Repeated mutation of a developmental enhancer contributed to human thermoregulatory evolution

Daniel Aldea1,2,*, Yuji Atsuta3,1,2,*, Blerina Kokalaris3,*, Stephen F. Schaffner3,*, Rexxi D. Prasasya4,*, Adam Aharoni3,*, Heather L. Dingwall3,*, Bailey Warder3,*, and Yana G. Kamberova5,6

1Department of Genetics, Perelman School of Medicine, Philadelphia, PA 19104; 2Genetics Department, Harvard Medical School, Boston, MA 02115; 3Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, MA 02138; and 4Epigenetics Institute, Department of Cell and Developmental Biology, Perelman School of Medicine, Philadelphia, PA 19104

Edited by Denis Duboule, University of Geneva, Geneva, Switzerland, and approved March 19, 2021 (received for review October 16, 2020)

Humans sweat to cool their bodies and have by far the highest eccrine sweat gland density among primates. Humans’ high eccrine gland density has long been recognized as a hallmark human evolutionary adaptation, but its genetic basis has been unknown. In humans, expression of the Engrailed 1 (EN1) transcription factor correlates with the onset of eccrine gland formation. In mice, regulation of ectodermal EN1 expression is a major determinant of natural variation in eccrine gland density between strains, and increased EN1 expression promotes the specification of more eccrine glands. Here, we show that regulation of EN1 has evolved specifically on the human lineage to promote eccrine gland formation. Using comparative genomics and validation of ectodermal enhancer activity in mice, we identified a human EN1 skin enhancer, hECE18. We showed that multiple epistatically interacting derived substitutions in the human ECE18 enhancer increased its activity compared with nonhuman ape orthologs in cultured keratinocytes. Repression of hECE18 in human cultured keratinocytes specifically attenuated EN1 expression, indicating this element positively regulates EN1 in this context. In a humanized enhancer knock-in mouse, hECE18 increased developmental EN1 expression in the skin to induce the formation of more eccrine glands. Our study uncovers a genetic basis contributing to the evolution of one of the most singular human adaptations and implicates multiple interacting mutations in a single enhancer as a mechanism for human evolutionary change.

Significance

One of the most distinctive physiological traits differentiating humans from other primates is a reliance on sweating to cool off. The effectiveness of human thermoregulatory sweating is underlain by the evolution of a dramatically increased density of water-secreting eccrine sweat glands in human skin relative to that of other primates. Here, we show that the accumulation of human-specific mutations in a developmental enhancer collectively promoted the production of eccrine glands in humans by up-regulating the expression of the Engrailed 1 transcription factor in the skin. This study reveals a mechanism that contributed to the evolution of humans’ signature thermoregulatory capabilities and underscores the importance of regulatory evolution in generating the modern human form.

Author contributions: D.A. and Y.G.K. designed research; D.A., Y.A., B.K., S.F.S., R.D.P., A.A., and B.W. performed research; D.A., S.F.S., H.L.D., and Y.G.K. analyzed data; and D.A. and Y.G.K. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution License 4.0 (CC BY).

1D.A. and Y.A. contributed equally to this work.
2Present address: Department of Biology, Kyushu University, Fukuoka, Japan 819-0395.
3To whom correspondence may be addressed. Email: yana2@pennmedicine.upenn.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2021722118/-/DCSupplemental.

Published April 13, 2021.

https://doi.org/10.1073/pnas.2021722118
specification (13). In light of these findings, we investigated whether evolutionary changes leading to increased ectodermal En1 expression during development could be an underlying mechanism for the adaptive increase in human eccrine gland density.

**Results**

**Identification of Candidate En1 Enhancers in the Skin.** There are no human-specific variants that distinguish the human En1 coding sequence from that of other Catarhines, the group of primates which includes humans, apes, and Old World monkeys (monkeys of Asia and Africa, OWMs), and most evolutionarily relevant changes are likely to reside in regulatory DNA (21). We therefore first investigated whether humans have altered En1 regulation by screening for enhancers with the potential to control this locus during eccrine gland development.

Direct interrogation of human developmental enhancers in this context is confounded by the fact that there are no in vitro systems that recapitulate human eccrine gland development or indeed the development of these appendages from any other species. Moreover, eccrine gland anlagen, or placodes, begin to form during the second trimester in humans (22). Accordingly, we used a candidate-based screen to identify putative enhancers. In general, patterns of ectodermal expression are similar across mammals in body regions where eccrine glands form (13, 18, 23), and therefore we sought to identify putative En1 enhancers based on evidence of evolutionary sequence conservation (Fig. L4). We restricted our screen to noncoding DNA within the genomic interval spanned by the En1 topologically associated domain (TAD) as defined by Hi-C in normal human cultured keratinocytes (Chr2: 118000000–118880000 [hg38]) since studies on the compartmentalization of mammalian genomes suggest that elements controlling the expression of En1 are likely to be located within this interval (24, 25). Importantly, the genomic interval spanned by the human En1 TAD is in a region of conserved synteny to the mapped mouse En1 TAD (Chr1: 120583423–121463423 [mm10]) (24). Using phastCons (26), we called 209 evolutionarily conserved elements within the En1 TAD that were at least 50 base pairs (bps) using an empirically determined threshold of ≥98% probability of being conserved in placental mammals (SI Appendix, Fig. S1A). We prioritized elements based on overlap with published datasets of genomic regions containing epigenomic marks suggestive of enhancer presence from both mouse (27, 28) and human datasets and also with published catalogs of human genomic regions showing evidence of accelerated evolution (29–31). Specifically, we prioritized elements overlapping deoxyribonucleic acid (DNA) I hypersensitivity, En1 Capture-C, and H3K27Ac chromatin immunoprecipitation sequencing (ChIP-seq) peaks from embryonic mouse limbs in which En1 is expressed throughout the ventral limb ectoderm (27, 28). Elements were also prioritized if they overlapped ChIP-seq peaks for DEAF1 in human keratinocytes since we previously reported that this transcription factor positively regulates Engrailed 1 expression in both human and mouse skin cells (20). In addition, we prioritized any elements that overlapped with annotated human accelerated regions (HARs), which are a class of genomic elements that are highly conserved across vertebrates but are exceptionally diverged in humans, suggestive of evolutionary importance in our species (29–31). This prioritization scheme generated a list of 41 priority conserved elements. Using each priority element as a kernel, we expanded these genomic regions to 1,000 to 1,500 bps to accommodate typical enhancer lengths, which tend to be hundreds of base pairs long; combining overlapping regions, this collapsed our list to 23 Engrailed 1 candidate enhancers (ECEs) for functional testing (Fig. 1 A and B and SI Appendix, Fig. S1B).

To determine if the ECEs could function as enhancers in En1-expressing keratinocytes during eccrine gland development, we tested whether each of the 23 ECEs could activate reporter expression in the basal keratinocyte layer of mouse volar skin on postnatal day (P) 2.5, which is the stage when eccrine gland placodes are being specified in this region (18). To this end, we used lentiviruses encoding the mouse ortholog of each ECE upstream of a minimal promoter and an eGFP reporter cassette to stably transduce embryonic mouse ectoderm and generate skin-specific transgenic ECE reporter mice (Fig. L4) (32–34).

Of the 23 tested ECEs, five consistently produced multiple eGFP-positive keratinocyte clones (clusters of keratinocytes that are derived from the same ancestral progenitor) in the basal En1-positive layer of mouse volar skin on P 2.5 (Fig. 1 B and C and SI Appendix, Fig. S1B). At this stage, En1 is critical for specifying the number of eccrine glands in the IFP and is expressed throughout basal keratinocytes of the distal volar skin and focally up-regulated in IFP eccrine placodes (Fig. L4) (18). The five positive ECEs consisted of a fragment of the En1 promoter (ECE2, Chr1: 120601976–120602486 [mm10]) and four elements located downstream of En1 (ECE8, Chr1: 120757766; ECE18, Chr1: 121097676; ECE19, Chr1: 121097762; ECE20, Chr1: 12117648; and ECE23, Chr1: 12139405-121395702 [mm10]) (Fig. 1C and SI Appendix, Figs. S1B and S2A). All positive ECEs also produced positive clones within the basal ectoderm of the volar footpads and differentiating eccrine glands therein, indicating these ECEs can function at later stages of eccrine gland development (SI Appendix, Figs. S2B and S2D). We noticed that the human ECE18 and ECE20 lenti-virus-mediated transgenic mice also had eGFP-positive clones in the dorsal limb skin, which is outside of the En1-positive domain in mice (SI Appendix, Fig. S2C). This observation is intriguing given that En1 expression in the basal ectoderm is expanded to the nonvolar skin in humans.

**Enhanced Activity of the Human ECE18 Ortholog.** Among the five positive ECEs identified in our in vivo screen, we focused our subsequent analyses on ECE18, which overlapped the most epigenomic signatures that are suggestive of enhancer activity from published datasets (SI Appendix, Fig. S1A). Moreover, since our goal was to identify ECEs with the potential to explain a human-specific phenotype, ECE18 additionally stood out as the element that contains the highest number of derived human mutations among the five positive ECEs and overlaps the HACNS56, HAR19, and HAR80 human accelerated regions (collectively named 2xHAR20) (29–31) (Fig. 1D and SI Appendix, Figs. S1C and S3A).

Consistent with having ectodermal enhancer capabilities during development, the human and chimpanzee orthologs of ECE18 (hECE18 and cECE18, respectively) induced eGFP-positive clones in limb skin in a pattern identical to that of mouse ECE18 (Fig. 1C and SI Appendix, Fig. S2 B and C). The mouse lentiviral transgenic system allows for comparisons of the spatial and temporal activities of candidate enhancers in the context of eccrine gland development. This assay is by nature qualitative since we cannot control for differences in transduction efficiency and in lentiviral expression due to integration site effects. Intriguingly, while all tested ECE18 orthologs behaved similarly in the mouse transgenic assay, the comparison of the activity of ECE18 orthologs in quantitative luciferase reporter assays in cultured human keratinocytes revealed dramatic differences in the potency of this enhancer between species (Fig. 24 and SI Appendix, Fig. S3 B and C). When cloned upstream of a minimal promoter driving luciferase expression, we found that anthropoid ECE18 orthologs produced the greatest-fold luciferase induction, while orthologs from species outside this primate infraclass induced little to no luciferase activity (Fig. 24). Among tested orthologs, hECE18 was the most potent enhancer, producing on average a 13-fold increase in luciferase over control. Within Catarrhines, the group consisting of OWMs and apes, hECE18 activity was 2.6-fold and 1.4-fold higher relative to chimpanzee and...
Fig. 1. Identification of Engrailed 1 candidate enhancer ECE18. (A) Strategy to identify putative developmental En1 enhancers. Staining for En1 mRNA (purple) in mouse volar limb skin at P2.5 is shown. The basal keratinocyte layer (arrow) and eccrine placode (double arrow) are shown. (B) The location in mm10 of ECEs tested in vivo (vertical gray lines). ECE18 is in red and boxed and other positive ECEs are highlighted in orange. (C) Representative images of mouse, chimpanzee, and hECE18 transgenic P2.5 volar limb stained with anti-GFP antibody. eGFP (black arrow) is visualized using HRP-DAB coupled immunohistochemistry. (D) Sequence alignment and evolutionarily features of ECE18, which contains four conserved elements by phastCons (blue rectangles) and overlaps the human accelerated regions HACNS56, HAR19, and HAR80 (collectively called 2xHAR20).
Fig. 2. Repeated mutation of ECE18 produced human-specific gains in enhancer activity. (A) Fold induction of luciferase reporter activity relative to empty vector (Control) by ECE18 orthologs in cultured human keratinocytes. (B) Schematic of ECE18 variants tested in C. Derived human bases at A and B are highlighted in red, and the positions of the remaining derived human-specific base substitutions are shown as dashed vertical lines. (C) Localization of hECE18 enhancer activity. The fold change normalized luciferase activities of full-length hECE18, chimpanzee ECE18 (cECE18), human 2xHAR20, hECE18 fragment (Frag) A and B, and mutant hECE18 in which the indicated derived base is replaced by the ancestral ape base are plotted. Mutated base (Mut). All 10 derived human substitutions are mutated to ancestral ape bases in hECE18MutA−J. (D) Species alignment and predicted binding affinities for SP1 at SP1A and SP1B. Derived human bases at A and B are in red. Dots indicate identity to human base. Blue arrows indicate the location of ChIP-qPCR primers used in E. (E) Enrichment of SP1 by ChIP-qPCR at ECE18 interval containing SP1A and SP1B motifs (SP1A−B) in human keratinocytes. Human γ-globin (HGB2) promoter is used as a negative control, and IgG or SP1 enrichment over the input for each set of primers is shown. Mean enrichment is reported across three independent experiments (line). (F) Fold induction of luciferase reporter by hECE18FragA upon deletion of SP1A and SP1B. (G) Fold induction of luciferase reporter by mouse ECE18FragA (mECE18FragA) in human keratinocytes upon knock-in of human SP1A and SP1B motifs. Normalized Firefly luciferase activity is plotted as the fold change relative to Control (empty reporter vector). Firefly luciferase values are normalized to Renilla luminescence. The dots represent an individual biological replicate. Median (line), box (bounds 25 to 75%) and whiskers (min and max) plotted. Significance by one-way ANOVA. Tukey-adjusted $P$ values are reported in heatmaps. Assays are performed in human GMA24F1A keratinocytes.
macaque orthologs, respectively (Fig. 2A). These findings suggest gains in ECE18 potency specific to the anthropoid lineage and are consistent with further human-specific evolutionary changes to the functionality of this enhancer.

Multiple Mutations Underlie Increased ECE18 Activity during Human Evolution. Because ECE18 activity is higher in humans, we reasoned that human-specific sequences might underlie the potency of this enhancer in our species. To investigate the specific changes driving the higher activity of the human enhancer ortholog, we first tested the 2xHAR20 genomic fragment that contains the highest number of derived human substitutions in hECE18 for enhancer activity (Fig. 1 and SI Appendix, Fig. S3A). Human 2xHAR20 did not induce reporter expression in cultured human keratinocytes or in mouse transgenic assays, indicating that this region is not sufficient for enhancer activity (Fig. 2B and C). We then split hECE18 into two partially overlapping fragments: hECE18FragB (hg38 Chr2:118309555–118310154 bp) and hECE18FragA (hg38 Chr2:118309952–118310531 bp) (Fig. 2B and SI Appendix, Fig. S3.A and C). hECE18FragB did not induce reporter expression in vivo or in vitro (Fig. 2C). In contrast, hECE18FragA recapitulated both the quantitative activity and spatial/temporal activity of full-length hECE18 in cultured human keratinocytes and P2.5 mouse transgenic skin, respectively (Fig. 2C and SI Appendix, Fig. S3.C–E). hECE18FragA is highly conserved among modern humans, and we found that only a handful of low frequency variants detectably altered enhancer activity in vitro (SI Appendix, Fig. S3.F and G). ECE18FragA orthologs from nonhuman anthropoids also recapitulated the quantitative activity of the respective full-length elements (SI Appendix, Fig. S3.E). As with full-length hECE18, hECE18FragA produced the same fold change increase in luciferase relative to tested nonhuman primate orthologs, consistent with

hECE18 positively regulates *Engrailed 1* to promote eccrine gland formation. (A) Fold change normalized *EN1* mRNA by qRT-PCR in human GMF24F1A cultured keratinocytes upon dCAS9-KRAB-mediated hECE18 repression. The schematic of strategy and relative positions of *EN1*, *Spat* and *Spat*, and Cr1 and Cr2 gRNA targets is shown. The fold change calculated relative to dCAS9-KRAB transduction alone is shown. (B) Fold luciferase induction relative to empty reporter vector (Control) by ECE18 orthologs in primary mouse keratinocytes. (C) Strategy to generate hECE18 knock-in (hECE18KI) mouse model. (D) Fold change in *En1* mRNA by qRT-PCR in P2.5 volar forelimb skin of wild-type (+/+), hECE18KI heterozygote (hECE18KI/+), and hECE18KI homozygote (hECE18KI/hECE18KI) mice relative to wild type. (E) Normalized ratio of C57BL/6J:FVB/N *En1* allelic expression in volar forelimb skin of wild-type (C57BL/6J/FVB/N) and hECE18KI (C57BL/6J(hECE18KI)/FVB/N) hybrid mice. The ratios were normalized to genomic DNA allelic ratio. (F) Representative stained epidermal preparations of volar forelimb skin from *En1KO/+;+/+* and *En1KO/+;+/hECE18KI* adult mice. The number of eccrine glands in the IFP area (outlined) and excluding the footpads (FP, circled) were quantified in analyses in G. Hair follicle (HF, *), eccrine gland (EG). (G) Quantification of IFP eccrine glands in *En1KO/+;+/+* and *En1KO/+;+/hECE18KI* mice. Each point represents the average number of eccrine glands in the IFP across both forelimbs of a mouse. (A, B, G) significance assessed by one-way ANOVA and Tukey-adjusted P values are reported. (E and G) Significance assessed by a two-tailed t test. **P < 0.0001, ***P < 0.01, *P < 0.05. (+) wild- type allele. (KO) knockout. (KI) knock-in. The dots represent an individual biological replicate.
the idea that this region is sufficient to explain enhancer activity in all species ([Fig. 2C and SI Appendix, Fig. S3E]).

We found that replacing even a portion of the genomic interval spanned by hECE18FragA in hECE18 with the orthogonal chimpanzee sequence was sufficient to reduce the activity of the enhancer in luciferase assays, indicating the importance of human sequence changes ([SI Appendix, Fig. S5H]).

hECE18FragA contains 12 derived human mutations relative to other apes, including two single nucleotide insertions and 10 single nucleotide substitutions (called here sites A to J) ([SI Appendix, Fig. S3J]). Deletion of the two derived insertions or individual mutations of substitutions C to J to the ancestral ape base in hECE18 produced little or no effect on enhancer activity ([SI Appendix, Fig. S3 J and K]). However, mutation of the derived human nucleotides at A and B, alone or in combination, to the ancestral ape bases resulted in substantial and significant reductions of enhancer activity, though not to chimp-like levels ([Fig. 2C]).

Interestingly, mutation of all 10 derived human substitutions (A to J) together to the ancestral ape bases reduced hECE18 activity to that of nonhuman apes ([Fig. 2C]). It is worth noting that five of the derived human-specific substitutions (namely C to G) that contribute to this increased potency are all located within the region overlapping 2xHAR20 ([Fig. 2B]). This indicates that while the accelerated region is not sufficient for enhancer activity, its evolution played a role in driving the full magnitude of increase in hECE18 enhancer function. Moreover, our data suggest that the epistatic effects of multiple human-specific variants located within the interval spanned by hECE18FragA are responsible for the increased potency of this enhancer in humans relative to other apes. From an evolutionary perspective, this means that ECE18 was repeatedly mutated over the course of human evolution, which collectively produced the human enhancer phenotype.

A Conserved Mechanism for ECE18 Evolution. In silico motif analysis revealed that sites A and B lie within tandem predicted binding motifs for the transcription factor SP1, which we named SP1A and SP1B, respectively ([Fig. 2D]). We confirmed SP1 occupancy by ChiP-qPCR in the region spanned by SP1A and SP1B ([Fig. 2E]). Deleting SP1A or SP1B in hECE18FragA reduced luciferase levels to less than half that of the wild-type human enhancer, with loss of both motifs reducing enhancer activity nearly to control levels ([Fig. 2F]).

Overall, in silico analysis predicted anthropoids to have the highest affinity SP1A and SP1B motifs among mammals, and ECE18 anthropoid orthologs did prove to have the highest activity in vitro ([Fig. 2A and D]). The correlation between higher affinity motifs at SP1A and SP1B and increased ECE18 potency was also observed within anthropoids, with humans having the highest affinity sites among apes and the quantitatively strongest enhancer in keratinocytes ([Fig. 2A]). This may also explain our observation that OWM enhancer orthologs were more active in luciferase assays than those of nonhuman apes, since OWMs have identical SP1A motif to humans ([Fig. 2A and D]).

We confirmed SP1 occupancy by ChIP-qPCR in the region spanned by SP1A and SP1B ([Fig. 2E]). Deleting SP1A or SP1B in hECE18FragA reduced luciferase levels to less than half that of the wild-type human enhancer, with loss of both motifs reducing enhancer activity nearly to control levels ([Fig. 2F]).

ECE18 Modulates Endogenous En1 Expression in Human Keratinocytes. Having identified an enhancer, weECE18, in which human-specific sequence evolution increased activity in keratinocytes, we investigated whether hECE18 regulates En1 in this context. We repressed the endogenous hECE18 enhancer in cultured human keratinocytes using a dCas9-Krüppel associated box (KRAB) domain fusion and guide RNAs (gRNAs) targeting the ECE18 genomic interval ([Fig. 3A]). We found that targeting the repressor complex to hECE18 reduced En1 messenger RNA (mRNA) by 40% on average relative to control ([Fig. 3A]). In contrast, repression of hECE18 did not decrease the expression of Insig2 and Ccdc93, the other protein-coding gene loci located within the En1 TAD ([SI Appendix, Fig. S44]). These data demonstrate that hECE18 acts to specifically up-regulate En1 expression in human keratinocytes ([Fig. 3A]).

Human ECE18 Up-Regulates En1 to Promote Eccrine Gland Specification. To determine if hECE18 can function as a developmental enhancer of ectodermal En1, we turned to the mouse, which is an in vivo model system amenable to targeted genetic manipulation. We first determined if the higher potency of hECE18 could be recapitulated in this model. Consistently, hECE18 produced the greatest-fold increase in luciferase in cultured mouse primary keratinocytes ([Fig. 3B]). We noted, however, the relative increase over control was four-fold and 1.5-fold on average by hECE18 and cECE18, respectively, which was markedly lower than that observed in human skin cells for both orthologs. This suggests that the relative activity of ECE18 may be attenuated in this system. As in human keratinocytes, mouse ECE18 did not induce luciferase expression significantly.

Moreover, in contrast to our findings in repressing the endogenous hECE18 enhancer in human keratinocytes, deletion of the homologous ECE18 sequence from the mouse genome did not alter ectodermal En1 levels or eccrine gland number in mouse vular skin ([SI Appendix, Fig. S5]). These data show that though ECE18 is not essential for En1 regulation in mouse skin cells, the human ECE18 ortholog is able to function in this context.

To investigate the functionality of hECE18 during eccrine gland development, we generated a humanized hECE18 knock-in (hECE18KI) by replacing the endogenous ECE18 enhancer of C57BL/6J mice with its human ortholog ([Fig. 3C and SI Appendix, Fig. S6]). hECE18KI mice were born at expected Mendelian ratios and were viable and fertile. qRT-PCR analysis of En1 expression in P2.5 vular skin revealed that ECE18KI heterozygotes and homozygotes each had on average a 1.6-fold increase in En1 relative to wild-type sibling controls ([Fig. 3D]).

Given the evidence that mouse ECE18 is not required for En1 expression in this context, competition with the endogenous mouse En1 enhancers may help to explain why we did not observe a copy-dependent increase in En1 expression between hECE18KI homozygous and heterozygous knock-in mice. The effect of hECE18 on En1 expression was restricted to the vular skin, and we did not observe ectopic expression in the haired dorsal limb ([SI Appendix, Fig. S7A]). Allele-specific analysis of En1 transcription from the vular skin of FVB/N:C57BL/6J and FVB/N:C57BL/6JhECE18KI F1 hybrids revealed that hECE18-mediated up-regulation of En1 occurs in cis, consistent with hECE18 acting as an enhancer of ectodermal En1 in developing mouse vular skin ([Fig. 3E]).

Since hECE18 exerts a positive effect on ectodermal En1 levels, we tested whether this enhancer could promote the specification of mouse eccrine glands. We did not observe a change in eccrine gland number in hECE18KI mice ([SI Appendix, Fig. S7B]). This may reflect the relative attenuation of hECE18 potency in the mouse. Alternatively, since our data indicate that ECE18 is not required for En1 regulation in mouse skin, the mouse may be insensitive to the human enhancer under endogenous conditions. Accordingly, we sensitized the number of vular eccrine glands formed to En1 levels by crossing in a single copy of the En1 knockout allele (En1K0) ([35]) and quantified the number of eccrine glands in the IFP of adult mice ([Fig. 3F]). Consistent with our previous studies ([18, 20], En1K0 heterozygotes had fewer IFP...
eccrine glands than wild-type controls (SI Appendix, Fig. S7C). Strikingly, we found that hECE18 could partially rescue the En1KO eccrine gland phenotype since En1KO/+; hECE18KO/+ mice carrying a copy of the human enhancer had on average 17% more IFP eccrine glands than En1KO alone animals (En1KO++; +/+ ) (Fig. 3G). This result not only shows that the hECE18 enhancer can function in a developmental context to promote the formation of more eccrine glands but also implicates the up-regulation of ectodermal EN1 levels as the underlying mechanism for this effect.

Discussion
Natural selection has shaped many unique human attributes, but the genetic bases of these features are largely unknown. The findings of this study implicate a genetic and developmental mechanism harnessed in the evolution of one of the defining physiological adaptations of humans. That human ECE18 (hECE18) not only positively regulates EN1 expression in cultured human skin cells but also in a developmental model of eccrine gland formation provides compelling evidence for a similar functionality for this enhancer in humans. Combined with the finding that human-specific mutations have increased the relative strength of hECE18, our study supports an evolutionary scenario in which a stronger hECE18 produced up-regulation of ectodermal EN1 during development and a concomitant increase in the number of eccrine glands specified in human skin relative to that of other apes (Fig. 4).

An intriguing feature of hECE18 is that the accumulation of multiple point mutations scattered throughout the enhancer have collectively increased this element’s activity in keratinocytes. Derived human bp substitutions at conserved primate SP1 binding sites SP1A and SP1B appear to have played the greatest individual role in increasing hECE18 activity. Intriguingly, five of the remaining derived substitutions which contribute to increased hECE18 activity lie within the human accelerated region 2xhHar20. While future additional work will be required to tease apart the epistatic interactions between the mutations within and outside the HAR and to unravel the specific steps the HAR played in the evolution of higher eccrine gland density, it is worth noting that our findings directly implicate a bioinformatically defined HAR in the evolution of an adaptive human trait.

That hECE18 gains evolved in a stepwise manner suggests that this element could have contributed to evolutionary increases in the eccrine gland density of Homo sapiens at multiple points during human evolution. As such, our study raises the possibility of sustained selection on the multiple functional derived human mutations which together account for the increased activity of this enhancer relative to other primates. To understand this, we must parse out the order in which hECE18 variants evolved and the threshold at which hECE18 activity gains affect EN1 and eccrine phenotypes. In terms of considering evolutionary mechanisms more broadly, our findings indicate that in addition to the traditional paradigm whereby human adaptive traits evolve as consequences of modifications to multiple components of a developmental program (36), sequential mutations in a single functional element that incrementally alter its activity is also a mechanism to produce gradual change in a trait over time.

A final point to consider is whether the evolution of hECE18 can by itself explain humans’ 10-fold increase in eccrine gland density over that of other apes. hECE18 produced a subtle increase in eccrine gland number in the En1-sensitized mouse knock-in. In light of the finding that ECE18 is not an essential part of the endogenous regulatory machinery controlling En1 levels in the mouse, the mild effects of introducing the human enhancer may be reflective of an attenuated functionality of hECE18 in mouse skin. Determining whether there are associations between present-day hECE18 polymorphisms and variation in human eccrine gland density will help shed light on the relative importance of this enhancer to the specification of eccrine gland density in its endogenous, human context (SI Appendix, Fig. S5G). Future studies to comprehensively map the enhancers that control ectodermal EN1 levels during eccrine gland specification and function and follow-up of the significance of derived human substitutions therein may implicate additional EN1-dependent mechanisms that played a role in the evolution of human eccrine gland density.

Materials and Methods
Identification of En1 Candidate Enhancers for In Vivo Testing.
Conserved elements were called using phastCons (26, 40) from the alignment of placental mammals using the phastCons60way dataset. We restricted our scan to the EN1 TAD defined by Hi-C in normal human cultured keratinocytes (NHEK) (Chr2: 118000000–118880000 [hg38]) since studies on the compartmentalization of the mammalian genome suggest that elements controlling expression of EN1, or indeed any locus, are likely to be located within this smaller interval (24, 25). We called 209 conserved elements based on strings of at least 50 bps of consecutive high phastCons (26) values using an empirically determined cutoff of ≥90% probability of being conserved, excluding coding exons and allowing for short gaps of missing data. We prioritized 41 of these conserved elements based on overlap with the following published epigenomic marks: DNase hypersensitivity peaks from embryonic mouse limbs (27), Capture-C peaks from mouse embryonic limbs.

Fig. 4. A genetic basis for the evolution of increased eccrine gland density in human skin. ECE18 is a developmental enhancer of the anthropoid Engrailed 1 locus. Over the course of human evolution, the ECE18 enhancer accumulated multiple activity-enhancing single nucleotide mutations, particularly at the SP1 sites SP1A and SP1B (boxed and bold), that collectively induced the specification of more eccrine glands in human skin by potentiating developmental EN1 expression. Derived human substitutions relative to other apes underlying higher activity of hECE18 are in red.

Aldea et al.
Repeated mutation of a developmental enhancer contributed to human thermoregulatory evolution.

https://doi.org/10.1073/pnas.2021722118
(28), H3K27Ac ChIP-seq peaks from mouse embryonic limbs (27), H3K27Ac ChIP-seq peaks from NHEK cells (27), coordinates of known HARS (25–31), and DEAF1 ChIP-seq peaks (20). Using each of these 41 priority elements as a kernel, we expanded these genomic regions to 1,000 to 1,500 bps because known enhancers tend to be hundreds of base pairs long, collapsing our list to 23 ECEs for functional testing.

Mice. CD1 (Crl:CD1) timed pregnant mice and FVB/NCrj mice were purchased to 23 ECEs for functional testing. CD1 (Crl:CD1) timed pregnant mice and FVB/NCrj mice were purchased to 23 ECEs for functional testing. CD1 (Crl:CD1) timed pregnant mice and FVB/NCrj mice were purchased to 23 ECEs for functional testing.

Generation of Lentiviral-Mediated Transient Transgenic Mice. The generation of lentivirus-mediated transient transgenic mice was performed as previously described (32, 41). In brief, each ECE or ECE fragment was cloned into the Stagia3 lentiviral enhancer reporter vector, upstream of the minimal TATA box promoter and eGFP reporter cassette (33, 34). Primers used for cloning are reported (SI Appendix, Table S1). A high-titer lentivirus for each construct was prepared in HEK293T cells according to established protocols using a second-generation packaging system (gift from Connie Cepko, Harvard Medical School, Boston, MA). The transfection of HEK293T cells was carried out using polyethylenimine (Polysciences, Inc.). Following concentration, the lentivirus was collected and mixed with a control virus expressing mCherry under a ubiquitous promoter (modified FuDrug in which tdTomato was replaced with an mCherry expression cassette was a gift from Connie Cepko, Harvard Medical School, Boston, MA). The expression of mCherry was used to identify successfully transduced mice at harvest. A virus for each tested element in this study, mixed with FuDrug, was injected into the amniotic cavity of embryonic day (E) 18 CD1/NCrj mouse embryos (Charles River Labs). Injections were carried out under ultrasound guidance using the Vevo 2100 ultrasound imaging system (Visualsonics) equipped with a 35 to 50 MHz mechanical transducer as described previously (32, 41). At E9, lentiviral particles stably integrate into the genomes of cells of the single layer mouse ectoderm but cannot pass beyond the underlying basement membrane, allowing for ectoderm-specific generation of viral transgenic mice (32). All layers of the skin, including the basal keratinocyte layer, derive from the single layer ectoderm at E9. At least six embryos were infected per tested DNA element. All survival surgeries were carried out in accordance with approved IACUC protocols.

On P2.5, pups were euthanized, and both forelimbs and both hindlimbs were processed for expression of eGFP in mCherry-positive animals. At this stage, En1 is expressed throughout basal keratinocytes of the distal volar skin, is focally up-regulated in eccrine glands placodes of the IFP, and is critical for specifying IFP eccrine gland number (Fig. 1A) (18). This spatial pattern of expression is maintained by other transcriptions encoded in the Th0 box, and elements were called based on the observation of multiple eGFP-positive clones in at least three individual transduced mice. Of note, since viral integration into the infected cell genome occurs independently and at random, each clone of positive cells constitutes an individual transgenic event.

In Situ Hybridization, Immunohistochemistry, and Imaging. Whole forelimb and hindlimb autopods (the distal segment of the limb) were embedded in OCT (Tissue Tek) and cryosectioned at a thickness of 10 to 12 μm. En1 in situ hybridization was performed as previously described (18). HRP (horseradish peroxidase)/DAB (3,3-diaminobenzidine) immunostaining was performed as follows. Tissue sections were fixed in 4% paraformaldehyde (PFA) and then washed in phosphate-buffered saline (PBS). Endogenous peroxidase was blocked using 0.3% hydrogen peroxide. The tissue was blocked in PBST (PBS + 0.1% Tween) + 10% normal donkey serum before incubating in chicken anti-GFP primary antibody (1:2,000, Jackson). After washing, samples were incubated with biotin-SP–conjugated rabbit anti-chicken secondary antibody (1:250, Jackson ImmunoResearch). Samples were washed and incubated in Vectastain ABC reagent (Vector Laboratories). Enzymatic detection was carried out using the DAB peroxidase substrate kit (Vector Laboratories) according to manufacturer's instructions. Images were acquired on a Leica DM5500 B microscope equipped with a Leica DFC500 camera.

Bidirectional Luciferase Vector and Luciferase Assays. Nonhuman primate genomic DNA samples were obtained from Christopher Brown (Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA). A domestic female cat (Felis catus), named Phila, served as the source of genomic DNA. All cloned genomic fragments were sequence verified against the current genome assembly. Generation of lentiviral-luciferase reporter was built based on refs. 42 and 43 and by using Stagia3 as a backbone (SI Appendix, Fig. S3B). Briefly, the different ECE18 orthologs and the various fragments described in this study (SI Appendix, Fig. S3C) were cloned upstream of a minimal TATA box promoter and upstream of the Firefly luciferase reporter gene into the bidirectional luciferase lentiviral vector and using the following primers (SI Appendix, Table S2). Lentivirus production and skin cell transduction were carried out as previously described (20, 44). Cells were harvested at 72 h post-transduction, and Firefly and Renilla (for normalization) luciferase activities were measured using the Dual-Luciferase Reporter Assay System (Promega). The assay was performed using the SpectraMax i3x (Molecular Devices). Experiments were done independently at least three times in biological triplicate each time. Mutagenesis of ECE18 orthologs was carried out using standard site-directed mutagenesis protocols, and primers are reported in SI Appendix, Table S3.

In Silico Motif Discovery Analysis. SNP DNA-binding sites were first identified by looking at conserved motifs across the ECE18 sequence and using the Evolutionary Conserved Regions (ECR) browser with default parameters (TRANSFAC professional V10.2 library) (45). The prediction of SNP DNA-binding motifs at the selected sites was confirmed using JASPAR 2018, and affinity scores were obtained from the JASPAR 2018 database (Matrix identification: MA0079.3) (46).

Cell Culture, Transfection, and Transduction. Primary mouse keratinocytes were isolated and cultured as previously described with minor modifications (47, 48). Briefly, CD1 P0 pups were euthanized, and then back skin was removed from the body and incubated over night with 0.25% trypsin-ethylenediaminetetraacetic acid (EDTA, Gibco) to detach the epidermis from the dermis. The next day, the epidermis was collected in a conical tube, and the keratinocytes were dissociated by mechanical shearing by nutation at a slow speed for 45 min at 4 °C. Cells were plated and maintained using Keratinoocyte Serum Free Medium (Gibco) containing 45 μM calcium, 4% of chelse-treated fetal bovine serum (HyClone), 10 ng/mL mouse epidermal growth factor (Coming), 50 μg/mL bovine pituitary extract (Gibco), and streptomycin/penicillin (Sigma).

GMA24F1A human keratinocytes were a gift from Howard Green (Harvard Medical School, Boston, MA) and were previously described (23). GMA24F1A cells are a clonal line of human keratinocytes that express EN1 and have basal keratinocyte characteristics, which we confirmed by the expression of the signature basal keratin Keratin 14 (KRT14) and by the expression of EN1. HEK293T cells were obtained from the laboratory of Connie Cepko (Harvard Medical School, Boston, MA). The transfection of HEK293T cells was carried out using polyethylenimine (Polysciences, Inc.). We used a second-generation packaging system to generate the lentiviruses used in this study (gift from Connie Cepko, Harvard Medical School, Boston, MA). All viruses were produced in HEK293T cells according to established protocols. The transduction of GMA24F1A cells and primary mouse keratinocytes was carried out as previously described (20, 44). Packaging plasmid psPAX2 was a gift from Didier Trono (École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, Addgene plasmid number 12260), and envelope plasmid pCL-VSV was a gift from Connie Cepko (Harvard Medical School, Boston, MA).

Chromatin immunoprecipitation and qPCR. ChIP-qPCR was performed as previously described (20). In brief, GMA24F1A cells were cultured in 150 mm dishes and harvested for immunoprecipitation using the EZ-Magna ChIP HSens Kit (EMD Millipore) according to the manufacturer’s instructions. DNA was cross-linked by adding formaldehyde (1% final concentration) for 10 min at room temperature. After cell lysis, the chromatin was sheared using a Covaris M220 ultrasonicator, and the time was optimized to get chromatin fragments between 100 and 500 bps. Immunoprecipitation was carried out using 1/125 SP1 Antibody (ab13370, Abcam) or 1 μg of normal rat IgG antibody (Sigma-Aldrich). Input and immune-precipitated samples were purified using a Qiagen PCR purification kit. qPCR was done using the Power SYBR PCR master mix (Thermo Fisher). SP1 occupancy of a given region after the immunoprecipitation was determined as the percentage of enrichment over the input using the equation 100 × 2^([CT(input)–CT(IP)]) and is represented in all graphs as enrichment of SP1 or IgG. The primers used for qPCR are reported in SI Appendix, Table S4. The ChIP-qPCR experiment was performed three times, and each reported value represents the average across three technical qPCR replicates for that biological sample in a single experiment.

Generation and Maintenance of Mouse ECE18 Knockout (mECE18del) and NCE18® Mice. For specific details of targeting design, strategy and validation, screening of founder animals, and establishment of mutant lines see
The membrane was hybridized with a DNA probe labeled using a Nick Translation Kit (Roche) and [γ-32P] deoxyadenosine-5'-triphosphate (PerkinElmer). The membrane was washed with 0.1× saline sodium citrate (SSC) buffer (1× SSC is 150 mM NaCl, 15 mM sodium citrate) and 0.1% SDS at 55 °C. Membrane was exposed to a phosphor screen and visualized using the Amersham Typhoon Biomolecular Imager (GE).

**Southern Blot Analysis.** Genomic DNA was isolated from mouse tail biopsies using tail lysis buffer (10 mM Tris pH: 8.0, 100 mM NaCl, 10 mM EDTA pH: 8.0, and 0.5% sodium dodecyl sulfate [SDS])/Proteinase K and extracted by a phenol:chloroform:isoamyl alcohol method. In brief, 10 μg of genomic DNA was digested with KpnI-HF and NotI-HF (New England Biolabs) overnight at 37 °C, followed by heat inactivation for 20 min at 65 °C. Digested genomic DNA was run on 1% agarose gels and transferred to a 0.45 μm nitrocellulose filter (Millipore). The membrane was hybridized with a 32P-labeled specific DNA probe (SI Appendix, Fig. S56). The probe was labeled using a Nick Translation Kit (Roche) and [γ-32P] deoxyadenosine-5'-triphosphate (PerkinElmer). The membrane was washed with 0.1× saline sodium citrate (SSC) buffer (1× SSC is 150 mM NaCl, 15 mM sodium citrate) and 0.1% SDS at 55 °C. Membrane was exposed to a phosphor screen and was imaged on the Amersham Typhoon Biomolecular Imager (GE).

**Long-Range PCR and DNA Digestion.** Genomic DNA fragments of interest were PCR-amplified using Q5 polymerase (New England Biolabs) (for primers, see SI Appendix, Fig. S56) and purified using gel extraction (Qiagen). PCR products were digested using NotI (New England Biolabs) (SI Appendix, Fig. S57).
16. M. Plikus et al., Morpho-regulation of ectodermal organs: Integument pathology and phenotypic variations in K14-Noggin engineered mice through modulation of bone morphogenetic protein pathway. Am. J. Pathol. 164, 1099–1114 (2004).

17. J. A. Mayer, J. Foley, D. De La Cruz, C.-M. Chuong, R. Widelitz, Conversion of the nipple to hair-bearing epithelia by lowering bone morphogenetic protein pathway activity at the dermal-epidermal interface. Am. J. Pathol. 173, 1339–1348 (2008).

18. Y. G. Kamberov et al., A genetic basis of variation in eccrine sweat gland and hair follicle density. Proc. Natl. Acad. Sci. U.S.A. 112, 9932–9937 (2015).

19. C. A. Loomis et al., The mouse Engrailed-1 gene and ventral limb patterning. Nature 382, 360–363 (1996).

20. D. Aldea et al., The transcription factor Deaf1 modulates enogrd1-1 expression to regulate skin appendage fate. J. Invest. Dermatol. 139, 2378–2381.e4 (2019).

21. M. C. King, A. C. Wilson, Evolution at two levels in humans and chimpanzees. Science 188, 107–116 (1975).

22. W. Montagna, P. F. Parakhel, The Structure and Function of Skin (Elsevier Inc., ed. 3, 1974).

23. G. Mainguy et al., Regulation of epidermal bullous pemphigoid antigen 1 (BPAG1) synthesis by homeoprotein transcription factors. J. Invest. Dermatol. 113, 643–650 (1999).

24. J. R. Dixon et al., Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature 485, 376–380 (2012).

25. J. R. Dixon et al., Chromatin architecture reorganization during stem cell differentia

26. A. Siepel et al., Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034–1050 (2005).

27. C. A. Davis et al., The encyclopedia of DNA elements (ENCODE): Data portal update. Nucleic Acids Res. 46, D794–D801 (2018).

28. G. Andrey et al., Characterization of hundreds of regulatory landscapes in developing limbs reveals two regimes of chromatin folding. Genome Res. 27, 223–233 (2017).

29. K. S. Pollard et al., Forces shaping the fastest evolving regions in the human genome. PLoS Genet. 2, e168 (2006).

30. J. A. Capra, G. D. Erwin, G. McKinsey, J. L. R. Rubenstein, K. S. Pollard, Many human accelerated regions are developmental enhancers. Philos. Trans. R. Soc. Lond. B Biol. Sci. 368, 20130025 (2013).

31. S. Prabhakar, J. F. Noonan, S. Paião, E. M. Rubin, Accelerated evolution of conserved noncoding sequences in humans. Science 314, 786 (2006).

32. B. Beronja, G. Livshits, S. Williams, E. Fuchs, Rapid functional dissection of genetic networks via tissue-specific transduction and RNAi in mouse embryos. Nat. Med. 16, 827 (2010).

33. M. M. Emerson, C. L. Cepko, Identification of a retina-specific Otx2 enhancer element active in immature developing photoreceptors. Dev. Biol. 360, 241–255 (2011).

34. S. Wang, C. Sengel, M. M. Emerson, C. L. Cepko, A gene regulatory network controls the binary fate decision of rod and bipolar cells in the vertebrate retina. Dev. Cell 30, 513–527 (2014).

35. R. A. Kimmel et al., Two lineage boundaries coordinate vertebrate apical ectodermal ridge formation. Genes Dev. 14, 1377–1389 (2000).

36. S. B. Carroll, Genetics and the making of Homo sapiens. Nature 422, 849–857 (2003).

37. Y. G. Kamberov et al., Modeling recent human evolution in mice by expression of a selected EDAR variant. Cell 152, 691–702 (2013).

38. R. A. Barton, C. Venditti, Rapid evolution of the cerebellum in humans and other great apes. Curr. Biol. 24, 2440–2444 (2014).

39. W. Wurst, A. B. Auerbach, A. L. Joyner, Multiple developmental defects in enogrd1-1 mutant mice: An early mid-hindbrain deletion and patterning defects in forelimbs and sternum. Development 120, 2065–2075 (1994).

40. K. S. Pollard, M. J. Hubisz, K. R. Rosenberg, A. Siepel, Detection of nonneutral substitution rates on mammalian phylogenies. Genome Res. 20, 110–121 (2010).

41. S. Beronja, E. Fuchs, RNAi-mediated gene function analysis in skin. Methods Mol. Biol. 961, 351–361 (2013).

42. I.-K. Na et al., Concurrent visualization of trafficking, expansion, and activation of T lymphocytes and T-cell precursors in vivo. Blood 116, e18–e25 (2010).

43. J. W. Kleinovink et al., A dual-color bioluminescence reporter mouse for simultaneous in vivo imaging of T cell localization and function. Front. Immunol. 9, 2097 (2019).

44. A. S. McNeal et al., CDKN2B loss promotes progression from benign melanocytic nevus to melanoma. Cancer Discov. 5, 1072–1085 (2015).

45. I. Ovcharenko, M. A. Nobrega, G. G. Loots, L. Stubs, ECR browser: A tool for visualizing and accessing data from comparisons of multiple vertebrate genomes. Nucleic Acids Res. 32, W280–W286 (2004).

46. A. Khan et al., JASPAR 2018: Update of the open-access database of transcription factor binding profiles and its web framework. Nucleic Acids Res. 46, D260–D266 (2018).

47. A. Pirrone, B. Hager, P. Fleckman, Primary mouse keratinocyte culture. Methods Mol. Biol. 289, 3–14 (2005).

48. U. Lichti, J. Anders, S. H. Yuspa, Isolation and short-term culture of primary keratinocytes, hair follicle populations and dermal cells from newborn mice and keratino
cytes from adult mice for in vitro analysis and for grafting to immunodeficient mice. Nat. Protoc. 3, 795–810 (2008).

49. P. I. Thakore et al., Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat. Methods 12, 1143–1149 (2015).

50. I. M. Carr et al., Inferring relative proportions of DNA variants from sequencing electropherograms. Bioinformatics 25, 3244–3250 (2009).

51. Jairo Navarro Gonzalez et al., The UCSC Genome Browser database: 2021 update. Nucleic Acids Res. 49, D1046–D1057, 10.1093/nar/gkaa1070 (2021).