Non-invasive human skin transcriptome analysis using mRNA in skin surface lipids

Takayoshi Inoue*, Tetsuya Kuwano, Yuya Uehara, Michiko Yano, Naoki Oya, Akira Hachiya, Yoshito Takahashi, Noriyasu Ota, Takatoshi Murase*

Biological Science Research, Kao Corporation, Tochigi, Japan

2606, Akabane, Ichikai-machi, Haga-gun, Tochigi, 321-3497, Japan

***Corresponding authors**

Takayoshi Inoue, PhD; E-mail: inoue.takayoshi@kao.com

ORCID ID: 0000-0003-1038-4127

Takatoshi Murase, PhD; E-mail: murase.takatoshi@kao.com

ORCID ID: 0000-0002-5496-0133

Abstract

Non-invasive acquisition of mRNA data from the skin would be extremely useful for understanding skin physiology and diseases. Inspired by the holocrine process, in which the sebaceous glands secrete cell contents into the sebum, we focused on the possible presence of mRNAs in skin surface lipids (SSLs). We found that measurable human mRNAs exist in SSLs, where sebum protects them from degradation by RNases. The AmpliSeq transcriptome analysis was modified to measure SSL-RNAs, and our results revealed that SSL-RNAs predominantly contained mRNAs derived from sebaceous glands, epidermis, and hair follicles. Analysis of SSL-RNAs non-invasively collected from patients with atopic dermatitis revealed significantly increased expression of inflammation-related genes and decreased expression of terminal differentiation-related genes, consistent with the results of previous reports. Further, we found that lipid synthesis-related genes were downregulated in the sebaceous glands of patients with atopic dermatitis. These results indicate that the analysis of SSL-RNAs is promising to understand the pathophysiology of skin diseases.

Introduction

Intra- and inter-organ communication mediated via various hormones, growth factors, cytokines, metabolites, and miRNAs play important roles in maintaining homeostasis in the human body (1). Several efforts have been made to establish comprehensive analytical methods for these mediators to monitor the physiological conditions of the body and explore predictive biomarkers for various diseases (2–5). Especially, the use of serum, urine, and saliva samples, which can be obtained in a non- or low-invasive manner, has been widely investigated.

The skin is often referred to as "the window to body's health" since the skin phenotypes, such as the cutaneous pathology, appearance, and its secretions reflect not only the skin condition but also the condition inside the body (6). Moreover, the skin forms the body surface and biomolecules can be easily collected from the sweat, hair, and stratum corneum samples, and thus, skin is a useful source of samples to monitor the skin and body conditions. For instance, the cortisol content in the scalp hair correlates with long-term cumulative cortisol exposure (7). The sweat can also be used as an indicator of internal physiological changes (8), and attempts have been made to monitor patients' conditions, for instance, tracking blood glucose levels by measuring glucose in sweat samples of patients with diabetes (9). Although metabolites, proteins, and DNA are relatively easy to collect from the sweat and hair samples (10, 11), it is difficult to collect measurable mRNAs from skin in a non- or low-invasive manner. So far, tape-

stripped stratum corneum has been used to collect mRNAs in a minimally invasive manner (12), however, the mRNA content is very low and highly degraded due to the RNase activity on the skin surface. Therefore, a skin biopsy is practically required to analyze mRNA expression; however, this method is invasive, which limits its application. More recently, a minimally invasive method for mRNA analysis via RNA-seq using AmpliSeq technology with 16–20 consecutive tape strips was reported (13–15). However, tape stripping of the stratum corneum is known to induce skin damage, including disruption of the skin barrier (16), epidermal hyperproliferation, and infiltration of CD3-positive T cells into the dermis (17), indicating that the problems related to the invasiveness of this technique remain to be fully resolved.

The sebaceous glands synthesize and accumulate lipids to produce sebum. The lipids accumulated in the cytosol of sebocytes are secreted into the sebaceous ducts following rupture of the plasma membrane; this mode of secretion is called holocrine secretion (18) and is unique among the exocrine glands such as lipid-secreting sebaceous and meibomian glands. The holocrine secretion of the cell contents led us to the idea that sebum may contain various biomolecules, including mRNAs, which may be useful for analyzing biological information. Therefore, in this study, we first investigated the presence of mRNAs in human sebum and established a non-invasive, comprehensive method of analyzing human mRNAs using skin surface lipids (SSLs) as samples. Further, the applicability of this method in skin

characterization was verified in healthy subjects and patients with atopic dermatitis (AD).

Methods

Subject recruitment and collection of SSLs

Thirty-two healthy male individuals (mean age: 34.6 years, range: 20–49 years, SD = 9.24) were recruited for the study that was conducted in October 2016. The individuals were evaluated by dermatologists prior to the commencement of the study to confirm no obvious skin disease or condition on their faces. Thirty male patients (mean age: 31.0 years, range: 20–48 years, SD = 8.82) diagnosed with mild or moderate facial AD were recruited during June–October 2017. All subjects were required not to remove facial sebum by washing or using wipes or shaving their face on the test day until the end of the test. Patients with AD were also restricted to use steroidal anti-inflammatory and immunosuppressive drugs on the facial skin 24 h prior to the study. The study was approved by the Human Research Ethics Committee, Kao Corporation (approval numbers: 792-2016082 and T003-170413), Japan Aesthetic Dermatology Symposium (approval number: KU-2017-05-003), and the Shinjukuminamiguchi Dermatologic Clinic (approval number: KU-2016-10-005).

SSLs were collected by wiping the whole face (forehead, cheek, face line, nose, and chin) using an oil blotting film (8.0 cm x 5.0 cm, 3M Japan, Tokyo, Japan) and samples were stored in glass vials at -80 °C until use.

88 **Skin tissue**

89 Surgically removed adult forehead skin from three Caucasian males, aged 62, 66, and 67
90 years, and nose skin from one adult (Caucasian female, 66-years-old) were provided by the
91 Colorado Dermatology Institute (Colorado Springs, CO, USA) for laser microdissection and
92 immunostaining, respectively. The procurement of skin tissues was approved by the Institutional
93 Review Board of IntegReview Ltd. (Austin, TX, USA; approval number: T046a-170829) and
94 was conducted according to the Declaration of Helsinki Principles. Informed consents were
95 obtained from the volunteers prior to surgery. After surgery, the skin tissues were stored in
96 William's E medium (Life Technologies, Carlsbad, CA, USA) at 4 ºC until embedding. All skin
97 tissues were embedded using the Tissue-Tek optimal cutting temperature compound (Sakura
98 Finetek, Tokyo, Japan) and kept frozen until sectioning.

99

100 **mRNA extraction and qPCR**

101 Human saliva, sweat, urine, and serum samples collected from two donors were purchased
102 from Cosmo Bio (Tokyo, Japan). Total RNA was extracted from 1 mL of each sample using the
103 TRIzol LS reagent (Thermo Fisher Scientific, Waltham, MA, USA). Total RNA was extracted
104 from the stratum corneum of the cheek of two healthy males according to a previous report (12).
105 Collected RNA was dissolved with 10 µL of nuclease-free water. RNAs in SSLs (SSL-RNAs)

106 was extracted using the TRIzol reagent (Thermo Fisher Scientific) as follows: 2.85 mL of
107 TRIzol was added to a finely cut oil blotting film containing sebum samples. Next, the solution
108 was divided equally into two tubes and 260 μL of chloroform was added to each tube and mixed
109 by vortexing. The tubes were centrifuged at 12,000 × *g* for 15 min at 4 ºC. The upper layer was
110 transferred to a fresh tube and precipitated with ethanol. The precipitates were washed with
111 70 % ethanol (v/v) and dissolved in 10 μL of nuclease-free water.
112 Reverse transcription was performed using the SuperScript IV First-Strand Synthesis
113 System and Oligo-dT primers (Thermo Fisher Scientific). The qPCR was performed using the
114 TaqMan Fast Universal PCR Master Mix (Thermo Fisher Scientific) and TaqMan probes for
115 each gene (Thermo Fisher Scientific).
116
117 **Evaluation of mRNA degradation in SSLs**
118 To prepare standard samples with different levels of degraded mRNAs, total RNA (1 μg)
119 extracted from normal human epidermal keratinocytes (NHEK) (Cascade Biologics, Portland,
120 OR, USA) was incubated with 30–1000 ng/mL of recombinant human RNase 7 (Novus
121 Biologicals, Littleton, CO, USA) in 10 mM Tris-HCl buffer (pH 8.0) (Nippon Gene, Tokyo,
122 Japan) for 30 min. QIAzol lysis reagent (Qiagen, Hilden, Germany) and chloroform were added
123 to the samples treated with RNase 7 or SSLs collected from six healthy males as described

above. After the solutions were vortexed and centrifuged, RNA in the supernatant was purified using the miRNeasy Mini Kit (Qiagen). The level of RNA degradation in the samples was determined using the High Sensitivity RNA ScreenTape (Agilent Technologies, Palo Alto, CA, USA) on an Agilent 4200 TapeStation system (Agilent Technologies). The level of degradation of human mRNAs in SSLs was estimated using the following calculated DV200 value, which evaluates the percentage of fragments containing > 200 nucleotides. mRNA was reverse-transcribed using the SuperScript IV First-Strand Synthesis System and Oligo-dT primers (Thermo Fisher Scientific) and was amplified in a thermal cycler using the PowerUP SYBR Green Master Mix (Thermo Fisher Scientific) and the following primers:

ACTB (forward primer 57 bp): 5'-GCTTTTGGTCTCCCTGGGAG-3'

ACTB (forward primer 363 bp): 5'-ACAATGTGGCCGAGGACTTT-3'

ACTB (reverse primer): 5'-AGTCAGTGTACAGGTAAGCCC-3'

Abundance of each amplicon was determined from the RNA standard curve. Further, DV200 values of the SSL-RNA samples were calculated from the standard curve of DV200 plotted against abundance ratio of the amplicons (363 bp/57 bp).

Immunostaining

Skin sections obtained from the nose of a Caucasian female were fixed in acetone at -20 °C

142 for 10 min and treated with 0.1 % Triton X-100/PBS for 5 min. The skin sections were
143 incubated with Protein Block serum-free (Agilent Technologies) for 30 min, then with the
144 primary antibodies against keratin/cytokeratin (mouse mono-clonal (AE-1/AE-3), original
145 solution, Nichirei, Tokyo, Japan) or RNase 7 (rabbit poly-clonal, 1:50, Cloud‑Clone Corp.,
146 Houston, TX, USA) for 1 h at 20–25 °C, and finally with anti-rabbit IgG (donkey poly-clonal,
147 Alexa Fluor 555, 1:1000, Thermo Fisher Scientific) or anti-mouse IgG (goat poly-clonal, Alexa
148 Fluor 647, 1:1000, Thermo Fisher Scientific) for 30 min. Samples were mounted on a glass slide
149 and imaged using fluorescence microscopy (BZ-X710, Keyence, Osaka, Japan).
150
151 **Western blotting**
152 Total protein was extracted from SSLs using 350 μL of RIPA buffer and was purified using
153 the Ready Prep 2-D cleanup kit (Bio-Rad, Hercules, CA, USA); the protein concentration was
154 measured using the BCA protein assay kit (Thermo Fisher Scientific). Afterwards, 70 μL of
155 trichloroacetic acid was added and the samples were incubated on ice for 30 min, followed by
156 centrifugation at 13,000 × *g* for 5 min at 4 °C. Chilled acetone (500 μL) was added to the
157 pellets, the tubes were centrifuged at 13,000 × *g* for 5 min at 4 °C, and the supernatant was
158 removed. The pellets were dissolved in SDS sample buffer (Novagen, Darmstadt, Germany) and
159 boiled at 95 °C for 5 min. Proteins (5 μg) were separated on a 4–15 % polyacrylamide gradient

gel (Bio-Rad) and then transferred to a polyvinylidene fluoride (PVDF) membrane (Bio-Rad) soaked in Tris-Glycine Buffer (25 mM Tris, 192 mM glycine; pH 8.2) containing 20 % (v/v) methanol. The PVDF membrane blocked with PVDF Blocking Reagent for Can Get Signal (Toyobo, Tokyo, Japan) was incubated with anti-RNase 7 antibody (rabbit polyclonal, 1:200, Cloud-Clone Corp, Katy, TX, USA) for 60 min followed by incubation with anti-rabbit IgG, horseradish peroxidase-linked secondary antibody (donkey monoclonal, 1:2000, GE Healthcare, Bucks, UK) for 45 min. The bands were visualized using ECL prime western blotting detection reagents (GE Healthcare).

Effect of sebum lipids on RNase activity

The oil blotting film used to collect SSLs from the face was finely cut and 1 mL distilled water and 4 mL tert-butyl methyl ether were added to the cut film. The solution was transferred to a new glass vial, and centrifuged at 2,050 × *g* at 4 °C for 10 min. The upper layer was transferred to a new glass vial, and the organic solvent was dried by blowing nitrogen over it. The remaining sebum lipids were dissolved in 100 μL dimethyl sulfoxide (DMSO) and used for analysis.

Cholesterol ester (cholesteryl palmitate, Sigma-Aldrich, St. Louis, MO, USA, C6072), wax ester (lauryl palmitoleate, Santa Cruz Biotechnology, Santa Cruz, CA, USA, sc-280908),

178 triacylglycerol (glyceryl trioleate, Sigma-Aldrich, T7140), free fatty acid (palmitoleic acid,
179 Sigma-Aldrich, P9417), squalene (Sigma-Aldrich, S3626), and cholesterol (Sigma-Aldrich,
180 C8667) were used as authentic samples to identify the lipid molecular species. The sebum lipids
181 extracted from the oil blotting films were dissolved in chloroform/methanol (2:1, v/v). The
182 authentic lipids and 5 mg sebum lipid samples were separated on a thin-layer chromatography
183 (TLC) plate using hexane: diethyl ether: acetic acid (70:30:1, v/v/v). After chromatography, the
184 plate was divided using a glass cutter into portions containing the authentic lipids and sebum
185 lipid samples, and the portion containing the authentic lipids was sprayed with a solution
186 containing 10 % (w/v) copper sulfate and 8 % (w/v) phosphoric acid, and then heated at 180 °C
187 for 3 min to visualize the location of each lipid. Based on the mobility of the authentic lipids,
188 the silica corresponding to portions of sebum lipids was scraped off. The silica was sonicated in
189 the chloroform/methanol (2:1, v/v) mixture to extract the lipids and centrifuged at 2,050 × *g* for
190 5 min. The supernatant was dried and dissolved in 15 μL DMSO; the solution was further
191 diluted 3- and 9-fold using DMSO.

192 Cholesterol ester (cholesteryl palmitoleate, Olbracht Serdary Research Laboratories,
193 Toronto, Canada, D-161), wax ester (behenyl palmitoleate, Nu Chek Prep, Elysian, MN, USA,
194 WE-1368), triacylglycerol (glyceryl tripalmitoleate, Sigma-Aldrich, T5888; glyceryl trioleate,
195 Sigma-Aldrich, T7140), free fatty acids (myristoleic acid, Sigma-Aldrich, M3525; palmitoleic

196 acid, Sigma-Aldrich, P9417; oleic acid, Sigma-Aldrich, O1008), squalene (Sigma-Aldrich,
197 S3626), and cholesterol (Sigma-Aldrich, C8667) were used to determine the active ingredients
198 in the sebum. NHEK RNA (20 μg/mL final concentration) was mixed with RNase 7 (final
199 concentration: 1 μg/mL) in 10 mM Tris-HCl buffer (pH 8.0) with or without sebum or the lipid
200 reagent (total 50 μL), and sonicated, followed by incubation at 20–25 °C for 30 min.
201 Subsequently, RNA was extracted using TRIzol LS reagent and the quality of the RNA was
202 determined using the High Sensitivity RNA ScreenTape on Agilent 4200 TapeStation System.
203
204 **Library preparation for Ion AmpliSeq**
205 Following the addition of 2.85 mL of QIAzol reagent (Qiagen) to a finely cut oil blotting
206 film containing sebum samples, QIAzol solution was divided equally into two tubes.
207 Chloroform (260 μL) was added to each tube and vortexed, and the tubes were centrifuged at
208 12,000 × *g* for 15 min at 4 ºC. The upper layer was transferred to a fresh tube. RNA was purified
209 using the RNeasy mini kit (performing DNase treatment in the purification step) (Qiagen) and
210 eluted from the resin twice using 50 μL of nuclease-free water followed by ethanol
211 precipitation, and then dissolved in 10 μL of nuclease-free water.
212 To improve the success rate of the library preparation by AmpliSeq protocol, we modified
213 the default protocol of the Ion AmpliSeq Transcriptome Human Gene Expression kit (Thermo

Fisher Scientific). Briefly, 1.75 µL RNA solution was mixed with 0.5 µL VILO Reaction Mix and 0.25 µL SuperScript III Enzyme. Reverse transcription was performed at 25 ºC for 10 min, 42 ºC for 90 min, and finally 85 ºC for 5 min. The target DNA amplification was performed by mixing 2.5 µL cDNA solution, 1.5 µL nuclease-free water, 2.0 µL Ion AmpliSeq HiFi Mix, and 4.0 µL Ion AmpliSeq Transcriptome Human Gene Expression Core Panel under the following conditions: 99 °C for 15 sec and 62 °C for 16 min for 20 cycles. The amplified DNA library was purified by mixing 10 µL AMPure XP beads (Beckman Coulter, Miami, FL, USA) according to the manufacturer's protocol and eluted using 10 µL nuclease-free water. The quality check of the DNA library was conducted using the High Sensitivity D1000 ScreenTape on Agilent 4200 TapeStation. If the DNA library had amplified, a band of approximately 170 bp was observed. After checking the DNA library quality, the reaction solution was prepared by mixing 3.5 µL purified library solution, 2.0 µL Ion AmpliSeq HiFi Mix, 4.0 µL Ion AmpliSeq Transcriptome Human Gene Expression Core Panel, and 0.5 µL VILO Reaction Mix. After adding 1.0 µL FuPa reagent to 10 µL reconstituted reaction solution, the primer sequence was partially digested under the following conditions: 50 °C for 10 min, 55 °C for 10 min, and 60 °C for 20 min. To ligate the adaptor sequence, 2 µL Switch solution, 1 µL Ion Xpress Barcode adapters, and 1 µL DNA ligase were added to 11 µL reaction solution, followed by incubation at 22 ºC for 60 min and 72 ºC for 5 min. The library (15 µL) ligated with the adaptor sequence was purified via

mixing with 18 µL AMPure XP beads according to the manufacturer's protocol. Libraries were eluted using 50 µL Library Amp Mix (Thermo Fisher Scientific) to which, 2 µL Library Amp Primers were added. The library amplification was conducted using following steps: 98 °C for 15 seconds and 64 °C for 1 min for five cycles. Next, 50 µL PCR product was mixed with 25 µL AMPure XP beads and the supernatant was transferred to fresh PCR tubes. The supernatants were mixed with 60 µL AMPure XP beads and purified; target fragments were eluted from the beads using 10 µL TE buffer. The quality check of the library was performed using the High Sensitivity D1000 ScreenTape on Agilent 4200 TapeStation.

Sequencing

The library was quantified using the Ion Library TaqManTM Quantitation Kit (Thermo Fisher Scientific). After an input of 50 pM DNA library in the Ion Chef System (Thermo Fisher Scientific), template preparation and chip loading were performed, and RNA-seq was conducted on the Ion S5 XL System (Thermo Fisher Scientific).

Verification of the correlation coefficient between AmpliSeq and qPCR results

For identifying *RPLP0*, *CDSN*, and *CCL17* expression in SSL-RNAs, cDNA was pre-amplified in 14 cycles using the TaqMan PreAmp Master Mix (Thermo Fisher Scientific) and

pooled TaqMan probe (*RPLP0, CDSN*, and *CCL17*) (Thermo Fisher Scientific) and then diluted 5-fold with nuclease-free water. The qPCR was performed according to the aforementioned "mRNA extraction and qPCR" method. Expression value of *RPLP0* was used as an internal control. Correlation between the value of reads per million mapped reads (RPM) of AmpliSeq and the relative expression value of qPCR in healthy subjects and patients with AD was analyzed.

Laser microdissection (LMD) and AmpliSeq transcriptome analysis

Frozen skin sections (thickness: 10 μm) from three Caucasian males were mounted on membrane slides (PEN-Membrane 2.0 μm, Leica Microsystems, Wetzlar, Germany) treated with 0.1 % (w/v) poly-L-lysine (Fujifilm Wako Pure Chemical, Osaka, Japan). In addition, two frozen sections were directly collected into 750 μL RLT buffer (Qiagen) containing 40 mM dithiothreitol (Sigma-Aldrich) to analyze the transcriptome of the whole tissue. The LMD sections were fixed with acetone at -20 °C for 10 min, stained with 0.05 % (v/v) toluidine blue, and finally dried. Epidermis, sebaceous glands, sweat glands, hair follicles, and dermis were carefully microdissected from the skin sections using LMD7000 (Leica Microsystems). Fifteen target regions from three skin tissues were dissected and dissolved in 50 μL RLT buffer containing 40 mM dithiothreitol. Total RNA was extracted and purified using the RNeasy mini

kit (performing DNase treatment in the purification step) (Qiagen). The RNA was concentrated via ethanol precipitation and dissolved in 10 µL nuclease-free water. The RNA quality was checked on the 4200 TapeStation System and cDNA library was prepared for AmpliSeq transcriptome sequencing using 75 pg total RNA according to the method described in "Library preparation for Ion AmpliSeq". The number of target amplification was changed from 20 to 18 cycles except for dermis samples.

Statistics, normalization, and differential expression analysis of the AmpliSeq whole transcriptome

All statistical analyses and normalization of the RNA-seq transcriptome data were performed using the R statistical language. Read counts were generated using the AmpliSeq RNA plugin in Ion Torrent Suite Software (Thermo Fisher Scientific) and normalized with the DESeq2 R package (Bioconductor). For differential expression analysis between healthy subjects and patients with AD, a likelihood ratio test was performed with DESeq2 using normalized counts. Heat maps were generated using the heatmap3 package. Dimensionality reduction using t-distributed stochastic neighbor embedding (t-SNE) was performed using the Rtsne function of the Rtsne package. All plots were generated using the tidyverse package in combination with the reshape2, gplots, ggplot2, grid, and cowplot packages.

Results

Measurable human mRNA is present in SSLs

qPCR was conducted to evaluate the mRNA abundance in SSL samples, stratum corneum, urine, serum, saliva, and sweat samples. The expression of *ACTB* and *GAPDH* mRNA in SSLs was comparable to their expression in 100 ng–500 pg and 1 ng–100 ng total RNA in NHEK, respectively, whereas these transcripts were largely undetectable in the other body fluid samples (Fig. 1a). The level of human mRNA degradation in SSLs could not be measured directly due to the presence of bacterial mRNA. Therefore, we established an assay to verify the quality of human mRNA in SSLs. The reverse primer was designed near the 3' end and was used with the forward primers to amplify 363 bp and 57 bp long human *ACTB* mRNA. After reverse transcription using oligo-dT primers, *ACTB* levels were quantified by performing qPCR using primers generating 363 bp and 57 bp human *ACTB* mRNA. In the case of *ACTB* mRNA of longer than 363 bp, both 363 and 57 bp fragments were amplified, whereas in case of *ACTB* mRNA of 57–363 bp, only the 57 bp fragment was amplified (Fig. 1b). This assay was validated using RNA subjected to artificial and gradual degradation using RNase. The gradually degraded RNA was analyzed to calculate the percentage of fragments containing > 200 nucleotides (DV200 value) (Fig. 1c, d). The DV200 values of gradually degraded RNAs negatively correlated with the abundance ratio of the 363 bp and 57 bp amplicons (Fig. 1e). The level of

304 human mRNA degradation in SSLs collected from six males was calculated using a standard
305 curve (Fig. 1e) and showed a mean DV200 value of 56.5 % (Fig. 1f).
306
307 **Sebum lipids inhibit RNase activity**
308 Consistent with the results of a study reporting RNase 7 expression in human skin (19), we
309 confirmed that RNase 7 was expressed in the sebaceous glands and epidermis and also detected
310 in SSLs (Fig. 2a, b). Because there is high abundance of RNase on the surface of human skin, it
311 was surprising to detect human mRNA in SSLs. This finding led us to hypothesize that sebum
312 lipids inhibit RNase activity. Based on these results, we evaluated the influence of sebum lipids
313 on recombinant RNase 7 activity. Intact cellular RNA and RNase 7 were incubated with or
314 without sebum lipids from four subjects at 37 ºC for 30 min. Interestingly, the 28S and 18S
315 ribosomal RNAs were completely degraded after incubation with RNase 7 in the absence of
316 sebum lipids; on the other hand, ribosomal RNAs were stable in the presence of lipids (Fig. 2c).
317 Next, we aimed to identify the key lipids that inhibit RNase 7 activity. The sebum lipids were
318 separated into fractions A–D on a TLC plate (Fig. 2d), and the lipids were recovered from each
319 fraction and subjected to the RNase 7 inhibition assay. We observed that while fractions A, B,
320 and C decreased RNase 7 activity, the inhibitory effect of the fraction D was weak (Fig. 2e).
321 Triglycerides and esters in the sebum are hydrolyzed by the skin microbiome to generate free

322 fatty acids (FFAs). The FFAs in human sebum are predominantly composed of 16 carbon atoms
323 (palmitic acid, 16:0; sapienic acid, 16:1Δ6; and palmitoleic acid, C16:1Δ9) (14). Therefore, we
324 evaluated the inhibitory effects of free palmitoleic acids and various lipids with palmitoleic acid
325 as their main fatty acid, on RNase activity. The free palmitoleic acids and other lipids was
326 dissolved in the reaction buffer at 1 mg/mL or 100 mg/mL concentration, except cholesterol and
327 wax esters, and cholesterol, which could not be dissolved at 100 mg/mL. Our results showed
328 that FFAs strongly suppressed RNase activity at 1 mg/mL compared to other lipids (Fig. 2f).
329 Moreover, FFAs of different chain lengths (myristoleic acid (C14:1) and oleic acid (C18:1)) also
330 suppressed RNase 7 activity (Fig. 2g).
331
332 **Global expression analysis of SSL-RNAs**
333 For specific and comprehensive quantification of human mRNAs in SSLs, we performed the
334 AmpliSeq transcriptome analysis that can perform multiplexed amplification of cDNA
335 amplicons for more than 20,000 genes. Moreover, this method can analyze even small amounts
336 of RNAs as well as degraded RNAs (20, 21). Although we attempted to prepare sequence
337 libraries from SSL-RNAs based on the default protocol, our success rate was low. Since the data
338 quality can be improved via optimization of the AmpliSeq protocol (22), we modified the
339 experimental conditions for preparing the sequence library. In the modified protocol, we altered

the volume of reagents, standardized the conditions for reverse transcription and target amplification, and added a purification step after the target amplification to remove the primer dimers (Supplementary Method 1). With the new protocol, the success rate of the library preparation improved significantly and the AmpliSeq library was prepared with samples obtained from healthy subjects (91 %, 29/32) and patients with AD (100 %, 30/30). To analyze the experimental bias of our protocol, we verified the correlation between the expression results of AmpliSeq and qPCR of thymus and activation-regulated chemokine (*TARC*/*CCL17*) and corneodesmosin (*CDSN*), and observed a high correlation (*CCL17*, R = 0.81, *CDSN*, R = 0.87) (Fig. 3a). Furthermore, the correlation coefficients for the technical replicate of the reverse transcription (0.94 and 0.90) confirmed that our protocol had low experimental bias (Fig. 3b).

The SSL-RNA expression profile predominantly reflects mRNA expression in sebaceous glands, epidermis, and hair follicles

The regions of the sebaceous glands, epidermis, sweat glands, hair follicles, and dermis were isolated from the human skin sections using LMD followed by AmpliSeq transcriptome analysis (Supplementary Fig. 1a). Each region generated distinct clusters when multidimensional scaling (MDS) was performed to analyze the similarity of the transcriptome profile in different regions (Supplementary Fig. 1b). Next, we focused on the genes highly

358 expressed in each region (Supplementary Fig. 1c) (23–38). Genes encoding ELOVL fatty acid
359 elongase 3 and 5 (*ELOVL3* and *ELOVL5*), perilipin 2 and 5 (*PLIN2* and *PLIN5*), and
360 microsomal glutathione S-transferase 1 (*MGST1*) are highly expressed in sebaceous glands (23–
361 26). In our study, these genes were highly expressed in sebaceous glands isolated via LMD than
362 in other regions (Supplementary Fig. 1c). Other regions isolated by LMD also expressed region-
363 characteristic genes, indicating that the LMD was performed successfully.

364 To gain further insight into the characteristics and origin of SSL-RNAs, we explored genes
365 highly expressed in each region isolated by LMD. In each region, genes with a mean of $\log_2$
366 (normalized counts + 1) > 10, and with more than 1.5-fold differential expression compared to
367 other region(s), were selected (Fig. 4). However, using these criteria, we did not find any gene
368 selectively expressed in epidermis, due to its similarity with the hair follicles. Therefore, we
369 listed epidermal genes with more than 1.5-fold differential expression than in sebaceous glands,
370 sweat glands, and dermis. To obtain information regarding the origin of SSL-RNAs, the SSL-
371 RNA profile of 29 healthy subjects was analyzed using the genes expressed characteristically at
372 each region. The results presented in the heat map indicate that SSL-RNAs predominantly
373 comprise mRNAs characteristic for sebaceous glands, epidermis, and hair follicles (Fig. 4).

374 The filaggrin (*FLG*), filaggrin 2 (*FLG2*) and aspartic peptidase retroviral like 1 (*ASPRV1*)
375 that were expressed in the granular layer of epidermis (27–29), were abundantly expressed in

376 SSL-RNAs (Fig. 4). The genes encoding keratin 25, 27, and 71 (*KRT25, KRT27*, and *KRT71*,
377 respectively) that are expressed in the inner root sheath of hair follicles (33, 34) were highly
378 expressed in SSL-RNAs (Fig. 4). These results suggest that SSL-RNAs provide significant
379 information regarding the granular layer of the epidermis and the inner root sheath of hair
380 follicles.
381
382 **Comparison of SSL-RNAs expression between healthy subjects and patients with AD**
383 We analyzed SSL-RNAs from 29 healthy subjects and 30 patients with AD. The major
384 output of the AmpliSeq data obtained with our modified method was as follows: i) the average
385 number of reads was 11,456,318 in healthy subjects and 11,137,677 in patients with AD; ii) the
386 average mapping ratio was 84.0 % in healthy subjects and 92.4 % in patients with AD; and iii)
387 the average of the ratio of target detected genes was 44.8 % in healthy subjects and 50.2 % in
388 patients with AD. First, we verified the expression of genes with significant differential
389 expression in AD. Consistent with previous reports (39–42), our results in SSL-RNAs analysis
390 showed that the expression of *CCL17,* interleukin 1β (*IL1B*), interleukin 13 (*IL13*), and S100
391 calcium binding protein A9 (*S100A9*) significantly increased in patients with AD, but the
392 expression of *FLG* and involucrin (*IVL*) was significantly decreased in the patients with AD
393 compared to healthy subjects (Fig. 5a). In a previous study reporting a global gene expression

394 analysis in skin biopsy samples, genes related to the terminal differentiation of keratinocytes
395 were significantly downregulated and those related to immune-mediated inflammation were
396 upregulated in patients with AD compared with their expression levels in healthy subjects (39).
397 Based on this report, we selected 12 genes related to terminal differentiation and 22 genes
398 related to immune-mediated inflammation that were detected in SSL-RNAs and compared their
399 expression patterns. Our results showed that the expression patterns of these genes in SSL-
400 RNAs of patients with AD and healthy subjects were largely consistent with the previous report
401 (Fig. 5b).

402 Moreover, the analysis of the dimensionality reduction using t-distributed stochastic
403 neighbor embedding (t-SNE) and variance stabilizing transformation (VST) values in all genes
404 showed that the healthy subjects and patients with AD could be distinctly classified into two
405 groups (Fig. 6a). To identify the differential biological functions between healthy subjects and
406 patients with AD, we extracted 918 upregulated and 1,033 downregulated genes in patients with
407 AD (Fig. 6b). In the 833 upregulated genes, GO terms of “mRNA splicing” and “stimulatory C-
408 type lectin receptor signaling pathway” were significantly enriched, while GO term “detection
409 of chemical stimulus involved in sensory perception of smell and keratinocyte differentiation”
410 was enriched in the 951 downregulated genes (Fig. 6c).

411 The sebum secretion is reduced in patients with AD than in healthy individuals (43);

412 however, the molecular mechanism underlying this reduction remains unknown. GO analysis
413 was performed on 46 genes highly expressed in the sebaceous glands selected from the results
414 of the LMD experiment (Fig. 4), which resulted in the enrichment of genes involved in lipid
415 metabolism (Fig. 7a). The expression of 25 genes involved in lipid metabolism (GO:0006629)
416 was downregulated in patients with AD compared to healthy individuals (Fig. 7b). Furthermore,
417 genes encoding peroxisome proliferator-activated receptor alpha (*PPARA*), peroxisome
418 proliferator-activated receptor gamma (*PPARG*), MYC proto-oncogene, bHLH transcription
419 factor (*MYC*), transforming growth factor beta 1 (*TGFB1*), tumor protein p53 (*TP53*), and
420 PR/SET domain 1 (*BLIMP1*) regulate sebocyte differentiation and sebum production *in vivo* and
421 *ex vivo* (44–47). Among these gene, the expression of *TGFB1* was significantly upregulated in
422 patients with AD than in healthy subjects (Fig. 7c).
423

Discussion

In this study, we found that mRNAs of measurable quantity and quality were present in skin surface lipids. Further, for the first time, we established a non-invasive and comprehensive method to profile skin mRNAs using SSLs conveniently collected from the skin surface with an oil blotting film.

A significant quantity of RNases is present on the skin surface, and extreme caution should be taken when handling mRNAs. However, unexpectedly, we found that mRNAs in SSLs escape RNase degradation due to the lipid components of the sebum and can be analyzed by AmpliSeq transcriptome sequencing. We analyzed the lipid components responsible for inhibiting the RNase activity and our results suggested that FFAs, triacylglycerol, and squalene contributed significantly to the RNase inhibitory activity of sebum. Due to inhibitory effects of lipids on RNases, we speculate that mRNAs may be less susceptible to RNases-mediated degradation in the lipid-rich/low-water environment of SSLs. In addition, the optimal pH for RNase activity is 6.5-8.0 (48), whereas the skin surface is generally weakly acidic (pH 4.1 to 5.8) due to the presence of organic acids such as lactic acid (49); hence, these factors may collectively reduce the RNase activity in SSLs. It is known that sebum is secreted by sebaceous glands in the form of fine granules (4–5 nm) (50); however, there is no knowledge on the spatial arrangement of lipids, organic acids, and mRNAs in SSLs. Identifying the distribution of these

components, as well as the molecular interactions between them, is necessary to understand the precise stabilization mechanism of mRNA on the skin.

We established the method for comprehensive analysis of SSL-RNAs. Recent advances in sequence technology have made it possible to analyze even degraded mRNA. To prepare RNA-seq libraries, a minimum DV200 value of 30 % is generally recommended. Using our modified method, the DV200 of human mRNAs in SSLs was approximately 56.5 %, making it suitable for transcriptome analysis. However, when the sequence libraries were prepared according to the standard protocol of Ion AmpliSeq Transcriptome Human Gene Expression Kit, the success rate was very low. Use of the optimized protocol (by standardizing conditions of reverse transcription and target amplification, and adding a purification step after target amplification) led to significant improvement in library production efficiency, and the transcriptome sequencing resulted in a success rate of 95 %. Further, the analytical error of this method was low and the results showed high correlation with qPCR results. Thus, our SSL-RNA analysis method using the improved AmpliSeq protocol enables profiling of the mRNA expression in SSLs in a reliable manner.

We investigated the origin of the SSL-RNAs by comparing their mRNA expression profile with those of different regions of the skin. Holocrine secretion of sebum made us speculate that the expression pattern of SSL-RNAs should be similar to that of the sebaceous glands. Indeed,

460 analysis of different regions of the skin obtained by LMD showed that the mRNAs derived from
461 the sebaceous glands were highly expressed in SSL-RNAs. However, SSL-RNAs were rich not
462 only in mRNAs derived from sebaceous glands but also in those derived from the epidermis and
463 hair follicles, the tissues in close contact with the sebum. In contrast, the mRNAs characteristic
464 of sweat glands and dermis, which are not in close contact with sebum, were absent in SSLs.
465 The mechanism underlying transfer of epidermal mRNAs into SSLs remain unclear, and
466 there could be several possible mechanisms. The mRNAs characteristic of the granular layer of
467 the epidermis (transcribed from *FLG*, *FLG2* and *ASPRV1*) were highly expressed in SSL-RNAs.
468 In addition, stratum corneum is reported to contain detectable amounts of mRNA (13–15). We
469 speculate that the epidermal mRNAs transferred to the surface of the stratum corneum due to
470 keratinization are mixed with the sebum on the skin surface, leading to their presence in SSLs.
471 The SSLs also contained hair follicle-derived RNAs, which may be related to the anatomical
472 features of the hair follicle. SSL-RNAs were rich in mRNAs of *KRT25*, *KRT27*, and *KRT71*, the
473 marker genes for the inner root sheath (33, 34). The inner root sheath detaches from the hair
474 shaft and degrades during hair growth, and the process occurs at the orifice of the sebaceous
475 duct (51). These observations indicate that the epithelial cells of the inner root sheath may get
476 mixed with sebum at the orifice and skin surface, and as a result, information pertaining to the
477 hair follicles may be reflected in the SSL-RNAs. In addition, since extracellular vesicles

478 released from various cells contain several biomolecules including mRNAs (52, 53), it is
479 possible that extracellular vesicles-derived mRNAs are also included in the SSL-RNAs.
480 Collectively, our results indicate that SSL-RNAs predominantly contains mRNAs derived from
481 the sebaceous glands, epidermis, and hair follicles, and are therefore, a useful resource for
482 analyzing the biological information related to the relevant regions of the skin.
483 Finally, we verified the applicability of this method by performing a comparative analysis of
484 the SSL-RNAs profiles of healthy subjects and patients with AD. We observed that the
485 transcriptome profile was markedly different between healthy subjects and patients with AD,
486 with differential expression of immune-mediated inflammation and terminal differentiation-
487 related genes, as shown in a previous skin biopsy report (39). Moreover, the GO term “detection
488 of chemical stimulus involved in sensory perception of smell” identified in our study was
489 consistent with a previous report based on patients with AD (13) indicating that the analysis of
490 SSL-RNAs successfully captured the characteristics of AD. Atrophy of the sebaceous glands
491 and reduction in sebum secretion have been reported in patients with AD (54). However, little is
492 known about the underlying mechanism, including the gene expression profile of sebaceous
493 glands, in patients with AD due to difficulty in obtaining facial skin tissue samples containing
494 sebaceous glands. Here, we showed that the expressions of 25 lipid metabolism-related genes
495 highly expressed in sebaceous glands were lower in AD patients than in healthy subjects.

496 Moreover, the expression of *TGFB1*, which suppresses sebocyte differentiation and lipid
497 accumulation (47), was significantly increased in patients with AD, suggesting that the
498 suppression of lipid synthesis via *TGFB1* may be one of the mechanisms responsible for
499 dysregulated sebum synthesis in these patients. Thus, the transcriptome analysis of SSL-RNAs
500 can evaluate the molecular profile of AD in a non-invasive manner, and is a promising method
501 for comprehensive understanding of AD pathology.

502 In summary, we established a non-invasive method for SSL-RNA analysis that utilizes SSL
503 samples collected by simply wiping the skin surface for less than a minute. This non-invasive
504 method has potential application in unraveling the molecular profile of skin diseases, such as
505 AD, which will be helpful for clinical management of these diseases in future. Understanding
506 the status and course of AD at the molecular level is essential not only to assess the
507 pathophysiology of AD but also to design its effective therapeutic treatments. The clinical
508 phenotypes of AD are extremely complex, warranting the need for identifying biomarkers that
509 can classify invisible endophenotypes (55). Further, since the skin is called “the disease-sensor
510 organ” (56) and is believed to reflect the conditions inside the body, the SSL-RNA analysis may
511 have wide applicability to understand various pathologies of human body.
512

Data availability

The datasets generated and analyzed in the current study are available from the corresponding author on reasonable request.

References

1) Droujinine, I. A. & Perrimon, N. Interorgan communication pathways in physiology: focus on *Drosophila*. *Annu. Rev. Genet.* **50,** 539-570 (2016).

2) Heikenfeld, J. et al. Accessing analytes in biofluids for peripheral biochemical monitoring. *Nat. Biotechnol.* **37,** 407-419 (2019).

3) Yokoi, A. et al. Integrated extracellular microRNA profiling for ovarian cancer screening. *Nat. Commun.* **9,** 4319 (2018).

4) Sudo, K. et al. Development and validation of an esophageal squamous cell carcinoma detection model by large-scale microRNA profiling. *JAMA Netw. Open* **2,** e194573 (2019).

5) Yasui, T. et al. Unveiling massive numbers of cancer-related urinary-microRNA candidates via nanowires. *Sci. Adv.* **3,** e1701133 (2017).

6) Paliwal, S., Hwang, B. H., Tsai, K. Y. & Mitragotri, S. Diagnostic opportunities based on skin biomarkers. *Eur. J. Pharm. Sci.* **50,** 546-556 (2013).

7) Wester, V. L. & van Rossum, E. F. C. Clinical applications of cortisol measurements in

hair. *Eur. J. Endocrinol.* **173,** M1-10 (2015).

8) Gao, W. et al. Fully integrated wearable sensor arrays for multiplexed in situ perspiration analysis. *Nature* **529,** 509-514 (2016).

9) Moyer, J., Wilson, D., Finkelshtein, I., Wong, B. & Potts, R. Correlation between sweat glucose and blood glucose in subjects with diabetes. *Diabetes Technol. Ther.* **14,** 398-402 (2012).

10) Park, J. H. P. et al. An antimicrobial protein, lactoferrin exists in the sweat: proteomic analysis of sweat. *Exp. Dermatol.* **20,** 369-371 (2011).

11) Bengtsson, C. F. et al. DNA from keratinous tissue. Part I: hair and nail. *Ann. Anat.* **194,** 17-25 (2012).

12) Wong, R. et al. Use of RT-PCR and DNA microarrays to characterize RNA recovered by non-invasive tape harvesting of normal and inflamed skin. *J. Invest. Dermatol.* **123,** 159-167 (2004).

13) Dyjack, N. et al. Minimally invasive skin tape strip RNA sequencing identifies novel characteristics of the type 2–high atopic dermatitis disease endotype. *J. Allergy Clin. Immunol.* **141,** 1298-1309 (2018).

14) He, H. et al. Tape strips detect distinct immune and barrier profiles in atopic dermatitis and psoriasis. *J. Allergy Clin. Immunol.* **147,** 199-212 (2021).

549 15) Pavel, A. B. et al. Tape strips from early-onset pediatric atopic dermatitis highlight disease
550 abnormalities in nonlesional skin. *Allergy.* **76,** 314-325 (2021).
551 16) Leung, D. Y. M. et al. The nonlesional skin surface distinguishes atopic dermatitis with food
552 allergy as a unique endotype. *Sci. Transl. Med.* **11**, eaav2685 (2019).
553 17) Peppelman, M., van den Eijnde, W. A. J., Jaspers, E. J., Gerritsen, M-J. P. & van Erp, P. E. J.
554 Combining tape stripping and non-invasive reflectance confocal microscopy : an *in vivo*
555 model to study skin damage. *Skin Res. Technol.* **21**, 474-484 (2015).
556 18) Schneider, M. R. & Paus, R. Sebocytes, multifaceted epithelial cells: lipid production and
557 holocrine secretion. *Int. J. Biochem. Cell Biol.* **42,** 181-185 (2010).
558 19) Köten, B. et al. RNase 7 contributes to the cutaneous defense against Enterococcus
559 faecium. *PLoS One* **4,** e6424 (2009).
560 20) Li, W. et al. Comprehensive evaluation of AmpliSeq transcriptome, a novel targeted whole
561 transcriptome RNA sequencing methodology for global gene expression analysis. *BMC*
562 *Genomics* **16,** 1069 (2015).
563 21) Wang, Z., Lyu, Z., Pan, L., Zeng, G. & Randhawa, P. Defining housekeeping genes suitable
564 for RNA-seq analysis of the human allograft kidney biopsy tissue. *BMC Med. Genomics* **12,**
565 86 (2019).
566 22) Fitzgerald, L. M. et al. Obtaining high quality transcriptome data from formalin-fixed,

paraffin-embedded diagnostic prostate tumor specimens. *Lab. Invest.* **98,** 537-550 (2018).

23) Westerberg, R. et al. Role for ELOVL3 and fatty acid chain length in development of hair and skin function. *J. Biol. Chem.* **279,** 5621-5629 (2004).

24) Kitago, M. et al. Localization of the candidate genes ELOVL5 and SCD1 for 'male effect' pheromone synthesis in goats (*Capra hircus*). *J. Reprod. Dev.* **53,** 1329-1333 (2007).

25) Dahlhoff, M. et al. PLIN2, the major perilipin regulated during sebocyte differentiation, controls sebaceous lipid accumulation in vitro and sebaceous gland size in vivo. *Biochim. Biophys. Acta* **1830,** 4642-4649 (2013).

26) Joost, S. et al. Single-cell transcriptomics reveals that differentiation and spatial signatures shape epidermal and hair follicle heterogeneity. *Cell Syst* **3,** 221-237.e9 (2016).

27) Zhang, Y. et al. Activation of beta-catenin signaling programs embryonic epidermis to hair follicle fate. *Development* **135,** 2161-2172 (2008).

28) Makino, T., Mizawa, M., Yamakoshi, T., Takaishi, M. & Shimizu, T. Expression of filaggrin-2 protein in the epidermis of human skin diseases: a comparative analysis with filaggrin. *Biochem. Biophys. Res. Commun.* **449,** 100-106 (2014).

29) Donovan, M. et al. Filaggrin and filaggrin 2 processing are linked together through skin aspartic acid protease activation. *PLoS One* **15,** e0232679 (2020).

30) Gao, Y. et al. Isolation, culture and phenotypic characterization of human sweat gland

epithelial cells. *Int. J. Mol. Med.* **34,** 997-1003 (2014).

31) Rieg, S., Garbe, C., Sauer, B., Kalbacher, H. & Schittek, B. Dermcidin is constitutively produced by eccrine sweat glands and is not induced in epidermal cells under inflammatory skin conditions. *Br. J. Dermatol.* **151,** 534-539 (2004).

32) Miura, K. et al. Homeobox transcriptional factor engrailed homeobox 1 is expressed specifically in normal and neoplastic sweat gland cells. *Histopathology* **72,** 1199-1208 (2018).

33) Langbein, L. et al. K25 (K25irs1), K26 (K25irs2), K27 (K25irs3), and K28 (K25irs4) represent the type I inner root sheath keratins of the human hair follicle. *J. Invest. Dermatol.* **126,** 2377-2386 (2006).

34) Langbein, L., Yoshida, H., Praetzel-Wunder, S., Parry, D. A. & Schweizer, J. The keratins of the human beard hair medulla: the riddle in the middle. *J. Invest. Dermatol.* **130,** 55-73 (2010).

35) Kutz, W. E. et al. ADAMTS10 protein interacts with fibrillin-1 and promotes its deposition in extracellular matrix of cultured fibroblasts. *J. Biol. Chem.* **286,** 17156-17167 (2011).

36) Le Goff, C. et al. Regulation of procollagen amino-propeptide processing during mouse embryogenesis by specialization of homologous ADAMTS proteases: insights on collagen biosynthesis and dermatosparaxis. *Development* **133,** 1587-1596 (2006).

603 37) Warrick, E. et al. Morphological and molecular characterization of actinic lentigos reveals
604 alterations of the dermal extracellular matrix. *Br. J. Dermatol.* **177,** 1619-1632 (2017).
605 38) Moustou, A. E. et al. Expression of lymphatic markers and lymphatic growth factors in
606 psoriasis before and after anti-TNF treatment. *An. Bras. Dermatol.* **89,** 891-897 (2014).
607 39) Suárez-Fariñas, M. et al. Nonlesional atopic dermatitis skin is characterized by broad
608 terminal differentiation defects and variable immune abnormalities. *J. Allergy Clin.*
609 *Immunol.* **127,** 954-64.e1-4 (2011).
610 40) Nomura, I. et al. Cytokine milieu of atopic dermatitis, as compared to psoriasis, skin
611 prevents induction of innate immune response genes. *J. Immunol.* **171,** 3262-3269 (2003).
612 41) Bianchi, P. et al. Analysis of gene expression in atopic dermatitis using a microabrasive
613 method. *J. Invest. Dermatol.* **132,** 469-472 (2012).
614 42) Gittler, J. K. et al. Progressive activation of T(H)2/T(H)22 cytokines and selective
615 epidermal proteins characterizes acute and chronic atopic dermatitis. *J. Allergy Clin.*
616 *Immunol.* **130,** 1344-1354 (2012).
617 43) Furuichi, M. et al. The usefulness of sebum check film for measuring the secretion of
618 sebum. *Arch. Dermatol. Res.* **302,** 657-660 (2010).
619 44) Trivedi, N. R. et al. Peroxisome proliferator-activated receptors increase human sebum
620 production. *J. Invest. Dermatol.* **126,** 2002-2009 (2006).

621 45) Cottle, D. L. et al. c-MYC-induced sebaceous gland differentiation is controlled by an
622 androgen receptor/p53 axis. *Cell Rep.* **3,** 427-441 (2013).
623 46) Horsley, V. et al. Blimp1 defines a progenitor population that governs cellular input to the
624 sebaceous gland. *Cell* **126,** 597-609 (2006).
625 47) McNairn, A. J. et al. TGFβ signaling regulates lipogenesis in human sebaceous glands cells.
626 *BMC Dermatol.* **13,** 2 (2013).
627 48) Sorrentino, S. The eight human "canonical" ribonucleases: molecular diversity, catalytic
628 properties, and special biological actions of the enzyme proteins. *FEBS Lett.* **584,** 2194-
629 2200 (2010).
630 49) Lambers, H., Piessens, S., Bloem, A., Pronk, H. & Finkel, P. Natural skin surface pH is on
631 average below 5, which is beneficial for its resident flora. *Int. J. Cosmet. Sci.* **28,** 359-370
632 (2006).
633 50) Sheu, H. M., Chao, S. C., Wong, T. W., Yu-Yun Lee, J. & Tsai, J. C. Human skin surface
634 lipid film: an ultrastructural study and interaction with corneocytes and intercellular lipid
635 lamellae of the stratum corneum. *Br. J. Dermatol.* **140,** 385-391 (1999).
636 51) Sundberg, J. P. et al. Asebia-2J (Scd1ab2J): A new allele and a model for scarring
637 alopecia. *Am. J. Pathol.* **156,** 2067-2075 (2000).
638 52) Xiao, D. et al. Identifying mRNA, microRNA and protein profiles of melanoma

639 exosomes. *PLoS One* **7,** e46874 (2012).

640 53) Wu, C.-X. & Liu, Z.-F. Proteomic profiling of sweat exosome suggests its involvement in

641 skin immunity. *J. Invest. Dermatol.* **138,** 89-97 (2018).

642 54) Shi, V. Y. et al. Role of sebaceous glands in inflammatory dermatoses. *J. Am. Acad.*

643 *Dermatol.* **73,** 856-863 (2015).

644 55) Muraro, A. et al. Precision medicine in patients with allergic diseases: airway diseases and

645 atopic dermatitis-PRACTALL document of the European Academy of Allergy and Clinical

646 Immunology and the American Academy of Allergy, Asthma & Immunology. *J. Allergy*

647 *Clin. Immunol.* **137,** 1347-1358 (2016).

648 56) Kozawa, S. et al. The body-wide transcriptome landscape of disease models. *iScience* **2,**

649 238-268 (2018).

650

Author Contributions

T.I. conceived the study. T.I., A.H., Y.T., and T.M. planned the study. T.I., T.K., Y.U., M.Y., and N.Oy. performed the experiments and analyzed the data. T.M., Y.T., and N.Ot. supervised the research. T.I. and T.M. wrote the manuscript. All authors reviewed the manuscript.

Competing Interests statement

A patent application related to this work has been filed (No. PCT/JP2017/021040,: “Method for preparing nucleic acid sample.” Status: patent granted (DE, FR, GB, KR, JP), patent pending (CN, US). Inventors: T.I. and A.H. Patent applicant: Kao Corporation). All other co-authors declare that they have no competing interests.

Figure 1

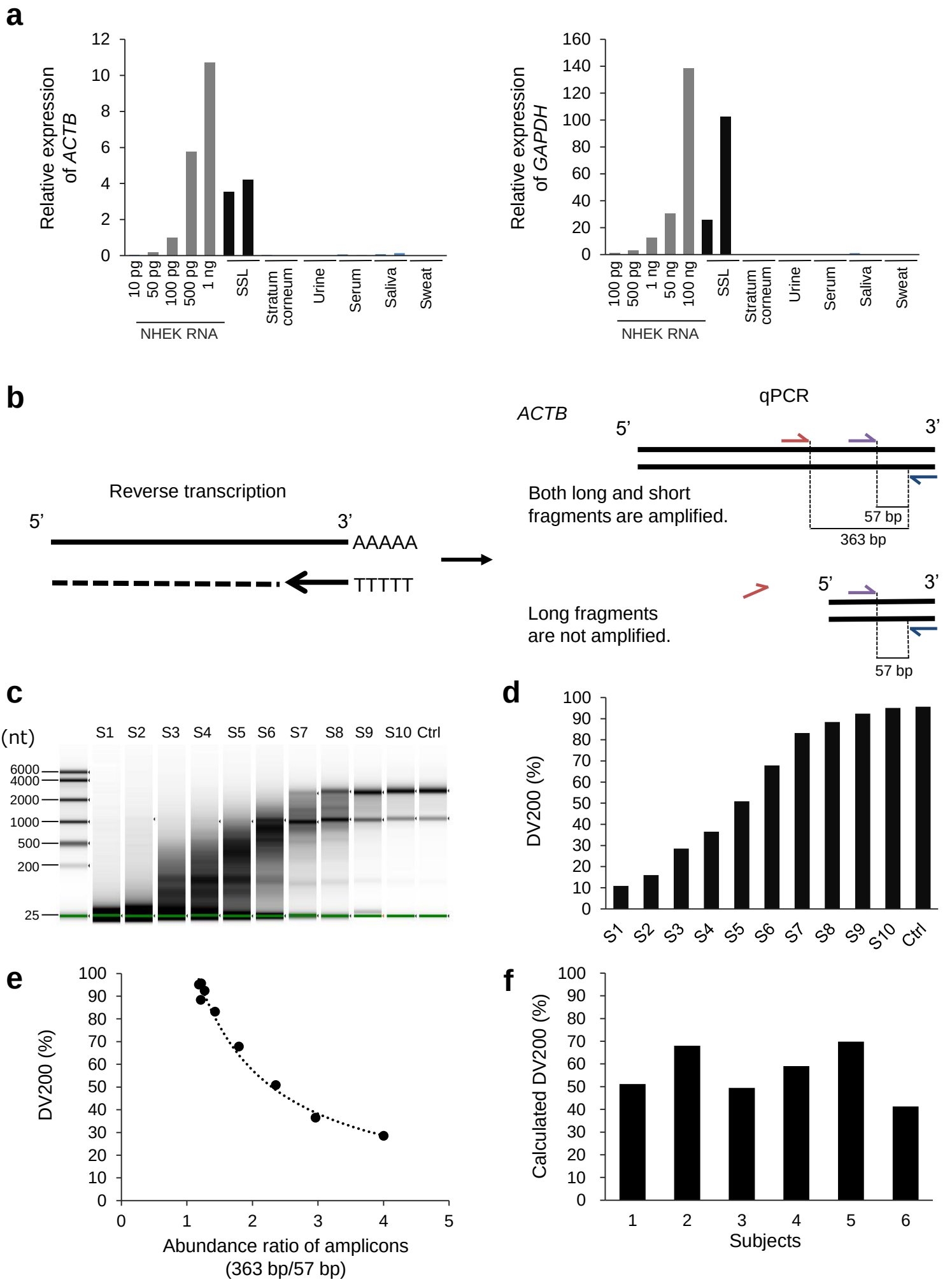


Figure 1. Evaluation of mRNA expression and RNA degradation in skin surface lipids (SSLs)

(a) Expression of *ACTB* and *GAPDH* mRNA in SSLs, stratum corneum, urine, serum, saliva, and sweat samples analyzed by qPCR. Gene expression is shown as the relative expression for 100 pg of RNA derived from normal human epidermal keratinocytes (NHEK). NHEK total RNA (10 pg to 100 ng) is the standard used. **(b)** Outline for assaying the degradation of human mRNA in SSLs. The extent of mRNA degradation was determined using the long (363 bp)/short (57 bp) amplicon ratio calculated from qPCR results. **(c)** Preparation of RNA samples with different levels of degradation. A series of standards (S1 to S10) are presented with known level of RNA degradation. **(d)** DV200 value indicating the percentage of fragments containing > 200 nucleotides, of 1**c**. **(e)** The relationship between the DV200 of 1**d** and the value of 363 bp/57 bp (abundance of 363 bp amplicon/abundance of 57 bp amplicon measured by qPCR). **(f)** The indirect assessment of the human mRNA degradation in SSLs, n = 6. DV200 of each SSL-RNA sample was calculated using the standard curve of 1**e**.

Figure 2

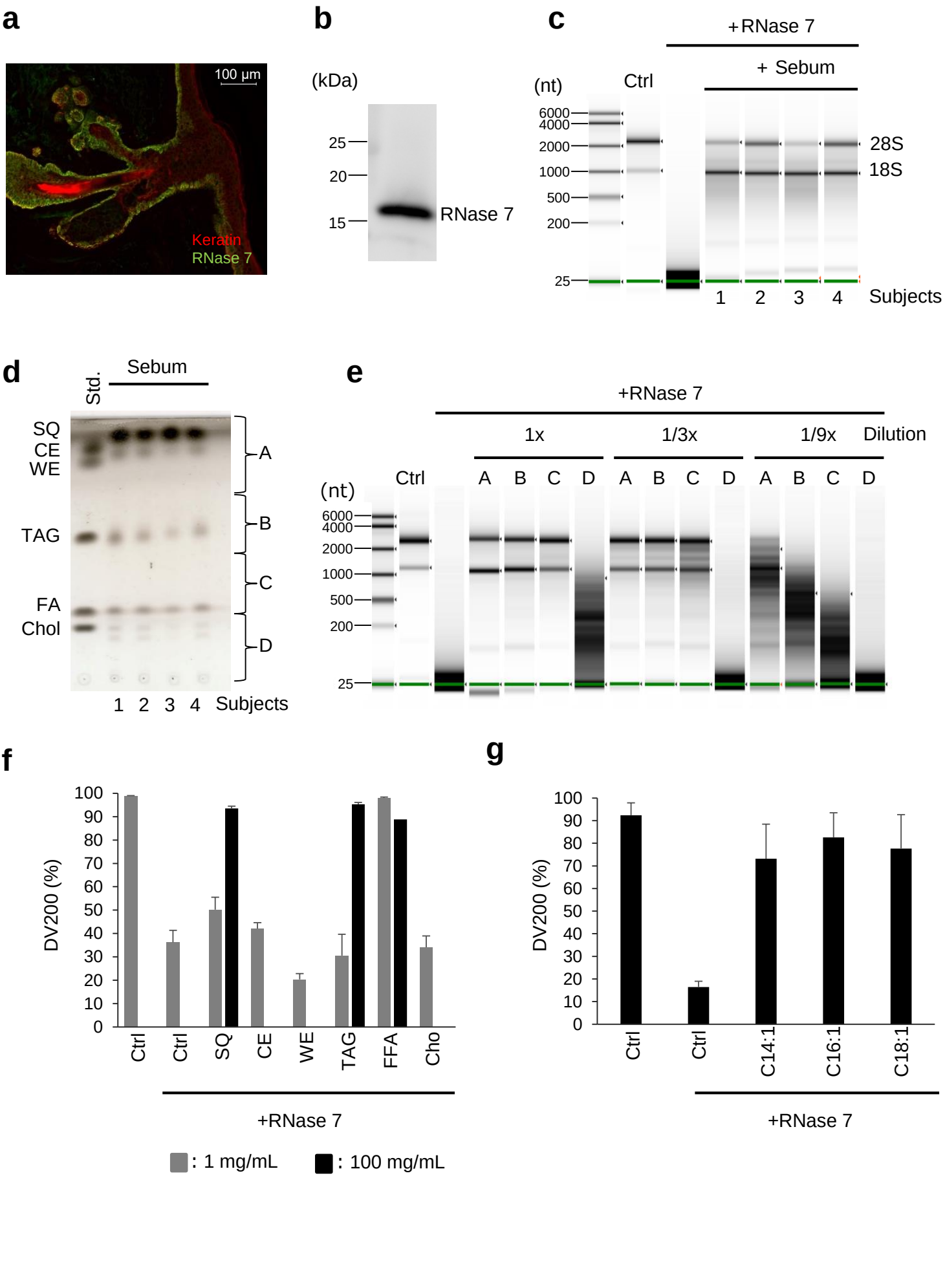


Figure 2. Effect of sebum lipids on RNase activity

(a) The localization of RNase 7 in the human skin. Green, RNase 7; red, keratin/cytokeratin. Bar: 100 μm. **(b)** The detection of RNase 7 in SSLs collected from healthy males by western blotting. **(c)** Determination of the effect of sebum lipids collected from four healthy males on RNase activity using NHEK total RNA. **(d)** Fractionation of sebum lipids collected from four healthy males by performing thin-layer chromatography (TLC). The standard lane (Std) includes authentic samples: SQ, squalene; CE, cholesterol ester (cholesteryl palmitate); WE, wax ester (lauryl palmitoleate); TAG, triacylglycerol (glyceryl trioleate); FA, free fatty acid (palmitoleic acid); and Chol, cholesterol. A to D sebum samples were collected, and lipids were extracted from the silica gel for the subsequent assays. **(e)** Effect of pooled sebum lipids collected from A to D on RNase activity using NHEK RNA. **(f)** Effect of each lipid on RNase activity using NHEK RNA. The DV200 values are shown as mean ± SE, n = 3. WE, wax ester (behenyl palmitoleate); TAG, triacylglycerol (glyceryl tripalmitoleate); Ctrl; EtOH. **(g)** Relationship between RNase activity and chain length of fatty acids. The DV200 values are shown as the mean ± SE, n = 6. C14:1, myristoleic acid; C16:1, palmitoleic acid; C18:1, oleic acid; Ctrl; EtOH.

Figure 3

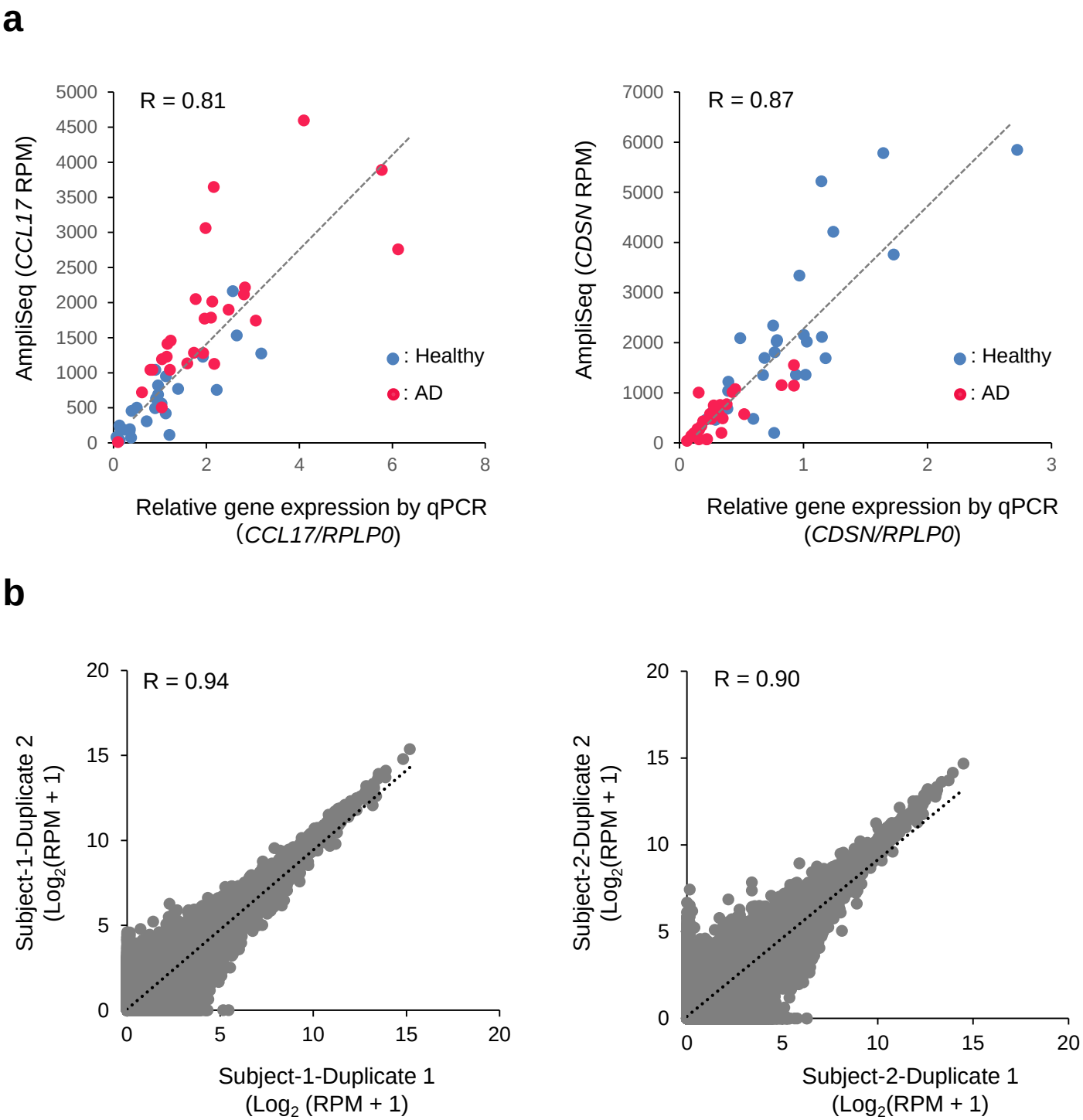


Figure 3. Accuracy of the AmpliSeq data output using the modified protocol

(a) The correlation between AmpliSeq and qPCR results for *CCL17* and *CDSN* expression in healthy subjects (n = 29, blue) and patients with AD (n = 30, red). **(b)** The correlation of the transcriptome profile when preparing libraries in duplicate from each SSL-RNA obtained from two healthy male subjects.

Figure 4

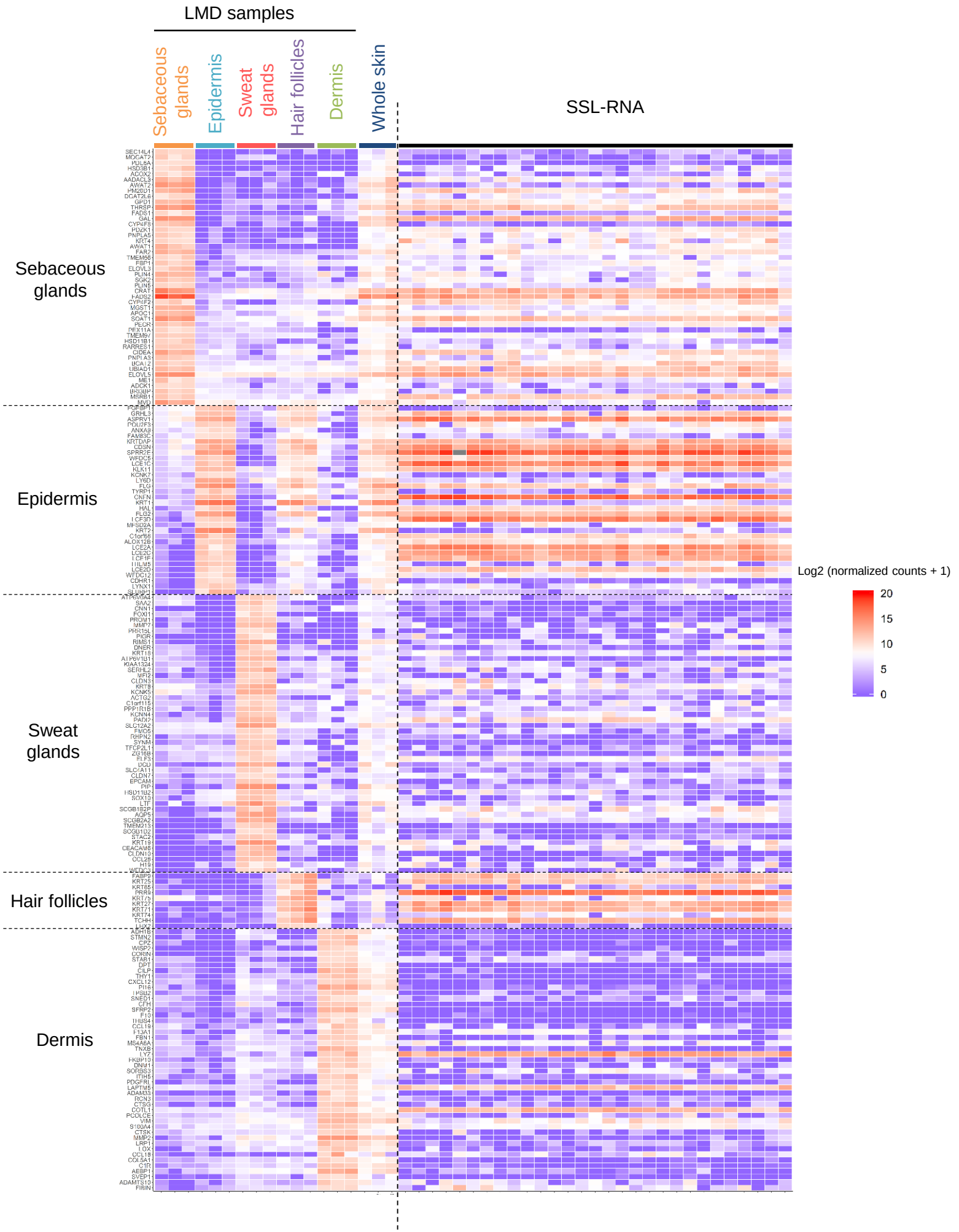
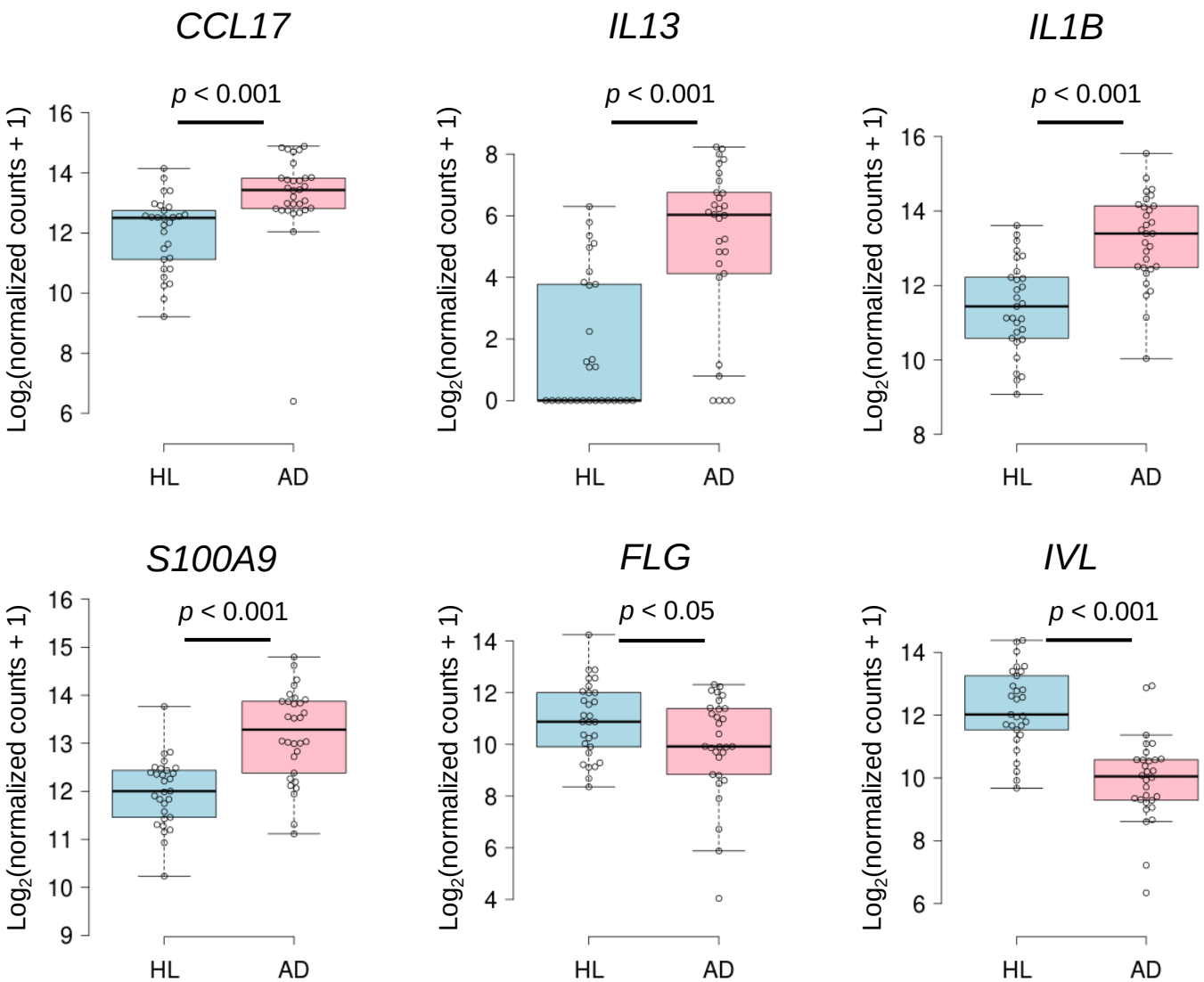


Figure 4. mRNA expression characteristic of different skin regions and its comparison with the expression profile of SSL-RNAs

Heatmap showing the expression profiles for each region (sebaceous glands, epidermis, sweat glands, hair follicles, and dermis) isolated using LMD, whole skin from three healthy male subjects, and SSL-RNAs obtained from 29 healthy male subjects.

Figure 5

a



b

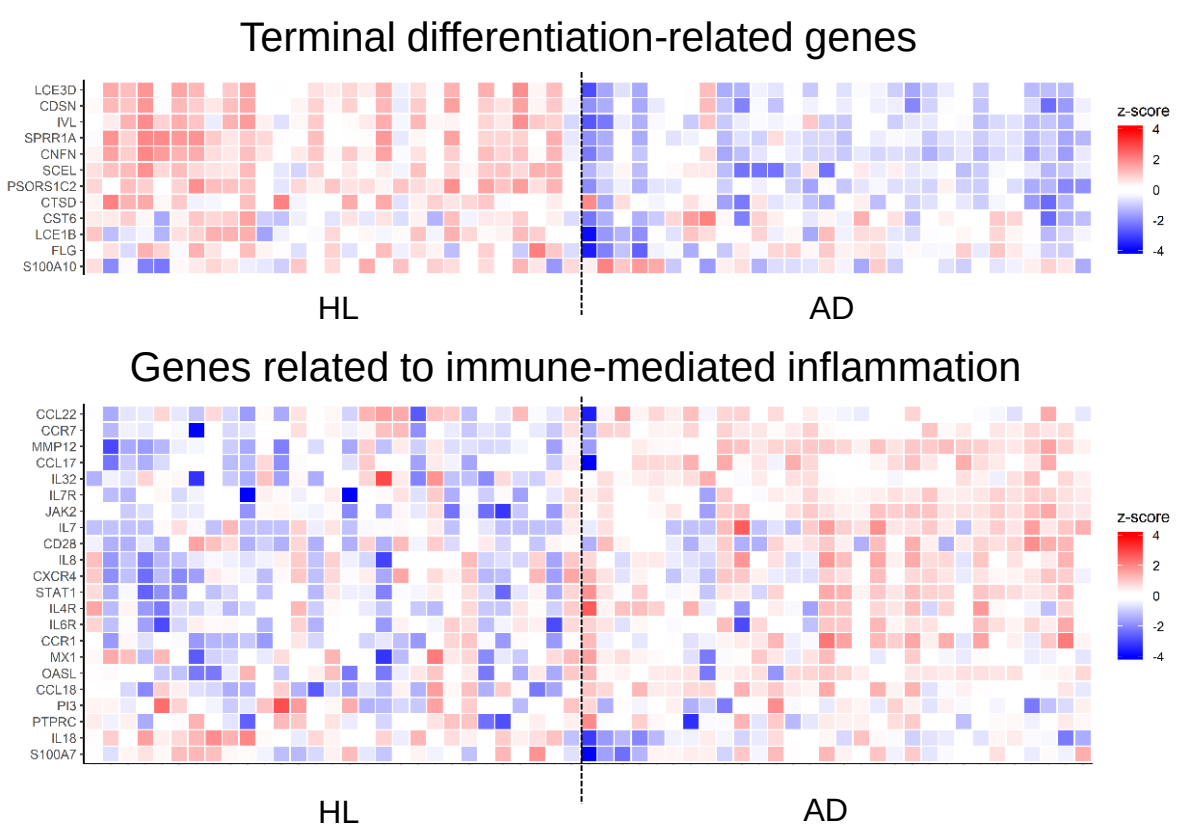


Figure 5. Comparison of AD marker genes in healthy subjects (HL) and patients with AD

(a) The differential expression of *CCL17*, *IL13*, *IL1B*, *S100A9*, *FLG*, and *IVL* in HL and AD. Boxes represent mean ± interquartile range (IQR), and whiskers represent 1st and 3rd quartile 1.5 * IQR. Benjamini-Hochberg adjusted *p*-values are shown from the likelihood ratio test between HL and AD. HL (n = 29), AD (n = 30). **(b)** Heatmaps using z-transformed $\log_2$ (normalized counts + 1) in 12 terminal differentiation-related genes (upper) and 22 genes related to immune-mediated inflammation (lower).

Figure 6

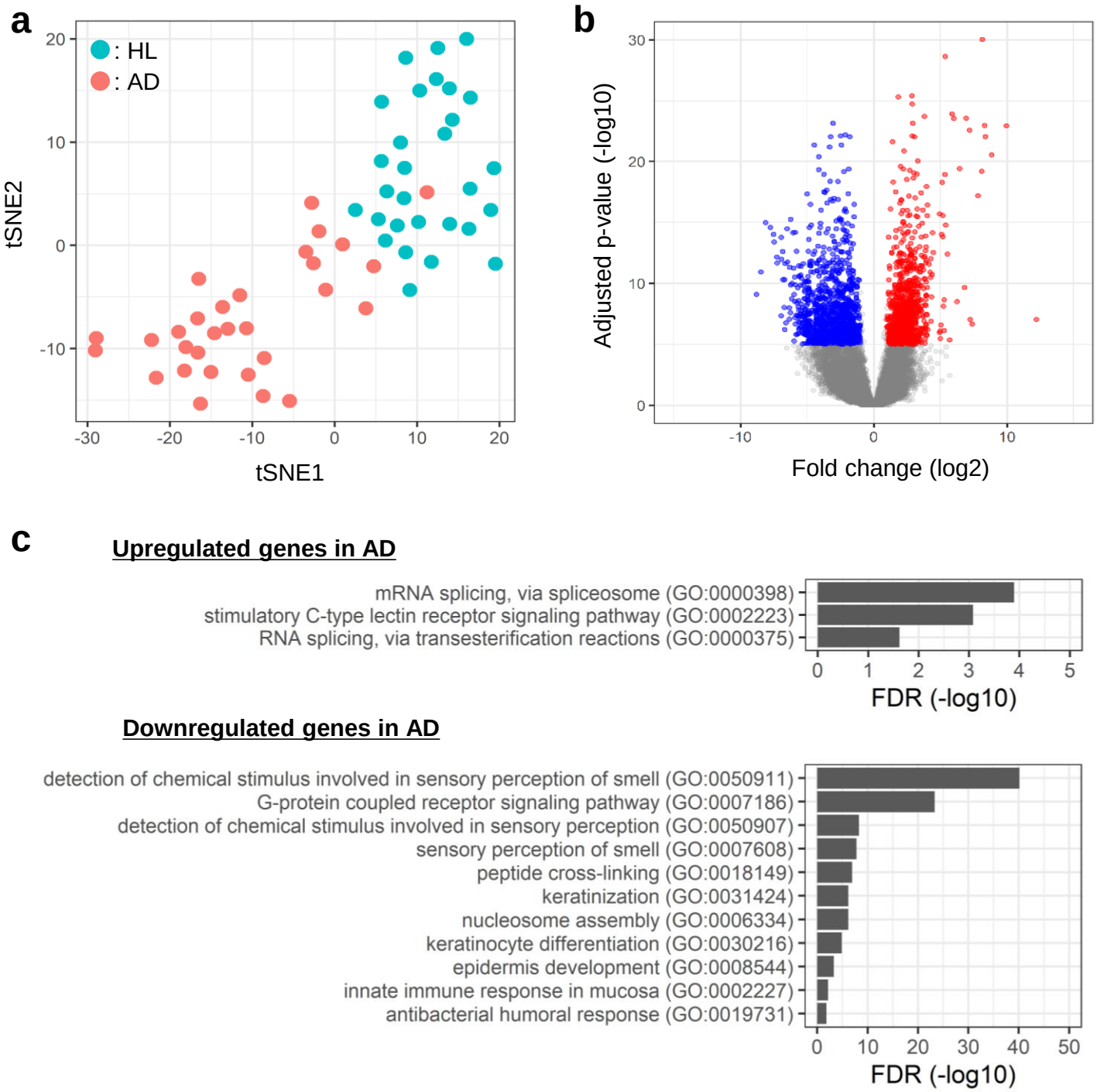
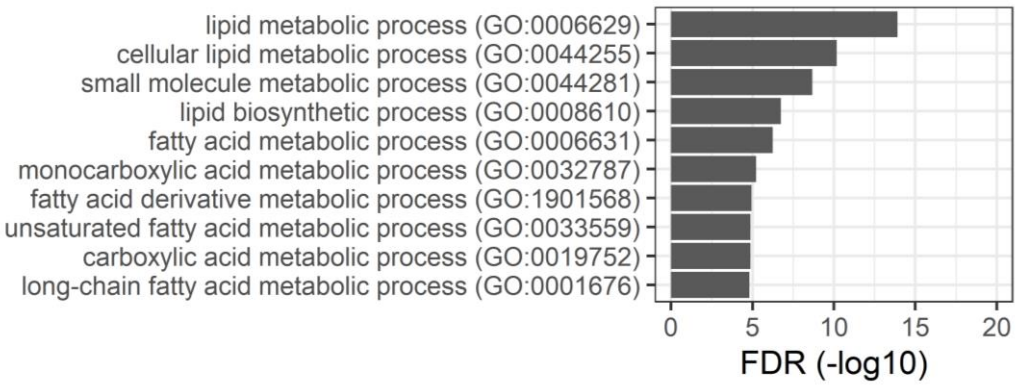


Figure 6. Characterization of SSL-RNAs profiles in healthy subjects (HL) and patients with AD

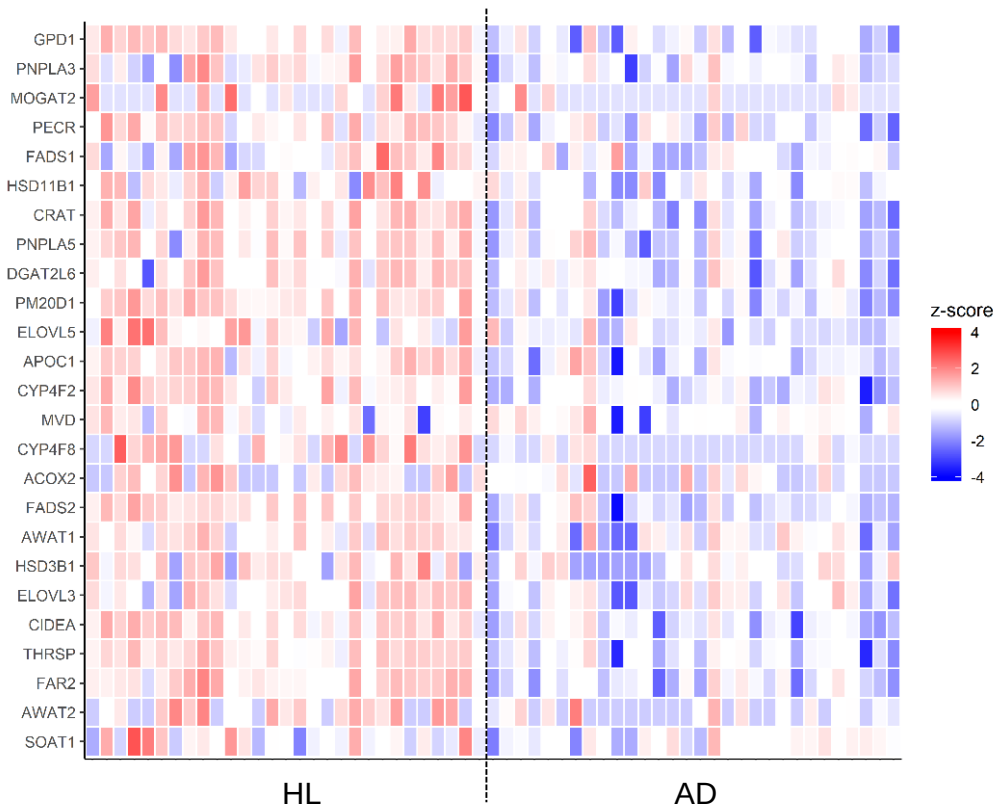
(a) t-SNE analysis using variance stabilizing transformation (VST) values for all genes (green, HL; red, AD). **(b)** Volcano plot of differentially expressed genes (DEGs) (red, upregulated; blue, downregulated) in patients with AD compared to HL subjects (Benjamini-Hochberg adjusted *p*-value < 10^{-5} and fold change > 2.0). **(c)** Gene ontology analysis of DEGs. The upper panel shows significant biological process (BP) of upregulated DEGs and the lower panel shows BP of downregulated DEGs in patients with AD (FDR < 0.05).

Figure 7

a



b



c

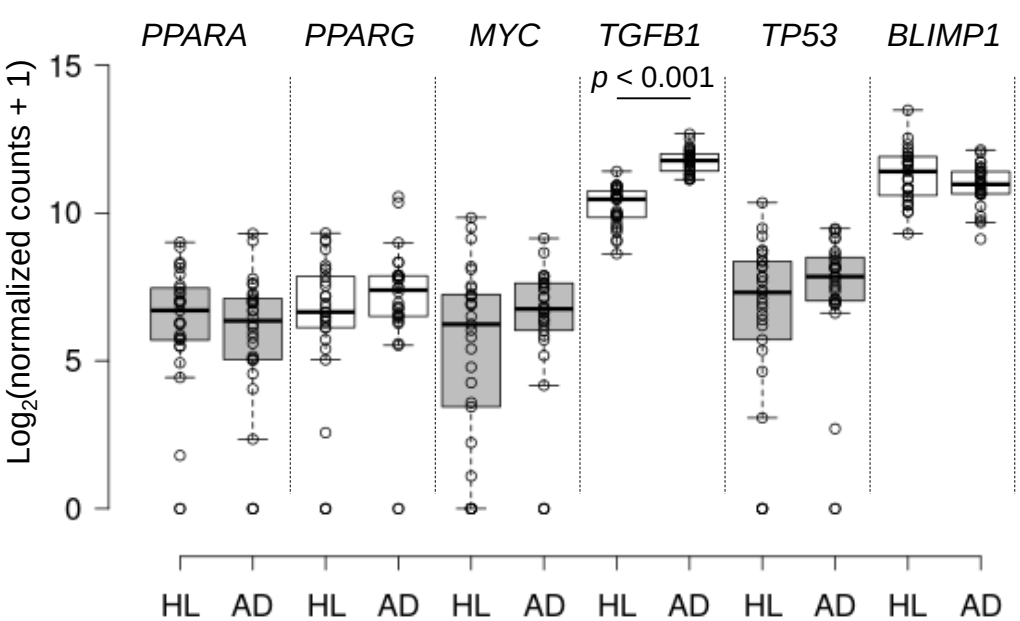


Figure 7. Comparison of SSL-RNAs profile representing highly expressed genes in sebaceous glands in healthy subjects (HL) and patients with AD

(a) Gene ontology analysis of 25 genes highly expressed in sebaceous glands (selected in Fig. 4). **(b)** Heatmaps using z-transformed $\log_2$ (normalized counts + 1) of 25 genes highly expressed in sebaceous glands. **(c)** The differential expression of *PPARA, PPARG, MYC*, *TGFB1*, *TP53*, and *BLIMP1* in HL and patients with AD. Boxes represent mean $\pm$ interquartile range (IQR), and whiskers represent 1st and 3rd quartile 1.5 * IQR. Benjamini-Hochberg adjusted *p*-values are shown from the likelihood ratio test between HL and AD. HL (n = 29), patients with AD (n = 30).