nature genetics

Genomic analysis identifies new drivers and progression pathways in skin basal cell carcinoma

Ximena Bonilla[1,19], Laurent Parmentier[2,19], Bryan King[3], Fedor Bezrukov[4,5], Gürkan Kaya[6], Vincent Zoete[7], Vladimir B Seplyarskiy[8–10], Hayley J Sharpe[11], Thomas McKee[12], Audrey Letourneau[1], Pascale G Ribaux[1], Konstantin Popadin[1], Nicole Basset-Seguin[13], Rouaa Ben Chaabene[1], Federico A Santoni[1,14], Maria A Andrianova[8–10], Michel Guipponi[14], Marco Garieri[1], Carole Verdan[12], Kerstin Grosdemange[6], Olga Sumara[15], Martin Eilers[15,16], Iannis Aifantis[3], Olivier Michielin[7,17], Frederic J de Sauvage[11], Stylianos E Antonarakis[1,14,18] & Sergey I Nikolaev[1,14]

Basal cell carcinoma (BCC) of the skin is the most common malignant neoplasm in humans. BCC is primarily driven by the Sonic Hedgehog (Hh) pathway. However, its phenotypic variation remains unexplained. Our genetic profiling of 293 BCCs found the highest mutation rate in cancer (65 mutations/Mb). Eighty-five percent of the BCCs harbored mutations in Hh pathway genes (*PTCH1*, 73% or *SMO*, 20% ($P = 6.6 \times 10^{-8}$) and *SUFU*, 8%) and in *TP53* (61%). However, 85% of the BCCs also harbored additional driver mutations in other cancer-related genes. We observed recurrent mutations in *MYCN* (30%), *PPP6C* (15%), *STK19* (10%), *LATS1* (8%), *ERBB2* (4%), *PIK3CA* (2%), and *NRAS*, *KRAS* or *HRAS* (2%), and loss-of-function and deleterious missense mutations were present in *PTPN14* (23%), *RB1* (8%) and *FBXW7* (5%). Consistent with the mutational profiles, N-Myc and Hippo-YAP pathway target genes were upregulated. Functional analysis of the mutations in *MYCN*, *PTPN14* and *LATS1* suggested their potential relevance in BCC tumorigenesis.

BCC of the skin is the most common cancer in humans, accounting for about 90% of all skin cancers[1]. BCC originates from stem cells in the interfollicular epidermis or hair follicle infundibulum undergoing ultraviolet light (UV)-induced mutagenesis[2]. Histologically, BCCs are classified into four major subtypes: nodular, morpheaform, metatypical and superficial. Early studies of Gorlin syndrome, a Mendelian disease with a high prevalence of BCC (MIM 109400), identified germline mutations in the key Hh receptor *PTCH1* (ref. 3). Additionally, sporadic BCCs exhibit aberrant activation of the Hh pathway resulting from either genetic inactivation of *PTCH1* or activating mutations in *SMO*, a key Hh signal transducer[4–6] that promotes the activity and nuclear localization of GLI transcription factors. A small fraction of tumors harbor loss-of-function mutations in the negative regulator of the Hh pathway *SUFU*[4,5,7]. Aberrant activation of the Hh signaling pathway through mutation of these components also occurs in medulloblastoma, meningioma and ameloblastoma[8–12]. The next most prevalent events in BCC are alterations in *TP53*, which are present in over one-half of tumors[13].

Hyperactivation of the Hh pathway is the most common driver of BCC; therefore, tumors generally respond well to SMO inhibitors such as vismodegib[14–17]. However, about 30% of treated tumors either do not respond or develop resistance to treatment and relapse[5,6,17,18] through mutations in *SMO* or a combination of *SUFU* inactivation and *GLI2* amplification[5,6].

Although the most common driver pathway in BCC is known, tumors display great variability in aggressiveness, morphology and response to treatment. We therefore hypothesized that additional genes and pathways may contribute to BCC formation and progression. To test this hypothesis, we analyzed a cohort of 293 BCCs and found that additional driver mutations beyond the *PTCH1-SMO* axis are highly prevalent in this tumor type. Here we report a high frequency of mutations in cancer-related genes not previously associated with BCC that could contribute to tumorigenesis or phenotypic diversity. In line with the identification of mutations in *MYCN*, *PTPN14* and *LATS1*, RNA sequencing (RNA-seq) analysis showed that Hippo-YAP pathway and N-Myc target genes are significantly upregulated in BCC.

[1]Department of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland. [2]Department of Dermatology, Hospital of Valais, Sierre, Switzerland. [3]Department of Pathology, New York University School of Medicine, New York, New York, USA. [4]Department of Physics, University of Connecticut, Storrs, Connecticut, USA. [5]RIKEN BNL Research Center, Brookhaven National Laboratory, Upton, New York, USA. [6]Department of Dermatology, University Hospitals of Geneva, Geneva, Switzerland. [7]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [8]Institute of Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia. [9]Pirogov Russian National Research Medical University, Moscow, Russia. [10]Lomonosov Moscow State University, Moscow, Russia. [11]Department of Molecular Oncology, Genentech, Inc., South San Francisco, California, USA. [12]Service of Clinical Pathology, University Hospitals of Geneva, Geneva, Switzerland. [13]Department of Dermatology, Saint Louis Hospital, Paris 7 University, Paris, France. [14]Service of Genetic Medicine, University Hospitals of Geneva, Geneva, Switzerland. [15]Department of Biochemistry and Molecular Biology, University of Würzburg, Würzburg, Germany. [16]Comprehensive Cancer Center Mainfranken, University of Würzburg, Würzburg, Germany. [17]Department of Oncology, University of Lausanne and Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland. [18]Institute of Genetics and Genomics of Geneva (iGE3), Geneva, Switzerland. [19]These authors contributed equally to this work. Correspondence should be addressed to S.I.N. (sergey.nikolaev@unige.ch) or S.E.A. (stylianos.antonarakis@unige.ch).

Received 14 September 2015; accepted 11 February 2016; published online 7 March 2016; doi:10.1038/ng.3525

RESULTS

We performed mutational screening of 293 BCCs from 236 patients; 225 BCCs were collected for this study, and sequencing data from 68 previously reported cases were added. The sample set consisted of 263 sporadic cases and 30 tumors from patients with Gorlin syndrome. The collection also comprised 23 vismodegib-resistant tumors, 11 vismodegib-sensitive tumors and 259 tumors naive to vismodegib treatment.

We analyzed whole-exome sequencing data from 126 BCCs and their matched blood samples (68 sequenced for this study and 58 previously reported; **Supplementary Table 1a**). For 163 sporadic and 4 Gorlin syndrome BCCs, we performed analysis of a targeted panel of 387 cancer-related genes (referred to hereafter as the cancer panel; **Supplementary Tables 1a** and **2**). We performed RNA-seq analysis of 61 BCCs and 25 individual-matched samples of normal epidermis (**Supplementary Table 1b**). Somatic copy number aberrations (SCNAs) were identified from exome sequencing data (**Supplementary Note**).

A total of 234,770 somatic coding mutations were identified (**Supplementary Table 3**). Consistent with previous observations[5,13], the average mutation rate for sporadic BCCs was 65 mutations/Mb, the highest mutation rate observed among cancers studied thus far[19], and the rate was 21 mutations/Mb for tumors from patients with Gorlin syndrome (**Fig. 1**, top).

Mutational signatures of BCC

The mutational profile of BCC was consistent with UV-induced mutagenesis: on average, 90% of single-nucleotide variants (SNVs) were C>T changes occurring in a YC or CY context (where Y = C or T), as previously reported[5,6,13,20] (**Fig. 1**, bottom), and 5% of all mutations were dinucleotide substitutions. To better understand the mutational mechanisms in BCC, we compared its mutational profile with that of melanoma[21,22]. We specifically investigated the prevalence of UV- and oxidative stress–induced mutagenesis and the effectiveness of transcription-coupled repair (TCR) in BCC.

The contribution of UV signature mutations was higher in BCC than in melanoma (90% versus 85% of mutations; $P < 2.2 \times 10^{-16}$, χ^2 test). We estimated the fraction of mutations with highly specific UV signatures[23–25] of the TCC>TTC and CC>TT type for all C>D and CC>DD mutations (where D = A, T or G). In BCC, these types of mutations were 2% (C>D, $P < 2.2 \times 10^{-16}$, χ^2 test) and 15% (CC>DD, $P < 2.2 \times 10^{-16}$, χ^2 test) more frequent than in melanoma.

We then estimated the efficiency of TCR through analysis of the ratio of UV-induced mutations on the transcribed and non-transcribed strands[26]. For TCC>TTC mutations, this ratio was lower in BCC than in melanoma (0.59 versus 0.71; $P < 2.2 \times 10^{-16}$, χ^2 test), and a similar pattern was observed for the genes expressed in both melanoma and BCC (0.56 versus 0.68; $P < 2.2 \times 10^{-16}$, χ^2 test), suggesting that TCR is more effective in BCC than in melanoma. Moreover, the effect of TCR was even stronger in BCC when only highly expressed genes were considered (0.21 versus 0.29; $P < 2.2 \times 10^{-16}$, χ^2 test) (**Supplementary Fig. 1a**).

We next investigated C>A mutations, which may occur as a result of either oxidative stress[27] or UV exposure[28]. We selected the subset

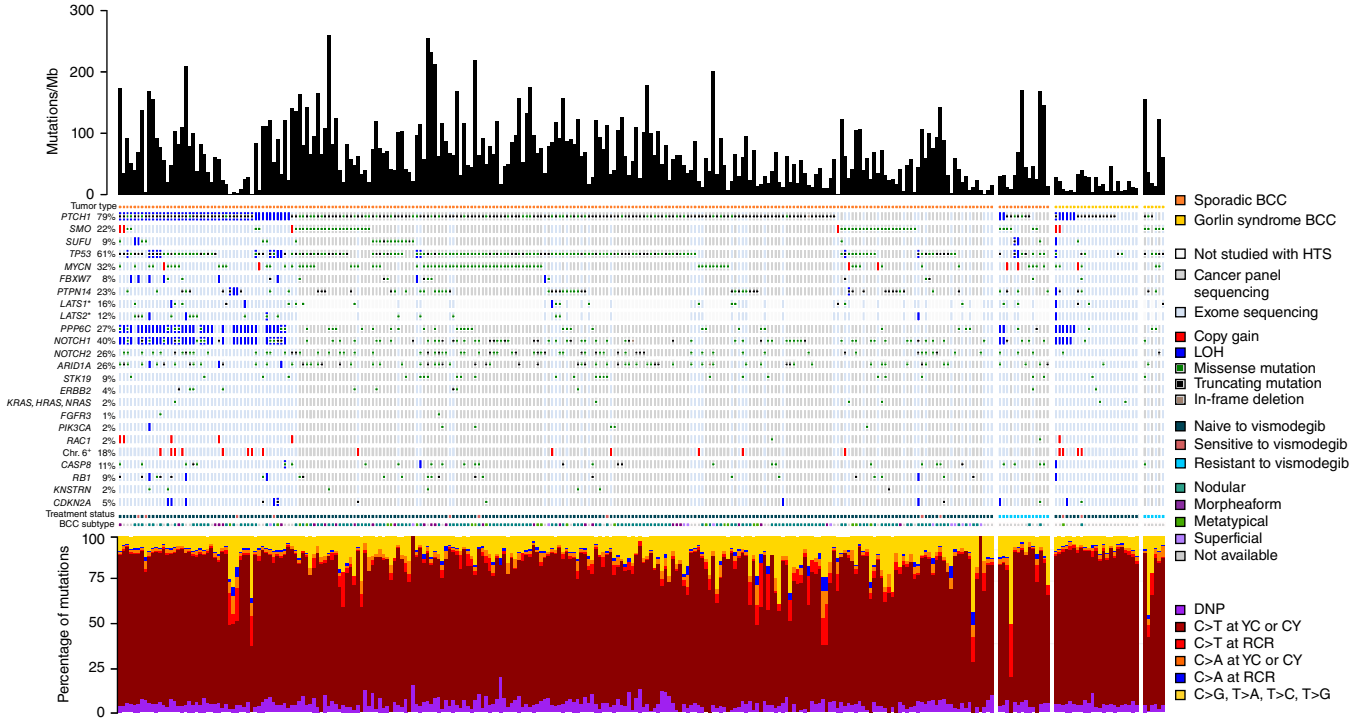

Figure 1 Mutational landscape of BCC. Data are shown for exome- and cancer panel–sequenced samples. Each column corresponds to one sample (283 non-clonal samples are displayed). Top, number of somatic mutations per megabase. Middle (top), tumors classification as sporadic BCC or a tumor from a patient with Gorlin syndrome. Middle (center), distribution of driver somatic mutations and SCNAs. SCNAs were only estimated for exome-sequenced samples (light blue background). When multiple mutations per gene were found in a sample, only the most damaging mutation is shown. The percentage of mutations in the data set for each of the genes is shown on the left. *, *LATS1* and *LATS2* were not included in the cancer panel (white background). *LATS1* mutations encoding p.Arg995Cys were screened with Sanger sequencing in cancer panel–sequenced samples; the identified mutations correspond to green squares over a white background in the *LATS1* track. +, chromosome 6 percentage corresponds to exome-sequenced samples only and not to the total number of samples displayed, as cancer panel–sequenced samples could not be evaluated for SCNAs. This part of the figure was generated with OncoPrinter in cBioPortal tools. HTS, high-throughput sequencing. Middle (bottom), tumor status with respect to treatment and histological subtype (nodular, morpheaform, metatypical or superficial), when available. Bottom, mutational spectrum of the displayed samples. DNP, dinucleotide polymorphism.

of C>A mutations in the RCR context (where R = A or G) that are only mildly affected by UV light because cytosines in this context are not involved in the formation of photodimers when exposed to UV light[29]. In this context, C>A changes may be enriched as a result of guanine oxidation. As expected, C>A mutations were enriched on the transcribed strand in both BCC and melanoma (1.26 and 1.27 times as compared to the non-transcribed strand, respectively; in both tumor types, $P < 0.001$, χ^2 test), suggesting that mutations indeed occurred at guanine bases, which were then repaired by TCR on the transcribed strand (**Supplementary Table 4a**).

We subsequently selected a subset of C>A mutations at CC sites (where the underlined residue is mutated), the context of UV-induced mutagenesis where these mutations are mostly enriched (**Supplementary Fig. 1b**). Indeed, in cutaneous T cell lymphomas, for which UV signature mutations were previously reported[30,31], C>A mutations specifically occurred in a CC context[31]. We observed no statistically significant bias between the transcribed and non-transcribed strands for C>A mutations in BCC (1.09; $P = 0.10$, χ^2 test), whereas in melanoma there was a significant excess of C>A mutations on the non-transcribed strand (0.89; $P < 0.001$, χ^2 test; **Supplementary Table 4a**). Taking into account the effect of TCR, it is possible that C>A mutations in the CC context in BCC represent a mixture of mutations at cytosine and guanine bases, whereas in melanoma cytosine mutations are more frequent.

Moreover, the fraction of C>A mutations in the RCR context (depleted of UV-induced mutations) was twice as high in BCC than in melanoma (21% in BCC versus 10% in melanoma, as compared to an expected fraction of 19%). Furthermore, the fraction of C>A mutations in the CC context (enriched in UV-induced mutations) was 1.6 times higher in melanoma than in BCC (53% in melanoma versus 35% in BCC, as compared to an expected fraction of 28%; **Supplementary Table 4b**).

These data suggest that the mechanism of C>A mutation is different for BCC and melanoma. In BCC, these mutations predominantly occur as a result of guanine oxidation, whereas in melanoma the contributions from UV-induced mutagenesis of cytosines and oxidation of guanines are similar.

BCC drivers

In concordance with previous studies[13], 85% of BCCs harbored somatic mutations in Hh pathway genes—*PTCH1* (73%), *SMO* (20%) and *SUFU* (8%)—and 61% harbored mutations in *TP53*. MutSigCV analysis of our cohort identified seven significantly mutated genes (SMGs) with q values <0.1 (**Supplementary Table 5a–c**). The known drivers of BCC, *PTCH1*, *TP53* and *SMO*, but not *SUFU*, were significant ($q = 0$, 2.1×10^{-14}, 4×10^{-2} and 1, respectively). In addition to the known BCC-associated genes, the SMGs included *PTPN14* ($q = 2 \times 10^{-6}$), *MYCN* ($q = 1.4 \times 10^{-4}$), *RPL22* ($q = 1.1 \times 10^{-2}$) and *PPIAL4G* ($q = 3.5 \times 10^{-2}$).

To identify significantly recurrent mutations (SRMs), we developed an algorithm that uses the background mutation rate per nucleotide (**Supplementary Note**). This method yielded 28 recurrent mutations in 14 genes ($q < 0.05$) (**Supplementary Table 6a**,**b**). Among these genes, we found one SRM site in *LATS1* (p.Arg995Cys, $q = 1.18 \times 10^{-3}$), two SRM sites in *SMO* (p.Trp535Leu, $q = 6.4 \times 10^{-7}$; p.Ala459Val, $q = 6.4 \times 10^{-9}$), three SRM sites in *MYCN* (p.Pro44Leu, p.Pro44Ser or p.Pro44Phe, $q = 2.4 \times 10^{-3}$; p.Pro59Leu, $q = 2.5 \times 10^{-2}$; p.Pro60Leu, $q = 5.3 \times 10^{-2}$), one SRM site in *STK19* (p.Asp89Asn, $q = 8.4 \times 10^{-7}$), five SRM sites in *TP53* (p.Glu286Lys, $q = 2 \times 10^{-2}$; p.Pro278Leu, $q = 1.8 \times 10^{-3}$; p.Arg248Gln or p.Arg248Trp, $q = 8.6 \times 10^{-4}$; p.Pro179Tyr, $q = 2.1 \times 10^{-5}$; p.Pro177Leu, $q = 2.2 \times 10^{-3}$), one SRM site in *ERBB2* (p.Ser310Phe, $q = 4.4 \times 20^{-3}$) and one SRM site in *PPP6C* (p.Arg264Cys, $q = 3.5 \times 10^{-5}$) (**Supplementary Fig. 2**).

Additionally, we identified known driver variants reported in public cancer databases (Catalogue of Somatic Mutations in Cancer (COSMIC) and The Cancer Genome Atlas (TCGA)) in our BCCs. We found mutations in *NRAS*, *HRAS* or *KRAS*, *PIK3CA* and *RAC1*, as well as deleterious and truncating mutations in *FBXW7*, *RB1*, *CDKN2A*, *ARID1A*, *NOTCH1*, *CASP8* and *NOTCH2* (**Fig. 1**, middle). Although these variants were not statistically significant in our analysis, they are potentially relevant drivers in BCC.

We next investigated the mutational profiles of the identified cancer-related genes in BCC. Truncating mutations were frequently observed in *PTCH1* (70%), *TP53* (36%) and *PTPN14* (61%), consistent with their roles as tumor-suppressor genes (**Fig. 1**, middle, and **Supplementary Fig. 2l**,**m**). Moreover, 32%, 48% and 19% of tumors harbored more than one tumorigenic event in these genes, respectively; these fractions are highly unlikely in a random distribution ($P < 0.02$; **Supplementary Note**). The most commonly observed SCNAs were the loss of one allele of *PTCH1* on chromosome 9q ($q = 5.80 \times 10^{-6}$) and of 17p ($q = 0.042$; overlapping *TP53*) (**Supplementary Fig. 3**). Pathogenic *SMO* mutations[5] were found in 20% of tumors (**Supplementary Fig. 2k**), and mutations in *SUFU* were found in 8% of tumors (**Supplementary Fig. 2n**). Interestingly, *TP53* displayed a high prevalence of recurrent mutations[32], with five SRMs identified in 8% of tumors (**Supplementary Fig. 2l** and **Supplementary Table 6**).

We subsequently investigated the interaction between driver mutations and loss of heterozygosity (LOH) in the *PTCH1* and *TP53* genes. We selected 23 cases where the driver mutations in these genes overlapped copy-neutral LOH events and calculated the variant allele frequencies (VAFs) of all mutations and driver mutations in these regions occurring before and after LOH (present in one or two alleles in tumor cells). We detected 585 mutations occurring before and 1,392 mutations occurring after the copy-neutral LOH (**Fig. 2**). However, of 23 driver mutations, 18 arose before the LOH (81% in *PTCH1* and 75% in *TP53*), indicating that the copy-neutral LOH occurred predominantly ($P = 0.002$, binomial test) after the driver mutation. This analysis confirmed that copy-neutral LOH events overlapping mutant tumor-suppressor genes remove the remaining wild-type allele. Moreover, we observed that the tumor suppressors *PTCH1* and *TP53* exhibit two mutually exclusive mechanisms of eliminating the remaining wild-type allele: a deleterious mutation of the second allele or loss of the wild-type allele ($P = 0.035$, one-sided Fisher's exact test).

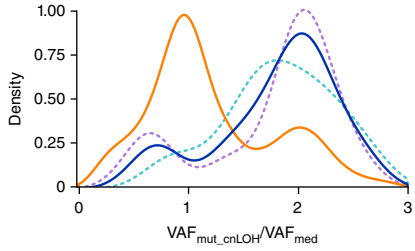

Figure 2 Mutations in copy-neutral LOH regions in BCC. The distribution of mutations occurring before or after copy-neutral LOH is shown. VAF$_{mut_cnLOH}$ is the VAF of mutations inside copy-neutral LOH regions, and VAF$_{med}$ is the median VAF of all somatic mutations. The VAF$_{mut_cnLOH}$/VAF$_{med}$ ratio represents VAF$_{mut_cnLOH}$ values normalized to the fraction of tumor cells in each sample. The orange line represents all mutations in regions with copy-neutral LOH, where most of the substitutions affect a single allele. The dark blue line represents driver mutations in *PTCH1* and *TP53*, where most mutations affect both alleles. The light blue dashed line represents only *TP53* mutations, and the violet dashed line corresponds to *PTCH1* mutations.

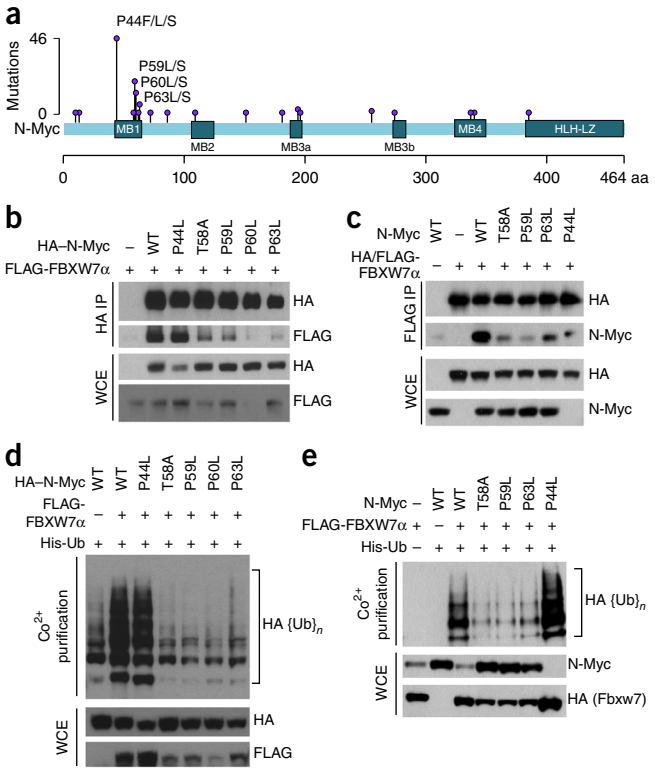

Figure 3 Functional characterization of N-Myc alterations in BCC. (**a**) Missense alterations in N-Myc cluster in the MB1 domain. Purple lollipops represent missense alterations. Amino acid substitutions occurring five or more times are labeled. The protein diagram was generated with cBioPortal tools for 104 mutations from 87 samples with *MYCN* mutations. Protein functional domains are represented by dark blue boxes. (**b**) Coimmunoprecipitation of wild-type (WT) or mutant HA-tagged N-Myc with FLAG-tagged FBXW7 from transient transfections in HEK293 cells. Immunoblots for HA and FLAG are shown following HA immunoprecipitation (IP) and for whole-cell extract (WCE). (**c**) Coimmunoprecipitation of wild-type or mutant N-Myc with FLAG/HA-tagged FBXW7 from transient transfections in HEK293 cells. Immunoblots for N-Myc and HA are shown following FLAG immunoprecipitation and for whole-cell extract. (**d**,**e**) *In vivo* ubiquitination assays to measure polyubiquitination of wild-type and mutant N-Myc in the presence of FBXW7. His-tagged ubiquitin (Ub) was coexpressed with FBXW7 and wild-type or mutant N-Myc in HEK293 cells. Ubiquitinated proteins were enriched on a Co^{2+} resin column under denaturing conditions, and the amount of polyubiquitinated N-Myc was measured by immunoblot analysis for HA. Autoubiquitinated FBXW7 is shown in **e**. For **b**–**e**, cells were treated with 10 μM MG-132 for 4 h before cells were collected to inhibit the proteasome. Data shown are representative of three (**b**,**c**) or two (**d**,**e**) independent experiments.

MYCN and *FBXW7*

Missense mutations in *MYCN*, an oncogene paralogous to *MYC* (43% sequence homology at the protein level[33]), were found in 30% of BCCs. Most mutations were clustered in the sequence encoding the Myc box 1 (MB1) region (**Fig. 3a**). Interestingly, a similar or overlapping set of MB1 alterations was previously identified in *MYC* in Burkitt lymphoma[34] and was shown to impair interaction with FBXW7, the substrate-binding component of the SCF^{Fbxw7} E3 ubiquitin ligase complex that ubiquitinates Myc and targets it for proteasome degradation. Thus, mutations mapping to this region of Myc lead to its stabilization and accumulation at the protein level[35].

In line with previous observations, copy-gain SCNAs overlapping the *MYCN* locus were detected in 12% of BCCs[36] (**Supplementary Fig. 3**). The recent report of a BCC susceptibility locus at 2q24 in the vicinity of the *MYCN* region[37] further supports the implication of *MYCN* in BCC.

We investigated how the mutations detected in *MYCN* would influence the protein structure and function of N-Myc. A substantial proportion of the N-Myc substitutions, such as p.Pro44Leu, p.Pro59Leu, p.Pro60Leu, p.Ser62Leu and p.Pro63Leu, mapped within the highly conserved Cdc4 phosphodegron (CPD), which is critical for interaction with FBXW7 upon phosphorylation[38].

Phosphorylation of Thr58 by GSK-3β was predicted to be impaired by the p.Pro59Leu and p.Pro60Leu substitutions using KinasePhos2 (ref. 39). Similarly, p.Ser62Leu and p.Pro63Leu are likely to interfere with phosphorylation at the priming +4 position of the GSK-3β target sequence. Modeling of N-Myc Pro44Leu yielded less clear results[40], suggesting that it may function through a mechanism independent of GSK-3.

We next tested whether the alterations identified in N-Myc would affect affinity for FBXW7 *in vivo*. Wild-type and mutant forms of N-Myc were expressed together with FBXW7, and interaction was assessed by coimmunoprecipitation. Whereas wild-type N-Myc showed a robust interaction with FBXW7, four of the N-Myc substitutions (p.Thr58Ala, p.Pro59Leu, p.Pro60Leu and p.Pro63Leu) impaired binding to FBXW7 (**Fig. 3b**,**c**). Accordingly, the mutants displayed reduced levels of polyubiquitination in the presence of FBXW7 (**Fig. 3d**,**e**). The Pro44Leu mutant was consistently expressed at lower levels but still interacted with FBXW7. In the presence of N-Myc Pro44Leu, we also detected enhanced autoubiquitination of FBXW7 (**Fig. 3e**), further suggesting a unique functional mechanism for the effects of N-Myc Pro44Leu.

The identification of deleterious mutations in *FBXW7* in 5% of BCCs also suggests selective pressure for enhanced N-Myc stability in BCC. Most of these mutations mapped upstream of the WD40 domain (**Supplementary Fig. 2a**), which is critical for substrate interaction[41,42]. We additionally observed LOH events overlapping *FBXW7* in 8% of samples (**Fig. 1**).

PTPN14

The mutations of *PTPN14* in this study suggest that this gene has a tumor-suppressor role in BCC, as *PTPN14* was mutated in 23% of BCCs and 61% of the mutations were truncating (**Figs. 1** and **4a**). It has been shown that PTPN14 promotes the nucleus-to-cytoplasm translocation of YAP1, the key effector of the Hippo pathway, preventing YAP1-mediated transcriptional activation[40]. Furthermore, YAP1 nuclear localization has been reported to occur in a fraction of BCCs[43]. To determine whether *PTPN14* mutations impair the cytoplasmic localization of YAP1, we performed immunostaining for YAP1 of eight BCCs with *PTPN14* loss-of-function mutations and three BCCs with wild-type *PTPN14*. We observed that BCCs with wild-type *PTPN14* presented with diffuse, cytoplasm-specific staining and almost no nuclear accumulation of YAP1, whereas tumors with mutant *PTPN14* accumulated YAP1 in the nucleus (**Fig. 4b**). Nuclear staining of YAP1 was observed in specific regions of the tumor (on average, in 26% of the tumor cells; **Fig. 4c**). Immunostaining for MIB1 (a cellular marker for proliferation) in *PTPN14*-mutated tumors showed considerable variation in the fraction of positive nuclei, and colocalization with YAP1 was only observed in two of the eight cases studied (data not shown). The lack of a consistent relationship between MIB1 staining and nuclear YAP1 is not surprising given the genetic complexity of these tumors and the presence of additional potential drivers. These results confirm that the nuclear localization of YAP1 may be associated with *PTPN14* loss-of-function mutations,

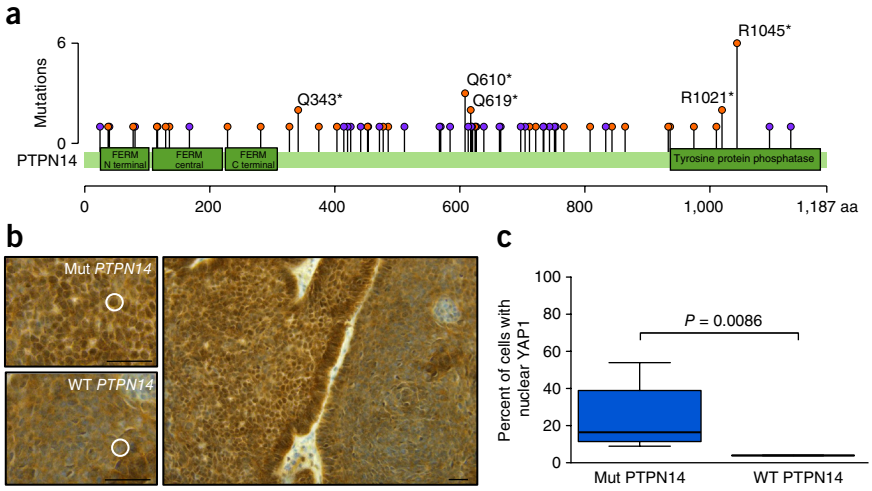

Figure 4 *PTPN14* mutations in BCC. (**a**) Distribution of alterations along the PTPN14 protein. The orange lollipops represent truncating alterations, and the purple lollipops represent missense alterations. Events occurring two or more times are labeled with the amino acid change. The protein diagram was generated with cBioPortal tools for 75 mutations from 63 samples with *PTPN14* mutations. Protein functional domains are represented by green boxes. (**b**) Immunohistochemistry for YAP1 in a *PTPN14*-mutated tumor. The panel on the top left shows a region of the tumor with mutant (Mut) *PTPN14*; the white circle highlights nuclear accumulation of YAP1. The panel on the bottom left shows a region of the tumor with wild-type *PTPN14*; the white circle highlights a nucleus depleted of YAP1. The panel on the right shows a *PTPN14*-mutated tumor with two adjacent tumoral nests, one with nuclear staining for YAP1 (left) and one with cytoplasmic localization of YAP1 (right). Scale bars, 100 µm. (**c**) Fraction of cells with YAP1 staining in BCCs with mutant (eight tumors) or wild-type (three tumors) *PTPN14* ($P = 0.0086$). Boxes represent the median (black middle line) limited by the 25th and 75th percentiles. Whiskers are the upper and lower adjacent values.

implicating *PTPN14* as a new tumor-suppressor gene in BCC that functions through activation of the Hippo-YAP pathway. Hippo-YAP activation in BCC is also consistent with the recent observation that Yap1 promotes proliferation in normal skin[43] and in cutaneous squamous cell carcinoma (SCC) in mice[44].

LATS1 and *LATS2*

We identified the *LATS1* mutation encoding p.Arg995Cys as being significant in the SRM analysis (**Supplementary Table 6**). Moreover, *LATS1* harbored deleterious mutations in 16% of the exome-sequenced tumors, with 24% of these being truncating mutations (**Fig. 5a**). LATS1 is a known tumor-suppressor kinase that can be activated by PTPN14 (ref. 45) and that also phosphorylates YAP1, preventing its translocation to the nucleus[46]. We assessed the effects of the LATS1 p.Arg995Cys substitution with *in silico* protein modeling. This residue was shown to be deeply buried in the core of the kinase domain, in the center of an important hydrogen-bond network (**Fig. 5b**). These analyses also predicted that Arg995 is important for structural stability, and its alteration is expected to have a severe impact on protein structure and activity. *LATS1*-inactivating mutations may represent an independent mechanism for activation of the Hippo-YAP pathway in BCC. Interestingly, a number of BCCs carried missense mutations in *LATS2* (12% of exome-sequenced tumors), a paralog of *LATS1*. The mutations clustered in the sequence encoding the C terminus, with ~10% of them affecting Pro1004 (**Supplementary Fig. 4a**). *In silico* modeling of the LATS2 protein showed that Pro1004 could have a role in substrate recognition and specificity (**Supplementary Fig. 4b**). A functional role for *LATS1* in BCC pathogenesis is further supported by the report of a patient with Gorlin syndrome harboring biallelic inactivation of *LATS1* in an infiltrative BCC but not in a superficial tumor[47].

Other putative drivers

We identified somatic driver SRMs in two new candidate BCC-related genes, *PPP6C* and *STK19*. We observed mutations in the serine/threonine protein phosphatase gene *PPP6C* in 15% of BCCs (**Supplementary Fig. 2b**); similar mutations were detected in 12% of melanomas[20,48] and impaired the phosphatase activity of the encoded protein[20,49,50]. PPP6C specifically inhibits the expression of cyclin D1, inactivates RB1 (ref. 51) and participates in LATS1 activation[52]. Ten percent of BCCs harbored mutations in the serine/threonine kinase gene *STK19*, all of which resulted in a p.Asp89Asn substitution (**Supplementary Fig. 2c**). This substitution has previously been described in 5% of melanomas[20].

Several known cancer-related genes were also recurrently mutated in BCC with low frequency[19]. Eleven percent of our tumors harbored mutations in *CASP8* (14% nonsense mutations), a gene potentially involved in BCC[37] (**Supplementary Fig. 2d**). *RB1* alterations were observed in 10% of the tumors studied (44% LOH or loss-of-function mutation)

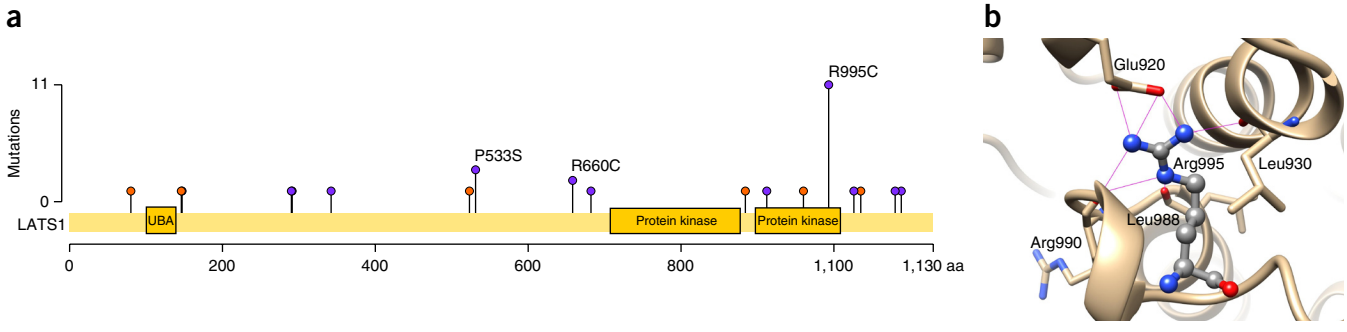

Figure 5 *LATS1* mutations in BCC. (**a**) Distribution of alterations along the LATS1 protein. The orange lollipops represent truncating alterations, and the purple lollipops represent missense alterations. Protein functional domains are represented by yellow boxes. Events occurring two or more times are labeled with the amino acid change. The protein diagram was generated with cBioPortal tools for 31 mutations from 24 samples with *LATS1* mutations. (**b**) Kinase domain structure of LATS1 highlighting the position and interactions of residue Arg995.

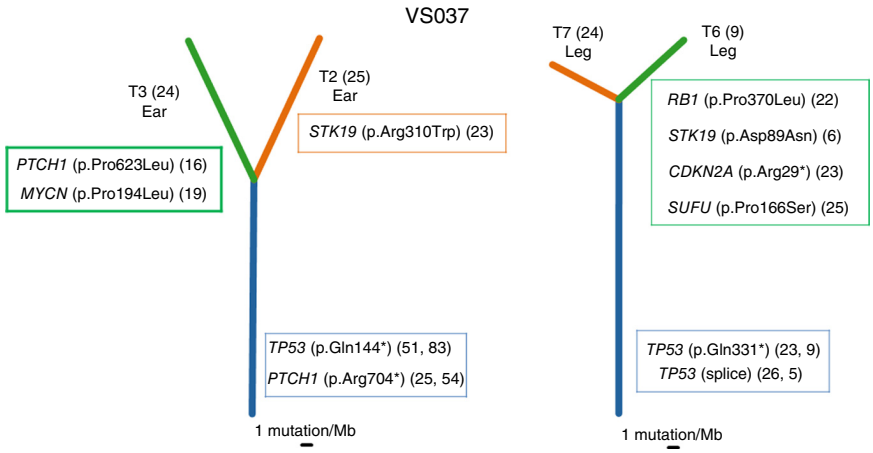

Figure 6 Clonality in sporadic BCCs. Common and tumor-specific mutations are shown for two pairs of tumors. The trunk of each tree corresponds to shared mutations (blue), and the branches (orange and green) correspond to mutations specific for each tumor. The length of the lines is proportional to the number of mutations per megabase. Protein changes are shown, and VAFs appear in parentheses. For common mutations, the first VAF corresponds to the left tumor and the second VAF corresponds to the right tumor. Patient IDs are at the top of the trees, and tumor IDs are marked for each of the branches. Each tumor's anatomical location is noted under the tumor ID.

(**Supplementary Fig. 2e**). We also identified recurrent mutations in *KNSTRN* (p.Arg11Lys and p.Ser24Phe) in 2% of samples[53,54] (**Supplementary Fig. 2f**). We observed oncogenic mutations in genes of the mitogen-activated protein kinase (MAPK) pathway in BCC. A p.Ser310Phe substitution encoded in *ERBB2* was observed in 4% of samples (**Supplementary Fig. 2g**), and we additionally observed gain-of-function mutations (p.Gly12Cys, p.Gly12Asp, p.Gly13Arg, p.Gly13Val, p.Gln22Lys and p.Gln61Arg) in the *KRAS*, *NRAS* and *HRAS* genes as well as in *PIK3CA* (p.Val344Lys, p.Asn345Lys, p.Glu542Lys and p.Glu726Lys), with these mutations cumulatively occurring in 4% of BCCs. One percent of our BCCs harbored p.Pro29Ser substitutions encoded in *RAC1*, which were previously reported in 9% of melanomas[20,48].

We identified mutations in Notch family genes, which suggests that these genes have a tumor-suppressor role in BCC. Twenty-six percent of BCCs had *NOTCH2* mutations, with 30% of these being loss of function and 22% of the affected tumors having paired loss-of-function mutations (**Supplementary Fig. 2h**). Twenty-nine percent of BCCs had *NOTCH1* mutations, where 25% of these mutations were loss of function (**Supplementary Fig. 2i**). Additionally, of the eight *NOTCH1* point mutations overlapping copy-neutral LOH events, six occurred before the LOH; therefore, in these cases, no wild-type allele remained in the tumor.

ARID1A harbored deleterious mutations in 26% of BCCs. Thirty-two percent of these were truncating and 25% of the tumors affected presented two mutations, a profile consistent with a tumor-suppressor gene (**Supplementary Fig. 2j**).

Finally, we observed chromosome 6 copy gain in 18% of the exome-sequenced BCCs ($q = 6.15 \times 10^{-6}$, GISTIC), 5% of which were limited to a 64-Mb region that was previously associated to tumors with activated Hh signaling and includes the *CCND3* and *E2F3* genes[55].

Analysis of gene mutual exclusivity

As expected, mutations in the Hh pathway tumor suppressors *PTCH1* and *SUFU* were unlikely to occur in tumors driven by mutations in *SMO* ($P = 5.5 \times 10^{-8}$, Fisher's exact test). Furthermore, mutations in *PTPN14* and *MYCN* tended to be mutually exclusive ($P = 0.06$, Fisher's exact test) (**Fig. 1**, middle).

Tumor clonality and germline variants

It is not uncommon for patients to develop several BCC tumors. Our BCC cohort included 38 individuals with more than one tumor. One patient had two pairs of tumors in anatomically close sites (patient VS037; distances of 0.4 and 0.5 cm between the tumors) that were

a

Pathway (KEGG pathways)	*P* value (Benjamini corrected)
Cell cycle	1.5×10^{-8}
Pathways in cancer	4.7×10^{-8}
ECM-receptor interaction	1.5×10^{-5}
Basal cell carcinoma	6.5×10^{-5}
Focal adhesion	3.6×10^{-4}
Hedgehog signaling pathway	1.2×10^{-3}
p53 signaling pathway	1.1×10^{-2}
DNA replication	1.2×10^{-2}
TGF-β signaling pathway	4.3×10^{-2}
Melanomagenesis	6.0×10^{-2}

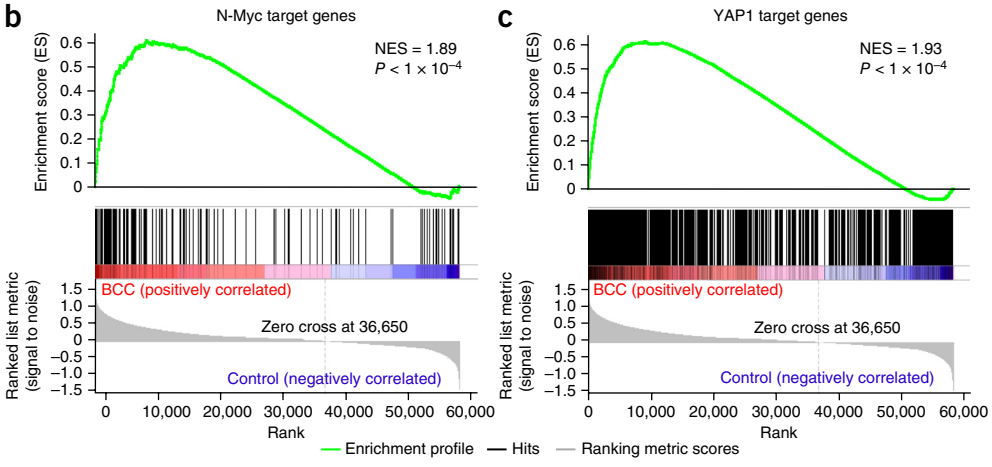

Figure 7 The expression of N-Myc and YAP1 target genes is upregulated in BCC. (**a**) KEGG signaling pathways significantly over-represented in the list of genes with upregulated expression in BCC (DAVID analysis). (**b**,**c**) Enrichment of N-Myc (**b**) and YAP1 (**c**) targets among the genes with upregulated expression in BCC when compared to normal skin by GSEA. Genes were sorted according to their fold change in expression between BCC and normal adjacent skin samples (*x* axis; 0 represents the most upregulated gene). Black vertical bars show the positions of the target genes in the ranked list. The enrichment scores (green lines; 0.6 in **b** and 0.59 in **c**) significantly deviate from 0 at the beginning of the distribution, showing that target genes are not randomly distributed in the ranked list but are enriched among the upregulated genes. *P* values are FDR-corrected *P* values.

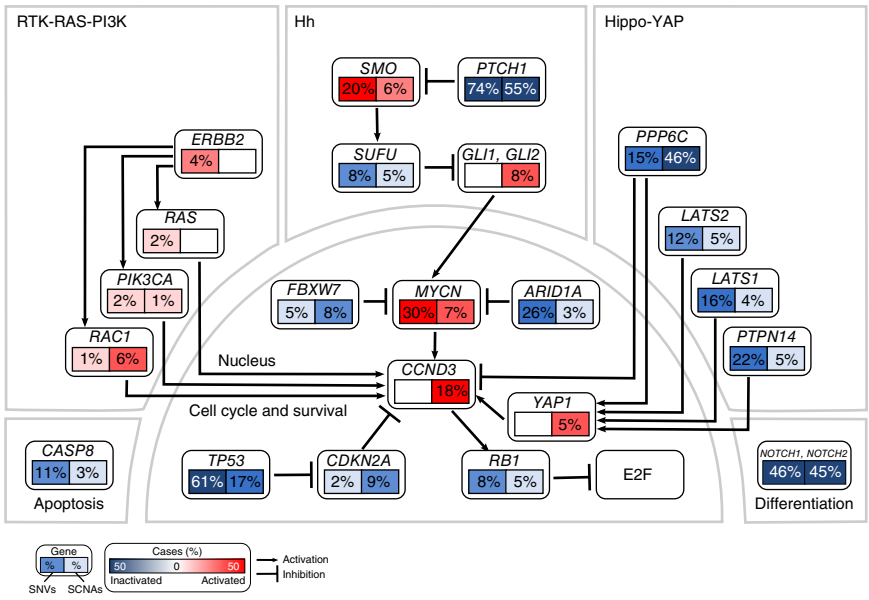

Figure 8 Signaling pathways involved in BCC. Key affected pathways and genes. The frequency of somatic point mutations and SCNAs for each gene is shown in the left and right panel, respectively, under each gene name. Genes inactivated in BCC are shown in blue, and genes activated in BCC are shown in red. The intensity of the color corresponds to the fraction of samples affected. RTK, receptor tyrosine kinase.

clonal, sharing a considerable fraction (14% for the first pair and 49% for the second pair) of passenger mutations and at least one driver mutation. Clonal *TP53* loss-of-function mutations were present in both cases, and clonal *PTCH1* mutations were observed in one of the pairs (**Fig. 6**). Conversely, none of the BCCs located in anatomically distant sites were clonal.

We identified germline variants in *PTCH1* in 19 of 20 patients with Gorlin syndrome. Ninety percent of these variants were truncating. In sporadic BCCs, we identified three missense germline *PTCH1* variants of unknown clinical relevance, as well as three missense germline variants in *TP53*. One of these (encoding p.Thr377Pro), reported in COSMIC as a somatic mutation, was detected in an individual with multiple BCCs, suggesting that it may confer predisposition.

Histological subtypes and drug resistance

We investigated the association of driver mutations with specific BCC phenotypes (**Supplementary Fig. 5**). In a subset of BCCs with high risk of recurrence (non-clonal metatypical and morpheaform BCCs), the profile of new driver mutations was different from that of BCCs with low risk of recurrence (non-clonal nodular and superficial BCCs) (**Supplementary Table 1a**). Specifically, in the group of BCCs with increased risk of recurrence, *MYCN*, *PPP6C* and *PTPN14* mutations were 1.8 ($P = 3.4 \times 10^{-3}$), 2.7 ($P = 2.3 \times 10^{-3}$) and 1.8 ($P = 5.3 \times 10^{-2}$) times more frequent than in BCCs with low recurrence risk, respectively. Moreover, for the most recurrent p.Pro44Ser and p.Pro44Leu substitutions encoded in *MYCN*, the difference increased to 2.1-fold ($P = 2.2 \times 10^{-2}$, Fisher's exact test). These correlations suggest a link between oncogenic mutations in *MYCN*, *PPP6C* and *PTPN14* and histological subtype.

Vismodegib-resistant BCCs had at least a 2.2-fold higher frequency of *SMO* mutations than treatment-naive BCCs in both sporadic (43% versus 20%, $P = 5 \times 10^{-2}$) and Gorlin syndrome (83% versus 0%, $P = 7.3 \times 10^{-5}$, Fisher's exact test) cases. Additionally, *PTPN14* mutations were more prevalent in BCCs driven by *SMO* mutations than in those driven by *PTCH1* mutations (37% versus 22%, $P = 6.18 \times 10^{-2}$; *PTPN14* truncating, 27% versus 11%, $P = 2.11 \times 10^{-2}$, Fisher's exact test) (**Supplementary Fig. 5**).

In Gorlin syndrome tumors, driver mutations occurred in the same new genes as in sporadic BCCs, including *MYCN*, *PTPN14*, *LATS1*, *KRAS*, *NRAS*, *HRAS*, *ERBB2*, *STK19* and *PPP6C* (**Fig. 1**); however, some differences were observed. In the 19 treatment-naive patients with Gorlin syndrome harboring germline *PTCH1* mutations, we found no somatic *SMO* mutations; this fraction is significantly less than in treatment-naive sporadic BCCs ($P = 0.025$, Fisher's exact test). These data suggest that tumorigenic pathways may vary among the BCC subtypes, between tumors resistant and naive to vismodegib treatment, and between Gorlin syndrome and sporadic BCC tumors.

Expression profiling

We performed RNA-seq analysis of BCCs and unaffected skin samples (Online Methods). Pathway analysis identified several pathways with highly significant upregulation, including 'basal cell carcinoma' ($P = 6.5 \times 10^{-5}$), 'Hedgehog signaling pathway' ($P = 1.2 \times 10^{-3}$) and 'p53 signaling pathway' ($P = 1.1 \times 10^{-2}$; **Fig. 7a**). A gene set enrichment analysis (GSEA) for N-Myc and YAP1 target genes[56,57] showed that genes whose expression is upregulated in BCC are significantly enriched in these gene sets (normalized enrichment score (NES) = 1.89, false discovery rate (FDR) $< 1 \times 10^{-4}$ and NES = 1.93, FDR $< 1 \times 10^{-4}$, respectively) (**Fig. 7b**,**c**). Most notably, the subsets of tumors with *MYCN* mutations or amplifications and with *PTPN14* mutations showed stronger enrichment for N-Myc and YAP1 target genes (NES = 2.11, FDR $< 1 \times 10^{-4}$ and NES = 2.30, FDR $< 1 \times 10^{-4}$, respectively). These observations are consistent with the expected effects of N-Myc stabilization and amplification and of YAP1 nuclear accumulation (caused by *PTPN14* mutations) on the BCC transcriptome. Moreover, subsets of sporadic and Gorlin syndrome vismodegib-resistant tumors showed higher enrichment for N-Myc (NES = 2.28, FDR $< 1 \times 10^{-4}$; NES = 2.18, FDR $< 1 \times 10^{-4}$) and YAP1 (NES = 2.43, FDR $< 1 \times 10^{-4}$; NES = 2.60, FDR $< 1 \times 10^{-4}$) target genes than all samples and than the treatment-naive samples (NES = 1.78, FDR $< 1 \times 10^{-4}$; NES = 1.77, FDR $< 1 \times 10^{-4}$, respectively).

DISCUSSION

The identification of new driver genes in BCC suggests a more complex genetic network of cancer-related genes (**Fig. 8**) than previously anticipated. Hh pathway mutations downstream of GLI as well as Hh-independent mechanisms, such as Hippo-YAP and MAPK pathway activation, are involved in BCC tumor progression. Damaging mutations in new SMGs in BCC may explain differences in response to pharmacological treatment.

Interestingly, we observed similarities in terms of activated cancer pathways between BCC and other tumor types, which may depend on the Hh pathway, such as in Hh-driven medulloblastoma[58] or neuroblastoma[59,60] and Wilms tumor[60]. In these tumor types, the list of driver events may include *MYCN*[56–58,60] and *YAP1* (refs. 56,61) amplification or mutations in *PTPN14* (ref. 62). We noted that activation of N-Myc and the Hippo-YAP pathway, which may predispose to BCC

recurrence, also renders poor prognosis in these tumor types[58,62–64]. The increased protein levels of N-Myc in neuroblastomas and medulloblastomas are usually caused by *MYCN* amplification, whereas in BCC the increase is mainly due to point mutations affecting the MB1 domain that prevent protein degradation. Such differences likely reflect the fact that BCCs are highly mutated tumors, whereas neuroblastomas and medulloblastomas show evidence of a qualitatively different genomic instability[62].

BCC is a very common tumor type with frequent cases of multiple lesions, providing a unique opportunity to study specific tumorigenic mechanisms of the Hh pathway and tumor evolution in general.

URLs. Catalogue of Somatic Mutations in Cancer (COSMIC), http://cancer.sanger.ac.uk/; TCGA and cBioPortal tools, http://www.cbioportal.org/; basal cell skin cancer National Comprehensive Cancer Network (NCCN) clinical practice guidelines in oncology, http://www.nccn.org/professionals/physician_gls/pdf/nmsc.pdf; Clinical Practice Guide for basal cell carcinoma of the Australian Cancer Network, http://www.cancer.org.au/content/pdf/HealthProfessionals/ClinicalGuidelines/Basal_cell_carcinoma_Squamous_cell_carcinoma_Guide_Nov_2008-Final_with_Corrigendums.pdf; Picard tools, http://broadinstitute.github.io/picard/.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. All sequences produced in this study have been deposited in the European Genome-phenome Archive (EGA) under accession EGAS00001001540.

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

ACKNOWLEDGMENTS

We thank Z. Modrusan (next-generation sequencing), R. Piskol (computational biology), G. Pau (computational biology), F. Peale (pathology) and S. Jillo (collaboration management) from Genentech, Inc. This work was supported by Swiss Cancer League (LSCC 2939-02-2012), Dinu Lipatti 2014 and Novartis (14B065) research grants to S.I.N.

AUTHOR CONTRIBUTIONS

S.I.N. and L.P. designed the project. S.I.N. and X.B. performed high-throughput sequencing data analysis. S.I.N., F.B., K.P., F.A.S. and R.B.C. performed statistical analysis. S.I.N., X.B., H.J.S., F.J.d.S. and S.E.A. wrote the manuscript. V.Z. and O.M. performed *in silico* modeling. X.B. and P.G.R. performed sample and high-throughput sequencing library preparation. L.P., N.B.-S., G.K. and K.G. collected samples. B.K. and I.A. performed *MYCN in vivo* experiments. T.M. and C.V. performed PTPN14 immunohistochemistry experiments. S.I.N., A.L. and M. Garieri performed RNA-seq analysis. V.B.S., M.A.A. and S.I.N. performed mutational signature analysis. M. Guipponi, P.G.R. and X.B. performed exome capture and high-throughput sequencing. M.E. and O.S. performed experiments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Iwasaki, J.K., Srivastava, D., Moy, R.L., Lin, H.J. & Kouba, D.J. The molecular genetics underlying basal cell carcinoma pathogenesis and links to targeted therapeutics. *J. Am. Acad. Dermatol.* **66**, e167–e178 (2012).
2. Youssef, K.K. *et al.* Identification of the cell lineage at the origin of basal cell carcinoma. *Nat. Cell Biol.* **12**, 299–305 (2010).
3. Hahn, H. *et al.* Mutations of the human homolog of *Drosophila patched* in the nevoid basal cell carcinoma syndrome. *Cell* **85**, 841–851 (1996).
4. Reifenberger, J. *et al.* Somatic mutations in the *PTCH*, *SMOH*, *SUFUH* and *TP53* genes in sporadic basal cell carcinomas. *Br. J. Dermatol.* **152**, 43–51 (2005).
5. Sharpe, H.J. *et al.* Genomic analysis of Smoothened inhibitor resistance in basal cell carcinoma. *Cancer Cell* **27**, 327–341 (2015).
6. Atwood, S.X. *et al.* Smoothened variants explain the majority of drug resistance in basal cell carcinoma. *Cancer Cell* **27**, 342–353 (2015).
7. Smith, M.J. *et al.* Germline mutations in *SUFU* cause Gorlin syndrome–associated childhood medulloblastoma and redefine the risk associated with *PTCH1* mutations. *J. Clin. Oncol.* **32**, 4155–4161 (2014).
8. Sweeney, R.T. *et al.* Identification of recurrent *SMO* and *BRAF* mutations in ameloblastomas. *Nat. Genet.* **46**, 722–725 (2014).
9. Laurendeau, I. *et al.* Gene expression profiling of the hedgehog signaling pathway in human meningiomas. *Mol. Med.* **16**, 262–270 (2010).
10. Brastianos, P.K. *et al.* Genomic sequencing of meningiomas identifies oncogenic *SMO* and *AKT1* mutations. *Nat. Genet.* **45**, 285–289 (2013).
11. Pomeroy, S.L. *et al.* Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415**, 436–442 (2002).
12. Clark, V.E. *et al.* Genomic analysis of non-NF2 meningiomas reveals mutations in *TRAF7*, *KLF4*, *AKT1*, and *SMO*. *Science* **339**, 1077–1080 (2013).
13. Jayaraman, S.S., Rayhan, D.J., Hazany, S. & Kolodney, M.S. Mutational landscape of basal cell carcinomas by whole-exome sequencing. *J. Invest. Dermatol.* **134**, 213–220 (2014).
14. Basset-Seguin, N. *et al.* Vismodegib in patients with advanced basal cell carcinoma (STEVIE): a pre-planned interim analysis of an international, open-label trial. *Lancet Oncol.* **16**, 729–736 (2015).
15. Von Hoff, D.D. *et al.* Inhibition of the hedgehog pathway in advanced basal-cell carcinoma. *N. Engl. J. Med.* **361**, 1164–1172 (2009).
16. Sekulic, A. *et al.* Efficacy and safety of vismodegib in advanced basal-cell carcinoma. *N. Engl. J. Med.* **366**, 2171–2179 (2012).
17. Dreno, B., Basset-Seguin, N., Caro, I., Yue, H. & Schadendorf, D. Clinical benefit assessment of vismodegib therapy in patients with advanced basal cell carcinoma. *Oncologist* **19**, 790–796 (2014).
18. Chang, A.L. & Oro, A.E. Initial assessment of tumor regrowth after vismodegib in advanced basal cell carcinoma. *Arch. Dermatol.* **148**, 1324–1325 (2012).
19. Lawrence, M.S. *et al.* Discovery and saturation analysis of cancer genes across 21 tumour types. *Nature* **505**, 495–501 (2014).
20. Hodis, E. *et al.* A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012).
21. Cancer Genome Atlas Network. Genomic classification of cutaneous melanoma. *Cell* **161**, 1681–1696 (2015).
22. Nikolaev, S.I. *et al.* Exome sequencing identifies recurrent somatic *MAP2K1* and *MAP2K2* mutations in melanoma. *Nat. Genet.* **44**, 133–139 (2012).
23. Drobetsky, E.A. & Sage, E. UV-induced G:C-->A:T transitions at the *APRT* locus of Chinese hamster ovary cells cluster at frequently damaged 5′-TCC-3′ sequences. *Mutat. Res.* **289**, 131–138 (1993).
24. Marionnet, C., Benoit, A., Benhamou, S., Sarasin, A. & Stary, A. Characteristics of UV-induced mutation spectra in human XP-D/ERCC2 gene–mutated xeroderma pigmentosum and trichothiodystrophy cells. *J. Mol. Biol.* **252**, 550–562 (1995).
25. Alexandrov, L.B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
26. Hanawalt, P.C., Donahue, B.A. & Sweder, K.S. Repair and transcription. Collision or collusion? *Curr. Biol.* **4**, 518–521 (1994).
27. Besaratinia, A., Kim, S.I., Bates, S.E. & Pfeifer, G.P. Riboflavin activated by ultraviolet A1 irradiation induces oxidative DNA damage–mediated mutations inhibited by vitamin C. *Proc. Natl. Acad. Sci. USA* **104**, 5953–5958 (2007).
28. Brash, D.E. UV signature mutations. *Photochem. Photobiol.* **91**, 15–26 (2015).
29. Hu, J., Adar, S., Selby, C.P., Lieb, J.D. & Sancar, A. Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. *Genes Dev.* **29**, 948–960 (2015).
30. McGregor, J.M. *et al.* Spectrum of p53 gene mutations suggests a possible role for ultraviolet radiation in the pathogenesis of advanced cutaneous lymphomas. *J. Invest. Dermatol.* **112**, 317–321 (1999).
31. da Silva Almeida, A.C. *et al.* The mutational landscape of cutaneous T cell lymphoma and Sézary syndrome. *Nat. Genet.* **47**, 1465–1470 (2015).
32. Strano, S. *et al.* Mutant p53: an oncogenic transcription factor. *Oncogene* **26**, 2212–2219 (2007).
33. Kato, G.J. & Dang, C.V. Function of the c-Myc oncoprotein. *FASEB J.* **6**, 3065–3072 (1992).
34. Bahram, F., von der Lehr, N., Cetinkaya, C. & Larsson, L.G. c-Myc hot spot mutations in lymphomas result in inefficient ubiquitination and decreased proteasome-mediated turnover. *Blood* **95**, 2104–2110 (2000).
35. Welcker, M. *et al.* The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. *Proc. Natl. Acad. Sci. USA* **101**, 9085–9090 (2004).
36. Freier, K. *et al.* Recurrent NMYC copy number gain and high protein expression in basal cell carcinoma. *Oncol. Rep.* **15**, 1141–1145 (2006).
37. Stacey, S.N. *et al.* New basal cell carcinoma susceptibility loci. *Nat. Commun.* **6**, 6825 (2015).
38. Knoepfler, P.S. & Kenney, A.M. Neural precursor cycling at sonic speed: N-Myc pedals, GSK-3 brakes. *Cell Cycle* **5**, 47–52 (2006).
39. Wong, Y.H. *et al.* KinasePhos 2.0: a web server for identifying protein kinase–specific phosphorylation sites based on sequences and coupling patterns. *Nucleic Acids Res.* **35**, W588–W594 (2007).
40. Wang, W. *et al.* PTPN14 is required for the density-dependent control of YAP1. *Genes Dev.* **26**, 1959–1971 (2012).

41. Welcker, M. & Clurman, B.E. FBW7 ubiquitin ligase: a tumour suppressor at the crossroads of cell division, growth and differentiation. *Nat. Rev. Cancer* **8**, 83–93 (2008).
42. King, B. *et al.* The ubiquitin ligase FBXW7 modulates leukemia-initiating cell activity by regulating MYC stability. *Cell* **153**, 1552–1566 (2013).
43. Zhang, H., Pasolli, H.A. & Fuchs, E. Yes-associated protein (YAP) transcriptional coactivator functions in balancing growth and differentiation in skin. *Proc. Natl. Acad. Sci. USA* **108**, 2270–2275 (2011).
44. Schlegelmilch, K. *et al.* Yap1 acts downstream of α-catenin to control epidermal proliferation. *Cell* **144**, 782–795 (2011).
45. Wilson, K.E. *et al.* PTPN14 forms a complex with Kibra and LATS1 proteins and negatively regulates the YAP oncogenic function. *J. Biol. Chem.* **289**, 23693–23700 (2014).
46. Hao, Y., Chun, A., Cheung, K., Rashidi, B. & Yang, X. Tumor suppressor LATS1 is a negative regulator of oncogene YAP. *J. Biol. Chem.* **283**, 5496–5509 (2008).
47. Tate, G., Kishimoto, K. & Mitsuya, T. Biallelic alterations of the large tumor suppressor 1 (*LATS1*) gene in infiltrative, but not superficial, basal cell carcinomas in a Japanese patient with nevoid basal cell carcinoma syndrome. *Med. Mol. Morphol.* **48**, 177–182 (2015).
48. Krauthammer, M. *et al.* Exome sequencing identifies recurrent somatic *RAC1* mutations in melanoma. *Nat. Genet.* **44**, 1006–1014 (2012).
49. Hammond, D. *et al.* Melanoma-associated mutations in protein phosphatase 6 cause chromosome instability and DNA damage owing to dysregulated Aurora-A. *J. Cell Sci.* **126**, 3429–3440 (2013).
50. Duman-Scheel, M., Weng, L., Xin, S. & Du, W. Hedgehog regulates cell growth and proliferation by inducing cyclin D and cyclin E. *Nature* **417**, 299–304 (2002).
51. Stefansson, B. & Brautigan, D.L. Protein phosphatase PP6 N terminal domain restricts G1 to S phase progression in human cancer cells. *Cell Cycle* **6**, 1386–1392 (2007).
52. Couzens, A.L. *et al.* Protein interaction network of the mammalian Hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci. Signal.* **6**, rs15 (2013).
53. Jaju, P.D. *et al.* Mutations in the kinetochore gene *KNSTRN* in basal cell carcinoma. *J. Invest. Dermatol.* **135**, 3197–3200 (2015).
54. Lee, C.S. *et al.* Recurrent point mutations in the kinetochore gene *KNSTRN* in cutaneous squamous cell carcinoma. *Nat. Genet.* **46**, 1060–1062 (2014).
55. Santos, G.C., Zielenska, M., Prasad, M. & Squire, J.A. Chromosome 6p amplification and cancer progression. *J. Clin. Pathol.* **60**, 1–7 (2007).
56. Zhang, H. *et al.* TEAD transcription factors mediate the function of TAZ in cell growth and epithelial-mesenchymal transition. *J. Biol. Chem.* **284**, 13355–13362 (2009).
57. Valentijn, L.J. *et al.* Functional *MYCN* signature predicts outcome of neuroblastoma irrespective of *MYCN* amplification. *Proc. Natl. Acad. Sci. USA* **109**, 19190–19195 (2012).
58. Pfister, S. *et al.* Outcome prediction in pediatric medulloblastoma based on DNA copy-number aberrations of chromosomes 6q and 17q and the *MYC* and *MYCN* loci. *J. Clin. Oncol.* **27**, 1627–1636 (2009).
59. Xu, L. *et al.* Sonic Hedgehog pathway is essential for neuroblastoma cell proliferation and tumor growth. *Mol. Cell. Biochem.* **364**, 235–241 (2012).
60. Oue, T., Yoneda, A., Uehara, S., Yamanaka, H. & Fukuzawa, M. Increased expression of the hedgehog signaling pathway in pediatric solid malignancies. *J. Pediatr. Surg.* **45**, 387–392 (2010).
61. Murphy, A.J. *et al.* Aberrant activation, nuclear localization, and phosphorylation of Yes-associated protein-1 in the embryonic kidney and Wilms tumor. *Pediatr. Blood Cancer* **61**, 198–205 (2014).
62. Pugh, T.J. *et al.* The genetic landscape of high-risk neuroblastoma. *Nat. Genet.* **45**, 279–284 (2013).
63. Ellison, D.W. *et al.* Medulloblastoma: clinicopathological correlates of SHH, WNT, and non-SHH/WNT molecular subgroups. *Acta Neuropathol.* **121**, 381–396 (2011).
64. Williams, R.D. *et al.* Multiple mechanisms of *MYCN* dysregulation in Wilms tumour. *Oncotarget* **6**, 7232–7243 (2015).

ONLINE METHODS

Sample collection and preparation. We collected 225 BCCs for this study. Sequencing data from 68 previously reported cases were integrated into the analysis (62 exomes from Sharpe *et al.*[5] and 6 exomes from Atwood *et al.*[6] are accessible through the European Genome-phenome Archive (EGA) under accession EGAS00001000845 and through the Gene Expression Omnibus (GEO) under accession GSE58377, respectively). Four samples from Sharpe *et al.* and the six exomes from Atwood *et al.* were analyzed as unmatched samples because of the unavailability or unsatisfactory sequencing quality of the matched germline samples. Only the cancer panel genes were investigated in these ten BCCs.

We obtained 125 fresh tumor samples (and matched blood samples for 104 of them) from patients with appropriate informed consent, and 100 archived formalin-fixed, paraffin-embedded (FFPE) samples (matched blood for 4 of them) were studied with approval from the relevant review board (Commission d'Ethique de la Recherche sur l'Etre Human, Switzerland, protocol 12-247 and Assistance Publique–Hôpitaux de Paris, Hôpital Saint-Louis, Paris, France, IE-2011-590(1)). On the basis of histological subtype, samples were classified as being at high risk (morpheaform and metatypical) or low risk (nodular and superficial) of recurrence according to the 2015 basal cell skin cancer National Comprehensive Cancer Network (NCCN) clinical practice guidelines in oncology and the 2008 Clinical Practice Guide for basal cell carcinoma of the Australian Cancer Network. Of the 104 matched samples, 68 were exome sequenced. The remaining 36 paired samples, along with the 21 unpaired samples and the 100 FFPE samples, underwent targeted sequencing of the cancer panel genes (**Supplementary Tables 1a** and **2**). Fresh BCC tissue sampling was performed by debulking the most nodular part of the tumor, under magnification, to enrich for tumor cells. When possible, adjacent normal skin was also sampled. Skin tissue was immediately snap frozen. DNA from tumors was extracted with the Qiagen DNeasy Blood and Tissue kit, and DNA from blood was purified on a Qiagen Autopure LS instrument. For FFPE samples, punch biopsies or whole excisates with at least 30% tumor cells by histological evaluation were selected for the study. From each block, the first 5 slices were discarded and the subsequent 10–20 slices were used for DNA extraction with the Qiagen QIAamp DNA FFPE Tissue kit. The quality for all DNA samples was evaluated by gel electrophoresis or with Agilent TapeStation Genomic DNA ScreenTape. Exome capture and sequencing library preparation were performed using Agilent SureSelect Human All Exon V5 with either 1 µg or 200 ng of DNA as input, following standard protocols. Cancer panel libraries were prepared with either Agilent SureSelect reagents or the KAPA Hyper Prep kit (Kapa Biosystems), using 50–200 ng of DNA as input. Targeted enrichment of cancer genes was achieved with an Agilent SureSelect kit custom designed for the genes of interest, called throughout the text the cancer panel (**Supplementary Table 2**). In total, 387 cancer-related genes were selected on the basis of their mutational profiles in the COSMIC v71 database, taking into account fractions of truncating and recurrent mutations in different tumor types[65]. All cancer-related genes reported in recent reviews[19,66] were included in the panel as well.

DNA sequencing: exome and cancer panel genes. Libraries were sequenced on an Illumina HiSeq 2000 instrument with 100-bp paired-end reads to a median coverage of 100× (range of 76–204×) for exomes and 171× (range of 34–952×) for the cancer panel. The average Illumina error rate for the FFPE samples was 0.23 ± 0.09% for read 1 and 0.4 ± 0.10% for read 2, with a cluster density of ~1 million clusters/mm^2. The mean percentage of bases with quality score ≥Q30 was 92.31%. For the fresh tissue and blood samples, the average Illumina error rate was 0.24 ± 0.05% for read 1 and 0.37 ± 0.08% for read 2, with a cluster density of ~900,000 clusters/mm^2. The mean percentage of bases with quality score ≥Q30 was 91.83%. Sequencing reads were mapped to the reference genome with the Burrows-Wheeler Aligner (BWA)[67] and then processed with the Genome Analysis Toolkit (GATK; version 3.3.0)[68] following best practices for exomes and specific indications for smaller targeted regions. Sequencing quality and target enrichment were verified with Picard tools metrics. For the tumors with a matching germline sample, MuTect[69] version 1.1.4 with standard parameters was used to identify somatic SNVs and the GATK HaplotypeCaller with the --comp function was used to call somatic indels. For the samples for which a matched germline sample was not available, a panel of normal samples (PON) was constructed from 50 germline samples. We filtered out all known polymorphisms from 1000 genomes. This PON was used in MuTect for somatic variant calling, and only stop-gain, splice-site and probably damaging nonsynonymous mutations were considered to be putative drivers. Additionally, variants were considered putative drivers if reported in the COSMIC v71 database. Some candidate variants were validated by Sanger sequencing. To estimate the fraction of false positive somatic variants in tumors without matching control samples, we reanalyzed 43 tumors for which we had a matched germline sample using a PON. Only 4.6% of the mutations that fit the criteria for driver mutations were false positives. *LATS1* mutation status at Arg995 in the samples sequenced with the cancer panel was evaluated with Sanger sequencing. All samples in which somatic mutations were identified were included in the study. The fraction of tumor cells in each sample was estimated from genetic data as twice the median fraction of reads for all somatic mutations (**Supplementary Table 7**). Mutation rates in samples with a low tumor fraction may be underestimated.

Copy number analysis. SCNAs were retrieved from the exome sequencing data of 100 samples, taking into account allelic coverage, percentage of reads for germline-heterozygous variants and VAF in a sample. Edges and types of SCNAs (loss of a copy, copy-neutral LOH and gain of copies) were detected with a hidden Markov model (HMM) algorithm[70] (**Supplementary Note**) followed by manual inspection (**Supplementary Table 8**). GISTIC 2.0.16 (ref. 71) was used with exome sequencing data to identify regions with significant copy gains and losses.

Significantly mutated genes. *MutSigCV analysis.* The search for SMGs was performed with MutSigCV[72]. As a covariate for the transcriptional profile, median fragments per kilobase per million mapped reads (FPKM) values were calculated from RNA-seq data on 17 BCCs. As a covariate for DNase I hypersensitivity, the profiles for penis foreskin keratinocyte primary cells (PFKPCs) from the Epigenomics Roadmap[73] were used. Reads for all samples belonging to PFKPCs were combined, and reads per kilobase of transcript per million mapped reads (RPKM) values were calculated for 1-Mb windows along the genome. These values were then assigned to genes. Replication time[74] was used as a covariate as well. According to consensus, genes with a q value <0.1 were deemed significant. The MutSigCV algorithm was run for the exome-sequenced BCCs (121 non-clonal samples), and four SMGs were found (**Supplementary Table 5**). To identify less common mutations and to increase the number of samples analyzed[75], we used MutSigCV to look for SMGs in exomes (121 non-clonal samples) and both exomes and cancer panel samples (283 non-clonal samples), taking into account only the genes in the cancer panel list (**Supplementary Table 2**). Three additional SMGs were found (**Supplementary Table 5**). We considered as SMGs the genes deemed significant by any of these three approaches.

Properties of cancer genes. To identify tumor-suppressor genes in our data set, we selected all loss-of-function mutations and LOH events (when available) for each gene and counted the tumors that had one mutation or more in each gene from the cancer panel (**Supplementary Table 2**). A binomial test was applied to calculate the probability of two or more loss-of-function events occurring in the same gene in the same tumor. Fisher's exact test was used to calculate mutual exclusivity between *PTCH1*, *SUFU* and *SMO* genes as well as for the analysis of colocalization or mutual exclusivity of LOH events with single or double driver mutations in the *PTCH1* and *TP53* genes.

Investigation of recurrent mutations. Significantly recurrently mutated nucleotides were identified using TumOnc-v0.01 (**Supplementary Note**). Estimations of the background mutation rate took into account the trinucleotide mutation context, patient mutation rates and the combination of gene covariates (the same covariates used to run MutSigCV were implemented). Calculation accuracy was increased by pooling the data for mutations with similar properties. Significance levels (P values) were determined by testing whether the observed number of mutations at a nucleotide position significantly exceeded the count expected from the background model. The threshold was set to more than four recurrent mutations. FDRs (q values) were then calculated, and genes with a q value <0.05 were reported to be significantly mutated.

TumOnc-v0.01 was run on the two data sets: the 283 non-clonal samples sequenced for cancer panel genes and the 121 non-clonal exome samples sequenced for all genes. Both analyses were performed using the background mutation rate estimated from the full exome set of the data (**Supplementary Table 2**).

The mutation input files for TumOnc-v0.01 included dinucleotide polymorphisms in a format with one single-nucleotide change per line. Full details on the methods used are provided in the **Supplementary Note**.

RNA sequencing. mRNA-seq libraries were prepared from 500 ng to 2 µg of total RNA using the Illumina TruSeq RNA sample prep kit (RS-122-2001), following the manufacturer's instructions, for 61 tumor-extracted RNA samples and 25 normal skin RNA samples. Libraries were sequenced on the Illumina HiSeq 2000 platform using 2 × 100-bp paired-end sequencing. Reads were mapped against the human genome (Build 38) using TopHat. Uniquely mapped reads were used in a custom pipeline to calculate the exon coverage per gene (GENCODE v21 annotation). The coverage was normalized in FPKM.

Analysis of differential expression was performed on the raw read counts using the default parameters of EdgeR[76] for the genes expressed in at least three samples. Differentially expressed genes (FDR < 0.03) were submitted to DAVID[77,78] for gene ontology analysis and KEGG pathway analysis. Gene set enrichment was analyzed with GSEA software[79] (corrected *P* values were obtained after 10,000 permutations of the gene sets).

***In silico* protein structure predictions.** Protein secondary structure predictions were made by the PSIPRED program[80]. Prediction of the phosphorylation sites of N-Myc and the putative kinases was performed with Kinasephos2 (ref. 39). Motifs of protein-protein interaction were retrieved with the ELM algorithm[81]. Structural models of LATS1 and LATS2 were obtained using the Modeller v9.1 program[82] with Protein Data Bank (PDB) IDs 4LQS, 2V55, 4WP7, 3V8S, 4L6Q and 2F2U corresponding to experimental structures as templates. Quantitative estimation of the contribution of LATS1 and LATS2 residues to the stability of the proteins was performed using FoldX[83].

N-Myc functional assays. The HEK293 cells used in this experiment were obtained from the American Type Culture Collection and were routinely tested for mycoplasma. HEK293 cells are not listed in the database of commonly misidentified cell lines.

N-Myc cDNA was cloned into a mammalian expression vector (pcDNA3.1), and mutations were introduced by PCR mutagenesis. HEK293 cells were transiently transfected with N-Myc and FBXW7 constructs (or empty vector controls). The following day, cells were treated with 10 µM MG-132 for 4 h and whole-cell extracts were prepared. FLAG-tagged complexes were immunoprecipitated with FLAG-M2 magnetic beads (Sigma) for 3 h. To isolate complexes containing N-Myc, the extracts were incubated with antibody to N-Myc (B8.4 B (ref. 84), Santa Cruz Biotechnology) overnight and immune complexes were captured with Protein G Dynabeads (Life Technologies). For the *in vivo* ubiquitination assays, these constructs were cotransfected into cells with a construct for His-ubiquitin, and cells were lysed in denaturing buffer containing 8 M urea and *N*-ethylmaleimide to inhibit cysteine proteases. Proteins incorporating His-ubiquitin were enriched with His-Pur columns (Thermo Fisher), according to the manufacturer's instructions, before immunoblot analysis.

Immunohistochemistry. FFPE blocks from eight BCC tumors with *PTPN14* mutations and three BCC tumors without *PTPN14* or *LATS1* mutations were sectioned and mounted for immunohistochemistry analysis with rabbit monoclonal antibody to YAP1 (ab52771, Abcam, EP1674Y; 1:1,000 dilution; refs. 85,86).

Eighteen additional tumors with damaging mutations in *PTPN14* were sectioned and mounted for immunohistochemistry analysis with the same antibody to YAP1 and with mouse monoclonal antibody to Ki-67 (clone MIB1, M7240 (ref. 87), Dako; 1:100 dilution). Staining was performed using a Ventana Ultrastainer following standard protocols. Antigen retrieval used Ventana buffer CC1 and samples were incubated at 95 °C for 52 min. Antibodies were applied manually, and samples were incubated for 1 h at 37 °C. Signal detection was performed using peroxidase with hematoxylin as a counterstain, and samples were then assessed by a clinical pathologist.

65. Nikolaev, S.I. *et al.* Frequent cases of *RAS*-mutated Down syndrome acute lymphoblastic leukaemia lack *JAK2* mutations. *Nat. Commun.* **5**, 4654 (2014).
66. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
67. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
68. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
69. Cibulskis, K. *et al.* Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat. Biotechnol.* **31**, 213–219 (2013).
70. Nikolaev, S. *et al.* Extrachromosomal driver mutations in glioblastoma and low-grade glioma. *Nat. Commun.* **5**, 5690 (2014).
71. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
72. Lawrence, M.S. *et al.* Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature* **499**, 214–218 (2013).
73. Roadmap Epigenomics Consortium. Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
74. Chen, C.L. *et al.* Impact of replication timing on non-CpG and CpG substitution rates in mammalian genomes. *Genome Res.* **20**, 447–457 (2010).
75. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, 576–582 (2015).
76. Nikolayeva, O. & Robinson, M.D. edgeR for differential RNA-seq and ChIP-seq analysis: an application to stem cell biology. *Methods Mol. Biol.* **1150**, 45–79 (2014).
77. Huang, W., Sherman, B.T. & Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
78. Huang, W., Sherman, B.T. & Lempicki, R.A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
79. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* **102**, 15545–15550 (2005).
80. Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K. & Jones, D.T. Scalable web services for the PSIPRED Protein Analysis Workbench. *Nucleic Acids Res.* **41**, W349–W357 (2013).
81. Dinkel, H. *et al.* The eukaryotic linear motif resource ELM: 10 years and counting. *Nucleic Acids Res.* **42**, D259–D266 (2014).
82. Eswar, N. *et al.* Comparative protein structure modeling using MODELLER. *Curr. Protoc. Protein Sci.* Chapter 2, Unit 2.9 (2007).
83. Guerois, R., Nielsen, J.E. & Serrano, L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **320**, 369–387 (2002).
84. Lodrini, M. *et al.* MYCN and HDAC2 cooperate to repress miR-183 signaling in neuroblastoma. *Nucleic Acids Res.* **41**, 6018–6033 (2013).
85. Xiao, W. *et al.* Mutual interaction between YAP and c-Myc is critical for carcinogenesis in liver cancer. *Biochem. Biophys. Res. Commun.* **439**, 167–172 (2013).
86. Li, L. *et al.* MEK1 promotes YAP and their interaction is critical for tumorigenesis in liver cancer. *FEBS Lett.* **587**, 3921–3927 (2013).
87. Gerdes, J., Becker, M.H., Key, G. & Cattoretti, G. Immunohistological detection of tumour growth fraction (Ki-67 antigen) in formalin-fixed and routinely processed tissues. *J. Pathol.* **168**, 85–86 (1992).

doi:10.1038/ng.3525