SCIENTIFIC REPORTS

Correspondence and
requests for materials
should be addressed to
M.-K.K. (kimmk@
chungbuk.ac.kr) or
S.W.K. (swkwon@snu.
ac.kr)

* These authors
contributed equally to
this work.

Functional interpretation of metabolomics data as a new method for predicting long-term side effects: treatment of atopic dermatitis in infants

Seul Ji Lee[1]*, Sung-il Woo[2]*, Soo Hyun Ahn[3], Dong Kyu Lim[1], Ji Yeon Hong[1], Jeong Hill Park[1], Johan Lim[3], Mi-kyeong Kim[4] & Sung Won Kwon[1]

[1]College of Pharmacy, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea, [2]Department of Pediatrics, College of Medicine, Chungbuk National University, 52 Naesudong-ro, Heungdeok-gu, Cheongju 361-763, Korea, [3]Department of Statistics, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 151-742, Korea, [4]Department of Internal Medicine, College of Medicine, Chungbuk National University, 52 Naesudong-ro, Heungdeok-gu, Cheongju 361-763, Korea.

Topical steroids are used for the treatment of primary atopic dermatitis (AD); however, their associated risk of serious complications is great due to the presence of vulnerable lesions in young children with AD. Topical calcineurin inhibitors (TCIs) are steroid-free, anti-inflammatory agents used for topical AD therapy. However, their use is prohibited in infants <2 years of age because of their carcinogenic potential. We conducted a randomized, double-blind trial to evaluate the efficacy of TCIs as a secondary AD treatment for children <2 years of age by comparing 1% pimecrolimus cream with 0.05% desonide cream. We performed urinary metabolomics to predict long-term side effects. The 1% pimecrolimus cream displayed similar efficacy and exceptional safety compared with the 0.05% desonide cream. Metabolomics-based long-term toxicity tests effectively predicted long-term side effects using short-term clinical models. This applicable method for the functional interpretation of metabolomics data sets the foundation for future studies involving the prediction of the toxicity and systemic reactions caused by long-term medication administration.

Atopic dermatitis (AD) is the most common chronic disease in children, and its prevalence is 10–20% in young children[1]. It is more prevalent in younger children than school-age children, and the occurrences of clinical symptoms and eczema primarily appear before two years of age. AD affects the quality of life of children and their parents, leading to significant social, economic and health care costs.

AD attacks specific regions of the body depending on age; in very young children, it particularly occurs around the outer regions of the face and limbs[2]. Currently, the majority of children use topical steroids to treat these conditions[3]. However, the use of steroids in children is limited due to the development of complications, such as skin atrophy and telangiectasia, which are related to thinned skin. In particular, the side effects occurring in facial areas cause serious cosmetic problems and result in additional future complications, such as a loss of self-confidence and cosmetic treatment costs.

As an alternative, topical calcineurin inhibitors (TCIs) have been widely used to treat AD in young children because they lack the side effects of topical steroids[4]. However, the US Food and Drug Administration (FDA) issued a boxed warning for TCIs in 2005 because of the carcinogenic potential of immunosuppressants. The use of TCIs in children younger than 2 years of age was restricted in 2006[5]. However, according to the data published to date, it remains unclear whether TCI use increases cancer risk. Therefore, the need to verify TCI efficacy and safety through clinical trials and elucidate its association with cancer risk is evident.

Previous pharmacokinetic-based studies of TCI safety and its relevance to carcinogenic effects have had various limitations. In earlier animal experiments, toxicity has been evident in potential target organs, such as the kidney, pancreas and lymphatic systems, following the oral administration of TCIs[6]. This toxicity is exclusively observed when the blood concentration exceeds 100 ng/mL. However, conventional methods are not appropriate for the prediction of the toxicity of topically administered TCIs. When a TCI formulation is applied to the skin,

undetectable or low concentrations are often measured in the blood[7]. Therefore, an advanced method to predict TCI toxicity following topical administration for AD treatment is inevitably required.

This study aimed to assess an alternative method to earlier approaches to predicting the side effects of immunosuppressive agents based on the functional interpretation of metabolomics data[8]. The greatest advantage of metabolomics is that alterations in metabolites, which are the ultimate products of gene expression, are more extensive and sensitive than are those of DNA, RNA and proteins[9]. Therefore, even if no changes in DNA, RNA or protein levels are evident, significant alterations in metabolite levels can still be observed.

In a generalized, short-term (6 months) clinical trial, the efficacy and side effects of a TCI formulation, pimecrolimus, was compared with the topical corticosteroid desonide. We detected metabolites in urine samples and selected those that were meaningful using statistical analysis optimized for experimental models. The output from the statistical analysis included a long list of metabolites that displayed noteworthy patterns under the various conditions. To interpret the results, group-based significance tests and pathway analysis[10] were performed, and side effects, including carcinogenicity, were determined.

This study demonstrates that metabolomics can be further applied to assess actual clinical data by functional interpretation rather than single marker detection. Our results suggest that metabolomics can potentially serve to predict the toxicity and systemic reactions of long-term medication administration. This study provides scientific evidence for the safe usage of medication, and it will further contribute to improving the quality of life of patients and caregivers by allowing for the determination of its appropriate treatments.

Methods

This randomized, double-blind trial was conducted for 6 months at a single center. The aim of the study was to compare the efficacy and tolerability of 1% pimecrolimus cream (Elidel®) and 0.05% desonide cream (DesOwen®) in infants younger than 2 years of age. The Institutional Review Board or Ethics Committee of the participating organizations approved this study protocol (IRB Registration Number: 2011-09-067). The study was performed according to the Declaration of Helsinki and Good Clinical Practices. Prior to study enrollment, the entire process was clearly explained to each patient's parent, caregiver or legal representative, and a concise written informed consent form was obtained.

Patients. We targeted male and female infants with AD who were 3–23 months of age for this study. The inclusion criteria were as follows: the presence of AD for greater than 3 months; a lesion covering over 5% of the patient's body surface area involving the face, head and neck; an Investigator's Global Assessment (IGA) score of greater than 2; and persistent AD lesions despite the use of existing potency class VII or greater topical steroids for more than two weeks. The key exclusion criteria included the following: inability to receive local formulation treatment due to the presence of an external wound; the presence of other infections, diseases or skin disorders; treatment with a systemic steroid, systemic immunosuppressant, or oriental herbal medicine within the previous 3 months; and general inappropriateness for the study as determined by the clinical trial investigator.

Study design. AD patients were managed with a standardized protocol. The skin management protocol was based on clinical practice guidelines for similar conditions. Prospective subject screening for the inclusion/exclusion criteria was completed 3 days before the baseline visit. In the double-blind phase, 1% pimecrolimus cream was applied to the AD lesions of one patient group, whereas 0.05% desonide cream was applied to the lesions of the other group. The treatments were applied twice daily (every 12 h) for 4 weeks or until the conclusion of the double-blind phase.

If clinical improvement was evident within the entire lesion (IGA score ≤ 1) prior to completion of the 4 weeks of treatment, the patients were switched to an open-label phase. Upon conclusion of the double-blind phase, a 24-week extension was initiated to assess long-term efficacy and safety. In the open-label phase, all patients applied the corresponding medication to their lesion based on authorization and off-label approval if topical administration was required due to the worsening of lesions.

AD severity, skin condition, medication usage and side effects were examined during the double-blind phase and open-label phase. At baseline and week 4 or upon completion of the double-blind phase, urine cortisol and morning plasma-free cortisol levels were measured and compared to confirm the presence of any endocrine-related side effects resulting from the topical steroid treatment. In addition, we compared the alterations in the IGA and EASI scores and the lesion to body surface area (BSA) ratio between the two groups.

Clinical evaluation and analysis. Study visits were performed at screening, baseline, and weeks 1, 2, 4, 8, 16 and 24. The IGA score served as the primary efficacy parameter. An IGA score of 0 (clear) or 1 (almost clear) was the endpoint. Upon achievement of an IGA score of 0 or 1, the patient was switched to the open-label challenge. The Eczema Area and Severity Index (EASI) score served as a secondary efficacy parameter. The clinical evaluation included the recording of individual data, such as date of birth, weight and height; vital signs, such as blood pressure and pulse rate; the use of topical agents; the IGA and EASI scores and other factors; and the administration of a survey about medical history and diet.

Safety assessments consisted of the monitoring and recording of all adverse events as well as dermoscopic assessments of face, head and neck lesions for skin atrophy and telangiectasia. Analysis of endogenous metabolites was also performed from the baseline. The following parameters were monitored and analyzed at baseline, the end of the double-blind phase and 24 weeks: hematology, blood chemistry, morning serum cortisol, T cell subset and urinalysis.

Metabolite extraction. Urine samples from each patient were stored at −80 °C until analysis. For the analysis, 150 μL of each urine sample was centrifuged at 16,000 rcf at 4 °C for 5 min. Next, 100 μL of supernatant was transferred to a new tube and mixed with 400 μL of methanol. After 5 min at −20 °C, the sample was centrifuged again at 16,000 rcf at 4 °C for 5 min, and 100 μL of supernatant was used as an analytical sample.

Ultraperformance liquid chromatography/quadrupole time-of-flight mass spectrometry (UPLC/QTOFMS). Analysis was performed using a UPLC system (Waters Corporation, Milford, CT, USA) equipped with a binary pump, autosampler, degasser and column oven. Reverse-phase LC separation was performed using an ACQUITY UPLC BEH C18 column (2.1 × 100 mm, 1.7 μm). For this procedure, 5 μL of each sample was injected and separated by an optimized gradient method. The mobile phase was water (J.T Baker) + 0.1% formic acid (A) and acetonitrile (J.T Baker) + 0.1% formic acid (B). The following gradient was applied: 0 min of 100% A, 5 min of 70% A, 15 min of 30% A, 25 min of 20% A and ~35 min of 0% A. The flow rate was set to 0.2 mL/min, and the column temperature was maintained at 40 °C. For data quality control, a quality control sample (a mixture of every sample) was injected at regular intervals in the middle of the sequence. In addition, a random sequence was used. We performed equilibration for 10 min to ensure for column equilibration. This process was performed using HyStar chromatography software (Bruker Daltonics, Bremen, Germany).

ESI-MS analysis was performed on a micrOTOF-QII mass spectrometer (Bruker Daltonics, Bremen, Germany). Mass spectrometry was conducted using the ESI positive ion mode. High-purity nitrogen gas was used as the collision and drying gases. The mass range of the instrument was set to 50~1000 m/z. The following MS parameters were used: 8.0 L/min dry gas flow rate at 200 °C of dry temperature, 4.5 kV capillary voltage, 1.2 bar nebulizer pressure, 15.0 eV quadrupole ion energy, 10.0 eV collision energy, 55.0 μs transfer time and 2.0 μs pre-pulse storage. The data were collected with DataAnalysis 4.0 (Bruker Daltonics, Bremen, Germany).

Data preprocessing. Raw data were analyzed based on MZmine version 2.5[11]. Peak detection was carried out by mass detection, chromatogram building and peak deconvolution. During the mass detection process, the conversion of each MS spectrum into a pair of m/z and intensity values (the noise level was set to 10^3, centroid algorithm) was performed. Using the chromatogram builder function, a chromatogram was created with the following parameter values: 0.1 m/z tolerance and $2 × 10^3$ minimum height. To achieve peak deconvolution, peak recognition was completed by the Savitzky-Golay function (minimum peak height of $2 × 10^3$ and derivative threshold level of 80%). For peak alignment, the RANdom SAmple Consensus (RANSAC) aligner method was used (m/z tolerance of 0.05, absolute retention time (RT) tolerance of 0.2 min, and RANSAC iterations of 10,000). The list of obtained peaks was converted into a .csv file.

Statistical analysis. We performed statistical analyses to test whether there was a significant change over time in metabolite levels for the patient group who applied 1% pimecrolimus cream and for the group who applied 0.05% desonide cream. To determine metabolite levels, we used the mixed effects model with fixed effects for time (weeks) and random effects for patients. Finally, we implemented χ^2-tests individually to identify the metabolites that changed over time for each group.

Functional interpretation of metabolomics data. Using MetaboAnalyst[12], we confirmed the mechanisms and diseases related to the metabolites affected by each agent. In general, further research aimed at marker selection was necessary for highly reliable metabolite selection at a significance level of approximately 0.05. However, in this study, because we attempted to widely confirm the involved mechanisms to predict every potential side effect, the metabolites with α < 0.2 were included in the functional analysis. Throughout this process, all of the mechanisms were monitored, even if only a slight alteration was observed. The data found to be statistically significant were exclusively interpreted. MetaboAnalyst (http://www.metaboanalyst.ca) was used as an analytical tool for interpretation. For the metabolites that displayed altered levels in response to 1% pimecrolimus cream and 0.05% desonide cream, we also applied metabolite set enrichment analysis (MSEA) and metabolic pathway analysis (MetPA)[12,13]. For this process, we selected the 'Disease-associated metabolite set (Urine)' option to confirm the results by targeting the metabolites in urine with an emphasis on disease relevance. The results are presented in the tables consisting of 8–

Table 1	Baseline demographics (*: 95% Confidence Interval)		
	Pimecrolimus (n = 28)	Desonide (n = 27)	p-value
Age; years, mean ± SD	0.7 ± 0.3	0.7 ± 0.5	0.971
Gender; male: female	17:11	20:7	0.291
IGA score (95% CI*)	3.9 (3.5–4.2)*	3.6 (3.2–4.0)	0.326
EASI score (95% CI*)	13.1 (10.6–15.6)	12.9 (10.3–15.4)	0.893
Area of lesion to BSA; % (95% CI*)	27.7 (23.3–32.2)	33.3 (27.0–39.5)	0.146
Total IgE count; IU/mL	176 ± 333	70 ± 209	0.163
Food Allergen Sensitization; n (%)	19 (68%)	14 (52%)	0.226
Significant Sensitization; n (%)	12 (43%)	5 (19%)	0.080

IGA: Investigator's Global Assessment total eosinophil;
EASI: Eczema Area and Severity Index;
BSA: Body Surface Area.

9 columns which include the metabolite set/pathway, the total number of metabolites in that set (Total), the number of metabolites in the experimental set matching the pathway set (Hits), the expectation value (Expected), the p-value, the Holm p-value, the false discovery rate (FDR) and additional details.

Results

Patients and treatment. Of the 95 patients screened, 55 who agreed to undergo the clinical study and complete a washout period were targeted for the double-blind study. During the study, one of the patients in the pimecrolimus group was withdrawn because of nonattendance at a scheduled visit; therefore, a total of 54 patients completed the randomized, double-blind study. No statistically significant differences were noted between the 2 groups based on gender, age, the presence of other atopic manifestations, sensitization to food allergens, total IgE levels, blood eosinophil count, IGA scores or EASI scores.

The baseline patient demographics are presented in Table 1. At the initiation of treatment, the mean IGA scores (95% confidence interval, CI) for the pimecrolimus and desonide treatment groups were 3.9 (3.5–4.2) and 3.6 (3.2–4.0), respectively. Moreover, the mean EASI scores for the pimecrolimus and desonide treatment groups were 13.1 (10.6–15.6) and 12.9 (10.3–15.4), respectively. Thus, no difference between the two groups was observed.

Because of withdrawals due to loss to follow up (one for the pimecrolimus group and two for the desonide group), 26 and 25 patients in the pimecrolimus and desonide groups, respectively, were targeted for the 6-month open-label phase (Figure 1).

Efficacy. Upon completion of the double-blind study, no statistically significant difference in IGA scores between the two groups was observed (Table 2). We measured the recovery rate according to the criterion of 'clear or almost clear lesion with IGA 0–1' in 54 patients. A total of 11 patients (41%) from the pimecrolimus and desonide groups displayed the same recovery rates until the end of the double-blind study. However, the EASI score, which reflects the extent and severity of the lesion, displayed significant differences at weeks 1 and 2 as well as at the end of the double-blind phase. Although the lesion to BSA ratio for the desonide group was reduced compared with that of the pimecrolimus group, no statistically significant difference ($p < 0.05$) was observed.

At the end of the open-label phase of the study, 14 patients (54%) from the pimecrolimus group and 13 (52%) from the desonide group displayed an IGA score of 0–1, with no significant difference. The facial IGA scores of the pimecrolimus and desonide groups were 1.5 (1.0–2.0) and 1.6 (1.3–2.0), respectively. The EASI scores for the two groups were 3.1 (1.7–4.6) and 2.3 (1.4–3.3), respectively. No significant differences were observed for both scores between the two treatment groups.
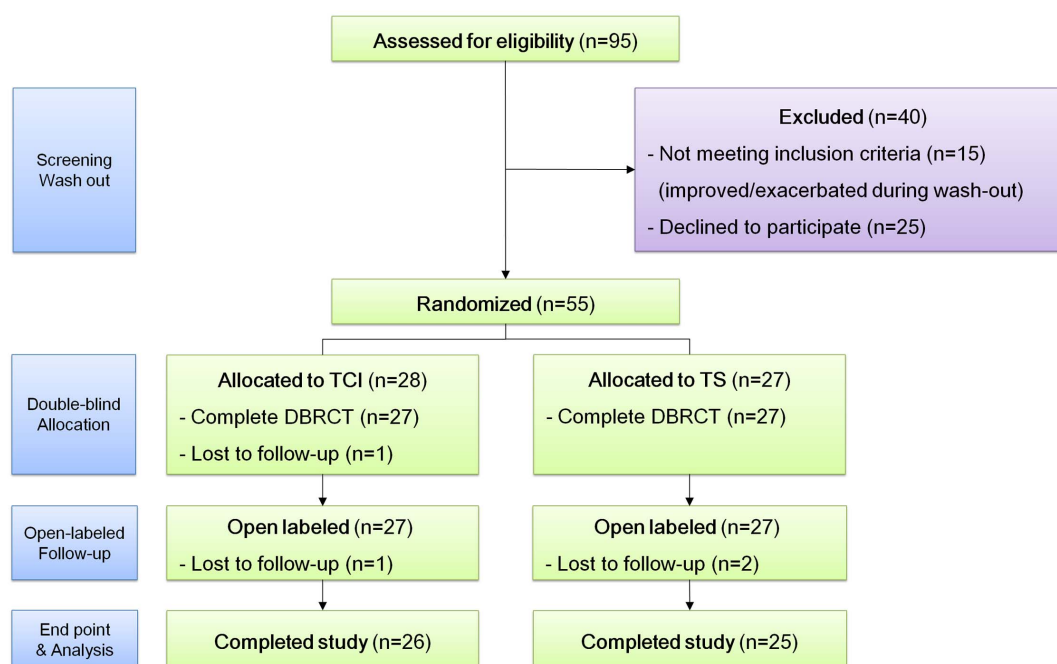

Figure 1 | Flow chart of the randomized clinical trial assessing the usages of 1% pimecrolimus cream and 0.05% desonide cream.

Table 2 | Alterations in IGA and EASI scores and area of lesion to BSA (*: 95% confidence interval) measured until week 24

	Pimecrolimus	Desonide	p-value
IGA score at week 1	2.9 (2.4–3.2)*	2.4 (2.0–2.9)	0.146
IGA score at week 2	2.5 (1.9–3.0)	2.0 (1.6–2.5)	0.185
IGA score at week 4 or open	2.4 (1.8–3.1)	2.1 (1.6–2.6)	0.346
IGA score at week 24, open-label phase	1.5 (1.0–2.0)	1.6 (1.3–2.0)	0.641
EASI score at week 1	8.3 (6.3–10.3)	5.1 (3.9–6.3)	0.007
EASI score at week 2	6.3 (4.4–8.1)	3.0 (2.1–4.0)	0.003
EASI score at week 4 or open	7.0 (4.4–9.6)	3.5 (2.1–4.0)	0.016
EASI score at week 24, open-label phase	3.1 (1.7–4.6)	2.3 (1.4–3.3)	0.340
Area of lesion to BSA at week 1; %	19.4 (15.0–23.8)	14.8 (10.9–18.8)	0.118
Area of lesion to BSA at week 2; %	13.7 (9.3–18.1)	9.1 (5.8–12.5)	0.097
Area of lesion to BSA at week 4 or open; %	13.6 (8.3–18.8)	8.7 (6.1–11.3)	0.093
Area of lesion to BSA at week 24, open-label phase; %	5.4 (3.2–7.6)	5.6 (2.8–8.5)	0.906

IGA: Investigator's Global Assessment total eosinophil;
EASI: Eczema Area and Severity Index;
BSA: Body Surface Area.

Safety. At the end of the double-blind period, infection in the application area was reported in 16 patients (59%) from the pimecrolimus group and 15 (55%) from the desonide group; no significant difference was observed. No serious adverse events or infection requiring hospitalization due to the presence of infectious disease were observed during the double-blind phase in either group. Known side effects reported by previous studies, such as the exacerbation of acute bronchiolitis or wheezing, were observed in one patient from the pimecrolimus group and four in the desonide group. The incidence of acute bronchiolitis did not significantly differ. The numbers of CD4 T cells in the pimecrolimus and desonide groups displayed no significant alterations (Table 3). The early morning free-cortisol levels (µg/dL) in the two groups were 13.1 ± 19.6 and 11.0 ± 7.6, respectively. At the end of the double-blind study, these levels were 9.9 ± 6.5 and 8.1 ± 6.7, respectively, with no significant difference observed between the groups. However, significant suppression (≤ 3 µg/dL) of the morning free-cortisol levels was exclusively observed in four patients (19%) in the desonide group ($p = 0.051$), indicating the need to consider the possibility of hypothalamic-pituitary-adrenal (HPA) axis inhibition upon application of 0.05% desonide cream in infants <2 years old. Additional blood tests, including C-reactive protein (CRP) and total eosinophil count (TEC), did not demonstrate any significant differences. Dermoscopy examination showed that skin atrophy or telangiectasia only occurred following the application of desonide. Four patients in the pimecrolimus group and eight in the desonide group experienced side effects, such as redness, hot flashes or irritability or displayed dead skin cells. However, the corresponding incidence rates did not significantly differ.

During the open-label phase, which is when the patients intermittently applied the creams for six months, most infants displayed skin infection. Among those infants who completed the 6-month follow-up, serious infections requiring hospitalization occurred in two patients from the desonide group and four from the pimecrolimus group, and no significant difference was observed. With the exception of infection, no serious adverse events were observed in the clinical subjects during the open-label phase. Blood tests, including a CD4 (+) T cell test, performed on the pimecrolimus and desonide treatment groups displayed no significant differences. No local side effects were observed in either group during the open-label phase. However, exacerbated skin atrophy or telangiectasia was reported in one patient from the pimecrolimus group and ten from the desonide group during dermoscopic evaluation of the application site. The difference between the two groups was statistically significant ($p = 0.002$).

MS data acquisition and preprocessing. To determine the changes in endogenous metabolites induced by the two drugs, urine samples were collected from the patients at the following time points: baseline and weeks 1, 2, 4, 8, 16 and 24. After sample extraction, the endogenous metabolites were screened by UPLC-QTOFMS. Additionally, analytical results from quality control samples demonstrated good reproducibility with a relative standard deviation of less than 20%. No specific tendency was observed in any of the chromatograms, and principal component analysis (PCA) was performed to derive more definitive results (Figure 2a). High-dimensional data were reduced to two dimensions with PCA, including principal component 1 (PC1) and PC2. The metabolite levels in a total of 235 samples were determined, and a distribution without specific clustering was observed. These results were attributed to the fact that extensive changes to metabolic patterns were not present. Moreover, sample variation resulting from individual differences in subjects or patient conditions at the time of sample collection could have played a role. Accordingly, an optimized statistical model was established to selectively monitor the effects of the drugs according to the duration of administration. To understand the variable range and

Table 3 | Blood test results during the double-blind period

	Pimecrolimus			Desonide		
	Baseline	Week 4 or open	p-value	Baseline	Week 4 or open	p-value
CD4 T cell (/mm³)	2250 ± 892	2107 ± 986	0.362	1585 ± 590	1532 ± 706	0.648
Morning free cortisol (µg/dL)	13.1 ± 9.6	9.9 ± 6.5	0.353	11.0 ± 7.6	8.1 ± 6.7	0.302
CRP (mg/dL)	0.24 ± 0.44	0.23 ± 0.38	0.906	0.19 ± 0.24	0.18 ± 0.22	0.973
TEC (/mm³)	946 ± 913	641 ± 691	0.059	572 ± 384	476 ± 362	0.265

CRP: C-Reactive Protein;
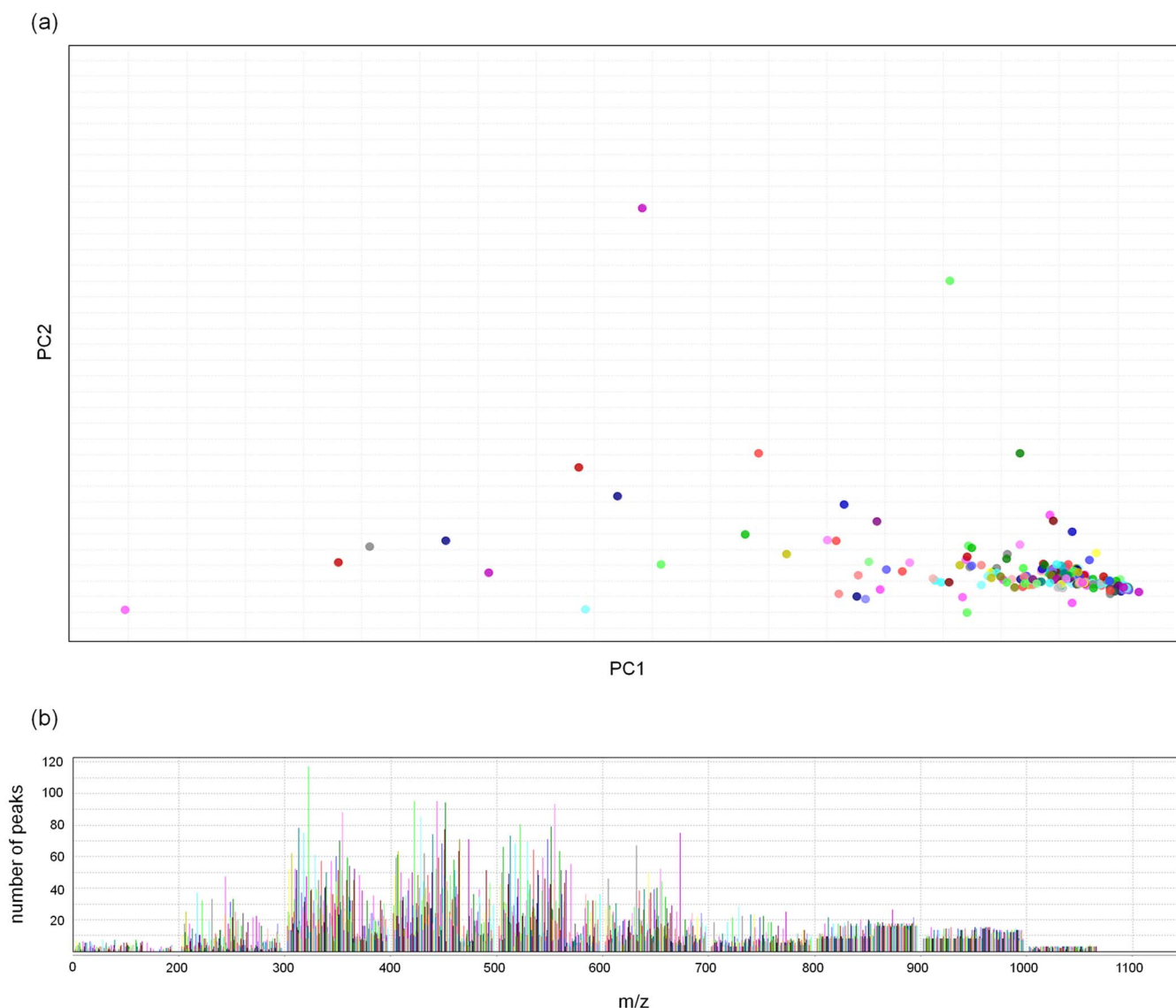TEC: Total Eosinophil Count.

(a)

(b)

Figure 2 | Results of data preprocessing. (a) Principal component analysis. One spot represents an individual sample. (b) Histogram of peak m/z values. One color represents one sample.

distribution, we confirmed peak intensities via a histogram (Figure 2b). The fact that most of the peaks occurred in the $200\sim600$ m/z range and the number of peaks displayed similarly normal distributions suggested that statistical modeling was possible without the necessity of additional normalization processes.

Statistical analysis. To determine whether meaningful patterns were evident over time in the metabolite levels of the pimecrolimus and desonide groups, we considered the mixed effects model for each metabolite individually. For the pimecrolimus group, we assumed the following mixed effects model:

$$A_i = \mu^a + \alpha_i^a + X_i\beta^a + \epsilon_i^a, \quad i = 1,2,\ldots,25, \qquad (1)$$

where $A_i = (a_{i1},a_{i2},\ldots,a_{in_i})^T$ denotes the $n_i \times 1$ vector of the meta-bolite measurements on the ith subject at time $t_i = (t_{i1},t_{i2},\ldots,t_{in_i})$, n_i is the number of measured time points of the ith subject, $\alpha_i^a \sim iid\ N(0,\sigma_a^2)$ is the random effect, $\beta^a = (\beta_0^a,\beta_1^a,\ldots,\beta_6^a)^T$ is the 7×1 vector of fixed time (or week) effects, χ_i is the $n_i \times 7$ design matrix, and $\epsilon_i^a \sim MVN_{n_i}(0,\sigma_{a\epsilon}^2 I_i)$ is the $n_i \times 1$ vector of errors. It is

important to note that n_i could be different for different subjects. All of the random effects and measurement errors were assumed to be independent of each other. In addition, we assumed the constraint $\beta_0^a = 0$. Thus, the hypothesis for the constancy of each metabolite over time can be written as $H_0^a : \beta_t^a = 0$ for all $t = 1, \ldots, 6$.

For the desonide group, we assumed the following mixed effects model:

$$B_i = \mu^b + \alpha_i^b + X_i\beta^b + \epsilon_i^b, \quad i = 1,2,\ldots,25, \qquad (2)$$

where $B_i = (b_{i1},b_{i2},\ldots,b_{in_i})^T$ denotes the $n_i \times 1$ vector of the meta-bolite measurements on the ith subject at time $t_i = (t_{i1},t_{i2},\ldots,t_{in_i})$. The other parameters and X_i are similar to those defined in the model for the pimecrolimus cream group. Here, the superscripts a and b are used to distinguish the parameters of the two groups. It is important to note that n_i could be different for different subjects. Additionally, the constraint $\beta_0^b = 0$ was assumed. We tested $H_0^b : \beta_t^b = 0$ for all $t = 1$, …, 6, which represents the constancy of each metabolite over time.

For the first analysis, we used the hypotheses $H_0^a : \beta_t^a = 0$ or $H_0^b : \beta_t^b = 0$ for all $t = 1$, …, 6 to assess the constancy of each meta-bolite over time. Using the contrast matrix

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \qquad (3)$$

the hypotheses can be written as $H_0^G : C\beta^G = 0$. We considered the following statistic:

$$T^G = \left(\widehat{\beta^G} - \beta^G\right)^T C^T \left(C\widehat{V_{\beta^G}}C^T\right)^{-1} C\left(\widehat{\beta^G} - \beta^G\right), \qquad (4)$$

where $V_{\beta^G} = Var\left(\widehat{\beta^G}\right)$ and $G = a$ or b. Under the null hypothesis (no change in intensity and no difference between the two groups), the statistics T^a and T^b theoretically follow the χ^2 distributions with $rank(C)$ degrees of freedom. Hence, we observed the statistics T_k^a and T_k^b and the corresponding p-values $p\left(T_k^a\right)$ and $p\left(T_k^b\right)$ for the kth metabolite.

We implemented the BH procedure by Benjamini and Hochberg (1995) to control for multiple testing error (FDR), which is the expected proportion of false positives among all rejected null hypotheses when testing many hypotheses simultaneously. An FDR level of $q = 0.20$ is commonly used to evaluate multivariate omics data, such as genomic data[14,15]. Sufficient numbers of selected metabolites were necessary to facilitate functional interpretations to achieve the final study goal. Hence, we applied the BH procedure using a q-value of 0.20 to model each metabolite individually. Finally, at an FDR level of 0.20, we considered the kth metabolite as displaying significant intensity in the pimecrolimus or desonide group if $p\left(T_k^a\right) \le 0.0126$ or $p\left(T_k^b\right) \le 0.0082$, respectively. Finally, we recorded the list of metabolites achieving the criteria corresponding to the models. In addition, we summarized the results by plotting the estimated coefficients over time in the figures. Figure 3 shows changes in metabolite intensity.

Functional interpretation of metabolomics data. The mechanistic meanings of the selected metabolites were precisely interpreted by statistical analyses. To predict long-term side effects, such as the carcinogenicity of the pimecrolimus formulation in the short-term clinical model, the metabolomics data were functionally interpreted and compared for the desonide and pimecrolimus groups. The MSEA results (Table 4) for the steroid-treated group indicated that metabolites related to cervical, colon and ovarian cancer exhibited altered levels during the 6-month clinical study. However, in the pimecrolimus group, no association between a particular disease and altered metabolites ($p < 0.05$) was observed. On the other hand, the MetPA results (Table 4) confirmed that the metabolites that displayed altered levels in the steroid-treated group were associated with tryptophan and phenylalanine metabolism, whereas those that were altered in the pimecrolimus-treated group were involved in phenylalanine metabolism.

Discussion

The application of topical steroids for the treatment of facial lesions in infants less than 2 years of age carries the possibility of potential side effects. Increasing concerns regarding side effects resulting from the use of steroids have been voiced by parents and guardians[16]. However, topical calcineurin inhibitors that have been suggested as alternatives to steroids are not permitted for use in infants <2 years of age. Thus, the actual uses for these drugs are limited[17]. Therefore, the need for a safe and effective secondary treatment that does not cause the skin-based side effects observed with topical steroids has emerged. In this study, we aimed to assess the efficacy of 1% pime-

crolimus cream compared with that of 0.05% desonide cream. Our data suggest that 1% pimecrolimus cream is an alternative to desonide formulations.

The carcinogenicity of pimecrolimus formulations is suspected given its long-term toxicity, which occurs via immunosuppressive mechanisms. In 2006, the use of pimecrolimus formulations was prohibited in infants under 2 years of age[18]. However, the efficiency of current carcinogenicity tests for the evaluation of compounds is low because these tests require a large number of animals. Additionally, high concentrations that are not used in clinical practice have been orally administered for test measurements. Thus, the direct application of these findings to clinical practice is complicated. Considering these findings, further studies are required to assess the carcinogenicity of pimecrolimus[19].

Our study suggests that an omics approach, which enables the dynamic examination of complete endogenous expression, would be useful to confirm the carcinogenicity of pimecrolimus[20]. This approach facilitated early predictions of carcinogenic potential using endogenous metabolites to determine both systemic reactions and side effects, which have been impossible to measure in previous short-term clinical studies[21]. To improve compliance in clinical practice, we profiled endogenous metabolites in urine samples, and this non-invasive method established an experimental model to effectively predict long-term side effects, such as carcinogenicity. Urine samples were collected from infants exposed to both creams at every study visit, and the endogenous metabolites were monitored in a time-dependent fashion. The metabolites with altered levels were examined to assess their relevance to carcinogenic mechanisms.

During the 6-month study, we assessed the reliability of our results with quality control samples to confirm the reproducibility of instrumental analysis. Additionally, with the exception of drug efficacy, additional factors, such as differences among subjects, were adjusted through more precise statistical modeling. Finally, the selected metabolites were functionally interpreted to examine their associations with actual carcinogenic mechanisms. The MSEA and MetPA methods integrate prior biological knowledge into analytical data processing, which significantly simplifies data interpretation[13].

The analytical results indicated that metabolites involved in cervical/colon/ovarian cancers were altered during the 6-month application period in the steroid-treated group. However, only one metabolite involved in these diseases was altered. Functional interpretation of metabolomics data is very useful to identify those metabolites with altered levels to predict medium- to long-term side effects at the metabolomic level. However, predictability is not perfect because probability is determined based on statistical analysis. Also, there is the possibility of false positives or false negatives. Therefore, the actual likelihood of the development of disease due to the 6-month topical application of these medications is likely to be low regardless of p-value. These results demonstrate that our approach can potentially prevent side effects by predicting the risks associated with drug application. In the pimecrolimus group, the altered metabolites were confirmed to be unrelated to specific diseases ($p < 0.05$). Accordingly, our results revealed that pimecrolimus did not initiate carcinogenic mechanisms during the 6-month topical application period, suggesting that this drug can potentially be used as a secondary therapeutic agent in infants <2 years of age.

MetPA indicated that the metabolites with altered levels in the steroid-treated groups were involved in tryptophan and phenylalanine metabolism, whereas those that were altered in the pimecrolimus-treated group were associated with phenylalanine metabolism. With regard to tryptophan metabolism, several studies have reported that oral administration of steroid formulations reduces the urinary excretion of tryptophan metabolites and increases the breast cancer phenotype[22,23]. Therefore, the possibility of the use of these metabolites as indicators of side effects has been suggested[24]. An illogical connection to previous studies can be made. It is important to note
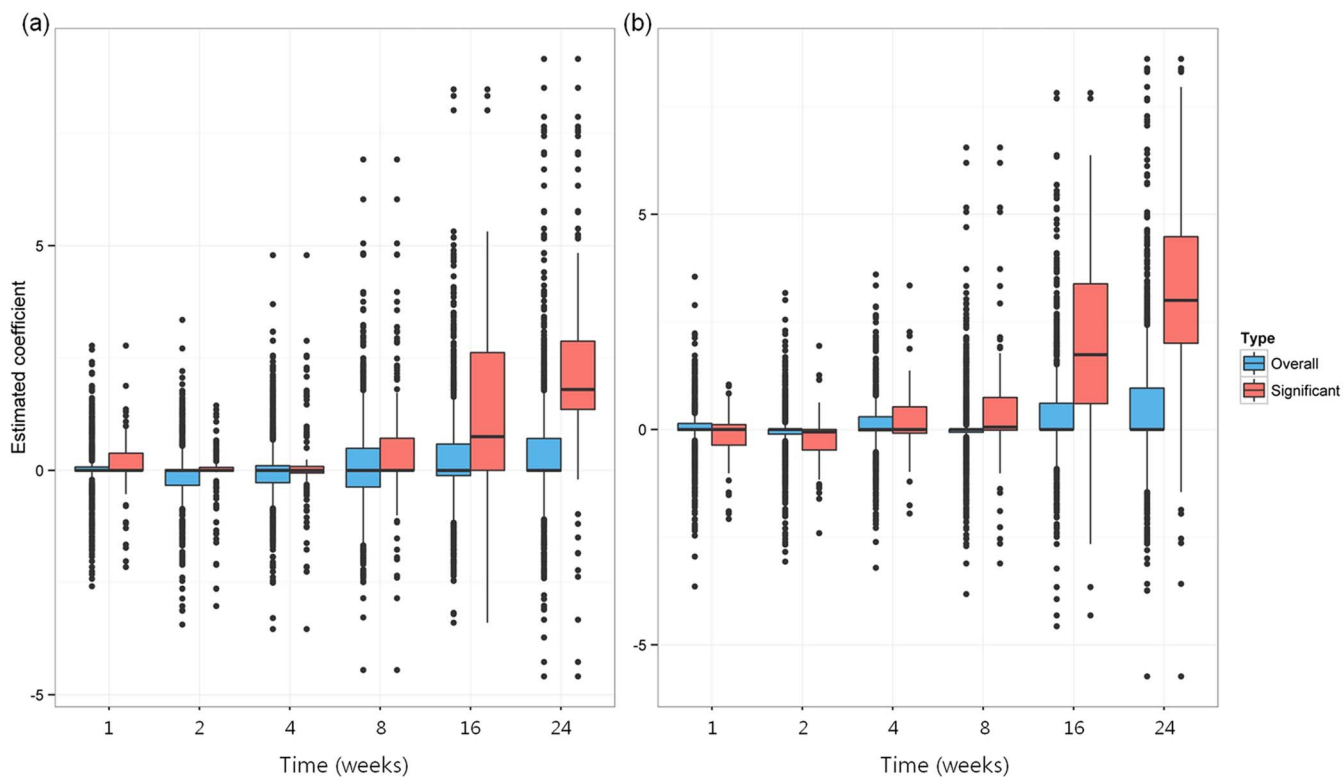
Figure 3 | Box plots demonstrating the endogenous metabolite levels in the two experimental groups. (a) Box plots of the estimated coefficients ($\beta_t^a, t = 1, 2, \ldots, 6$) over time for the group of patients who applied 1% pimecrolimus cream. (b) Box plots for the estimated coefficients ($\beta_t^b, t = 1, 2, \ldots, 6$) over time for the group of patients who applied 0.05% desonide cream. Blue box: the complete endogenous metabolite levels; red box: the endogenous metabolite levels selected for by statistical analyses.

that adults have received an oral steroid formulation in previous studies. However, we presented these studies as evidence for the feasibility of predicting possible side effects during steroid treatment. With regard to phenylalanine metabolism, the direct relevance of each component in desonide and pimecrolimus has not been previously reported; however, pre-stage catecholamine biosynthesis, which is initiated by L-phenylalanine, can be altered due to alterations in the enzymatic activities of phenylalanine hydroxylase in atopic eczema/dermatitis patients[25]. Therefore, these pathway alterations are affected by disease progression rather than drug effects.

Conclusions

Currently, steroids are the only AD treatment permitted in infants <2 years of age. However, given the possibility of potential side effects, numerous studies have compared the use of steroids versus pimecrolimus. Unfortunately, only short-term efficacy and safety

have been investigated in most of the studies, and evidence in support of long-term side effects has not been provided. In this study, we confirmed the efficacy and safety of desonide and pimecrolimus for AD treatment in infants <2 years of age. The off-label study was conducted using 51 infants with AD over 6 months of age, and we demonstrated that pimecrolimus displayed non-inferior efficacy compared with desonide as assessed by IGA and EASI scores and other measurements. In addition, we simultaneously performed various safety tests to demonstrate the superior safety of pimecrolimus compared with desonide in terms of side effects. In particular, we used a metabolomics approach to assess long-term toxicity because no effective measurements are currently available. Our results demonstrated a lack of association with carcinogenic mechanisms. This technique has the potential to predict side effects in short-term clinical models in the future. In this study, we established effective analytical techniques and statistical algorithms optimized for use

Table 4 | Results from MSEA and MetPA. This table shows the metabolite set/pathway, the total number of metabolites in that set (Total), the expectation value (Expected), the number of metabolites in the experimental set matching the pathway set (Hits), the *p*-value, the Holm *p*-value, the false discovery rate (FDR) and additional details

MSEA	Total	Expected	Hits	Raw p	-ln(p)	Holm p	FDR	
Pimecrolimus								
-								
Desonide								
Cervical/colon/ovarian cancer	1	0.041	1	0.041	3.199	1	1	
MetPA	Total	Expected	Hits	Raw p	-ln(p)	Holm p	FDR	Impact
Pimecrolimus								
Phenylalanine metabolism	45	0.168	2	0.011	4.481	0.906	0.906	0.061
Desonide								
Tryptophan metabolism	79	0.394	3	0.006	5.108	0.484	0.484	0.044
Phenylalanine metabolism	45	0.224	2	0.020	3.911	1	0.801	0.061

7

with clinical samples and demonstrated the relevance of endogenous metabolites to disease progression and phenotype. Hence, we suggest the use of endogenous metabolites as predictors of drug efficacy and side effects.

1. Walling, H. W. & Swick, B. L. Update on the management of chronic eczema: new approaches and emerging treatment options. *Clin Cosmet Investig Dermatol* **3**, 99–117 (2010).
2. Darsow, U. *et al*. Difficult to control atopic dermatitis. *World Allergy Organ J* **6**, 6 (2013).
3. Mandelin, J. M. *et al*. Long-term efficacy and tolerability of tacrolimus 0.03% ointment in infants:* a two-year open-label study. *Int J Dermatol* **51**, 104–10 (2012).
4. Luger, T. *et al*. Recommendations for pimecrolimus 1% cream in the treatment of mild-to-moderate atopic dermatitis: from medical needs to a new treatment algorithm. *Eur J Dermatol* **23**, 758–66 (2013).
5. Carr, W. W. Topical calcineurin inhibitors for atopic dermatitis: review and treatment recommendations. *Paediatr Drugs* **15**, 303–10 (2013).
6. Wooltorton, E. Eczema drugs tacrolimus (Protopic) and pimecrolimus (Elidel): cancer concerns. *CMAJ* **172**, 1179–80 (2005).
7. Paul, C. *et al*. Safety and tolerability of 1% pimecrolimus cream among infants: Experience with 1133 patients treated for up to 2 years. *Pediatrics* **117**, E118–E128 (2006).
8. Zhang, A. H. *et al*. Ultraperformance liquid chromatography-mass spectrometry based comprehensive metabolomics combined with pattern recognition and network analysis methods for characterization of metabolites and metabolic pathways from biological data sets. *Anal Chem* **85**, 7606–12 (2013).
9. Mattarucchi, E., Baraldi, E. & Guillou, C. Metabolomics applied to urine samples in childhood asthma; differentiation between asthma phenotypes and identification of relevant metabolites. *Biomed Chromatogr* **26**, 89–94 (2012).
10. Xia, J. & Wishart, D. S. Web-based inference of biological patterns, functions and pathways from metabolomic data using MetaboAnalyst. *Nat Protoc* **6**, 743–60 (2011).
11. Pluskal, T., Castillo, S., Villar-Briones, A. & Oresic, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).
12. Xia, J., Mandal, R., Sinelnikov, I. V., Broadhurst, D. & Wishart, D. S. MetaboAnalyst 2.0–a comprehensive server for metabolomic data analysis. *Nucleic Acids Res* **40**, W127–33 (2012).
13. Xia, J. & Wishart, D. S. Metabolomic data processing, analysis, and interpretation using MetaboAnalyst. *Curr Protoc Bioinformatics* **Chapter 14**, Unit 14 10 (2011).
14. Dalgliesh, G. L. *et al*. Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
15. Sharov, A. A. *et al*. Identification of Pou5f1, Sox2, and Nanog downstream target genes with statistical confidence by applying a novel algorithm to time course microarray and genome-wide chromatin immunoprecipitation data. *BMC Genomics* **9**, 269 (2008).
16. Korfitis, C., Gregoriou, S., Rallis, E. & Rigopoulos, D. Pimecrolimus versus topical corticosteroids in dermatology. *Expert Opin Pharmacother* **8**, 1565–73 (2007).
17. Manthripragada, A. D. *et al*. Off-label topical calcineurin inhibitor use in children. *Pediatrics* **132**, e1327–32 (2013).
18. Patel, T. S., Greer, S. C. & Skinner, R. B., Jr. Cancer concerns with topical immunomodulators in atopic dermatitis: overview of data and recommendations to clinicians. *Am J Clin Dermatol* **8**, 189–94 (2007).
19. Svensson, A., Chambers, C., Ganemo, A. & Mitchell, S. A. A systematic review of tacrolimus ointment compared with corticosteroids in the treatment of atopic dermatitis. *Curr Med Res Opin* **27**, 1395–406 (2011).
20. Assfalg, M. *et al*. An exploratory (1) H-nuclear magnetic resonance metabolomics study reveals altered urine spectral profiles in infants with atopic dermatitis. *Br J Dermatol* **166**, 1123–5 (2012).
21. Suzuki, N. Mass spectrometry-based quantitative analysis and biomarker discovery. *Yakugaku Zasshi* **131**, 1305–9 (2011).
22. Rose, D. P. The influence of oestrogens on tryptophan metabolism in man. *Clin Sci* **31**, 265–72 (1966).
23. Donner, N. & Handa, R. J. Estrogen receptor beta regulates the expression of tryptophan-hydroxylase 2 mRNA within serotonergic neurons of the rat dorsal raphe nuclei. *Neuroscience* **163**, 705–18 (2009).
24. Rose, D. P. Aspects of tryptophan metabolism in health and disease: a review. *J Clin Pathol* **25**, 17–25 (1972).
25. Schallreuter, K. U. *et al*. Altered catecholamine synthesis and degradation in the epidermis of patients with atopic eczema. *Arch Dermatol Res* **289**, 663–6 (1997).

Acknowledgments

Author contributions

S.J.L., S.I.W., M.K.K., J.H.P. and S.W.K. designed the experiments. S.J.L. and S.I.W. performed the experiments. S.J.L., S.H.A. and J.L. analyzed the data. S.J.L., S.I.W. and J.Y.H. wrote the manuscript.

Additional information