chemical & Biomolecular Engineering, Johns Hopkins University, Baltimore, MD

[2] The Johns Hopkins Institute for NanoBioTechnology, Johns Hopkins University, Baltimore, MD

[3] Department of Dermatology, Johns Hopkins University, Baltimore, MD

[4] Department of Illustration Practice, Maryland Institute College of Art, Baltimore, MD

[5] Department of Medicine, University of Maryland School of Medicine, Baltimore, MD

[6] Center for Cancer Research, National Cancer Institute, Frederick, MD

[7] Department of Medicine, Johns Hopkins School of Medicine, Baltimore, MD

[8] Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD

[9] Department of Plastic and Reconstructive Surgery, Johns Hopkins School of Medicine, Baltimore, MD

[10] Department of Epidemiology, Johns Hopkins School of Public Health, Baltimore, MD

[11] Department of Oncology, Johns Hopkins School of Medicine, Baltimore, MD

Abstract

Skin is the most visible indicator of aging, in part because it is the largest organ exposed to the outer environment. It is well recognized that age-related cumulative damage to our skin tissue degrades its structural integrity, yet how age affects the skin at the micro-anatomical and cellular scales is poorly understood. Here, we develop an Automated detection of Skin Cellular features and Microanatomy (AutoSCM) to characterize the microanatomical tissue and cellular features that change in the skin with age. We automatically label twelve distinct skin tissue components and nuclei contours on whole slide images from skin histological sections stained by hematoxylin and eosin (H&E). Then, we extract 914 objective structural features, from which we determine 108 are significantly affected by age and 806 are not. In addition to identifying decreases in epidermal and dermal thickness and hair follicle appendage density with age, we identify four previously unknown biomarkers of human aging in skin including: increased interstitial space, reduced sebaceous gland size, hair loss in back skin, and progressive horizontal alignment of the extracellular matrix and stromal cells. We further find major sex-based differences in aging skin, including increased stromal cell density by 28%, fibroblast density by 24%, epidermal thickness by 15%, and decreased epidermal cell density by 21%, sweat gland size by 23%, and 28% less content of arrector pili muscles in female compared to male. This rigorously curated and segmented atlas of normal skin microarchitecture constitutes important reference maps for future studies in age-associated diseases of the skin.

Introduction

Human skin and its appendages comprise the largest organ in the body[1]. In addition to its barrier function, the skin is critical for maintaining homeostasis, including temperature regulation, mechano-structural integrity and resilience, host pathogen defense, and wound repair[1,2]. Mechanistic studies have shown that deleterious changes to these functions stem from combinations of both intrinsic (i.e., genetic/molecular/extracellular matrix ECM)) and extrinsic (i.e., environmental, dietary) factors[3]. As humans age, nearly all skin functions are affected. These include the loss of structural integrity and resilience, impaired self-renewal capabilities, poor wound repair and scarring, inefficient temperature regulation from deficiencies in eccrine sweat glands, decreased hormone production, impaired pigmentation and susceptibility to environmental damage, increased inflammation and senescence, and a transformed cellular ecosystem [4–8] [9,10]. The mechanisms that underlie these age-related defects are broad and range from cell senescence[11], proteostasis[12–15], autophagy[16–18], mitochondrial and metabolic functions[12], stem/progenitor cell functions[19], nutrient sensing[20], intercellular communication[21], inflammation[22], stem cell exhaustion[17,23], macromolecular damage[24], and epigenetic regulation[11,18,25–27]. However, it remains poorly understood how the progressive decline in homeostatic and regenerative capacities of the aging skin is associated with changes in tissue micro-architecture and cellular composition.

Here, we developed a deep-learning workflow to segment tissue compartments and nuclei in H&E images of FFPE archived skin specimen for quantitative microanatomy of human skin. We examine how changes in morphological features are related to normal human skin aging. We examined the age-related structural features of skin with a thorough analysis of total of 7.8 million nuclei and 12 types of skin tissue compartments from 17,673 mm^2 area of tissue collected from 99 donors. This analysis allowed the identification of morphometric markers of human aging in skin. We then explore the utility of skin morphometric features as biomarkers of aging.

Results

Overall workflow of human skin tissue analysis

To explore age-related changes in non-diseased skin, we obtained 379 formalin-fixed, paraffin-embedded (FFPE) skin biopsy specimens from patients at Johns Hopkins Hospital collected between 2015 and 2021. These samples comprised histologically normal adjacent tissues from excision specimen from patients diagnosed with various conditions such as melanoma, seborrheic keratosis, epidermal cysts, fibrosis, basal cell carcinoma, and squamous cell carcinoma. A board-certified dermatopathologist reviewed the H&E slides of these tissue specimen and identified 167 patient cases with substantial healthy margin, defined as absent any sign of the diagnosed disease or other damage such as solar elastosis, excessive inflammation, or cicatrix (**Figure 1a and Supplementary Figure 1).** Among these, the majority (N=130) were skin tissue blocks from the back, with the remainder from the abdomen, chest, head, neck, arms, and legs. Out of the 130 back tissues, 99 originated from Caucasian donors. 56 are male, and 43 are female. The samples are split into three age groups: young (age<30), middle-aged (30≤age≤60), and old (60<age). Of 99 donors, 17 are young, 41 are middle-aged, and 41 are old. (**Supplementary Table 1**)

To comprehensively analyze the micro-architecture of skin and its constituent cells, we spatially mapped the architecture of the skin's tissue and nuclear components in these tissue specimens. We identified twelve primary microanatomical compartments distinguishable in H&E images: 1. Stratum corneum, 2. Stratum spinosum, 3. Hair shaft, 4. Hair follicle, 5. Smooth muscle, 6. Sebaceous gland, 7. Eccrine coils (sweat gland), 8. Nerve, 9. Vasculature, 10. Extracellular matrix (ECM), 11. Fat, and 12. Interstitial space **(Figure 1b** and **Supplementary Figure 2**). Here, we established a deep learning-based workflow that precisely delineated these tissue microanatomical compartments (i.e. tissue compartment labeling) and the nuclei of the cells contained in these compartments. This was achieved by combining the DeepLabV3+ tissue labeling pipeline[28] with HoVerNet for nuclei segmentation[29] (**Figure 1b**). To train and validate our deep learning model for tissue compartment labeling, we performed manual annotations on 50 instances of each tissue microanatomical compartment in 20 skin samples. Out of these, 15 samples were allocated for training, and the model's performance was assessed on the remaining 5 unseen tissue images. The trained model achieved a precision of 96.6% and an average recall of 93.9% across all tissue types (**Supplementary Figure 3**).

To segment the nuclei in skin tissues, we first tested the performance of the pretrained HoVerNet model trained on CoNSeP pancreatic tissue images[29]. The pretrained model achieved a DICE

score of 0.66, DICE is defined as $\frac{2 \times Area\ of\ overlap}{Area\ of\ union}$ when the predicted nuclei is compared to the manually annotated nuclei. To improve nuclei segmentation, we developed a semi-supervised learning workflow (**Supplementary Figure 4**) based on the concept described by Chapelle *et al*.[30] that did not require additional manual annotations or curation to re-train the model on the skin tissue images. Our results showed that this semi-supervised model substantially improved nuclei segmentation on skin with a DICE score of 0.73 (**Supplementary Figure 4**).

Tissue and cell morphometric analysis

To quantitatively assess the association between tissue micro-architecture and cellular content within these skin tissues, we developed a panel consisting of 914 morphological features categorized into five groups: 1) tissue composition, 2) tissue and cell morphology, 3) dermal ECM structure, 4) epidermal and dermal lamellar structure, and 5) spatial distributions of tissues/cells (**Figure 1c**). The detailed features list is found in **Supplementary Table 2**. Of note, our tissue analysis focuses on the epidermis and dermis, and does not include the subcutaneous tissue.

Qualitative assessment of the skin images readily showed changes in skin tissue architecture, such as dermal thinning and epidermal flattening, indicating that changes in tissue composition and architecture are associated with aging (**Figure 2a-b**). Quantitative analysis of skin tissue composition, based on the deep learning labeled tissue component map, indicated that the average human back skin was composed of approximately 2.6% epidermis, 1.5% pilosebaceous unit (including the hair shaft, hair follicle, sebaceous gland, and arrector pili muscle), 0.6% sweat gland, 1.3% neurovascular bundle (consisting of nerve and blood vessel), 4.6% intradermal fat, and 89.3% ECM and interstitial space (**Supplementary Figure 5**).

To identify potential aging-associated tissue changes in the microarchitecture of skin, we used a ranked-based Spearman correlation (r) of the 914 microarchitectural features with aging across all tissue samples (**Figure 2b**). In addition, we measured the Cohen's distance (*d*), which describes the difference between two sample populations: samples from patients < 30 years and > 60 years old. We found that 108 features were significantly associated with aging with $|r| > 0.3$, P value < 0.05, and $d > 1$ (**Figure 2c, Supplementary Table 3**). Among these 108 features, we identified several features previously known aging-associated skin microarchitectural changes, such as epidermal thinning[31–35] and loss of hair follicles and associated appendageal structures[36,37], which validated our morphometrics analysis (**Figure 2d**). This approach allowed for the precise quantitation of those changes. Epidermal thickness decreased from 117μm to

100µm from young (less than 30 years old) to old (more than 60 years old) age groups. Across the tissue samples of all ages, this metric had spearman's rho of -0.36 and *P* value of 2.4×10^{-4}. The average reduction of epidermal thickness is 0.28 microns per year. Additionally, our results showed a significant decrease in the density of hair follicles, decreasing from 0.4% in young patient samples to 0.1% in old patient samples (Spearman's rho -0.55, *P* value 5.11×10^{-9}), and sebaceous glands, which decrease from 1.0% in young patient samples to 0.3% in old patient samples (Spearman's rho -0.42, *P* value 1.3×10^{-5}).

Beside these highly correlated features, we identified several features that are previously believed as aging-associated skin microarchitectural features; however, the changes in these 806 unselected features are not significant. For example, dermal thickness decreased with age from 5.0mm to 4.8mm (Spearman's rho -0.16, *P* value 0.10). The full list of unselected features is found in **Supplementary Table 3**.

We also found several novel aging biomarkers of the skin, such as increased interstitial space, reduced sebaceous gland size, reduced hair density in back skin, and increased alignment of stromal cells to ECM components (**Figure 2e**). We performed unsupervised hierarchical clustering analysis to group these 108 aging-associated features (**Figure 2f**). This analysis revealed that these features could be further grouped into ten distinct core aging processes, including: 1). Reduced stromal cell density and nuclei size, 2). Interstitial space alignment, 3). Hair and sebaceous gland shrinking, 4a). Hair shortening, 4b). Narrower sebaceous gland, 4c). Shorter dermal white adipose tissue, 4d). stromal cell alignment, 4e). Shorter smooth muscle, 4f). Epidermal thinning, 5). Smaller immune cells, 6). Stromal cell local alignment, 7). Hair loss, 8). Reduced inflammatory cell density, 9). Elongated fibroblasts, 10a). Extracellular matrix alignment, and 10b). Interstitial space widening (**Supplementary Table 4**). The analysis showed that core 3 (Hair and sebaceous gland shrinking) and core 4 (Shorter smooth muscle and stromal cell alignment) had the largest numbers of features, comprising 25 and 28 features, respectively (**Figure 2g**).

Skin tissue microarchitecture as biomarker of aging

To assess whether the above aging-related morphometric features within skin tissue might serve as predictive biomarkers of chronological age, we conducted an analysis on the mean absolute error (MAE) associated with each individual feature predict age. This was carried out using leave-one-out validation combined with univariate generalized linear regression (GLM) (**Figure 3a**).

Overall, the MAE from these individual features ranged from 13 to 29 years with an average of 17 years and a standard deviation of 0.7 years (**Figure 3a and Supplementary Table 5**). The ten best predictors are listed in **Supplementary Table 5**. Among all aging-associated features, the reduction in blood vessel radius was the most accurate aging predictor with MAE of 13.1 years (**Figure 3b**). This result may explain the reduced blood flow observed with aging[38]. Comparing the ten distinct core aging processes, we found that features from core 2 (Interstitial space alignment) and core 3 (Hair and sebaceous gland shrinking) had better univariate prediction with their univariate MAE of 14.9 years and 14.6 years, which were significantly lower than the MAE of other cores (**Figure 3c, Supplementary Table 4**).

To explore the connection between features, we primarily evaluated the prediction power of bivariate pairs among the aging-associated features using the bivariate generalized linear regression (GLM). Our results showed that the MAE derived from a bivariate model was generally lower than the univariate MAE (**Figure 3d**). The best predictive bivariate pair was ECM alignment and hair follicle composition with an MAE of 11.7 years (**Figure 3e**). Among the top 100 most predictive bivariate pairs, 41 were pairs between core 3 (Hair and sebaceous gland shrinking) and core 10 (ECM matrix alignment and Interstitial space widening) (**Figure 3f, Supplementary Table 6**).

Finally, we explored the age predictive power of these features using a multivariate model. Principal component analysis (PCA) was used to reduce the 108 features into 29 principal components (PCs), capturing 95% variance within the data set. The reduced PCs space readily exhibited a strong association with age (**Figure 3g**). We evaluated the predictive power of the reduced feature space with various machine learning (ML) algorithms including the generalized regression model (GLM), support vector machines (SVMs), Lasso (least absolute shrinkage and selection operator), which drops some features by penalizing coefficients and driving them to zero (L2 penalty). The results showed that the SVMs displayed the most accurate prediction with an MAE value of 7.7 years compared to 9.3 years from GLM, and 8.9 years from Lasso. (**Figure 3h**, **Supplementary Table 7**). Overall, this analysis indicates that morphometric features of skin tissue are quantitatively linked to the skin aging process and could be further developed as biomarkers for aging.

No significant effect of biological sex on aging in skin tissue

Next, we investigated whether the predictive power of the aging-associated morphometric features of skin was associated with sex. We found that the age predictive accuracy from the multivariate model was comparable between males and females, with an average MAE of 8.45 years for males and 7.76 years for females, respectively (**Figure 4a**). Analysis of sex-discrimination capability among 108 aging features, revealed seven features which had absolute effect size difference larger than 10% and P-value lower than 0.1 from two-tailed t-test (**Figure 4b** and **Supplementary Table 3**). Among the seven sex discriminant features, three were tissue/cell spatial distribution feature, three were dermal ECM fine structure features, and one was a tissue shape feature.

We ranked the aging biomarkers based on the GLM accuracy in classifying the sex (**Figure 4c**). Stromal cell density was the top feature that differentiated biological sex with a classification accuracy of 66%, where samples from females had higher stromal cell density (169 N/mm^2) than observed in samples from males (132 N/mm^2) with a P-value of 0.78×10^{-5} (**Figure 4d**). Sex classification accuracy narrowly increased to 67% when using the multivariate GLM using the principal components reduced from the seven predictive features (**Figure 4e**). Overall, these results indicate that sex is not strongly associated with skin aging features. This observation is further supported by the minimal difference in aging predicting accuracy when incorporating sex as an additional feature to establish a multivariate aging prediction model (**Figure 4f**).

Discussion

Using a deep learning-based framework, we present a rigorously curated and segmented atlas of normal skin microarchitecture at both tissue and cell levels from FFPE skin tissue sections. With the fully segmented skin tissue atlas, we captured the comprehensive tissue, cell morphometrics and associated spatial distributions. We identified ten core processes that occur in aging skin tissue. Among identified age-associated architecture changes, several features were consistent with previous studies, such as epidermis thinning[42–44], hair loss[4536], decreased collagen content[46], and collagen alignment and bundling[47]. Our study highlights five new features of skin aging: 1) fewer and smaller pilosebaceous units; 2) narrower blood vessels; 3) increased distance between the blood vessel and epidermis; 4) Aligned ECM with epidermis; and 5) expanded interstitial space.

These findings suggest that the dominant model of back skin aging is progressive aligned fibrosis of the dermis and subsequent loss of supporting tissue structures. Fibrosis drives an increasingly aligned and shift towards a rigid dermal ECM during biological aging. A direct result of these architectural changes is a flattening and therefore thinning of the epidermis, squeezing of interstitial cells and structures (i.e., hair follicles, sebaceous glands, blood vessels and dermal adipose) and a gradual alignment of stromal cells as they squeeze into rigid spaces (**Figure 5**). Of 914 morphological features, we identified decreased blood vessel radius (**Figure 3b**) as the top predictor of skin age. The physical decrease in vessel radius highlights the substantial transformation vasculature architecture undergoes with age. Reduced vessel radius likely contributes to previously identified age-associated decrease in blood circulation (REFS) Interestingly, we also observed that blood vessels are locate farther away from the epidermis with age (**Supplementary Figure 6**). Blood vessels supply oxygen and nutrients while providing circulatory routes for immune cells[48]. A physically deeper vessel translates to reduced nutrient availability at distal tissue sites. Dermal fibrosis combined with smaller blood vessels may facilitate the observed epidermal thinning (**Figure 2d**), reduced ECM deposition, reduced stromal cell density (**Figure 4d**), and increased interstitial space with age.

The identification of valid, reliable, and accurate markers of human aging in humans would greatly facilitate the ability of observational studies to assess the influence of these mechanisms on a variety of aging outcomes. Aging research has been primarily focused on the progressive changes that occur at either the molecular scale or the tissue scale where dysfunctional, cosmetic changes are observed[39,40]. A recent study of epigenetics of skin aging has revealed that epigenetics clocks are tightly associated with aging. Epigenetic clocks can predict age with

the mean absolute error (MAE) of 1.94 years[41]. Epigenetic clocks are accurate, yet they lack biophysical representation of the skin aging processes.

While all skin follows the same general architecture, significant variations occur throughout the body[50–53]. This study focused on evaluating samples from the human back, a large area of the body that suffers relatively less from body location variance[53]. However, other areas of the body show significantly different histologic features; scalp, for example, has an abundance of hair follicles and other dermal appendages[54,55], whereas volar skin of the palms and soles of the feet has no hair, but presents a thickened epidermis and stratum corneum to withstand pressure[56]. Future directions include comparing the unique morphometric features of aging back skin with body sites. A combination of morphometric feature analysis with spatially resolved proteomics or genomic approaches may provide significant insight on molecular drivers of age-related biophysical changes.

Methods

Sample acquisition.

Excisional skin specimens with substantial margin of normal tissue around the lesions were identified and collected from the archives at the Johns Hopkins Hospital. We compiled a cohort of patients whose skin samples were acquired as part of clinical routine from year 2015 to 2019 with underlying disease, such as melanoma, keratosis, epidermal cyst, fibrosis, basal cell carcinoma, and squamous cell carcinoma. This cohort was compiled based on the volume of sample, evaluation on clinical notes whether the margin is free of disease, and their age distributions. The H&E stained tissue section of archive skin tissue samples from the cohorts were than obtained from the archive and imaged. The images of H&E stained sections were than evaluated by a board-certified dermatopathologist to identify the sections with grossly normal skin margins with free of solar elastosis, inflammation, cicatrix, and the primary underlying disease.

Each slide contains between 1 and 6 tissue sections, depending on the size of the biopsy. Larger biopsies are split and placed into a same tissue cassette to form a formalin-fixed paraffin embedded (FFPE) block. All tissue sections on a slide are evaluated, and only the sections free of diseases are used for further analysis. Out of 295 patient samples, 165 samples are selected. Among these samples, the majority (N=130) of tissue blocks are from the skin from the back, and the rest are from the abdomen, chest, head, neck, arms, and legs. Out of 130 back tissues, 101 are from Caucasian patients.

Image acquisition

Each selected H&E tissue slide is scanned at 20x magnification using a Hamamatsu Nanozoomer S210 scanner with a resolution of 0.454 mm per pixel. Nine focus points are used per tissue section.

Tissue semantic segmentation

To segment skin tissue objects (e.g., glands, smooth muscles, etc.), we used a DeepLabV3+ CNN architecture[60] pretrained on ImageNet[61]. To train the CNN for tissue labeling, we first establish the training data set. We randomly selected 20 out of the above 165 tissue images and manually annotated them using Aperio ImageScope. We labeled twelve tissue components in the skin including: 1. stratum corneum, 2. stratum spinosum, 3. hair shaft, 4. hair follicle, 5. smooth muscle, 6. sebaceous gland, 7. sweat gland, 8. Nerve, 9. blood vessel, 10. ECM, 11. fat, and 12. interstitial space. The annotated tissue components are validated by a board-certified dermal pathologist. The DeepLabV3+ model was trained, cross-validated, and applied to all 165 whole slide images for pixel-wise classification of these 12 tissue components. Our skin tissue segmentation is validated using 5 independent, randomly selected samples. The model's precision is 96.61%, and its recall is 93.94% (**Supplementary Figure 3**).

Rotational alignment of skin tissue

The epidermis was labeled by a deep learning model. The labeled epidermal is converted to 2D point clouds[57,58] (**Supplementary Figure 6**). The Y coordinate of the point clouds is subtracted by their means in Y-axis direction to be flattened along the horizontal X-axis. These flattened point clouds are the target for the alignment. Next, we compute the rotational angle needed to minimize the error between the original point clouds and the target using singular value decomposition. This type of problem is called an orthogonal Procrustes problem[59]. Once the rotational angle is found, we can use it to rotate each skin tissue section so that the horizontal axis of the resulting image is aligned with the skin surface. Detailed mathematical method is the following. Given that original point cloud $Q = \{q_1, \dots, q_N\}$ and target cloud $P = \{p_1, \dots, p_N\}$ has correspondences $C = \{(i, j)\}$, the error between the two are defined as $E(R, t) = \sum_{(i,j)\in C} \left\| q_i - Rp_j - t \right\|^2$. Singular value decomposition (SVD) is used to decompose the cross-covariance matrix $W = \sum_{(i,j)\in C} q_i{}' p_j{}'^T$ to

$W = \mathrm{UDV}^T$, providing the values of U, D, V^T. Then, the rotational matrix can be found as $R = \mathrm{UV}^T$. This matrix is 2x2 dimensions with the following elements, $R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Finally, we take arctan of $\frac{\sin\theta}{\cos\theta}$ from the rotational matrix to compute the rotational angle. When the transpose of the rotational matrix has negative determinant, the matrix is in a reflection position, resulting in a negative value of rotational angle. To compute the correct angle, we multiply -1 to the rotational angle to flip its sign.

Tissue composition and* architecture *quantification

The fully labeled skin tissue images allowed us to assess the composition of tissue subtypes and their respective organization. The composition is calculated by counting the number of pixels occupied by each tissue component and dividing it by the total number of pixels in the whole tissue section. The tissue architectural features such as thickness, waviness, roughness, etc. A detailed list of tissue architecture features can be found in the **supplementary table 2**.

Nuclei detection

A deep neural network model was applied to detect hematoxylin-stained nuclei. HoVerNets, a convolutional neural network for nuclear segmentation in H&E[63], was trained using several images selected arbitrarily from each group of tissue component for a total of N samples. Detailed workflow for training the model is illustrated in **Supplementary Figure 3**. Our implementation of HoVerNet takes a different approach from the original workflow in training the model to associate the Human Skin H&E images with nuclear regions. After training for 100 epochs, the HoVerNet was applied to detect nuclei for all 165 whole slide images, producing output images of predicted nuclei regions along with the coordinates of nuclei contour. Each tissue section size from 15mm^2 ~ 35mm^2 generated 5000~12000 detected nuclei. From the entire Skin tissue dataset, we detected 2.1 million nuclei. To determine the cell type of the nuclei, we use the location of nuclei centroid with regards to the tissue component label to identify which tissue component is a nucleus reside in. For instance, a nucleus detected within sweat gland region is almost certainly from a sweat gland cell. The spatial correlation between neighboring nuclei is quantified using metrics such as distance between neighboring cells using K-nearest neighbor algorithm (KNN), alignment between adject cells, and local nuclei population density within a tissue component object.

Nuclear morphology

The nuclei contours are used to quantify morphological features of the nuclei, including area, circularity, aspect ratio, orientation, and extent. The area A is the total number of pixels within the contour and then converted to micrometer squared given that width and height of pixel are both 0.454µm. Perimeter P is defined by the length of contour line. The circularity is computed using the equation: $\frac{4\pi A}{P^2}$. The aspect ratio is defined as the major axis length divided by minor axis length. The orientation is quantified by a conic fitting algorithm described by Fitzgibbon et al[64]. The extent is defined as area of contour divided by area of bounding box. A detailed list of nuclear morphology features can be found in the **supplementary table 2**.

Software

The following software tools were used for the predictor and related analysis: python 3.8.10, tensorflow-gpu 2.4.1, tensorflow-probability 0.11.1, scipy 1.6.2, statsodels 0.12.0, scikit-learn 0.23.2, pandas 1.1.1, seaborn 0.11.1 and numpy 1.21.6, R 4.0.2, MATLAB R2024a.

Statistics and reproducibility

A correlation between a sample's age and a morphometric feature was made using spearman rank coefficient. Linear regression methods were evaluated with P value statistics.

Figure captions

Fig. 1: Overall workflow of human skin tissue analysis a. Skin excisional sample typically with an elliptical shape is cut into multiple pieces before embedding into a paraffin block. The skin pieces with abnormal tissue architecture (such as lesion, sun damage, …etc.) are excluded for further analysis through examining the image of H&E-stained tissue section. **b**. Whole slide H&E images are labeled with tissue components automatically using a trained CNN with DeepLab V3+ framework. Cell nuclei within each labeled tissue component are segmented using HoVerNet. Scale bar, 1mm in tissue, 100µm and 20µm in zoom-in **c**. Five groups of morphometric features with a total of 914 features are established to characterize the tissue and cellular architecture of skin. Tissue scale bar, 100µm. Cell scale bar, 20µm. H&E scale bar, 50µm.

Fig. 2: Core aging processes through extensive morphometric analysis. a. Representative image of H&E-stained tissue section from donors and at young, middle age and old (top). The corresponding tissue component maps labeled by the trained CNN (bottom). **b.** Histogram of age distribution of the sample cohort (56 male and 43 female) obtained from the back of skin. **c.** Age associated features are selected based on the spearman correlation coefficient with aging and Cohen's d from comparing young (30 years old or younger) vs old (60 years old or older). A scatter plot showing each features spear correlation coefficient and Cohen's d. A total of 108 features with Cohen's d >1 and |spearman's rho| >0.3 are identified as aging-associated features. **d.** Representative H&E images of skin and scatter plots showing the hair follicle content and epidermis content are among the identified aging-associated features. Hair follicle content and epidermis content are features known to be correlated with age. Scale bar, 1mm. **e.** Representative H&E images of skin and scatter plots show the increment of interstitial gap and shrinkage of the sebaceous gland, which are novel features of skin aging, among others. Scale bars, 100µm. Wide interstitial gap ratio is measure by ratio of the interstitial gap exceeding 4µm thickness among total by length. **f-g.** Hierarch clustering analysis groups the aging-associated features into ten distinct cores of the aging process using pair-wise correlation distance (**f**). Summary of general characteristics of each core based on the morphometric features of skin in each core (**g**). Cores 4 and 10 are further divided into subgroups to describe distinct characteristics of the assigned features better.

Fig. 3: Morphometric features are predictive of chronological age. a. A scatter plot shows correlation coefficient with aging and mean absolute error of aging prediction for each individual aging associated feature. Mean absolute error is calculated from leave-one-out validation using univariate generalized linear regression model (GLM). **b.** The blood vessel radius is the best predictive feature with MAE of 13.1 years. Two representative H&E images from a young and an old donor are shown. Scale bar = 20µm. A scatter plot shows a negative correlation between blood vessel radius from each sample and their corresponding ages. Predicted age with blood vessel radius from leave-one-out validation and actual chronological ages for each sample is shown as a scatter plot. **c.** Violin plot shows MAE distribution for features in each core of aging process. Features from core 2 (interstitial space alignment) and core 3 (Hair and sebaceous gland shrinking) show the overall best performance in aging prediction as single variable among others. **d-f.** Aging prediction uses bivariate features. A scatter plot shows the MAEs from all pairs (N=5,778) of features using bivariate aging prediction with GLM compared to the average MAE of each pair with univariate prediction with GLM. The MAE is calculated using the leave-one-out validation. In most pairs, bivariate MAE is lower than the mean univariate MAE (**d**). The best predictive bivariate pair is ECM alignment and hair follicle composition with an MAE of 11.7 years. Representative H&E images of skin in young and old highlight decrement in ECM alignment and loss of hair follicles with aging. Scale bar, 100µm for ECM and 1mm for hair follicle composition. A 3d scatter plot shows the association of these two features with aging among all samples (**e**). Analysis of the occurrence of predictive paired features associated with cores of the aging process. In the top 100 most predictive bivariate pairs, the most frequently occurred pairs are between core 3 (hair and sebaceous gland shrinking) and core 10 (ECM alignment and Interstitial space widening) (**f**). **g.** The heatmap shows 99 tissue specimens (ranked in order by age) and the corresponding 108 age-associated features. Principal component analysis shows the first 29 principal components contain 95% of the total variance. A scatter plot shows the distribution of the sample along the first two principal components (PCs) and readily shows the chronological age of samples is associated with PC1 and PC2. **h,** Implementation of multivariate models with first 29 PCs for aging prediction. A violin plot shows the MAE of each sample from leave-one-out validation using GLM, Lasso, and SVM. A scatter plot shows predicted age (leave-one-out validation) vs chronological age using SVM in each sample. SVM achieved the lowest MAE of 7.7 years.

Fig. 4: Skin aging features are not strongly associated with biological sex. a. A scatter plot of predicted ages and chronological ages in males and females. Predicted age is derived from leave-one-out using the multivariate SVM. MAE is 7.3 years among female samples compared to 8.1 years among male samples. **b.** A scatter plot illustrating the association between gender and 108 aging features with relative differences in mean and *p*-value (two-way t-test) between genders. Seven features have a relative difference of 10%+ with a p-value less than 0.1. **c.** Summary of the gender-associated features and their corresponding accuracy in classifying gender. **d.** A scatter plot shows the relation between stromal cell density and the age of donors in both genders. Violin plots comparing stromal cell density between genders in both young (30 years old or younger) and old (60 years old or older) show that stromal cell density is significantly higher in females than in males in both age ranges. **e.** The accuracy of gender classification is based on an increasing number of principal components derived from gender-associated features. The highest accuracy, reaching 68%, is attained using the first 4 principal components (PCs). **f.** Incorporating gender information into the aging prediction using the multivariate SVM model as depicted in Fig. 3h, a scatter plot illustrates the predicted age versus chronological age for donors of both genders. The Mean Absolute Error (MAE) is 7.8 years when gender information is considered in the prediction model. The scatter plot further depicts the difference in prediction errors for each sample, comparing models with and without considering gender. On average, there is a +0.8 years change in males and a -0.5 years change in females. The overall MAE increases by 0.1 year with the inclusion of gender information when compared to the MAEs obtained without considering sample gender information.

Fig. 5: the effect of aging on the microanatomy of human skin a. illustration of distinct macroscopic differences between young (age 10~30) and old (age 60~90) human skin, labeled with the corresponding number in the following list: 1 and 2. flatter and thinner epidermis, 3. thinner dermis, 4. fewer pilosebaceous units (a unit includes hair shaft, hair follicle, sebaceous gland, and arrector pili muscle), 5. smaller pilosebaceous units, 6. narrower blood vessels, 7. blood vessel distant from epidermis, 8. ECM more aligned with epidermis, 9. thicker ECM bundles, 10. wider interstitial space

References

1. Yanez, D. A., Lacher, R. K., Vidyarthi, A. & Colegio, O. R. The Role of Macrophages in Skin Homeostasis. *Pflugers Arch* **469**, 455 (2017).

2. Rittié, L. & Fisher, G. J. Natural and Sun-Induced Aging of Human Skin. *Cold Spring Harb Perspect Med* **5**, (2015).

3. Farage, M. A., Miller, K. W., Elsner, P. & Maibach, H. I. Intrinsic and extrinsic factors in skin ageing: a review. *Int J Cosmet Sci* **30**, 87–95 (2008).

4. Phillip, J. M., Aifuwa, I., Walston, J. & Wirtz, D. The Mechanobiology of Aging. *Annu Rev Biomed Eng* **17**, 113 (2015).

5. Zouboulis, C. C., Adjaye, J., Akamatsu, H., Moe-Behrens, G. & Niemann, C. Human skin stem cells and the ageing process. *Exp Gerontol* **43**, 986–997 (2008).

6. Fuchs, E. Scratching the surface of skin development. *Nature* **445**, 834 (2007).

7. Blanpain, C. & Fuchs, E. Epidermal homeostasis: a balancing act of stem cells in the skin. *Nat Rev Mol Cell Biol* **10**, 207 (2009).

8. Wang, A. S. & Dreesen, O. Biomarkers of Cellular Senescence and Skin Aging. *Front Genet* **9**, 247 (2018).

9. Solé-Boldo, L. *et al.* Single-cell transcriptomes of the human skin reveal age-related loss of fibroblast priming. *Commun Biol* **3**, (2020).

10. Farage, M. A., Miller, K. W., Elsner, P. & Maibach, H. I. Characteristics of the Aging Skin. *Adv Wound Care (New Rochelle)* **2**, 5 (2013).

11. Zhu, X. *et al.* Inflammation, epigenetics, and metabolism converge to cell senescence and ageing: the regulation and intervention. *Signal Transduction and Targeted Therapy 2021 6:1* **6**, 1–29 (2021).

12. Jensen, M. B. & Jasper, H. Mitochondrial Proteostasis in the Control of Aging and Longevity. *Cell Metab* **20**, 214–225 (2014).

13. Labbadia, J. & Morimoto, R. I. The Biology of Proteostasis in Aging and Disease. *http://dx.doi.org/10.1146/annurev-biochem-060614-033955* **84**, 435–464 (2015).

14. Steffen, K. K. & Dillin, A. Cell Metabolism Review A Ribosomal Perspective on Proteostasis and Aging. *Cell Metab* **23**, 1004–1012 (2016).

15. Kaushik, S. & Cuervo, A. M. Proteostasis and aging. *Nature Medicine 2015 21:12* **21**, 1406–1415 (2015).

16. Barbosa, M. C., Grosso, R. A. & Fader, C. M. Hallmarks of aging: An autophagic perspective. *Front Endocrinol (Lausanne)* **10**, 790 (2019).

17. Revuelta, M. & Matheu, A. Autophagy in stem cell aging. *Aging Cell* **16**, 912–915 (2017).

18. Lapierre, L. R., Kumsta, C., Sandri, M., Ballabio, A. & Hansen, M. Transcriptional and epigenetic regulation of autophagy in aging. *http://dx.doi.org/10.1080/15548627.2015.1034410* **11**, 867–880 (2015).

19. Zouboulis, C. C., Adjaye, J., Akamatsu, H., Moe-Behrens, G. & Niemann, C. Human skin stem cells and the ageing process. *Exp Gerontol* **43**, 986–997 (2008).

20. Kapahi, P. *et al.* With TOR less is more: a key role for the conserved nutrient sensing TOR pathway in aging. *Cell Metab* **11**, 453 (2010).

21. Marar, C., Starich, B. & Wirtz, D. Extracellular vesicles in immunomodulation and tumor progression. *Nature Immunology 2021 22:5* **22**, 560–570 (2021).

22. Olivieri, F. *et al.* MicroRNAs linking inflamm-aging, cellular senescence and cancer. *Ageing Res Rev* **12**, 1056–1068 (2013).

23. Oh, J., Lee, Y. D. & Wagers, A. J. Stem cell aging: mechanisms, regulators and therapeutic opportunities. *Nature Medicine 2014 20:8* **20**, 870–880 (2014).

24. Rattan, S. I. S. Increased molecular damage and heterogeneity as the basis of aging. *Biol Chem* **389**, 267–272 (2008).

25. Schmeer, C., Kretz, A., Wengerodt, D., Stojiljkovic, M. & Witte, O. W. Dissecting Aging and Senescence-Current Concepts and Open Lessons. *Cells* **8**, (2019).

26. Pollina, E. A. & Brunet, A. Epigenetic regulation of aging stem cells. *Oncogene 2011 30:28* **30**, 3105–3126 (2011).

27. Grönniger, E. *et al.* Aging and Chronic Sun Exposure Cause Distinct Epigenetic Changes in Human Skin. *PLoS Genet* **6**, e1000971 (2010).

28. Chen, L.-C. *et al.* DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs.

29. Graham, S. *et al.* HoVer-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images. *Med Image Anal* **58**, (2018).

30. Semi-Supervised Learning (Chapelle, O. et al., Eds.; 2006) [Book reviews] | IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/abstract/document/4787647.

31. Sun, C.-K. *et al.* Determination of chronological aging parameters in epidermal keratinocytes by in vivo harmonic generation microscopy. *Biomedical Optics Express, Vol. 4, Issue 1, pp. 77-88* **4**, 77–88 (2013).

32. Koehler, M. J. *et al.* In vivo measurement of the human epidermal thickness in different localizations by multiphoton laser tomography. *Skin Research and Technology* **16**, 259–264 (2010).

33. Longo, C., Casari, A., Beretti, F., Cesinaro, A. M. & Pellacani, G. Skin aging: In vivo microscopic assessment of epidermal and dermal changes by means of confocal microscopy. *J Am Acad Dermatol* **68**, e73–e82 (2013).

34. Sandby-Møller, J. & Wulf, H. C. Epidermal Thickness at Different Body Sites: Relationship to Age, Gender, Pigmentation, Blood Content, Skin Type and Smoking Habits. (2003) doi:10.1080/00015550310015419.

35. Gambichler, T., Matip, R., Moussa, G., Altmeyer, P. & Hoffmann, K. In vivo data of epidermal thickness evaluated by optical coherence tomography: Effects of age, gender, skin type, and anatomic site. *J Dermatol Sci* **44**, 145–152 (2006).

36. Trueb, R. M. Aging of hair. *J Cosmet Dermatol* **4**, 60–72 (2005).

37. Kamenisch, Y. & Berneburg, M. Progeroid syndromes and UV-induced oxidative DNA damage. *J Investig Dermatol Symp Proc* **14**, 8–14 (2009).

38. Waller, J. M. & Maibach, H. I. Age and skin structure and function, a quantitative approach (I): blood flow, pH, thickness, and ultrasound echogenicity. *Skin Research and Technology* **11**, 221–235 (2005).

39. Belsky, D. W. *et al.* Quantification of biological aging in young adults. doi:10.1073/pnas.1506264112.

40. Xue, Y. *et al.* Mechanical tension mobilizes Lgr6 + epidermal stem cells to drive skin growth. *Sci. Adv* **8**, 8698 (2022).

41. Boroni, M. *et al.* Highly accurate skin-specific methylome analysis algorithm as a platform to screen and validate therapeutics for healthy aging. *Clin Epigenetics* **12**, 1–16 (2020).

42. Gambichler, T., Matip, R., Moussa, G., Altmeyer, P. & Hoffmann, K. In vivo data of epidermal thickness evaluated by optical coherence tomography: effects of age, gender, skin type, and anatomic site. *J Dermatol Sci* **44**, 145–152 (2006).

43. Gambichler, T. *et al.* In vivo optical coherence tomography of basal cell carcinoma. *J Dermatol Sci* **45**, 167–173 (2007).

44. Longo, C., Casari, A., Beretti, F., Cesinaro, A. M. & Pellacani, G. Skin aging: in vivo microscopic assessment of epidermal and dermal changes by means of confocal microscopy. *J Am Acad Dermatol* **68**, (2013).

45. Shin, W. *et al.* Dysfunction of Hair Follicle Mesenchymal Progenitors Contributes to Age-Associated Hair Loss. *Dev Cell* **53**, 185-198.e7 (2020).

46. Wu, S. *et al.* Quantitative analysis on collagen morphology in aging skin based on multiphoton microscopy. *J Biomed Opt* **16**, 040502 (2011).

47. Wu, S. *et al.* Quantitative analysis on collagen morphology in aging skin based on multiphoton microscopy. *https://doi.org/10.1117/1.3565439* **16**, 040502 (2011).

48. Alberts, B. *et al.* Epidermis and Its Renewal by Stem Cells. (2002).

49. Tsuchida, Y. The effect of aging and arteriosclerosis on human skin blood flow. *J Dermatol Sci* **5**, 175–181 (1993).

50. Tagami, H. Location-related differences in structure and function of the stratum corneum with special emphasis on those of the facial skin. *Int J Cosmet Sci* **30**, 413–434 (2008).

51. Kleesz, P. *et al.* Full-Body Skin Mapping for Six Biophysical Parameters: Baseline Values at 16 Anatomical Sites in 125 Human Subjects. *Skin Pharmacol Physiol* **25**, 25–33 (2011).

52. Firooz, A. *et al.* Variation of biophysical parameters of the skin with age, gender, and body region. *The Scientific World Journal* **2012**, (2012).

53. Smalls, L. K., Randall Wickett, R. & Visscher, M. O. Effect of dermal thickness, tissue composition, and body site on skin biomechanical properties. *Skin Research and Technology* **12**, 43–49 (2006).

54. Shwartz, Y. *et al.* Cell Types Promoting Goosebumps Form a Niche to Regulate Hair Follicle Stem Cells. *Cell* **182**, 578-593.e19 (2020).

55. NICOLAIDES, N. & ROTHMAN, S. Studies on the Chemical Composition of Human Hair Fat: II. The Overall Composition with Regard to Age, Sex and Race. *Journal of Investigative Dermatology* **21**, 9–14 (1953).

56. Oltulu, P., Ince, B., Kökbudak, N., Findik, S. & Kiliç, F. Measurement of epidermis, dermis, and total skin thicknesses from six different body regions with a new ethical histometric technique. *Turkish Journal of Plastic Surgery* **26**, 56–61 (2018).

57. Myronenko, A. & Song, X. Point set registration: Coherent point drifts. *IEEE Trans Pattern Anal Mach Intell* **32**, 2262–2275 (2010).

58. Oomori, S., Nishida, T. & Kurogi, S. Point cloud matching using singular value decomposition. *Artif Life Robot* **21**, 149–154 (2016).

59. Higham, N. J. The symmetric procrustes problem. *BIT Numerical Mathematics 1988 28:1* **28**, 133–143 (1988).

60. Chen, L.-C., Papandreou, G., Schroff, F. & Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. (2017) doi:10.48550/arxiv.1706.05587.

61. Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. 248–255 (2010) doi:10.1109/CVPR.2009.5206848.

62. Breiman, L. Bagging Predictors. (1994).

63. Graham, S. *et al.* HoVer-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images.

64. Fitzgibbon, A. W. & Fisher, R. B. A Buyer's Guide to Conic Fitting.

Acknowledgements

This work was supported through grants from the National Cancer Institute (U54CA143868 and U54CA268083 to DW, PW), National Cancer Institute (UG3CA275681 to PW, DW) the National Institute of Arthritis and Musculoskeletal and Skin Diseases (U54AR081774 to DW) and the National Institute on Aging (U01AG060903 to DW, JMP)

Conflict of Interest

The authors declare no financial/commercial conflicts of interest.

Contributions

KSH, IS, CJ, PW, DW conceived, developed, and designed the study. KSH, ER, CB performed sample collection. KSH, IS, JS developed skin tissue model. KSH, BS, AK, PW, DW wrote and edited the manuscript with input from all co-authors. PW, DW supervised study and secured funding.

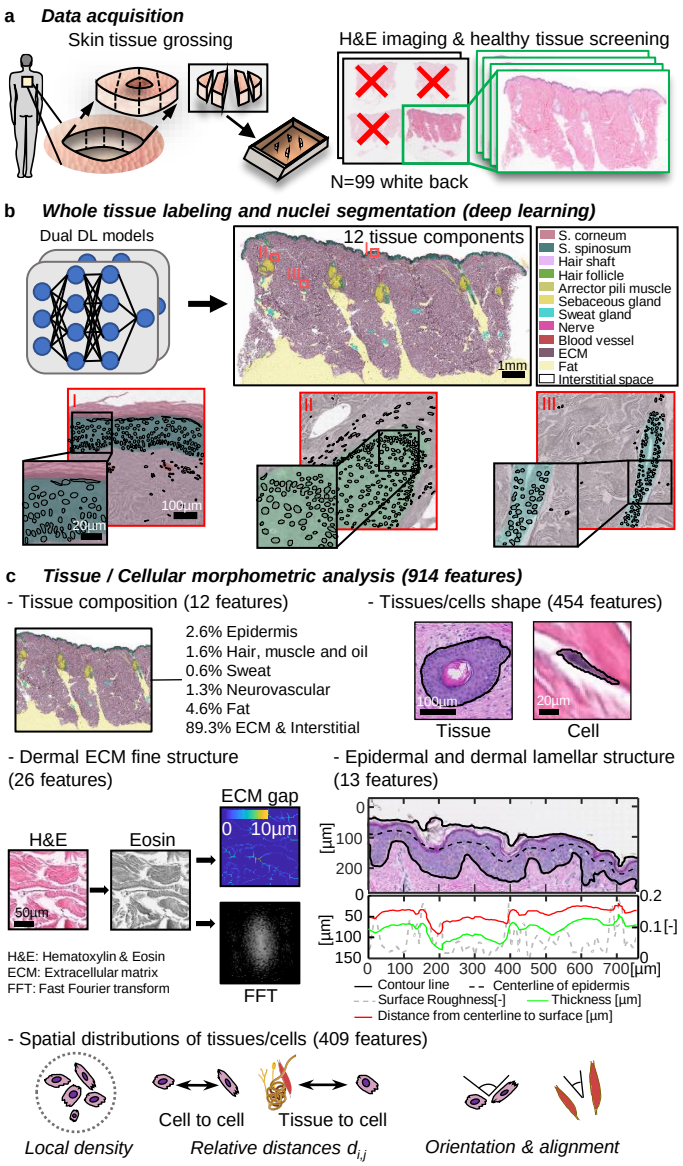
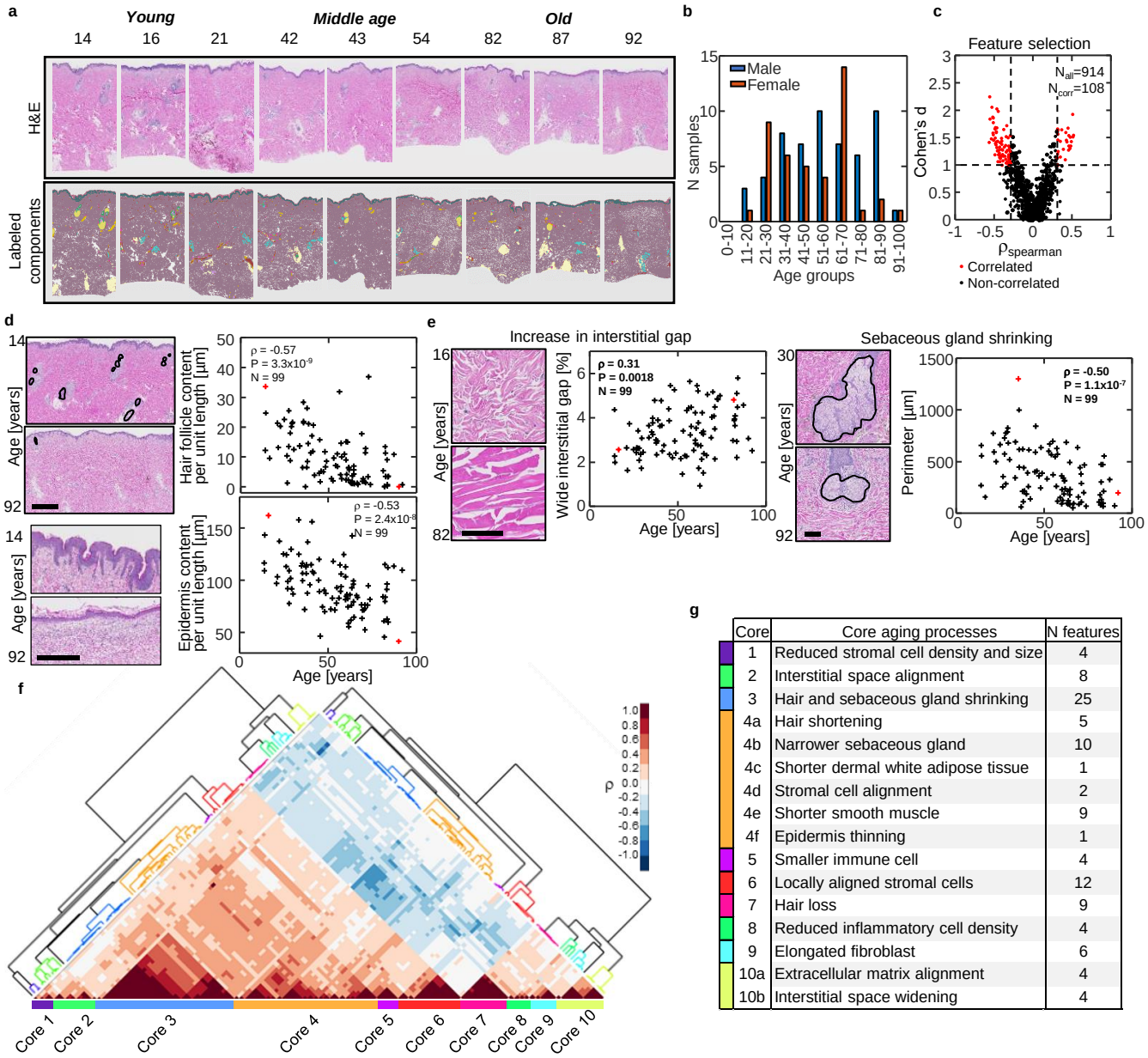


Fig. 1: Overall workflow of human skin tissue analysis a. Skin excisional biopsy sample typically with an elliptical shape is cut into multiple pieces before mounting into a paraffin block. Tissue section. The skin pieces with abnormal tissue architecture (such as lesion, sun damage, etc.) are excluded for further analysis through examining the image of H&E-stained tissue section. **b**. Whole slide H&E images are labeled with tissue components automatically using a trained CNN with DeepLab V3+ framework. Cell nuclei within each labeled tissue component are segmented using HoVerNet. Scale bar, 1mm in tissue, 100µm and 20µm in zoom-in. **c**. Five groups of morphometric features with a total of 914 features are established to characterize the tissue and cellular architecture of skin. Tissue scale bar, 100µm. Cell scale bar, 20µm. H&E scale bar, 50µm.



Core	Core aging processes	N features
1	Reduced stromal cell density and size	4
2	Interstitial space alignment	8
3	Hair and sebaceous gland shrinking	25
4a	Hair shortening	5
4b	Narrower sebaceous gland	10
4c	Shorter dermal white adipose tissue	1
4d	Stromal cell alignment	2
4e	Shorter smooth muscle	9
4f	Epidermis thinning	1
5	Smaller immune cell	4
6	Locally aligned stromal cells	12
7	Hair loss	9
8	Reduced inflammatory cell density	4
9	Elongated fibroblast	6
10a	Extracellular matrix alignment	4
10b	Interstitial space widening	4

Fig. 2: Core aging processes through extensive morphometric analysis. a. Representative image of H&E-stained tissue section from donors and at young, middle age and old (top). The corresponding tissue component maps labeled by the trained CNN (bottom). **b**. Histogram of age distribution of the sample cohort (56 male and 43 female) obtained from the back of skin. **c**. Age associated features is selected based on the spearman correlation coefficient with aging and Cohen's d from comparing young (30 years old or younger) vs old (60 years old or older). A scatter plot showing each features spear correlation coefficient and Cohen's d. A total of 108 features with Cohen's d >1 and |spearman's rho| >0.3 are identified as aging-associated features. **d**. Representative H&E images of skin and scatter plots showing the hair follicle content and epidermis content are among the identified aging-associated features. Hair follicle content and epidermis content are features known to be correlated with age. Scale bar, 1mm. **e**. Representative H&E images of skin and scatter plots show the increment of interstitial gap and shrinkage of the sebaceous gland, which are novel features of skin aging, among others. Scale bars, 100μm. Wide interstitial gap ratio is measure by ratio of the interstitial gap exceeding 4μm thickness among total by length. **f-g**. Hierarch clustering analysis groups the aging-associated features into ten distinct cores of the aging process using pair-wise correlation distance (**f**). Summary of general characteristics of each core based on the morphometric features of skin in each core (**g**). Cores 4 and 10 are further divided into subgroups to describe distinct characteristics of the assigned features better.

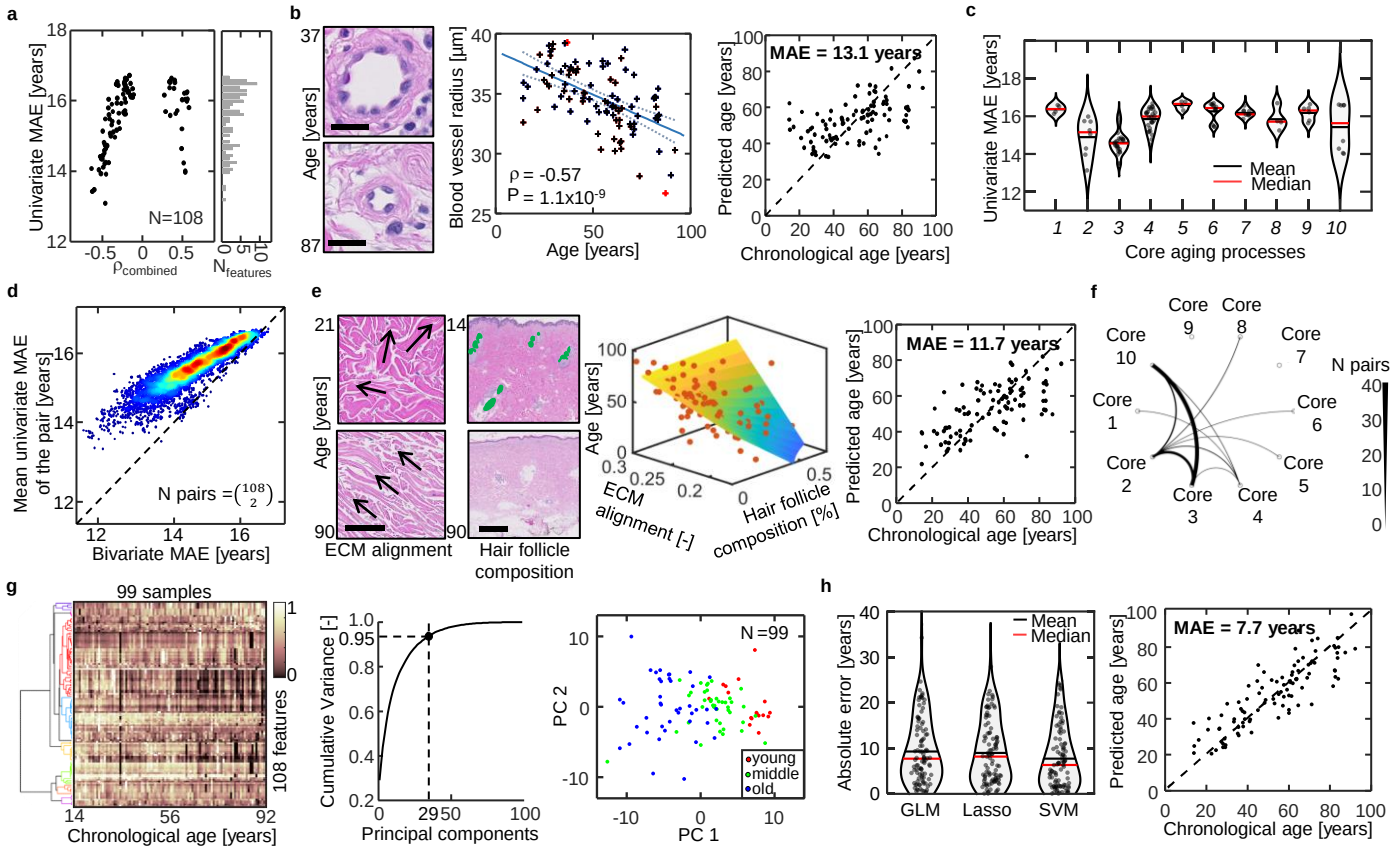
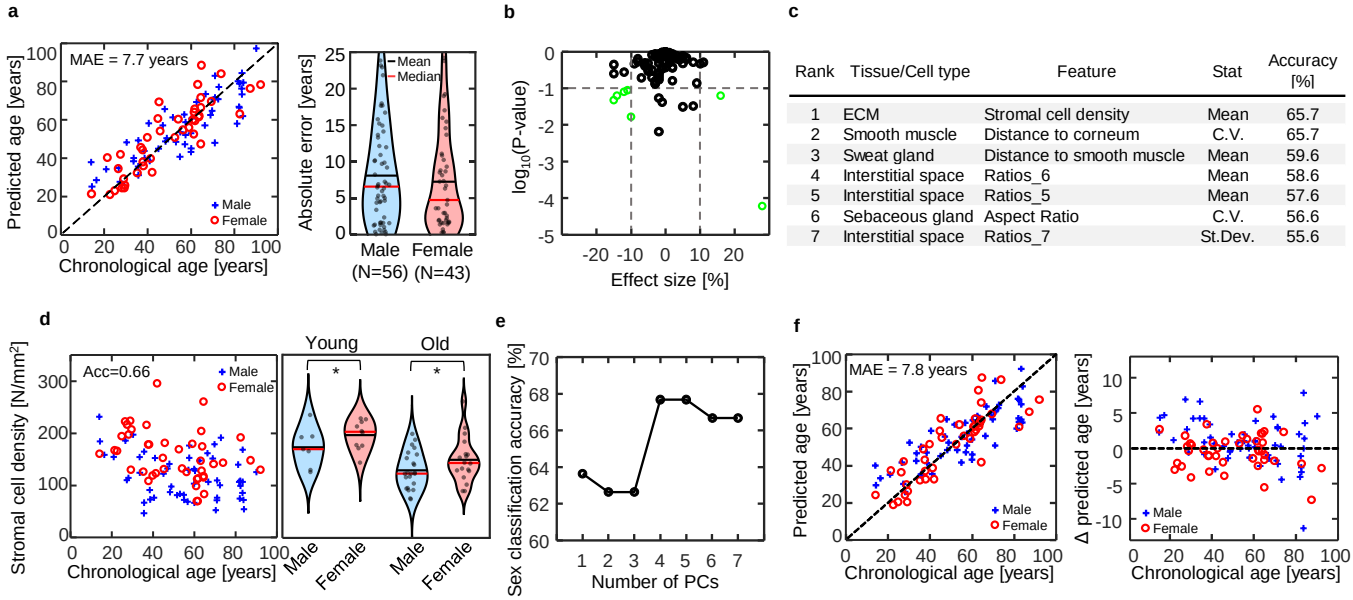


Fig. 3: Morphometric features are predictive of chronological age. a. A scatter plot shows correlation coefficient with aging and mean absolute error of aging prediction for each individual aging associated feature. Mean absolute error is calculated from leave-one-out validation using univariate generalized linear regression model (GLM). **b**. The blood vessel radius is the best predictive feature with MAE of 13.1 years. Two representative H&E images from a young and an old donor are shown. Scale bar = 20µm. A scatter plot shows a negative correlation between blood vessel radius from each sample and their corresponding ages. Predicted age with blood vessel radius from leave-one-out validation and actual chronological ages for each sample is shown as a scatter plot. **c**. Violin plot shows MAE distribution for features in each core of aging process. Features from core 2 (interstitial space alignment) and core 3 (Hair and sebaceous gland shrinking) show the overall best performance in aging prediction as single variable among others. **d-f**. Aging prediction uses bivariate features. A scatter plot shows the MAEs from all pairs (N=5,778) of features using bivariate aging prediction with GLM compared to the average MAE of each pair with univariate prediction with GLM. The MAE is calculated using the leave-one-out validation. In most pairs, bivariate MAE is lower than the mean univariate MAE (**d**). The best predictive bivariate pair is ECM alignment and hair follicle composition with an MAE of 11.7 years. Representative H&E images of skin in young and old highlight decrement in ECM alignment and loss of hair follicles with aging. Scale bar, 100µm for ECM and 1mm for hair follicle composition. A 3d scatter plot shows the association of these two features with aging among all samples (**e**). Analysis of the occurrence of predictive paired features associated with cores of the aging process. In the top 100 most predictive bivariate pairs, the most frequently occurred pairs are between core 3 (hair and sebaceous gland shrinking) and core 10 (ECM alignment and Interstitial space widening) (**f**). **g**. The heatmap shows 99 tissue specimens (ranked in order by age) and the corresponding 108 age-associated features. Principal component analysis shows the first 29 principal components contain 95% of the total variance. A scatter plot shows the distribution of the sample along the first two principal components (PCs) and readily shows the chronological age of samples is associated with PC1 and PC2. **h**, Implementation of multivariate models with first 29 PCs for aging prediction. A violin plot shows the MAE of each sample from leave-one-out validation using GLM, Lasso, and SVM. A scatter plot shows predicted age (leave-one-out validation) vs chronological age using SVM in each sample. SVM achieved the lowest MAE of 7.7 years.



Rank	Tissue/Cell type	Feature	Stat	Accuracy [%]
1	ECM	Stromal cell density	Mean	65.7
2	Smooth muscle	Distance to corneum	C.V.	65.7
3	Sweat gland	Distance to smooth muscle	Mean	59.6
4	Interstitial space	Ratios_6	Mean	58.6
5	Interstitial space	Ratios_5	Mean	57.6
6	Sebaceous gland	Aspect Ratio	C.V.	56.6
7	Interstitial space	Ratios_7	St.Dev.	55.6

Fig. 4: Skin aging features are not strongly associated with biological sex. a. A scatter plot of predicted ages and chronological ages in males and females. Predicted age is derived from leave-one-out using the multivariate SVM. MAE is 7.3 years among female samples compared to 8.1 years among male samples. **b**. A scatter plot illustrating the association between gender and 108 aging features with relative differences in mean and p-value (two-way t-test) between genders. Seven features have a relative difference of 10%+ with a p-value less than 0.1. **c**. Summary of the gender-associated features and their corresponding accuracy in classifying gender. **d**. A scatter plot shows the relation between stromal cell density and the age of donors in both genders. Violin plots comparing stromal cell density between genders in both young (30 years old or younger) and old (60 years old or older) show that stromal cell density is significantly higher in females than in males in both age ranges. **e**. The accuracy of gender classification is based on an increasing number of principal components derived from gender-associated features. The highest accuracy, reaching 68%, is attained using the first 4 principal components (PCs). **f**. Incorporating gender information into the aging prediction using the multivariate SVM model as depicted in Fig. 3h, a scatter plot illustrates the predicted age versus chronological age for donors of both genders. The Mean Absolute Error (MAE) is 7.8 years when gender information is considered in the prediction model. The scatter plot further depicts the difference in prediction errors for each sample, comparing models with and without considering gender. On average, there is a +0.8 years change in males and a -0.5 years change in females. The overall MAE increases by 0.1 year with the inclusion of gender information when compared to the MAEs obtained without considering sample gender information.

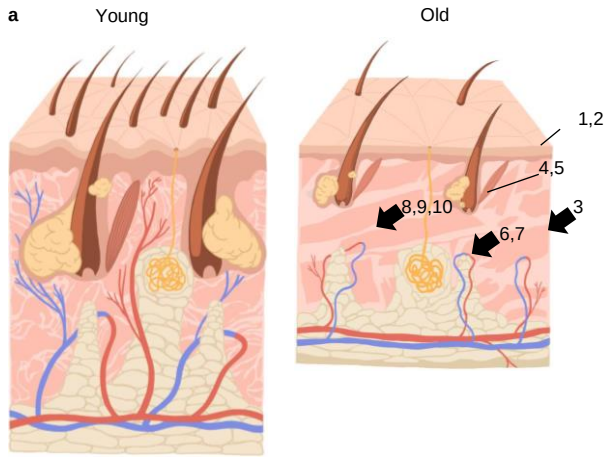


Fig. 5: the effect of aging on the microanatomy of human skin a. illustration of distinct macroscopic differences between young (age 10~30) and old (age 60~90) human skin, labeled with the corresponding number in the following list: 1 and 2. flatter and thinner epidermis, 3. thinner dermis, 4. fewer pilosebaceous units (a unit includes hair shaft, hair follicle, sebaceous gland, and arrector pili muscle), 5. smaller pilosebaceous units, 6. narrower blood vessels, 7. blood vessel distant from epidermis, 8. ECM more aligned with epidermis, 9. thicker ECM bundles, 10. wider interstitial space

Rank	Feature ID	Tissue component	Feature name	Feature type	Core process	MAE [years]
1	218	Blood vessel	Mean bounding box height	4	6	13.1
2	116	Hair follicle	Content per length scale	4	5	13.5
3	4	Hair follicle	Composition	1	5	13.5
4	76	S. Spinosum	Content per length scale	4	6	14.0
5	121	Hair follicle	Occurrence per length scale	4	5	14.0
6	27	ECM	Anisotropy_B	3	1	14.0
7	29	ECM	Anisotropy_B_masked	3	1	14.0
8	160	Sebaceous gland	Bounding box height st.dev.	4	5	14.1
9	153	Sebaceous gland	Size st.dev.	4	5	14.1
10	163	Sebaceous gland	Major axis length st.dev.	4	5	14.2

Table 1: Top 10 univariate features

Rank	Feature ID	Tissue component	Feature name	Feature type	Core process	MAE [years]
1	4	Hair follicle	Composition	1	5	11.7
	29	ECM	Anisotropy_B_masked	3	1	
2	4	Hair follicle	Content per length scale	1	5	11.7
	27	ECM	Anisotropy_B	3	1	
3	160	Sebaceous gland	std bounding box height	4	5	11.8
	218	Blood vessel	mean bounding box height	4	6	
4	29	ECM	anisotropy_B_masked	3	1	11.8
	116	Hair follicle	content per length scale	4	5	
5	29	ECM	anisotropy_B_masked	3	1	11.8
	218	Blood vessel	mean bounding box height	4	6	
6	171	Sebaceous gland	std perimeter	4	5	11.8
	218	Blood vessel	mean bounding box height	4	6	
7	27	ECM	anisotropy_B	3	1	11.8
	116	Hair follicle	content per length scale	4	5	
8	153	Sebaceous gland	std size	4	5	11.9
	218	Blood vessel	mean bounding box height	4	6	
9	27	ECM	anisotropy_B	3	1	11.9
	218	Blood vessel	mean bounding box height	4	6	
10	153	ECM	anisotropy_A	3	1	11.9
	218	Blood vessel	mean bounding box height	4	6	

Table 2: Top 10 bivariate features