Dnmt3a deficiency in the skin causes focal, canonical DNA hypomethylation and a cellular proliferation phenotype

David Y. Chen1, Ian M. Ferguson1, Krista A. Braun1,2, Leslie A. Sutton1, Nichole M. Helton1,2, Sai Mukund Ramakrishnan3, Amanda M. Smith1, Christopher A. Miller1, and Timothy J. Ley1,4

1Division of Dermatology, Department of Medicine, Washington University School of Medicine, Saint Louis, MO 63110; 2Division of Oncology, Department of Medicine, Washington University School of Medicine, Saint Louis, MO 63110; and 3Section of Stem Cell Biology, Commonwealth University.

Contributed by Timothy J. Ley, February 27, 2021 (sent for review November 4, 2020; reviewed by Stephen B. Baylin and Gordon D. Ginder)

DNA hypomethylation is a feature of epidermal cells from aged and sun-exposed skin, but the mechanisms responsible for this methylation loss are not known. Dnmt3a is the dominant de novo DNA methyltransferase in the skin; while epidermal Dnmt3a deficiency creates a premalignant state in which keratinocytes are more easily transformed by topical mutagens, the conditions responsible for this increased susceptibility to transformation are not well understood. Using whole genome bisulfite sequencing, we identified a focal, canonical DNA hypomethylation phenotype in the epidermal cells of Dnmt3a-deficient mice. Single-cell transcriptomic analysis revealed an increased proportion of cells with a proliferative gene expression signature, while other populations in the skin were relatively unchanged. Although total Dnmt3a deficiency has not been described in human disease states, rare patients with an overgrowth syndrome associated with behavioral abnormalities and an increased risk of cancer often have heterozygous, germline mutations in Dnmt3a that reduce its function (Tatton-Brown Rahman syndrome [TBR5]). We evaluated the DNA methylation phenotype of the skin from a TBR5 patient with a germline Dnmt3a nonsense mutation, which encodes a dominant-negative protein that reduces its methyltransferase function by ~80%. We detected a focal, canonical hypomethylation phenotype that revealed considerable overlap with hypomethylated regions found in Dnmt3a-deficient mouse skin. Together, these data suggest that Dnmt3a loss creates a premalignant epigenetic state associated with a hyperproliferative phenotype in the skin and further suggest that Dnmt3a acts as a tumor suppressor in the skin.

Significance

Keratinocyte carcinomas are the most common malignancies in fair-skinned patients. Increased age and sun exposure are important skin cancer risk factors and have been associated with a loss of DNA methylation in the human epidermis. Complete Dnmt3a deficiency in mouse skin creates a premalignant state, but the mechanisms responsible for this phenomenon are not clear. This study demonstrates that Dnmt3a deficiency in murine epidermis causes focal, canonical DNA hypomethylation and an increased proportion of cells with a proliferative gene expression signature. Understanding the mechanisms responsible for the premalignant state may be important for defining novel preventative or therapeutic strategies for skin cancers.

Author contributions: D.Y.C. and T.J.L. designed research; D.Y.C., I.M.F., K.A.B., L.A.S., N.M.H., and T.J.L. performed research; D.Y.C., I.M.F., L.A.S., S.M.R., A.M.S., C.A.M., and T.J.L. analyzed data; and D.Y.C. and T.J.L. wrote the paper.

Reviewers: S.B.B., Johns Hopkins University School of Medicine; and G.D.G., Virginia Commonwealth University.

The authors declare no competing interest.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

1To whom correspondence may be addressed. Email: timley@wustl.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2022760118/-/DCSupplemental.

Published April 12, 2021.
cases and DNMT1 in 15% (45/304) (13-15). Many of the mutations in DNMT3A in these tumors were localized to the methyltransferase domain, or were predicted to cause protein truncations that would be expected to disrupt its function and potentially cause haploinsufficiency.

Dnmt3a deficiency in murine bone marrow cells results in a focal, canonical DNA hypomethylation phenotype (16-18) that is associated with the development of hematopoietic malignancies (17, 19, 20). DNA hypomethylation in Dnmt3a-deficient marrow is reversible with DNMT3A reexpression, and remethylation occurs in an orderly fashion with predictable kinetics and high fidelity (18). Germline mutations in DNMT3A are strongly associated with an overgrowth syndrome (Tatton-Brown-Rahman syndrome [TBRs]) that is associated with a potentially increased risk of acute myeloid leukemia (AML) (21, 22). Finally, acquired mutations of DNMT3A in older individuals are the most common cause of hematopoietic clonal expansion, which is associated with an increased risk of AML development (23-26). Indeed, DNMT3A mutations are among the most common initiating mutations in adults with de novo AML; mutations at amino acid position R882 (most frequently R882H) represent the most common mutational hot spot in this gene (27).

In this study, we evaluate the premalignant epigenetic phenotype of Dnmt3a-deficient skin and identify changes that could potentially contribute to AML, both through increased susceptibility to transformation. We then relate these findings to the skin methylation phenotype of a TBRs patient with a germline Dnmt3aR882H mutation.

Results

Dnmt3a Deficiency in Epidermal Keratinocytes Results in a Focal, Canonical DNA Hypomethylation Phenotype. Mice with homozygous Dnmt3a deficiency develop normally but die within 3 wk of birth (28). To study the effects of Dnmt3a deficiency in the skin, we therefore used transgenic mice expressing Cre recombinase under the control of the Keratin 14 (Krt14) promoter, which exhibits a well-established pattern of expression that affects the epidermal skin and its appendages (29). With appropriate crosses, we obtained mice that were wild type (WT) for Dnmt3a (Krt14-Cre × Dnmt3aWT/T), heterozygous (Krt14-Cre × Dnmt3aHET/WT), or homozygous for the floxed allele (Krt14-Cre × Dnmt3aKOFlox/floxC). The floxing efficiency of the Dnmt3aKOFlox allele in epidermal cells with Krt14-Cre was $>90$%, resulting in a complete loss of Dnmt3a protein expression (SI Appendix, Fig. S1A).

Whole genome bisulfite sequencing (WGBS) was used to assess DNA methylation of isolated murine epidermal cells. Dnmt3aKO mice exhibited a small but significant decrease in global DNA methylation in all annotated regions examined, including CpG islands, island shores, island shelves, gene bodies, promoters, enhancers, and transcriptional start sites (Fig. 1A, Top). Loss of one allele of Dnmt3a (Dnmt3aHET) resulted in a statistically significant change in methylation from the Dnmt3aWT in all regions except gene bodies. On a global scale, the overall decrease in CpG methylation was very small, similar to what has been observed in hematopoietic tissues (Fig. 1B, Top) (18, 30).

We identified differentially methylated regions (DMRs) between Dnmt3aWT and Dnmt3aKO samples using previously described approaches (18, 30). DMRs were defined as having greater than 10 CpGs with a mean methylation difference of more than 0.2 (i.e., 20%) between Dnmt3aWT and Dnmt3aKO skin, with a false discovery rate (FDR) of less than 0.05. Adjacent DMRs within 50 base pairs were merged. Based on these criteria, we identified 23,983 DMRs with an average size of 0.95 kilobases (Kb), encompassing 23,53 megabases (Mb) of DNA, which represents 0.87% of the genome (Dataset S1). Virtually all DMRs were hypomethylated (23,981/23,983; 99.99%), and the degree of hypomethylation did not vary based on their location in annotated genomic regions (Fig. 1A, Bottom). Skin from Dnmt3aHET mice demonstrated a more subtle methylation phenotype at these DMRs (Fig. 1B, Bottom). The DMRs in Dnmt3a- deficient skin are represented by a heatmap, where each column represents an independent mouse, and each row, the mean methylation value of an individual DMR for all samples (Fig. 1C). There is remarkable consistency (i.e., canonicality) among the methylation phenotypes at each DMR among independent mice with the same genotypes. The methylation values of the Dnmt3aHET samples are plotted passively and exhibit an intermediate level of methylation at many DMRs (Fig. 1C), demonstrating that haploinsufficiency for Dnmt3a also has a methylation phenotype in the skin (Fig. 1A and B).

Dnmt3a-dependent DMRs occur most commonly within gene bodies; 16.4% of all annotated protein-coding genes were found to have at least one DMR. DNA methylation within most CpG islands is low in Dnmt3aWT skin, as expected. A small fraction of these islands (142/16,923; 0.89%) exhibit significantly higher methylation than total CpG islands in Dnmt3aWT skin, and these rare islands generally become hypomethylated in Dnmt3aKO skin (Fig. 1D and A, Bottom). These DMRs usually represent intragenic CpG islands, of which ~65% are methylated in human embryonic stem cells (31). Indeed, 129/142 (91%) Dnmt3aKO DMRs associated with CpG islands are located within gene bodies. The shores and shelves flanking CpG islands exhibited greater proportions of DMRs (3.7% and 8.85%, respectively), while the methylation of transcriptional start sites (TSSs) was minimally affected by Dnmt3a deletion (0.86%, Fig. 1D). The distribution of DMRs across the genome is shown in a Dnmt3aHET DMR that occurs in a gene body (Fig. 1E). This gene was found to be up-regulated in Dnmt3aKO skin tumors (11).

Dnmt3a Deficiency Results in a Distinct Transcriptional Phenotype. Dnmt3a deficiency does not result in any gross abnormalities in unperturbed skin (11). Further, flow cytometric evaluation of cellular populations within the skin revealed minimal differences between Dnmt3aWT and Dnmt3aKO (SI Appendix, Fig. S1B). To interrogate the transcriptomic consequences of Dnmt3a deficiency, we employed single-cell RNA sequencing to analyze adult murine epidermis in telogen (the resting phase of the hair cycle). We analyzed two mice from each genotype (Dnmt3aWT and Dnmt3aKO) at 9 to 10 wk of age; 10,482 cells from Dnmt3aWT and 15,403 cells from Dnmt3aKO met quality standards. Unbiased clustering based on expression signatures defined subpopulations of cells in the skin (Fig. 2A) that were remarkably similar among all four samples (Fig. 2B).

Unique populations segregated into distinct clusters with graph-based clustering and are represented by uniform manifold approximation and projection (UMAP) dimension reduction plots shown in Fig. 2A (32). In total, 19 clusters were identified by unbiased methods (Louvain method graph-based clustering, Dataset S3), which were subsequently categorized into 10 populations based on well-established, experimentally validated markers representing known populations in the epidermis (Fig. 2A, Right). The hair follicle bulge (HFB) contains follicular stem cells and is marked by Cd34 expression (33), while the upper hair follicle (uHF) is marked by beta defensin (Defb6) expression (34, 35). Subdivided into the isthmus and infundibulum, the upper hair follicle bridges the follicle with the interfollicular epidermis and sebaceous gland, which is marked by the expression of Mgst1 (34). The basal interfollicular epithelium (bIFE) contains slow cycling, proliferative cells that act as an epidermal stem cell compartment (36), and is marked by “basal” keratin gene expression (Krt5 and Krt14), as well as Thbs1 (37). The suprabasal interfollicular epithelium (sIFE) is postmitotic and marked by the expression of the genes that encode keratins Krt1 and Krt10 and an early differentiation marker metallothionein Mt4 (34). Melanocytes were marked by expression of the genes encoding tyrosinase (Tyr) and dopachrome tautomerase (Dct). Inflammatory cells, including T cells and Cd207-expressing Langerhans cells, comprised 1.3% and 1.1% of total cells in the samples, respectively. Finally, one population
Fig. 1. Dnmt3a deficiency results in focal, canonical CpG methylation loss in murine epidermal cells. (A) WGBS of isolated murine epidermal cells, comparing independent samples from 11 Dnmt3a<sup>WT</sup> vs. 5 Dnmt3a<sup>KO</sup> vs. 3 Dnmt3a<sup>HET</sup> mice. Dot plots represent average methylation values over annotated regions for all CpGs (Top) and for DMRs within each annotated region (Bottom). Significance was determined by Student’s t test with denotation of “ns” for not significant and levels of significance as * adjusted P < 0.05, ** P < 0.01, *** P < 0.001. (B) Methylation density plots for the indicated genotypes across all genomic CpGs (Top) and at DMRs (Bottom). (C) Targeted deficiency of Dnmt3a (Krt14-Cre<sup>+</sup> × Dnmt3a<sup>fl/fl</sup>, Dnmt3a<sup>KO</sup>) results in 23,983 DMRs in epidermal cells compared to wild-type control, Dnmt3a<sup>WT</sup>. Heatmap representation of WGBS data demonstrates DMRs defined by differences between Dnmt3a<sup>WT</sup> vs. Dnmt3a<sup>KO</sup> epidermal cells. Heterozygosity for Dnmt3a (Krt14-Cre<sup>+</sup> × Dnmt3a<sup>fl/+</sup>, Dnmt3a<sup>HET</sup>) results in an intermediate phenotype. Each column represents an independent mouse. (D) Annotated regions associated with DMRs. *Of 16,315 DMRs involving gene bodies, 3,612 unique genes are involved. (E) Integrated genome viewer (IGV) view of the G0s2 gene CpG methylation of Dnmt3a<sup>WT</sup> versus Dnmt3a<sup>KO</sup> epidermal cells. Normalized methylated CpG reads are represented in bars (scale 0 to 1.0 for each row). Differentially methylated regions are indicated in the bottom track and highlighted in yellow.

Chen et al. Dnmt3a deficiency in the skin causes focal, canonical DNA hypomethylation and a cellular proliferation phenotype.
Fig. 2. Single-cell RNA sequencing reveals that Dnmt3a deficiency causes discrete gene expression changes in murine keratinocytes. (A) UMAP representation of single-cell RNA sequencing data from whole epidermis of two pairs of Krt14-Cre− × Dnmt3afl/fl (Dnmt3aWT; 10,783 cells) and Krt14-Cre+ × Dnmt3afl/fl mice (Dnmt3aKO; 16,324 cells) with unbiased graph-based clustering demonstrating known skin populations represented by lineage defining genes (Right), including HFB, uHF, bIFE, sbIFE, and sebaceous gland (Seb), Krt18-expressing cells (Krt18), proliferative cells (Pro), melanocytes (Mel), T cells (T-cell), and Langerhan’s cells (Lang). (B) Comparison of keratinocyte lineage cells in the Dnmt3aWT versus Dnmt3aKO mice, excluding inflammatory cells and melanocytes (excluded clusters 6, 8, 9, and 10). UMAP projections for each pair of samples with the indicated genotypes are shown. (C) No major population shifts are observed in epidermal lineages affected by Krt14-Cre mediated Dnmt3a deficiency, with the exception that the cells dominated by a proliferative signature (Pro, cluster M) are increased (6.1% in Dnmt3aKO versus 1.8% in Dnmt3aWT), as indicated by the bar graphs of keratinocyte lineage subclusters. Subclusters of the hair follicle bulge (A and B), upper hair follicle (C and D), basal interfollicular epidermis (E–G), suprabasal interfollicular epidermis (H–K), and sebocytes (L) are relatively unchanged. (D) Canonical gene expression changes were noted in single-cell RNA sequencing comparing keratinocyte lineage cells in Dnmt3aWT versus Dnmt3aKO epidermis (n = 2 mice per genotype). (E) There are 174 differentially expressed genes between Dnmt3aWT and Dnmt3aKO keratinocyte lineage cells with 137 up-regulated in the Dnmt3aKO and 37 down-regulated genes with fold change >2 and FDR <0.01.
was dominated by expression of several genes associated with cellular proliferation (“Pro”), including genes encoding cyclins, cyclin-dependent kinases, and later cell cycle components like Ube2c, which encodes part of the anaphase promoting complex (see Fig. 4 below); this suggests that cells committed to mitosis have a signature dominated by components necessary for execution of the cell cycle (38).

The Krt14-Cre conditional knockout of Dnmt3a targets the epidermal keratinocyte lineages; we therefore excluded skin-infiltrating hematopoietic cells from subsequent analyses. Keratinocyte lineage UMAP representations of single-cell data demonstrate a high degree of similarity among samples, and unbiased clustering demonstrated similar distributions between populations (Fig. 2B, clusters A–M). Small, but statistically significant differences in population sizes between Dnmt3a WT and Dnmt3a KO epidermal keratinocytes were noted in the follicular bulge cluster 2 and upper hair follicle 1 (clusters B and C), as well as basal and suprafollicular interfollicular epidermis (clusters E and I); however, the most dramatic difference was in the proliferative (Pro) cluster (cluster M), which is expanded 3.4-fold in Dnmt3a KO skin (1.8% vs. 6.1%, P = 2.2 × 10−16, Fig. 2C). In addition to this population shift, there were canonical gene expression changes associated with Dnmt3a deficiency (Fig. 2D). In total, there were 174 differentially expressed genes (DEGs) exhibiting a fold change of 2 or greater with a FDR of 0.01 or less. Most of these genes (137) were more highly expressed in the Dnmt3a KO cells, while 37 were expressed at lower levels (Fig. 2D and E).

Differentially Methylated Regions Are Associated with Gene Expression Changes. Of the 23,983 total DMRs identified by WGBS, 16,772 were within 10 kb of a gene, representing 5,573 unique genes. Of these, 2,181 genes (39%) were differentially expressed, as defined by an FDR of < 0.01, and 9 down-regulated in Dnmt3a KO cells, Fig. 3A; however, the magnitude of methylation change was not predictive of expression changes. Further, some genes (e.g., Krt7 and Meis2) exhibited dramatic but opposing expression changes, despite associated DMRs that were hypomethylated to a similar extent (Fig. 3B). The location of DMRs within annotated regions (e.g., enhancers, CpG islands, etc.) did not predict the magnitude or direction of expression changes (SI Appendix, Fig. S2).

Meis2 is a homeobox transcription factor involved in developmental processes and regulation of tumor proliferation (39), and Hoxc8 modulates Wnt signaling in the hair follicle niche to control hair follicle regeneration (40). Both are DMR-associated genes and are expressed at significantly lower levels in most Dnmt3a KO epidermal cells (Fig. 3C and D and SI Appendix, Fig. S3A), suggesting that the normal expression of these two genes may be linked to DNA methylation marks added by Dnmt3a. In contrast, some genes are differentially expressed only in subpopulations of skin cells. For example, Keratin 7 (Krt7), which is frequently overexpressed in aggressive subtypes of cutaneous squamous cell carcinoma, and associated with poor survival in patients with esophageal squamous cell carcinoma (41, 42), is predominantly expressed in the upper hair follicle in Dnmt3a WT mice; in Dnmt3a KO mice, its expression increases in the upper hair follicle as well as the basal and suprabasal interfollicular epidermis compartments (Fig. 3E and SI Appendix, Fig. S3B). Similarly, Irx1 is primarily expressed in the hair follicle bulge and upper hair follicle cells of Dnmt3a WT mice, but its expression increases in all compartments in Dnmt3a KO mice (Fig. 3F and SI Appendix, Fig. S3B), a finding that was validated by Western blotting (Fig. 3G). Decreased methylation in the Irx1 promoter is associated with increased expression and metastatic capability in osteosarcoma (43).

Deficiency in Dnmt3a Is Associated with an Increased Fraction of Cycling Cells in the Skin. The most significant population shift between Dnmt3a WT and Dnmt3a KO keratinocyte lineage cells occurred in a cluster that was dominated by the expression of genes associated with cell division (Figs. 2B and C and 4A). This population (cluster M) is defined by the expression of late interphase cyclins (e.g., Ccnb2 and Ccnb1), as well as later cell cycle genes like Cenpf; these genes are rarely expressed in the cells of any other clusters (Fig. 4B). The size of this “proliferative cluster” expanded from 1.8% of total epidermal cells in Dnmt3a WT mice to 6.1% in Dnmt3a KO mice (3.4-fold, Fig. 4A), similar to the increase in the fraction of cells expressing Mki67 (from 2.8 to 10.2%, 3.6-fold, Fig. 4C). The expression of cell cycle genes in cluster M was equally high in Dnmt3a WT and Dnmt3a KO cells, as expected, since all cells in this cluster are likely to be actively dividing, regardless of their genotype (Fig. 4D); however, Dnmt3a KO cells in clusters A–L demonstrated increased expression levels of cell cycle genes, including Mki67 (although the expression levels of these genes were far lower than that of proliferating cells), as well as the essential cyclin-dependent kinase Cdk1, which is a differentially expressed gene associated with a DMR (Fig. 3A). To validate the proliferative phenotype with an orthogonal method, we injected 5-bromo-2-deoxyuridine (BrdU) intraperitoneally into young (8 to 9 wk old) Dnmt3a WT and Dnmt3a KO mice, and generated single-cell suspensions of epidermal cells 24 h later for flow cytometric analysis of BrdU uptake. As predicted by the single-cell RNA-sequencing (scRNA-seq) data, the fraction of actively cycling, Cd45-negative epidermal cells was significantly increased in the Dnmt3a KO mice (Fig. 4E). Epidermal subpopulations likewise demonstrated significantly increased BrdU incorporation in Dnmt3a KO basal interfollicular epidermis cells (and trended upwards in the upper hair follicle); however, no change was detected in the follicular bulge stem cell compartment, or the postmitotic suprabasal interfollicular epidermis, neither of which proliferates in the telogen state (SI Appendix, Fig. S4 A and B).

Focal, Canonical DNA Hypomethylation in the Skin of a Patient with a Dominant-Negative DNMT3A R882H Mutation. Patients with TBRs have germline mutations in DNMT3A and exhibit several common clinical features, including overgrowth, characteristic facies, and neurodevelopmental disorders (44, 45) as well as an increased risk of developing AML and perhaps other malignancies. We previously demonstrated that the peripheral blood cells from a nonleukemic patient with a heterozygous germline DNMT3A R882H allele exhibited a focal, canonical hypomethylation phenotype similar to that of AML samples initiated by the same mutation, suggesting that hypomethylation precedes transformation (30).

To determine whether this mutation causes a methylation phenotype in human skin, we analyzed DNA methylation in two skin samples from the same TBRs patient (at ages 9 and 14) and his unaffected brother (at ages 13 and 17). Global DNA methylation levels were minimally altered (similar to Dnmt3a KO skin). Nearly all of the 1,234 DMRs identified (1,221,1234; 98.9%) were hypomethylated (Fig. 5B); the smaller number of DMRs (compared to Dnmt3a KO mouse skin) probably reflects the fact that the R882H mutation is dominant negative (80% reduced activity) rather than null, and also the reduced power to detect DMRs with only two samples from each genotype. Plotting the mean methylation values of each DMR between both samples from the two brothers demonstrated the canonical nature of most of these DMRs (Fig. 5C). These DMRs were most likely to occur in gene bodies, similar to that seen in Dnmt3a KO mice (Figs. 1D and 5E).

At many loci, there was a striking similarity in DNA methylation patterns between the DMRs of homologous genes in mice and humans. For example, DMRs near the homeobox gene IRX3 in human TBRs skin are similar to DMRs near Irx3 in mouse skin (Fig. 5 D, Left). Another gene important for development, Fzd1, shares a similar methylation phenotype between the TBRS.

Chen et al.

Dnmt3a deficiency in the skin causes focal, canonical DNA hypomethylation and a cellular proliferation phenotype

PNAS | 5 of 10
https://doi.org/10.1073/pnas.2022760118

MEDICAL SCIENCES
Fig. 3. Differentially methylated regions are associated with differentially expressed genes. (A) Heatmap from single-cell RNA sequencing represents the 47 differentially expressed genes between Dnmt3aWT and Dnmt3aKO keratinocyte lineage cells that are associated with a Dnmt3a DMR and exhibit an expression fold change $>2$ with an FDR $<0.01$. Grayscale bars represent cells from independent samples, two from each genotype. (B) Relationship between gene expression and methylation changes between Dnmt3aKO and Dnmt3aWT keratinocytes. There are 5,968 differentially expressed genes with an FDR $<0.01$, of these there are 2,181 genes associated with DMRs. This plot represents expression versus methylation of these 2,181 genes. Significant differential expression, with fold change $>2$ denoted as red dots. Gray dots represent fold change between 1 and 2. (C) IGV views of differentially methylated regions associated with Meis2, which has annotated DMRs in CpG islands, shelves, shores, and gene body. Gene expression of Meis2 changes by $-5.77$-fold (FDR $5.7 \times 10^{-249}$). Dnmt3aKO expresses Meis2 in 172/15,403 (1.11%) of cells compared to Dnmt3aWT 1,106/10,482 (10.6%). (D) Differentially methylated regions associated with Hoxc8, which has DMRs in shores and gene body. Hoxc8 expression changes by $-6.72$-fold (FDR $2.51 \times 10^{-300}$) and is expressed in 234/15,403 (1.52%) Dnmt3aKO cells compared to 1,375/10,482 (13.1%) Dnmt3aWT. (E) Differentially methylated regions associated with Irx1, which has DMRs in CpG islands, shores, promoter, and gene body. Expression is changed by $7.19$-fold (FDR $9.2 \times 10^{-182}$) in Dnmt3aKO cells compared to $1,375/10,482 (13.1\%)$ Dnmt3aWT. (F) Differentially methylated regions associated with Krt7, which has DMRs in shores, promoters, and gene body. Expression is changed by $2.02$-fold (FDR $1.16 \times 10^{-109}$) in Dnmt3aKO keratinocyte lineage cells, and cells expressing increases from 2,873/15,403 (18.7%) Dnmt3aKO cells to 885/10,482 (8.44%) Dnmt3aWT. (G) Simple Western demonstrates protein level expression changes in Irx1 and Krt7 from epidermal cells isolated from the indicated Dnmt3aWT and Dnmt3aKO littermates (bracketed). Actin serves as a loading control.
The purpose of this study was to define the baseline epigenetic characteristics of Dnmt3a-deficient skin. Using WGBS, we defined the number of cells expressing the proliferative marker gene Mki67 increases from 250/10,485 (2.8%) cells in Dnmt3aWT to 1,509/14,815 (10.2%) cells in the Dnmt3aKO. (D) Single-cell RNA-sequencing expression data of multiple cell cycle relevant genes, comparing the proliferative cells (cluster M) versus the rest (clusters A–L). *FDR < 0.01 and ns, not significant. Expression values are represented by least-squares (LS) means. (E) BrdU uptake in vivo in keratinocytes assessed by flow cytometry (Cd45-negative, single-cell epidermal suspension) comparing Dnmt3aWT (n = 7) and Dnmt3aKO (n = 5) mice injected 24 h earlier with BrdU. BrdU uptake is significantly higher in Dnmt3aKO epidermal keratinocytes (5.6% BrdU+ in Dnmt3aWT vs. 7.5% BrdU+ in Dnmt3aKO, P = 0.023).

**Discussion**

The purpose of this study was to define the baseline epigenetic characteristics of Dnmt3a-deficient skin. Using WGBS, we defined the focal, canonical DNA methylation phenotype of Dnmt3a-deficient mouse epidermal cells and of a human patient with a germline DNMT3AHS2H mutation. The small changes in the global methylation state of epidermal cells reflects the restricted places in the genome where Dnmt3a acts, and the small sizes of these regions (on average, less than 1 Kb). There were 1,234 DMRs identified in the skin of the TBRS patient, which affected 3.19% of annotated gene bodies. Similarly, Dnmt3a-dependent methylation in murine epidermis occurred most often in gene bodies—more than in promoters or CpG island shores, which were the most common regions affected. However, the magnitude and location of these differentially methylated regions were not predictive of gene expression changes detected by scRNA-seq. Overall, there were 5,968 DEGs in Dnmt3aKO skin cells, but the number of DEGs that were close to a DMR (i.e., within 10 Kb) represented a subset of these genes (2,181/5,968; 36.5%). Only a small subset of these exhibited an expression fold change of >2 (2.15% of DMR associated DEGs, 0.8% of all DEGs). Like multiple previous studies (18, 30, 46), the relationship of DMRs to gene expression patterns was unpredictable at a global level, again demonstrating that DNA methylation is but one factor in determining the expression of a specific gene in a specific cellular context. Prior studies demonstrated accurate remethylation of Dnmt3a-dependent DMRs in Dnmt3aKO bone marrow with induced expression of wild-type DNMT3A, as well as correction of a subset of gene expression changes (18). However, it is not yet clear whether remethylation was directly responsible for altering the expression of these differentially expressed genes.

Aging and sun exposure have been demonstrated to cause DNA methylation changes, which were particularly apparent comparing skin and mouse Dnmt3aKO skin (Fig. 5 D, Right). The Wnt signaling pathway is important both for skin and bone development, and multiple genes of the Wnt pathway also exhibited similar Dnmt3a-dependent methylation patterns between mouse and human skin (SI Appendix, Fig. S5). Of the 1,034 DMR-associated genes identified in the TBRS patient, 432 overlapped with Dnmt3aKO. These data represent the aggregate of the remaining keratinocyte lineage clusters (clusters A–L). Each dot is an individual cell. Expression values are represented by least-squares (LS) means. (E) BrdU uptake in vivo in keratinocytes assessed by flow cytometry (Cd45-negative, single-cell epidermal suspension) comparing Dnmt3aWT (n = 7) and Dnmt3aKO (n = 5) mice injected 24 h earlier with BrdU. BrdU uptake is significantly higher in Dnmt3aKO epidermal keratinocytes (5.6% BrdU+ in Dnmt3aWT vs. 7.5% BrdU+ in Dnmt3aKO, P = 0.023).
**Fig. 5.** A germline, dominant-negative DNMT3A mutation causes focal DNA methylation alterations in human skin. (A) Whole genome bisulfite sequencing analysis for whole skin samples comparing two samples from a TBRS patient and his brother, comparing CpG methylation globally (Top) or at DMRs (Bottom). Level of significance * adjusted $P < 0.05$ and **$P < 0.01$. (B) Density plots demonstrate global CpG methylation (Top) and CpG methylation at DMRs (Bottom) in the TBRS patient and his unaffected sibling. (C) Heatmap illustrates 1,234 differentially methylated regions in the skin of the TBRS patient, compared with his sibling. (D) IGV visualization of CpG methylation at the IRX3/Irx3 and FZD1/Fzd1 loci. Each row represents an independent sample with $Dnmt3a^{WT}$ vs. $Dnmt3a^{KO}$ represented in blue and red (Top) and sibling vs. TBRS patient in purple and yellow (Bottom). Gray boxes highlight areas of differential methylation. (E) Gene bodies are the most frequently involved annotated region by TBRS DMRs. *There are 963 DMRs that reside in annotated gene bodies, representing 637 unique genes. (F) There are 23,983 DMRs in the $Dnmt3a^{WT}$, 16,315 of them are within 10 kb of 3,612 unique genes. Of the 1,234 TBRS DMRs, 1,086 are located within 10 kb of genes in the human genome. Of the 1,086 TBRS-DMR associated genes, 707 have a mouse homolog, with 432 genes overlapping DMR-associated genes in $Dnmt3a^{KO}$ mice and the TBRS patient.

### Table: DMRs and CpG methylation

| Region       | Total | DMRs | % Total |
|--------------|-------|------|---------|
| CpG island   | 31144 | 640  | 2.05    |
| Shore        | 56324 | 391  | 0.69    |
| Shelf        | 78636 | 57   | 0.07    |
| Gene body    | 19951 | 637* | 3.19*   |
| Promoter     | 38459 | 185  | 0.48    |
| Enhancer     | 121622| 13   | 0.01    |
| TSS          | 28988 | 29   | 0.10    |

*Note: * indicates significance adjusted for multiple comparisons.*
UV-exposed skin in older patients to UV-protected skin in younger patients. Large, predominantly hypomethylated blocks were found to cover 99 Mb of the genome in older, UV-exposed skin (5). Although the cause of this phenotype is unknown, a number of possibilities exist, including a progressive decline in expression or function of the DNA methyltransferases, or increased activity of the demethylases (i.e., TIM gene family members). In mice, Dnmt1 and Dnmt5a are expressed in the basal layer of keratinocytes during development; however, in adults, the expression of Dnmt1 is restricted to the hair follicle, and Dnmt3a expression is restricted to the interfollicular epidermis (8, 11). Notably, Dnmt3a expression decreases in this compartment in adult mice (11), though it is not yet clear whether these findings contribute to the increased risk of skin cancers in elderly patients.

To begin to address this question, Rinaldi et al. demonstrated that epidermal Dnmt3a deficiency resulted in decreased tumor latency and increased tumor numbers in mouse skin treated with DMBA/TPA, suggesting that the threshold for oncogenic transformation is lower in mice with Dnmt3a deficiency (11). In the same study, mice with Dnmt3b deficiency had no phenotype, and compound deficiency of Dnmt3a and Dnmt3b (double knockout/DKO) resulted in tumor latency that was indistinguishable from that of Dnmt3a deficiency; however, a broader spectrum of tumors arose in the DKO mice, including basal cell carcinomas, as well as more aggressive tumors (including squamous cell and spindle cell carcinomas). These tumors developed a reliance on Pparγ; we did not detect up-regulation of Pparγ in the premalignant state, suggesting that dysregulation may occur during or after transformation. In contrast, certain gene expression changes (i.e., increased expression of G62δ) persisted from the Dnmt3a-deficient preneoplastic state to the resulting tumors (11) (Fig. 1E and Dataset S2). Together, these data strongly suggested that Dnmt3a deficiency creates a premalignant state for keratinocyte carcinomas in mice, with some gene expression changes that persist in tumors. We observed that Dnmt3b deficiency causes only 20 DMRs in the skin, and that Dnmt3a and Dnmt3b compound deficiency (DKO) adds little to the Dnmt3a KO methylation phenotype (SI Appendix, Fig. S6). Therefore, in mouse skin, Dnmt3a is the dominant de novo DNA methyltransferase, and the loss of its function in aging skin could likewise be relevant for human skin cancer progression.

An alternative, but more likely cause of the hyperproliferative state is down-regulation of Dnmt3a expression (or function) of the DNA methyltransferases, or inactivation of their activity by epigenetic mechanisms (48). Additionally, a model of Dnmt3a deficiency in KRASG12D-driven lung tumors demonstrated that at 24 wk after induction, Dnmt3a-KO tumors were six times larger than their Dnmt3a-WT counterparts and exhibited a markedly higher proliferative index (49). Together, these findings suggest that epigenetic changes caused by Dnmt3a deficiency can result in a hyperproliferative state in several cellular contexts. The specific genes responsible for initiating this phenotype are not yet clear, but several good candidates (i.e., Cdk1, Hras, E2f1, and Cdkn1a) are defined in this study (SI Appendix, Fig. S7B–D). Proliferative priming may therefore represent one mechanism that is relevant for the premalignant state associated with DNMT3A haploinsufficiency or deficiency in multiple tissues. The resulting expanded proliferative capacity may exist in equilibrium until cooperating mutations occur, which appear to be essential for transformation in many model systems.

In addition to clonal hematopoiesis, premalignant mutation-driven clonal expansion has now been shown to occur in many other systems, including the skin (2), liver (50), esophagus (51, 52), and bronchial epithelium (53). Mutationally driven, expanded clones usually do not progress to frank malignancies, suggesting that additional epigenetic or genetic alterations are required to cause transformation. In clonal hematopoiesis, the expansion of hematopoietic stem/progenitor cells is usually caused by mutations in epigenetic regulatory genes, most commonly in DNMT3A (23, 24, 26, 54). These clonally expanded cells are at increased risk of transformation to myelodysplastic syndromes or AML, likely by altering the epigenetic “fitness” of these cells for transformation (55–57). Similar epigenetic fitness alterations may modify tumor susceptibility in skin and other organs by a variety of mechanisms.

We identified similarities between the methylation phenotypes of mice and humans with decreased DNMT3A activity: DMRs in the epidermis of Dnmt3a KO mice shared considerable gene-level similarity with a TBR5 patient with a germline DNMT3A R802H mutation. In fact, 432 out of the 1,086 gene-associated DMRs identified in the skin of this patient overlapped with gene-associated DMRs found in Dnmt3a KO skin. Several of these overlapping genes have known roles in development, including IRX3/Irx3; this locus shares a high degree of evolutionary conservation in noncoding sequences between human, mouse, and zebrafish, suggesting the presence and conservation of both functional and regulatory elements (58). Additionally, important pathways in skin development (like the WNT pathway) have multiple members with similar DNMT3A/Dnmt3a-dependent methylation patterns (SI Appendix, Fig. S5). Recent descriptions of AMLs and central nervous system (CNS) tumors in young patients with germline loss-of-function DNMT3A mutations strongly suggest that DNMT3A acts as a tumor suppressor in both hematopoietic and CNS cells (59). Although the skin phenotype and skin cancer susceptibility of children with TBR5 has not yet been described, our data suggest that these patients may be at elevated risk, and that appropriate primary prevention measures should be taken to protect these patients from sustained UV light exposure. Finally, these data also suggest a role for DNMT3A in keratinocyte carcinomas arising in aging and sun-exposed skin; additional studies to define this role are in progress.

Materials and Methods

Human Samples. Samples were obtained as part of a study approved by the Human Research Protection Office at Washington University School of Medicine (approval #20111765), which explicitly allows for potentially identifying genomic studies, including whole genome sequencing. Written informed consent was obtained from the parents of the patient with TBR5 (and his unaffected brother) for whole genome bisulfite sequencing studies, in accordance with the institutional review board approved protocol at Washington University School of Medicine.

Animals and Epidermal Cell Isolation. All mouse work was performed in accordance with institutional guidelines and approved by the Animal Studies Committee at Washington University. For all studies, we harvested single-cell suspensions of dorsal epidermis in second telogen (7 to 10 wk of age), or in
Whole Genome Bisulfite Sequencing and Analysis. WGBS was performed as described previously (17) with the exception that DNA was generated from epidermal cells. Alignment and DMR calling details are in SI Appendix, Supplementary Materials and Methods.

Single-Cell RNA Sequencing and Analysis. Single-cell suspensions of murine epidermis were freshly prepared and libraries generated using the 10x Chromium Controller system. Data analysis was performed using the standard tools available in the Partek Flow software suite (Partek, Inc.). Full details of the analysis are in SI Appendix, Supplementary Materials and Methods.

Immunoblotting and Flow Cytometry. Freshly isolated single-cell suspensions of murine epidermis were lysed in radioimmunoprecipitation assay (RIPA) buffer or suspended in staining buffer for further analysis by flow cytometry. Lysis and staining conditions (as well as antibody information) are described in SI Appendix, Supplementary Materials and Methods.

Data Availability. DNA and RNA sequences (mouse data) have been deposited in NCBI Sequence Read Archive (Bioproject PRJNA674614). Anonymized DNA sequences (human data) have been deposited in dbGaP (https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000159.v1.p5).

ACKNOWLEDGMENTS. This study was supported by grants from the Dermatology Foundation (Dermatologist Investigator Research Fellowship and Physician Scientist Career Development Award) and NIH Grants CA237277 (to D.Y.C.), CA211782 (to C.A.M.), and CA101937 and CA197561, and a Barnes Jewish Foundation Award #5169 (to T.J.L.). The Siteman Cancer Center Flow Cytometry Core (National Cancer Institute P30CA19842) provided expert support for all flow sorting studies. We thank Michael Hoeck for excellent animal husbandry.

1. H. W. Rogers, M. A. Weinstock, S. R. Feldman, B. M. Coldiron, Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the U.S. Population, 2012. JAMA Dermatol. 151, 1081–1086 (2015).
2. I. Martincorena et al., Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. Science 348, 880–886 (2015).
3. S. Baylin, T. H. Beseor, Altered methylation patterns in cancer cell genomes: Cause or consequence? Cancer Cell 9, 293–305 (2001).
4. W. Timp, A. P. Feinberg. Cancer as a dysregulated epigenome allowing cellular growth advantage at the expense of the host. Nat. Rev. Cancer 13, 497–510 (2013).
5. S. A. R. Vandervelde et al., Age and sun exposure-related widespread genomic blocks of hypomethylation in a nonmelanoma skin cancer. Genomics 16, 80 (2015).
6. T. Chen, Y. Ueda, J. E. Dodge, Z. Wang, E. Li. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol. Cell. Biol. 23, 5594–5605 (2003).
7. P. A. Jones, G. Liang, Rethinking how DNA methylation patterns are maintained. Nat. Biotechnol. 37, 38–48 (2019).
8. C. S. Tremps et al., Enrichment for living murine keratinocytes from the hair follicle bulge with the cell surface marker CD34. J. Invest. Dermatol. 120, 501–511 (2003).
9. S. Jost et al., Single-cell transcriptomics reveals that differentiation and spatial signaling shape epidermal skin cell homeostasis. Cell 13, 221–237.e5 (2016).
10. A. Füllgrabe et al., Dynamics of Lgr6+ progenitor cells in the hair follicle, sebaceous gland, and interfollicular epithelium. Stem Cell Reports 5, 843–855 (2015).
11. C. Blanpain, E. Fuchs. Epidermal homeostasis: A balancing act of stem cells in the skin. Nat. Rev. Mol. Cell Biol. 10, 207–217 (2009).
12. J. Varani et al., Thrombospondin-induced adhesion of human keratinocytes. J. Clin. Invest. 81, 1537–1544 (1988).
13. M. L. Whitfield et al., Identification of genes specifically expressed in the human cell cycle and their expression in tumors. Mol. Biol. Cell 13, 1977–2000 (2002).
14. A. Groß et al., Tumorigenic and anti-proliferative properties of the TALE-transcription factors MEIS2 and MEIS2A in neuroblastoma. Cancer Res. 78, 1395–1497 (2018).
15. Z. Yu et al., Hoxc-dependent mesenchymal niche heterogeneity drives regional hair follicle regeneration. Cell Stem Cell 23, 487–500.e5 (2018).
16. M. Pulitzer, G. Desman, K. J. Busam, CK7 expression in primary cutaneous squamous cell carcinoma. J. Cutan. Pathol. 37, 966–972 (2010).
17. A. Yamada et al., Expression of cytokeratin 7 predicts survival in stage IIIA/IIIB squamous cell carcinoma of the esophagus. Oncol. Rep. 37, 1081–1085 (2012).
18. I. Martincorena et al., Clonal hematopoiesis and blood cancer risk. N. Engl. J. Med. 370, 1357–1361 (2019).