

Selection of Oncogenic Mutant Clones in Normal Human Skin Varies with Body Site



Joanna C. Fowler¹, Charlotte King¹, Christopher Bryant¹, Michael W.J. Hall^{1,2}, Roshan Sood¹, Swee Hoe Ong¹, Eleanor Earp¹, David Fernandez-Antoran¹, Jonas Koeppel¹, Stefan C. Dentre^{1,3}, David Shorthouse², Amer Durrani⁴, Kate Fife⁴, Edward Rytina⁴, Doreen Milne⁴, Amit Roshan^{4,5}, Krishnaa Mahubani⁶, Kourosh Saeb-Parsy⁶, Benjamin A. Hall⁷, Moritz Gerstung^{3,8}, and Philip H. Jones^{1,2}

ABSTRACT

Skin cancer risk varies substantially across the body, yet how this relates to the mutations found in normal skin is unknown. Here we mapped mutant clones in skin from high- and low-risk sites. The density of mutations varied by location. The prevalence of *NOTCH1* and *FAT1* mutations in forearm, trunk, and leg skin was similar to that in keratinocyte cancers. Most mutations were caused by ultraviolet light, but mutational signature analysis suggested differences in DNA-repair processes between sites. Eleven mutant genes were under positive selection, with *TP53* preferentially selected in the head and *FAT1* in the leg. Fine-scale mapping revealed 10% of clones had copy-number alterations. Analysis of hair follicles showed mutations in the upper follicle resembled adjacent skin, but the lower follicle was sparsely mutated. Normal skin is a dense patchwork of mutant clones arising from competitive selection that varies by location.

SIGNIFICANCE: Mapping mutant clones across the body reveals normal skin is a dense patchwork of mutant cells. The variation in cancer risk between sites substantially exceeds that in mutant clone density. More generally, mutant genes cannot be assigned as cancer drivers until their prevalence in normal tissue is known.

See related commentary by De Dominici and DeGregori, p. 227.

¹Wellcome Sanger Institute, Hinxton, United Kingdom. ²MRC Cancer Unit, University of Cambridge, Hutchison-MRC Research Centre, Cambridge Biomedical Campus, Cambridge, United Kingdom. ³European Molecular Biology Laboratory, European Bioinformatics Institute, Cambridge, United Kingdom. ⁴Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Cambridge, United Kingdom. ⁵Cancer Research UK Cambridge Institute, University of Cambridge, Li Ka Shing Centre, Cambridge, United Kingdom. ⁶Department of Surgery and Cambridge NIHR Biomedical Research Centre, Biomedical Campus, Cambridge, United Kingdom. ⁷Department of Medical Physics and Biomedical Engineering, University College London, London, United Kingdom. ⁸European Molecular Biology Laboratory, Genome Biology Unit, Heidelberg, Germany.

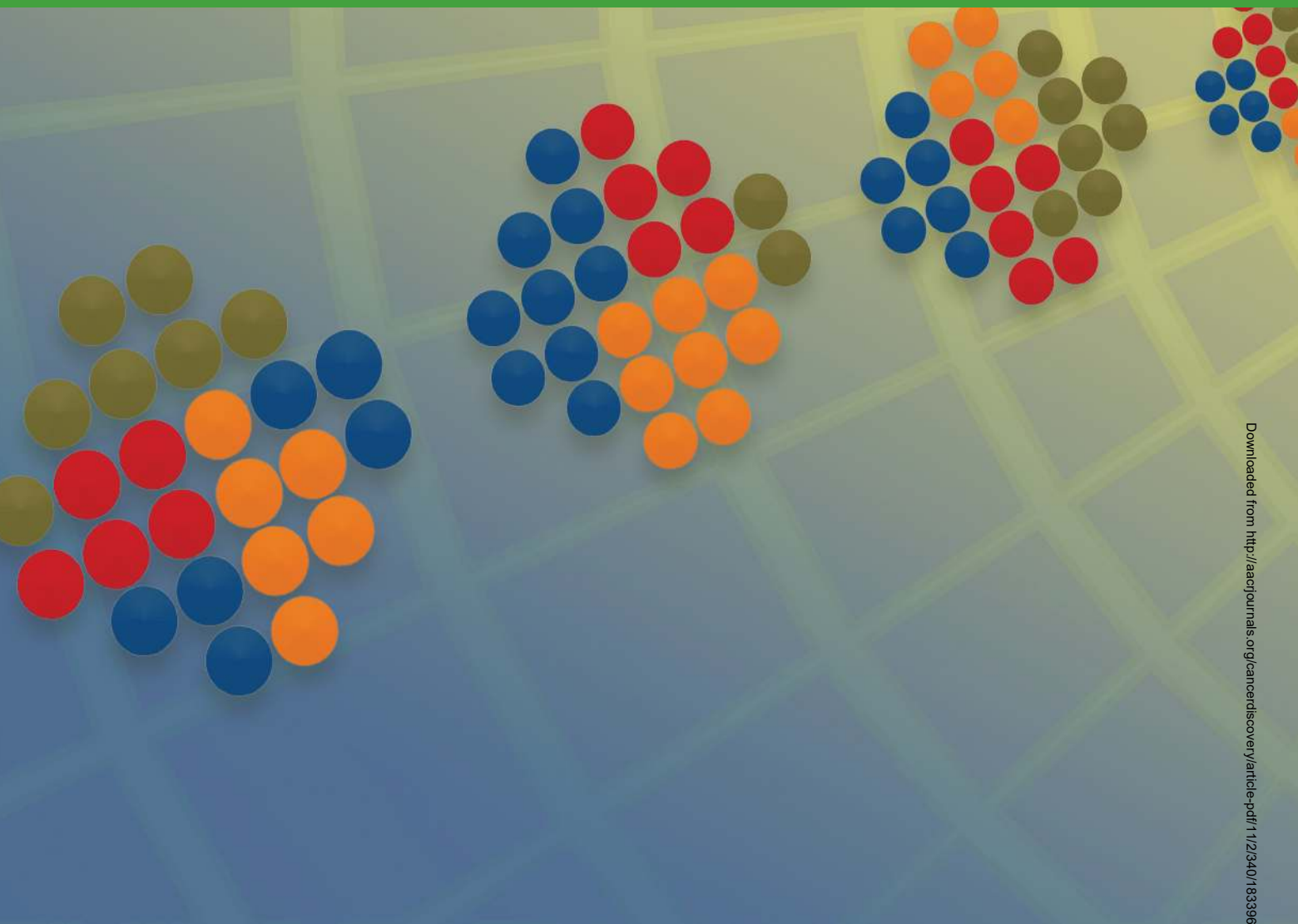
Note: Supplementary data for this article are available at Cancer Discovery Online (<http://cancerdiscovery.aacrjournals.org/>).

Corresponding Author: Philip H. Jones, Wellcome Sanger Institute, Genome Campus, Hinxton, Cambridge CB10 1SA, United Kingdom. Phone: 44-1223499917; E-mail: pj3@sanger.ac.uk

Cancer Discov 2021;11:340–61

doi: 10.1158/2159-8290.CD-20-1092

©2020 American Association for Cancer Research.



INTRODUCTION

Normal facial skin from middle-aged humans has a substantial burden of clones carrying cancer-associated mutations generated by ultraviolet (UV) light (1). Several mutant genes are under strong positive genetic selection, evidenced by an excess of protein altering relative to silent mutations (the dN/dS ratio), suggesting these mutations promote clonal expansion rather than being neutral “passengers” (2). The positively selected mutant genes include those commonly mutated in basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) of the skin, which are among the commonest cancers worldwide. Both cancers contain frequent mutations in *TP53*, *NOTCH1*, *NOTCH2*, and *FAT1*, all of which are positively selected in facial skin (1). These findings are consistent with cancer emerging from a subset of clones within normal tissue.

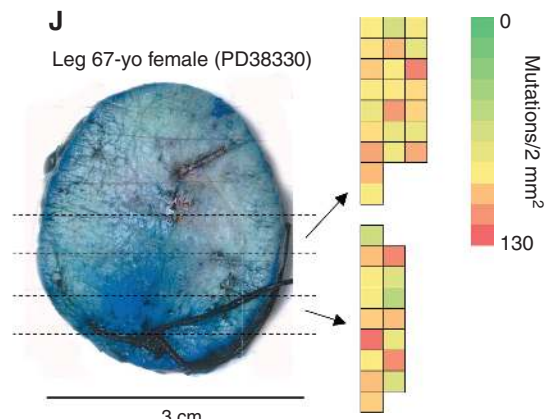
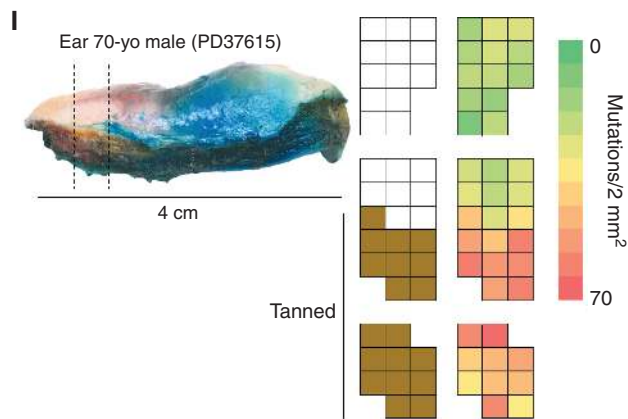
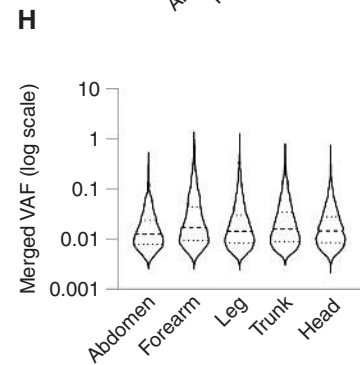
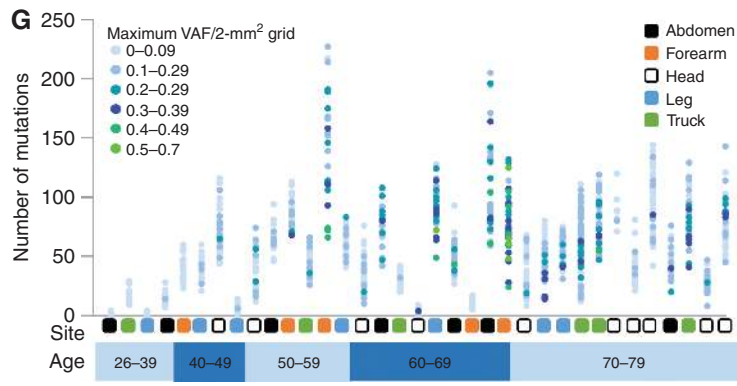
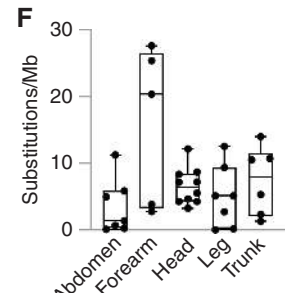
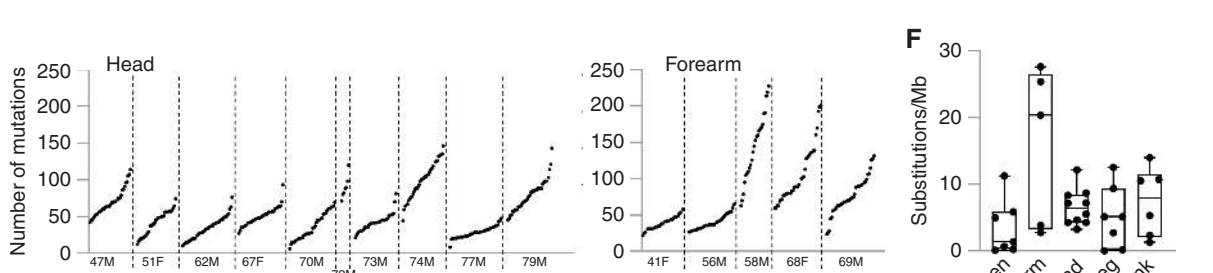
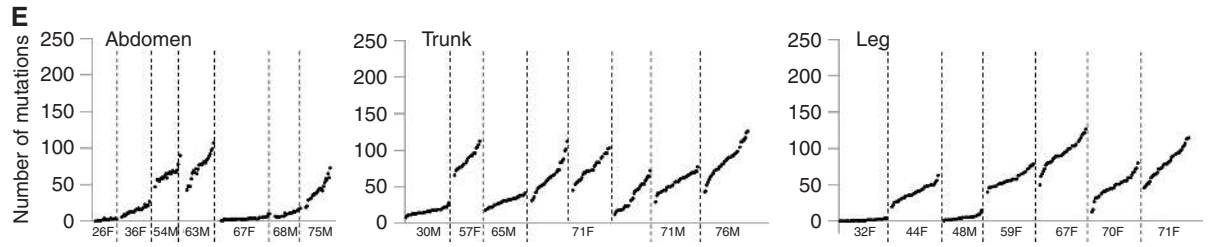
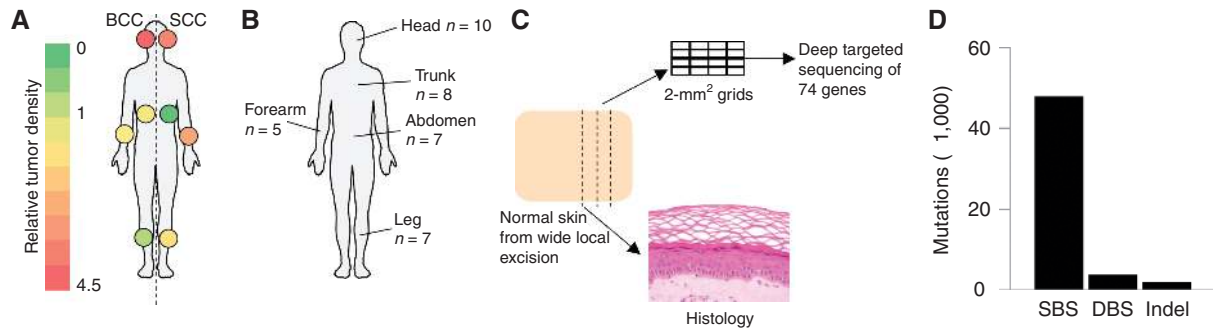
However, as the risk per unit area of developing BCC or SCC differs substantially between body sites (Fig. 1A), a more systematic investigation across different parts of the skin beyond the face is warranted to understand the influence of

frequency of exposure to UV light, sensitivity to UV injury (sunburn), and also the histologic structure of the skin (3). These observations motivated us to use deep targeted and whole-genome sequencing to perform broad and fine-scale mapping of mutations across different locations. Clonal burden, mutational processes, and competitive selection varied both across body sites and within millimeters in individual skin samples. The comprehensive clonal map of the skin at different locations enables evolutionary characteristics to be compared with UV exposure and cancer risk.

RESULTS

Mapping Mutations in Normal Skin

Normal skin from chronically and intermittently sun-exposed sites was collected from 35 Caucasian donors whose ages ranged from 26 to 79 with a balance of males and females (Fig. 1B; Supplementary Table S1). Information on skin type was available for 24 donors, 19 of whom had pale skin, a history of severe sunburn, and limited ability to tan (Fitzpatrick type 1 and 2; Supplementary Table S1). Eight of



Downloaded from <http://aacrjournals.org/cancerdiscovery/article-pdf/11/2/340/1833965/340.pdf> by guest on 27 August 2022

28 donors with an occupational history had worked outside for five years or more, whereas of the indoor workers six had lived in subtropical or tropical countries.

Samples collected for sequencing appeared normal under a dissecting microscope and the adjacent skin had normal histology as assessed by a specialist skin pathologist. Sheets of epidermis were detached from the underlying dermis and cut into a gridded array of 2-mm² pieces, each containing approximately 10⁵ nucleated cells, of which approximately 3 × 10⁴ were in the proliferative basal cell layer (4). A total of 1,261 2-mm² pieces were sequenced at an average coverage of 690× using a bait set of 74 cancer-associated genes (Fig. 1C; Supplementary Tables S2 and S3; ref. 1). Mutations were called using the ShearwaterML algorithm, which detects mutations present in 1% or less of nucleated cells in the sample (5, 6). In a total area of 25.2 cm² of skin sampled across all donors, we identified 47,977 single-base substitutions (SBS), 3,824 double-base substitutions (DBS), and 2,090 small (<200 bp) insertion or deletion events (indels) after merging mutations shared between adjacent samples (Fig. 1D; Supplementary Table S4; Methods).

Variation in the Density of Mutant Clones

The large difference in SCC and BCC risk and the variation in UV exposure across the body led us to hypothesize there would be substantial differences in the mutational landscape of different sites. However, although some samples from intermittently sun-exposed sites (e.g., abdomen and leg) did show a very low number of mutations, others had a similar number of mutations/2 mm² [clonal mutation density (CMD)] and mutational burden (synonymous SNV per megabase) as frequently sun-exposed sites such as the head (Fig. 1E and F). Most mutations had a small variant allele frequency (VAF), indicating the skin is colonized by a large number of small clones (median VAF = 0.015; Fig. 1G and H), though there is a tendency for clone size to increase with age (Supplementary Fig. S1A).

Although large differences were not seen between body sites, we noted intradonor variability in the number of mutations detected per 2-mm² sample (Fig. 1E; Supplementary Fig. S1B). The accumulation of mutations and the growth of mutant clones are both believed to be stochastic processes, which may lead to differences between samples (7). In stochastic spatial simulations (Methods), we found the observed intradonor range in CMD was consistent with such stochastic variation (Supplementary Fig. S1B and S1C).

We also investigated if the environment differed across skin from individual donors. If this was the case, we would expect

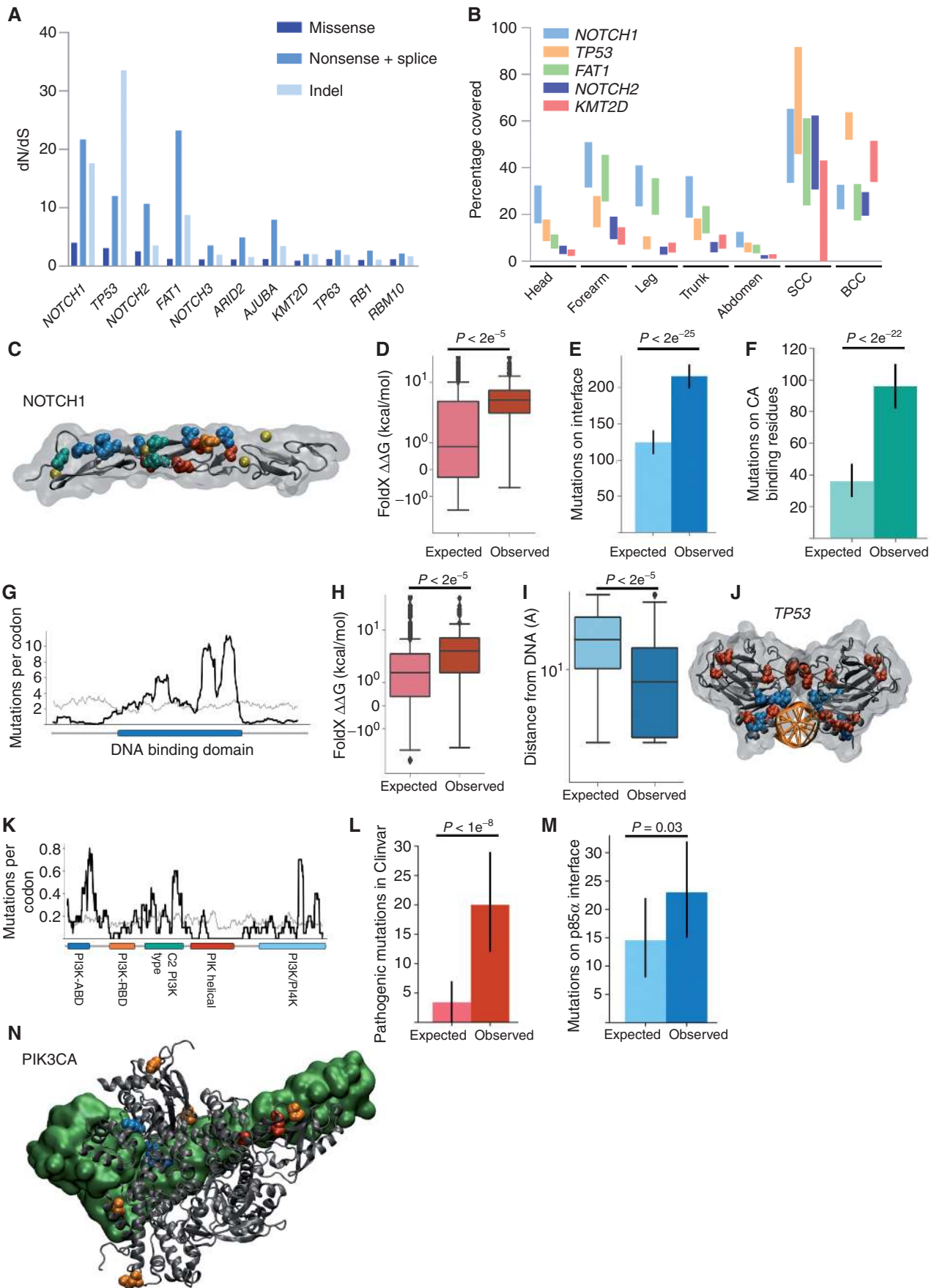
CMD in adjacent samples to be more similar than in more distant samples. We found six individuals who showed distinct gradients in the number of mutations across their epidermal tissue (Supplementary Fig. S1D). In one such specimen from the ear, the sun-exposed, pigmented skin on the outside of the ear had 3-fold higher CMD than the unpigmented skin at the back of the ear (Fig. 1I). Similar variability was seen in five other samples from forearm, breast, earlobe skin, eyebrow, and eyelid (Supplementary Fig. S1E). In the other 29 donors, there was no significant spatial clustering of CMD (Fig. 1J). Thus, the local level of mutations may vary due to both local differences in UV exposure and stochastic fluctuation.

Positive and Negative Genetic Selection

The organization of the skin allows mutant clones to spread laterally and compete for space both with wild-type cells and clones carrying other mutations, a potential selective mechanism in which only the fittest mutants persist (2, 7). Analysis of the dN/dS ratio, which calculates the ratio of protein altering to silent mutations, is a test for competitive selection. Genes with a dN/dS ratio >1 show that nonsynonymous mutations in the gene increase cellular fitness, permitting clonal expansion. Eleven genes were found to be under positive selection after correcting for genomic context, mutational spectrum, and multiple testing (refs. 1, 5; Fig. 2A; Supplementary Fig. S2A and S2B; Supplementary Table S5). Six of these genes (*NOTCH1*, *TP53*, *NOTCH2*, *NOTCH3*, *FAT1*, and *RBM10*) have been previously identified as being positively selected in eyelid skin (1); however, five are novel drivers. Of the newly identified mutant genes, the p53 family transcription factor *TP63* regulates Notch signaling and keratinocyte differentiation, whereas the *KMT2D* histone-modifying lysine methyl transferase interacts with *TP63* and promotes expression of its target genes. The SWI/SNF complex component *ARID2* and *AJUBA*, which regulate Notch and Hippo signaling, are both recurrently mutated in squamous cancers of the skin and other organs. Thus, positively selected mutant genes are predominant regulators of epidermal proliferation and/or differentiation.

The proportion of skin colonized by mutant clones in each sample may be estimated by summing their VAFs for each mutant gene, within upper and lower bounds that allow for uncertainty over copy number (5). Five positively selected mutant genes, *NOTCH1*, *TP53*, *FAT1*, *NOTCH2*, and *KMT2D*, occupied a substantial proportion of normal skin (Fig. 2B; Supplementary Fig. S2C). At several sites, the prevalence of mutant *NOTCH1* and *FAT1* was similar to that found in BCC

Figure 1. Human epidermis is a patchwork of small competing mutant clones. **A**, Tumor density (cancers/unit area) for BCC and SCC across different body sites; whole body = 1.0. Data from ref. 3. **B**, Samples were collected from a variety of body sites with traditionally differing levels of sun exposure. **C**, Experimental outline: normal peeled epidermis was cut into 2-mm² grids and sequenced at high depth using a 74-gene panel bait set. Mutations were called using the ShearwaterML algorithm. A total of 35 patients were sequenced across 37 positions clustered into five main sites: abdomen, head, forearm, leg, and trunk. An average of 39 2-mm² grids were sequenced per position. **D**, Total number of single-base substitutions (SBS), double-base substitutions (DBS), and insertion/deletion (indel) events across all 2-mm² grid samples from all body sites. **E**, Number of mutations per 2-mm² grid from all 1,261 samples from the 35 patients. Each point represents a single 2-mm² grid. Patients are ordered by age and divided by body site. M, male; F, female. **F**, Average mutational burden (substitutions/Mb) per patient per site. Identical mutations in adjacent samples of a patient are merged (Methods). **G**, Distribution of the number of mutations and maximum VAF of all 2-mm² samples per patient. Each point represents a single 2-mm² sample, and point colors represent the maximum VAF in that 2-mm² sample. Patients are ordered by age. **H**, VAF distribution ordered by body site. Where adjacent grids share the same mutation, this is assumed to be the same clone, so the allele frequencies are summed as described in the Methods. **I** and **J**, Heat maps showing the variation in the number of mutations across the epidermis. **I** shows an example from a contiguous strip of skin divided into three blocks from an ear, where 2-mm² samples carrying a high number of mutations cluster together. This is not the case for the leg sample shown in **J**. In **J**, the two blocks of samples are approximately 1 cm apart. In **I**, the increase in the number of mutations correlates with an increase in the pigmentation seen in the epithelia.



Downloaded from <http://aacrjournals.org/cancerdiscovery/article-pdf/11/2/340/1833965/340.pdf> by guest on 27 August 2022

and SCC, hinting that these genes may appear as cancer drivers by virtue of colonizing normal skin rather than promoting tumor formation (Fig. 2B). In contrast, *TP53*, *NOTCH2*, and *KMT2D* mutants are enriched in both SCC and BCC compared with normal tissue.

The large number of missense mutations in *NOTCH1* and *TP53* allowed us to explore mutant selection in more detail. Consistent with previous data sets, missense mutations in *NOTCH1* are concentrated in the ligand-binding region (Supplementary Fig. S2D; ref. 5). Missense mutations that destabilize the protein, likely to be functionally equivalent to nonsense mutations, or mutations that disrupt the ligand-binding interface or calcium-binding sites, which are expected to interfere with ligand-mediated NOTCH activation, are significantly selected (Fig. 2C–F). Most missense mutations in *TP53* occur in the DNA-binding domain (Fig. 2G), and were either protein destabilizing or clustered close to the DNA-binding surface (Fig. 2H–J).

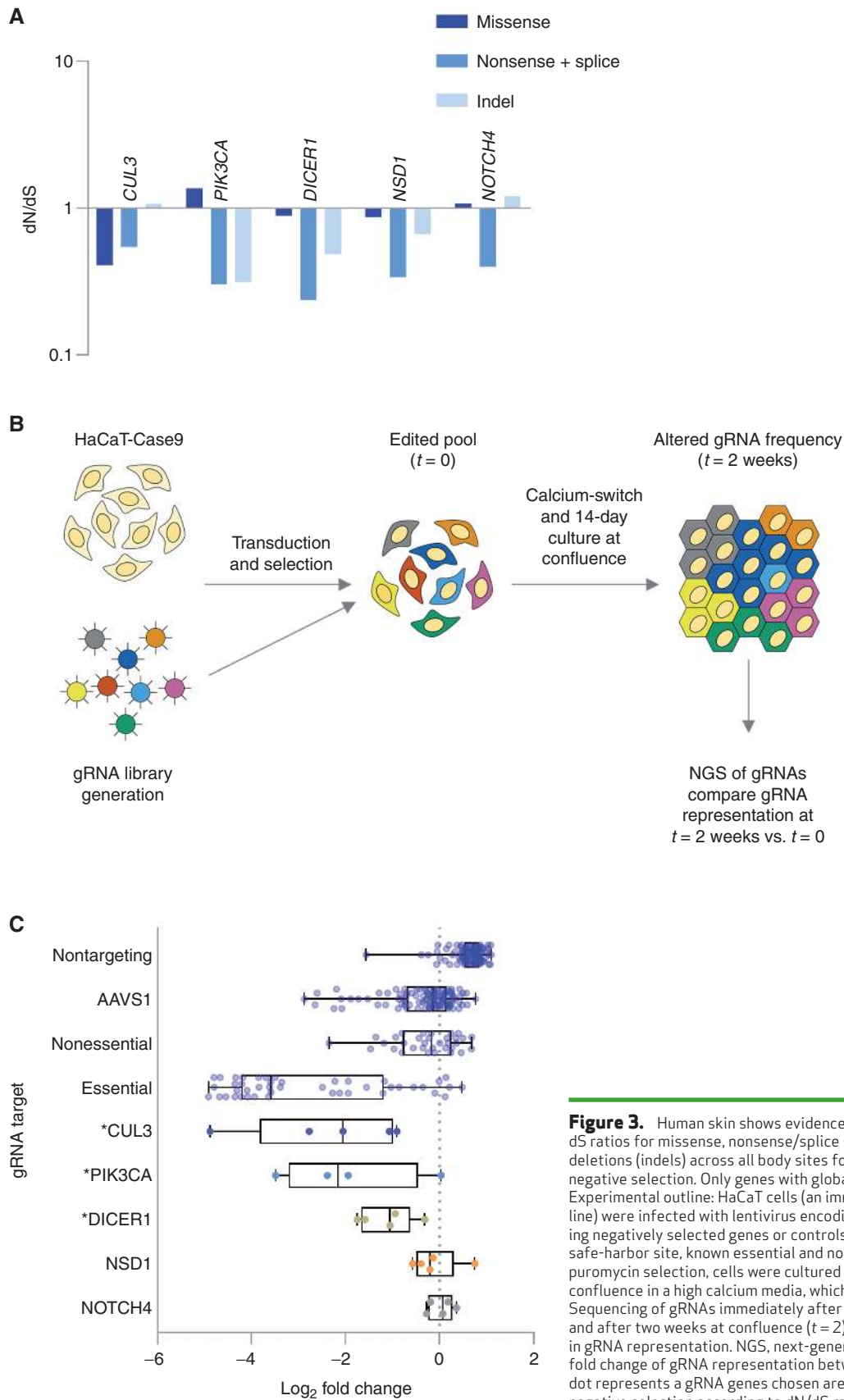
Alongside missense mutations that inactivate proteins, there are also activating point mutations. These may not be detected by dN/dS analysis if they are outnumbered by neutral or inactivating mutations in a gene. We found examples of canonical “hotspot” activating mutations in receptor tyrosine kinases (*EGFR*, *ERBB2*, *ERBB3*, and *FGFR3*), signaling intermediates (*KRAS*, *HRAS*, *AKT1*, and *PIK3CA*), and the redox response regulatory transcription factor *NFE2L2*. The most frequent activating mutations were in *PIK3CA*, which encodes the catalytic p110 α subunit of PI3K (Fig. 2K). The dN/dS results showed that nonsense mutations in *PIK3CA* were negatively selected, but that missense selection was not significantly different from neutral. However, there was a significant enrichment of specific *PIK3CA* mutations associated with various cancers or overgrowth diseases that are known to be driven by *PIK3CA*-activating mutations (Fig. 2L; ref. 8). This suggests that there is negative selection of inactivat-

ing *PIK3CA* mutations and positive selection of activating *PIK3CA* mutations. To test this hypothesis, we assessed the selection of mutations in *PIK3CA* predicted to alter the binding of p110 α protein to the inhibitory p85 α PI3K subunit, finding a small, but statistically significant, enrichment of mutations at the interface of p110 α with p85 α ($P = 0.03$, two-tailed binomial test; Fig. 2M and N). We conclude that activating *PIK3CA* mutants are positively selected, but mutations that inhibit its function are under negative selection.

In addition to *PIK3CA*, we also found a further four mutant genes that appeared to be under negative selection, with a dN/dS ratio below 1 (Fig. 3A). *PIK3CA*, *DICER1*, *CUL3*, *NSD1*, and *NOTCH4* were all depleted for protein-truncating mutations and indels relative to the number of silent mutations. This was not due to variations in sequencing depth at the gene level (Supplementary Fig. S3A, methods). Synonymous mutations are generally distributed evenly across protein-coding genes. However, in the case of *CUL3*, two synonymous mutations within the first exon were highly enriched (Supplementary Fig. S3B). This may reflect altered translational start sites or transcription factor binding due to the nucleotide changes (Supplementary Fig. S3C). Such synonymous mutations with the potential to alter cellular fitness may render the dN/dS ratio an unreliable metric of selection.

To test if the negatively selected mutant genes altered keratinocyte proliferative potential or survival, we performed a pooled lentiviral CRISPR/Cas9 knockout screen with a targeted library of 317 guide RNAs (gRNA) in HaCaT epidermal keratinocyte cells. Whole-genome sequencing of HaCaT cells showed that *PIK3CA*, *CUL3*, *NOTCH1*, *NSD1*, and *NOTCH4* were wild-type, and there was a p.S1631A (c.4891T>G) mutation with a VAF of 0.49 in *DICER1*, which represents a conservative change in a poorly conserved region of the protein and therefore seems likely to be nonpathogenic (<https://www.ncbi.nlm.nih.gov/clinvar/variation/221072/>). gRNA abundance was determined

Figure 2. Positive selection of genes linked with cancer in normal skin. **A**, dN/dS ratios for missense, nonsense/splice substitutions, and insertions/deletions (indel) across all body sites for genes under significant (global $q < 0.01$ and dN/dS > 2) positive selection. **B**, Estimated percentage of cells carrying mutations in the most strongly positively selected genes (*NOTCH1*, *FAT1*, *TP53*, and *NOTCH2*) for each body site as well as for basal and squamous cell carcinomas (see Methods). Upper and lower bound range allows for uncertainty in copy number and biallelic mutations. Upper bound represents no CNA and one mutant allele per gene. **C–F**, Positive selection of categories of missense mutations in *NOTCH1* EGF repeats 11–12 that form part of the ligand-binding domain. **C**, Structure of *NOTCH1* EGF11–13 (PDB 2VJ3). Residues containing missense mutations that occur >10 times are highlighted. Ligand-binding interface residues, blue; calcium-binding residues, green; destabilizing residues, red; D464N, orange, does not fit into the previous categories. Calcium ions are shown in yellow. **D**, Missense mutations that are not on ligand-interface or calcium-binding residues are significantly more destabilizing than would be expected under neutral selection ($P < 2e^{-5}$, $n = 452$, two-tailed Monte Carlo test, Methods). **E**, Non-calcium-binding missense mutations with $\Delta\Delta G < 2$ kcal/mol (i.e., are not highly destabilizing) occur on the ligand-binding interface significantly more than would be expected under neutral selection ($P = 2e^{-25}$, $n = 315$, two-tailed binomial test, error bars show 95% confidence intervals, Methods). **F**, Missense mutations with $\Delta\Delta G < 2$ kcal/mol (i.e., are not highly destabilizing) and that are not on the ligand-binding interface occur on calcium-binding residues significantly more than would be expected under neutral selection ($P = 2e^{-22}$, $n = 195$, two-tailed binomial test, error bars show 95% confidence intervals, Methods). **G** and **H**, Positive selection of missense mutations in *TP53*. **G**, Sliding window plot of missense mutations per codon in *TP53*. Observed counts shown by the black line. Expected counts assuming that missense mutations were distributed across the gene according to the mutational spectrum (Methods) shown in gray. DNA-binding domain (DBD) of *TP53* shown in blue below the x-axis. **H**, Missense mutations in the *TP53* DBD that are more than 5 Å from the DNA are significantly more destabilizing than would be expected under neutral selection ($P < 2e^{-5}$, $n = 760$, two-tailed Monte Carlo test, Methods). **I**, Missense mutations with $\Delta\Delta G < 2$ kcal/mol (not highly destabilizing) in the *TP53* DBD are significantly closer to the DNA than would be expected under neutral selection ($P < 2e^{-5}$, $n = 395$, two-tailed Monte Carlo test, Methods). **J**, Structure of the *TP53* DNA-binding domain (PDB 2AC0) bound to DNA (orange). Residues containing missense mutations that occur at least 10 times are highlighted. Highly destabilizing mutations ($\Delta\Delta G > 2$ kcal/mol) shown in red. Non-destabilizing mutations are shown in blue. **K–N**, Positive selection of missense mutations in *PIK3CA*. **K**, Sliding window plot of missense mutations per codon in *PIK3CA*. Observed counts shown by the black line. Expected counts assuming that missense mutations were distributed across the gene according to the mutational spectrum (Methods) shown in gray. Domains of *PIK3CA*-encoded protein are shown below the x-axis. **L**, Significantly more single-nucleotide substitutions in *PIK3CA* are annotated as pathogenic/likely pathogenic in the Clinvar database than would be expected under neutral selection ($q = 1e^{-8}$, $n = 216$, two-tailed binomial test; error bars show 95% confidence intervals, Methods). **M**, Significantly more missense mutations in *PIK3CA* occur in codons at the interface binding PIK3R1 (defined as PIK3CA residues with atoms within 5 Å of PIK3R1 in PDB 4L1B) than would be expected under neutral selection ($P = 0.03$, $n = 157$, two-tailed binomial test; error bars show 95% confidence intervals, Methods). **N**, Structure of PIK3CA protein, gray, bound to PIK3R1, green (PDB 4L1B). Residues with mutations occurring at least three times are highlighted. Mutations close to PIK3R1 shown in blue; other mutations that are annotated as pathogenic/likely pathogenic are shown in red; all others are shown in orange.



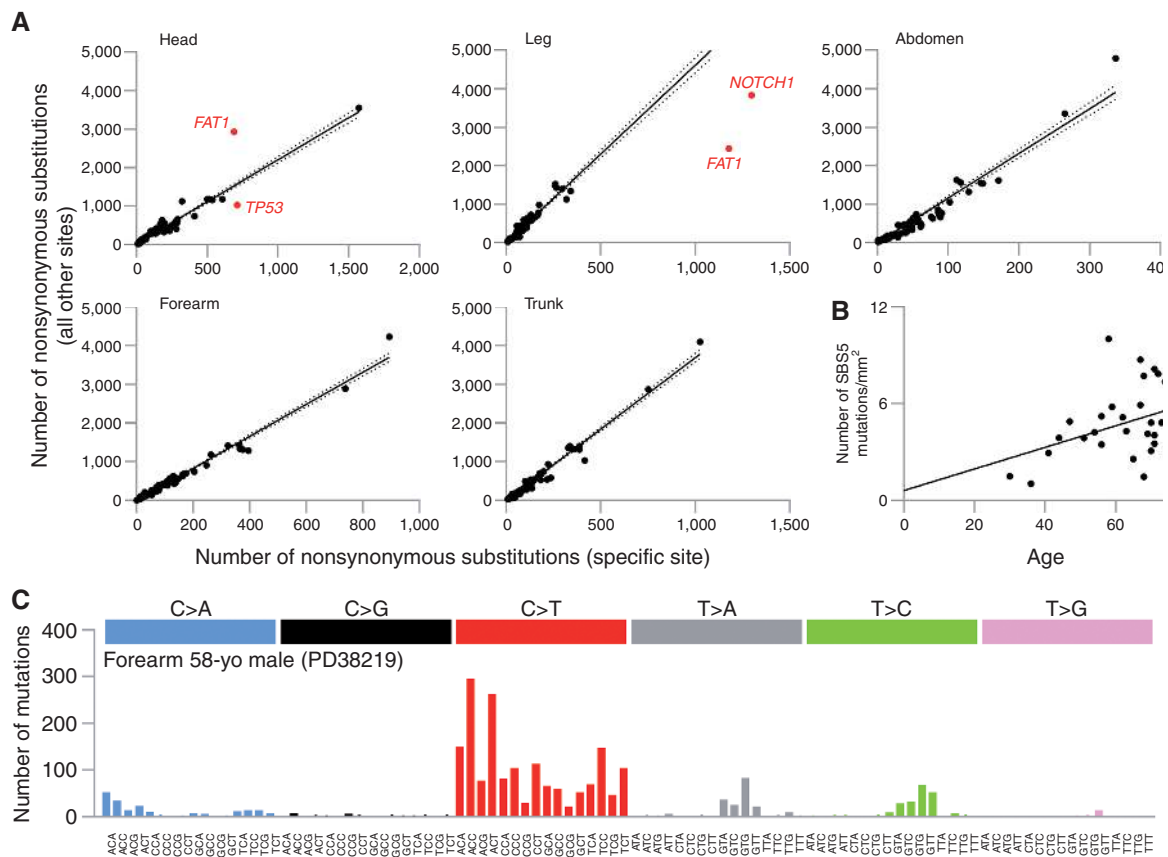


Figure 4. Human epidermis shows selection and signature variation between body sites. **A**, Differential selection of *TP53*, *NOTCH1*, and *FAT1* across different body sites. Number of nonsynonymous mutations per gene from indicated body site versus all nonsynonymous mutations in that gene from all other body sites. Each dot represents a gene from the 74-gene targeted bait panel. Solid line represents trend line with 95% CI marked (dotted line). For points highlighted in red, $q < 1 \times 10^{-6}$ using a likelihood ratio test of dN/dS ratios (Methods). **B**, Correlation between patient age and number of mutations attributed to signature SBS5. Linear regression $P = 0.0299$, slope 0.067, intercept 0.608. **C**, Trinucleotide spectrum for SBS of donor PD38219, which shows considerable differences to all other donors. This signature is consistent with SBS32 (cosine similarity 0.95) linked with azathioprine treatment. (continued on next page)

following lentiviral infection and after 14 days in culture (Fig. 3B). The specific depletion of gRNAs targeting known essential genes compared with those targeting known nonessential genes, or the safe-harbor AAVS1 locus, confirmed the ability of this assay to detect the dropout of gRNAs affecting cell growth or survival (Fig. 3C). Of the genes under negative selection, gRNAs targeting *CUL3*, *PIK3CA*, and *DICER1* were significantly depleted, arguing knockout of these genes is detrimental to keratinocyte growth or survival *in vitro* (Fig. 3C; Supplementary Table S6). gRNAs targeting *NSD1* and *NOTCH4* were not significantly depleted, which may reflect differences between the cultured HaCaT cells and keratinocytes in the epidermis. Although a CRISPR screen will, for any given single guide RNA (sgRNA), generate a cell population containing inactivating indels in either one or both alleles and therefore imperfectly models the monoallelic loss we see in human epidermis, it does provide evidence that these genes are detrimental to keratinocyte fitness. In addition, developmental mutations argue these genes are haploinsufficient (Decipher database haploinsufficiency scores: *PIK3CA*, 0.26%; *DICER1*, 1.99%; and *CUL3*, 4.57%; <https://decipher.sanger.ac.uk>), suggesting that a single wild-type allele is insufficient to maintain cellular fitness.

Negative selection has not been observed in other epithelial tissues including esophagus and colon (5, 9). To show a mutant gene is negatively selected, a large number of silent mutations must be detected for reduction in the dN/dS ratio below 1 to reach statistical significance. It is possible that in normal tissues other than skin, the density of mutations has not been high enough to reveal negative selection.

Mutant Gene Selection and Mutational Signature Vary across the Body

To investigate gene selection between body sites, we compared the number of nonsynonymous mutations in each gene at each site with that found at all other sites (Fig. 4A). This revealed differential selection with protein-altering *FAT1* and *NOTCH1* mutants overrepresented in the leg. In contrast, *FAT1* mutants were depleted in the head and *TP53* mutants enriched over other sites. This possibly reflects environmental differences between sites such as the frequency of UV exposure.

The majority of mutations observed are consistent with damage induced by UV light, with C>T (or G>A) changes accounting for 77% of all SBS and CC>TT (or GG>AA) changes accounting for 88% of all double-base substitutions

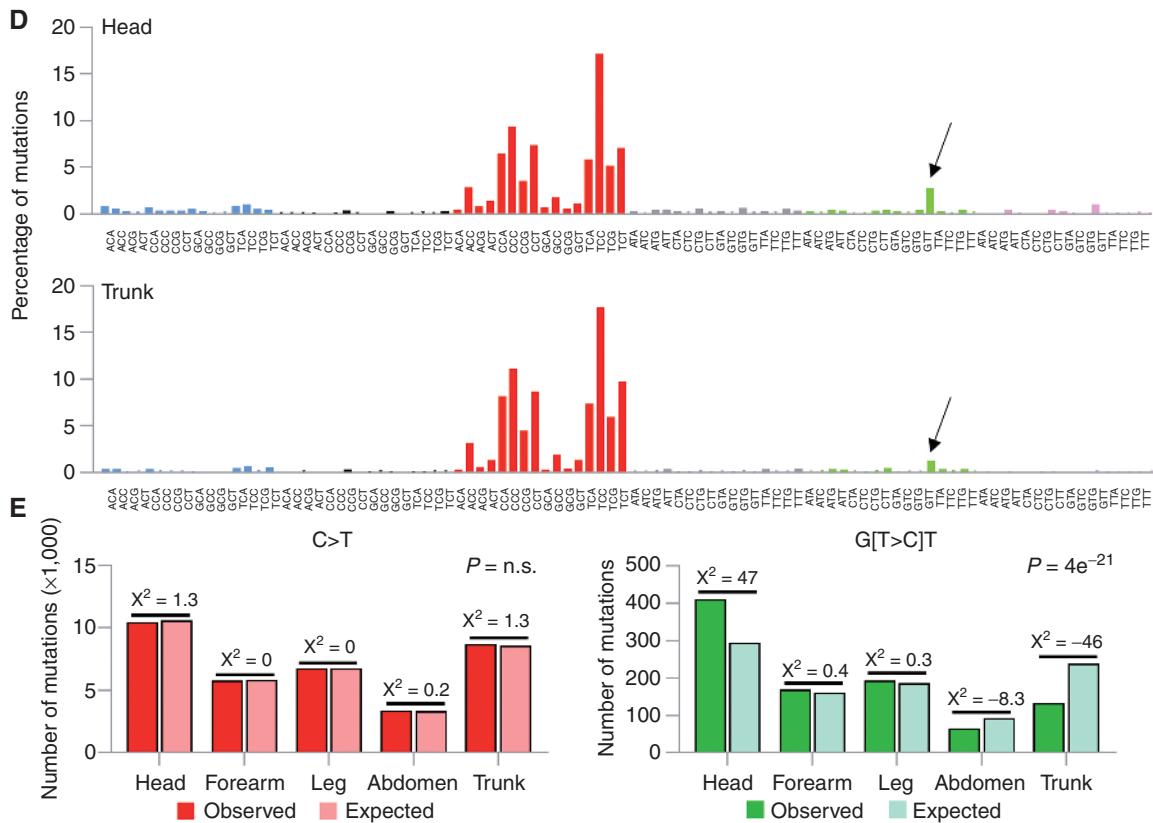


Figure 4. (Continued) D. Combined trinucleotide spectra for SBS in all samples from the head (top) and trunk (bottom) from targeted sequencing. Arrow indicates G(T>C)T peak, which contributes to a higher proportion of overall burden in the head. Head = 14,561 mutations; trunk = 10,515 mutations. **E.** Chi-squared test for all C>T mutations versus G(T>C)T mutations by body site (χ^2 105, df = 4, $P = 0$). A higher proportion of G(T>C)T mutations are observed at sites with higher relative risk of cutaneous squamous cell carcinoma.

(Supplementary Fig. S2B). Although C>T changes can be linked with the aging signature SBS1 as well as the SBS7 UV signatures, once the trinucleotide context was taken into account, only a small fraction of the C>T changes can be attributed to SBS1. In addition, both C>T and CC>TT substitutions were more frequently found on the untranscribed strands of genes, suggesting a link with transcription-coupled nucleotide excision repair, an observation again consistent with SBS7A and SBS7B rather than SBS1 (10). In common with other epithelial tissues, there was a positive correlation between SBS5, a signature associated with cellular aging and patient age (Fig. 4B; linear regression $P = 0.0299$, slope 0.067, intercept 0.608; refs. 5, 9). However, the contribution of SBS5 is relatively modest compared with that of SBS7 (average 19.8% and 76%, respectively).

Using the trinucleotide context of each mutation, four single-base substitution reference signatures associated with damage by UV light (SBS7A-d) have recently been characterized in cancers of the skin (10). SBS7A and SBS7B are likely to arise from UV-induced cyclobutane pyrimidine dimers or 6–4 photoproducts, with recent work showing that SBS7B is enriched in open chromatin and SBS7A in quiescent regions, perhaps caused by differences in repair due to genomic context (11). It has been hypothesized that SBS7C and SBS7D may be due to translesion DNA synthesis by error-prone polymerases inserting T rather than A, or G rather than A,

opposite UV damage, respectively (10). In all but one donor (Fig. 4C), trinucleotide spectra were characteristic of SBS7A, SBS7B, and SBS7D composition (Fig. 4D); however, mutations consistent with SBS7C (T[T>A]T) were not observed at a frequency any greater than other T>A mutation contexts.

Forearm skin from a 58-year-old male (PD38219) was an outlier in terms of mutational spectrum (Fig. 4C) and also this was the donor with the highest mutation burden (Fig. 1E). This spectrum has a cosine similarity of 0.95 with the reference signature SBS32 (10). SBS32 was first identified in patients treated with azathioprine to induce immunosuppression and is characterized by a predominance of C>T mutations (75%) with a very strong transcriptional strand bias favoring higher rates of C>T mutations on the transcribed strand. The strand bias is consistent with azathioprine interfering with purine synthesis leading to incorporation of guanine derivatives into DNA, which lead to C>T (G>A) mutations, especially with guanosine alterations on the coding strand, that escape nucleotide excision repair. This strand bias is in contrast to that caused by UV damage, which induces cyclobutane pyrimidine dimers, which are more efficiently repaired by transcription-coupled nucleotide excision repair on the transcribed strand, leading to higher rates of C>T in template strand orientation. Here, this donor was on medication for immunosuppression

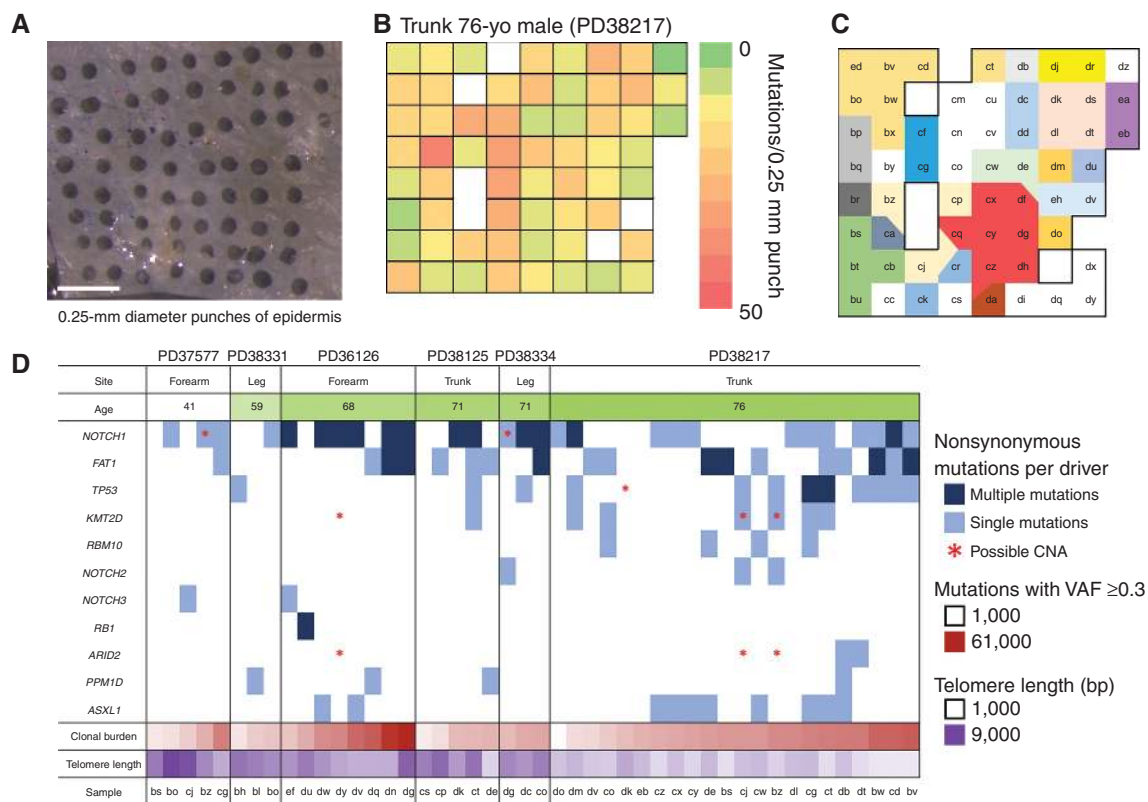


Figure 5. Variation in mutational load, mutational signatures, and telomere length at fine-scale resolution. **A**, 0.25-mm diameter punches were taken in a gridded array format from peeled epidermis. Samples were sequenced using a 324-gene bait panel and mutations identified using ShearwaterML as previously described. Two hundred thirty-two punches from six individuals from leg, trunk, and forearm were sequenced. A subset of samples were subsequently whole-genome sequenced. Scale bar = 1 mm. **B**, Heat map of a single individual (Trunk 76-year-old male PD38217; shown in **A**) showing the number of mutations per 0.25 mm punch detected from targeted sequencing. **C**, Clonal map of the same individual as in **B**. A filtered set of mutations with a variant allele fraction ≥ 0.2 was used to spatially map the clones; letters indicate individual samples shown in panel **D**; each color denotes a separate clone. White is used for polyclonal samples. Samples with too low DNA yield for sequencing have been removed from the map. **D**, Plot summarizing the mutations (VAF ≥ 0.3) and copy-number aberrations for genes identified as being under positive selection in targeted sequencing data for 46 whole-genome punch samples. Age of donor, number of clonal (VAF ≥ 0.3) mutations, and telomere length for each sample is shown. Not all events are independent as some samples are part of the same clone (Fig. 5E; Supplementary Fig. S5). (continued on next page)

since age 24 following a kidney transplant. Despite the high cosine similarity of this donor's spectrum to SBS32, it is likely that SBS32 itself is contaminated by the SBS7 (UV) signatures as it was first characterized in cutaneous SCCs (12). The majority (58%) of DBS in this donor were CC>TT substitutions but with the opposite transcriptional strand bias to UV-induced DBS. Long-term azathioprine treatment may generate most of the mutations in the normal skin of transplant recipients.

In all other donors, we noticed a visible difference in the frequency of G[T>C]T mutations (Fig. 4D). A χ^2 test revealed a significant difference in the proportions of G[T>C]T mutations and all C>T mutations by body site ($\chi^2 = 105$, $df = 4$, $P = 0$). In fact, the proportion of G[T>C]T mutations was found to increase with the increased relative risk of cutaneous squamous cell carcinoma (cSCC) at that site (Fig. 4E; ref. 3). G[T>C]T mutations make up the predominant peak of SBS7D, suggesting a lower proficiency for DNA repair after UV damage at sites at higher risk of cSCC. We conclude that exposure to UV light dominates over aging in generating mutations in normal skin, and that the nature of UV-induced

DNA damage and repair differs between high and low cancer risk sites.

Higher Resolution Mapping and Clonal Genomes

To map mutant clones at increased spatial resolution, 232 additional samples were taken from six donors with a 250- μ m diameter circular punch, collected in a rectangular array with approximately a punch diameter between each sample (Fig. 5A). Each sample corresponds to approximately 2,400 nucleated cells, of which 750 were in the basal layer. Although 23 samples could not be processed due to low DNA yield, the remaining samples were sequenced for 324 cancer-associated genes and the mutations called with ShearwaterML as described above (Fig. 1C; Supplementary Tables S2 and S3). A total of 2,805 mutations were identified with a similar mutational spectrum and distribution of SBS, double-base substitutions, and indels to the larger skin samples (Supplementary Fig. S4A–S4C; Supplementary Table S7). The number of mutations per sample was highly variable within each individual site (Fig. 5B; Supplementary Figs. S4B and S5A). Two additional genes in the expanded gene set, *PPM1D*

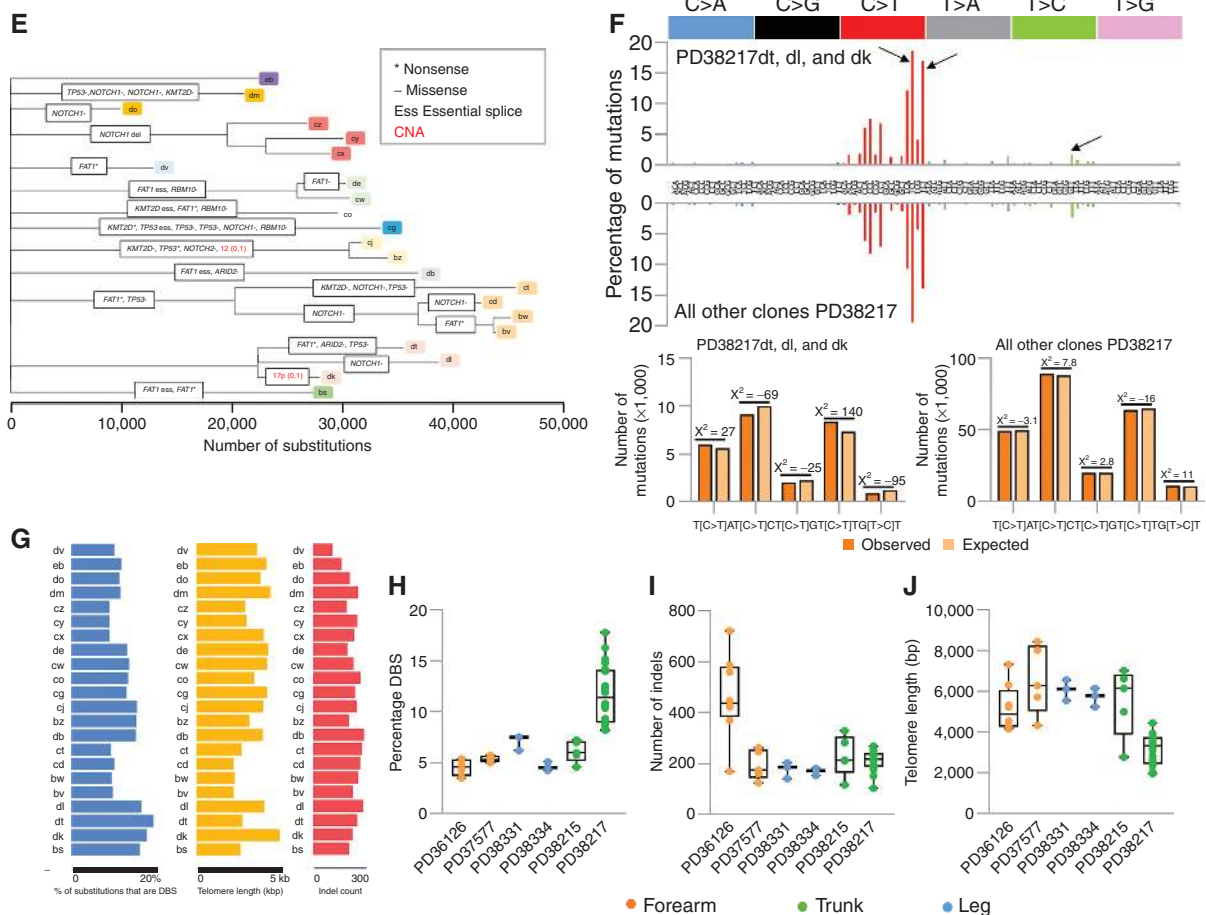


Figure 5. (Continued) **E**, Maximum parsimony tree of clonal substitutions detected in 32 whole-genome punch samples of trunk skin from a 76-year-old male (PD38217). Branch lengths are equivalent to the number of clonal single- and double-base substitutions and are annotated with clonal nonsynonymous mutations detected in the 13 genes found to be under positive selection. Within each branch, driver mutations are arbitrarily ordered. CNAs are shown in red. **F**, Combined trinucleotide spectra for SBS assigned to branches of the dt/dl/dk clade (top) versus those assigned to all other branches (bottom) from the same individual as in **E** (PD38217). Arrows show trinucleotide contexts found to differ the most between these two groups ($\chi^2 = 397$, $df = 4$, $P = 0$). Clone dt, dl, dk = 49,144 mutations; all other clones = 458,479 mutations. **G**, Variation in a single individual (PD38217) in the percentage of substitutions that are DBS, telomere length, and number of clonal (VAF of at least 0.3) insertion/deletions (indel) per whole-genome sample. **H–J**, Variation in the percentage of substitutions that are double base (**H**), number of indels (**I**), and telomere length (**J**) per whole-genome sample, by donor. Three different body sites are shown: forearm (orange), trunk (green), and leg (blue).

and *ASLX1*, were under positive selection (Supplementary Fig. S4D and S4E). *PPM1D* is a regulator of *TP53* under positive selection in normal esophagus whereas mutations of the epigenetic regulator *ASLX1* are selected in squamous head and neck cancer. A filtered set of mutations with a variant allele fraction of at least 0.2 was used to spatially map the clones present in the tissue (Fig. 5C; Supplementary Fig. S5B). The density of clones varied substantially between individuals and within sites; for example, trunk samples from PD38217 showed the highest density of clones (Fig. 5C), whereas the trunk skin from donor PD38215 was lightly mutated (Supplementary Fig. S5B).

In order to better understand the genome-wide mutational burden, mutational processes, and copy-number aberrations, 46 samples from 34 different clones were submitted for whole-genome sequencing (WGS). All but three clones harbored one or more of the mutant genes under positive

selection in the skin; no additional genes under selection were detected (Figs. 2A and 5D; Supplementary Fig. S4D). The majority of clones (58%) carried two or more positively selected mutant genes (Supplementary Fig. S4F). In terms of copy-number alteration (CNA), we found two genomes with clonal nonsynonymous mutations in *NOTCH1* that displayed loss of heterozygosity at 9q, the *NOTCH1* locus. In addition, a clone spanning samples PD38217cj and bz had a clonal deletion of one whole allele of chromosome 12. This clone harbored a missense mutation in *KMT2D*, a positively selected driver of clonal expansion located on chromosome 12. In contrast, PD36126dy had an amplification for one whole allele of chromosome 12 (this sample had clonal nonsynonymous mutations in 18 genes located on this chromosome, but none were under selection or seem to be relevant to skin biology or carcinogenesis). Finally, one sample, PD38217dk, has a deletion of one allele at 17p, the *TP53* locus (Fig. 5D).

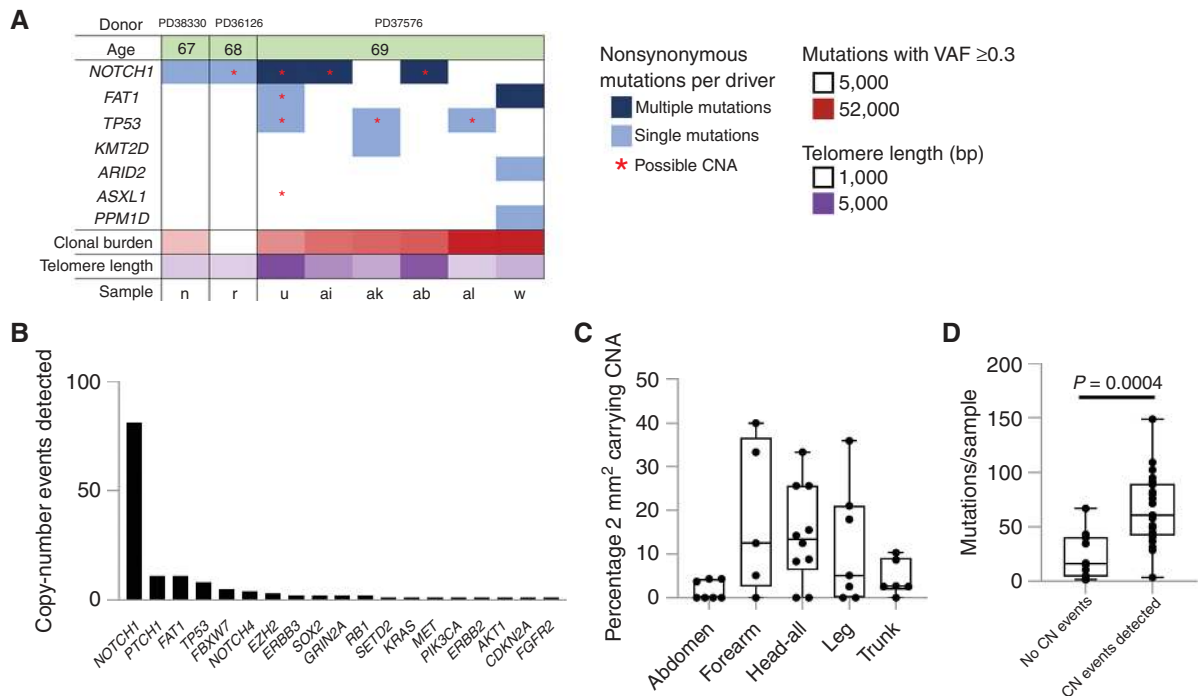


Figure 6. Normal human skin shows frequent copy-number changes including loss of heterozygosity of *PTCH1*. **A**, Plot summarizing the mutations (VAF ≥ 0.3) and CNAs for genes identified as being under positive selection in targeted sequencing data for eight clonal whole-genome 2 mm² grid samples. Age of donor, number of clonal (VAF ≥ 0.3) mutations, and telomere length for each sample is also shown. Not all events are independent because some samples are part of the same clone. Sample PD37576u shows multiple changes with loss of heterozygosity at 4q (*FAT1*), 9q (*NOTCH1*), 17p (*TP53*) and 20q (*ASXL1*). **B**, Number of copy-number events per gene detected by SNP phasing of targeted sequence data in all 1,261 2mm² grid samples. **C**, Percentage of grids carrying a copy-number event detectable by SNP phasing segregated by body site. Each dot represents an individual. **D**, Average number of mutations per 2 mm² grid in patients either carrying or not carrying a copy-number (CN) event $P = 3.8 \times 10^{-4}$. Student two-tailed *t* test. (continued on next page)

Phylogenetic trees were constructed using shared clonal single-base and double-base substitutions across all whole-genome samples for each donor (Fig. 5E; Supplementary Fig. S5C). We observed a 2- to 3-fold variation in the number of clonal substitutions per genome in the same donor, despite the clones being almost adjacent in the skin.

Comparison of the proportion of the mutations in each trinucleotide context revealed differences between the branches. The spectrum of the clone that covers samples dt, dl, and dk in PD38217 showed several differences when compared with all other clones sequenced in this donor ($\chi^2 = 397$, $df = 4$, $P = 0$, Fig. 5F). Each branch of this clade was found to have an increase of T[C>T]T mutations and fewer T[C>T]C and G[T>C]T mutations, with the G[T>C]T mutations being characteristic of SBS7D. Because the differences we observe in the frequency of mutations in these trinucleotide contexts were common to branches of the same clone, it is likely they are caused by biological differences in the cells of that clone (see Methods). In the whole-genome data, we do detect an increased frequency of T[T>A]T mutations, characterized as SBS7C, that was not observed in the targeted data. However, the contribution of these mutations to the overall burden is small and less than that previously described in skin cancers (10). We conclude that genuine differences in UV damage or repair exist between clones in very close spatial proximity.

We found double-base substitutions, indels, and telomere length also varied between clones, from the same individual in

some cases by nearly 2-fold (Fig. 5G). The intraindividual range in DBS and indels was as large as that between donors (Fig. 5H and 5I, respectively). Telomere length varied in a similar manner; indeed, the variation seen in individuals was similar to that across cancers derived from different tissues (Fig. 5J).

CNA in Normal Skin

Given the presence of CNA in 6 of 46 genomes from micro-punch biopsies, we explored other samples for evidence of genomic instability. Eight (0.4%) 2-mm² samples were also clonal and were submitted for WGS. We found nonsynonymous mutations in at least one driver gene in all samples, and CNA of one or more of *NOTCH1*, *FAT1*, *ASXL1*, and *TP53* in six of the eight samples (Fig. 6A).

To get further insight into the extent of CNA, we used a heterozygous SNP-phasing approach to detect subclonal allele loss of the genes in the targeted sequencing of 2-mm² samples (Fig. 6B; ref. 5). This revealed recurrent allele loss of *NOTCH1*, *NOTCH1* and *PTCH1*, *TP53*, *FAT1*, and *FAT1* and *FBXW7*. There was a higher prevalence of CNA in frequently sun-exposed head and forearm skin that reached borderline significance ($P = 0.06$ one-way ANOVA), though changes were seen at all body sites (Fig. 6C). Across all samples, there were significantly more mutations in samples containing CNAs (Fig. 6D, Student two-tailed *t* test, $P = 3.8 \times 10^{-4}$).

PTCH1 is of particular interest as it mutated in almost all cases of BCC but is not under selection in the epidermis (13).

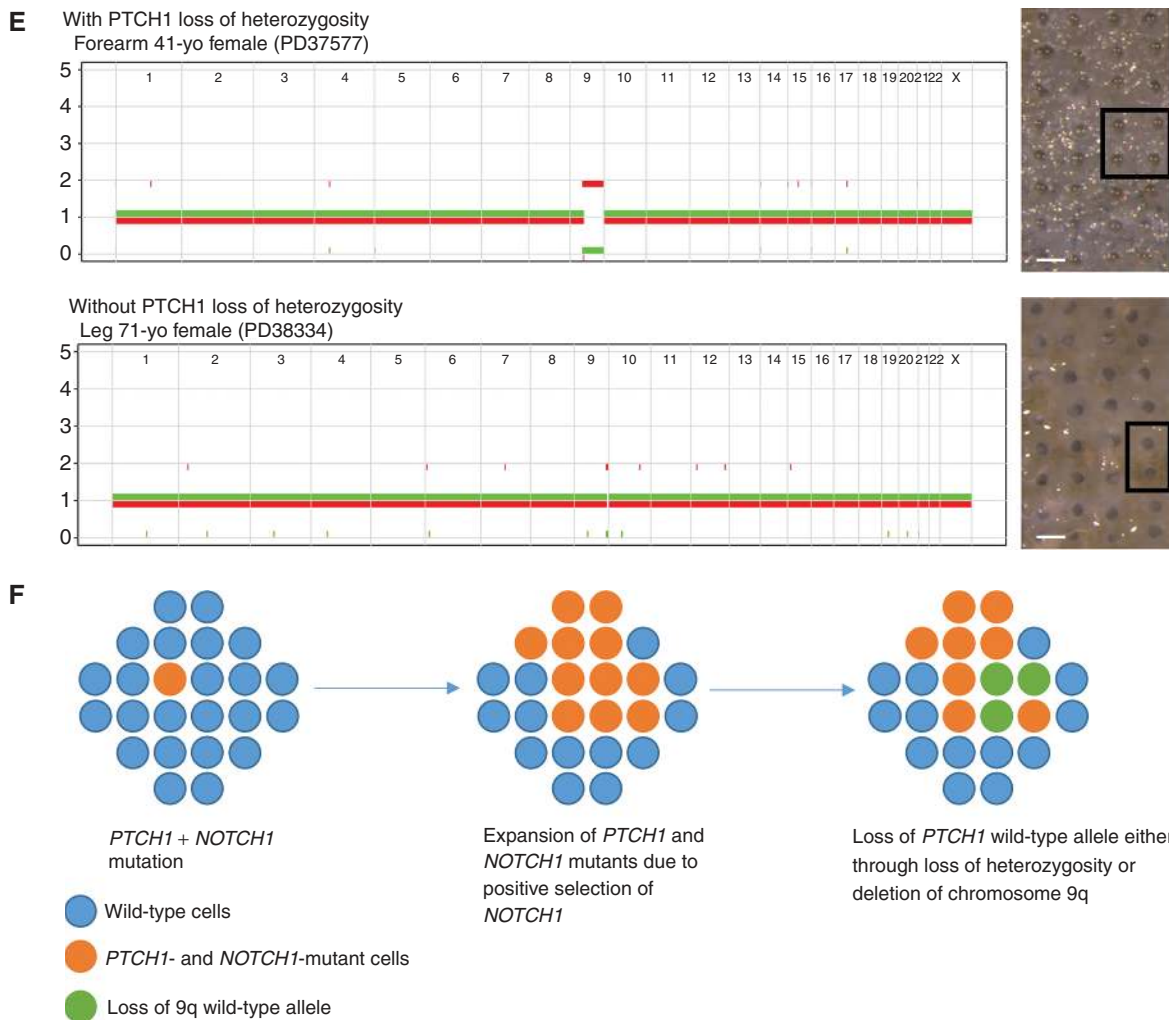


Figure 6. (Continued) E, Copy-number profile of whole-genome sequenced punch samples showing chromosome 9 loss of heterozygosity. The top example shows loss of heterozygosity for both *NOTCH1* and *PTCH1*; the bottom example shows just *NOTCH1* LOH. Scale bar, 0.5 mm. **F**, Possible origin of BCC. In a wild-type population of cells (blue circles), a single cell acquires *PTCH1* and then *NOTCH1* nonsynonymous mutations (marked orange), and the clone size expands and persists due to *NOTCH1*-positive selection. At a later time point the wild-type *PTCH1* allele is lost either through deletion or LOH (marked green), thus leaving a clone lacking functional *PTCH1* expression and primed for BCC transformation.

We identified examples of clones that had undergone LOH of 9q, some including *PTCH1*, a gene frequently showing CNA in BCC (Fig. 6E). In addition, the relatively close proximity of *PTCH1* to *NOTCH1* suggests a potential mechanism for the persistence of *PTCH1* mutations, despite their not being under positive selection (Fig. 6F). Should a *PTCH1*-mutant cell acquire a *NOTCH1* mutation, it will be carried as a passenger in a clonal expansion driven by *NOTCH1*. Subsequent loss of the wild-type 9q allele involving *PTCH1* as well as *NOTCH1* would result in biallelic disruption of *PTCH1* in a highly competitive clone, with the potential to generate a BCC. We conclude that normal skin harbors substantial numbers of clones with CNA, with recurrent LOH of positively selected mutants linked to cancer, particularly in frequently sun-exposed sites.

Mutations within Hair Follicles

Finally, we assayed mutations in hair follicles, appendages that are discrete self-maintaining units and have been

proposed as a site of origin of BCC and SCC (Fig. 7A; ref. 14). Each follicle undergoes cyclical expansion and regression throughout life. Coarse (intermediate type) follicles at the largest (anagen) stage of the hair cycle were isolated from forearm, leg, and trunk skin of seven donors. We took a 250- μ m diameter micro biopsy from adjacent epidermis and then dissected each follicle into thirds that were sequenced as described for the punch biopsies (Fig. 7B). The upper section included the funnel-shaped infundibulum that is in continuity with the epidermis. The mid follicle included the bulge region that contains stem cells that maintain the follicle base as it undergoes cyclical expansion and regression. Confocal imaging of adjacent follicles suggested there was minimal contamination with underlying dermis (Fig. 7C).

The total number of mutations varied along the follicle; the number of mutations in the upper follicle was similar to that in the adjacent epidermis, but lower in the mid follicle and base, consistent with the limited depth penetration of UV light into

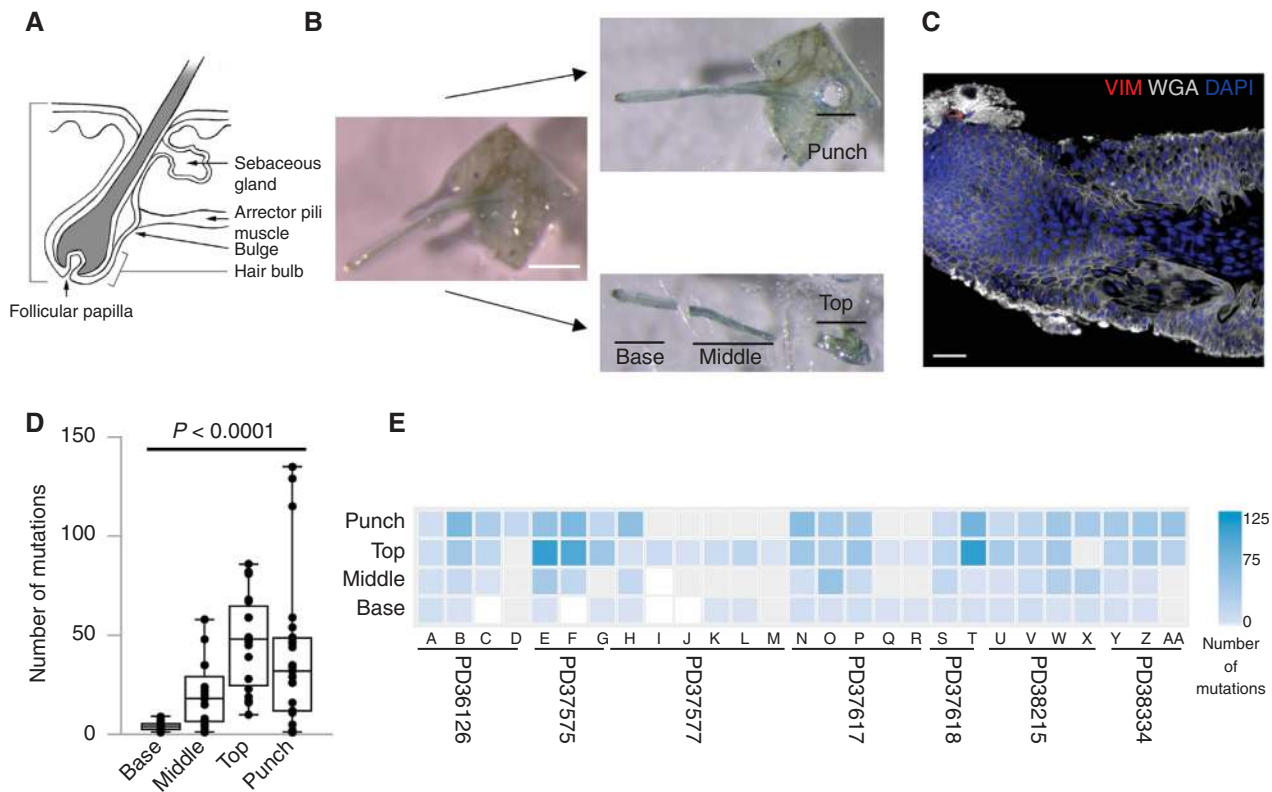


Figure 7. Human hair follicles are polyclonal with the base of the follicle showing differences in mutational load compared with the top. **A**, Structure of the human hair follicle (<https://emedicine.medscape.com/article/835470-overview>). **B**, Experimental outline: Intact hair follicles are dissected from the epidermis and cut into three (designated base, middle, and top). A 0.25-mm punch of epidermis was taken adjacent to the follicle (designated punch). All samples were sequenced at high depth using a 324-gene bait panel and mutations identified using ShearwaterML as previously. Scale bar, 0.5 mm. **C**, Example confocal image of hair follicle stained with WGA (white), vimentin (red), and DAPI (blue). Scale bar, 38 μ m. **D**, Distribution of the number of mutations in different parts of the follicle. Each dot represents a sample. Solid line indicates the median. $P < 0.0001$ Kruskal-Wallis test. **E**, Number of exonic mutations per follicle across 324 genes. Each column represents a follicle with patient noted below. Each row is either the punch, top, middle, or base. Samples that had insufficient DNA for sequencing are shown in gray. (continued on next page)

the skin (Fig. 7D and E; Supplementary Table S8). The type of mutations and the mutational spectrum were similar to those of the epidermis (Supplementary Figs. S2A and S2D; S4C and S4E; S6A and S6B). There was considerable variation in the number of mutations in each segment of the different follicles ranging from 3- to 6-fold in the middle and top areas (Fig. 7D and E).

The maximum VAF across all mutations indicates the clonality of each sample. In all three regions, VAFs above 0.4 (or 0.8 in mutations with LOH) were seen, indicating it is possible for a clone to take over all or most of each region of the follicle (Supplementary Fig. S6C). The distribution of all VAFs was similar in the upper and mid follicles, but VAFs were significantly lower in the follicle base, indicating the bulk of clones were relatively small in this region ($P < 0.0001$, Kruskal-Wallis; Fig. 7F). dN/dS analysis showed *NOTCH1*, *TP53*, and *FAT1* were positively selected in both the upper follicle and the adjacent epidermal punch, whereas *NOTCH1* was selected in the mid follicle (Fig. 7G). There were too few mutations to analyze selection in the base, but no *TP53* and only three *NOTCH1* mutations were detected.

Most mutations were private to each region (Fig. 7H); however, some mutations spanned multiple segments (Fig. 7I, methods). One mutation in the nonselected *ACVR* gene was

present with a VAF of more than 0.4 in the base and mid follicles. Several follicles had one or more mutations spanning the mid and top follicles and the top follicle to the epidermal sample. Seven follicles had clones spreading from mid follicle to the epidermis and two clones spanned from the base to the epidermis. These clones reveal major migrations of cells perhaps in response to local injury.

DISCUSSION

This study reveals normal-appearing human epidermis as a competitive battleground between mutant clones generated by UV light. We find multiple lines of evidence supporting strong competitive selection of mutant genes, both positive and negative. There is strong genetic selection and marked enrichment of protein destabilizing or activating missense mutations of a subset of mutant genes. Using fine-scale spatial mapping, we observe multiple tightly packed clones indicating the skin is a dense patchwork of mutant clones. These findings are consistent with the hypothesis that selected mutant clones expand laterally within the proliferative layer of the epidermis as they outcompete neighboring cells. Mutant clones then collide and compete at clonal

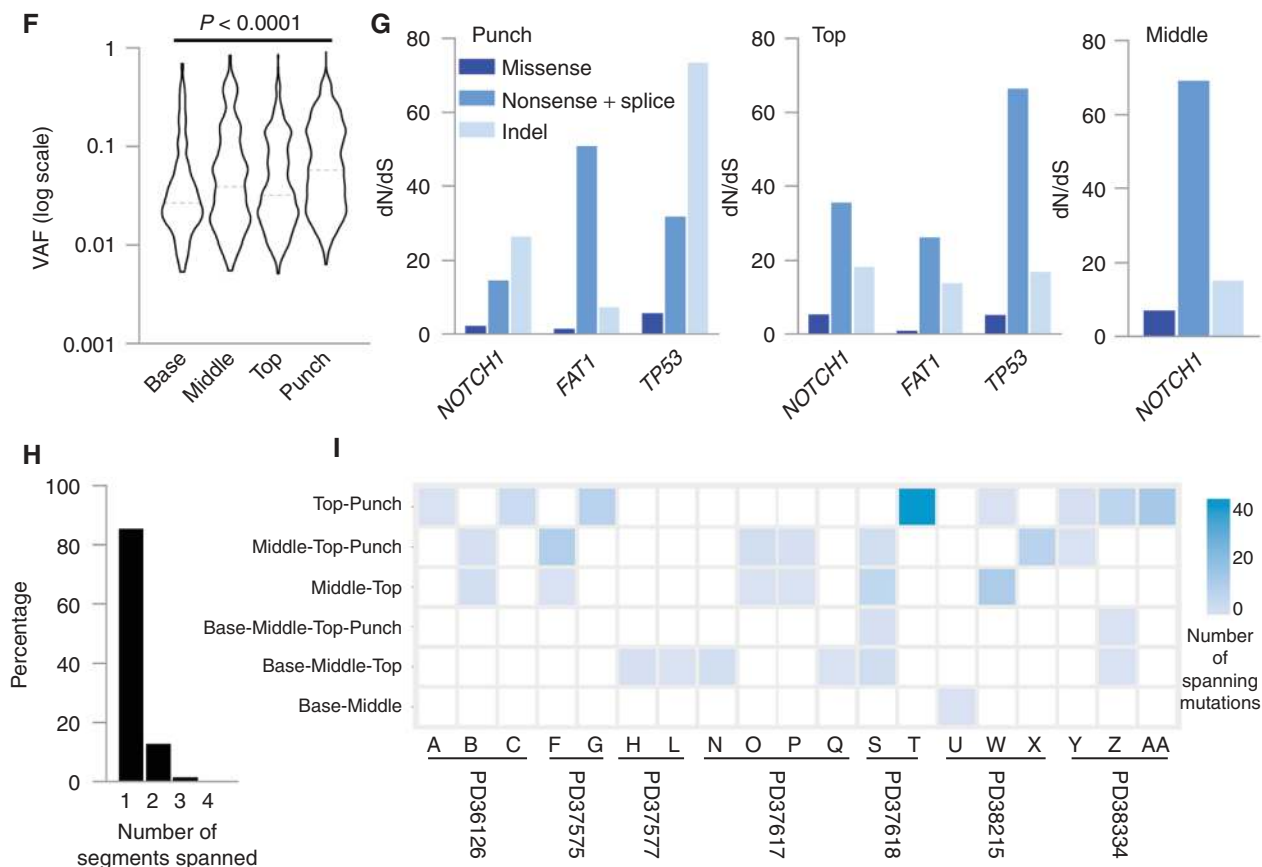


Figure 7. (Continued) **F**, Violin plot of the VAF at each part of the follicle. Dashed line indicates the median. $P < 0.0001$ Kruskal-Wallis test. **G**, dN/dS ratios for missense, nonsense/splice substitutions, and insertions/deletions (indel) across different parts of the follicle. Only genes with global $q < 0.01$ are shown. **H**, Percentage of mutations spanning more than one location ($n = 2009$ mutations). **I**, Heat map of number of mutations spanning different segments of the hair follicle. Each column is a follicle with the patient indicated below. Only follicles with spanning mutations are shown (Methods). Spanning segments are detailed on the left.

boundaries, resulting in the elimination of less fit mutants (7). That this all happens within a normal-appearing tissue argues even the fittest mutant clones are constrained eventually and revert toward homeostatic behavior (2).

Although competitive selection is strong, it is not uniform over the body, illustrated by the different prevalence of *TP53* and *FAT1* mutants in head and leg skin. Long-term low-dose UV exposure promotes the expansion of preexisting *TP53*-mutant clones as well as generating new mutations, but the impact of UV on *FAT1* clonal fitness is unknown (7). We also see variation in the UV signature 7d between frequently sun-exposed and intermittently sun-exposed sites, hinting at differences in UV lesion-repair processes.

Hair follicles have been the subject of intense study in mice, but far less is known about human hair dynamics. The upper follicle resembles the adjacent epidermis in terms of gene selection and the distribution of mutant VAFs. The follicle base is maintained by a small population of cells in the bulge region of the mid follicle, but remains highly polyclonal, arguing multiple clones sustain the hair cycle over several decades. This differs from other tissues with spatially confined niches containing small numbers of stem cells, such as colonic crypts that become monoclonal with age (9).

What insights into carcinogenesis does this study provide? First, the differences we see in mutational landscape across sites are small relative to those in cancer risk. One possible explanation for this apparent discrepancy may be variations in environment between body sites; indeed, this is hinted at by differences seen in mutational signatures and the strength of selection of some mutant genes. If environmental factors that promote transformation can be defined, they may represent opportunities to intervene to reduce cancer risk.

In addition to these broad patterns, WGS revealed differences in mutation burden, indels, double-base substitution, and telomere length in individual clones from the same donor in very close proximity. In addition, differences in mutational signatures point to variation in DNA damage and repair in clones with near-identical UV exposure. What may cause such heterogeneity? It may reflect stochastic fluctuations, “neutral drift,” as it is unclear if the extent to which the burden of mutations, indels, DBS, and telomere length, or indeed the differences in DNA damage/repair alter clonal fitness. Another source of variation is time. WGS reports the genome of the common ancestor of the clone, which reflects the history of that cell up until the time the clone was founded. A clone originating in the donors’ 20s may

differ in parameters such as mutational burden and telomere length from a clone founded decades later. Finally, the fitness of clones and their likelihood of surviving to expand to a detectable size may also depend not merely on the genome/epigenome of the clone itself but on the fitness of its immediate neighbors during clone expansion (15).

The high prevalence of some positively selected mutant genes in normal skin, similar to that in cancers, suggests that *FAT1* and *NOTCH1* may have little role in transformation, whereas *TP53* and *NOTCH2*, which are enriched in cancers, play a part in transformation. One mutant gene that is not selected and striking by its low prevalence is *PTCH1*, mutated in almost all BCC. If the vast majority of mutations that have no effect on competitive fitness are eliminated by neutral drift, as is the case in mouse epidermis, heterozygous *PTCH1*-mutant clones are unlikely to persist for long enough to undergo biallelic inactivation (7). A potential mechanism that would allow heterozygous *PTCH1* mutants to expand and persist follows from the proximity of *PTCH1* and *NOTCH1* on chromosome 9q. If a *PTCH1*-mutant cell acquires a *NOTCH1* mutation, it will have a strong fitness advantage and will expand and persist. Subsequent LOH affecting both genes would result in biallelic *PTCH1* inactivation (Fig. 6E and F). The similarities between the upper hair follicle and the epidermis make it possible that some cancers may arise from follicles, but the sparsity of clonal mutations in the base of larger follicles suggests cancers arising from this area are likely to be rare.

In conclusion, the skin carries a far higher burden of mutant clones than other epithelial tissues, due to the effects of UV exposure that both generates mutations and promotes clonal expansion (7). In the light of this, the high incidence of BCC and SCC compared with other epithelial cancers seems unsurprising, but the resilience of the tissue to mutation is remarkable (5, 9). The selected mutant genes in skin partially overlap with those in esophageal epithelium, another squamous tissue, but there are key differences; for example, *NOTCH1*-mutant clones are unable to colonize the majority of the skin as they do in the esophagus, perhaps because their expansion is restricted by the presence of other strongly competitive mutants. Understanding the environmental differences that promote the transformation of mutants in high- and low-risk sites may give insights that guide skin cancer prevention.

METHODS

Ethics

Written informed consent was obtained in all cases under ethically approved protocols (Research Ethics Committee references 15/EE/0152 NRES Committee East of England—Cambridge South and

15/EE/0218 NRES Committee East of England—Cambridge East). The study was conducted in accordance with the Declaration of Helsinki.

Human Samples

Normal skin samples were collected from patients undergoing wide local excision after initial melanoma excision, patients undergoing browplexy, or deceased organ donors from whom organs were being retrieved for transplantation. For every sample, as much underlying fat and dermis as possible was removed and samples were cut into approximately 0.5 cm × 0.5 cm pieces. Samples were incubated in 20 mmol/L EDTA for two hours at 37°C. The epidermis was then peeled away from the dermis using fine forceps and a dissecting microscope. The epidermis was fixed for 30 minutes with 4% paraformaldehyde (PFA; FD Neurotechnologies) before being washed three times in 1× PBS. For sequencing, the epidermis was cut into 2 × 1 mm grid samples and DNA extracted using the QIAamp micro DNA extraction kit (Qiagen) by digesting overnight and following the manufacturer's instructions. DNA was eluted using prewarmed AE buffer where the first eluent was passed through the column two further times. DNA was extracted from flash-frozen fat and dermis as for the epidermal samples and was used as the germline control. Patient PD38218 contributed three independent trunk samples. For the purpose of mutational burden, signature, and selection, these were treated as a single site.

0.25-mm diameter epidermal samples were collected using a brain punch biopsy (Stoelting Europe).

Hair follicles were dissected away from the peeled epidermis and cut into equal thirds. DNA from both the punches and hair follicles was extracted using the Arcturus Picopure Kit (Applied Biosystems) following the manufacturer's instructions.

Immunofluorescence of Hair Follicles

PFA-fixed wholemounts were blocked for two hours in blocking buffer (0.5% bovine serum albumin, 0.25% fish skin gelatin, 0.5% Triton X-100, and 10% donkey serum) dissolved in PHEM buffer (60 mmol/L PIPES, 25 mmol/L HEPES, 10 mmol/L EGTA, and 4 mmol/L MgSO₄·7H₂O). Follicles were stained with Vimentin (ab92547, Abcam RRID:AB_10562134) diluted 1:500 in blocking buffer for 24 hours at room temperature with continuous rocking. Samples were washed for a minimum of 24 hours with 0.2% Tween-20 in PHEM buffer, changing daytime washes every two to three hours. Follicles were then stained with donkey anti-rabbit Alexa-conjugated 555 (1:500 dilution), wheat germ agglutinin 657 (1:500), and 1 μg/mL DAPI. All secondary antibodies were diluted in blocking buffer and samples were incubated for 24 hours at room temperature with continuous rocking. Samples were washed for a minimum of 24 hours with 0.2% Tween-20 in PHEM buffer, changing daytime washes every two to three hours before imaging on a Leica SP8 confocal microscope.

CRISPR-Cas9 Knockout Screening

The plasmids used are listed below.

Plasmid name	Description	Supplier/Source
pKLV2-EF1a-Cas9Bsd-W	Cas9 expression vector	Addgene #68343 RRID:Addgene_68343
pKLV2-U6gRNA5(Bbs1)-ccdb-PGKpuro2ABFP-w	gRNA vector, modified to contain ccdB gene between Bbs1 sites for gRNA subcloning	Gene Editing, Cellular operations, Wellcome Sanger Institute
psPAX2	VSV-G envelope expressing plasmid	Addgene #12260 RRID:Addgene_12260
pMD2.G	Second-generation lentiviral packaging plasmid	Addgene #12259 RRID:Addgene_12259
pKLV2-U6gRNA5(gGFP)-PGKBFP2AGFP-W	Cas9 activity reporter plasmid	Addgene #67980 RRID:Addgene_67980
pKLV2-U6gRNA5(Empty)-PGKGFP2ABFP-W	Cas9 activity reporter plasmid (control)	Addgene #67983 RRID:Addgene_67983

Cell Culture

HaCaT keratinocytes are a spontaneously immortalized but nontumorigenic cell line obtained from DKFZ (now distributed by CLS). Short-tandem repeat profiling by the Eurofins Cell Line Authentication Service confirmed the cells had the correct marker profile as listed in Cellosaurus (https://web.expasy.org/cellosaurus/CVCL_0038). HaCaT cells were maintained in calcium-free DMEM (Invitrogen) with 10% FBS (calcium-chelated using 5% chelex resin) and supplemented to 0.03 mmol/L CaCl₂ ("low calcium DMEM"). HaCaT cells maintained in low-calcium DMEM were passaged twice weekly to maintain cells in an undifferentiated, proliferating state. Following lentiviral delivery of Cas9 and the gRNA library, cells were switched to DMEM with 10% FBS and supplemented with 1.8 mmol/L CaCl₂, a concentration permissive to keratinocyte differentiation. Cells tested negative for *Mycoplasma* on completion of experiments using a LookOut Mycoplasma Detection Kit (Sigma-Aldrich; cat. # MP0035).

gRNA Library Design and Synthesis

A minipool gRNA library was designed to include gRNAs targeting five genes under apparent negative selection (global $q < 0.01$; see dN/dS analysis). gRNA sequences were selected from the previously described Human CRISPR Library v.1.0 (16). Control gRNAs targeting known essential and nonessential genes (17) and the AAVS1 safe-harbor locus were also included.

An 89-mer oligo pool (oPool) was purchased from IDT in the format ATATATCTTGTGGAAAGGACGAAACACCGN19GTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATA, where N19 specifies the variable gRNA sequence. gRNAs were PCR-amplified from the oPool under the following conditions:

Component	25 μ L reaction	Final concentration
2X Q5 Mastermix (NEB M0492)	12.5 μ L	1 \times
10 μ mol/L forward primer	1.25 μ L	0.5 μ mol/L
10 μ mol/L reverse primer	1.25 μ L	0.5 μ mol/L
Template DNA (1 ng/ μ L)	1.0 μ L	0.04 ng/ μ L
Nuclease-free water	9.0 μ L	Up to 25 μ L

Step	Temperature	Time
Initial denaturation	98°C	30 s
14 Cycles	98°C	10 s
	67°C	10 s
	72°C	15 s
Final extension	72°C	2 min
Hold	4°C	Hold

Forward primer: ATCATATGCTTACCCTGTAACCTGAAAGTATTTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACC

Reverse primer: TGCCACTTTTCAAGTTGATAACGGACTAGCC TTATTTAAACTTGCTATGCTGTTTCCAGCATAGCTCTTAAAC

To avoid the need for gel extraction of digested fragments, we utilized a modified version of pKLV2-U6gRNA5(BbsI)-PGKpuro2ABFP-W in which the ccdB cassette was inserted between the BbsI sites. pKLV2-U6gRNA5(BbsI)-ccdb-PGKpuro2ABFP-W was digested with FastDigest BbsI (Fermentas) for 1 hour at 37°C. The vector digest and PCR products (4 pooled 25- μ L PCR reactions) were purified using the DNA clean and concentrator kit (Zymo Research) according to the manufacturer's instructions.

Gibson assembly reactions were set up using 100 ng each of the PCR-amplified oligo library and the BbsI-digested gRNA vector, and

incubated at 50°C for one hour. Gibson assembly products were purified and concentrated by ethanol precipitation (1 hour at 4°C) and resuspended in RNase-free water to a final concentration of 100 ng/ μ L. 100 μ L of Lucigen Endura electrocompetent cells were combined with 1 μ L of Gibson assembly product and electroporated using the BTX ECM630 (Harvard Apparatus; 2000 V, 200 Ω , 25 μ F). Cells were recovered in 1 mL of Lucigen recovery medium for one hour at 37°C. Bacterial plates were prepared using a serial dilution of recovered cells to estimate the total number of transformants and calculate library coverage. The remaining cells were grown overnight in a 250-mL low-salt 2 \times LB + ampicillin liquid culture at 37°C. Plasmid DNA was extracted using the Endofree Maxi Kit (Qiagen) according to the manufacturer's instructions. A 317 gRNA library was transformed at an estimated coverage of >100,000 \times . Control Gibson assembly reactions without oligo insert yielded a negligible number of transformants, indicating a low background due to undigested vector and efficient suppression of growth by the ccdB cassette.

For quality control of the generated plasmid library, PCR amplification, Illumina sequencing (19-bp single-end sequencing with custom primers on the MiSeq platform), and sgRNA counting were performed as described previously (18). 97.1% (309/317) of gRNAs were recovered, with a skew ratio of 2.2, indicating a favorable gRNA distribution when compared with broadly used genome-wide gRNA libraries (19).

The "ineq" R package was used to generate Lorenz curves and to calculate a Gini index for the generated library and widely used genome-wide libraries, as shown below. The Gini index is a measure of statistical dispersion and reflects the evenness of read count distribution. A lower Gini index is indicative of a more even distribution (Supplementary Fig. S7).

gRNA library	Skew ratio	Gini index
SMS (this study)	2.2	0.19
Avana	4.9	0.36
Brunello	4.0	0.29
Yusa_V1	7.9	0.34
Yusa_V1.1	3.2	0.23

Lentivirus Production and Transduction

Lentiviruses were generated by transfection of HEK293^{FT} cells using Lipofectamine 3000 (Invitrogen). One day before transfection, 4 million HEK293^{FT} cells were plated in a 10-cm dish to reach 80% to 90% confluence after 24 hours. On the day of transfection, 6 μ g of lentiviral vector, 7.5 μ g of the second-generation lentiviral packaging plasmid psPAX2, 2.5 μ g of VSV-G envelope expressing plasmid pMD2.G, and 35 μ L of P3000 Enhancer were added to 1.5 mL of serum-free OptiMEM (Invitrogen). Lipofectamine 3000 (41 μ L) was diluted in 1.5 mL serum-free OptiMEM and incubated at room temperature for five minutes, before combining with the prepared DNA mixture and incubating at room temperature for a further 30 minutes. HEK293^{FT} media were replaced with 6 mL serum-free OptiMEM, before adding the transfection complexes. After six hours, the transfection medium was replaced with complete DMEM. Lentiviral supernatants were collected 24 and 48 hours after transfection. Harvests were pooled together, centrifuged to remove cellular debris, and filtered (0.45- μ m pore, PES; Merck). Virus aliquots were stored at -80°C.

Generation and Validation of Cas9-Expressing HaCaT Keratinocytes

HaCaT cells were transduced with Cas9-Bsd lentivirus at a low MOI and expanded in the presence of blasticidin (10 μ g/mL) for seven days. Cas9 activity was confirmed in >90% of cells using a previously described dual fluorescence reporter vector system². Briefly, HaCaT or HaCaT-Cas9 keratinocytes were transduced with a

lentiviral vector to express GFP and BFP in addition to a gRNA targeting GFP. Cells were also transduced with vector lacking a targeting gRNA in parallel to serve as a control. More than 90% of transduced (BFP⁺) cells were efficiently edited (GFP⁻), specifically in HaCaT-Cas9 cells. No GFP loss was detected in parental HaCaT cells, or in either cell population transduced with the control vector.

Generation of Genome-Wide Mutant Libraries and Screening

1.5×10^6 Cas9-expressing HaCaT keratinocytes were transduced in a total volume of 15 mL low-calcium DMEM containing polybrene (8 μ g/mL) and gRNA lentivirus at a 1/30 dilution. The cell/virus suspension was plated at 20,000 cells/cm² in a T75 cell culture flask. After six hours, cells had adhered, and the transduction medium was replaced with low-calcium DMEM.

Transduction efficiencies were measured 72 hours after transduction by cytometric analysis of the BFP⁺ fraction. Transduction efficiencies varied between 10% and 15% between replicates, corresponding to an initial library coverage of ~500–700 \times . gRNA-expressing cells were enriched to >95% by 72-hour puromycin selection (2 μ g/mL).

Puromycin-selected cells were replated at 10,000 cells/cm² and grown for 72 hours under low-calcium conditions until confluent. Cells were then switched to DMEM + 1.8 mmol/L CaCl₂ and maintained for a further 14 days. Media were refreshed every three days. Cells were trypsinized, and genomic DNA was extracted from ~1.7 million cells (5,000 \times coverage) using a previously described method using salt precipitation and purification by isopropanol precipitation.

PCR amplification, Illumina sequencing (19-bp single-end sequencing with custom primers on the MiSeq platform), and sgRNA counting were performed as described previously (18).

Identification of Positively and Negatively Selected Guide RNAs

Read counts generated from samples collected two weeks after library transduction were compared with those from samples collected immediately after puromycin selection ($t=0$) using the MAGeCK package to identify positively and negatively selected gRNAs and genes (20).

Sequencing

The two sequence capture bait sets used in this study have been described previously (5, 9). The “grid” bait set contains a set of 74 genes recurrently mutated in SCC and BCC as well as genes commonly mutated in other epithelial cancers. The “punches + follicles” bait set is a broader range of genes frequently mutated in a range of cancers based on the COSMIC cancer gene census (<https://cancer.sanger.ac.uk/census>). Samples were sequenced with each bait set as detailed below using fat/dermis from the same patient as a germline control. A list of all genes covered by the bait sets can be found in Supplementary Table S3, and metrics of the bait sets are summarized below.

Bait set name	Size (Mb)	Size (bases)	Number of genes	Usage
Grid	0.39	391,028	74	2-mm ² samples
Punches + follicles	2.00	1,995,559	324	0.25-mm diameter punch samples and hair follicles

Target-enriched samples were multiplexed and sequenced on HiSeq 2000 (Illumina) to generate 75-bp paired-end reads. WGS was performed on either HiSeq X Ten or NovaSeq 6000 machines (Illumina) to generate 150-bp paired-end reads.

BAM files were mapped to the GRCh37d5 reference genome using BWA-mem (version 0.7.17; ref. 21) and targeted sequencing

was aligned using the GATK tool IndelRealigner (version 3.6.0; ref. 22). Duplicate reads were marked using Biobambam2 (Biobambam2 version 2.0.86. <https://gitlab.com/german.tischler/biobambam2>, <https://www.sanger.ac.uk/science/tools/biobambam>). Depth of coverage was calculated using Samtools (version 0.1.18) to exclude reads which were unmapped, not in the primary alignment, failing platform/vendor quality checks, or were PCR/Optical duplicates. BEDTools (version 2.23.0) coverage program was then used to calculate the depth of coverage per base across samples (Supplementary Table S2).

To determine which samples were suitable for WGS, we used the VAF from targeted sequencing to determine which samples were clonal. For the 0.25-mm samples, WGS was performed on all clonal samples in the individuals studied. For the 2-mm² samples, only eight out of 1,261 samples were clonal, and all of these were sequenced.

Variant Calling

For targeted sequencing data, subclonal mutation variant calling was made using the deepSNV R package (also commonly referred to as ShearwaterML), version 1.21.3, available at <https://github.com/gerstung-lab/deepSNV>, used in conjunction with R version 3.3.0 (2016-05-03; ref. 5).

deepSNV makes use of statistical testing to differentiate sequencing errors from true low-frequency mutations and has been shown to be reliable down to a detection limit of 1/10,000 alleles (23). The statistical tests compare by position and strand between skin samples and a panel of control samples to estimate how likely an observed nucleotide is a sequencing error or a true variant. Combining the information for each strand generates a single value used for filtering false-positive variants. It was noted in development that the performance of deepSNV was not strongly dependent upon P , q -values, or PCR amplifications, and its sensitivity can be increased through higher sequencing depths. A q -value of 0.01 was used to filter the variant calls.

Fat and dermis was used as the germline sample for each donor, and sequenced as outlined previously. Aligned germline BAM files for each corresponding sample type (grids, punches, and follicles) were provided to deepSNV, excluding the germline for the sample being analyzed, to form a normal sample panel used for statistical testing and false-positive variant identification. Variants called from a donor’s germline sample were subtracted from the list of variants called from nongermline samples belonging to the donor.

Mutations called by ShearwaterML were filtered using the following criteria:

- Positions of called SNVs must have a coverage of at least 100 reads (10 reads for 0.25 mm diameter punch biopsy samples).
- Germline variants called from the same individual were removed from the list of called variants.
- The P values of the putative mutations were adjusted with FDR and filtered with a q -value threshold of 0.01.
- Mutations not present in at least one read from both strands were removed.
- Pairs of SNVs on adjacent nucleotides within the same sample are merged into a dinucleotide variant if at least 90% of the mapped DNA reads containing at least one of the SNV pair contained both SNVs.
- Identical mutations found in multiple contiguous tissue biopsies are merged and considered as a single clone in order to prevent duplicate clone counting.

For the hair follicle samples, due to the very low input quantity of DNA, particularly toward the middle and base of the follicle, false-positive variant calls attributed to sequencing artifacts are more likely. We therefore applied a strict method of variant calling with ShearwaterML in order to be as conservative as possible. This involved, for all samples within each follicle, calling variants against a

custom normal panel consisting of all other follicle samples from all donors, plus the dermis/fat samples from all other donors. The corresponding dermis/fat sample for the donor of that follicle was then used to remove germline variants. For follicle-spanning mutations, only follicles where adjacent segments were successfully sequenced or where the same mutation is present within the same follicle, i.e., base and top, are shown.

Shearwater was run with a normal panel of >24,000, >31,000, and >12,000× mean coverage depth for the 2-mm² grid samples, hair follicles, and punch biopsy samples, respectively.

For WGS, data variants were called using the CaVEMan and Pindel algorithms (24, 25). For SNVs, CaVEMan was run with the major copy number set to 10 and the minor copy number set to 2. Only SNVs that passed all CaVEMan filters were kept. Additional filtering to remove mapping artifacts associated with BWA-MEM were: the median alignment score of reads supporting a variant had to be at least 140 and the number of clipped reads equal to zero. In addition, the proportion of mutant reads present in the matched sample also had to be zero. Variants with at least one mutant read present in the matched sample were also removed. Two SNVs called at adjacent positions within the same sample were merged to form a doublet-base substitution if at least 90% of the mapped DNA reads containing at least one of the SNV pair contained both SNVs. Small (<200 bp) insertions and deletions were called using Pindel. Only indels that passed all Pindel filters were kept. For the punch samples only, variants were filtered to remove a large excess of single base pair insertions at homopolymers of length five or more, an artifact likely caused by PCR amplification of low-input DNA concentrations during WGS. Indels were then classed as clonal if the VAF was at least 0.3.

Variants were annotated using VAGrENT (26). Full lists of called variants from 2-mm² gridded samples, 0.25 mm diameter punch samples, and hair follicles are shown in Supplementary Tables S4, S7, and S8, respectively.

Clonal Maps

Clonal maps were plotted using exonic mutations called from the targeted sequencing data. Colored squares indicate samples dominated by a clone, after correcting for copy number. Polygons are used where more than two dominant clones occupy a sample; the clone with the larger VAF is shown as the larger polygon. White squares indicate polyclonal samples not dominated by a large clone.

Mutational Signature Analysis

Mutational spectra and signatures are described using the notation used by the PCAWG Mutational Signatures working group (10). The frequency of mutations within each trinucleotide context was calculated using deconstructSigs (<https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-0893-4>). SigProfiler, a method of nonnegative matrix factorization, was used to calculate transcriptional strand bias and was also used to estimate the contribution of 49 single-base-substitution and 11 doublet-base-substitution reference signatures (originally characterized in human cancers) to the combined mutational spectrum of each donor (10).

Copy-Number Calling

For WGS data, the copy number was inferred using ascatNgs using dermis/fat from the same donor as a germline control (27).

SNP Phasing

For targeted sequenced samples, the copy number was estimated on a per-gene basis by phasing heterozygous SNPs, as described previously (1, 5). For each sample, a panel of all other samples from the same donor was used to call SNP loci.

Telomere Length Estimation

Telomere length was estimated using the Telomerecat software package (28). The telomere length given is a median of all chromosomes for all cells in that sample.

Phylogeny Construction of Whole Genome-Sequenced Punches

Substitutions were genotyped prior to phylogeny construction. For each sample in a donor, a pile-up of all quality reads was constructed, and the number of mutant and wild-type reads was counted for every locus that had a mutation called in any sample of that donor. Only reads with a mapping quality of at least 30 and bases with a base quality of at least 25 were included. Mutations with a VAF less than 0.2 were given a genotype of 0 and those with a VAF greater than or equal to 0.3 a genotype of 1. For mutations with a VAF between 0.2 and 0.3, the genotype was set to NA (not applicable) for the purposes of phylogeny construction. Maximum parsimony trees were drawn for each donor after 1,000 iterations with MPBoot (29). Phylogenies for each donor were drawn with single-base and double-base substitutions combined and with SBS only. The addition of double-base substitutions did not alter the structure of the tree, only the branch lengths. Branch lengths of the final phylogenies were adjusted so that each double-base change counted as a single substitution. Substitutions were then reassigned to each branch of a phylogeny. Each branch was annotated with exonic nonsynonymous (missense, nonsense, essential splice, insertion, or deletion) mutations in genes previously identified as being under positive selection by *dNdScv* in the targeted grid data set. To reduce the number of false negatives, each phylogeny was annotated with a driver mutation in these genes if it was present in either the filtered whole-genome CaVEMan and Pindel calls, the targeted ShearwaterML calls and/or the targeted CaVEMan and Pindel calls with a VAF of at least 0.3. Copy-number changes were assigned to each phylogeny manually.

Trinucleotide spectra were drawn for all SBS that had been assigned to each branch of a phylogeny. A visible difference was noticeable in the frequency of T[C>T]A, T[C>T]T, and G[T>C]T peaks in branches of the same donor. There are four single-base-substitution reference signatures associated with damage by UV light (SBS7a–d) that have been characterized in cancers of the skin (10). We used a χ^2 test to quantitatively compare the proportions of mutations in the most prominent peaks of these reference signatures (SBS7a and b: C[C>T]A, C[C>T]C, C[C>T]T, T[C>T]A, T[C>T]C and T[C>T]T; SBS7c: T[T>A]T; SBS7d: G[T>C]T) for each branch. Any branches with fewer than five mutations in any of these categories were removed from the analysis. The differences observed between clones of the same donor are unlikely to be due to sequencing artifact as they are common to branches of samples that share thousands of substitutions attributed to UV damage and are spatially adjacent in the tissue (i.e., are of the same clone). Because each branch is treated as an independent sample in the analysis, we can be confident that the differences observed in mutational spectra are due to biological differences of the cells within each clone.

Calculations of Percentage Mutant Epithelia

The prevalence of BCC and SCC mutations was assessed using variant calls from previously published studies. Variant calls for BCC samples were taken from the study by Bonilla and colleagues (13). One hundred twenty-six whole exome-sequenced samples, Agilent SureSelect Human All Exon V5 capture kit (50 Mb), and 157 targeted samples capturing 387 cancer genes were used for BCCs. Variant calls from a total of 103 whole exome-sequenced SCC samples were collated from five studies (12, 30–33). Two studies used Agilent SureSelect Human All Exon V5 (50 Mb; refs. 12, 31). The remaining studies each used NimbleGen HGSCVCR EXome 2.1 design1 (42 Mb; ref. 30); Illumina Nextera Rapid Capture Exome Library Prep (37 Mb;

ref. 32); and an unspecified version of the Agilent SureSelect Human All Exon Capture Kit (33). Variant calls were filtered to remove low-confidence variants below a coverage depth of 10 \times . Variant annotation was performed using VAGrENT to enable the calculation of the percentage of samples containing nonsynonymous mutations for genes of interest. A weighted mean and the confidence interval (95%) were calculated for each gene.

For the current study, percentage of mutant tissue and dN/dS analysis was performed as described previously (5). Mutation burden was calculated as per section 5.1 of the supplementary methods of Martincorena and colleagues (5). This method accounts for the synonymous footprint (number of possible synonymous mutations) over the baitset.

To test for variation in selection between body sites, dN/dS ratios were compared using likelihood ratio tests, as used previously to compare selection between individuals (5).

Simulations of Mutation Number per Biopsy

Simulations were run using a Wright-Fisher process (34) constrained to a 2-D hexagonal grid with periodic boundary conditions. At the start of a simulation, all cells in the grid are assigned a fitness of 1. In each step of the simulation, a new generation of cells pick their parents from neighboring cells in the previous generation (35). Each cell inherits the fitness of its parent unless a mutation occurs. Mutations are introduced at random to a number of cells in each generation. Most mutations are “neutral” and do not alter the cell fitness inherited from the parent cell, but a small proportion of mutations introduce a fitness advantage drawn from an exponential distribution. The new cell fitness is then the product of the parental cell fitness and the fitness advantage of the mutation. The fitter a cell compared with its neighbors, the more likely it is to be picked as a parent by cells in the next generation, and hence fitter cells produce more offspring (35). In this way, a fit cell is more likely to expand into a larger clone. Aside from the use of a Wright-Fisher style model to simulate one generation of cells in a single step, the simulations follow the same principles as a Moran-style model (single pairs of cell birth/stratification events per step) described previously (2). New generations of cells are assumed to occur every two weeks. Biopsies are then taken from the simulated grids after 3,000 or 4,000 weeks (~58, ~77 years), with a simulated coverage of 690 reads, a minimum of 5 mutant reads required for a clone to be observed, and 3×10^4 cells per 2-mm² biopsy (4). Simulations were run for grids of 40 adjacent biopsies (8 \times 5 grid).

The simulations were run for a wide range of mutation rates and mutation fitness advantages. All combinations of the following parameters were run. The mutation rates used were 0.0000001, 0.000001, 0.00001, 0.0001, 0.001, or 0.01 mutations per cell per generation. The proportion of mutations that were nonneutral (changed the cell fitness) was 0.00001, 0.0001, 0.001, 0.01, or 0.1. The fitness advantages of the nonneutral mutations were drawn from exponential distributions with mean 0.01, 0.05, 0.1, 0.2, 0.5, 1, or 4. If the pre-mutation fitness of the cell was f and the fitness advantage drawn was s , then the post-mutation cell fitness was $f \times (1 + s)$.

Testing Clustering of Mutation Counts

We used the Moran I test of spatial autocorrelation (36) to investigate whether environmental variation across each sampled region might be affecting mutation density. Clones that spread between multiple samples can create local clusters of higher mutation density if they contain many passenger mutations, but these clusters could be unrelated to environmental factors. To limit the effect of these clusters due to clonal spread, all mutations spanning multiple samples were excluded from this test.

We used the python package PySAL (37) to run Moran I. We used the P value under the randomization assumption, and multiple test correction was applied using the Benjamini-Hochberg method (38).

Protein Structure Analysis

$\Delta\Delta G$ Calculation. We used FoldX 5 (39) to calculate $\Delta\Delta G$, a measure of the change in protein folding energy caused by a mutation.

For each pdb file, the FoldX command *RepairPDB* was run to minimize steric clashes and optimize residue orientation. Then the FoldX command *PositionScan* was run for every residue of the proteins of interest in the structure, which mutates each residue to every other amino acid and calculates the $\Delta\Delta G$ value. Default FoldX settings were used for both the *RepairPDB* and *PositionScan* commands.

Calcium-Binding Sites. The calcium-binding residues in EGF11-12 of NOTCH1 were identified defined using MetalPDB (40) and the 2VJ3 (41).

NOTCH1 Ligand-Binding Interface. We used the residues previously identified as on the ligand-binding interface (42).

Distance Measurements in Protein Structures. Distances from residues in p53 to the DNA molecule were measured using the python package MDAnalysis (43) and PDB 2AC0 (44).

Enrichment of Pathogenic Variants

We used the variant summary file of Clinvar annotations, https://ftp.ncbi.nlm.nih.gov/pub/clinvar/tab_delimited/variant_summary.txt.gz, downloaded April 7, 2020) to identify missense variants annotated as pathogenic or likely pathogenic.

The proportions of variants annotated as pathogenic/likely pathogenic in each gene were tested using the method described in “Statistical testing for bias in mutation composition” and a binomial test. Multiple test correction was applied using the Benjamini-Hochberg method (38).

Many pathogenic mutations are likely to be missing from the Clinvar database, so the numbers of annotated pathogenic mutations we find per gene may be an underestimate of the true number of pathogenic mutations.

Statistical Testing for Bias in Mutation Composition

If certain types of mutations convey a fitness advantage to their cell of origin, then they will be more likely to form clones large enough to detect through sequencing. The observed mutations can be compared with the mutations expected under a neutral assumption (that all mutations in a gene have an equivalent impact) to test whether certain mutation features are selected for.

The trinucleotide mutational spectrum was used to estimate how frequently each possible single-nucleotide mutation in a gene would appear in the data if there was no bias due to selection. Each possible mutation was scored for a particular metric and weighted using the spectrum to form the “neutral” null distribution. The observed distribution was then compared with the null to test for a significant bias in the metric scores. For example, if destabilizing mutations provided a fitness advantage, we would expect the observed distribution of a destabilization metric (like FoldX $\Delta\Delta G$) to be skewed toward higher values than would be expected under the neutral null hypothesis.

The test detects differences in relative selection in a gene or section of a gene. Therefore, if there are multiple selected features of mutations in the same region, it becomes harder to detect selection of each individual feature. This can be mitigated by testing for one feature after segmenting the mutations using the other features.

Full details of the method have been described previously (45).

For continuous metrics ($\Delta\Delta G$, distance between the mutated residue and a particular set of atoms in the protein structure), a two-tailed Monte Carlo test was used to compare the null and observed distributions. 100,000 random samples were taken from the null distribution, meaning the minimum P value possible in this test was slightly less than $2e^{-5}$.

For true/false metrics [is the mutation annotated as pathogenic/likely pathogenic in Clinvar (<https://www.ncbi.nlm.nih.gov/clinvar/>); is the mutation on an interface or on a calcium-binding site], a two-tailed binomial test was used. Where 95% confidence intervals are shown for the true/false distributions, the intervals were calculated by taking 10,000 random samples from the null distribution or by bootstrapping 10,000 random samples with replacement from the observed data.

Visualization of Protein Structures

Images of protein structures were generated using VMD (46).

Data and Code Availability Statement

The sequencing data sets generated during study are available at the European Genome-phenome Archive (EGA). 2-mm² grid samples (Accession number EGAS00001001933, EGAS00001002165), 0.25-mm diameter punch and hair follicle samples (Accession number EGAS00001002708), and WGS (Accession number EGAS00001002416).

Code for the simulations of mutation counts per sample:

https://github.com/michaelhall28/Fowler_et_al_scripts

https://github.com/michaelhall28/darwinian_shift

<https://github.com/michaelhall28/clone-competition-simulation>

Authors' Disclosures

K. Fife reports personal fees from IPSEN, Eisai, Novartis, Merck, Eusa, Pfizer, and BMS, grants and personal fees from Roche and MSD, and grants from Exelixis outside the submitted work; and is a Member of Guideline development group for Basal and squamous cell cancer (British Assoc Dermatologists). D. Milne reports personal fees from Pierre Fabre outside the submitted work. A. Roshan reports grants from Cancer Research UK Clinician Scientist Award (C64667/A27958) during the conduct of the study; personal fees from Bepak Europe Ltd outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

J.C. Fowler: Conceptualization, supervision, funding acquisition, investigation, methodology, writing—original draft, project administration, writing—review and editing. **C. King:** Formal analysis, investigation, methodology, writing—original draft, project administration, writing—review and editing. **C. Bryant:** Formal analysis, investigation, methodology, and writing—original draft. **M.W.J. Hall:** Software, formal analysis, investigation, methodology, and writing—original draft. **R. Sood:** Software, formal analysis, and writing—original draft. **S.H. Ong:** Software, formal analysis, and writing—original draft. **E. Earp:** Software, formal analysis, investigation, and writing—original draft. **D. Fernandez-Antoran:** Investigation and writing—original draft. **J. Koepfel:** Investigation and writing—original draft. **S.C. Dentro:** Software, formal analysis, and investigation. **D. Shorthouse:** Software, formal analysis, and investigation. **A. Durrani:** Resources and investigation. **K. Fife:** Conceptualization and resources. **E. Rytna:** Conceptualization and investigation. **D. Milne:** Resources and investigation. **A. Roshan:** Resources. **K. Mahububani:** Resources. **K. Saeb-Parsy:** Resources. **B.A. Hall:** Resources, software, formal analysis, and supervision. **M. Gerstung:** Software, formal analysis, supervision, writing—review and editing. **P.H. Jones:** Conceptualization, supervision, funding acquisition, writing—original draft, writing—review and editing.

Acknowledgments

This work was supported by the Wellcome Trust grants to the Wellcome Sanger Institute (098051 and 296194) and Cancer Research UK Programme Grants to P.H. Jones (C609/A17257 and C609/A27326). B.A. Hall and Michael W.J. Hall are supported by Medical Research

Council grants MC_UU_12022/9 and MR/S000216/1. Michael W.J. Hall was supported by Clare College, Cambridge. B.A. Hall acknowledges support from the Royal Society (grant no. UF130039). S.C. Dentro benefited from an ESPOD fellowship, 2018–21, from the Wellcome Sanger and European Bioinformatics Institutes EMBL-EBI. A. Roshan is supported by a Cancer Research UK Clinician Scientist Fellowship (C64667/A27958). K. Mahububani is supported by the Chan Zuckerberg Initiative. We thank the Cambridge Biorepository for Translational Medicine for access to human tissue. This study was supported by Cancer Research UK, the Wellcome Trust, the Medical Research Council, the Royal Society, the Chan Zuckerberg Initiative, Clare College, Cambridge University, and the European Bioinformatics Institute EMBL-EBI.

Received July 23, 2020; revised September 24, 2020; accepted October 15, 2020; published first October 21, 2020.

REFERENCES

- Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* 2015;348:880–6.
- Hall MWJ, Jones PH, Hall BA. Relating evolutionary selection and mutant clonal dynamics in normal epithelia. *J Roy Soc Interface* 2019;16:20190230.
- Subramaniam P, Olsen CM, Thompson BS, Whiteman DC, Neale RE. Anatomical distributions of basal cell carcinoma and squamous cell carcinoma in a population-based study in Queensland, Australia. *JAMA Dermatol* 2017;153:175–82.
- Bergstresser PR, Pariser RJ, Taylor JR. Counting and sizing of epidermal cells in normal human skin. *J Invest Dermatol* 1978;70:280–4.
- Martincorena I, Fowler JC, Wabik A, Lawson ARJ, Abascal F, Hall MWJ, et al. Somatic mutant clones colonize the human esophagus with age. *Science* 2018;362:911–7.
- Gerstung M, Papaemmanuil E, Campbell PJ. Subclonal variant calling with multiple samples and prior knowledge. *Bioinformatics* 2014;30:1198–204.
- Murai K, Skrupskelyte G, Piedrafitra G, Hall M, Kostiou V, Ong SH, et al. Epidermal tissue adapts to restrain progenitors carrying clonal p53 mutations. *Cell Stem Cell* 2018;23:687–99.
- Madsen RR, Vanhaesebroeck B, Semple RK. Cancer-associated PIK3CA mutations in overgrowth disorders. *Trends Mol Med* 2018;24:856–70.
- Lee-Six H, Olafsson S, Ellis P, Osborne RJ, Sanders MA, Moore L, et al. The landscape of somatic mutation in normal colorectal epithelial cells. *Nature* 2019;574:532–7.
- Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al. The repertoire of mutational signatures in human cancer. *Nature* 2020;578:94–101.
- Vöhringer H, Gerstung M. Learning mutational signatures and their multidimensional genomic properties with TensorSignatures. *bioRxiv* 2019:850453.
- Inman GJ, Wang J, Nagano A, Alexandrov LB, Purdie KJ, Taylor RG, et al. The genomic landscape of cutaneous SCC reveals drivers and a novel azathioprine associated mutational signature. *Nat Commun* 2018;9:3667.
- Bonilla X, Parmentier L, King B, Bezrukov F, Kaya G, Zoete V, et al. Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma. *Nat Genet* 2016;48:398–406.
- Page ME, Lombard P, Ng F, Gottgens B, Jensen KB. The epidermis comprises autonomous compartments maintained by distinct stem cell populations. *Cell Stem Cell* 2013;13:471–82.
- Colom B, Alcolea MP, Piedrafitra G, Hall MWJ, Wabik A, Dentro SC, et al. Spatial competition shapes the dynamic mutational landscape of normal esophageal epithelium. *Nat Genet* 2020;52:604–14.
- Behan FM, Iorio F, Picco G, Gonçalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* 2019;568:511–6.
- Chen C-H, Li W, Xiao T, Xu H, Jiang P, Meyer CA, et al. Integrative analysis and refined design of CRISPR knockout screens. *bioRxiv* 2017. doi: <https://doi.org/10.1101/106534>.

18. Tzelepis K, Koike-Yusa H, De Braekeleer E, Li Y, Metzakopian E, Dovey Oliver M, et al. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep* 2016;17:1193–205.
19. Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD, et al. Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat Protoc* 2017;12:828–63.
20. Li W, Xu H, Xiao T, Cong L, Love MI, Zhang F, et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol* 2014;15:554.
21. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *ArXiv* 2013:1303.3997.
22. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, et al. The genome analysis toolkit: a mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 2010;20:1297–303.
23. Gerstung M, Beisel C, Rechsteiner M, Wild P, Schraml P, Moch H, et al. Reliable detection of subclonal single-nucleotide variants in tumour cell populations. *Nat Commun* 2012;3:811.
24. Jones D, Raine KM, Davies H, Tarpey PS, Butler AP, Teague JW, et al. *cgpCaVEManWrapper*: simple execution of *CaVEMan* in order to detect somatic single nucleotide variants in NGS data. *Curr Protoc Bioinformatics*. 2016;56:15.0.1–0.8.
25. Raine KM, Hinton J, Butler AP, Teague JW, Davies H, Tarpey P, et al. *cgpPindel*: identifying somatically acquired insertion and deletion events from paired end sequencing. *Curr Protoc Bioinformatics* 2015;52:15.7.1–7.2.
26. Menzies A, Teague JW, Butler AP, Davies H, Tarpey P, Nik-Zainal S, et al. *VAGrENT*: variation annotation generator. *Curr Protoc Bioinformatics* 2015;52:15.8.1–8.1.
27. Raine KM, Van Loo P, Wedge DC, Jones D, Menzies A, Butler AP, et al. *ascatNgs*: identifying somatically acquired copy-number alterations from whole-genome sequencing data. *Curr Protoc Bioinformatics* 2016;56:15.9.1–9.7.
28. Farmery JHR, Smith ML, Lynch AG. *Telomerecat*: a ploidy-agnostic method for estimating telomere length from whole genome sequencing data. *Sci Rep* 2018;8:1300.
29. Hoang DT, Vinh LS, Flouri T, Stamatakis A, von Haeseler A, Minh BQ. *MPBoot*: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. *BMC Evol Biol* 2018;18:11.
30. Pickering CR, Zhou JH, Lee JJ, Drummond JA, Peng SA, Saade RE, et al. Mutational landscape of aggressive cutaneous squamous cell carcinoma. *Clin Cancer Res* 2014;20:6582–92.
31. South AP, Purdie KJ, Watt SA, Haldenby S, den Breems N, Dimon M, et al. *NOTCH1* mutations occur early during cutaneous squamous cell carcinogenesis. *J Invest Dermatol* 2014;134:2630–8.
32. Yilmaz AS, Ozer HG, Gillespie JL, Allain DC, Bernhardt MN, Furlan KC, et al. Differential mutation frequencies in metastatic cutaneous squamous cell carcinomas versus primary tumors. *Cancer* 2017;123:1184–93.
33. Durinck S, Ho C, Wang NJ, Liao W, Jakkula LR, Collisson EA, et al. Temporal dissection of tumorigenesis in primary cancers. *Cancer Discov* 2011;1:137–43.
34. Wright S. Evolution in mendelian populations. *Genetics* 1931;16:97–159.
35. Etheridge A. Some mathematical models from population genetics. Switzerland: Springer Science & Business Media; 2011:119.
36. Moran PAP. Notes on continuous stochastic phenomena. *Biometrika* 1950;37:17–23.
37. Rey SJ, Anselin L. *PySAL*: A Python library of spatial analytical methods. In: Fischer MM, Getis A, editors. *Handbook of applied spatial analysis: software tools, methods and applications*. Berlin/Heidelberg: Springer; 2010:175–93.
38. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B* 1995;57:289–300.
39. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L. The *FoldX* web server: an online force field. *Nucleic Acids Res* 2005;33:W382–W8.
40. Putignano V, Rosato A, Banci L, Andreini C. *MetalPDB* in 2018: a database of metal sites in biological macromolecular structures. *Nucleic Acids Res* 2017;46:D459–D64.
41. Cordle J, Johnson S, Tay JZ, Roversi P, Wilkin MB, de Madrid BH, et al. A conserved face of the jagged/serrate *DSL* domain is involved in Notch trans-activation and cis-inhibition. *Nat Struct Mol Biol* 2008;15:849–57.
42. Luca VC, Kim BC, Ge C, Kakuda S, Wu D, Roein-Peikar M, et al. Notch-Jagged complex structure implicates a catch bond in tuning ligand sensitivity. *Science* 2017;355:1320–4.
43. Gowers RJ, Linke M, Barnoud J, Reddy TJE, Melo MN, Seyler SL, et al. *MDAnalysis*: a python package for the rapid analysis of molecular dynamics simulations. In Benthall S, Rostrup S, editors, *Proceedings of the 15th Python in science conference*, Austin, TX; 2016:98–105. *SciPy*, doi:10.25080/majora-629e541a-00e.
44. Kitayner M, Rozenberg H, Kessler N, Rabinovich D, Shaulov L, Haran TE, et al. Structural basis of DNA recognition by p53 tetramers. *Mol Cell* 2006;22:741–53.
45. Hall MWJ, Shorthouse D, Jones PH, Hall BA. Investigating structure function relationships in the *NOTCH* family through large-scale somatic DNA sequencing studies. *bioRxiv* 2020. doi: <https://doi.org/10.1101/2020.03.31.018325>.
46. Humphrey W, Dalke A, Schulten K. *VMD*: visual molecular dynamics. *J Mol Graph* 1996;14:33–8.