Large-scale pan-cancer analysis reveals broad prognostic association between TGF-β ligands, not Hedgehog, and GLI1/2 expression in tumors

Aurélien de Reyniès1,*, Delphine Javelaud2,3, Nabila Elarouci1, Véronique Marsaud2,3, Cristèle Gilbert2,3 & Alain Mauviel2,3*

GLI1 expression is broadly accepted as a marker of Hedgehog pathway activation in tumors. Efficacy of Hedgehog inhibitors is essentially limited to tumors bearing activating mutations of the pathway. GLI2, a critical Hedgehog effector, is necessary for GLI1 expression and is a direct transcriptional target of TGF-β/SMAD signaling. We examined the expression correlations of GLI1/2 with TGFβ and HH genes in 152 distinct transcriptome datasets totaling over 23,500 patients and representing 37 types of neoplasms. Their prognostic value was measured in over 15,000 clinically annotated tumor samples from 26 tumor types. In most tumor types, GLI1 and GLI2 follow a similar pattern of expression and are equally correlated with HH and TGFβ genes. However, GLI1/2 broadly share prognostic value with TGFβ genes and a mesenchymal/EMT signature, not with HH genes. Our results provide a likely explanation for the frequent failure of anti-Hedgehog therapies in tumors, as they suggest a key role for TGF-β, not Hedgehog, ligands, in tumors with elevated GLI1/2-expression.

Elevated expression and nuclear localization of GLI1, reminiscent of Hedgehog (HH) pathway activation, has been reported in a wide variety of tumor types1. Although extensive experimental evidence exists for a protumorigenic and pro-metastatic role of GLI1, efficacy of HH inhibitors is restricted to a handful of cancers with genetic activation of upstream components of the pathway2,3. The sole FDA-approved indication for HH inhibitors is advanced cutaneous basal cell carcinoma4. Understanding their lack of efficacy in other tumors despite high GLI1 expression remains a challenge5,6.

Members of the HH family of growth factors, Sonic (SHH), Indian (IHH) and Desert (DHH) control tissue patterning, limb and skeletal polarity during embryonic life, and broadly contribute to tissue homeostasis and repair processes during adulthood, by controlling cell proliferation, migration, as well as stem cell maintenance and self-renewal1,7. Ligand binding to the 12-transmembrane receptor PATCHED-1 (PTCH1) allows activation of the 7-transmembrane G-coupled receptor Smoothened (SMO), and HH signal transduction proceeds towards activation and nuclear accumulation of GLI1 transcription factors with activator or repressor functions dependent upon proteolytic cleavage8. GLI1 is a prototypic HH target gene and its expression is widely considered a read-out of HH pathway activation. GLI2 is the primary substrate and effector of the pathway that largely contributes to GLI1 induction by HH, as well as that of other HH target genes in cooperation with GLI19,10. GLI3 has weak transcriptional activity and is considered an inhibitor of HH activity9.

Ligand-independent HH pathway activation as a result of mutations in genes encoding upstream pathway components, such as loss-of-function mutations in PTCH1 or SUFU (suppressor of fused) and activating mutations of SMO are rare in cancers and only found in cutaneous basal cell carcinoma, rhabdomyosarcoma, and medulloblastoma, cancers which exhibit notable therapeutic response to HH inhibitors11.
Clinical studies on HH inhibitors have often relied on increased expression or nuclear localization of GLI1 as a marker of HH pathway activation. It may be argued that in a number of cases, HH activation was stated empirically, with no direct evidence of active upstream HH signaling. Similarly, studies targeting GLI1/2 expression or function in tumor cells in vitro or in mouse models of cancer have shown remarkable anti-tumor efficacy and concluded to a pathogenic role for the HH pathway, although most studies targeted its main downstream effectors, not the pathway itself. Thus, while there is little doubt that GLI transcription factors contribute substantially to cancer progression, direct evidence that would link GLI activity in a given tumor setting to HH ligands activating their receptors is often missing. This may explain the overall lack of anti-tumor therapeutic efficacy of HH inhibitors, most of which target SMO activity1.

TGF-β is secreted abundantly by both tumor and stromal cells, allowing tumor evasion from immune surveillance, peri-tumoral angiogenesis and EMT, processes that all contribute to tumor progression2,3. TGF-β signals via ubiquitously expressed membrane-bound heteromeric serine-threonine kinase receptor complexes2. We identified TGF-β as a powerful inducer of GLI2 and GLI1 expression as well as GLI-dependent transcription, independent from SMO activity2,3. We established a role for both TGF-β signaling and GLI2 in driving melanoma invasion and metastasis, that could be targeted with TGF-β receptor inhibitors, the latter inhibiting GLI2 expression in tumor cells, or by knocking down GLI2 expression14,15. High GLI2 expression in invasive melanoma cells depends largely upon autocrine TGF-β signaling and is associated with a mesenchymal transition and loss of E-cadherin expression, events associated with enhanced cell motility and capacity to metastasize5,17. Similar observations have since been reported for other tumor types, including ovarian and oral squamous cell carcinomas, that link GLI2 and TGF-β expression to tumor aggressiveness via various mechanisms such as induction of PTHrP, leading to enhanced osteolytic bone metastases, or that of a stemness-like phenotype that also promotes metastatic progression16–20.

Herein, we hypothesized that the lack of efficacy of SMO antagonists in numerous tumors occurs because high GLI1 expression and activity may not be linked to HH, but rather to TGF-β, ligand expression, taken as surrogates for HH and TGF-β signaling in tumors, irrespective of the cellular compartment. We compiled data from publicly available gene expression datasets from over 23,500 cancer patients, of which over 15,000 with survival annotations, well above those available from The Cancer Genome Atlas (TCGA). While GLI1/2 expression is correlated with both HH and TGFβ expression, their prognostic value is tightly correlated with that of TGFβ, not HH. High GLI1/2 and TGFβ expression, associated with a mesenchymal/EMT signature, often represent parallel markers of poor clinical outcome. Inversely, high expression of HH is mostly associated with increased survival.

**Results**

Pan cancer correlation between GLI1, GLI2, TGFβ, and HH genes. We hypothesized that the correlation between GLI1 and GLI2 expression with HH (SHH, IHH, DHH) and TGFβ ligands (TGFβ1, TGFβ2, TGFβ3) transcript levels represent adequate surrogates for the respective pathogenic implication of HH and TGFβ ligands in GLI1/2 expression and activity in tumors. Pan-cancer analysis of the correlation of GLI1 and GLI2 with 19,540 genes expressed in 30 tumor types revealed that GLI1 and GLI2 are mutually the most correlated genes (Fig. 1A), and that all HH and TGFβ genes are in the top 20 percentiles of most correlated genes with GLI1 and GLI2, with one exception. Expression of at least one of the TGFβ genes was more closely related than that of any HH genes to both GLI1 and GLI2 expression.

Positive correlation (arbitrary threshold: r > 0.25) between the GLI1 and GLI2 genes was observed in 33/37 tumor types (Fig. 1B), representing 92.5% (21,825/23,587) of patients. Most tumor types exhibited similar correlation values between GLI1 expression and that of at least one of either HH or TGFβ genes (Fig. 1B). The correlation pattern between HH and TGFβ genes with GLI2 (Fig. 1B) was similar to that with GLI1, yet tumors with high GLI2/TGFβ correlation could be discriminated into two subgroups: one exhibiting low GLI2/HH correlation (tumor types from ovarian down to bladder), the other exhibiting high GLI2/HH correlation (tumor types from cervix down to thyroid). We did not identify a single neoplasm for which GLI1/2 expression was correlated with that of HH genes without a simultaneous correlation with that of at least one of the TGFβ genes.

Expression of GLI1, GLI2, HH and TGFβ genes differentially associates with key oncogenic signatures. Cell cycle progression, acquisition of a mesenchymal phenotype through epithelial-to-mesenchymal transition (EMT), and cell stemness are cellular traits considered hallmarks of cancer progression21, to which both the HH and TGF-β pathways are linked2,3. For each of the 37 tumor types, we generated a multivariate linear model based on the expression of the eight genes of interest taken together (three HH and three TGFβ genes, GLI1 and GLI2) to determine whether it may be predictive of these metagenes. To assess the goodness-of-fit of these models, correlations between the predicted values and observed values for each metagene were calculated in each tumor type. As shown in Fig. 2A, strong correlations were observed in most tumor types for the mesenchymal/EMT and cancer cell stemness metagenes albeit to a lesser extent, while it was seldom observed with the cell cycle metagene.

A simplified multivariate linear model using compounded expression of either GLI1/2, the HH or TGFβ genes was next calculated for each metagene. Coefficients for these three predictive variables within each model, presented in Fig. 2B, demonstrate the dominant role of TGFβ gene expression, followed by that of GLI1/2, not HH, in predicting mesenchymal and cancer cell stemness metagene expression in a broad array of tumor types. None of them was associated with the cell cycle metagene. Data for each GLI, HH and TGFβ gene taken individually in each model are provided in Supplementary Figure S1. It should be noted that the higher correlation observed with the mesenchymal/EMT metagene is consistent with the fact that the latter comprises a number of known TGF-β/SMAD target genes.
Figure 1. Pan-cancer expression correlations with GLI1 and GLI2. (A) Data from 152 expression public datasets from 37 cancer types, spanning over 23,500 patients were sorted and the correlation of 19,540 genes (expressed in at least 30 tumor types) with that of GLI1 (left panel) and GLI2 (right panel) was calculated. The respective position of HH and TGFB genes is indicated. (B) Heatmap representation of the expression correlation between GLI1 and GLI2 with each other and between HH and TGFB genes in 37 cancer types. Corresponding numerical values are provided in Supplementary Table 2. The number of patients for each neoplasm is indicated.
Pan-cancer prognostic values associated with GLI1/2, HH and TGFB transcript levels and select oncogenic signatures. Univariate Cox survival analysis from over 15,000 clinically annotated tumor samples brought critical information (Fig. 3A). At odds with a generalized assumption in the clinical setting that HH signaling is deleterious, expression of all HH genes in tumors was associated with good prognosis (H.R. < 1, green color). Not a single tumor type was found for which expression of any HH gene was associated with morbidity (H.R. > 1, red color) without a parallel pejorative prognostic value for at least one TGFB gene and either GLI1, or GLI2. Also, for each occurrence when high GLI1 or GLI2 expression was of bad prognosis, the same held true for at least one of the TGFB genes. On the other hand, when either GLI1 and GLI2 expression were of good prognosis, the same applied for at least one of the TGFB genes. Notably, in bladder, colorectal and kidney papillary carcinoma, GLI2 (and GLI1) expression shared bad prognostic value with that of TGFB genes, while high HH expression was associated with a positive outcome. As expected, the three metagenes were largely of bad prognosis, with mesenchymal/EMT and cancer cell stemness metagenes exhibiting a largely overlapping, yet cancer type-specific pattern of prognostic significance, while the cell cycle metagene was almost universally associated with poor outcome.

In breast cancer, HH expression had no prognostic value while GLI1/2 and TGFB expression were associated with better survival, together with the mesenchymal/EMT and cell stemness metagenes. These results are at odds with most neoplasms where GLI1/2 and TGFB genes share pejorative prognostic value. To understand this discrepancy, a large cohort of breast cancer patients’ data was further analyzed. Intra-dataset z-score GLI1(1/2), TGFB(1/2/3) and HH(S/I/D) and metagene expression values were sorted according to increasing GLI2 expression, then aligned to the molecular subtypes. High GLI1/2/TGFB and mesenchymal/EMT metagene expression was associated with the normal-like subgroup of tumors with better prognosis, and inversely correlated with luminal-type tumors of poor prognosis (Supplementary Figure S2). A hypothesis may be that the mesenchymal/
Figure 3. Prognostic value associated with GLI(1/2), HH(S/I/D) and TGFβ(1/2/3) genes and the mesenchymal/EMT, cell stemness and cell cycle metagenes in human cancers. (A) Data were derived from a meta-analysis for the univariate prognostic value for overall survival in 26 types of human cancers for which sufficient events were available. Bigger circles represent lower p values. Marked colors represent p value below 0.001, dull colors represent p values below 0.05. Green and red colors represent H.R. below and above 1, respectively (see scale bar). (B) Correlations between pan-cancer profiles of prognostic scores of GLI(1/2), HH(S/I/D) and TGFβ(1/2/3) genes and the mesenchymal/EMT, cell stemness and cell cycle metagenes. Prognostic scores are defined as the log2(H.R.) if the related p-value is below 0.05, or 0 otherwise.
whether driven by TGF-β, not HH, which would explain the lack of efficacy of anti-HH approaches. Large-scale datamin-

GLI1 2 significant correlation was found between the expression of GLI1 and oncogenic function metagenes were calculated. The results, presented in Fig. 3B, overwhelmingly demonstrate that GLI(1/2), TGF(1/2/3), Mesenchymal/EMT and stemness metagenes have highly correlated pan-cancer prognostic profiles, with no or modest correlation to that of HH(S/I/D). Little or no correlation between GLI(1/2), TGF(1/2/3), (S/I/D)HH was found with the cell cycle metagene H.R., consistent with the expression correlation data from Fig. 2.

**Discussion**

HH inhibitors have failed to fulfill expectations placed on them as powerful anti-cancer drugs. Despite widespread expression of GLI1 in tumors, considered to be a marker of HH activation, HH inhibitors have solely been granted FDA approval for the treatment of advanced basal cell carcinoma of the skin, a type of tumor that bears activating mutations of the HH pathway. Clinical trials on other solid (or non-solid) tumors have overall failed. Based on our earlier work that identified TGF-β as a potent transcriptional inducer of GLI1, consequently leading to SMO-independent induction of GLI1, we hypothesized that GLI1 expression in tumors may be driven by TGF-β, not HH, which would explain the lack of efficacy of anti-HH approaches. Large-scale datamin-

EMT metagene, which follows GLI2 and TGFβ expression pattern, may not be representative of actual EMT in normal-like breast tumors, but rather represents that of fibroblastic heterogeneity found in breast tumors. Clinical trials on other solid (or non-solid) tumors have overall failed. Based on our earlier work that identified TGF-β as a potent transcriptional inducer of GLI1, consequently leading to SMO-independent induction of GLI1, we hypothesized that GLI1 expression in tumors may be driven by TGF-β, not HH, which would explain the lack of efficacy of anti-HH approaches. Large-scale datamin-

Our broad pan-cancer analysis herein indicates that GLI1/2 functions parallel or match those of TGF-β ligands, not HH, as identified both in linear models of oncogenic functions and in prognostic analyses of tumors. Multivariate analysis demonstrated the dominant role of TGFβ, followed by that of GLI1/2, not HH, in predicting mesenchymal/EMT and cell stemness metagene expression in a broad array of tumor types. Noteworthy, these analyses were performed using gene expression data from whole tumors, that do not discriminate between tumor and stromal cells. These data fit a mechanistic model whereby TGF-β, acts in a paracrine or autocrine manner to control GLI2 expression which, in turn and depending upon context, allows for GLI1 expression in either a HH-dependent or -independent fashion, as proposed previously45. Our data hint that HH ligand-driven signaling in tumors leading to GLI1 expression without an overlap and contribution of the TGF-β pathway is not only a rare event but is also unlikely to be pathogenic. The divergence between the prognostic value of GLI1/2 expression and that of HH ligands indicates that there is no sensible justifica-

Our data hint that HH ligand-driven signaling in tumors leading to GLI1 expression without an overlap and contribution of the TGF-β pathway is not only a rare event but is also unlikely to be pathogenic. The divergence between the prognostic value of GLI1/2 expression and that of HH ligands indicates that there is no sensible justifica-

Our data hint that HH ligand-driven signaling in tumors leading to GLI1 expression without an overlap and contribution of the TGF-β pathway is not only a rare event but is also unlikely to be pathogenic. The divergence between the prognostic value of GLI1/2 expression and that of HH ligands indicates that there is no sensible justifica-

Our broad pan-cancer analysis herein indicates that GLI1/2 functions parallel or match those of TGF-β ligands, not HH, as identified both in linear models of oncogenic functions and in prognostic analyses of tumors. Multivariate analysis demonstrated the dominant role of TGFβ, followed by that of GLI1/2, not HH, in predicting mesenchymal/EMT and cell stemness metagene expression in a broad array of tumor types. Noteworthy, these analyses were performed using gene expression data from whole tumors, that do not discriminate between tumor and stromal cells. These data fit a mechanistic model whereby TGF-β, acts in a paracrine or autocrine manner to control GLI2 expression which, in turn and depending upon context, allows for GLI1 expression in either a HH-dependent or -independent fashion, as proposed previously45.

Our data hint that HH ligand-driven signaling in tumors leading to GLI1 expression without an overlap and contribution of the TGF-β pathway is not only a rare event but is also unlikely to be pathogenic. The divergence between the prognostic value of GLI1/2 expression and that of HH ligands indicates that there is no sensible justifica-

Our data hint that HH ligand-driven signaling in tumors leading to GLI1 expression without an overlap and contribution of the TGF-β pathway is not only a rare event but is also unlikely to be pathogenic. The divergence between the prognostic value of GLI1/2 expression and that of HH ligands indicates that there is no sensible justifica-

Our data hint that HH ligand-driven signaling in tumors leading to GLI1 expression without an overlap and contribution of the TGF-β pathway is not only a rare event but is also unlikely to be pathogenic. The divergence between the prognostic value of GLI1/2 expression and that of HH ligands indicates that there is no sensible justifica-
Limitations

1. Both tumor and stromal cells can express TGF-β ligand with is able to act in a paracrine or autocrine manner. This is highly context-specific and highly variable. Since the data originate from non-dissected tumors, we have to assume that the cellular origin of ligands does not affect the outcome of our analyses.

2. Driver mutations in either HH or TGF-β signaling are likely to influence GLI1/2 expression. Yet, we would like to contend that the outcome of HH signaling from driver mutations would not modify the conclusions drawn from our analyses, as those driver mutations are rare and restricted to a limited number of cancer types.

3. While a number of published mechanistic studies fit with our model, it would be interesting for our analyses to be functionally validated at the protein level in patients’ samples.

Material and methods

Transcriptome series. A set of 135 transcriptome series related to 37 cancer types was collected from public repositories (ArrayExpress, GEO, TCGA). Series that included multiple cancer types were split accordingly, yielding to a total of 152 distinct transcriptome datasets representing over 23,500 patients. Details and accession numbers are provided in Supplementary Table S1.

Pre-treatment, normalization. Datasets based on Affymetrix microarrays were normalized independently using the justRMA function from the Affy R package (with default parameter). For TCGA RNA-Seq datasets, raw counts were normalized using the upper quartile method 26. Datasets from other sources were used as furnished, after log2 transformation of expression values. All probe sets data were aggregated by HUGO Gene Symbol.

Correlation analyses. Transcripts with available measures in at least 30 cancer types were selected for further analyses (n = 19,540). The correlation between their expression and that of GLI1 and GLI2 was calculated independently in each of the 152 datasets, yielding 152 correlation matrices (dimension: 19,540 × 2). Correlations were then averaged across all 152 dataset matrices, yielding a unique (19,540 × 2) matrix used to plot Fig. 1A. The 152 matrices were also averaged per cancer type, yielding 37 sub-matrices that were reduced to the 8 genes of interest (GLI1, GLI2, TGFB1(1/2/3) and HH(S/I/D), Supplementary Table S2) and used to plot Fig. 1B.

Metagene calculation. Three published gene signatures corresponding to critically important oncogenic activities were selected: (i) mesenchymal-EMT27, (ii) Stemness28; (iii): Cell Cycle: https://www.genome.jp/dbget-bin/www_bget?pathway+hsa04110. Gene content for each signature is listed in Supplementary Table S3. For each of these signatures and each dataset, the average zero-centered expression of the genes, both measured and included in the signature, was calculated for each sample.

Linear models. Within each dataset, based upon the metagene values for each of the three oncogenic signatures, the lm function from the stats R package was used to perform a linear regression of the metagene variable, using GLI1, GLI2, SHH, IHH, DIH, TGFB1, TGFB2 and TGFB3 expression as predictive variables. To allow for inter-datasets and inter-variables comparisons, all variables were z-scored within each dataset before linear modeling (common unit = standard deviation). For each model in each dataset, correlations between predicted values and observed metagene values (z-scores) was recorded and averaged across datasets by cancer type, then represented as heatmaps. For each linear model in each dataset, coefficients of the predictive variables were recorded and averaged across datasets by cancer type and metagene. Numerical values are provided per series and per cancer type (Supplementary Tables S4A and S4B, respectively).

Survival analyses. Univariate Cox models of overall survival in 26 tumor types with cohorts of more than 50 patients and at least 10 death events were calculated using the survival R package. Aggregation of Hazard Ratios (H.R.) and related confidence intervals across datasets of a given cancer type for a given variable were calculated using the meta R package. All genes and metagenes were z-scored intra-dataset prior to modeling, to allow for inter-datasets comparisons (common unit = standard deviation). Numerical values are provided per series and per cancer type (Supplementary Tables S5A and S5B, respectively).

Received: 18 September 2019; Accepted: 13 August 2020
Published online: 02 September 2020

References
1. Wu, F., Zhang, Y., Sun, B., McMahon, A. P. & Wang, Y. Hedgehog signaling: from basic biology to cancer therapy. Cell Chem. Biol. 24, 252–280 (2017).
2. Perrot, C. Y., Javelaud, D. & Mauviel, A. Overlapping activities of TGF-beta and Hedgehog signaling in cancer: therapeutic targets for cancer treatment. Pharmacol. Ther. 137, 183–199 (2013).
3. Javelaud, D., Pierrat, M. J. & Mauviel, A. Crosstalk between TGF-beta and hedgehog signaling in cancer. FEBS Lett. 586, 2016–2025 (2012).
4. Rimkus, T. K., Carpenter, R. L., Qasem, S., Chan, M. & Lo, H. W. Targeting the sonic Hedgehog signaling pathway: review of smoothened and GLI inhibitors. Cancers (Basel) https://doi.org/10.3390/cancers8020022 (2016).
5. Laukkanen, M. O. & Castellone, M. D. Hijacking the Hedgehog pathway in cancer therapy. Anticancer Agents Med. Chem. 16, 409–317 (2016).

6. Justilien, V. & Fields, A. P. Molecular pathways: novel approaches for improved therapeutic targeting of Hedgehog signaling in cancer stem cells. Clin. Cancer Res. 21, 505–513 (2015).

7. Pak, E. & Segal, R. A. Hedgehog signal transduction: key players, oncogenic drivers, and cancer therapy. Dev. Cell 38, 333–344 (2016).

8. Ryan, K. E. & Chiang, C. Hedgehog secretion and signal transduction in vertebrates. J. Biol. Chem. 287, 17905–17913 (2012).

9. Bai, C. B., Auerbach, W., Lee, J. S., Stephen, D. & Joyner, A. L. Gli2, but not Gli1, is required for initial Shh signaling and ectopic activation of the Shh pathway. Development 129, 4753–4761 (2002).

10. Bai, C. B. & Joyner, A. L. Gli1 can rescue the in vivo function of Gli2. Development 128, 5161–5172 (2001).

11. Teglund, S. & Toftgard, R. Hedgehog beyond medulloblastoma and basal cell carcinoma. Biochim. Biophys. Acta 1805, 181–208 (2010).

12. Dennler, S. et al. Induction of sonic hedgehog mediators by transforming growth factor-beta: Smad3-dependent activation of Gli2 and Gli1 expression in vitro and in vivo. Cancer Res. 67, 6981–6986 (2007).

13. Dennler, S., Andre, J., Verrecchia, F. & Mauviel, A. Cloning of the human GLI2 Promoter: transcriptional activation by transforming growth factor-beta via SMAD3/beta-catenin cooperation. J. Biol. Chem. 284, 31523–31531 (2009).

14. Mohammad, K. S. et al. TGF-beta/RI kinase inhibitor SD-208 reduces the development and progression of melanoma bone metastases. Cancer Res. 71, 175–184 (2011).

15. Alexaki, V. I. et al. GLI2-mediated melanoma invasion and metastasis. J. Natl. Cancer Inst. 102, 1148–1159 (2010).

16. Javelaud, D. et al. Stable overexpression of Smad7 in human melanoma cells impairs bone metastasis. Cancer Res. 67, 2317–2324 (2007).

17. Perrot, C. V. et al. GLI2 cooperates with ZEB1 for transcriptional repression of CDH1 expression in melanoma cells. Pigment Cell Melanoma Res. 26, 861–873 (2013).

18. Steg, A. D. et al. Stem cell pathways contribute to clinical chemoresistance in ovarian cancer. Clin. Cancer Res. 18, 869–881 (2012).

19. Johnson, R. W. et al. TGF-beta promotion of Gli2 induced PTHrP expression is independent of canonical Hedgehog signaling. Cancer Res. 71, 822–831 (2011).

20. Cannomier, S. A., Gonzales, C. B., Ely, K., Guelcher, S. A. & Sterling, J. A. Hedgehog and TGFbeta signaling converge on Gli2 to control bony invasion and bone destruction in oral squamous cell carcinoma. Oncotarget 7, 76062–76075 (2016).

21. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. Cell 144, 646–674 (2011).

22. Guedj, M. et al. A refined molecular taxonomy of breast cancer. Oncogene 31, 1196–1206 (2012).

23. Costa, A. et al. Fibroblast heterogeneity and immunosuppressive environment in human breast cancer. Cancer Cell 33, 463–479 (2018).

24. Javelaud, D. et al. TGF-beta/SMAD/GLI2 signaling axis in cancer progression and metastasis. Cancer Res. 71, 5606–5610 (2011).

25. Sekulkis, A. et al. Efficacy and safety of vismodegib in advanced basal-cell carcinoma. N. Engl. J. Med. 366, 2171–2179 (2012).

26. Bullard, J. H., Purdom, E., Hansen, K. D. & Dudoit, S. Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinform. 11, 94 (2010).

27. de Reynies, A. et al. Molecular classification of malignant pleural mesothelioma: identification of a poor prognosis subgroup linked to the epithelial–mesenchymal transition. Clin. Cancer Res. 20, 1323–1334 (2014).

28. Ben-Porath, I. et al. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. Nat. Genet. 40, 499–507 (2008).

Acknowledgements
This work was supported by the French national program Cartes d’Identité des Tumeurs (CIT) funded and developed by the Ligue Nationale Contre le Cancer, grants from Ligue Nationale Contre le Cancer (Equipe Labellisée LIGUE-2016), Institut National du Cancer (INCa MELA13-002) and Donation Henriette et Emile Goutière (to A.M), Fondation ARC (to D.J.), and institutional funding from Institut National de la Santé et de la Recherche Médicale, Centre National de la Recherche Scientifique, Institut Curie, and Université Paris-Sud.

Author contributions
A.R. and A.M. conceptualized the project, performed the bioinformatics analyses, and wrote the manuscript. N.E. formatted several of the datasets used in this study. V.M. and C.G. provided help in presenting the data and writing the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-71559-w.

Correspondence and requests for materials should be addressed to A.R. or A.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020