Mutations in ACTRT1 and its enhancer RNA elements lead to aberrant activation of Hedgehog signaling in inherited and sporadic basal cell carcinomas

Elodie Bal, Hyun-Sook Park, Zakia Belaid-Choucair, Hulya Kayserili, Magali Naville, Marine Madrange, Elena Chiticariu, Smail Hadj-Rabia, Nicolas Cagnard, Francois Kuonen, et al.

nature medicine

Mutations in *ACTRT1* and its enhancer RNA elements lead to aberrant activation of Hedgehog signaling in inherited and sporadic basal cell carcinomas

Elodie Bal[1,2], Hyun-Sook Park[3], Zakia Belaid-Choucair[1,2,4,29], Hülya Kayserili[5,6,29], Magali Naville[7–9], Marine Madrange[1,2], Elena Chiticariu[3], Smail Hadj-Rabia[1,2,10], Nicolas Cagnard[11], Francois Kuonen[3], Daniel Bachmann[3], Marcel Huber[3], Cindy Le Gall[1,2], Francine Côté[1,2,4], Sylvain Hanein[1,2], Rasim Özgür Rosti[6,12], Ayca Dilruba Aslanger[6], Quinten Waisfisz[13], Christine Bodemer[1,2,10], Olivier Hermine[1,2,4,14,15], Fanny Morice-Picard[16], Bruno Labeille[17], Frédéric Caux[18], Juliette Mazereeuw-Hautier[19], Nicole Philip[20,21], Nicolas Levy[20,21], Alain Taieb[16,22], Marie-Françoise Avril[23], Denis J Headon[24], Gabor Gyapay[25], Thierry Magnaldo[26], Sylvie Fraitag[27], Hugues Roest Crollius[7–9], Pierre Vabres[28], Daniel Hohl[3], Arnold Munnich[1,2] & Asma Smahi[1,2]

Basal cell carcinoma (BCC), the most common human cancer, results from aberrant activation of the Hedgehog signaling pathway[1]. Although most cases of BCC are sporadic, some forms are inherited, such as Bazex–Dupré–Christol syndrome (BDCS)—a cancer-prone genodermatosis with an X-linked, dominant inheritance pattern[2]. We have identified mutations in the *ACTRT1* gene, which encodes actin-related protein T1 (ARP-T1), in two of the six families with BDCS that were examined in this study. High-throughput sequencing in the four remaining families identified germline mutations in noncoding sequences surrounding *ACTRT1*. These mutations were located in transcribed sequences encoding enhancer RNAs (eRNAs)[3–5] and were shown to impair enhancer activity and *ACTRT1* expression. ARP-T1 was found to directly bind to the *GLI1* promoter, thus inhibiting *GLI1* expression, and loss of ARP-T1 led to activation of the Hedgehog pathway in individuals with BDCS. Moreover, exogenous expression of *ACTRT1* reduced the *in vitro* and *in vivo* proliferation rates of cell lines with aberrant activation of the Hedgehog signaling pathway. In summary, our study identifies a disease mechanism in BCC involving mutations in regulatory noncoding elements and uncovers the tumor-suppressor properties of *ACTRT1*.

BDCS (MIM 301845) is an X-linked, dominantly inherited condition predisposing to BCC[2,6,7] (**Supplementary Fig. 1a**). By studying six families affected by BDCS (**Fig. 1a**), we mapped the gene associated with BDCS to a 7.5-Mb region at Xq25-q26.2 (**Supplementary Fig. 1b**,**c**) and identified an insertion in the *ACTRT1* gene, encoding ARP-T1, that segregated with the disease in two of the six families (c.547_548insA, p.Met183Asnfs*17 in families C and D; **Supplementary Fig. 1d**,**e**). The mutant cDNA encodes a 25-kDa truncated protein (**Supplementary Fig. 1f**). Yet, no mutations in the coding region of *ACTRT1* were found in the other four families in which disease was linked to the same interval. No rearrangements in the candidate region were identified by high-density tiling-path array-based comparative genomic hybridization.

Immunohistochemical, RT–PCR and western blot analyses of control skin samples taken from healthy individuals detected expression of ARP-T1 in epidermal layers and skin appendages involved in BDCS but not in dermal connective tissues (**Supplementary Fig. 2a–d**). Interestingly, immunohistochemical analyses failed to detect any specific ARP-T1 staining in BCC tumors from individuals with BDCS and detected only a weak signal in unaffected epidermis from these individuals (**Supplementary Fig. 2e**). Immunofluorescence (**Fig. 1b**,**c**) and qRT–PCR (**Fig. 1d**) analyses of the epidermis showed low ARP-T1 expression in all individuals with BDCS, regardless of the presence or absence of mutations in *ACTRT1*, suggesting that the remaining unexplained cases involved hitherto unknown *ACTRT1* regulatory elements.

Conserved noncoding elements (CNEs), which are known to control expression of neighboring genes[8,9], are concentrated in gene deserts[10]. Taking into consideration the fact that *ACTRT1* is located in a 2.6-Mb gene desert, we examined 17 CNEs located either upstream or downstream of the *ACTRT1* coding sequence (**Supplementary Fig. 3a**). Sanger sequencing detected a g.127372937A>T variation of CNE12 only in families E and F (CNE12, chr. X: 127,371,674–127,374,249; **Supplementary Fig. 3b**,**c**). Comparative genomic approaches aimed at predicting regulatory DNA sequences are known to have limitations, as regulatory elements are not necessarily conserved across species[11,12]. We therefore performed systematic array-based capture and high-throughput sequencing of the 7.5-Mb candidate region and used a specific genome browser to select candidate BDCS-associated variants (**Supplementary Fig. 4a**). We selected three additional candidate variants in family A (A1, g.125960325A>T;

A full list of affiliations appears at the end of the paper.

Received 5 May; accepted 14 June; published online 4 September 2017; doi:10.1038/nm.4368

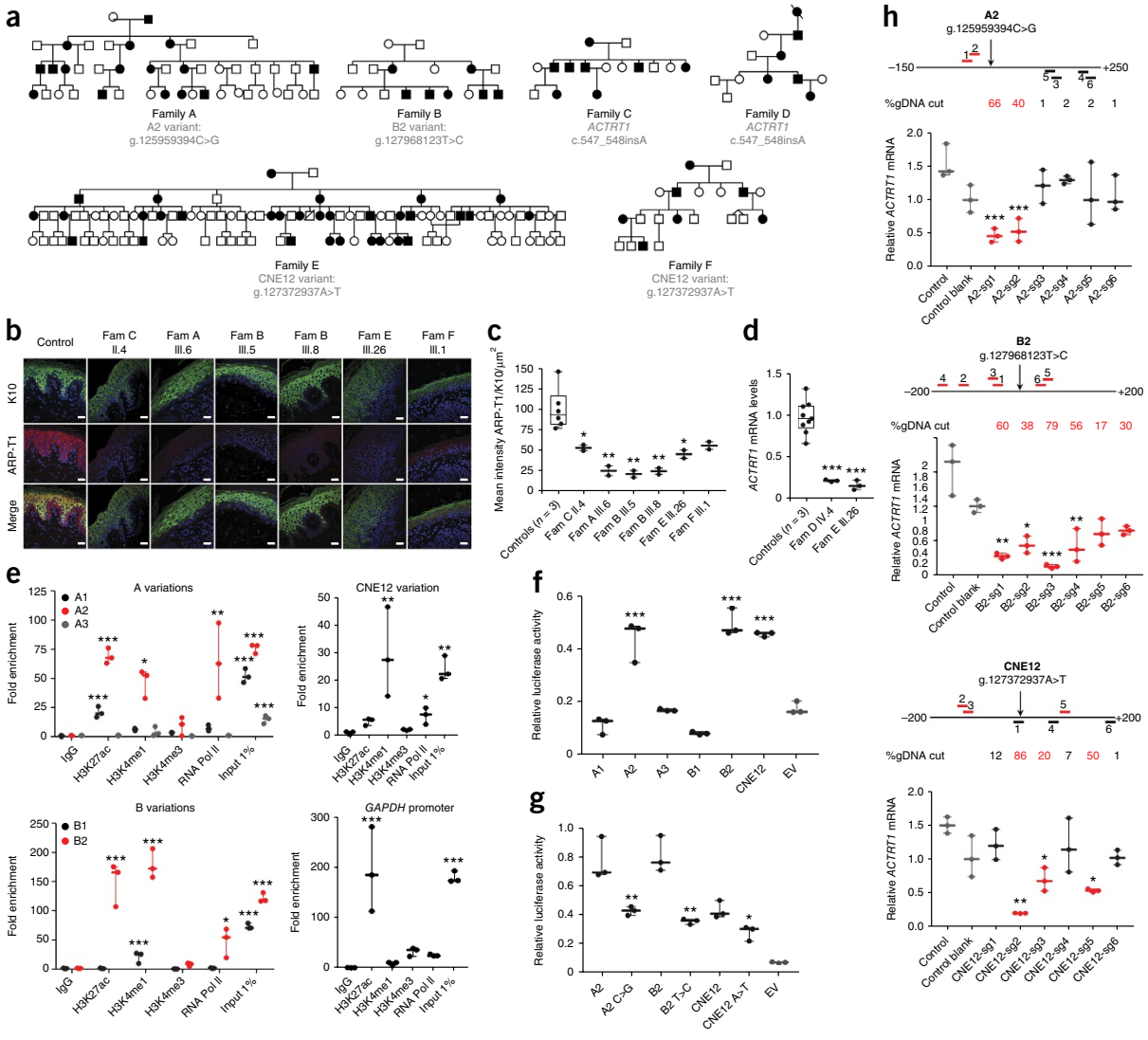

Figure 1 Genetic analysis of BDCS families. (**a**) Pedigrees of six BDCS families (including three families reported previously, A–C[2]). Mutations associated with BDCS are indicated under each pedigree. Circles, females; squares, males; filled shapes, individuals with BDCS; open shapes, unaffected individuals. (**b**) Immunofluorescent double staining of skin biopsies taken from a healthy individual (control) and six individuals with BDCS with anti-cytokeratin 10 (K10; Alexa Fluor 488; green) and anti-ARP-T1 (Alexa Fluor 546; red) antibodies. Cell nuclei were stained with DAPI (blue). Scale bar, 10 μm. Fam, family. (**c**) Quantification of the relative expression of ARP-T1 (normalized to that of cytokeratin 10) using immunofluorescence and confocal microscopy. Quantification was performed in duplicate. (**d**) qRT–PCR analysis of *ACTRT1* mRNA expression in skin biopsies from individuals with BDCS ($n = 2$) and controls ($n = 3$). Results were normalized to *PGK1* mRNA levels. Experiments were performed in triplicate. Results are representative of two independent experiments. (**e**) Levels of indicated histone marks in the sequences surrounding the candidate variants in families A and B and the CNE12 variant in control epidermis. The experiments were performed in triplicate. Results obtained from one control skin biopsy are shown but are representative of two independent experiments on two different control samples. Input 1%, sonicated but nonimmunoprecipitated DNA. The chromatin signature of the *GAPDH* promoter was assessed as a control. Note that the sequence of the *GAPDH* promoter was enriched in H3K4me3, a promoter-specific histone mark. The CNE12, A2 and B2 sequences were not tagged with H3K4me3, but they were enriched with the enhancer-specific markers H3K27ac and H3K4me1. Wild-type A1, A3 and B1 were not enriched in any specific histone marks. (**f**) An enhancer luciferase reporter assay based on the wild-type sequences surrounding candidate variants in HaCaT keratinocytes. Enhancer activity was observed with the A2, B2 and CNE12 sequences. No significant luciferase activity was observed with wild-type A1, A3 and B1 sequences in comparison to an empty-vector (EV) control. (**g**) Enhancer luciferase reporter assays demonstrating the impact of the variants in the A2 (g.125959394C>G), B2 (g.127968123T>C) and CNE12 (g.127372937A>T) sequences in HaCaT keratinocytes. In **f** and **g**, luciferase activity was normalized to that of *Renilla* luciferase; the experiments were performed in triplicate, and results in each panel are representative of three independent experiments. (**h**) Schematic representation of the target sites of sgRNAs around the A2, B2 and CNE12 sequences. sgRNAs that created indels in enhancer regions are shown in red, and sgRNAs that failed to delete enhancer regions are shown in black. The editing efficiency in genomic DNA (gDNA cut) was measured using tracking of indels by decomposition (TIDE) software (**Supplementary Figs. 6–8**). The expression of *ACTRT1* mRNA was measured in RT–PCR assays, and results were normalized to *PGK1* mRNA levels. The experiment was performed in triplicate, and results are representative of two independent experiments. Data in **c**–**h** are expressed as the mean ± s.d. For box-and-whisker plots, midlines represent the median; upper and lower perimeters extend from the 25th to the 75th percentile; and whiskers extend from minimum to maximum. *P* values were calculated by one-way ANOVA with Bonferroni multiple comparisons for **c** and **d** (relative to controls), **e** (relative to IgG), **f** (relative to EV) and **h** (relative to control blank) and by *t*-test for **g** (wild-type sequence versus mutant sequence). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

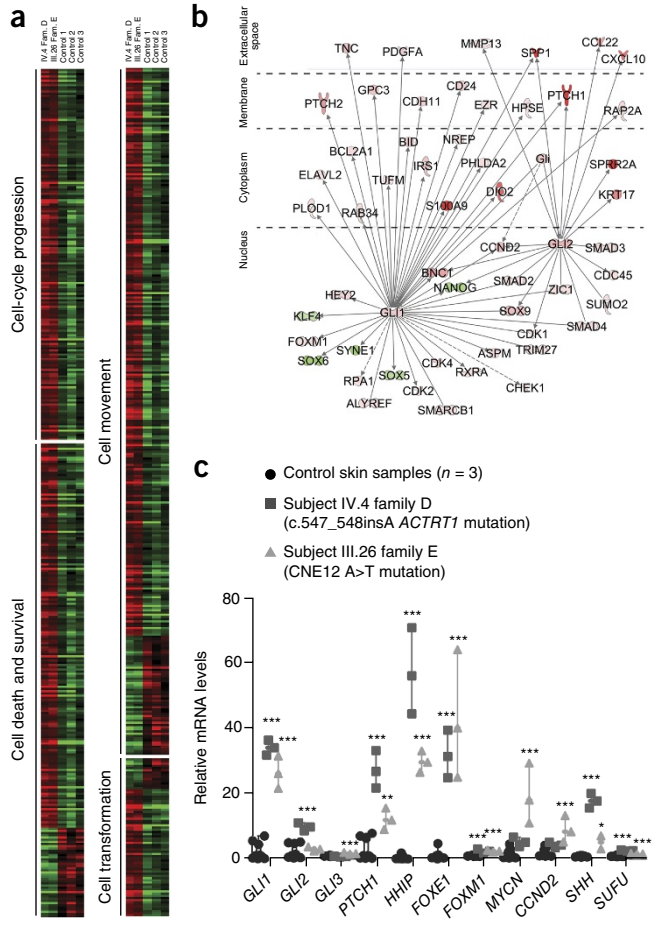

Figure 2 Transcriptomic analyses on skin biopsies from individuals with BDCS. (**a**) Heat maps showing genes deregulated in skin biopsies from individuals with BDCS in families D and E in comparison to control skin biopsies as determined by transcriptomic analyses. The differentially expressed genes are mostly involved in cell proliferation, cell death or survival, and cell migration (Online Methods). (**b**) Ingenuity Pathway Analysis of genes found to be deregulated in the skin biopsies of individuals with BDCS as compared to controls. 56 genes and their encoded proteins are directly linked to the Hedgehog pathway transcription factors GLI1 and GLI2. In **a** and **b**, the intensity of green or red indicates the degree of down- and upregulation, respectively. (**c**) qRT–PCR analysis of mRNA expression by selected targets of the Sonic Hedgehog pathway in skin from individuals with BDCS ($n = 2$) and controls ($n = 3$). Results were normalized to mRNA expression of *PGK1*. qRT–PCR was performed in triplicate, and results are representative of three independent experiments. Data are expressed as the mean ± s.d. *P* values were calculated by one-way ANOVA with Bonferroni multiple comparisons (relative to control skin samples). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

A2, g.125959394C>G; A3, g.126494053_126494054insT) and two variants in family B (B1, g.127061005G>C; B2, g.127968123T>C; **Supplementary Fig. 4b**,**c**). To identify putative disease-causing variants from among these candidates, we performed (i) a chromatin signature analysis to identify variants that mapped to active regulatory regions[13,14], (ii) enhancer luciferase reporter assays to identify sequences capable of activating transcription[15] and (iii) *in situ* hybridization to identify regions that are transcribed in skin[16].

Chromatin immunoprecipitation (ChIP) and targeted qPCR assays were performed on protein–DNA complexes extracted from normal human epidermis. Interestingly, an enhancer signature, as determined by enrichment of H3K27ac and H3K4me1 marks and absence of the H3K4me3 mark, was only found for the A2, B2 and CNE12 sequences (**Fig. 1e**). In these assays, a difference in amplification efficiency between primer sets constitutes a limitation for comparison of the data obtained using different primers. Luciferase reporter assays demonstrated that only the wild-type sequences encompassing the A2, B2 and CNE12 variant positions had enhancer activity in HaCaT keratinocytes (**Fig. 1f**). This activity was reduced when mutated constructs were used (**Fig. 1g**). Given that a subset of eRNAs are broadly transcribed[3,4] and that the A2, B2 and CNE12 sequences bound RNA polymerase II (RNA Pol II)[5] (**Fig. 1e**), we used RT–PCR and *in situ* hybridization to study transcription at the candidate loci. In agreement with the *ACTRT1* expression pattern, the A2, B2 and CNE12 sequences were specifically expressed in the epidermis and its appendages (**Supplementary Fig. 5a**–**d**). Interestingly, no staining for these sequences was detected when skin biopsies from the corresponding individuals were tested, demonstrating the dramatic impact of the A2, B2 and CNE12 variations on eRNA expression and/or stability (**Supplementary Fig. 5b**).

Lastly, in order to provide conclusive *in vitro* evidence of a link between these enhancer sequences and *ACTRT1* expression, we used CRISPR–Cas9 technology to disrupt the three enhancer regions[17]. We generated keratinocytes with insertions and/or deletions (indels) in the A2, B2 or CNE12 enhancer regions by independently introducing six single-guide RNAs (sgRNAs) specific for these enhancer regions (**Supplementary Figs. 6**–**8**). Disruption of these enhancers resulted in a decrease in *ACTRT1* expression relative to cells without enhancer region editing, and decreased expression was correlated with editing efficiency (**Fig. 1h** and **Supplementary Fig. 9**). We also observed that enhancer mutagenesis increased the rate of keratinocyte proliferation (**Supplementary Fig. 10**). Taken together, our results suggest that BDCS is directly associated with loss-of-function mutations either altering the coding region of *ACTRT1* or affecting specific enhancers located in transcribed noncoding surrounding regions.

In further immunohistological studies of sporadic BCC cases, no ARP-T1 signal was detected in BCC tissues from 51 of 60 unrelated sporadic cases (representative of one tumor from a single individual), whereas normal ARP-T1 staining was detected in BCC tissues from 5 individuals with xeroderma pigmentosum (**Supplementary Fig. 11a**). Further sequencing of a series of 20 BCC tissue samples identified three mutations in the coding sequence of *ACTRT1* (**Supplementary Fig. 11b**) and one mutation in the B2 enhancer region (g.127968192G>A) located 69 bp downstream of the mutation found in BDCS family B (g.127968123T>C). These results emphasize the genetic heterogeneity of these skin tumors and suggest that both inherited and sporadic BCCs are associated with loss-of-function mutations in the *ACTRT1* locus.

In order to elucidate the consequences of loss-of-function mutations in *ACTRT1*, we performed comparative transcriptomic analyses of skin samples from individuals with BDCS and identified 1,771 genes with deregulated expression relative to that of three skin samples from healthy individuals (for a description of the statistical analysis, see the Online Methods). Using Ingenuity Pathway Analysis (see URLs), we showed that these genes are mostly involved in regulation of cell-cycle progression, cell death and survival, and cell migration (**Fig. 2a**). As constitutive activation of the Hedgehog pathway is found in more than 70% of BCCs[1,18,19], particular attention was given to transcription factors activated by this pathway. Indeed, among the genes with deregulated expression, 56 are directly controlled by the Hedgehog transcription factors GLI1 and GLI2 (**Fig. 2b**).

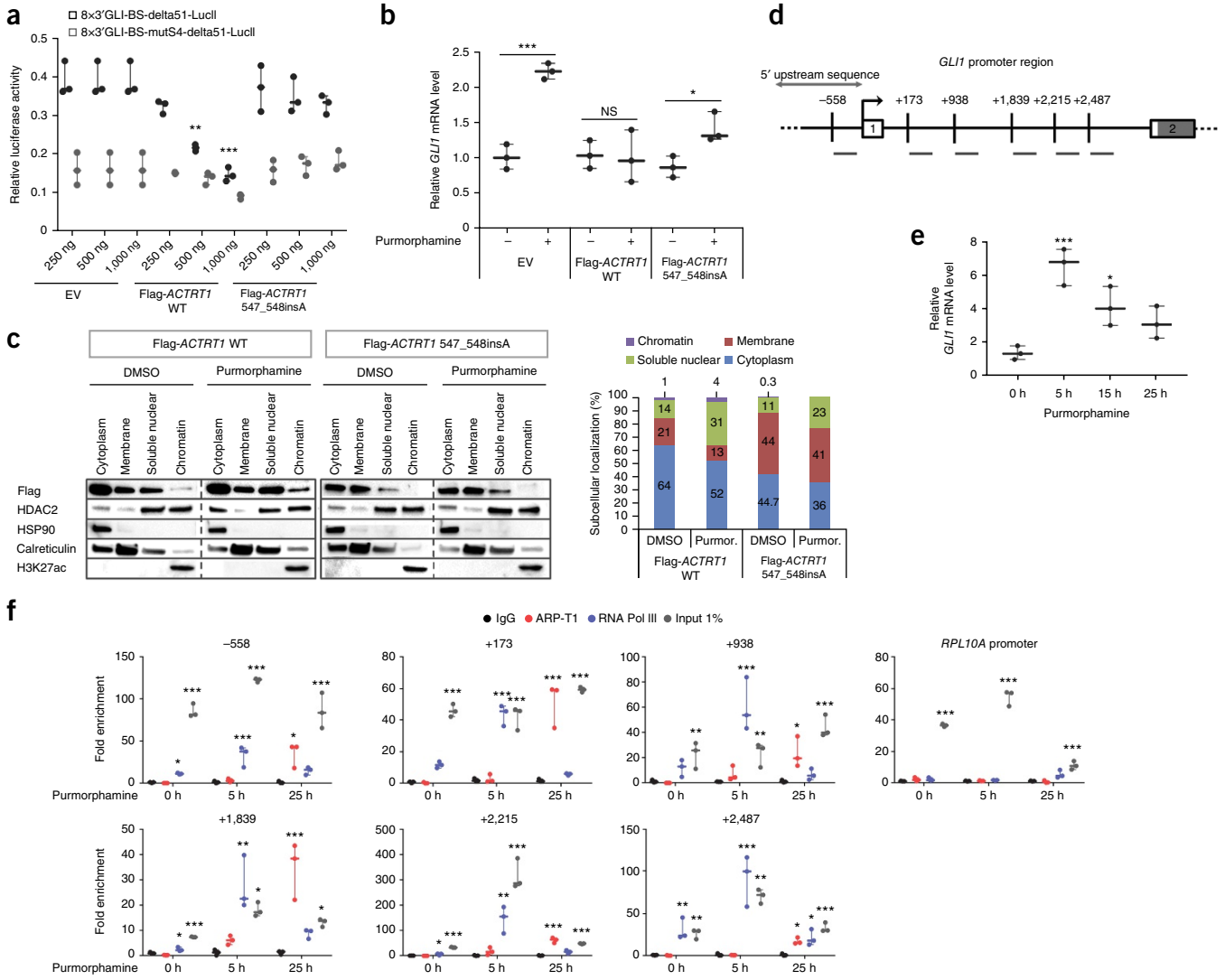

Figure 3 Functional impact of ARP-T1 on the Sonic Hedgehog signaling pathway. (**a**) Luciferase activity of a GLI luciferase reporter (8×3′Gli-BS-delta51-LucII) in HaCaT keratinocytes transfected with wild-type (WT) or mutant (c.547_548insA) *ACTRT1* construct after 24 h of stimulation with purmorphamine (3 μM). A luciferase reporter construct containing mutations in GLI-binding sites (8×3′Gli-BS-mutS4-delta51-LucII) was used as a negative control for activation of the Hedgehog pathway. Luciferase activity was normalized to that of *Renilla* luciferase. The experiment was performed in triplicate, and results are representative of three independent experiments. (**b**) qRT–PCR analysis of *GLI1* mRNA levels normalized to *PGK1* mRNA levels in primary keratinocytes after transfection with empty vector or with vector encoding wild-type or mutant *ACTRT1* following 5 h of stimulation with purmorphamine (3 μM) or DMSO. qRT–PCR was performed in triplicate, and results are representative of two independent experiments. (**c**) Western blot analyses of cytoplasmic, membrane, soluble nuclear and chromatin-bound proteins from HEK293T cells transfected to express Flag-tagged wild-type or mutant ACTRT1 (following 24 of h stimulation of cells with DMSO or purmorphamine (3 μM)). The nuclear marker histone deacetylase 2 (HDAC2), the cytoplasmic marker heat-shock protein 90 (HSP90), the membrane marker calreticulin and the chromatin marker H3K27ac were used to assess protein loading and fraction purity. The plot to the right quantifies the fraction of Flag-tagged protein in each cellular compartment under the different stimulation conditions as determined by band intensity in the western blots (see **Supplementary Data** for full blots). (**d**) Schematic representation of the human *GLI1* promoter region, showing the locations of the primers used in ChIP experiments (in **f**) relative to the transcriptional start site (TSS; indicated by the black arrow). The *GLI1* promoter region includes a 5′ flanking sequence (5′ UTR), an untranslated exon (1) and part of the first intron (see **Supplementary Fig. 12** for details on chromatin marks in the *GLI1* promoter region in NHEK keratinocytes). Untranslated exon sequences are represented by a white square; translated exon sequences are depicted in gray. Numbering above the line indicates nucleotide locations relative to the position of the TSS. (**e**) Levels of *GLI1* mRNA expression in primary keratinocytes over the course of stimulation with purmorphamine (3 μM) at the indicated times. Results were normalized to *PGK1* mRNA levels. qPCR was performed in triplicate, and results are representative of two independent experiments. (**f**) ChIP–qPCR analysis was performed with primers binding to the locations depicted in **d**, showing the fold enrichment of RNA Pol II and ARP-T1 at the *GLI1* promoter sites in primary keratinocytes at the times indicated on the *x* axis over the course of stimulation with purmorphamine (3 μM). Primers for the *RPL10A* promoter were used as a negative control for ARP-T1 binding. qPCR was performed in triplicate, and results are representative of two independent experiments. Data in **a**, **b**, **e** and **f** are expressed as the mean ± s.d. *P* values were calculated by one-way ANOVA with Bonferroni multiple comparisons for **a** (relative to EV), **e** (relative to stimulated cells at 0 h) and **f** (relative to IgG) and by *t*-test for **b** (unstimulated cells versus cells stimulated with purmorphamine). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. NS, not significant.

Consistent with these findings, qPCR showed that Hedgehog target genes were overexpressed in skin samples from individuals with BDCS carrying either a mutation in *ACTRT1* (c.547_548insA) or the CNE12 variant (**Fig. 2c**). Transactivation assays in cells from the HaCaT keratinocyte cell line employing a reporter construct with GLI1-binding sites upstream of the luciferase gene showed

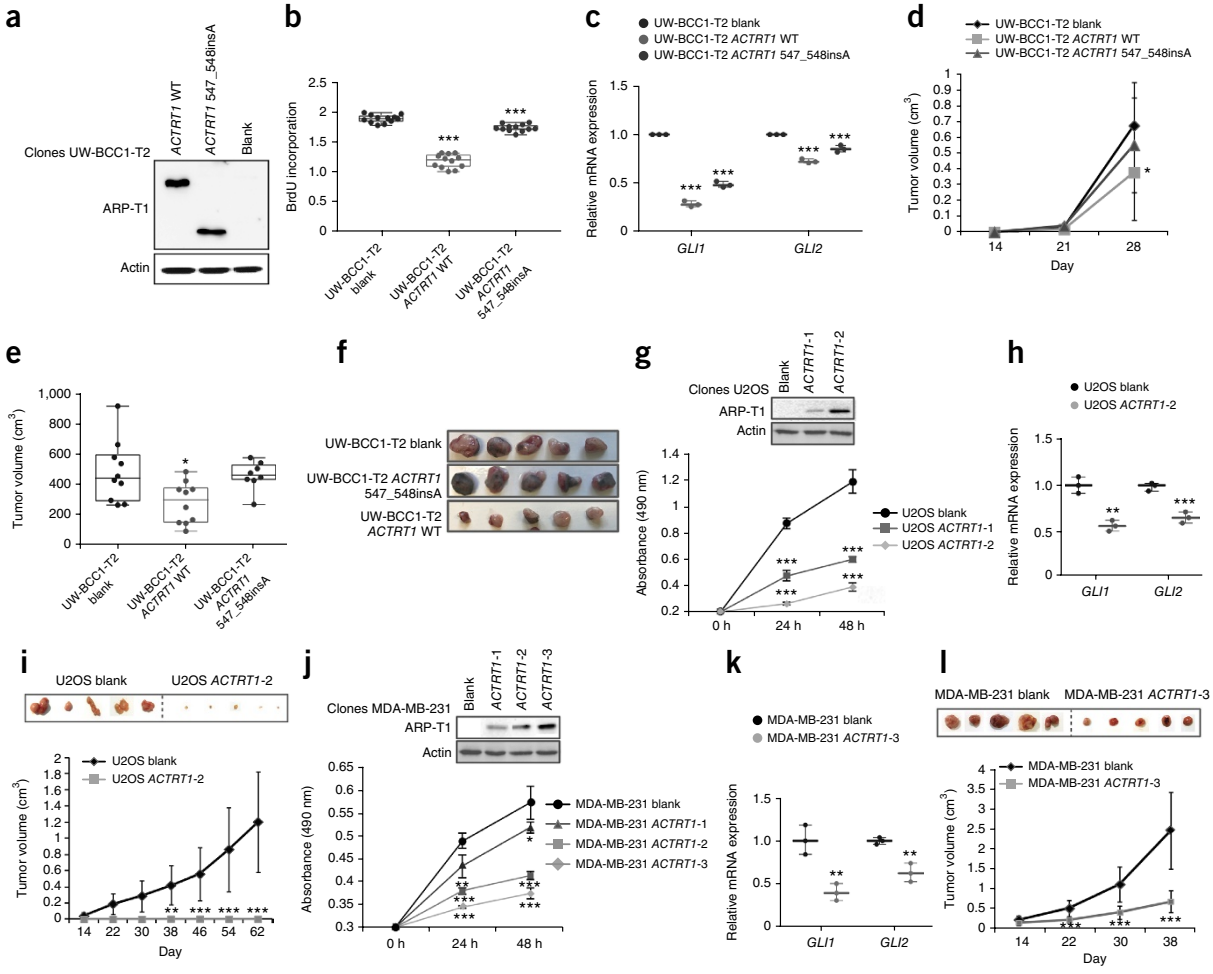

Figure 4 *ACTRT1* functions as a tumor-suppressor gene *in vivo*. (**a**) Western blot analysis of ARP-T1 expression in UW-BCC1-T2 cells transfected with wild-type or mutant (c.547_548insA) *ACTRT1* constructs. Blank indicates transfection with empty vector. (**b**) Proliferation of UW-BCC1-T2 cells expressing wild-type or mutant *ACTRT1* constructs in a BrdU incorporation assay ($n = 12$ in each group). (**c**) qRT–PCR analysis of *GLI1* and *GLI2* mRNA expression normalized to *RPL13A* mRNA expression in UW-BCC1-T2 cells stably expressing wild-type *ACTRT1*, mutant *ACTRT1* or empty vector. qPCR was performed in triplicate, and results are representative of three independent experiments. (**d**) Tumor volume measured at the indicated time points after subcutaneous injection of UW-BCC1-T2 cells stably expressing wild-type or mutant *ACTRT1* into AGR129 mice ($n = 5$ mice per group, 2 tumors per mouse). (**e**) Tumor volumes measured at the final time point indicated in **d**. (**f**) Mice injected with UW-BCC1-T2 cells were euthanized and dissected, and their tumors were photographed ($n = 5$ tumors, representative of the 10 tumors in each group). (**g**) Time course of the proliferation of U2OS cells expressing *ACTRT1* constructs in a methylthiazoltetrazolium (MTT) assay (performed in triplicate; one of three independent experiments is shown). Inset, western blot analysis of ARP-T1 expression in two different clones expressing wild-type *ACTRT1*; actin was used as a loading control. (**h**) qRT–PCR analysis of *GLI1* and *GLI2* mRNA expression normalized to *PGK1* mRNA levels in U2OS cells stably expressing wild-type *ACTRT1* or empty vector. qPCR was performed in triplicate, and results are representative of two independent experiments. (**i**) Time-course analysis of tumor volume for tumors resulting from injection of U2OS cells into NMRI nude mice ($n = 10$ mice per group). Mice were injected subcutaneously with either U2OS control cells or U2OS cells stably expressing *ACTRT1* ($n = 5$ mice per group, 2 tumors per mouse). Mice were euthanized and dissected, and tumors were photographed. Inset, assembly of five representative tumors of each group. (**j**) Time course of the proliferation of MDA-MB-231 cells expressing *ACTRT1* in an MTT assay (performed in triplicate; one of three independent experiments is shown). Inset, western blot analysis of *ACTRT1* expression in various cell clones. Actin was used as a loading control. (**k**) qRT–PCR analysis of *GLI1* and *GLI2* mRNA expression normalized to *PGK1* mRNA levels in MDA-MB-231 cells stably expressing wild-type *ACTRT1* or an empty vector. qPCR was performed in triplicate, and results are representative of two independent experiments. (**l**) Time-course analysis of tumor volume following injection of MDA-MB-231 cells into NMRI nude mice ($n = 10$ mice per group). Mice were injected subcutaneously with either MDA-MB-231 control cells or MDA-MB-231 cells stably expressing *ACTRT1* ($n = 5$ mice per group, 2 tumors per mouse). Mice were euthanized and dissected, and tumors were photographed. Inset, assembly of five tumors representative of each group. Data in **b**–**l** are expressed as the mean ± s.d. For box-and-whisker plots, midlines represent the median; upper and lower perimeters extend from the 25th to the 75th percentile; and whiskers extend from minimum to maximum. *P* values were calculated by one-way ANOVA with Bonferroni multiple comparisons in **b**, **c** and **e** (relative to UW-BCC1-T2 blank), by two-way ANOVA with Bonferroni multiple comparisons in **d**, **g**, **i**, **j** and **l** (relative to cell line blank), and by *t*-test in **h** and **k** (relative to cell line blank). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$. See **Supplementary Data** for full immunoblots.

that expression of wild-type but not truncated ARP-T1 inhibited the Hedgehog pathway (**Fig. 3a**). Moreover, expression of wild-type but not truncated ARP-T1 in primary keratinocytes inhibited *GLI1* expression after stimulation with the Smoothened (Smo) agonist purmorphamine (**Fig. 3b**). Taken together, these results strongly suggest that ARP-T1 has a role in regulating the activity of the Hedgehog signaling pathway.

ARP-T1 belongs to the actin-related protein (ARP) family[20]. In the nucleus, ARP proteins are essential elements of the macromolecular machinery that controls nucleosome remodeling, histone acetylation, histone variant exchange, transcription and DNA repair[21,22]. Using ultrathin sections of normal epidermis processed for transmission electron microscopy, we detected ARP-T1 in both the nucleus and cytoplasm (**Supplementary Fig. 12a**). Western blot analyses of proteins from subcellular fractions of transfected HEK293T cells or control primary keratinocytes showed that ARP-T1 was mostly located in the nucleus and was bound to chromatin following stimulation of the Hedgehog pathway using purmorphamine (**Fig. 3c** and **Supplementary Fig. 12b**). Conversely, the truncated ARP-T1 protein was absent from the chromatin-bound fraction of transfected HEK293T cells (**Fig. 3c**). It has been reported that *GLI1* transcription and Hedgehog pathway activity are controlled by chromatin regulators, including the tumor-suppressors BRG1 (refs. 23,24) and SNF5 (ref. 25), two components of the mammalian SWI–SNF chromatin-remodeling complex that directly bind to *GLI1* regulatory domains. To determine whether ARP-T1 binds to the *GLI1* promoter (**Fig. 3d** and **Supplementary Fig. 13**), we performed ChIP analysis with an anti-ARP-T1 antibody in control primary keratinocytes. An anti–RNA Pol II antibody was used as a positive control for transcriptional activation. Interestingly, stimulation of keratinocytes for 5 h with purmorphamine triggered an increase in *GLI1* expression (**Fig. 3e**) correlating with enrichment of RNA Pol II (but not ARP-T1) binding in *GLI1* promoter regions (**Fig. 3f**). Longer stimulation of keratinocytes with purmorphamine (25 h) resulted in enrichment of ARP-T1 binding in *GLI1* promoter regions (**Fig. 3f**), which was concomitant with a decrease in *GLI1* expression (**Fig. 3e**). Taken together, our results support the notion that ARP-T1 exerts negative control over *GLI1* expression in a manner similar to that observed for SNF5 and BRG1.

In order to investigate both the tumor-suppressor activity of *ACTRT1 in vivo* and the impact of *ACTRT1* mutations, we generated populations of cells that stably expressed wild-type or mutant ARP-T1 from the long-term human BCC cell line UW-BCC1-T2 (ref. 26; **Fig. 4a**), which is characterized by enhanced activation of the Hedgehog signaling pathway. Interestingly, expression of wild-type ARP-T1 significantly decreased the cell proliferation rate as compared to that of UW-BCC-T2 cells expressing the empty vector (blank) (**Fig. 4b**) concomitantly with a reduction in *GLI1* and *GLI2* expression (**Fig. 4c**). Conversely, mutant ARP-T1 (c.547_548insA) only partially inhibited cell proliferation and had a less pronounced effect in reducing *GLI1* and *GLI2* expression. Similarly, in tumor xenografts of these cells in AGR129 mice, wild-type ARP-T1, but not the truncated variant, attenuated tumor development (**Fig. 4d–f**) and Ki-67 (proliferation marker) expression (**Supplementary Fig. 14**). These results further confirm the tumor-suppressor role of ARP-T1 in limiting BCC development *in vivo*.

Interestingly, a rare deletion in the *ACTRT1* locus has been reported in 17 of 63 familial cases of early-onset hereditary breast cancer that lack *BRCA1* and *BRCA2* mutations[27]. We hypothesized that ARP-T1 might have more general tumor-suppressor activity in addition to limiting BCC growth, so we selected two human cancer cell lines in which the Hedgehog signaling pathway is aberrantly active (the U2OS osteosarcoma cell line[28,29] and the MDA-MB-231 breast cancer cell line[30,31]) for further study; notably, ARP-T1 is highly expressed in glandular breast lobules and weakly expressed in osteocytes (**Supplementary Fig. 15**), but its expression is absent in these two cancer cell lines (**Fig. 4g,j**). In addition, using the specific GLI inhibitor GANT61, we confirmed the GLI-dependent growth of U2OS and MDA-MB-231 cells[31] (**Supplementary Fig. 16**). Stable expression of ARP-T1 in these cell lines inhibited *in vitro* proliferation (**Fig. 4g,j**) and reduced *GLI1* and *GLI2* expression (**Fig. 4h,k**). Consistent with these findings, in tumor xenograft experiments in NMRI nude mice, ARP-T1 overexpression prevented *in vivo* growth of injected U2OS cells (**Fig. 4i**) and reduced the development of xenograft tumors from MDA-MB-231 cells (**Fig. 4l**). Previously, *GLI1* inhibition has been shown to attenuate growth and migration of MDA-MB-231 cells by increasing apoptosis and decreasing cell proliferation[30]. In accordance with these observations, expression of exogenous wild-type ARP-T1 increased the number of apoptotic MDA-MB-231 cells, increased the expression of pro-caspase-3 and cleaved, active caspase-3 (**Supplementary Fig. 17a–c**), and decreased Ki-67 expression in MDA-MB-231 cells (**Supplementary Fig. 18**) in comparison to expression of the empty vector (blank). Expression of the wild-type protein also reduced the metastatic potential of these MDA-MB-231 cells *in vitro* and *in vivo* (**Supplementary Fig. 19a–e**).

Germline mutations in the *ACTRT1* coding sequence and its surrounding noncoding elements constitute a hitherto unreported causative mechanism of inherited predisposition to BCC. Here, to our knowledge, we report for the first time the involvement of *ACTRT1* in human disease. We suggest potential functions of ARP-T1 and mechanisms through which it could regulate *GLI1* expression on the basis of our results and some models in the literature describing the ARP family. First of all, we demonstrated that wild-type ARP-T1, but not the truncated protein identified in BDCS families (C and D, carrying the c.547_548insA *ACTRT1* mutation), binds to chromatin, suggesting that ARP-T1 has nuclear functions. Nuclear ARP proteins are components of the four main chromatin-remodeling complexes (CRCs): INO80, SRCAP, BAF (the human analog of the SWI–SNF complex) and TIP60/TRRAP[21,22]. ARP5 and ARP8 can bind to core histones to facilitate interaction of the CRCs with nucleosomes[21,22]. Mutations in the genes encoding the major mammalian SWI–SNF (BAF) CRC subunits are present in over 20% of human cancers[31]. SNF5 (the core component of the SWI–SNF CRC) and BRG1 (the ATPase subunit) are encoded by bona fide tumor-suppressors genes, in which mutations are responsible for various types of cancer[23–25]. Interestingly, both SNF5 and BRG1 act through direct inhibition of GLI genes[24,25], which is also the mechanism of ARP-T1 action. Further studies are needed to determine whether ARP-T1 acts specifically at the level of regulatory elements through recognition of specific histone marks, repositioning of nucleosomes, histone exchange or binding to transcription factors. Also, we cannot exclude the possibility that ARP-T1 could interact with key proteins involved in various Hedgehog-interconnected pathways related to tumor proliferation and progression, such as the p16–RB, Wnt, and Polycomb pathways, as has been demonstrated for SNF5 (refs. 32,33).

The discovery of distal regulatory noncoding elements, known as enhancers, with critical functions in gene expression has added a new dimension to transcriptional regulation. eRNAs are potent transcription units, and their alteration can impact biological processes involved in human diseases, including cancer[34]. They cover a broad

spectrum of molecular and cellular functions by implementing different modes of action. Studies have confirmed that long noncoding RNAs (lncRNAs) contribute to cancer initiation and progression by regulating gene transcription[35]. For instance, a link between lncRNAs and the SWI–SNF complexes has been reported in various tumoral conditions[36]. Another mechanism by which eRNAs act in gene regulation involves their interaction with cohesin and Mediator complexes to stabilize enhancer–promoter looping, causing chromatin stabilization and gene expression[37]. eRNAs are also involved in the recruitment of RNA Pol II to gene promoters and facilitate the access of specific transcription factors to enhancer sequences[38]. In addition to DNA methylation and histone modification, a role for noncoding RNA has recently emerged in epigenetic control. Alterations of eRNAs have been found to result in epigenomic reprogramming during tumor initiation and progression[39]. Further investigation is needed to place ARP-T1 and its noncoding regulatory elements in such complex mechanisms of gene regulation and cancer development. More broadly, our findings shed light on the functional relevance of genomic alterations in noncoding regions and their contribution to tumor development. Indeed, the clinical integration of noncoding RNAs as functionally relevant elements in conjunction with additional predictive biomarkers could improve the management of individuals with cancer.

URLs. Ingenuity Pathway Analysis, http://www.ingenuity.com/.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

The authors wish to thank P. Guigue, A. Vincent-Salomon, L. Fuhrmann, M. Oufadem, S. Thomas, S. Petit, F. Mouchrik, V. Geoffroy, A. Julien, O. Duchamp de Lageneste, L. Galmiche-Rolland, J. Cottineau, J. Rossignol, S. Saunier and D. Pouly for technical assistance and advice; M. Sahbatou for performing linkage analysis; M.-E. Huang for providing cell lines for this study; A. Schmitt and J.-M. Masse at the Cochin Imaging Facility (CNRS UMR 8104) for their assistance and help with transmission electron microscopy; C. Bole-Feysot at the Genomic Platform of IMAGINE Institute for acquiring transcriptomic data; M. Garfa-Traoré, N. Goudin and R. Desvaux at the Cell Imaging Platform of IMAGINE Institute for technical assistance and advice; P. Guigue-Rodet and V. Martin-Bouret for providing skin biopsies from some individuals with BDCS; and M.-C. Hors-Cayla and J.-L. Bonafé for initiating genetic research on BDCS. We are grateful to all subjects and their families for their participation in the study. H.E. was supported by TUBITAK (grant no. 112S398). H.-S.P., E.C., F.K., D.B., M.H. and D.H. were supported by the Swiss National Science Foundation, the Placide Nicod Foundation and the Dind Cottier Foundation. This work was funded by the Association pour la Recherche contre le Cancer and the Société Française de Dermatologie. This work was supported by funding from the Agence Nationale de la Recherche (ANR-10-IAHU-01).

AUTHOR CONTRIBUTIONS

E.B. designed and conducted most of the experiments and analyzed all the data. H.-S.P., E.C., F.K., D.B., M.H. and D.H. performed experiments on UW-BCC1 cells. Z.B.-C. contributed to the *in vivo* studies. H.K. collected BDCS families, provided clinical data, and performed and carried out analysis for linkage studies. M.N. and H.R.C. searched for CNEs and performed *in silico* analyses of the region around the *ACTRT1* gene. M.M. assisted with all experiments. N.C. performed transcriptomic and Ingenuity analyses. C.L.G. contributed to sequencing of candidate genes. F.C. contributed to flow cytometry analyses and cell fractionation experiments. S.H. contributed to electron microscopy analyses. S.F. performed immunohistochemical analyses of skin biopsies. O.H. contributed to discussions on some aspects of the project. R.O.R. and A.D.A. extracted DNA from families E and F and performed linkage analyses. Q.W. helped with some *in silico* analysis on Turkish families. S.H-R., C.B., F.M.-P., B.L., F.C., J.M.-H., N.P., N.L., A.T. and M.-F.A. provided well-characterized patient samples. T.M. contributed to discussions and provided tools for experiments on the Hedgehog signaling pathway. G.G. performed targeted high-throughput sequencing. D.J.H. contributed to discussions on some aspects of the project. P.V. provided well-characterized patient samples and participated in drafting the manuscript. A.M. contributed to the study design and participated in drafting the manuscript. A.S. designed the study and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Epstein, E.H. Basal cell carcinomas: attack of the hedgehog. *Nat. Rev. Cancer* **8**, 743–754 (2008).
2. Vabres, P. *et al.* The gene for Bazex–Dupré–Christol syndrome maps to chromosome Xq. *J. Invest. Dermatol.* **105**, 87–91 (1995).
3. Natoli, G. & Andrau, J.C. Noncoding transcription at enhancers: general principles and functional models. *Annu. Rev. Genet.* **46**, 1–19 (2012).
4. Lam, M.T., Li, W., Rosenfeld, M.G. & Glass, C.K. Enhancer RNAs and regulated transcriptional programs. *Trends Biochem. Sci.* **39**, 170–182 (2014).
5. Mousavi, K., Zare, H., Koulnis, M. & Sartorelli, V. The emerging roles of eRNAs in transcriptional regulatory networks. *RNA Biol.* **11**, 106–110 (2014).
6. Bazex, A., Dupré, A. & Christol, B. Follicular atrophoderma, baso-cellular proliferations and hypotrichosis. *Ann. Dermatol. Syphiligr. (Paris)* **93**, 241–254 (1966).
7. Rapelanoro, R., Taïeb, A. & Lacombe, D. Congenital hypotrichosis and milia: report of a large family suggesting X-linked dominant inheritance. *Am. J. Med. Genet.* **52**, 487–490 (1994).
8. Visel, A., Rubin, E.M. & Pennacchio, L.A. Genomic views of distant-acting enhancers. *Nature* **461**, 199–205 (2009).
9. Nelson, A.C. & Wardle, F.C. Conserved non-coding elements and *cis* regulation: actions speak louder than words. *Development* **140**, 1385–1395 (2013).
10. Nobrega, M.A., Ovcharenko, I., Afzal, V. & Rubin, E.M. Scanning human gene deserts for long-range enhancers. *Science* **302**, 413 (2003).
11. Blow, M.J. *et al.* ChIP–Seq identification of weakly conserved heart enhancers. *Nat. Genet.* **42**, 806–810 (2010).
12. Yang, S. *et al.* Functionally conserved enhancers with divergent sequences in distant vertebrates. *BMC Genomics* **16**, 882 (2015).
13. Visel, A. *et al.* ChIP–seq accurately predicts tissue-specific activity of enhancers. *Nature* **457**, 854–858 (2009).
14. Andersson, R. *et al.* An atlas of active enhancers across human cell types and tissues. *Nature* **507**, 455–461 (2014).
15. Inoue, F. & Ahituv, N. Decoding enhancers using massively parallel reporter assays. *Genomics* **106**, 159–164 (2015).
16. Wu, H. *et al.* Tissue-specific RNA expression marks distant-acting developmental enhancers. *PLoS Genet.* **10**, e1004610 (2014).
17. Korkmaz, G. *et al.* Functional genetic screens for enhancer elements in the human genome using CRISPR–Cas9. *Nat. Biotechnol.* **34**, 192–198 (2016).
18. Xie, J. *et al.* Activating Smoothened mutations in sporadic basal-cell carcinoma. *Nature* **391**, 90–92 (1998).
19. Gailani, M.R. *et al.* The role of the human homologue of *Drosophila* patched in sporadic basal cell carcinomas. *Nat. Genet.* **14**, 78–81 (1996).
20. Heid, H. *et al.* Novel actin-related proteins Arp-T1 and Arp-T2 as components of the cytoskeletal calyx of the mammalian sperm head. *Exp. Cell Res.* **279**, 177–187 (2002).
21. Oma, Y. & Harata, M. Actin-related proteins localized in the nucleus: from discovery to novel roles in nuclear organization. *Nucleus* **2**, 38–46 (2011).
22. Dion, V., Shimada, K. & Gasser, S.M. Actin-related proteins in the nucleus: life beyond chromatin remodelers. *Curr. Opin. Cell Biol.* **22**, 383–391 (2010).
23. Medina, P.P. *et al.* Frequent *BRG1*/*SMARCA4*-inactivating mutations in human lung cancer cell lines. *Hum. Mutat.* **29**, 617–622 (2008).
24. Zhan, X., Shi, X., Zhang, Z., Chen, Y. & Wu, J.I. Dual role of Brg chromatin remodeling factor in Sonic hedgehog signaling during neural development. *Proc. Natl. Acad. Sci. USA* **108**, 12758–12763 (2011).
25. Jagani, Z. *et al.* Loss of the tumor suppressor Snf5 leads to aberrant activation of the Hedgehog–Gli pathway. *Nat. Med.* **16**, 1429–1433 (2010).
26. Noubissi, F.K. *et al.* Role of CRD-BP in the growth of human basal cell carcinoma cells. *J. Invest. Dermatol.* **134**, 1718–1724 (2014).
27. Krepischi, A.C. *et al.* Germline DNA copy number variation in familial and early-onset breast cancer. *Breast Cancer Res.* **14**, R24 (2012).
28. Lo, W.W. *et al.* Involvement and targeted intervention of dysregulated Hedgehog signaling in osteosarcoma. *Cancer* **120**, 537–547 (2014).

29. Yang, W. *et al.* Targeting hedgehog–GLI-2 pathway in osteosarcoma. *J. Orthop. Res.* **31**, 502–509 (2013).
30. Kwon, Y.J. *et al.* Gli1 enhances migration and invasion via up-regulation of MMP-11 and promotes metastasis in ERα negative breast cancer cell lines. *Clin. Exp. Metastasis* **28**, 437–449 (2011).
31. St Pierre, R. & Kadoch, C. Mammalian SWI/SNF complexes in cancer: emerging therapeutic opportunities. *Curr. Opin. Genet. Dev.* **42**, 56–67 (2017).
32. Wilson, B.G. *et al.* Epigenetic antagonism between Polycomb and SWI/SNF complexes during oncogenic transformation. *Cancer Cell* **18**, 316–328 (2010).
33. Kohashi, K. & Oda, Y. Oncogenic roles of SMARCB1/INI1 and its deficient tumors. *Cancer Sci.* **108**, 547–552 (2017).
34. Ørom, U.A. *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* **143**, 46–58 (2010).
35. Bartonicek, N., Maag, J.L. & Dinger, M.E. Long noncoding RNAs in cancer: mechanisms of action and technological advancements. *Mol. Cancer* **15**, 43 (2016).
36. Tang, Y. *et al.* Linking long non-coding RNAs and SWI/SNF complexes to chromatin remodeling in cancer. *Mol. Cancer* **16**, 42 (2017).
37. Yang, L. *et al.* lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. *Nature* **500**, 598–602 (2013).
38. Zabidi, M.A. & Stark, A. Regulatory enhancer-core-promoter communication via transcription factors and cofactors. *Trends Genet.* **32**, 801–814 (2016).
39. Wei, J.W., Huang, K., Yang, C. & Kang, C.S. Non-coding RNAs as regulators in epigenetics (Review). *Oncol. Rep.* **37**, 3–9 (2017).

[1]Paris Descartes University, Sorbonne Paris Cité, Paris, France. [2]IMAGINE Institute, INSERM UMR 1163, Paris, France. [3]Department of Dermatology, Lausanne University Hospital, Hôpital de Beaumont, Lausanne, Switzerland. [4]Department of Hematology, Hôpital Necker–Enfants Malades, Paris, France. [5]Medical Genetics Department, Koç University School of Medicine (KUSOM), Istanbul, Turkey. [6]Medical Genetics Department, Istanbul Medical Faculty, Istanbul University, Istanbul, Turkey. [7]Institut de Biologie de l'ENS (IBENS), École Normale Supérieure, Paris, France. [8]CNRS, UMR 8197, Paris, France. [9]INSERM U1024, Paris, France. [10]Department of Dermatology, Hôpital Necker–Enfants Malades, Paris, France. [11]Plateforme Bio-informatique, Structure Fédérative de Recherche Necker, INSERM US24/CNRS, UMS 3633, Paris, France. [12]Laboratory of Genome Maintenance, Rockefeller University, New York, New York, USA. [13]Department of Clinical Genetics, Vrije Universiteit Medical Center, Amsterdam, the Netherlands. [14]GR-Ex Laboratory of Excellence, IMAGINE Institute, Paris, France. [15]Centre Référence Nationale pour les Mastocytoses, Hôpital Necker–Enfants Malades, Paris, France. [16]Centre de référence pour les maladies rares de la peau, Service de Dermatologie, Centre Hospitalier Universitaire de Bordeaux, Bordeaux, France. [17]Department of Dermatology, Centre Hospitalier Universitaire Nord, Saint-Etienne, France. [18]Department of Dermatology, Hôpital Avicenne, Bobigny, France. [19]Department of Dermatology, Centre de Référence des Maladies Rares de la Peau, Hôpital Larrey, Toulouse, France. [20]Department of Medical Genetics, Hôpital de la Timone, Marseille, France. [21]AMU-INSERM, UMR_S910, Faculté de Médecine de Marseille, Marseille, France. [22]INSERM U1035, Université de Bordeaux, Bordeaux, France. [23]Department of Dermatology, Hôpital Cochin, Paris, France. [24]Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, UK. [25]Genoscope (CEA), CNRS UMR 8030, University of Evry, Evry, France. [26]Institute for Research on Cancer and Aging, CNRS UMR 7284, INSERM U1081, University of Nice Sophia Antipolis, Nice, France. [27]Department of Pathological Anatomy, Hôpital Necker–Enfants Malades, Paris, France. [28]Department of Dermatology, Centre Hospitalier Universitaire, Hôpital du Bocage, Dijon, France. [29]These authors contributed equally to this work. Correspondence should be addressed to A.S. (asma.smahi@inserm.fr) or E.B. (elodie.bal@inserm.fr).

ONLINE METHODS

Subjects and samples. A total of 48 individuals (from six BDCS families) and many of their unaffected relatives underwent a comprehensive clinical examination. All affected individuals had two or more of the following symptoms of BDCS upon clinical examination or in their personal medical history: hypotrichosis, facial milia, follicular atrophoderma and BCC. All individuals provided their written, informed consent for genetic studies. The study was approved by the local investigational review board (Hôpital Necker–Enfants Malades, Paris, France). DNA was extracted from peripheral blood lymphocytes by conventional phenol-chloroform purification. Paraffin-embedded skin tumor samples were obtained from six individuals who underwent surgical excision of biopsy-confirmed BCCs. Frozen skin samples were obtained for two individuals. In addition, 81 paraffin-embedded skin tumor samples were obtained from the specimen collection at Necker Hospital's Pathology Department (French Ministry of Research reference number: DC-2009-955).

Targeted high-throughput sequencing. Next-generation sequencing with targeted enrichment of the 7,633,224-bp interval spanning the BDCS-associated gene was performed in four affected family members and one unaffected family member at the Genoscope facility (Evry, France). A custom sequence capture array (Roche NimbleGen) encompassing the region chr. X: 123,576,802–131,210,128 was used to hybridize shotgun fragment libraries obtained from selected subjects. Massively parallel sequencing was performed on this enriched library using a Solexa sequencer (Illumina). Sequence data were aligned to the hg19 reference version of the human genome. On average, 29,485,436 sequencing reads were obtained, with an average length of 200 bp. 95% of all reads could be mapped and 77% overlapped with the enriched regions. 93% of the enriched regions were covered entirely. The mean coverage for the five family members was 246×.

ChIP–qPCR. ChIP assays were performed on frozen epidermis to determine the chromatin signature or on primary keratinocytes for the study of ARP-T1 recruitment to the *GLI1* promoter, using an EZ-Magna ChIP A/G kit (Millipore). Briefly, epidermis (1 mg) or keratinocytes (1×10^7 cells) were treated with 1% formaldehyde for 10 min. Cross-linked chromatin was then sonicated to yield a mean fragment size of 300–500 bp. To determine the chromatin signature, DNA fragments were immunoprecipitated with anti-H3K27 (Millipore, 17-683), anti-H3K4me1 (Millipore, 07-436), anti-H3K4me3 (Millipore, 17-614), anti–RNA Pol II (Millipore, 05-623B) and IgG (provided with the EZ-Magna ChIP A/G Chromatin Immunoprecipitation Kit, Millipore, 17-186). These antibodies were validated for ChIP experiments by the manufacturer. For the study of ARP-T1 recruitment to the *GLI1* promoter, DNA fragments were immunoprecipitated with anti-ARP-T1 (PA5-31691, immunogen sequence residues 174–376, Thermo Scientific), anti–RNA Pol II and IgG antibodies. Sequences of interest were amplified by qPCR using Power SYBR Green PCR Master Mix (PE Applied Biosystems). Reactions were performed in triplicate on an ABI Prism 7000 machine (PE Applied Biosystems). Input recovery was calculated using a ChIP–qPCR data analysis calculation shell (Sigma-Aldrich). The oligonucleotides used in these analyses are listed in **Supplementary Table 1**. In determining the chromatin signature, the efficiency of the primers was calculated on sonicated DNA (input) after several dilutions and small differences between efficiencies were observed between the studied sequences (A1, 87%; A2, 87%; A3, 90%; B1, 90%; B2, 91%; CNE12, 101%).

Deletion of enhancer regions using the CRISPR-associated RNA-guided endonuclease Cas9. A lentiviral-based single vector (LentiCRISPR v2) that simultaneously delivers Cas9, sgRNA and a puromycin selection marker engineered by the Zhang laboratory was purchased from Addgene (http://www.addgene.org/). Guide RNA sequences were designed using an online tool (http://crispr.mit.edu/; **Supplementary Table 1**). Cloning of a guide sequence into the LentiCRISPR v2 vector was performed according to protocols available from the Zhang laboratory (http://www.genome-engineering.org/). To produce lentivirus particles, LentiCRISPR v2 vector encoding a specific guide RNA sequence was cotransfected into HEK293T cells with packaging vectors psPAX2, pMD2.G and pRSV-REV using jetPRIME reagent (Polyplus Transfection). Infectious lentiviruses were collected at 24 and 48 h after transfection and filtered through 0.8-µm cellulose acetate filters. Recombinant lentiviruses were concentrated by ultracentrifugation (2 h at 20,000*g*) and resuspended in HBSS buffer. Primary keratinocytes were transduced with virus particles and selected with puromycin (1 µg/ml) for 2 d. Deletion of targeted regions was determined using the online tool Tide-calculator (https://tide.nki.nl) after PCR amplification and Sanger sequencing.

Transcriptomic analyses. Microarray experiments were performed on Affymetrix GeneChip Human Transcriptome Array 2.0 (a genome-wide array with 70,523 probe sets). CEL files with raw data were imported into R/Bioconductor using the Oligo package. Expression levels were calculated using the RMA algorithm from the affy package, and flags were computed using a custom algorithm within R. Assuming that a maximum of 80% of genes are expressed, we set the 20% of probes with the lowest values for each microarray as background. A threshold was fixed at 2 s.d. above the mean intensity for the background. All probes with normalized intensity measures lower than the computed threshold were flagged as 0 instead of 1. An unsupervised analysis step was done by clustering before any supervised statistical comparison to identify natural groups among the tested samples and to detect potential outliers. For each comparison, a list was created by filtering to include probes flagged as "background" for no more than half of the samples according to flagged measurements for the relevant chips. Group comparisons were performed using Student's *t*-tests, and lists were filtered to include probes with P value ≤5% and fold change ≥1.2. Heat maps were created using a custom R script and the Java TreeView software. Functional analyses of the resulting lists of genes were performed using Ingenuity Pathway Analysis (http://www.ingenuity.com/).

Tumor growth assays using UW-BCC1-T2 cells. Animal experiments were performed in accordance with Swiss guidelines and regulations for the care and use of laboratory animals. Adult (aged 5–7 weeks) male AGR129 mice (deficient for IFN-α/β receptor, IFN-γ receptor and RAG-2) (specific-pathogen-free (SPF)-housed homozygous 129/C57BL/6, B6.129S2-Rag2tm1.1Cgn-Ifnar1tm1Agt-Ifngr1tm1Agt/J; according to the nomenclature available from http://www.informatics.jax.org/mgihome/nomen/) were used as host animals for grafted tumors[39]. Establishment of UW-BCC1 cells was described in Noubissi *et al.*[28]. To improve both tumor establishment and growth in AGR129 mice, UW-BCC1-T2 cells were isolated after two successive *in vivo* passages. Primary tumors were initiated by subcutaneous injection of UW-BCC1-T2–ACTRT1, UW-BCC1-T2–ACTRT1 c.547_548insA or UW-BCC1-T2–Blank cells (2×10^6 cells in 100 µl of PBS) into both the right and left lateral flanks of AGR129 mice. Each group injected was composed of five mice. Tumor growth was monitored weekly by measuring tumor volume using a caliper. Tumor volume (V) was calculated as $V = \pi/6 \times a \times b^2$, where a is the longer and b is the shorter of two orthogonal diameters. Mice were killed by cervical dislocation, and tumors were excised and embedded in paraffin. Eight-micrometer sections were stained with H&E and Ki-67 antiserum (Dako, M7249, clone MIB-1, mouse, 1:100 dilution).

Tumor growth assays using MDA-MB-231 and U2OS cells. Athymic NMRI female nude mice (*Foxn1*^{nu}/*Foxn1*^{nu}; Elevage Janvier) aged 6 weeks were housed in filtered-air laminar flow cabinets and handled under aseptic conditions. Procedures involving animals were approved by INSERM's Institutional Animal Care and Use Committee. MDA-MB-231-*ACTRT1*, U2OS-*ACTRT1*, MDA-MB-231-Blank and U2OS-Blank cells (10^6 cells per injection in 50 µl of 1× PBS and 50 µl of Matrigel (Becton-Dickinson)) were injected subcutaneously into the lateral flanks of each animal (two injections per mouse and five mice per group, chosen as a statistically robust sample size in accordance with the Institutional Animal Care and Use Committee's recommendation). Every week, tumor size was measured (using calipers) in a blinded manner by two investigators (neither of whom performed the injections). Tumor volume (V) was calculated using the equation $V = l \times w \times d$, where l is the length, w is the width and d is the depth of the tumor. Mice were killed by cervical dislocation, and tumors were excised, weighed, fixed in 5% acetic acid:10%

doi:10.1038/nm.4368

formaldehyde in 37% ethanol, and embedded in paraffin. Four-micrometer sections were stained with H&E reagent for histological analysis or with the following primary antibodies for immunohistochemical analysis: anti-Ki-67 (KI67PCE, Leica) and anti-cytokeratin (clone 34Be12, Dako) antibodies. Apoptosis assays were performed using an *In Situ* Cell Death Detection Kit, Fluorescein (Roche Diagnostics).

Statistical analysis. Results were expressed as the mean ± s.d. Statistical significance was determined using unpaired, two-sample equal-variance *t*-tests or one-way or two-way ANOVA with Bonferroni multiple comparisons. All data were normally distributed, and the variance was similar in all compared groups. The threshold for statistical significance was set to $P < 0.05$.

Additional methods. The procedures are described in detail in the **Supplementary Note**.

A **Life Sciences Reproducibility Summary** for this paper is available.

Data availability. Transcriptomic data are available in the ArrayExpress database under accession number E-MTAB-5597. Source data files for **Figures 1**–**4** are available online.

nature research

Corresponding author(s): Asma Smahi

☐ Initial submission ☐ Revised version ☒ Final submission

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

Experimental design

1. Sample size

Describe how sample size was determined.

Sample size has been chooses until to obtain statistical significance

2. Data exclusions

Describe any data exclusions.

there is no data exclusions

3. Replication

Describe whether the experimental findings were reliably reproduced.

the experimental findings were reliably reproduced throught repeated experiments

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Participants were allocated and experimental groups were designed depending the study

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

blinding were used in data acquisition and analysis

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☒ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ A statement indicating how many times each experiment was replicated
☐	☒ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
☐	☒ A description of any assumptions or corrections, such as an adjustment for multiple comparisons
☐	☒ The test results (e.g. *P* values) given as exact values whenever possible and with confidence intervals noted
☐	☒ A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
☐	☒ Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

Nature Medicine: doi:10.1038/nm.4368

Software

Policy information about availability of computer code

7. Software

Describe the software used to analyze the data in this study.	*PRISM*

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* guidance for providing algorithms and software for publication provides further information on this topic.

Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.	*There are no restrictions on availability of unique material*

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).	*Data are provided in the manuscript*

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.	*Data on eucaryotic cell lines are available in the manuscripy*
b. Describe the method of cell line authentication used.	*origin authtification has been used*
c. Report whether the cell lines were tested for mycoplasma contamination.	*All cell lines were assessed for mycoplasma contamination*
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.	*No misdentified cell lines were used in the current study .*

Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals

Details on animals used in the study are provided in the manuscript	*All details on animals are included in the manuscript*

Policy information about studies involving human research participants

12. Description of human research participants

the ethnic origin and provenance of human participants are provided in the manuscript	*clinical and genetic informations of human participants are provided in the manuscript*