See related commentary on pg 1455

# Pathogenic variants affecting peptidyl arginine deiminase 3 and its major substrates underlie central centrifugal cicatricial alopecia

JID Open

Noy Keller-Rosenthal[1], Ofer Sarig[1], Moshe Giladi[2,3], Kiril Malovitski[1,2], Rotem Rubinstein[4], Yoni Haitin[2], Jorge Larrondo[5], Yolanda Lenzy[6,7], Ncoza Dlova[8], Amy McMichael[9,10] and Eli Sprecher[1,2,10]

**Journal of Investigative Dermatology** (2026) **146,** 1596–1604; doi:10.1016/j.jid.2025.10.609

Central centrifugal cicatricial alopecia (CCCA) is the most common form of primary scarring alopecia in women of African descent, typically characterized by progressive hair loss originating at the vertex of the scalp. Although genetic susceptibility has been implicated in the pathogenesis of CCCA, only 1 gene (*PADI3*, encoding peptidyl arginine deiminase 3) has been thus far associated with CCCA. This study aimed to broaden our understanding of the genetic basis of CCCA by analyzing whole-exome sequences from 75 patients with clinically and histologically confirmed CCCA. We identified 9 pathogenic heterozygous variants in *PADI3*, including, to our knowledge, 4 previously unreported missense variants, all predicted to disrupt protein function. Functional analyses revealed reduced expression, abnormal intracellular localization, and diminished enzymatic activity in cells transfected with constructs expressing the *PADI3* variants. More interestingly, pathogenic variants were identified in 2 additional genes, *S100A3* and *TCHH*, which encode the main substrates of PADI3, S100 calcium-binding protein A3 and trichohyalin. Both proteins play critical roles in hair shaft integrity. The *S100A3* variant was found to cause reduced citrullination by PADI3, whereas *TCHH* variants altered intracellular localization and resulted in significantly reduced expression of the protein. These findings provide further insights into disease mechanisms and may inform future strategies for genetic testing and targeted therapies.

**Keywords:** CCCA, PADI3, S100A3, Scarring alopecia, Trichohyalin

## INTRODUCTION

Central centrifugal cicatricial alopecia (CCCA) is the most prevalent form of primary scarring alopecia in women of African descent, commonly presenting as progressive hair loss starting at the vertex of the scalp and progressing centrifugally; several clinical subtypes and manifestations have been described (Sow et al, 2024). Prevalence is estimated at 3–6% among African American and South African women; however, there are few well-designed epidemiological studies (Callender and Onwudiwe, 2011; Khumalo et al, 2007). It is classified as a lymphocytic cicatricial alopecia and often shares overlapping clinical and histopathologic features with lichen planopilaris, making differentiation challenging in some cases. Despite these similarities, a hallmark feature of CCCA is the premature desquamation of the inner root sheath in both inflamed and noninflamed follicles (Miteva and Tosti, 2014; Ogunleye et al, 2014; Tan et al, 2019). This process is believed to contribute to the disruption of follicular integrity and subsequent scarring; however, the underlying trigger for this premature inner root sheath desquamation remains unknown. CCCA pathogenesis is now recognized to encompass complex interactions between genetic predisposition, hair grooming practices, and immune-mediated pathways (Dlova et al, 2014; Gadre et al, 2024; Malki et al, 2019; Onamusi et al, 2023).

The genetics of CCCA is complex. A dominant mode of inheritance has been suggested in CCCA on the basis of clustering of familial cases (Dlova et al, 2014). More than 20% of patients with CCCA have been shown to carry variants in *PADI3*, encoding peptidyl arginine deiminase type III (Malki et al, 2019), an enzyme mediating post-translational

[1]*Division of Dermatology, Tel Aviv Medical Center, Tel Aviv, Israel;* [2]*Gray Faculty of Medicine and Health Sciences, Tel Aviv University, Tel Aviv, Israel;* [3]*Department of Internal Medicine D, Tel Aviv Medical Center, Tel Aviv, Israel;* [4]*The George S. Wise Faculty of Life Science, Tel Aviv University, Tel Aviv, Israel;* [5]*Department of Dermatology, Clínica Alemana-Universidad del Desarrollo, Santiago, Chile;* [6]*Lenzy Dermatology & Hair Loss Center, Chicopee, Massachusetts, USA;* [7]*Department of Dermatology UCONN Health, Framington, Connecticut, USA;* [8]*University of KwaZulu–Natal, Durban, South Africa; and* [9]*Wake Forest University School of Medicine, Winston-Salem, North Carolina, USA*

[10]*These authors contributed equally to this work.*

*Correspondence: Eli Sprecher, Tel Aviv Sourasky Medical Center, 6 Weizmann Street, Tel Aviv 64239, Israel. E-mail: elisp@tlvmc.gov.il*

*Abbreviations: CCCA, central centrifugal cicatricial alopecia; HEK, human embryonic kidney; KC, keratinocyte; WT, wild-type*

*Received 31 July 2025; revised 2 October 2025; accepted 20 October 2025; accepted manuscript published online 14 November 2025; corrected proof published online 8 December 2025*

modifications critical for proper hair shaft formation (Basmanav et al, 2022; Méchin et al, 2007). These variants in *PADI3* were shown to result in disrupted protein function, abnormal intracellular localization, and decreased enzymatic activity, supporting their pathogenic role in CCCA. In this study, we studied an extended cohort of patients with CCCA to broaden our understanding of the genetic basis of CCCA.

## RESULTS

### Variant analysis

Exome sequencing was performed on a cohort of 75 patients (details are provided in Acknowledgments) of African or African American descent with clinically and/or histologically confirmed CCCA; 58 of these patients have been previously described (Malki et al, 2019). Pathogenic variants were identified on the basis of their predicted deleteriousness using various bioinformatics tools and subsequently filtered on the basis of their possible involvement in hair biology. This led to the identification of 9 missense heterozygous variants in *PADI3*, including, to our knowledge, 4 previously unreported variants as well as 1 splice-altering variant, all of which were predicted to be pathogenic (Supplementary Table S1). In addition, 4 pathogenic heterozygous missense variants were detected in *TCHH*, and 1 heterozygous missense variant was identified in *S100A3* (Supplementary Table S1). The distribution of each identified variant among the 75 patients is summarized in Table 1. *TCHH* codes for trichohyalin, and *S100A3* codes for the S100 calcium-binding protein A3. Both proteins play critical roles in hair shaft integrity (Guan et al, 2015; Kizawa et al, 2002, 1998; Steinert et al, 2003).

### Functional consequences of the previously unreported variants of PADI3

The pivotal role of PADI3 in hair shaft formation has been demonstrated in previous studies (Méchin and Simon, 2023). To investigate the consequences of the 4 additional missense variants in *PADI3*, we used protein modeling. As previously demonstrated (Malki et al, 2019), all variants were localized to the protein Ig domains. These domains are assumed to play a role in protein dimerization. Variants within the Ig domains of PADI3 were shown to significantly affect its catalytic activity (Liu et al, 2011). p.Val48Met (c.142G>A), p.Arg60Gln (c.179G>A), p.Val171Met (c.511G>A), and p.Arg248Cys (c.742C>T) are predicted to confer conformational changes affecting dimerization as well as calcium binding and, consequently, reduce PADI3 catalytic activity (Figure 1a and b).

To further investigate the effect of the additional CCCA-associated variants, HaCaT cells were transfected with constructs encoding wild-type (WT) and variant *PADI3* (Supplementary Figure S1). Immunoblotting of the cell lysate demonstrated a reduced expression of all 4 variant PADI3 constructs compared with the WT (Figure 2a). Immunofluorescence analysis revealed a homogenous cytoplasmic distribution of PADI3 in cells transfected with the WT construct, whereas cells transfected with each of the 4 variant PADI3 constructs displayed an abnormal intracellular localization, characterized by the formation of aggregates within the cytoplasm (Figure 2b), as described for other CCCA-associated *PADI3* variants (Malki et al, 2019). We next assessed the impact of the 4 pathogenic variants on PADI3 enzymatic activity, using an antibody-based PAD activity assay (details are provided in Material and Methods). All 4 CCCA-associated variant-expressing constructs exhibited a significant decrease in enzymatic activity compared with cells transfected with the WT *PADI3* construct (Figure 2c).

### Functional consequences of the S100A3 variant

S100A3 is a member of the S100 family of calcium-binding proteins and functions as a homodimer (Eckert et al, 2004). The citrullination of S100A3 by PADI3 has been suggested to play a role in hair shaft development and the maintenance of hair structure, particularly in the hair cuticle, where S100A3 is most abundantly expressed (Guan et al, 2015; Kizawa et al, 2008, 1996). This post-translational modification is thought to affect the protein's interaction with other molecules influencing hair texture and resilience (Kizawa et al, 2011).

**Table 1. Distribution of Variants in Patient Cohort**

| Gene | Transcript | dbSNP | Codon Change | Amino Acid Change | Carriers ID | Total Carriers |
|---|---|---|---|---|---|---|
| *PADI3* | NM_016233.2 | rs150695491 | c.142G>A | p.Val48Met | N2,N6[1] | 2 |
| | | rs115356766 | c.179G>A | p.Arg60Gln | 5,16,44, 57,58, 63,N9 | 7 |
| | | rs2272629 | c.511G>A | p.Val171Met | 13,38,54,56, 63,N10 | 6 |
| | | rs139876092 | c.628C>T | p.Arg210Trp | 54,55 | 2 |
| | | rs145296609 | c.742C>T | p.Arg248Cys | 54 | 1 |
| | | rs1557508308 | c.832-2A>G | Splicing | 11,52 | 2 |
| | | rs139426141 | c.856A>G | p.Thr286Ala | 1,3,9,16,47,50 | 6 |
| | | rs140482516 | c.1669C>T | p.Arg557Trp | 11,31,52 | 3 |
| | | rs34097903 | c.1744G>A | p.Ala582Thr | 3,17,58 | 3 |
| | | rs1437225536 | c.1955G>A | p.Arg652Lys | 49 | 1 |
| *TCHH* | NM_007113.4 | rs114824000 | c.2057C>G | p.Ala686Gly | N8 | 1 |
| | | rs776762916 | c.2174A>C | p.Lys725Thr | N9 | 1 |
| | | rs1487387151 | c.3355G>A | p.Glu1119Lys | N16 | 1 |
| | | rs376831805 | c.5452G>C | p.Asp1818His | 46 | 1 |
| *S100A3* | NM_002960.2 | rs61731550 | c.64C>T | p.Arg22Cys | 5,9,54, N1, N13 | 5 |

Abbreviation: ID, identification.
[1]Patients from the new cohort are designated with an 'N' (N1–N17).

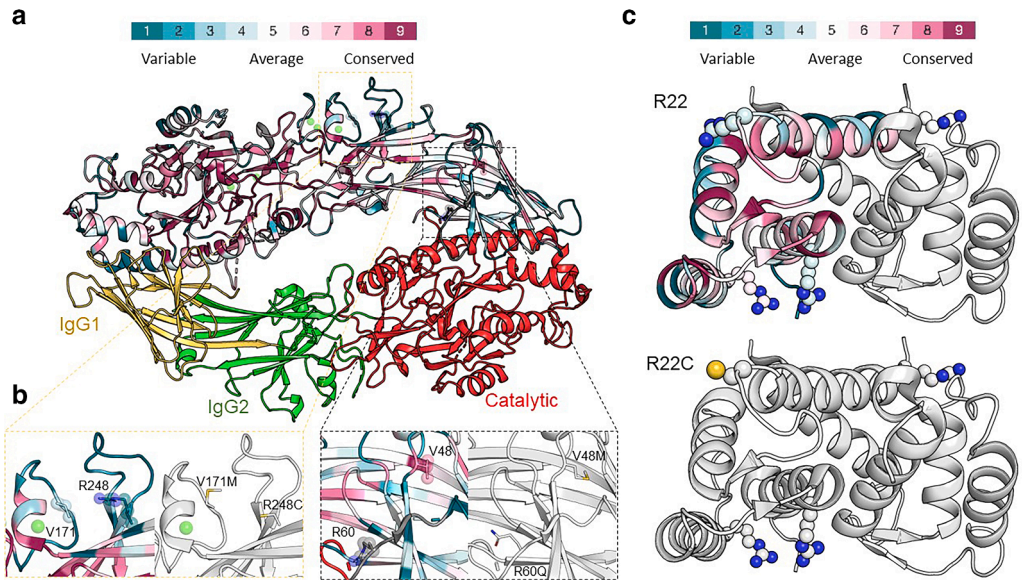



**Figure 1. PADI3 protein structural modeling.** **(a)** The crystal structure of a human PADI3 (PDB 6CEI) homodimer is presented. One monomer of the homodimeric PADI3 is colored according to the conservation score, whereas the second monomer is colored according to the structural domains (IgG1, IgG2, and catalytic domains, colored yellow, green, and red, respectively). **(b)** The mutated positions identified are magnified. Mutated positions and neighboring residues within 3 Å are shown as sticks. The mutated positions are further highlighted. The 4 variants localize to the IgG domains that presumably play a role in dimerization. Previously discovered variants within these domains were shown to compromise catalytic activity (Basmanav et al, 2022; Malki et al, 2019); p.Val171Met is adjacent to the $Ca^{2+}$ binding sites and stabilizes the loop in place by hydrophobic interactions with Leu177, Met180, and Phe214. An adjacent residue variant, Asp176Ala, exhibits 10% of the wild-type catalytic efficiency (Liu et al, 2011); p.Arg248Cys is located in the vicinity of the conserved calcium binding site, potentially affecting the local electrostatic potential and thereby calcium binding; p.Arg60Gln is predicted to perturb the monomer and dimer stability (predicted DDG = −0.45 and −0.89 kcal/mol, respectively); p.Val48Met affects a highly conserved residue. **(c)** The conserved S100A3 positions undergoing deimination by PADI3 are highlighted in blue across a crystal structure of S100A3 (upper panel); P. R22C is one of the 4 citrullination sites (lower panel). $Ca^{2+}$, calcium ion; PDB, Protein Data Bank.

Initially, protein modeling demonstrated that the CCCA-associated variant p.Arg22Cys (c.64C>T) does not affect calcium ion binding or the oligomeric state (Figure 1c).

S100A3 contains 4 arginine residues, 3 of which undergo citrullination by PADI3 (Kizawa et al, 2011). We therefore ascertained the impact of the CCCA-associated *S100A3* variant on its *PADI3*-mediated citrullination. Human embryonic kidney (HEK) 293 cells were transfected with constructs encoding WT and variant *S100A3* (p. Arg22Cys, c.64C>T) (Supplementary Figure S2). HEK293 cells were chosen for this experiment because they do not express either PADI3 or S100A3. Immunoblotting of the cell extract demonstrated significantly reduced expression of the variant S100A3 compared with that of WT S100A3 (Figure 3a). Immunofluorescence analysis demonstrated significantly reduced citrullination by PADI3 in cells transfected with variant S100A3 compared with those transfected with WT (Figure 3b). In addition, to confirm these results, we used an ELISA-based PADI3 activity assay to measure citrullination in the presence of WT and variant *S100A3*. We hypothesized that WT S100A3 should compete more efficiently for PADI3-mediated enzyme activity than the S100A3 variant protein; therefore, it should result in lower levels of citrullination of the assay's arginine-containing peptides than the variant S100A3 protein. As predicted, lysates from cells transfected with the variant *S100A3* along with the PADI3 enzyme showed higher deimination signal than the cells transfected with WT *S100A3* (Figure 3c).

**The role of TCHH in CCCA disease pathogenesis**

TCHH (trichohyalin) is a structural protein that colocalizes with PADI3 in the inner root sheath of the hair follicle and the medulla of the hair shaft (Holthaus and Eckhart, 2024; Steinert et al, 2003). PADI3-mediated deimination reduces TCHH overall charge, enabling it to associate with keratin intermediate filaments (Méchin and Simon, 2023). Subsequently, TCHH and keratin intermediate filament are crosslinked by TGM3 (transglutaminase 3). This crosslinking stabilizes and hardens keratin intermediate filaments, which are further reinforced and linked to cornified envelopes through additional crosslinking by transglutaminases, particularly TGM3 (Steinert et al, 2003; Tarcsa et al, 1997; Westgate et al, 2013). Together, TCHH and its modifications by PADI3 and TGM3 play a crucial role in structuring the mechanical integrity of the hair. We assessed 2 of the pathogenic *TCHH* variants identified in our cohort p.Ala686Gly (c.2057C>G) and p. Glu1119Lys(c.3355G>A). Owing to technical sequence–related constraints, we were unable to generate expression plasmids for the 2 other trichohyalin variants. Primary keratinocytes (KCs) were transfected with constructs encoding WT or variant TCHH. As a control, KCs were transfected with an empty vector. TCHH demonstrated cytoplasmic distribution in cells transfected with the WT construct, whereas cells transfected with each of the variant TCHH constructs displayed a significantly reduced signal with perinuclear distribution (Figure 4).

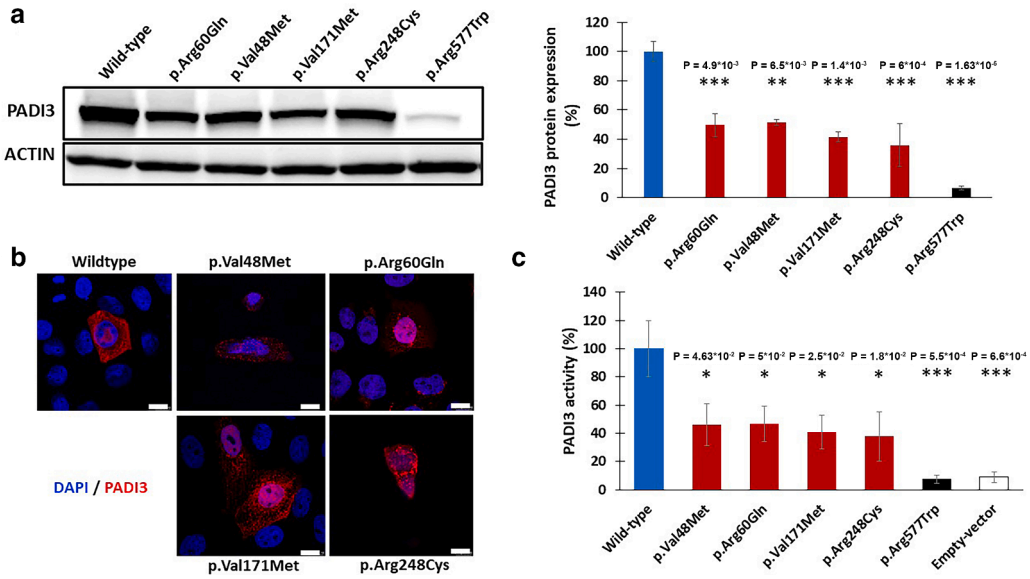



**Figure 2. Consequences of CCCA-associated variants in PADI3.** **(a)** HaCaT cells were transfected with constructs expressing either wild-type PADI3 or PADI3 harboring the p.Val48Met, p.Arg60Gln, p.Val171Met, or the p.Arg248Cys variants. The p.Arg577Trp variant, which was previously shown to be associated with CCCA (Malki et al, 2019), was used as a positive control. Immunoblotting analysis was used to ascertain the expression of the 4 PADI3 constructs in transiently transfected HaCaT cells 48 hours after transfection. Results are expressed as a percentage of PADI3 expression relative to the expression of the wild-type PADI3. Results represent the mean of 3 independent experiments ± standard error. **(b)** Immunofluorescence analysis was performed in HaCaT cells that were transiently transfected with the 5 PADI3-expressing constructs. Image analysis demonstrated homogeneous cytosolic localization of wild-type PADI3. In contrast, the 4 variant proteins formed large aggregates (red staining, bar = 10 μm; DAPI = blue staining). Of note, WB detects linear epitopes under denaturing conditions (total abundance), whereas IF depends on epitope exposure; the weak IF staining of p.Val48Met likely reflects altered epitope accessibility due to a conformational change. **(c)** HaCaT cells were transfected with constructs expressing either wild-type PADI3 or PADI3 harboring p.Val48Met, p.Arg60Gln, p.Val171Met, or p.Arg248Cys. The p.Arg577Trp variant, which was previously found to be associated with CCCA (Malki et al, 2019), was used as a positive control. PADI3 activity was measured with the use of an antibody-based assay for PAD activity according to the manufacturer's instructions (details are provided in Materials and Methods). Results are expressed as a percentage of PADI3 activity relative to the activity of the nonvariant PADI3. Results represent the mean of 3 independent experiments ± standard error (*$P < .05$, **$P < .01$, and ***$P < .005$; 1-way ANOVA). CCCA, central centrifugal cicatricial alopecia; IF, immunofluorescence; WB, western blot.

### CCCA-associated gene variant distribution

Overall, of 75 patients, 25 patients carried at least a heterozygous pathogenic variant in *PADI3* (33%), 5 patients had a heterozygous variant in *S100A3* (7%), and 4 patients had a heterozygous variant in *TCHH* (5%). Three patients harbored pathogenic changes in 2 different genes (PADI3 and TCHH or S100A3). We did not observe any significant changes in the phenotype displayed by patients carrying variants in 1 versus 2 genes. In total, 40% of the patients had a pathogenic variant in at least 1 of the 3 CCCA-associated genes.

## DISCUSSION

Our study provides further insights into the genetic basis of CCCA, identifying pathogenic variants in genes encoding PADI3 and 2 of its targets: S100A3 and TCHH. These findings confirm prior studies that revealed the importance of genetic susceptibility in CCCA and point to citrullination as a key pathway involved in the pathogenesis of this condition.

Given its crucial role in post-translational modification of structural hair proteins (Méchin and Simon, 2023), PADI3 has been previously identified as a key player in CCCA (Malki et al, 2019). In this study, we present, to our knowledge, 4 previously unreported missense pathogenic variants in *PADI3*, which were shown to affect catalytic activity alongside reduced protein expression and abnormal intracellular aggregation.

Interestingly, PADI3 functions by mediating citrullination of key proteins in the hair follicle, especially S100A3, expressed in the hair cuticle, and TCHH, expressed in the inner root sheet (Kizawa et al, 2011; Méchin and Simon, 2023; Steinert et al, 2003, 1999). Another important related element is TGM3, which cross-links TCHH and keratin intermediate filaments (Steinert et al, 2003; Tarcsa et al, 1997; Westgate et al, 2013). The colocalization of PADI3 and TCHH in the inner root sheath and medulla and of PADI3 and S100A3 in the hair shaft's cuticle (Méchin et al, 2007; Takizawa et al, 1999) underscores the functional interplay between these proteins and highlights their collective importance in hair shaft integrity (Figure 5). In this study, we detected additional variants in the genes encoding S100A3 and TCHH. These variants were found to affect the protein function and/or expression pattern. These results are reminiscent of similar data obtained in the context of a rare recessive hair disorder, the uncombable hair syndrome (MIM [Mendelian inheritance in man] 191480). Uncombable hair syndrome can result from deleterious biallelic variants in *PADI3* or in *TCHH* or *TGM3* (Ü Basmanav et al, 2016). In both conditions, the causative variants affect different elements of the hair follicle–associated citrullination pathway. Abnormal function of any component of this pathway results in the same phenotype owing to compromised mechanical stability of the hair shaft, associated with subsequent hair

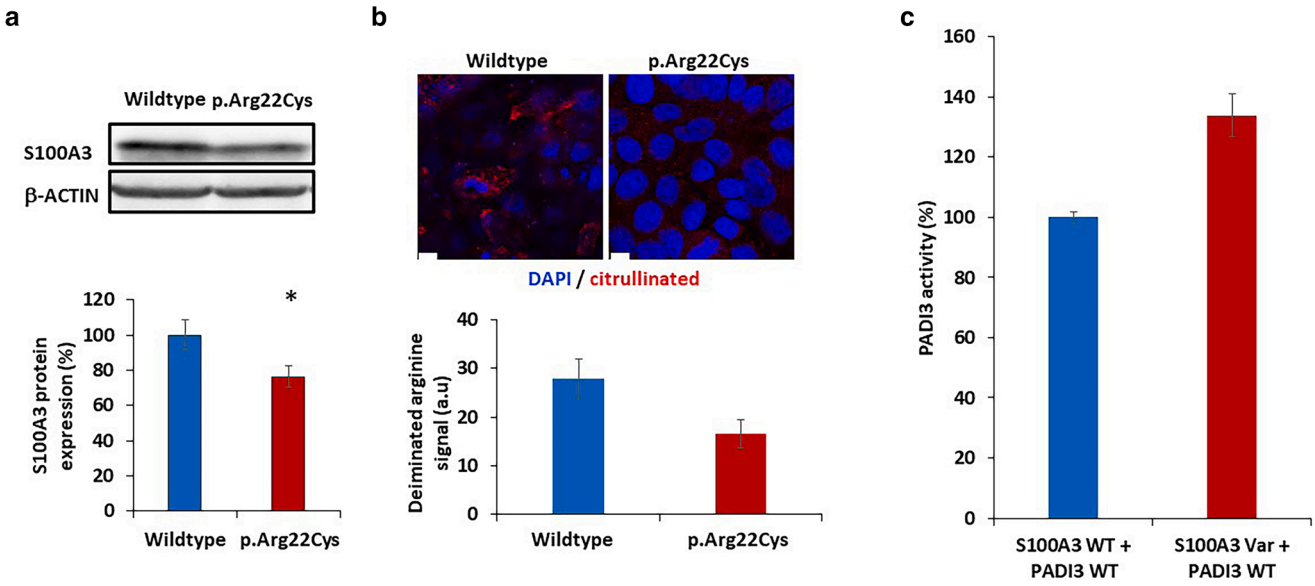



**Figure 3. Consequences of CCCA-associated S100A3 variant.** **(a)** HEK293 cells were transfected with expression constructs expressing either wild-type or mutant c.64C>T-harboring *S100A3* cDNA. As a control, cells were transfected with an EV. *S100A3* expression was ascertained using immunoblotting with anti-V5 tag antibodies followed by quantification with ImageJ software (bar = 25 μm, nuclei are stained in blue by DAPI). Results represent the mean of 2 independent experiments ± standard error (2-tailed unpaired Student's *t*-test; $*P < .05$). **(b)** HEK293 cells were transiently transfected with the wild-type and variant S100A3 constructs and stained with an antibody detecting citrullination as described in Materials and Methods. The deimination (citrullination) signal was significantly decreased in cells transfected with the variant protein. Results represent the mean of 3 independent experiments ± standard error (2-tailed unpaired Student's *t*-test; $*P < .05$). **(c)** An ELISA-based PADI3 activity assay was used to measure deamination in HEK293 cells expressing either the WT or the mutant S100A3 (p.Arg22Cys) along with the PADI3 enzyme. Cells transfected with the variant S100A3 showed higher deimination signal than cells transfected with the WT construct because they are less likely to compete with endogenous substrates. Results are expressed as a percentage of PADI3 activity in cells transfected with variant S100A3 relative to cells transfected with nonvariant S100A3. Results represent the mean of 2 independent experiments ± standard error (2-tailed unpaired Student's *t*-test; $P = .057$). CCCA, central centrifugal cicatricial alopecia; EV, empty vector; HEK, human embryonic kidney; WT, wild-type.

fragility and breakage as observed in CCCA (Basmanav et al, 2022; Ü Basmanav et al, 2016). The reduced mechanical resistance of the hair shaft predisposes it to breakage, and the resulting mechanical stress and follicular injury are likely to trigger inflammation, apoptosis, and fibrosis, providing a plausible mechanism for the development of scarring alopecia. A limitation of our study is the absence of expression analyses in patient-derived hair follicles because such biopsies were not available.

Overall, our study identified pathogenic variants in at least 1 of these 3 genes in 40% of the cohort. This highlights the genetic heterogeneity of CCCA and suggests that the disease likely results from the cumulative effects of both genetic and environmental factors.

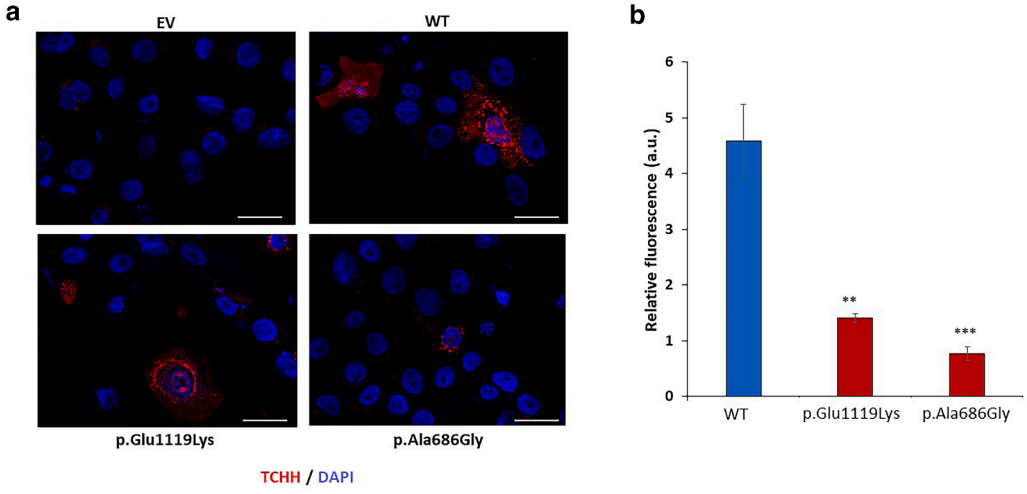



**Figure 4. Impact of TCHH variants.** **(a)** Primary keratinocytes were transfected with expression constructs harboring either *TCHH* WT sequence or the *TCHH* variants c.3355G>A (p.Glu1119Lys) or c.2057C>G (p.Ala686Gly) variants. As a control, keratinocytes were transfected with an EV. *TCHH* expression was ascertained using immunofluorescence with anti-V5 tag antibodies (bar = 25 μm, nuclei are stained in blue by DAPI). **(b)** Staining intensity was quantified using ImageJ software. Results represent the mean of 2 independent experiments ± standard error (1-way ANOVA; $**P < .01$ and $***P < .005$). a.u., arbitrary unit; EV, empty vector; WT, wild-type.

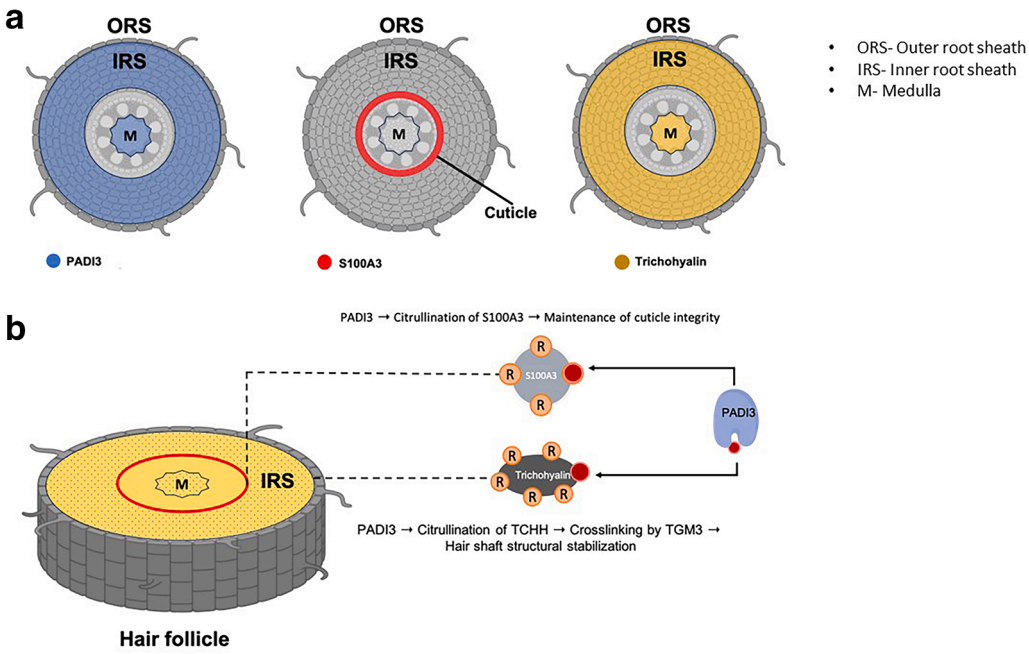



**Figure 5. Schematic representation of PAD3, S100A3, and trichohyalin expression patterns within the hair follicle. (a)** Site of maximal expression of PADI3 (blue, left panel), S100A3 (red, middle panel), and trichohyalin (mustard, right panel) in the hair follicle (ORS and IRS) (M denotes the medulla). **(b)** PADI3 essential role in the modification through residue (denoted as R) citrullination (red dots) of S100A3 in the cuticle and of trichohyalin in the IRS and medulla (denoted as M) where TGM3 cross-links modified trichohyalin. IRS, inner root sheath; ORS, outer root sheath; TGM3, transglutaminase 3.

The variants identified in association with CCCA are relatively frequent in the African population, suggesting that environmental factors or other genetic factors may codetermine the propensity to develop CCCA. Finally, our findings pave the way for therapeutic and preventive approaches in CCCA in genetically susceptible individuals, such as small molecules aimed at enhancing PADI3 enzymatic activity.

## MATERIALS AND METHODS

### Variant analysis

Whole-exome sequencing was performed at the Tel Aviv Sourasky Genomic Center. In short, targeted capture of protein-coding regions was performed using the xGen hybridization capture core reagent (Integrated DNA Technologies, Coralville, IA). Paired-end libraries were prepared from the captured fragments and sequenced on the NovaSeq 6000 system (Illumina, San Diego, CA). All whole-exome sequencing data were analyzed using the Genoox platform (http://www.franklin.genoox.com). The next-generation sequencing pipeline used was based on Burrows–Wheeler Aligner (Li and Durbin, 2010) and the 2 variant callers, namely GATK HaplotypeCaller (Mckenna et al, 2010) and FreeBayes (Miller and Bishop, 2021). Variants were classified according to their predicted protein and splicing effect with the use of the PolyPhen-2 (Polymorphism Phenotyping, version 2) tool (Adzhubei et al, 2013), the SIFT (Sorting Intolerant from Tolerant) algorithm (Kumar et al, 2009), the CADD (Combined Annotation Dependent Depletion) (Rentzsch et al, 2021) tool, Splice AI (Jaganathan et al, 2019), GERP ++ (Genomic Evolutionary Rate Profiling) (Davydov et al, 2010), and DANN (Deleterious Annotation of genetic variants using Neural Networks) (Quang et al, 2015). We then filtered the variants according to their involvement in hair biology. Validation of the variants was performed through direct sequencing except for the c.2057C>G, c.5452G>C, and c.2174A>C TCHH variants, which could not be directly sequenced for technical reasons.

### Direct sequencing

Genomic DNA was amplified through PCR using oligonucleotide primer pairs spanning the variants of interest within the *PADI3*, *S100A3*, and *TCHH* gene sequences (Supplementary Table S2) with Taq polymerase (Qiagen, Hilden, Germany). Cycling conditions were as follows: 94 °C, 2 minutes; 94 °C, 40 seconds; 62 °C, 30 seconds; and 72 °C, 30 seconds, for 35 cycles, and then 72 °C, 10 minutes. Gel-purified (QIAquick Gel Extraction Kit, Qiagen) amplicons were subjected to bidirectional DNA sequencing with the BigDye terminator system on an ABI Prism 3100 sequencer (Applied Biosystems, Foster city, NY) and the oligonucleotides used for PCR amplification. Sequencing results were analyzed using the Sequencher 5.4 software (Gene Codes).

### Conservation analysis

Evolutionary conservation analysis was performed using ConSurfDB (Ben Chorin et al, 2020) on the basis of the crystal structure of apo PADI3 (Protein Data Bank: 6CE1) and apo S100A3 (Protein Data Bank: 1KSO). Briefly, sequences are collected and multiply aligned using HMMER (https://www.ebi.ac.uk/Tools/hmmer/home) and MAFFT (https://www.ebi.ac.uk/jdispatcher/msa/mafft?stype=protein), respectively. Next, the evolutionary conservation of each position is calculated using the Rate4Site algorithm (Yariv et al, 2023), followed by projection of the conservation level for each residue onto the structure.

### Mutagenesis

We used a vector containing the reference (nonvariant) *PADI3* sequence to engineer the new mutant constructs. Primers for site-directed mutagenesis were designed using the NEBuilder online tool (New England Biolabs, Ipswich, MA), which provides optimized

primer sequences and recommended annealing temperatures for use with the Q5 Site-Directed Mutagenesis Kit (Supplementary Table S3).

The mutagenesis reaction was performed using the Q5 Site-Directed Mutagenesis Kit (New England Biolabs) according to the manufacturer's instructions. Briefly, PCR amplification was carried out in a 25-μl reaction containing 12.5 μl Q5 Hot Start High-Fidelity 2X Master Mix, 0.5 μM of each primer, and 1−10 ng of template plasmid DNA. The PCR program included initial denaturation at 98 °C for 30 seconds, followed by 25 cycles of 98 °C for 10 seconds, primer-specific annealing temperature for 10−30 seconds, and 72 °C for 20−30 seconds per kb, with a final extension at 72 °C for 2 minutes. The PCR product was then treated with the KLD enzyme mix (kinase, ligase, and DpnI) at room temperature for 5 minutes to phosphorylate and ligate the amplified DNA and digest the parental methylated template. The reaction was transformed into New England Biolabs 5-alpha chemically competent *E coli* cells, and colonies were selected on antibiotic plates. Successful mutagenesis was verified by Sanger sequencing of the target region using primers that are described in Supplementary Table S2.

### Immunoblotting

HaCaT and HEK293 cells were grown in 10 cm Petri dishes up to a confluency of 60%. Transfection with the various *PADI3*, *S100A3*, and *TCHH* constructs was carried out with Lipofectamine 3000 kit (Thermo Fisher Scientific, Waltham, MA), according to the manufacturer's instructions. For each plate, 15 μg of construct, 22.5 μl of Lipofectamine 3000, and 30 μl P3000 were used. A plate treated with the transfection buffer but without any construct was used as negative control. After 48 hours, cells were collected and lysed with RIPA buffer (Cell Signaling Technology, Danvers, MA) supplemented with protease inhibitor cocktail 100X (Thermo Fisher Scientific). After multiple sonication rounds and centrifugation at 10,000*g* for 10 minutes, the protein contents of the supernatant were quantified through Bradford assay. From each sample, 40 μg were diluted with Leammli buffer 4X (Bio-Rad Laboratories, Hercules, CA) and β-mercaptoethanol and then boiled at 95 °C for 5 minutes and loaded on 4−15% TGX stain-free gels (Bio-Rad Laboratories). Gels were run at 150 V for 40 minutes and activated for 2.5 minutes through ChemiDoc MP and the Image Lab software (Bio-Rad Laboratories). Blotting was performed on polyvinylidene difluoride membranes at 25 V and 100 mA for 1.5 hours. The following steps were carried out using an anti-V5 primary antibody (R960-25; Thermo Fisher Scientific;1:5000) and the Western Breeze chemiluminescent anti-mouse kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Finally, visualization of the bands and quantification were performed using the Image Lab software. Normalization was performed considering the total protein content of each lane, as enabled by the TGX stain-free technology.

### Immunofluorescence

HaCaT and HEK293 cells were grown on coverslips in 24-well plates. Transfection was carried out in triplicates with the Lipofectamine 3000 kit (Thermo Fisher Scientific), according to the manufacturer's instructions. For each well, 500 ng of construct, 1.5 μl of Lipofectamine 3000, and 1 μl P3,000 were used. After 48 hours, cells were fixed with 4% paraformaldehyde for 10 minutes, washed with PBS, and permeabilized with PBS-Tween20 for 10 minutes. Blocking was performed for 1 hour at room temperature, with a solution of 1% BSA supplemented with 10% normal goat serum and 10% PBS Tween20. All the washing steps were performed with PBS. Incubation with the primary antibody was performed for 2 hours at room temperature, followed by multiple washings and incubation with the secondary antibody and 200 ng DAPI for 40 minutes. The following antibodies were used: anti-V5 mouse (R960-25, 1:5000) or goat antimouse Cy3 (A10521, 1:500), both from Thermo Fisher Scientific, or anticitrullinated residues antibody from the PAD activity assay (ABAP, ModiQuest Research, AJ Nijmegen, The Netherlands). The coverslips were finally mounted with Mowiol mounting medium and visualized using an LSM 700 confocal microscope (Carl Zeiss, Oberkochen, Germany) or SP8 (Leica, Wetzlar, Germany).

### PAD activity assay

HaCaT cells were cultured in high-glucose DMEM medium containing 0.075 calcium chloride supplemented with 10% fetal calf serum, 1% L-glutamine, and 1% penicillin/streptomycin (Biological Industries, Beit-Haemek, Israel). Cells were seeded into 12 plates at 37 °C in 5% carbon dioxide in a humidified incubator, adhered overnight, washed, incubated with DMEM medium, and cultured up to 60% confluency. Cells were then transfected with various PADI3 constructs using 50 ng DNA and Lipofectamine 3000 transfection reagent (Invitrogen, Carlsbad, CA). After 48 hours, cells were collected. Cells were then homogenized on ice in Cellytic MT lysis/extraction reagent (Sigma-Aldrich, St. Louis, MO) supplemented with protease inhibitor cocktail (Sigma-Aldrich, P8340) using cell scraper and centrifuged at 10,000*g* for 10 minutes at 4 °C. The supernatant was collected and stored at −80 °C until used. PADI3 enzyme activity was measured in triplicates using an antibody-based assay for PAD activity (ABAP, ModiQuest Research) according to the manufacturer's instructions. After the second antibody incubation and washing, 100 μl of the horseradish peroxidase substrate, TMB B (Cell Signaling Technology, Danvers, MA), was added. After development of the color, the reaction was aborted with 100 μl stop solution (Cell Signaling Technology), and the optical density at 450 and 570 nm was measured with a Tecan Infinite M200 absorbance reader. We repeated the same process for the measurement of PADI3 activity in cells transfected with variant and nonvariant S100A3.

### Cell cultures

HaCaT cells were kindly provided by Dina Ron (Technion, Haifa, Israel). The cells were maintained in MEM media supplemented with 10% fetal calf serum, 1% L-glutamine, 1% streptomycin, and 1% amphotericin (Biological Industries). HEK293 cells were cultured in high-glucose DMEM with 10% fetal calf serum, 1% L-glutamine, and 1% penicillin/streptomycin (Biological Industries).

KC cell cultures were established from skin biopsies after written informed consent had been obtained as previously described (Samuelov et al, 2013). Primary KCs were maintained in KC Growth Medium supplemented with 0.4% bovine pituitary extract, 0.1% human epidermal GF, 0.1% insulin, 0.1% hydrocortisone, and 0.1% gentamicin/amphotericin B. All cell lines were confirmed to be free of mycoplasma contamination using the MycoBlue Mycoplasma Detector kit (Vazyme). Mycoplasma testing was performed prior to the initiation of experiments and at regular intervals.

### Statistical analysis

To ascertain differences in protein expression levels, enzymatic activity, deimination levels, and fluorescence signals, we used the 1-way ANOVA test. Values are reported as means ± SDs. All *P*-values are 2 tailed, and a $P < .05$ was considered to indicate statistical significance.

## ETHICS STATEMENT

This study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to sample collection and data use. Genomic analyses were performed on DNA samples previously collected and archived from patients with clinically and histologically confirmed central centrifugal cicatricial alopecia. All samples and associated clinical data were fully deidentified prior to analysis to ensure participant confidentiality.

The study protocol was approved by the Wake Forest Baptist Health Institutional Review Board (IRB00093829)

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon request.

## ORCIDs

Noy Keller-Rosenthal: http://orcid.org/0000-0002-2087-0252
Ofer Sarig: http://orcid.org/0000-0003-2987-2091
Moshe Giladi: http://orcid.org/0000-0002-8589-6920
Kiril Malovitski: http://orcid.org/0000-0002-4171-6668
Yoni Haitin: http://orcid.org/0000-0002-1648-6792
Jorge Larrondo: http://orcid.org/0000-0003-3600-8824
Yolanda Lenzy: http://orcid.org/0009-0005-8264-2748
Ncoza Dlova: http://orcid.org/0000-0002-2425-3905
Amy McMichael: http://orcid.org/0000-0002-5342-2181
Eli Sprecher http://orcid.org/0000-0002-4093-1032

## CONFLICT OF INTEREST

The authors state no conflict of interest.

## ACKNOWLEDGMENTS

This study was supported in part by a grant from the American Academy of Dermatology (Hair Loss and Alopecia Initiative in Research) (AMM and ES) and the Kahn Foundation's Orion project (MG). Some of the described central centrifugal cicatricial alopecia specimens were provided by the Black Women Health Study (https://www.bu.edu/bwhs/). This study was presented as an oral presentation at the European Society for Dermatological Research Annual Meeting (Lisbon, Portugal, 2024), with the abstract published in the official conference abstract book.

## AUTHOR CONTRIBUTIONS

Conceptualization: NK-R, MG, YH, ND, AMM, ES; Data Curation: NK-R, OS, ND, AMM; Formal Analysis: NK-R, OS, MG, YH, RR; Funding Acquisition: AMM, ES; Investigation: NK-R, OS, MG, KM, YH, RR; Methodology: NK-R, OS, ES; Project Administration: OS, AMM, ES; Resources: OS, JL, YL, AMM, ES; Supervision: OS, AMM, ES; Writing - Original Draft Preparation: NK-R, OS, ES; Writing - Review and Editing: NK-R, OS, MG, YH, RR, ND, AMM, ES

## SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at www.jidonline.org, and at 10.1016/j.jid.2025.10.609.

## REFERENCES

Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. Curr Protoc Hum Genet 2013. Chapter 7:Unit7.20.

Basmanav FB, Cesarato N, Kumar S, Borisov O, Kokordelis P, Ralser DJ, et al. Assessment of the genetic spectrum of uncombable hair syndrome in a cohort of 107 individuals. JAMA Dermatol 2022;158:1245–53.

Ben Chorin A, Masrati G, Kessel A, Narunsky A, Sprinzak J, Lahav S, et al. ConSurf-DB: an accessible repository for the evolutionary conservation patterns of the majority of PDB proteins. Protein Sci 2020;29:258–67.

Callender VD, Onwudiwe O. Prevalence and etiology of central centrifugal cicatricial alopecia. Arch Dermatol 2011;147:972–4.

Davydov EV, Goode DL, Sirota M, Cooper GM, Sidow A, Batzoglou S. Identifying a high fraction of the human genome to be under selective constraint using GERP++. PLoS Comput Biol 2010;6:e1001025.

Dlova NC, Jordaan FH, Sarig O, Sprecher E. Autosomal dominant inheritance of central centrifugal cicatricial alopecia in black South Africans. J Am Acad Dermatol 2014;70:679–82.e1.

Eckert RL, Broome AM, Ruse M, Robinson N, Ryan D, Lee K. S100 proteins in the epidermis. J Invest Dermatol 2004;123:23–33.

Gadre A, Dyson T, Jedrych J, Anhalt G, Byrd AS, Aguh C. Proteomic profiling of central centrifugal cicatricial alopecia reveals role of humoral immune response pathway and metabolic dysregulation [published correction appears in JID Innov 2024;4:100312]. JID Innov 2024;4:100263.

Guan W, Deng Q, Yu XL, Yuan YS, Gao J, Li JJ, et al. Blockade of S100A3 activity inhibits murine hair growth. Genet Mol Res 2015;14:13532–44.

Holthaus KB, Eckhart L. Development-associated genes of the epidermal differentiation complex (EDC). J Dev Biol 2024;12:4.

Jaganathan K, Kyriazopoulou Panagiotopoulou S, Mcrae JF, Darbandi SF, Knowles D, Li YI, et al. Predicting splicing from primary sequence with deep learning. Cell 2019;176:535–48.e24.

Khumalo NP, Jessop S, Gumedze F, Ehrlich R. Hairdressing and the prevalence of scalp disease in African adults. Br J Dermatol 2007;157:981–8.

Kizawa K, Takahara H, Troxler H, Kleinert P, Mochida U, Heizmann CW. Specific citrullination causes assembly of a globular S100A3 homotetramer: a putative Ca2+ modulator matures human hair cuticle. J Biol Chem 2008;283:5004–13.

Kizawa K, Takahara H, Unno M, Heizmann CW. S100 and S100 fused-type protein families in epidermal maturation with special focus on S100A3 in mammalian hair cuticles. Biochimie 2011;93:2038–47.

Kizawa K, Troxler H, Kleinert P, Inoue T, Toyoda M, Morohashi M, et al. Characterization of the cysteine-rich calcium-binding S100A3 protein from human hair cuticles. Biochem Biophys Res Commun 2002;299:857–62.

Kizawa K, Tsuchimoto S, Hashimoto K, Uchiwa H. Gene expression of mouse S100A3, a cysteine-rich calcium-binding protein, in developing hair follicle. J Invest Dermatol 1998;111:879–86.

Kizawa K, Uchiwa H, Murakami U. Highly-expressed S100A3, a calcium-binding protein, in human hair cuticle. Biochim Biophys Acta 1996;1312:94–8.

Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–81.

Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 2010;26:589–95.

Liu YL, Chiang YH, Liu GY, Hung HC. Functional role of dimerization of human peptidylarginine deiminase 4 (PAD4). PLoS One 2011;6:e21314.

Malki L, Sarig O, Romano MT, Méchin MC, Peled A, Pavlovsky M, et al. Variant PADI3 in central centrifugal cicatricial alopecia. N Engl J Med 2019;380:833–41.

Mckenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–303.

Méchin MC, Sebbag M, Arnaud J, Nachat R, Foulquier C, Adoue V, et al. Update on peptidylarginine deiminases and deimination in skin physiology and severe human diseases. Int J Cosmet Sci 2007;29:147–68.

Méchin MC, Simon M. Deimination in epidermal barrier and hair formation. Philos Trans R Soc Lond B Biol Sci 2023;378:20220245.

Miller HE, Bishop AJR. Correlation AnalyzeR: functional predictions from gene co-expression correlations. BMC Bioinformatics 2021;22:206.

Miteva M, Tosti A. Dermatoscopic features of central centrifugal cicatricial alopecia. J Am Acad Dermatol 2014;71:443–9.

Ogunleye TA, Mcmichael A, Olsen EA. Central centrifugal cicatricial alopecia: what has been achieved, current clues for future research. Dermatol Clin 2014;32:173–81.

Onamusi T, Larrondo J, Mcmichael AJ. Clinical factors and hair care practices influencing outcomes in central centrifugal cicatricial alopecia. Arch Dermatol Res 2023;315:2375–81.

Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics 2015;31:761–3.

Rentzsch P, Schubach M, Shendure J, Kircher M. CADD-Splice-improving genome-wide variant effect prediction using deep learning-derived splice scores. Genome Med 2021;13:31.

Samuelov L, Sarig O, Harmon RM, Rapaport D, Ishida-Yamamoto A, Isakov O, et al. Desmoglein 1 deficiency results in severe dermatitis, multiple allergies and metabolic wasting. Nat Genet 2013;45:1244–8.

Sow YN, Jackson TK, Taylor SC, Ogunleye TA. Lessons from a scoping review: clinical presentations of central centrifugal cicatricial alopecia. J Am Acad Dermatol 2024;91:259–64.

Steinert PM, Marekov LN. Multiple roles for trichohyalin in the inner root sheath. Exp Dermatol 1999;8:331–2.

Steinert PM, Parry DA, Marekov LN. Trichohyalin mechanically strengthens the hair follicle: multiple cross-bridging roles in the inner root sheath. J Biol Chem 2003;278:41409–19.

Takizawa T, Takizawa T, Arai S, Kizawa K, Uchiwa H, Sasaki I, et al. Ultrastructural localization of S100A3, a cysteine-rich, calcium binding protein, in human scalp hair shafts revealed by rapid-freezing immunocytochemistry. J Histochem Cytochem 1999;47:525–32.

Tan T, Guitart J, Gerami P, Yazdan P. Premature desquamation of the inner root sheath in noninflamed hair follicles as a specific marker for central centrifugal cicatricial alopecia. Am J Dermatopathol 2019;41:350–4.

Tarcsa E, Marekov LN, Andreoli J, Idler WW, Candi E, Chung SI, et al. The fate of trichohyalin. Sequential post-translational modifications by peptidyl-arginine deiminase and transglutaminases. J Biol Chem 1997;272:27893–901.

Ü Basmanav FB, Cau L, Tafazzoli A, Méchin MC, Wolf S, Romano MT, et al. Mutations in three genes encoding proteins involved in hair shaft formation cause uncombable hair syndrome. Am J Hum Genet 2016;99:1292–304.

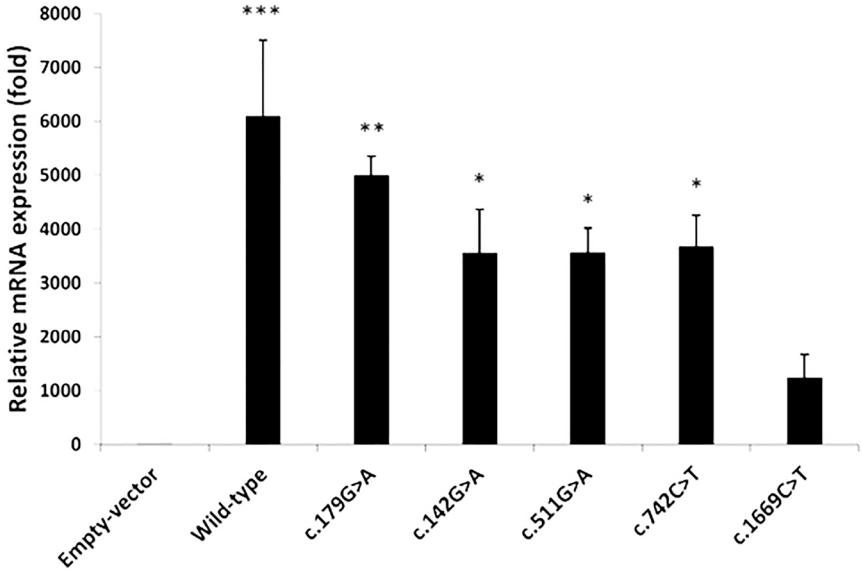
Westgate GE, Botchkareva NV, Tobin DJ. The biology of hair diversity. Int J Cosmet Sci 2013;35:329–36.

Yariv B, Yariv E, Kessel A, Masrati G, Chorin AB, Martz E, et al. Using evolutionary data to make sense of macromolecules with a "face-lifted" ConSurf. Protein Sci 2023;32:e4582.

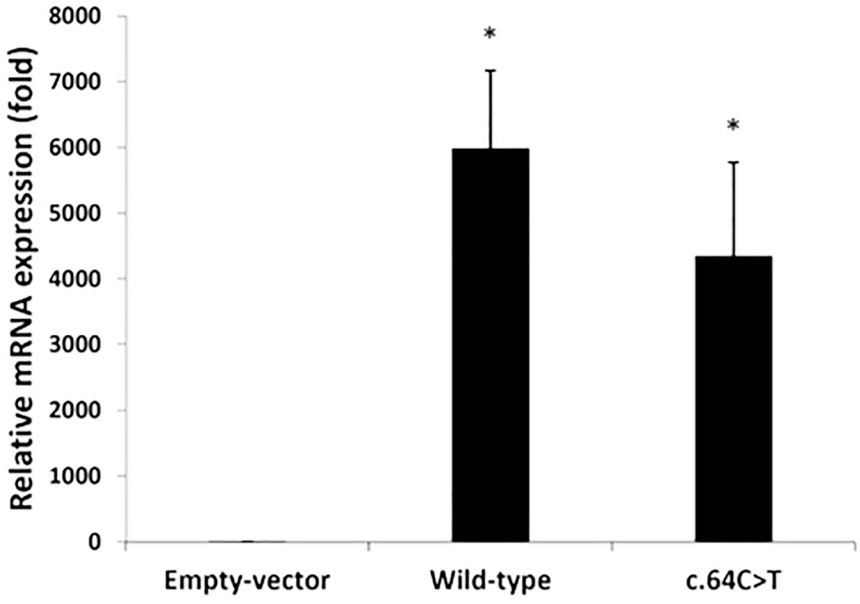
## SUPPLEMENTARY MATERIALS

## SUPPLEMENTARY REFERENCE

Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. Bioinformatics 2015;31:761–3.




**Supplementary Figure S1. *PADI3* mRNA expression in transfected cells.** *PADI3* mRNA levels were quantified using RT-qPCR in HaCaT cells transfected with expression vectors harboring the wild-type *PADI3* cDNA sequence or *PADI3* constructs harboring the c.179G>A, c.142G>A, c.511G>A, c.742C>T, or c.1669C>T variants. As a control, cells were transfected with an empty vector. Results were normalized to *GAPDH* mRNA, represent the mean of 3 experiments, and are expressed as a fold of expression of *PADI3* mRNA expression in cells transfected with an empty vector (*P*-values were calculated using 1-way ANOVA) (***$P < .001$, **$P < .01$, and *$P < .05$).




**Supplementary Figure S2. *S100A3* mRNA expression in transfected cells.** *S100A3* mRNA levels were quantified using RT-qPCR in HaCaT cells transfected with expression vectors harboring the wild-type *S100A3* cDNA sequence or *S100A3* construct harboring the c.64C>T variant. As a control, cells were transfected with an empty vector. Results were normalized to *GAPDH* mRNA, represent the mean ± SE of 3 experiments, and are expressed as a fold of expression of *S100A3* mRNA expression in cells transfected with an empty vector (*P*-values were calculated using 1-way ANOVA) (*$P < .05$). SE, standard error.

## Supplementary Table S1. Bioinformatics Analysis of *PADI3*, *S100A3*, and *TCHH* Variants

| Gene | dbSNP | Nucleotide Change | Amino Acid Change | Total AF | African AF | CADD[1] | SIFT[2] | polyPhen-2[3] | GERP ++2[4] | DANN[5] | ConSeq[6] |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *PADI3* | rs150695491 | c.142G>A | p.Val48Met | 2.2e-4 | 2.1e-3 | 23.7 | 0.001 | 1 | 5.2 | 0.999 | 8 |
| | rs115356766 | c.179G>A | p.Arg60Gln | 2.4e-3 | 4.2e-2 | 22.6 | 0.109 | 0.999 | 5.2 | 0.997 | 3 |
| | rs2272629 | c.511G>A | p.Val171Met | 8.7e-2 | 3.4e-2 | 23.2 | 0.05 | 0.986 | 4.85 | 0.998 | 6 |
| | rs139876092[7] | c.628C>T | p.Arg210Trp | 1.6e-4 | 1.8e-3 | 20.1 | 0.015 | 0.043 | 0.744 | 0.993 | 7 |
| | rs145296609 | c.742C>T | p.Arg248Cys | 6.6e-3 | 1.2e-3 | 21.3 | 0.054 | 0.005 | 5.69 | 0.990 | 3 |
| | rs1557508308[7] | c.832-2A>G | Splicing | 6.2e-7 | 1.3e-5 | 32 | NA | NA | 5.55 | NA | NA |
| | rs139426141[7] | c.856A>G | p.Thr286Ala | 2e-3 | 3.6e-2 | 24.2 | 0.034 | 0.953 | 5.55 | 0.998 | 8 |
| | rs140482516[7] | c.1669C>T | p.Arg557Trp | 4.3e-4 | 6.4e-3 | 25.5 | 0.001 | 1 | 3.64 | 0.999 | 6 |
| | rs34097903[7] | c.1744G>A | p.Ala582Thr | 1.2e-3 | 2.3e-2 | 25.7 | 0.067 | 0.957 | 4.57 | 0.999 | 9 |
| | rs1437225536[7] | c.1955G>A | p.Arg652Lys | 1.2e-6 | 2.7e-5 | 26.8 | 0.0000 | 1 | 4.84 | 0.997 | 9 |
| *TCHH* | rs114824000 | c.2057C>G | p.Ala686Gly | 8.5e-4 | 1.6e-2 | 21.8 | 0 | 0.982 | 4.3 | 0.875 | NA |
| | rs776762916 | c.2174A>C | p.Lys725Thr | 6.8e-6 | 1.3e-4 | 22.7 | 0.007 | 0.436 | 5.1 | 0.931 | NA |
| | rs1487387151 | c.3355G>A | p.Glu1119Lys | 6.2e-6 | 1.1e-4 | 21.7 | 0 | 0.799 | 3.16 | 0.89 | NA |
| | rs376831805 | c.5452G>C | p.Asp1818His | 2.6e-5 | 5.5e-4 | 22.7 | 0 | 0.995 | 3.32 | 0.874 | NA |
| *S100A3* | rs61731550 | c.64C>T | p.Arg22Cys | 1.7e-3 | 3.3e-2 | 22.4 | 0.001 | 0.994 | 5.42 | 0.998 | 6 |

Abbreviations: AF, allele frequency; CCCA, central centrifugal cicatricial alopecia; NA, not available; n.d., not determined.
African denotes African/African American.
[1]CADD: range = 0–99, and pathogenic > 20 (https://cadd.gs.washington.edu).
[2]SIFT: range = 1–0, and pathogenic < 0.05 (http://sift.jcvi.org/www/SIFT_enst_submit.html).
[3]PolyPhen-2: range = 0–1, and pathogenic > 0.85 (http://genetics.bwh.harvard.edu/pph2/index.shtml).
[4]GERP++2: highly conserved > 2 (https://genome.ucsc.edu/index.html).
[5]DANN: rang = 0–1, and pathogenic > 0.96 (Quang et al, 2015).
[6]ConSeq: range = 1–9, and moderately conserved > 5 (http://conseq.tau.ac.il).
[7]Previously published variants associated with CCCA.

## Supplementary Table S2. Oligonucleotide Sequences Used for Direct Sequencing

| Gene | Variant | Primers |
|---|---|---|
| *PADI3* | c.142G>A | F: ACCTCAAGTGATCCATCCGC<br>R: AAGAGAATGTTGGGCCCTCC |
| *PADI3* | c.179G>A | F: ACCTCAAGTGATCCATCCGC<br>R: AAGAGAATGTTGGGCCCTCC |
| *PADI3* | c.511G>A | F: ACAGGGTTGGGTGGTTACAG<br>R: CTCTGGCCTCAGAGTCTCGC |
| *PADI3* | c.742C>T | F: CCACTGTGTATTACCTGTCCT<br>R: TAGGTGTACACCACTATACC |
| *S100A3* | c.64C>T | F:GTGTGGTCTGGGAGGGCTA<br>R:GTGCTCATGCTTGTTGTCCC |
| *TCHH* | c.5452G>C | F:CCAGGAGCGCGACAGAAAAT<br>R:CGTTTCTGCTCCTCTTCTTGCC |

Abbreviations: F, forward; R, reverse.

## Supplementary Table S3. Oligonucleotide Sequences Used for Mutagenesis

| Gene | Variant | Primers |
|---|---|---|
| *PADI3* | c.142G>A | F-GACGCCTGGCATGGACATCTACA<br>R-CCATAGACCTCAAACATTTCTGTGC |
| *PADI3* | c.179G>A | F-GGAGAGGGGCCAGGAGCGTGCAG<br>R-ATGTTGGGAGAGATGTAGATGTCCACGC |
| *PADI3* | c.511G>A | F-TGACCAGCACATGCACTGCCTGC<br>R-CAATTGTCCTGGACATCACAGCTCG |
| *PADI3* | c.742C>T | F-GGATGAGGAGTGCTTCTTCGTGG<br>R-CCATGCAAGCGGGGTAC |

Abbreviations: F, forward; R, reverse.