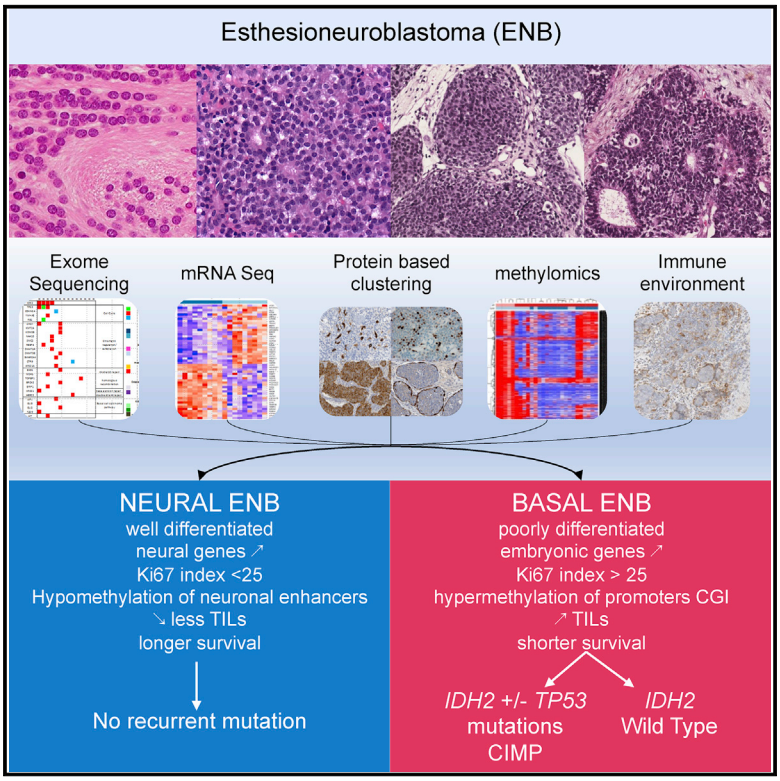
Cell Reports

Resource

Integrated Multi-omic Analysis of Esthesioneuroblastomas Identifies Two Subgroups Linked to Cell Ontogeny

Graphical Abstract



Highlights

- Integrative analysis of ENBs identifies neural-like and basal-like subgroups
- Approximately 35% of basal-like ENBs display *IDH2* R172 hotspot mutations
- ENBs with *IDH2* mutations display a CpG island methylator phenotype (E-CIMP)
- Basal-like ENBs display higher intratumoral-infiltrating lymphocytes

Authors

Marion Classe, Hui Yao, Roger Mouawad, ..., David Khayat, Xiaoping Su, Gabriel G. Malouf

Correspondence

marionclasse@hotmail.fr (M.C.), gabriel.malouf@igbmc.fr (G.G.M.)

In Brief

Classe et al. report an integrative multi-omics analysis of esthesioneuroblastomas (ENBs) and identify two subgroups of ENBs: neural-like and basal-like. These subgroups are linked to cell ontogeny and are associated with distinct clinicopathological features and patient outcomes. Notably, one-third of basal ENBs harbor an *IDH2* R172 mutation with a CpG island methylator phenotype.

Classe et al., 2018, Cell Reports *25*, 811–821
October 16, 2018 © 2018 The Author(s).
https://doi.org/10.1016/j.celrep.2018.09.047

CellPress

Cell Reports
Resource

OPEN ACCESS
CellPress

Integrated Multi-omic Analysis of Esthesioneuroblastomas Identifies Two Subgroups Linked to Cell Ontogeny

Marion Classe,[1,14,*] Hui Yao,[2,14] Roger Mouawad,[1] Chad J. Creighton,[2,3] Alice Burgess,[4] Frederick Allanic,[1] Michel Wassef,[5] Xavier Leroy,[6] Benjamin Verillaud,[4] Geoffrey Mortuaire,[7] Franck Bielle,[8] Christophe Le Tourneau,[9,10,11] Jean-Emmanuel Kurtz,[12] David Khayat,[1] Xiaoping Su,[2,15] and Gabriel G. Malouf[1,12,13,15,16,*]

[1]Department of Medical Oncology, Groupe Hospitalier Pitié-Salpêtrière, Assistance Publique Hôpitaux de Paris, Sorbonnes-Universités, University Pierre and Marie Curie, Paris, France
[2]Department of Bioinformatics and Computational Biology, The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[3]Division of Biostatistics, Department of Medicine and Dan L. Duncan Comprehensive Cancer Center, Baylor College of Medicine, Houston, TX 77030, USA
[4]Department of Otolaryngology-Head and Neck Surgery, Lariboisière Hospital, Assistance Publique Hôpitaux de Paris, Université Paris-Diderot Paris VII, Paris, France
[5]Department of Pathology, Lariboisière, Assistance Publique-Hôpitaux de Paris, Université Paris-Diderot Paris VII, Paris, France
[6]Department of Pathology, CHRU de Lille, Université Lille 2, Lille, France
[7]Department of Otolaryngology-Head and Neck Surgery, CHRU de Lille, Université Lille 2, Lille, France
[8]Department of Neuropathology Raymond Escourolle, AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, Paris, 75013, France
[9]Department of Drug Development and Innovation, Institut Curie, Saint-Cloud, France
[10]INSERM U900 Research Unit, Saint-Cloud, France
[11]Versailles-Saint-Quentin-en-Yvelines University, Montigny-le-Bretonneux, France
[12]Department of Hematology and Medical Oncology, CHRU Strasbourg, Hôpital Hautepierre, Strasbourg, France
[13]Institut Génomique et de Biologie Moléculaire et Cellulaire (IGBMC), Illkirch-Graffenstaden, France
[14]These authors contributed equally
[15]Senior author
[16]Lead Contact
*Correspondence: marionclasse@hotmail.fr (M.C.), gabriel.malouf@igbmc.fr (G.G.M.)
https://doi.org/10.1016/j.celrep.2018.09.047

SUMMARY

Esthesioneuroblastoma (ENB) is a rare cancer of the olfactory mucosa, with no established molecular stratification to date. We report similarities of ENB with tumors arising in the neural crest and perform integrative analysis of these tumors. We propose a molecular-based subtype classification of ENB as basal or neural, both of which have distinct pathological, transcriptomic, proteomic, and immune features. Among the basal subtype, we uncovered an *IDH2* R172 mutant-enriched subgroup (~35%) harboring a CpG island methylator phenotype reminiscent of *IDH2* mutant gliomas. Compared with the basal ENB methylome, the neural ENB methylome shows genome-wide reprogramming with loss of DNA methylation at the enhancers of axonal guidance genes. Our study reveals insights into the molecular pathogenesis of ENB and provides classification information of potential therapeutic relevance.

INTRODUCTION

Esthesioneuroblastoma (ENB), also known as olfactory neuroblastoma, is a rare tumor that arises in the skull base and expands into the nasal cavities. The annual incidence is estimated at 4 cases per 10 million people, and ENB accounts for approximately 3% of all sinonasal tumors (Su et al., 2014). Hyams grading establishes four categories of ENB associated with outcomes, although ENB is often divided into low- and high-grade tumors (Bell et al., 2015; Malouf et al., 2013). Patients with low-grade tumors typically experience late loco-regional recurrence, whereas those with high-grade tumors frequently develop distant metastasis, with 2-year survival less than 40% (Malouf et al., 2013). Thus, determining the optimal management of these tumors remains an unmet medical need.

Several prognostic factors have been associated with poor outcomes for patients with ENB, including TNM stage, presence of lymph node metastasis, and tumor grade (Czapiewski et al., 2016). However, management of aggressive ENBs is challenging, because no targetable oncogenic driver has been identified. Two recent reports analyzed ENB using next-generation targeted sequencing of a panel of cancer genes (Gay et al., 2017; Lazo de la Vega et al., 2017). Although Gay et al. (2017) identified *TP53* mutations as the most frequent mutation, Lazo de la Vega et al. (2017) found no gene to be frequently mutated. Tumor grade was not reported in either study, and several tumors were pretreated prior to sequencing.

Thus, we sought to classify these tumors by molecular subtype and identify their cells of origin within olfactory mucosa. We performed integrative analysis of primary ENB samples that were well annotated regarding their pathological and clinical

This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

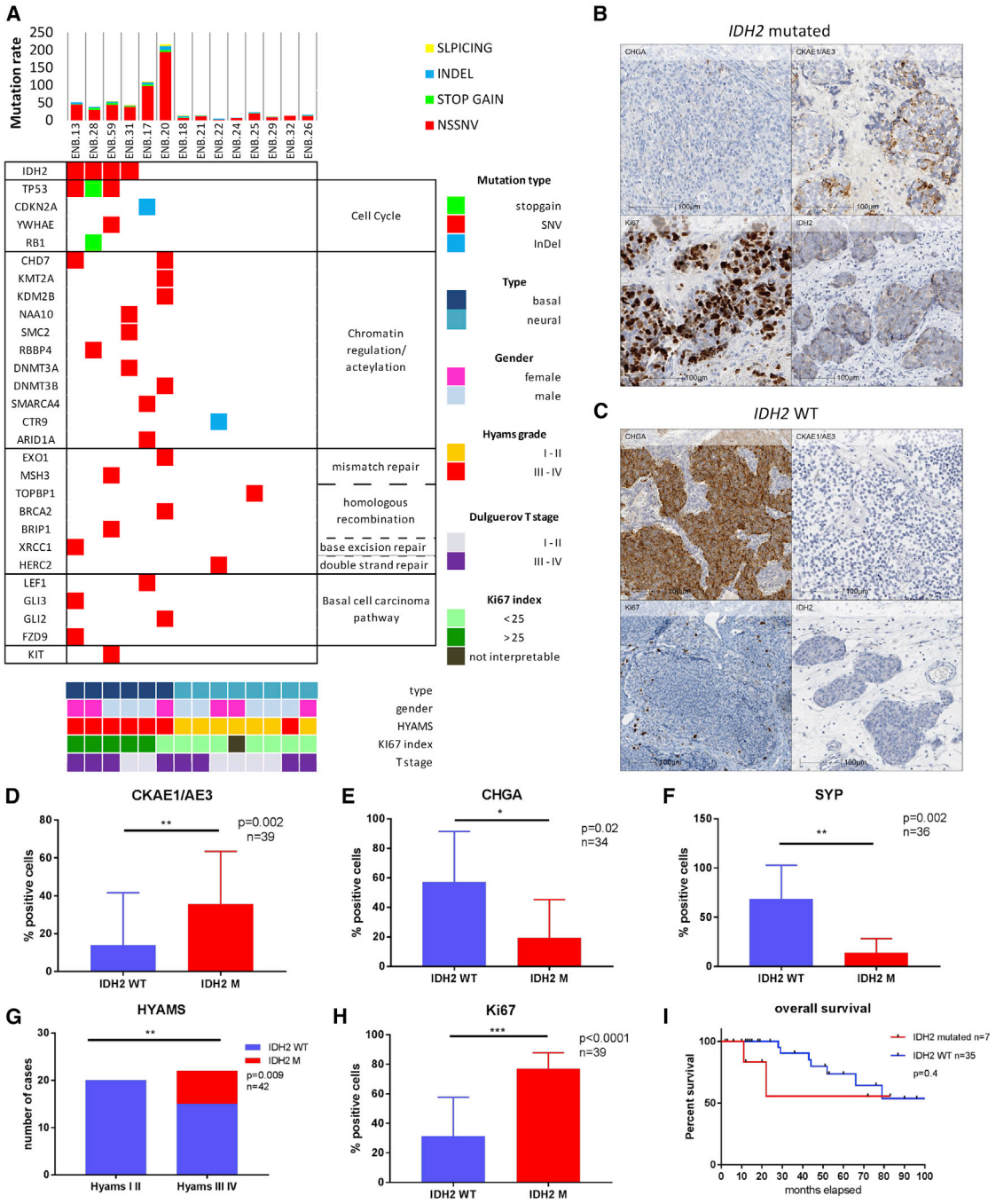


Figure 1. Landscape of Somatic Mutations in ENBs and Clinicopathological Characteristics of *IDH2*-Mutated Cases

(A) Heatmap showing individual mutations in 14 patient samples, color-coded by type of mutation. Upper panel: histogram showing the number of mutations in each sample. Rightmost column: annotation of clinicopathological tumor characteristics and patient gender. Bottom panel: heatmap of clinicopathological tumor characteristics and patient gender.

(B) Immunohistochemical profile of *IDH2*-mutated case showing focal expression of chromogranin A (CHGA), heterogeneous expression of cytokeratins in tumor cells, high Ki67 proliferation index, and cytoplasmic expression of the IDH2 mutant protein.

(C) Immunohistochemical profile of *IDH2* wild-type case showing a diffuse granular cytoplasmic expression of CHGA, no expression of cytokeratins, low-to-moderate Ki67 proliferation index, and no expression of the IDH2 mutant protein.

(D) Bar graph showing significantly higher expression of CKAE1/AE3 in *IDH2*-mutated cases compared to wild-type ENBs (Mann-Whitney test, p = 0.002).

(E) Bar graph showing a significantly lower expression of CHGA in *IDH2*-mutated cases compared to wild-type ENBs (Mann-Whitney test, p = 0.02).

(legend continued on next page)

features. We discovered two robust subtypes linked to cell ontogeny that displayed specific transcriptomic and epigenetic signatures and were associated with distinct outcomes.

RESULTS

Samples, Clinical Data, and Analytic Approach

Fifty-nine surgically resected primary ENBs were collected and re-evaluated histopathologically by two head and neck pathologists to confirm the diagnosis and establish a consensus Hyams grade (M.W. and M.C.). Immunohistochemical staining for neuroendocrine markers (i.e., chromogranin and synaptophysin), the Ki67 proliferation marker and cytokeratin were performed using whole slides. DNA (n = 42) and RNA (n = 21) were extracted from fresh-frozen samples when available; otherwise, DNA was extracted from formalin-fixed paraffin-embedded samples (n = 17). Matched germline DNA from blood or adjacent normal tissue was also collected in 14 cases. Whole-exome sequencing was performed for 27 ENBs, including a training set of matched tumor-normal samples (n = 14) and a validation set of 13 ENB samples without matching normal samples. DNA extracted from the remaining samples was used for targeted validation. RNA sequencing was performed on 19 ENB cases; in addition, DNA methylation was analyzed using Infinium EPIC arrays on 27 ENB cases. Detailed clinical and pathologic characteristics of the cohort matched those of the general population of patients with surgically resected ENB (Malouf et al., 2013) (Table S1). The median follow-up time for the 59 patients was 79.1 months; 12 patients had died at last follow-up. Among the clinical variables, only higher Hyams grades (III–IV versus I–II) were associated with significant poor overall survival (OS) time (hazard ratio [HR] = 0.19; 95% confidence interval [CI]: 0.06–0.59; p = 0.02); no difference in terms of OS was observed according to the patient's sex, age, or tumor stage.

Landscape of Genomic Alterations

Among the 14 matched ENBs for which we performed paired whole-exome sequencing, 623 somatic non-silent mutations implicating 575 different genes were identified (Table S2). The median mutational load was 20 somatic mutations per sample (range, 5–217). Nonsynonymous single nucleotide variants were the most frequent mutations identified (86%). Hotspot *IDH2* R172 codon (n = 4, 28.6%) and *TP53* (n = 3; 21.4%) mutations were the most recurrent mutations identified in more than 10% of cases and in those that co-occurred (p = 0.01) (Figure 1A). *TP53* stopgain mutation in ENB-28T was confirmed by immunohistochemistry, showing a loss of expression (Figure S1A). Conversely, immunohistochemistry showed that the two remaining nonsynonymous *TP53* mutations in ENB-59T and ENB-13T were responsible for the overexpression of the protein (Figures S1B and S1C). Exploring the biological functions related to the most relevant mutations, we also identified somatic mutations in DNA repair genes, which were altered in 35.7% (n = 5) of cases. Similar to other cancer subtypes, mutations of genes involved in chromatin regulation were identified in 42.8% (n = 6) of cases, including one sample which harbored deleterious mutations of members of SWI/SNF complex *ARID1A* and *SMARCA4*. In addition to *TP53*, we identified mutations in genes related to the cell cycle and basal cell carcinoma pathways that were altered in 28.6% (n = 4) and 35.7% (n = 5) of cases, respectively (Figure 1A). For two cases, we were able to analyze intratumoral heterogeneity by sequencing two distinct sections from the primary tumor (ENB-18T and ENB-59T). We found 56% and 93% of shared mutations, respectively (Table S2). ENB-59T displayed not only both *IDH2* mutations but also *TP53* and *KIT* mutations, which were common in both samples, indicating these mutations could be truncal events. To identify recurrent focal copy number alterations, GISTIC analysis was performed; no recurrent copy numbers gain and loss alteration were identified (Figure S2).

Sequencing an independent dataset of 13 ENB cases for which germline DNA was not available, we identified one additional case of *IDH2* hotspot mutation (Figure S1D; Table S3). After filtering out mutations not previously reported in polymorphism databases and belonging to the Cancer Gene Census catalog (n = 719), 4 genes were mutated in more than 10% of cases. These include *MUC16* (n = 5; 18.5%), *ARID1A* (n = 4; 14.8%), *KMT2D* (n = 3; 11.1%), and *NUMA1* (n = 3; 11.1%) (Figures S1D and S1E). Thus, 22.2% (n = 6) of ENB harbored mutations affecting SWI/SNF members (*ARID1A*, *SMARCA4*, and *SMARCC1*) (Figures S1D and S1E); those were predicted to lead to deleterious proteins and were mutually exclusive with *IDH2* mutations. Likewise, mutations affecting several H3K4 methyltransferases *(KMT2D*, *KMT2A*, and *KMT2C*) were identified in 22.2% (n = 6) of cases (Figures S1D and S1E; Table S3).

Prevalence and Clinicopathological Features of *IDH2*-Mutant ENBs

To explore the prevalence of *IDH2* mutations, pyrosequencing of an independent dataset of 15 ENB cases identified 2 additional cases; thus, *IDH2* R172 mutations were present in 7 of 42 ENB cases (16.7%). All *IDH2* mutations were validated by immunohistochemistry. At the pathological level, *IDH2*-mutant tumors harbored higher expression of cytokeratins (p = 0.002) and lower expression of neuroendocrine markers chromogranin A (p = 0.02) and synaptophysin (p = 0.002) (Figures 1B and 1C). *IDH2* mutations were found exclusively in tumors with higher grade (p = 0.009) and were characterized by higher Ki67 proliferation index (p < 0.0001) (Figures 1D–1H) and frequent necrosis (p = 0.04) (Table S4). Age, sex, and clinical stage were not different between *IDH2* wild-type and *IDH2*-mutated cases. No difference in overall survival was observed between cases with or without *IDH2* mutations (Figure 1I).

(F) Bar graph showing a significantly lower expression of synaptophysin in the *IDH2*-mutated cases compared to wild-type ENBs (Mann-Whitney test, p = 0.002).
(G) Stacked bar graph showing significant enrichment of *IDH2* mutations in ENB cases displaying a high Hyams histological grade (Fisher's exact test, p = 0.009).
(H) Bar graph showing a significantly higher Ki67 proliferation index in the *IDH2*-mutated ENB compared to *IDH2* wild-type (Mann-Whitney test, p < 0.0001).
(I) Kaplan-Meier survival curves showing no difference in overall survival between patients with tumors harboring *IDH2* wild-type and those with *IDH2* mutations (p = 0.4)

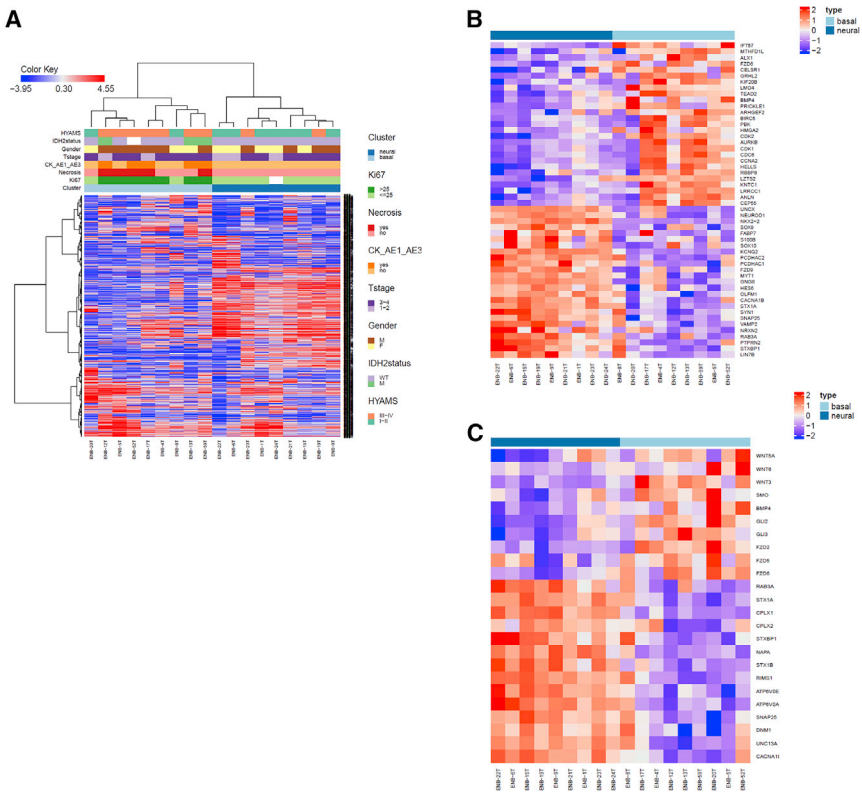


Figure 2. Subtype Classification of Esthesioneuroblastomas

(A) Unsupervised hierarchical clustering using the top-1,500 most variable genes, revealing two subgroups: basal (left) and neural (right). Upper panel: clinicopathological tumor characteristics according to the subtype classification.

(B) Heatmap depicting genes enriched for GO terms that were more significantly enriched in the neural ENB subtype than in basal ENB subtype.

(C) Heatmap depicting genes implicated in the KEGG pathways that were more significantly enriched in the neural ENB subtype than in basal ENB subtype.

Subtype Transcriptomic-Based Classification

We then considered whether ENBs display a unique gene expression signature compared to that of other sinonasal carcinomas and normal mucosa. We performed an unsupervised hierarchical clustering of gene expression of 19 ENBs, 3 intestinal-type adenocarcinomas, 3 squamous cell carcinomas, and 4 normal mucosa samples. All but one ENB sample (ENB-14T) gathered indicated that ENBs are unique at the transcriptomic level (Figure S3A). We used principal component analysis to confirm these results (Figure S3B). To clarify ENB cell ontogeny, we performed unsupervised clustering using a dataset from The Cancer Genome Atlas (TCGA) comprising 29 various cancer subtypes arising within different organs. We found that ENB clustered with all tumor types derived from the neural crest, namely glioblastomas, low-grade gliomas, pheochromocytomas, and paragangliomas (Figures S3C and S3D).

To establish a subgroup classification of ENB, we performed unsupervised clustering of ENBs (excluding ENB-14T), which revealed two distinct subgroups that we named basal and neural based on gene ontology (GO) and functional annotation analysis (Figure 2A). Overall, 1,220 genes were upregulated in the neural subtype compared to the basal subtype (fold change [FC] ≥ 2; false discovery rate [FDR] < 0.1). GO analysis revealed that those genes were enriched for neuron modules, such as neurotransmitter secretion ($p = 2.9 \times 10^{-14}$) and nervous system development ($p = 8.2 \times 10^{-10}$) (Figure 2B). Consistent with these findings, KEGG pathway analysis revealed enrichment of those genes for the synaptic vesicle cycle ($p = 1.1 \times 10^{-6}$) and neuroactive ligand-receptor interaction ($p = 1.2 \times 10^{-4}$) (Figure 2C). Conversely, 614 genes were overexpressed in basal compared to neural ENBs (FC ≥ 2; FDR < 0.1). GO analysis revealed that those genes were enriched for cell division ($p = 4.3 \times 10^{-19}$), cell proliferation ($p = 1.05 \times 10^{-7}$), embryonic morphogenesis ($p = 5.7 \times 10^{-5}$), neural tube closure ($p = 1.75 \times 10^{-4}$), and embryonic cranial skeleton morphogenesis ($p = 4.03 \times 10^{-4}$) (Figure 2B). KEGG pathway analysis was consistent with activation of the cell cycle ($p = 4.6 \times 10^{-5}$) and basal cell carcinoma pathways ($p = 0.002$) in basal ENBs (Figure 2C). Interestingly, only four genes (4.9%) belonging to proliferation gene set shared an embryonic signature.

Associations between the two schemes for transcriptomic classification and clinicopathological tumor features revealed that neural ENBs were often well differentiated, with high expression of the endocrine markers chromogranin ($p = 0.01$) and synaptophysin ($p = 0.08$) and the presence of S100 protein-positive sustentacular cells ($p = 0.0005$) and absence of necrosis ($p = 0.01$) (Table 1). In contrast, basal ENB tumors, which comprise less differentiated cells, were often high-grade tumors ($p = 0.06$) characterized by high mitotic rate ($p = 0.01$) and increased Ki67 proliferation marker ($p = 0.003$) as compared to neural ENBs (Table 1). The expression of CKAE1/AE3 was observed only in basal-type ENBs ($p = 0.03$).

Putative Cells of Origin of ENB Subtypes

Molecular differences between the ENB transcriptomic subtypes should provide clues to their origin. Thus, we examined the top-700 differential mRNAs associated with ENB transcriptomic subtypes, using a public dataset from the FANTOM consortium of 850 profiles representing various human cell and tissue specimens (Figure 3A). The expression pattern showed that the neural ENB subtype was associated with tissues related to the CNS (Figure 3A). To infer a link between ENB subtypes and cells composing olfactory mucosa, we took advantage of transcriptomic profiles generated in mice olfactory cells that described 13 putative cells ranging from the more undifferentiated horizontal basal cells (HBCs), globose basal cells (GBCs), and immature neuron progenitor type 1 (INP1) to more mature cells, such as the

Table 1. Association between Clinicopathologic Features of ENB and the Transcriptomic Subtype Classification in Neural and Basal Categories

	Neural (n = 9)	Basal (n = 9)	p Value
Age			p = 0.7
Mean	46.1 ± 5.1	49.2 ± 5.71	
Sex			p = 0.6
Male	4	6	
Female	5	3	
Ki67			p = 0.003
Mean	15 ± 3.2 (n = 8)	54.6 ± 9.7 (n = 9)	
NA	n = 1	n = 0	
Mitoses/2.5 mm^2			p = 0.01
Median	2 ± 2 (n = 9)	27 ± 13 (n = 8)	
NA	n = 0	n = 1	
Necrosis			p = 0.01
Yes	0	5	
No	9	4	
Chromogranin A (% tumor cells + by IHC)			p = 0.01
Median	75 ± 10 (n = 8)	5 ± 2.5	
NA	n = 1	n = 1	
Synaptophysin (% tumor cells + by IHC)			p = 0.08
Median	80 ± 10 (n = 9)	20 ± 19	
NA	n = 0	n = 1	
S100 sustentacular cells (% + by IHC)			p = 0.0005
Median	70 ± 30 (n = 9)	5 ± 2.5	
NA	n = 0	n = 1	
CKAE1/AE3 (% tumor cells + by IHC)			p = 0.03
Median	0 ± 0	5 ± 5	
NA	n = 0	n = 0	
Hyams grade			p = 0.06
I, II	7	2	
III, IV	2	7	
Intratumoral CD20+ cells/mm^2			p = 0.04
Median	0 ± 0 (n = 9)	4 ± 4	
NA	n = 0	n = 1	
Intratumoral CD4+ cells/mm^2			p = 0.02
Median	11 ± 8 (n = 9)	59.5 ± 43	
NA	n = 0	n = 1	
Intratumoral CD8+ cells/mm^2			p = 0.004
Median	4 ± 4 (n = 9)	50 ± 39.5 (n = 8)	
NA	n = 0	n = 1	
Modified Kadish			p = 0.6
A–B	2	1	
C–D	7	8	
Dulguerov stage			p = 0.9
T1–T2	3	3	
T3–T4	6	6	

NA, not applicable.

(*Continued on next page*)

immature neuron progenitor types 2 and 3 (INP2-3) (Fletcher et al., 2017). Notably, inter-correlations between those olfactory mucosal cells and ENB profiles identified similarities between basal ENB from one side and GBC and INP1 immature cells from the other side (n = 4/9), all having relatively higher *IDH2*, *NOTCH2*, *MYC*, *MYB*, *GRHL2*, *MKi67*, and *KIT* expression (Figures 3B and S4A). Overexpression of *KIT* in basal ENB was confirmed by immunohistochemistry (Figure S4B). We also observed similarities between a subset of neural ENBs (n = 4/9) from one side and INP2-3 cells, iOSN, and mOSN from the other side, all having relatively higher chromogranin (*CHGA*), synaptophysin (*SYP*), neural cell-adhesion molecule 1 (*NCAM1*), and neuronal differentiation 1 (*NEUROD1*) expression.

Proteomic-Based ENB Subtype Classification

Strengthened by the findings that basal ENB expressed cytokeratin (CKAE1/AE3) and a high level of Ki67 (≥25%) compared to neural ENB, which showed high expression levels for chromogranin (CHGA) and the presence of S100 sustentacular cells, we performed unsupervised protein-based clustering of the 18 ENB cases assessed by RNA-seq using the expression levels of the four following markers: CHGA, cytokeratins, Ki67, and S100 protein. Consistent with transcriptome-based classification, protein-based clustering identified two subgroups that were highly correlated with transcriptome-based clustering ($p = 0.002$) (Figure 4A). We then performed unsupervised protein-based clustering using the four markers in the whole cohort (n = 51). We observed that the ENB samples clustered into two groups, consistent with the basal (43.1%; n = 22) and neural (56.9%; n = 29) subtype classification (Figure 4B). Importantly, basal ENBs were less differentiated ($p < 0.0001$), had a higher tumor grade ($p < 0.0001$), and increased Ki67 index ($p < 0.0001$), as well as an overexpression of cytokeratins ($p < 0.0001$) compared to neural ENBs. Conversely, the latter ones displayed higher chromogranin and synaptophysin expression ($p < 0.0001$) and were enriched for the presence of S100 sustentacular cells ($p < 0.0001$). There was no difference in terms of the TNM stage between the two subgroups. Puzzlingly, the basal-type cluster was enriched in males ($p = 0.01$) (Figure 4C). Finally, using data for ENB cases where paired exome sequencing was performed, we found that the basal ENB samples harbored a higher mutation load compared to the neural ENB samples ($p = 0.02$). Patients with basal ENBs had shorter median overall survival of 73 months compared to patients with neural ENBs (median not reached) ($p = 0.008$) (HR = 6.2; 95% CI: 1.7–22.1) (Figure 4D).

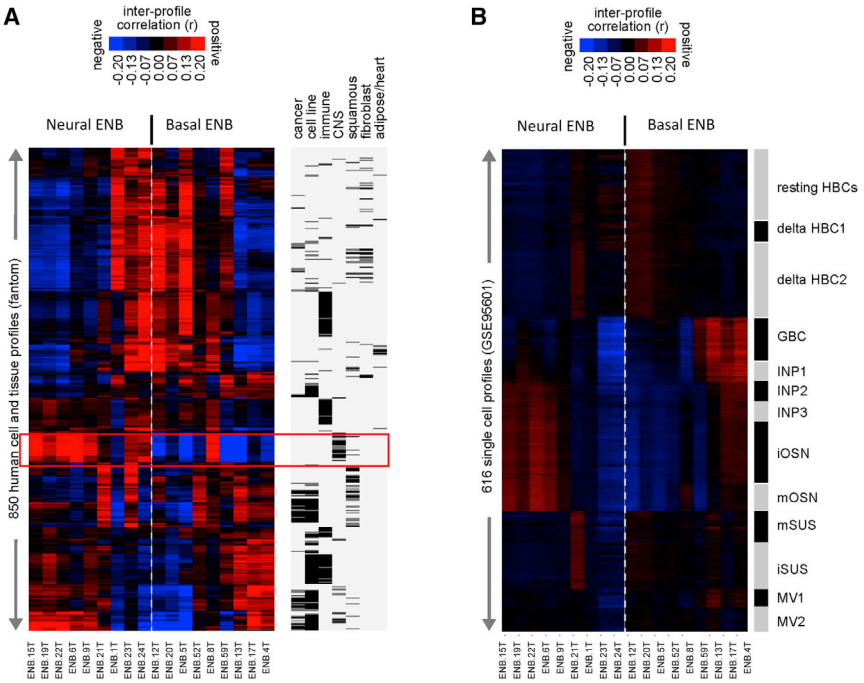


Figure 3. Gene Signatures Distinguishing ENB Transcriptomic Subtypes Are Linked to Specific Tissue Types or Cells of Origin within Olfactory Mucosa

(A) Heatmap showing inter-sample correlations (orange, positive) between ENB differential expression profiles (columns) and profiles from the Fantom consortium expression dataset of various cell types or tissues from human specimens (rows, n = 850 profiles). Membership of the Fantom profiles in general categories of "cancer," "cell line," "immune" (immune cell types or blood or related tissues), "CNS" (related to central nervous system including brain), "squamous" (including bronchial, trachea, oral regions, throat and esophagus regions, nasal regions, urothelial, cervix, sebocyte, keratin/skin/epidermis), "fibroblast," or "adipocyte/heart" is indicated. The red rectangle underlies the similarities between ENB profile and CNS cells profile.

(B) Heatmap showing inter-sample correlations (orange, positive) between differential expression profiles of ENBs (columns) and differential profiles of olfactory mucosal cell types (rows, GSE95601 dataset). HBC, horizontal basal cell; GBC, globose basal cell; INP, immature neuron progenitors; iOSN, immature olfactory sensory neuron; mOSN, mature olfactory sensory neuron; mSUS, mature sustentacular cell; iSUS, immature sustentacular cell; MV, microvillous cell.

Discovery of E-CpG Island Methylator Phenotype in *IDH2* R172 ENB

To identify specific and global differences between the genome-wide methylation profiles of ENB subtypes, we analyzed 27 ENB samples (5 of which displayed *IDH2* mutation) using Infinium EPIC arrays. The Illumina EPIC DNA methylation array was chosen because it provides a genome-scale interrogation of ~850,000 probes in the human methylome, both in CpG islands (CGIs) and outside CGIs. First, we focused our analysis on CGIs because of their functional significance and the strong correlations between CGIs and transcription initiation (Deaton and Bird, 2011). Hierarchical clustering analysis using M-values of the 1000 most differentially methylated probes located in CGIs revealed two epi-clusters that were highly correlated with our protein-based classification (p = 0.0001) (Figure 5A). Remarkably, all the *IDH2*-mutant cases clustered together, consistent with an ENB CGI methylator phenotype (E-CIMP) that is reminiscent of CIMP observed in other tumors such as glioblastomas and colon cancers. Overall, using stringent criteria (β-value < 0.2 in *IDH2* wild-type and ≥ 0.5 in *IDH2*-mutant), 6827 probes were found to be hypermethylated in E-CIMP, compared to only 7 probes in E-CIMP-negative samples. GO analysis revealed that hypermethylated genes in E-CIMP were enriched for chemical synaptic transmission ($p = 7.9 \times 10^{-17}$), negative regulation of neuron differentiation ($p = 8.4 \times 10^{-9}$) and neuropeptide signaling pathways ($p = 9.4 \times 10^{-9}$), suggesting that *IDH2* mutation inhibits neural differentiation through epigenetic genome-wide reprogramming. To determine whether methylome reprogramming of *IDH2*-mutant ENB was specific to CGI regions, we analyzed the enrichment of differentially methylated probes according to their CpG genomic and functional annotations in *IDH2*-mutant versus wild-type ENBs. Importantly, no enrichment was observed according to genomic distributions of CpG sites in islands, shores or shelves (Figure 5B). In contrast, we observed an enrichment of CpG sites located in exon 1 (3.2-fold enrichment) (Figure 5B) that was tightly linked to transcriptional gene silencing. GO analysis of methylated genes enriched in exon 1 were enriched for chemical synaptic transmission ($p = 3.5 \times 10^{-12}$), neuropeptide signaling pathway ($p = 6.7 \times 10^{-9}$) and nervous system development ($p = 1.4 \times 10^{-5}$) pathways, consistent with blockade of neuronal differentiation. To determine whether E-CIMP has a methylome similar to that of *IDH1-2*-mutant glioblastoma (G-CIMP), we performed supervised clustering using differentially methylated probes between E-CIMP and other ENB samples. We found that G-CIMP samples clustered according to their *IDH1-2* status, which implied that similar epigenetic programing occurs in *IDH1-2*-mutant tumors (Figure S5).

Methyl Divergence of ENB Subtypes

To analyze methylation changes related to cell ontogeny, we excluded E-CIMP samples from further analysis and performed unsupervised clustering using differentially methylated probes of the remaining ENB samples (n = 22) (Figure 5C). We observed that the basal samples showed greater clustering than the neural samples. Overall, 55,824 (6.7%) autosomal genes were significantly differentially methylated between the neural and basal clusters (Δβ-value ≥ 0.2 or ≤ −0.2, FDR < 0.1); 30,576 (3.6%) probes were hypermethylated and 25,248 (2.9%) probes were hypomethylated in the basal compared to the neural subgroup. Among the probes that lost methylation in the neural

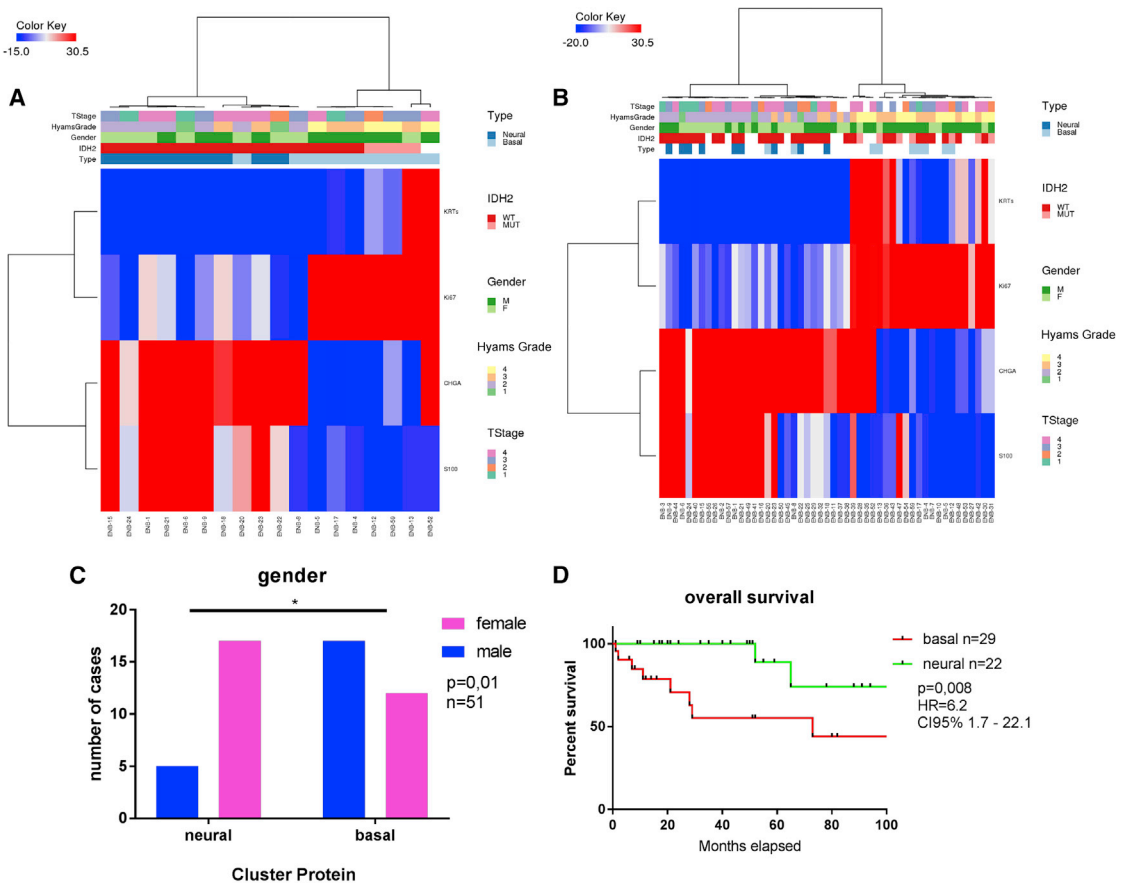


Figure 4. Protein-Based Expression Clustering of ENBs

(A) Hierarchical unsupervised clustering based on 4 protein markers (Ki67, S100, cytokeratins, and chromogranin A) applied to 18 ENBs assessed by RNA-seq. Each column represents one ENB sample, and each row represents one protein marker. Upper panel: clinicopathological tumor features, patient gender, *IDH2* status, and the mRNA-based classification subtype (neural or basal).

(B) Hierarchical unsupervised clustering based on the same 4 protein markers applied to 51 ENB samples with available data. Upper panel: clinicopathological tumor features, patient gender, *IDH2* status, and the mRNA-based classification subtype whenever available.

(C) Interval bar graph showing proportion of male and female patients with each ENB subtype within the dataset (Fisher's exact test, p = 0.01).

(D) Kaplan-Meier curves showing shorter overall survival times for a greater proportion of patients with tumors belonging to the basal cluster compared to those belonging to the neural cluster (p = 0.008).

ENB subgroup, no enrichment was identified according to genomic distribution of CpG sites. Conversely, we identified striking enrichment for probes located in enhancers (FC = 2.2) (Figure 5D). GO analysis revealed that those enhancers were enriched for signal transduction ($p = 2.2 \times 10^{-12}$), platelet activation ($p = 1.1 \times 10^{-6}$), and axonal guidance ($p = 5.2 \times 10^{-6}$).

Next, we concentrated on probes located in CGIs because of their known correlation with gene expression. Overall, for the 1,377 probes differentially methylated between the ENB subtypes ($\Delta\beta$-value ≥ 0.2 or ≤ -0.2, FDR < 0.1), we observed a higher methylation of probes located in CGIs in the basal ENB (mean $\beta = 0.38 \pm 0.005$) than in the neural ENB (mean $\beta = 0.30 \pm 0.006$) ($p < 0.0001$) (Figure S6A). We also observed significant enrichment for probes with low-level methylation in the neural subtype (Figure S6B). GO enrichment analysis showed that promoters that lost DNA methylation in the neural subgroup compared to the basal subgroup were related to genes involved in chemical synaptic transmission ($p = 6.9 \times 10^{-8}$); furthermore, this was associated with their increased expression (Figure S6C).

Charting Immune Tumor Profiles

The significance of T lymphocyte infiltration in ENB is not known. To investigate the immune microenvironment of ENBs, we explored CD4 (+) and CD8 (+) tumor-infiltrating lymphocytes (TILs) in 42 ENBs classified as neural (n = 25) or basal (n = 17). We found that the mean counts of both CD4+ and CD8+ T cells were significantly higher in basal ENBs than in neural ENBs ($p = 0.007$ and $p < 0.0001$, respectively) (Figures 6A–6F). *CD8A* mRNA expression levels assessed by RNA sequencing (RNA-seq) were highly correlated with CD8+ TILs (Pearson r = 0.8; p = 0.0001). We then investigated the expression of cytotoxicity markers, inflammatory cytokines, and immune checkpoints according to our two subtypes. Consistent with CD8 infiltration

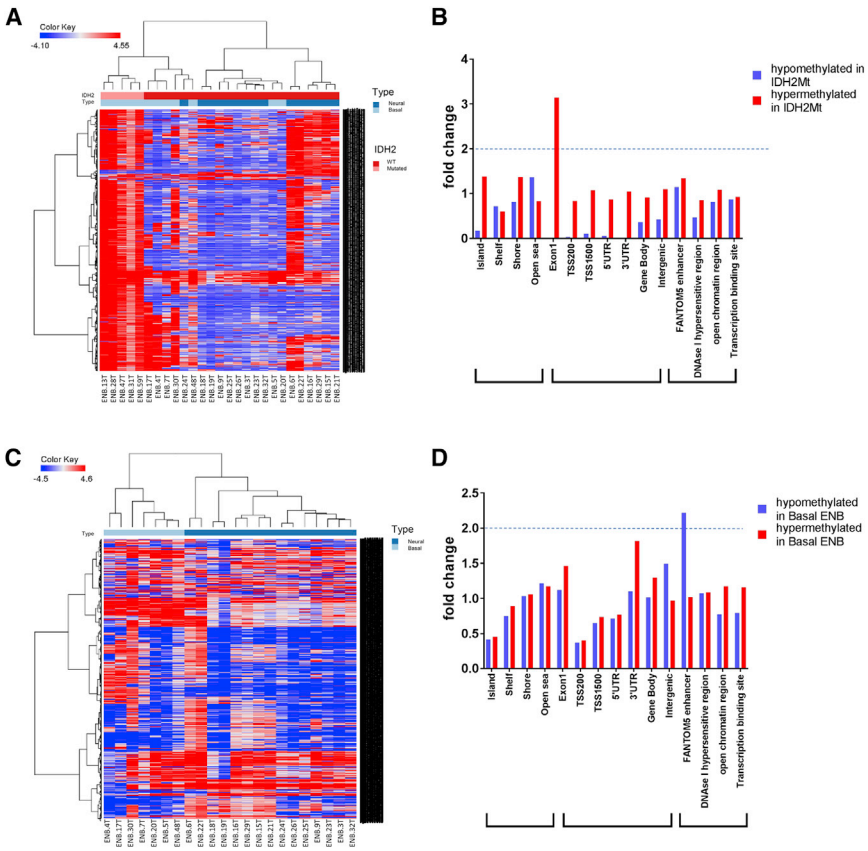


Figure 5. Methylation Landscape of ENB Subtypes

(A) Unsupervised hierarchical clustering of ENB samples (n = 27) using M-values of 1000 most variable probes in CpG islands. All *IDH2* mutated ENB cases clustered together and display a CpG island methylator phenotype (E-CIMP).

(B) Bar graph showing enrichment of genomic and functional distribution of differentially methylated probes between *IDH2* mutated and *IDH2* wild-type ENBs.

(C) Unsupervised hierarchical clustering of ENB samples without *IDH2* mutations (n = 22) according to the most variable 8,000 probes (using M values), showing that neural and basal ENB samples cluster together.

(D) Bar graph showing the enrichment of genomic and functional distribution of methylated probes that differentiates between neural and basal ENBs. Note the striking enrichment for enhancers in probes losing methylation in neural ENBs compared to basal ENBs.

in the basal subtype, we also identified higher expression levels of cytotoxic cell markers (*GZMA, GZMB*), T cell invasion chemokine factors (*CXCL13*, *CXCL10*, and *CXCL9*), immune checkpoints (notably PD1 [*PDCD1*] and its ligands PD-L1 [*CD274*], PD-L2 [*PDCD1LG2*], *CTL4A*, *ICOS*, TIM3 [*HAVCR2*], and *LAG3*), and suppressive factors (notably *IDO1*, *TGFB*, *IL-10*, and *FOXP3*) (Figure 6G).

To further estimate tissue-infiltrating immune and other stromal populations' abundances using gene expression, we used MCP counter. We found higher cytotoxic cells lymphocytes ($p = 0.06$), fibroblast ($p = 0.01$), and more monocytic cells ($p = 0.01$) in basal versus neural ENB. These data highly corroborate the findings observed by immunohistochemistry (IHC) for CD8+ cells. To consider possible drug targets, we explored the PD-L1 expression levels in our dataset and found that PD-L1 was expressed by both tumor cells and immune cells, predominantly within tumor areas (Figure 6H). Of 36 evaluable cases, 14 cases (39%) showed positive staining ($\geq$1%), including 9 cases with PD-L1 expression levels ranging from 5%–50%. Importantly, 64% of basal ENBs showed high PD-L1 expression ($\geq$1%), compared to 27% of neural ENBs ($p = 0.06$) (Figure 6I).

DISCUSSION

This study represents, to our knowledge, the first integrative analysis of ENBs using a multi-omic approach, which led to our classifying these tumors into two major subtypes with distinct clinicopathological features. In addition, it provides a link between distinct development stages of normal olfactory mucosa and ENB tumor subtypes defined by gene expression profiling. This model is consistent with carcinogenesis in other organs, where normal development of the primary organ has been shown to meet cancer "intrinsic" subtypes (Prat and Perou, 2009). We note that in a study of young rats, GBCs, but not HBCs, are the neuronal progenitors of the olfactory epithelium (Caggiano et al., 1994). Herein, the immature GBC cell signature was found to be highly enriched in the basal subtype. Conversely, the more mature neural signature was found to be enriched in the neural ENB subtype. These findings provide support for the cancer stem cell (CSC) paradigm. According to this model, a heterogeneous population of cancer cells can develop from the transformation of normal stem cells or progenitor cells (tumors cells that exhibits stem cells properties) (Ebben et al., 2010). Alternatively, a differentiated cancer cell may acquire stem cell properties and develop into a heterogeneous tumor population (Prat and Perou, 2009). However, in both situations, the CSC paradigm states that the bulk of the tumor cells exhibit limited proliferation potential; whereas enrichment with a CSC compartment favors tumor aggressiveness and contributes to treatment resistance, which is a characteristic of the basal ENB subtype. Thus, the different profiles displayed by neural and basal ENBs could be explained by two hypotheses. The first one is that basal ENBs derive from basal cells (HBC and/or GBC) while neural ENBs derive from immature olfactory neuron progenitors (INP). The second one is that basal ENBs are stuck in a basal state, whereas neural ENBs manage to initiate an incomplete neuron maturation process. Mechanistic data ore needed to decipher the molecular basis linking ontogeny with oncogenesis in these tumors. In addition, we

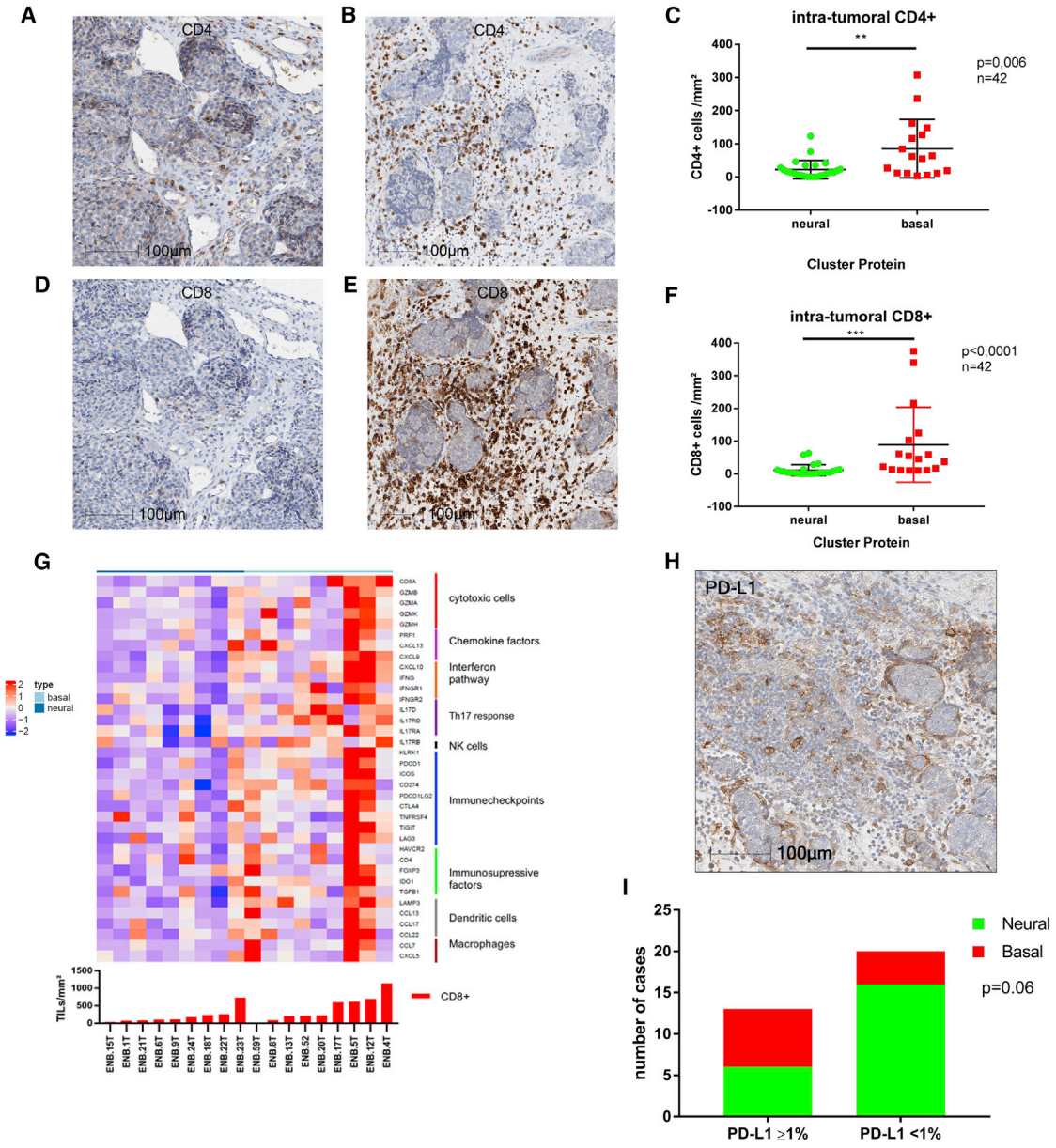


Figure 6. Immune Microenvironment of ENBs

(A–C) Immunohistochemical illustrations (A and B) and scatterplot (C) showing a tendency toward greater number of intratumoral CD4+ T cells in basal ENBs than in neural ENBs ($p = 0.06$).

(D–F) Immunohistochemical illustrations (D and E) and scatterplot (F) showing a greater number of intratumoral CD8+ T cells in basal ENBs than in neural ENBs ($p < 0.0001$).

(G) Heatmap of immune genes in neural and basal ENBs. ENB cases are ordered by CD8+ cell count in each subtype. The number of tumor-infiltrating lymphocytes (TILs) was not available for ENB-59T case.

(H) Example of the immunohistochemical expression of PD-L1 in a basal-type ENB.

(I) Bar graph showing enrichment of PD-L1 expression in basal compared to the neural ENB subtype ($p = 0.06$).

must keep in mind the experimental limitations of the study of normal stem cells and CSCs, specifically in ENBs because gene expression profiles were obtained from subpopulations of cells with different tumor grades and at different putative stages of differentiation in the neuronal hierarchy within the human olfactory mucosa.

At the genetic level, we discovered that ~17% of ENBs harbored the *IDH2* R172 hotspot mutation previously reported in a subset of glioblastomas (Balss et al., 2008), leukemia (Marcucci et al., 2010), and osteosarcomas, expanding the spectrum of cancerous organs with such oncogenic mutations (Dang and Su, 2017). We note that ~5% of ENBs in the cohort evaluated by

Gay et al. (2017) also harbored *IDH2* mutations. However, as tumor grade was not reported in that study, we cannot rule out the possibility that this might confound the prevalence of *IDH2* mutations (Gay et al., 2017). More importantly, in our cohort, *IDH2* mutations were found exclusively in almost one third of basal ENBs, and those samples displayed unique pathological features. Recently, *IDH2* mutations were also reported in the vast majority of sinonasal undifferentiated carcinomas and in variable proportions of other poorly differentiated sinonasal carcinomas (Dogan et al., 2017; Jo et al., 2017). As these tumors might overlap pathologically, future efforts are needed to better classify the *IDH2*-mutated subset. Beyond diagnostic implications, these findings have significant implications for therapy with IDH inhibitors, which have been recently approved to treat acute myeloid leukemia with a companion diagnostic along with the drug for *IDH2* mutation detection (Mullard, 2017).

In our study, we found that *IDH2* mutation underlies global epigenomic divergence in ENB with pervasive DNA hypermethylation enriched in enhancers of axonal guidance genes, possibly leading to failure of GBCs to differentiate into the neuronal lineage, which is consistent with previous studies that demonstrated that *IDH* mutations result in a block of cell differentiation *in vitro* and *in vivo* (Lu et al., 2012, 2013). Furthermore, our results suggest a mechanism for cancer-related DNA hypermethylation that may be similar to the one proposed in glioblastomas and leukemia. Generation of a novel genetically engineered mice model by inducing *IDH2* mutations on the relevant targeted cell populations might help clarify the cell of origin of these tumors and allow mechanistic studies.

A key finding of our study is the similarity between gene expression in ENBs and that in most tumors arising from neural crest cells (Maguire et al., 2015). Those tumors are also characterized by mutations of genes involved in the Krebs cycle, as is the case for *SDH* mutations in gastrointestinal stromal tumors and paragangliomas and pheochromocytomas.

Finally, investigation of the immune microenvironment in ENBs sheds a light on the divergent infiltration of TILs between basal and neural ENBs. Despite the higher rate of intratumoral CD8+ T cells, basal ENBs were associated with worse prognosis, which might be due to activation of the checkpoint inhibitors, providing a rationale for the use of immune checkpoint inhibitors in this setting. Whether a higher mutational load in basal ENBs can be linked to immune infiltration is worthy of analysis in a larger dataset.

We describe the multi-omic analysis of ENBs, which allowed us to chart a two-subtype classification of ENBs. This study provides a framework for future studies in this area, with clinical implications for the diagnosis, treatment, and management of patients with ENBs. We linked ontogeny with oncogenesis, which might have implications for our understanding of cancer development in general.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODELS AND SUBJECTS DETAILS
- METHOD DETAILS
  - Sample collection and histopathological analysis
  - Immunohistochemistry procedures
  - Quantification of IHC markers
  - Nucleic acid extraction
  - Whole-exome sequencing and somatic mutation detection
  - Validation of IDH2 mutations
  - RNA sequencing and bioinformatics analysis
  - Unsupervised classification of gene expression
  - Analysis of differentially expressed genes
  - Pathway enrichment analysis
  - Estimation of immune infiltrate
  - DNA methylation and bioinformatics analysis
- QUANTIFICATIONS AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and four tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.09.047.

ACKNOWLEDGMENTS

This work was sponsored by Fondation AVEC. The authors thank Professor Homa Adle-Biasette and the Lariboisière Biological Ressources Center. The authors also thank Dr. Karima Mokthari, a neuropathologist at the La Pitié Salpétière Pathology Department. This work was supported, in part, by grants from Fondation AVEC, the University of Texas MD Anderson Cancer Center via a CCSG (P30 CA016672) (to X.S.), the University of Texas MD Anderson Cancer Center via a Leukemia SPORE grant (P50 CA100632) (to X.S.), and the U.S. NIH through the CCSG Bioinformatics Shared Resource at MD Anderson Cancer Center.

AUTHOR CONTRIBUTIONS

Conceptualization, M.C., A.B., M.W., X.L., B.V., G.M., F.B., C.L., J.-E.K., D.K. X.S., and G.G.M.; Methodology and Investigation, M.C., H.Y., R.M, C.J.C., and F.A.; Formal Analysis, M.C., H.Y., X.S., C.J.C., M.C., R.M, and G.G.M.; Resources, A.B., M.W., X.L., B.V., G.M., F.B., J.-E.K. C.L., and D.K.; Writing – Original Draft, M.C. and G.G.M.; Bioinformatics Analysis, H.Y., X.S., and C.J.C.; Writing – Review & Editing, M.C., H.Y., R.M., C.J.C., A.B., F.A., M.W., X.L., B.V., G.M., F.B., C.L., J.-E.K., D.K., X.S., and G.G.M.; Supervision, G.G.M.; Funding Acquisition, G.G.M. and D.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: February 26, 2018
Revised: July 18, 2018
Accepted: September 12, 2018
Published: October 16, 2018

REFERENCES

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Arita, H., Narita, Y., Matsushita, Y., Fukushima, S., Yoshida, A., Takami, H., Miyakita, Y., Ohno, M., Shibui, S., and Ichimura, K. (2015). Development of a robust and sensitive pyrosequencing assay for the detection of IDH1/2 mutations in gliomas. Brain Tumor Pathol. *32*, 22–30.

Balss, J., Meyer, J., Mueller, W., Korshunov, A., Hartmann, C., and von Deimling, A. (2008). Analysis of the IDH1 codon 132 mutation in brain tumors. Acta Neuropathol. *116*, 597–602.

Becht, E., Giraldo, N.A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., Selves, J., Laurent-Puig, P., Sautès-Fridman, C., Fridman, W.H., and de Reyniès, A. (2016). Estimating tissue-infiltrating immune and other stromal populations' abundances using gene expression. Genome Biol. *20*, 218.

Bell, D., Saade, R., Roberts, D., Ow, T.J., Kupferman, M., DeMonte, F., and Hanna, E.Y. (2015). Prognostic utility of Hyams histological grading and Kadish-Morita staging systems for esthesioneuroblastoma outcomes. Head Neck Pathol. *9*, 51–59.

Caggiano, M., Kauer, J.S., and Hunter, D.D. (1994). Globose basal cells are neuronal progenitors in the olfactory epithelium: a lineage analysis using a replication-incompetent retrovirus. Neuron *13*, 339–352.

Campan, M., Weisenberger, D.J., Trinh, B., and Laird, P.W. (2009). MethyLight. Methods Mol. Biol. *507*, 325–337.

Chen, F., Zhang, Y., Gibbons, D.L., Deneen, B., Kwiatkowski, D.J., Ittmann, M., and Creighton, C.J. (2018). Pan-cancer molecular classes transcending tumor lineage across 32 cancer types, multiple data platforms, and over 10,000 cases. Clin. Cancer Res. *24*, 2182–2193.

Czapiewski, P., Kunc, M., and Haybaeck, J. (2016). Genetic and molecular alterations in olfactory neuroblastoma: implications for pathogenesis, prognosis and treatment. Oncotarget *7*, 52584–52596.

Dang, L., and Su, S.M. (2017). Isocitrate dehydrogenase mutation and (R)-2-hydroxyglutarate: from basic discovery to therapeutics development. Annu. Rev. Biochem. *86*, 305–331.

Deaton, A.M., and Bird, A. (2011). CpG islands and the regulation of transcription. Genes Dev. *25*, 1010–1022.

Dogan, S., Chute, D.J., Xu, B., Ptashkin, R.N., Chandramohan, R., Casanova-Murphy, J., Nafa, K., Bishop, J.A., Chiosea, S.I., Stelow, E.B., et al. (2017). Frequent IDH2 R172 mutations in undifferentiated and poorly-differentiated sinonasal carcinomas. J. Pathol. *242*, 400–408.

Du, P., Zhang, X., Huang, C.C., Jafari, N., Kibbe, W.A., Hou, L., and Lin, S.M. (2010). Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics *11*, 587.

Ebben, J.D., Treisman, D.M., Zorniak, M., Kutty, R.G., Clark, P.A., and Kuo, J.S. (2010). The cancer stem cell paradigm: a new understanding of tumor development and treatment. Expert Opin. Ther. Targets *14*, 621–632.

Fletcher, R.B., Das, D., Gadye, L., Street, K.N., Baudhuin, A., Wagner, A., Cole, M.B., Flores, Q., Choi, Y.G., Yosef, N., et al. (2017). Deconstructing olfactory stem cell trajectories at single-cell resolution. Cell Stem Cell *20*, 817–830.e8.

Gay, L.M., Kim, S., Fedorchak, K., Kundranda, M., Odia, Y., Nangia, C., Battiste, J., Colon-Otero, G., Powell, S., Russell, J., et al. (2017). Comprehensive genomic profiling of esthesioneuroblastoma reveals additional treatment options. Oncologist *22*, 834–842.

Huang, W., Sherman, B.T., and Lempicki, R.A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. *4*, 44–57.

Jo, V.Y., Chau, N.G., Hornick, J.L., Krane, J.F., and Sholl, L.M. (2017). Recurrent IDH2 R172X mutations in sinonasal undifferentiated carcinoma. Mod. Pathol. *30*, 650–659.

Lazo de la Vega, L., McHugh, J.B., Cani, A.K., Kunder, K., Walocko, F.M., Liu, C.J., Hovelson, D.H., Robinson, D., Chinnaiyan, A.M., Tomlins, S.A., and Harms, P.W. (2017). Comprehensive molecular profiling of olfactory neuroblastoma identifies potentially targetable *FGFR3* amplifications. Mol. Cancer Res. *15*, 1551–1557.

Lu, C., Ward, P.S., Kapoor, G.S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C.R., Khanin, R., Figueroa, M.E., Melnick, A., et al. (2012). IDH mutation impairs histone demethylation and results in a block to cell differentiation. Nature *483*, 474–478.

Lu, C., Venneti, S., Akalin, A., Fang, F., Ward, P.S., Dematteo, R.G., Intlekofer, A.M., Chen, C., Ye, J., Hameed, M., et al. (2013). Induction of sarcomas by mutant IDH2. Genes Dev. *27*, 1986–1998.

Maguire, L.H., Thomas, A.R., and Goldstein, A.M. (2015). Tumors of the neural crest: common themes in development and cancer. Dev. Dyn. *244*, 311–322.

Malouf, G.G., Casiraghi, O., Deutsch, E., Guigay, J., Temam, S., and Bourhis, J. (2013). Low- and high-grade esthesioneuroblastomas display a distinct natural history and outcome. Eur. J. Cancer *49*, 1324–1334.

Malouf, G.G., Su, X., Yao, H., Gao, J., Xiong, L., He, Q., Compérat, E., Couturier, J., Molinié, V., Escudier, B., et al. (2014). Next-generation sequencing of translocation renal cell carcinoma reveals novel RNA splicing partners and frequent mutations of chromatin-remodeling genes. Clin. Cancer Res. *20*, 4129–4140.

Marcucci, G., Maharry, K., Wu, Y.Z., Radmacher, M.D., Mrózek, K., Margeson, D., Holland, K.B., Whitman, S.P., Becker, H., Schwind, S., et al. (2010). IDH1 and IDH2 gene mutations identify novel molecular subsets within de novo cytogenetically normal acute myeloid leukemia: a Cancer and Leukemia Group B study. J. Clin. Oncol. *28*, 2348–2355.

Moran, S., Arribas, C., and Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. Epigenomics *8*, 389–399.

Mullard, A. (2017). FDA approves first-in-class cancer metabolism drug. Nat. Rev. Drug Discov. *16*, 593.

Prat, A., and Perou, C.M. (2009). Mammary development meets cancer genomics. Nat. Med. *15*, 842–844.

Su, S.Y., Bell, D., and Hanna, E.Y. (2014). Esthesioneuroblastoma, neuroendocrine carcinoma, and sinonasal undifferentiated carcinoma: differentiation in diagnosis and treatment. Int. Arch. Otorhinolaryngol. *18* (*Suppl 2*), S149–S156.

Triche, T.J., Jr., Weisenberger, D.J., Van Den Berg, D., Laird, P.W., and Siegmund, K.D. (2013). Low-level processing of Illumina Infinium DNA Methylation BeadArrays. Nucleic Acids Res. *41*, e90.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Mouse monoclonal CK AE1/AE3 clone AE1/AE3	Dako	Cat# M351501-2; RRID: AB_2631307
Mouse monoclonal CD 56 (NCAM) clone 1B6	Novocastra_Leica Biosystem	Cat# NCL-L-CD56-1B6; RRID: AB_563906
Mouse monoclonal Chromogranin A clone DAK-A3	Dako	M086901-2
Mouse monoclonal Synaptophysin clone DAK-SYNAP	Dako	Cat# M731529-2; RRID: AB_2687942
Rabbit polyclonal PS100	Dako	Z031101-2
Mouse monoclonal Ki67 clone MIB1	Dako	Cat# M724001-2; RRID: AB_2631211
Mouse monoclonal P53 clone DO-7	Dako	M700101-2
Mouse monoclonal IDH2 Ms-Mab1	Millipore	MABC1103
Rabbit monoclonal CD4 clone SP35	Ventana Medical Systems	Cat# 790-4423; RRID: AB_2335982
Mouse monoclonal CD8 clone 144B	Dako	M710301-2
Mouse monoclonal Anti-CD20 clone L26	Dako	IS60430-2
Rabbit monoclonal PD-L1 clone E1L3N	Cell Signaling Technology	Cat# 13684; RRID: AB_2687655
Biological Samples		
Primary tumor samples	Biological Ressource Center of 4 French Hospitals: CHRU-Lille, Institut Curie, Lariboisière and La Pitié-Salpétrière - Assistance Publique Hôpitaux de Paris	N/A
Matched blood samples	Biological Ressource Center of Lariboisière hospital,	N/A
Critical Commercial Assays		
ChargeSwitch gDNA Micro Tissue Kit	Invitrogen	Cat# CS11203
Infinium MethylationEPIC BeadChip kit	Illumina	N/A
Illumina barcoded Paired-end Library Preparation kit	Illumina	N/A
Deposited Data		
RNA sequencing raw and analyzed data	Illumina	GSE118995
GSE95601 single cell RNASeq signature	GEO	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE95601
FANTOM5 data repository	FANTOM project	http://fantom.gsc.riken.jp/5/data/
Oligonucleotides		
PCR IDH2: Forward primer: ACATCCCACGC CTAGTCCC	Arita et al., 2015	https://www.ncbi.nlm.nih.gov/pubmed/24748374
PCR IDH2: Reverse primer: bio-TCTCCACCC TGGCCTACCTG	Arita et al., 2015	https://www.ncbi.nlm.nih.gov/pubmed/24748374
PCR IDH2: Pyrosequencing primer: CCCATC ACCATTGGC	Arita et al., 2015	https://www.ncbi.nlm.nih.gov/pubmed/24748374
Software and Algorithms		
Venny 2.1	Juan Carlos Oliveros (oliveros@cnb.csic.es) BioinfoGP Service Centro Nacional de Biotecnología, (CNB-CSIC)	http://bioinfogp.cnb.csic.es/tools/venny/
DAVID Bioinformatics resources database (v6.7)	N/A	https://david.ncifcrf.gov/
Mutation assessor	cBio@MSKCC	http://mutationassessor.org/r3/
Aperio ImageScope	Leica, Biosystem	https://www.leicabiosystems.com/digital-pathology/manage/aperio-imagescope/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
GraphPad	GraphPad Software, La Jolla, CA, USA	https://www.graphpad.com/scientific-software/prism/
Gene set enrichment Analysis	N/A	http://software.broadinstitute.org/gsea/index.jsp
Pathcards	Weizmann Institute of Science	http://pathcards.genecards.org/
Deseq package	N/A	https://bioconductor.org/packages/release/bioc/html/DESeq.html
MCP counter	N/A	https://omictools.com/mcp-counter-tool

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Gabriel G. Malouf (gabriel.malouf@igbmc.fr).

EXPERIMENTAL MODELS AND SUBJECTS DETAILS

All tumors and clinical information were collected through a French collaborative network (see the STAR Methods) with consent as per protocol approved by the ethical committee of the Pitié-Salpêtrière Hospital (IDF-6, Ile de France). In total, 59 ENBs were collected for genomic, transcriptomic and/or proteomic analyses (Table S1). All ENBs were diagnosed according to the World Health Organization classification criteria and reviewed by two pathologists. DNA or RNA from snap frozen tumor were investigated with one or more of WES (n = 27), RNA-seq (n = 19) and methylation array (n = 26) analyses; 15 samples with DNA from formalin-fixed, paraffin-embedded materials were analyzed for *IDH2* mutations by pyrosequencing and immunohistochemistry. Exome-sequencing and RNA sequencing were done using previously described pipelines (Malouf et al., 2014). Analysis of CD8 and PD-L1 expression was also done in all samples for which tissues were available (n = 42). Significance of differences in gender, stage and immunohistochemistry markers between ENB subgroups were analyzed using a two-sided Fisher's exact test. Recurrence-free survival and overall survival were estimated using Kaplan-Meier curves. All analyses were done using GraphPad Prism. A p value of < 0.05 was regarded as significant for all analyses.

METHOD DETAILS

Sample collection and histopathological analysis

Patient samples with confirmed diagnosis of ENB were collected from several pathology departments in France after local committee approval. All cases were reviewed by two experts head and neck pathologists (M.C and M.W) who confirmed the diagnosis and reported Hyams grading; in case of discrepancy for the grading between the two pathologists, the final grade was reached by consensus. The diagnosis was confirmed using H&E stain and a panel of immunohistochemical markers performed on formalin-fixed paraffin-embedded tumor blocks. The IHC panel included at least 2 neuroendocrine (NE) markers (comprising chromogranin A, synaptophysine and CD56 (*NCAM1*)), S100 protein, pan-cytokeratin (AE1/AE3) and the proliferation marker Ki67. Cases for which the expression of CKAE1/AE3 was observed were considered as ENB if at least one NE marker was expressed by tumor cells and focal expression of S100 was observed in sustentacular type cells. When necessary, immunohistochemistry for selected markers was performed (i.e., *c-KIT, TP53, IDH2*) using the most representative paraffin block. For the cases displaying available cryopreserved material, a frozen section stained by hematoxylin eosin was used to assess tumor purity; only cases with tumor purity greater than 70% were used for nucleic acid extraction.

All patients in the participating centers provided informed written consent for tumor collection and analysis. Clinical data were collected in each participating institution and correlated with pathological and molecular tumor features. The study was approved by the ethical committee of the Pitié-Salpêtrière Hospital (IDF-6, Ile de France). The collection and use of tissues followed procedures in accordance with the ethical standards formulated in the Declaration of Helsinki.

Immunohistochemistry procedures

All immunohistochemical analyses were performed on serial freshly cut 5 μm sections of the selected block. All IHC were performed with a Ventana Benchmarck automated method. Markers used for a diagnostic purpose were chromogranin A (clone DAK-A3, Dako), synatophysin (clone DAK-SYNAP, Dako), CD56 (NCAM1) (Clone 1B6, Novocastra), S100 protein (Z 0311, polyclonal, Dako) and pan-

cytokeratin (clone E1/AE3, Dako). The Ki67 proliferation index was studied using the clone MIB1 (Dako). To characterize TILs, whole sections were stained for CD4 (clone SP35 Ventana Medical System) and CD8 (clone 144B Dako). The expression of PD-L1 was studied using the clone E1L3N (Cell Signaling).

Quantification of IHC markers

All slides were digitalized using Aperio AT2 slide scanner (Leica Biosystems). Quantification of the Ki67 index was achieved by counting hotspots. On digitalized slides, hotspot areas were determined at low magnification and the area of interest was delimited at x20 magnification. A minimum of 1000 nuclei was necessary to determine the percentage of positive nuclei. Quantification was then achieved by counting the positive cells among all tumor cells present in the area of interest with Aperio Imagescope Software (Leica Biosystems) (counting tool). To consider Ki67 relevance, we compared it to a mitotic count for 10 high-power fields (= 2.5 mm^2).

For the quantification of TILs, hotspot areas were determined at low magnification. On digitalized slides, 5 fields of 0.2 μm^2 were delimited in the hotspots using Aperio Imagescope Software (Leica Biosystems). Lymphocytes were counted in the delimited areas using Aperio Imagescope counting tool. Lymphocytes were counted, blinded to the clinical data, within the stromal area of the tumor and within the tumor lobules themselves. CD4+ cells displaying macrophages features were excluded. PD-L1 expression was evaluated on digital slides, globally without distinction between tumor or immune cells. It was quantified by a pathologist (M.C.) in the following five categories: negative, positive in less than 1% of cells, positive in 1% to 5% of cells, positive in 5% to 50% of cells or positive in more than 50% of cells. The expression level of cytokeratin, synoptophysin and chromogranin proteins was analyzed in all ENB samples; the percentage of positive tumor cells was used for protein based clustering.

Nucleic acid extraction

DNA extraction was performed using the DNeasy Blood & Tissue Kit (QIAGEN) according to the manufacturer's instructions. RNA extraction was performed using the RNeasy Kit (QIAGEN) according to the manufacturer's instructions. Quality control of extracted nucleic acids was done using Agilent® bioanalyzer.

Whole-exome sequencing and somatic mutation detection

Exome capture was performed using Agilent SureSelect Human All Exon 50 Mb according to the manufacturer's instructions. Briefly, 3 μg of DNA from each sample were used to prepare the sequencing library through shearing of the DNA followed by ligation of sequencing adaptors. Whole-exome sequencing was performed, and paired-end sequencing (2 × 76 bp) was carried out using the Illumina HiSeq 2000; the resulting data were analyzed with the Illumina pipeline to generate raw fastq files. The coverage of our germline samples and tumor samples varied between 43x-80x and 79x-158x, respectively. The technical details and mutation detection were done according to the pipeline we previously reported (Malouf et al., 2014). When germinal DNA was not available, the comparison was done according to the reference genome. We filtered out all known single-nucleotide variants (SNVs)/indels in the UCSC dbSNP 135 and 1000 Human Genome Project SNP databases, and kept any mutations, which are in the Catalogue of Somatic Mutations in Cancer (COSMIC) database, curated by the Wellcome Trust Sanger Institute. Each somatic mutation or indel was annotated with its functional effect by SIFT to determine whether a mutation candidate was synonymous or nonsynonymous (benign or deleterious).

Validation of IDH2 mutations

IDH2 R172 mutations were confirmed by pyrosequencing, as previously described (Arita et al., 2015). In addition, immunohistochemistry for *IDH2* was performed on 5-μm–thick formalin-fixed, paraffin-embedded sections with an antibody specific to the mutant *IDH2* (R132/172) (clone MsMab-1, Millipore) according to the manufacturer's instructions. A labeled streptavidin biotin kit was used as a detection system (DAB Detection Kit - Ventana Medical Systems, Inc.). Combined cytoplasmic and nuclear staining was interpreted as immunopositive.

RNA sequencing and bioinformatics analysis

Total RNA for each sample was converted into a library of template molecules for sequencing on the Illumina HiSeq 2000 according to the NuGen Ovation RNA-Seq System V2 protocol. In brief, first, single-stranded cDNA was synthesized from 100 ng of DNase1-treated total RNA using a mix of DNA/RNA chimeric primers that hybridize to both the 5′ portion of the poly (A) sequence and randomly across the transcript. Second, strand synthesis produced double-stranded cDNA, which was amplified using single-primer isothermal strand-displacement amplification. The resultant cDNA was fragmented to 200 bp (mean fragment size) with the S220 Focused-ultrasonicator (Covaris) and used to make barcoded sequencing libraries on the SPRI-TE Nucleic Acid Extractor (Beckman-Coulter). Libraries were quantitated by qPCR (KAPA Systems), multiplexed and sequenced, 4 samples per lane, on the HiSeq2000 using 75 bp paired-end sequencing. The resulting data were analyzed with the current Illumina pipeline to generate raw fastq files. Overall, a median of 5.8 million reads was obtained per sample

Unsupervised classification of gene expression

The variance stabilizing transformations implemented in the DESeq package were performed on the count data to conduct sensible distance calculation (Anders and Huber, 2010). For each gene, the dispersion was calculated to measure its variance among samples, and thus the 2,000 genes with highest dispersions were selected for clustering analysis. Regarding the 29 distinct TCGA tumor samples used to characterize the lineage of ENB, 15 samples from each cancer dataset were randomly selected. To remove the systematic difference between ENB and TCGA samples, the median expression values of each batch per gene were scaled to the same level. Hierarchical clustering analysis was performed using the Pearson correlation coefficient as the distance metric and Ward's linkage rule. Principal component analysis was also applied to investigate the multivariate pattern. The consensus-clustering algorithm with the hierarchical clustering method was used to perform clustering analysis for ENBs, sinonasal carcinomas and normal sinonasal mucosa.

Analysis of differentially expressed genes

To perform normalization and test for differential expression with a negative binomial model between conditions, we chose to use the Bioconductor package DESeq (version 1.11.0) (Anders and Huber, 2010). Specifically, for each gene, a generalized linear model was fit to compare the expressions of basal and neural ENBs. The Benjamini-Hochberg method was used to control the false discovery rate (FDR). The Holm method was applied to calculate the adjusted P values of pairwise comparisons.

The Fantom human dataset of gene expression by cell type were analyzed using a previously utilized approach (Chen et al., 2018). Briefly, logged expression values (base 2) for each gene in the fantom dataset were centered on the median of sample profiles. For each fantom differential expression profile (genes centered within the fantom dataset), the inter-profile correlation (Pearson's, based on 15990 genes shared between the two datasets) was taken with that of each ENB differential expression profile (with genes centered on the median across samples). The same type of analysis was also carried out using the GSE95601 gene expression dataset.

Pathway enrichment analysis

Gene Ontology (GO) enrichment for biological processes and pathways (KEGG) were performed though the DAVID Bioinformatics Resources database (Huang et al., 2009), using the default setting. Gene categories associated with p value < 0.05 and FDR < 0.05 were considered statistically significant.

Estimation of immune infiltrate

To estimate tissue-infiltrating immune and other stromal populations' abundances using our mRNA dataset, we used the MCP counter pipeline using default setting (Becht et al., 2016). In total, the abundance of 10 cell populations (8 immune populations, endothelial cells and fibroblasts) is predicted by this method.

DNA methylation and bioinformatics analysis

Global DNA methylation was assessed using the Infinium HumanMethylation850® (HM850) BeadChip Array. Briefly, genomic DNA (500-1000 ng) was bisulfite converted using the Zymo EZ DNA methylation kit (Zymo Research, Irvine, CA) according to the manufacturer's recommendations. The amount of bisulfite-converted DNA and the completeness of bisuflite conversion for each sample were assessed using a panel of MethyLight-based real-time PCR quality control assays as described previously (Campan et al., 2009). Bisulfite-converted DNA was then used as a substrate for the Illumina EPIC BeadArrays, as recommended by the manufacturer and first described by Moran et al. (2016). Specifically, each sample was whole-genome amplified (WGA) and then enzymatically fragmented. Samples were then hybridized overnight to an 8-sample BeadArray, in which the WGA-DNA molecules annealed to locus-specific DNA oligomers linked to individual bead types. After the chemical processes, BeadArrays were scanned and the raw signal intensities were extracted from the *.IDAT files using the 'noob' function in the minfi R package. The 'noob' function corrects for background fluorescence intensities and red–green dye bias (Triche et al., 2013). The beta (β) value for each probe was calculated as (M/(M+U)), in which M and U respectively refer to the (pre-processed) mean methylated and unmethylated probe signal intensities. Probes with measurements in which the fluorescent intensity was not statistically significantly above the background signal (detection p value > 0.05) were removed from the dataset.

The average β value reports a methylation signal ranging from 0 to 1, respectively representing completely unmethylated to completely methylated values. In addition to β values, we used M-values in this report (M-value = $\log(\beta/1-\beta)$ because of stronger signals for quantifying methylation levels (Du et al., 2010). Hierarchical unsupervised clustering analysis was performed using the Pearson correlation coefficient as the distance metrics and Ward's linkage rule. First, we selected the top 1000 most variable probes located in CGIs to perform unsupervised clustering analysis. To avoid bias related to E-CIMP samples, those samples were excluded and unsupervised clustering analysis using the M values of the top 5000 most variable probes was performed.

Probes with adjusted p values < 0.05, and $\Delta\beta \geq 0.2$ or ≤ -0.2 were considered statistically significant and differentially methylated between the two epi-clusters identified. Probes with β-value ≤ 0.2 were considered unmethylated and those with β-value ≥ 0.3 were considered methylated. GO enrichment analysis was performed using differentially methylated probes from the DAVID Bioinformatics Resources database. Categories that displayed $\geq$ 2-fold enrichment, p value $< 10^{-3}$ and FDR < 0.05 were considered statistically significant.

QUANTIFICATIONS AND STATISTICAL ANALYSIS

Statistical analysis was performed using Fisher's exact test for categorical variables and the Mann Whitney test for continuous variables. Correlation between two continuous parameters was evaluated by Pearson's coefficient when the distribution was normal and by Spearman's coefficient otherwise. Cumulative survival time was calculated by the Kaplan-Meier method and analyzed by the log-rank test. All statistical tests were two-sided and conducted at the significance level of 0.05 using Prism v.7 (GraphPad Software, Inc., La Jolla, CA, USA).

DATA AND SOFTWARE AVAILABILITY

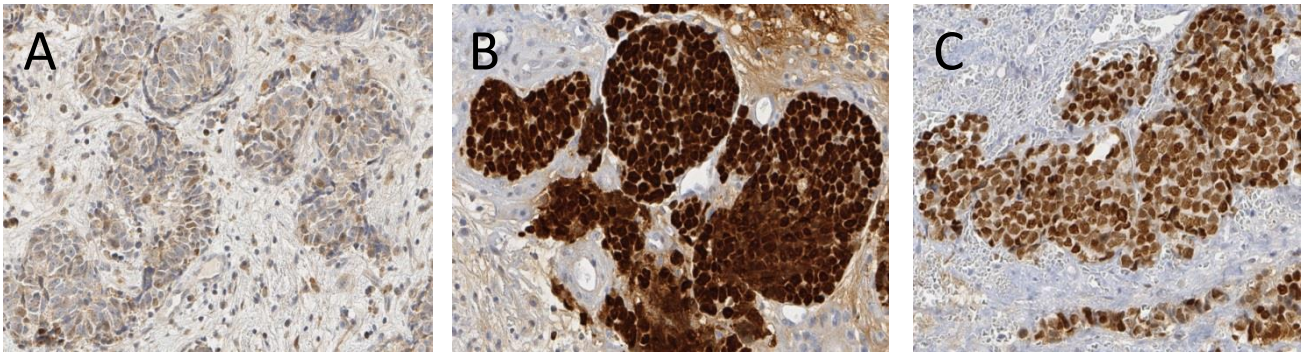
The accession number for the RNA-seq data reported in this study is GEO: GSE118995.

Cell Reports, Volume *25*

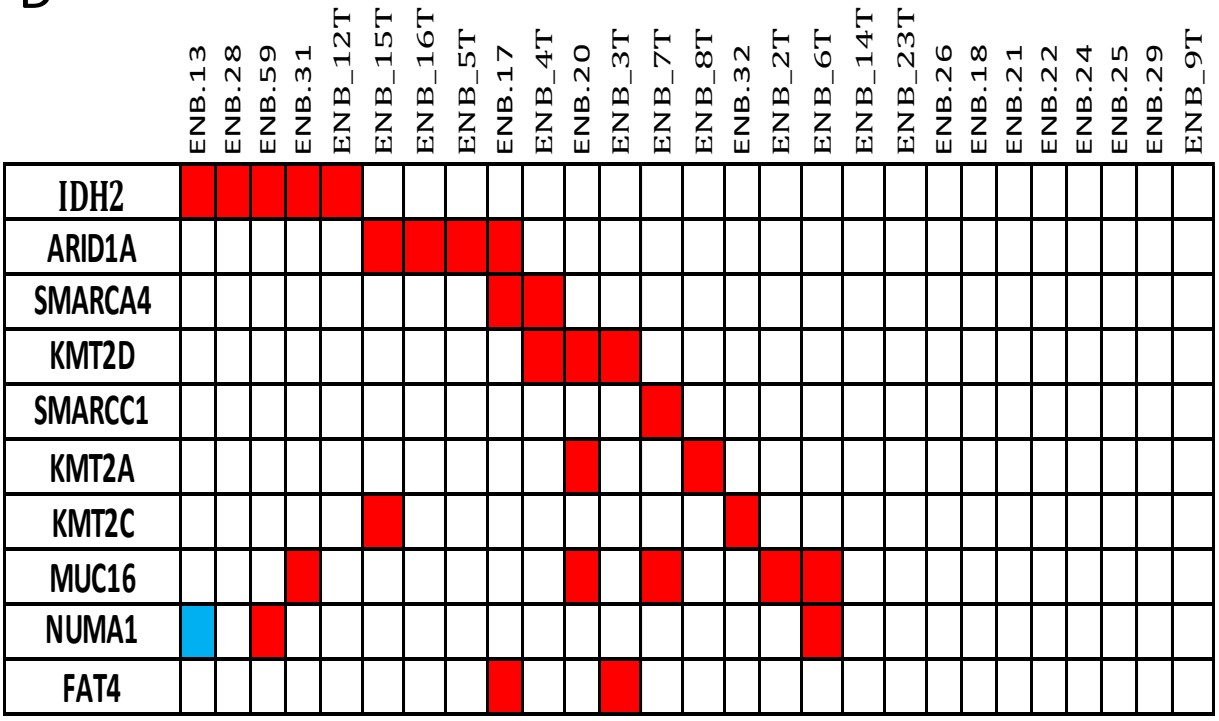
Supplemental Information

Integrated Multi-omic Analysis of Esthesioneuroblastomas Identifies Two Subgroups Linked to Cell Ontogeny

Marion Classe, Hui Yao, Roger Mouawad, Chad J. Creighton, Alice Burgess, Frederick Allanic, Michel Wassef, Xavier Leroy, Benjamin Verillaud, Geoffrey Mortuaire, Franck Bielle, Christophe Le Tourneau, Jean-Emmanuel Kurtz, David Khayat, Xiaoping Su, and Gabriel G. Malouf


A
B
C

D


ENB.13
ENB.28
ENB.59
ENB.31
ENB_12T
ENB_15T
ENB_16T
ENB_5T
ENB.17
ENB_4T
ENB.20
ENB_3T
ENB_7T
ENB_8T
ENB.32
ENB_2T
ENB_6T
ENB_14T
ENB_23T
ENB.26
ENB.18
ENB.21
ENB.22
ENB.24
ENB.25
ENB.29
ENB_9T
IDH2
ARID1A
SMARCA4
KMT2D
SMARCC1
KMT2A
KMT2C
MUC16
NUMA1
FAT4

E

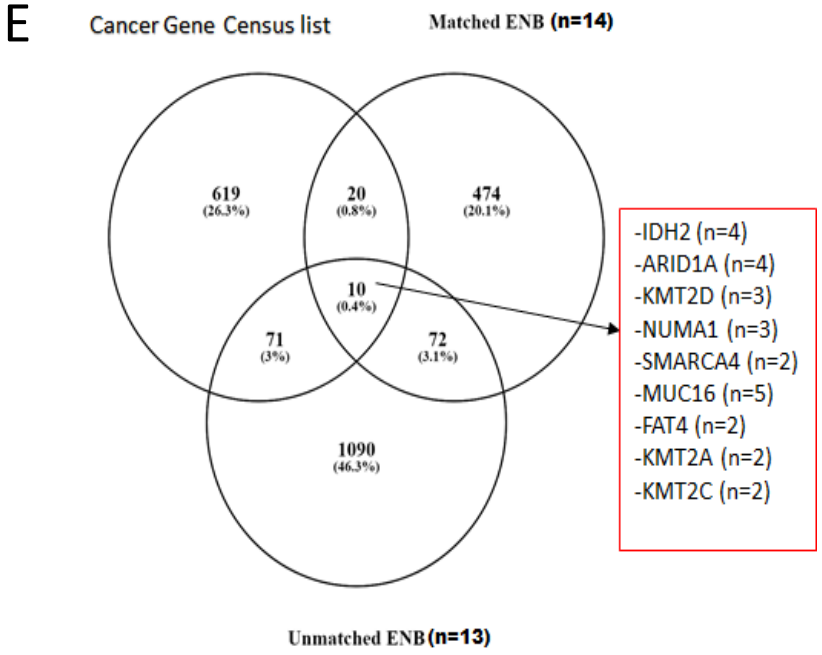

Cancer Gene Census list
Matched ENB (n=14)
619
(26.3%)
20
(0.8%)
474
(20.1%)
10
(0.4%)
71
(3%)
72
(3.1%)
1090
(46.3%)
-IDH2 (n=4)
-ARID1A (n=4)
-KMT2D (n=3)
-NUMA1 (n=3)
-SMARCA4 (n=2)
-MUC16 (n=5)
-FAT4 (n=2)
-KMT2A (n=2)
-KMT2C (n=2)
Unmatched ENB (n=13)

Figure S1, related to Figure 1: Immunohistochemical features of *TP53* mutated cases and landscape of somatic mutations in the validation set.

(A) *TP53* stopgain mutation in ENB-28T associated with loss of expression.

(B) *TP53* mutation R248W in ENB-13T associated with overexpression.

(C) *TP53* mutation C135F in ENB-59T associated with overexpression.

(D) Heatmap of somatic mutations in frequently mutated genes and family members in the whole esthesioneuroblastoma dataset assessed by whole-exome sequencing

(E) Venn diagram showing the intersection of genes belonging to Cancer Gene List (n=719) list, the matched (n=14) and the unmatched esthesioneuroblastomas samples (n+13)

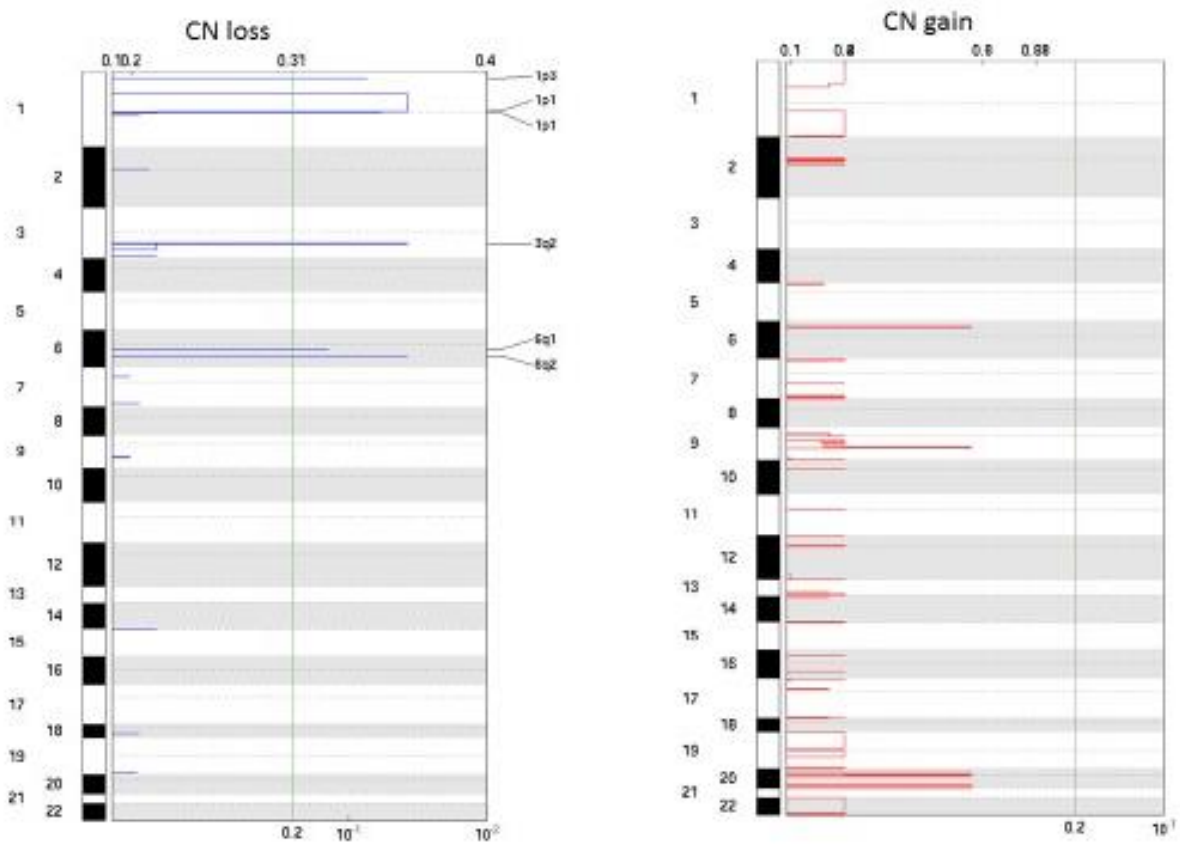


Figure S2, related to Figure 1: GISTIC analysis showing no recurrent focal copy-number changes in the 14 matched esthesioneuroblastomas

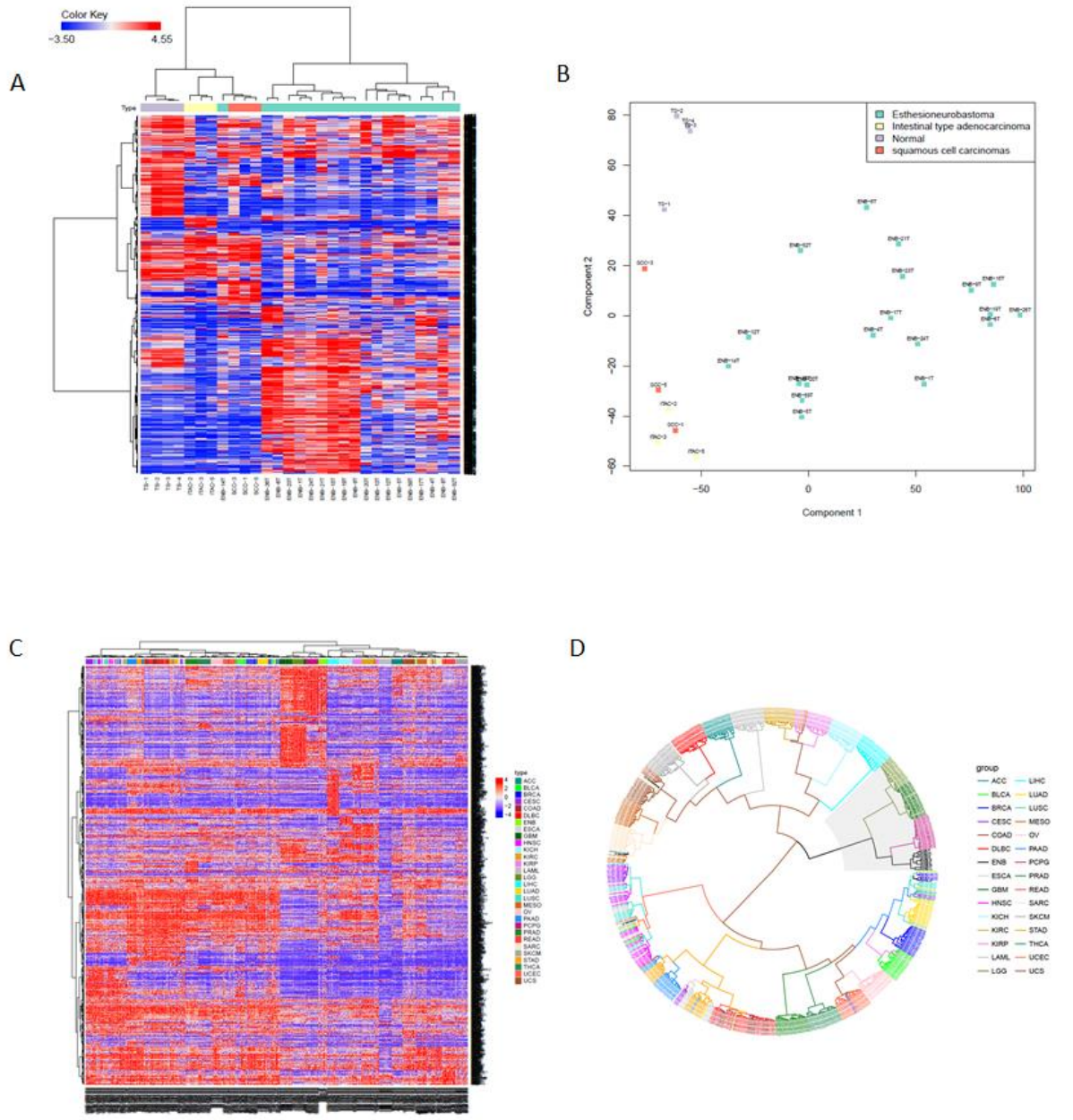


Figure S3, related to Figure 2: Unique profiles of esthesioneuroblastomas across cancer subtypes.

(A) Unsupervised hierarchical clustering using the 1500 most highly variable genes between esthesioneuroblastoma (ENB), squamous cell carcinoma (SCC), intestinal-type adenocarcinoma (ITAC) and healthy sinonasal mucosa.

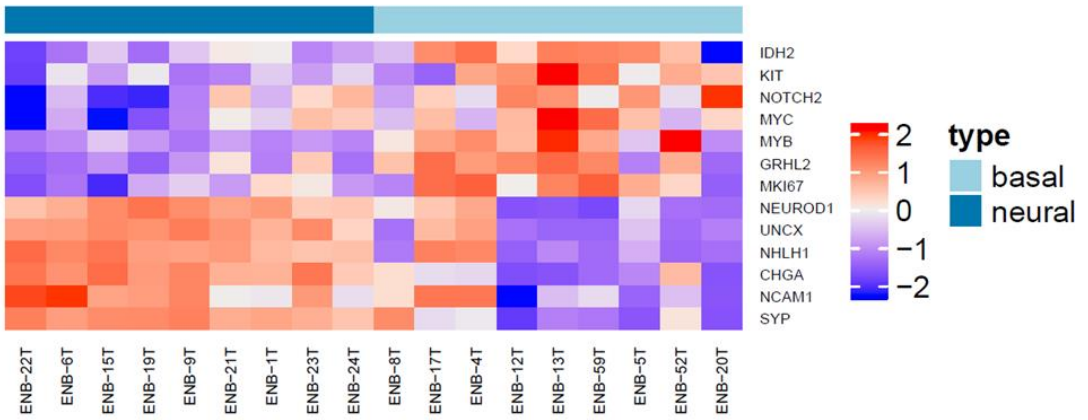
(B) Principal component analysis using the 1500 genes most highly variable between ENB, SCC, ITAC and healthy sinonasal mucosa.

(C) Unsupervised hierarchical clustering using ENB and 29 tumor subtypes extracted from TCGA database (adrenocortical carcinoma [ACC], bladder urothelial carcinoma [BLCA], breast invasive carcinoma [BRCA], cervical squamous cell carcinoma and endocervical adenocarcinoma [CESC], colon adenocarcinoma [COAD], lymphoid neoplasm diffuse large B-cell lymphoma [DLBC], esophageal carcinoma [ESCA], glioblastoma multiforme [GBM], head and neck squamous cell carcinoma [HNSC], kidney chromophobe [KICH], kidney renal clear cell carcinoma [KIRC], kidney renal papillary cell

carcinoma [KIRP], acute myeloid leukemia [LAML], brain lower grade glioma [LGG], liver hepatocellular carcinoma [LIHC], lung adenocarcinoma [LUAD], lung squamous cell carcinoma [LUSC], mesothelioma [MESO], ovarian serous cystadenocarcinoma [OV], pancreatic adenocarcinoma [PAAD], pheochromocytoma and paraganglioma [PCPG], prostate adenocarcinoma [PRAD], rectum adenocarcinoma [READ], sarcoma [SARC], skin cutaneous melanoma [SKCM], stomach adenocarcinoma [STAD], thyroid carcinoma [THCA], uterine corpus endometrial carcinoma [UCEC], uterine carcinosarcoma [UCS]). Note that ENB, glioblastoma and paraganglioma clustered together.

(D) Circular dendrogram representing ENB and 29 tumor subtypes extracted from TCGA showing that tumors derived from the neural crest and ENB cluster together.

A



B

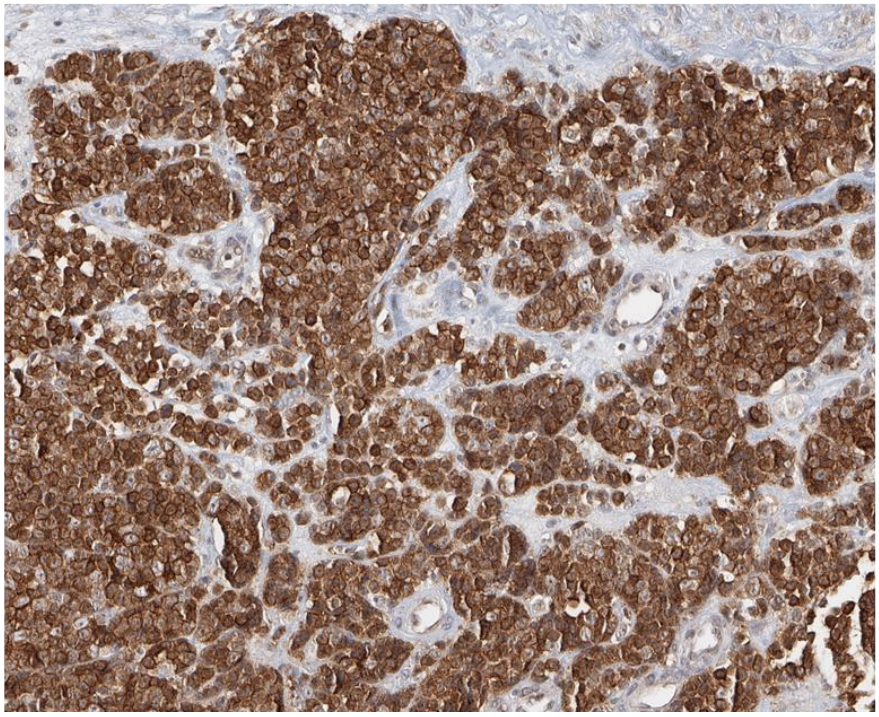


Figure S4, related to Figure 3: Focus on overexpression of key genes differentiating neural from basal ENBs.

(A): Heatmap showing the expression level (z-score) in each ENB sample (column) of genes (row) highly expressed in globose basal cells (*IDH2, KIT, NOTCH2, MYC, MYB, GRHL2* and *MKi67*) and immature neuron progenitors (INP2-3) (*NEUROD1, UNCX, NHLH1, CGHA, NCAM1*, and *SYP*).

(B) KIT overexpression shown by immunohistochemistry in a basal ENB (cytoplasmic stain with membrane strengthening).

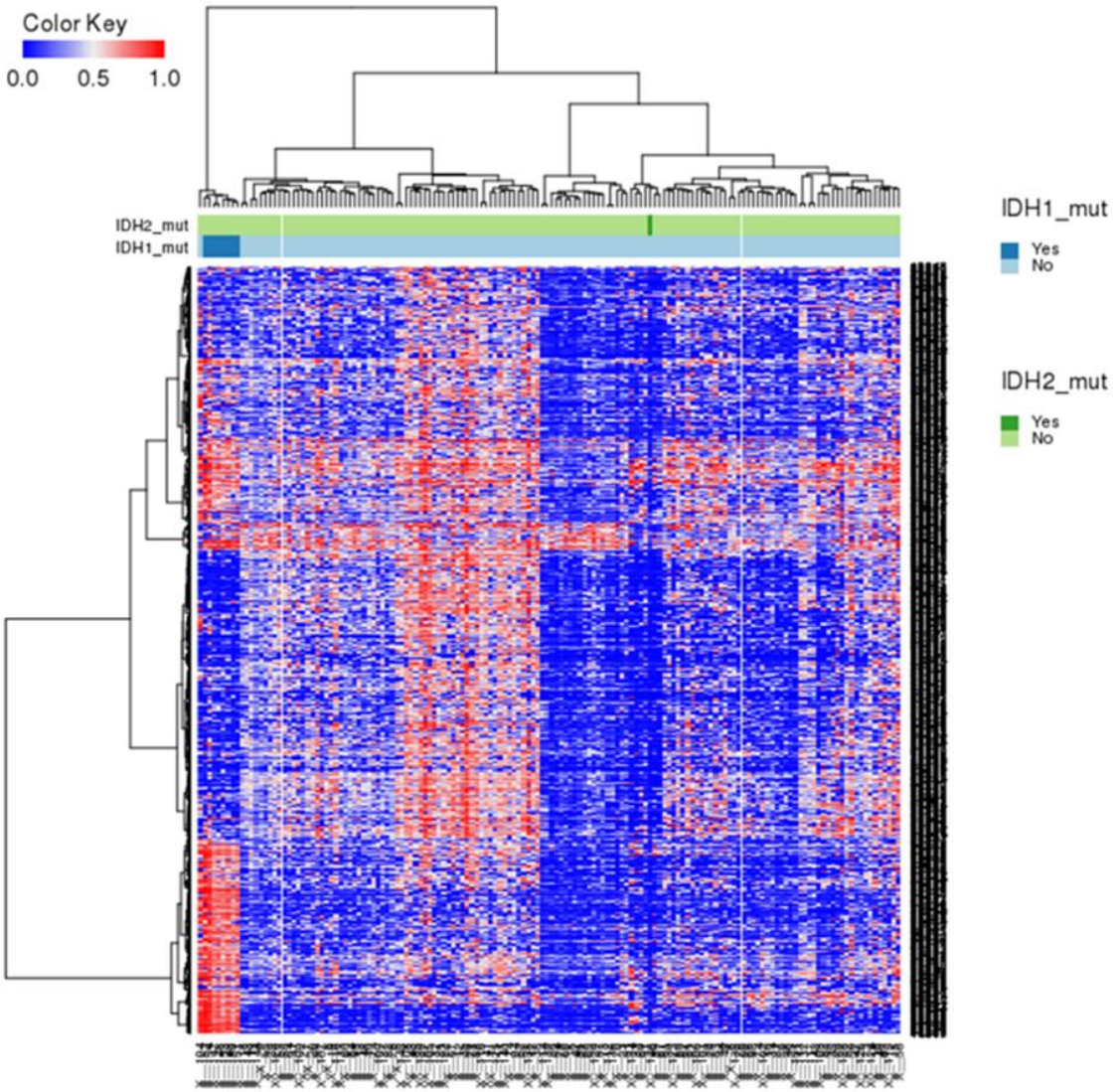
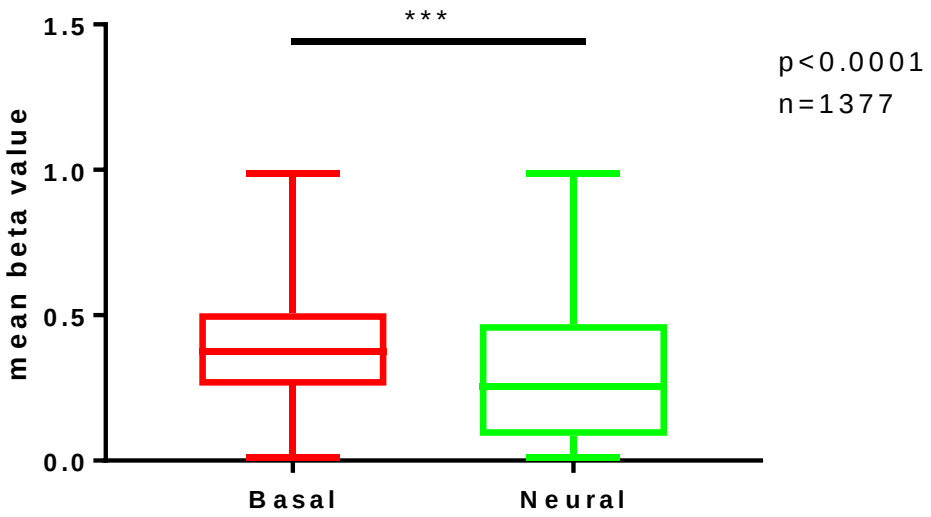
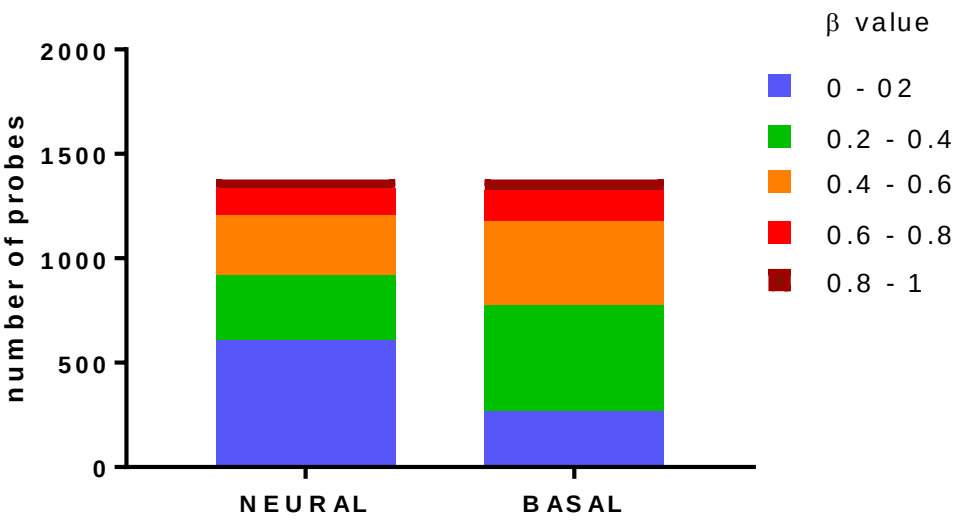


Figure S5, related to Figure 5: Supervised clustering of TCGA glioblastoma (GBM) samples using the 1500 most differentially methylated probes between *IDH2*-mutated and wild-type esthesioneuroblastomas. This signature was capable of distinguishing GBM with *IDH1-2* mutations from the remaining cases.

A


***
p<0.0001
n=1377
mean beta value
1.5
1.0
0.5
0.0
Basal
Neural

B


β value
0 - 02
0.2 - 0.4
0.4 - 0.6
0.6 - 0.8
0.8 - 1
number of probes
2000
1500
1000
500
0
NEURAL
BASAL

C

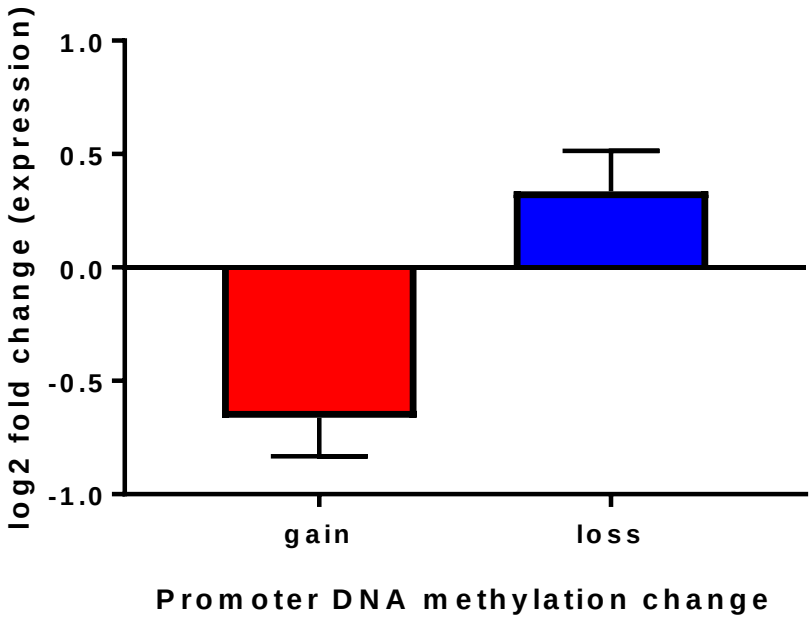

log2 fold change (expression)
1.0
0.5
0.0
-0.5
-1.0
gain
loss
Promoter DNA methylation change

Figure S6, related to Figure 5: Methylation divergence between neural and basal ENB subtypes.

(A) Box-plot of mean methylation level of probes located in CGI in both basal and neural ENB subtypes.

(B) Distribution of probes located in CGI in both neural and basal ENB subtypes.

(C) Gene expression changes associated with gain or loss of DNA methylation in neural as compared to basal ENB subtype.