

ARTICLE

Open Access

# Regulation of sarcomagenesis by the empty spiracles homeobox genes *EMX1* and *EMX2*

Manuel Pedro Jimenez-García<sup>1,2</sup>, Antonio Lucena-Cacace<sup>3</sup>, Daniel Otero-Albiol<sup>1,2</sup> and Amancio Carnero<sup>1,2</sup> 

## Abstract

The EMX (Empty Spiracles Homeobox) genes *EMX1* and *EMX2* are two homeodomain gene members of the EMX family of transcription factors involved in the regulation of various biological processes, such as cell proliferation, migration, and differentiation, during brain development and neural crest migration. They play a role in the specification of positional identity, the proliferation of neural stem cells, and the differentiation of certain neuronal cell phenotypes. In general, they act as transcription factors in early embryogenesis and neuroembryogenesis from metazoans to higher vertebrates. The *EMX1* and *EMX2*'s potential as tumor suppressor genes has been suggested in some cancers. Our work showed that *EMX1/EMX2* act as tumor suppressors in sarcomas by repressing the activity of stem cell regulatory genes (*OCT4*, *SOX2*, *KLF4*, *MYC*, *NANOG*, *NES*, and *PROM1*). EMX protein downregulation, therefore, induced the malignance and stemness of cells both in vitro and in vivo. In murine knockout (KO) models lacking *Emx* genes, 3MC-induced sarcomas were more aggressive and infiltrative, had a greater capacity for tumor self-renewal, and had higher stem cell gene expression and *nestin* expression than those in wild-type models. These results showing that EMX genes acted as stemness regulators were reproduced in different subtypes of sarcoma. Therefore, it is possible that the EMX genes could have a generalized behavior regulating proliferation of neural crest-derived progenitors. Together, these results indicate that the *EMX1* and *EMX2* genes negatively regulate these tumor-altering populations or cancer stem cells, acting as tumor suppressors in sarcoma.

## Introduction

Sarcomas are a type of tumor affecting a low percentage of the population and are malign and commonly highly aggressive. They are a heterogeneous group of tumors with more than 50 different histopathological subtypes according to WHO Classification of Tumours Editorial Board, 2020, giving complexity to their study<sup>1,2</sup>. As many sarcomas have their origin in the differentiation of neural crest-derived multipotent cells<sup>3–6</sup>, certain genes involved in the differentiation of this main embryological feature, such as *EMX1/2* genes, may be deregulated in sarcomas.

The neural crest is a transient embryonic structure of multipotent cells that contributes to the formation of

multiple tissues<sup>7</sup>. Cells derived from the neural crest have developed specialized mechanisms to promote their transition from an epithelial to a mesenchymal phenotype, a process called mesenchymal–epithelial transition<sup>8,9</sup>. Deregulations at the level of these characteristics can become a risk by contributing to the processes of tumorigenesis and metastasis<sup>3,4,8–11</sup>. Cells derived from the neural crest contribute to the development of multiple cell and tissue types, such as melanocytes, Schwann cells, cells of the nervous system, adrenal medulla, and derivatives of cartilage and bones of the cephalic region, which require a highly orchestrated transcriptomic gene program. This program depends on cellular and extracellular signals that regulate migration, proliferation, differentiation, and long-distance survival to give rise to derived histotypes<sup>8,9,12</sup>. In general terms, sarcomas that derive from the neural crest have been proposed to be schwannoma sarcomas, malignant peripheral nerve sheath tumors, neurothekeomas, rhabdoid tumors, and synovial sarcomas<sup>2,13–16</sup>.

Correspondence: Amancio Carnero (acarnero-ibis@us.es)

<sup>1</sup>Instituto de Biomedicina de Sevilla (IBIS), Hospital Universitario Virgen del Rocío, Universidad de Sevilla, Consejo Superior de Investigaciones Científicas, Sevilla, Spain

<sup>2</sup>CIBER de Cancer, IS Carlos III, Madrid, Spain

Full list of author information is available at the end of the article

Edited by A. Stephanou

© The Author(s) 2021



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

The EMX (Empty Spiracles Homeobox) genes *EMX1* and *EMX2* are two homeodomain genes homologous to the *Drosophila melanogaster* *ems* (empty spiracles) gene<sup>17</sup>. The EMX1/EMX2 proteins are members of the EMX family of transcription factors, primarily confined to the developing brain, where they play a role in the specification of positional identity, the proliferation of neural stem cells, and the differentiation of certain neuronal cell phenotypes<sup>17</sup>. Specifically, these genes are involved in establishing the regional pattern of the anterior brain, directly intervening in neural precursors. EMX1/EMX2 could have a role as inhibitors of cell proliferation, of the migration of cortical neuroblasts, and of the regulation of their differentiation<sup>18–21</sup>. In addition, *Emx2* is expressed at high levels in adult neural stem cells (ANSCs) in vitro and is downregulated after differentiation. The overexpression of the *Emx2* gene in ANSCs has an anti-proliferative effect but does not influence a particular differentiation pathway.

The *EMX1/EMX2* genes have been associated with cancer in a few solid tumors of epithelial origin, such as lung cancer<sup>22,23</sup>, malignant pleural mesothelioma<sup>24</sup>, gastric cancer<sup>25</sup>, endometrial cancer<sup>26,27</sup>, and liver metastases from colorectal adenocarcinoma<sup>28</sup>. In addition, these genes were associated with glioblastoma<sup>29,30</sup> and melanoma<sup>29,31</sup>, cancer whose origin comes from melanocytes, a derivative of the neural crest.

In the tumors tested, restoration of EMX2 expression levels suppressed cell proliferation and invasive phenotypes, sensitized lung cancer cells for treatment, and inhibited the canonical Wnt pathway. In vitro EMX2 silencing promoted cell proliferation, invasive phenotypes, and activation of the canonical Wnt pathway, suggesting that EMX2 may be a tumor suppressor gene<sup>25–30</sup>. Overexpression of *Emx2* under the Nes promoter (Nestin) suppressed cell proliferation in glioblastoma cell lines. The promoter was activated in neural stem cells, indicative of stem cell antagonism, and *Emx2* expression<sup>29,30,32–34</sup>. However, the role of EMX1/EMX2 in the initiation and progression of human sarcoma, as a tumor model derived from the mesoderm and neural crest, has not yet been elucidated. Nor have the molecular pathways altered and dependent on EMX1/EMX2 levels been studied in depth. In the present work, we linked the formation of sarcomas in general with a possible common origin in the neural crest where EMX would maintain the migrating progenitors derived from this tissue with proliferative arrest.

## Results

### EMX represses tumorigenic properties in sarcoma

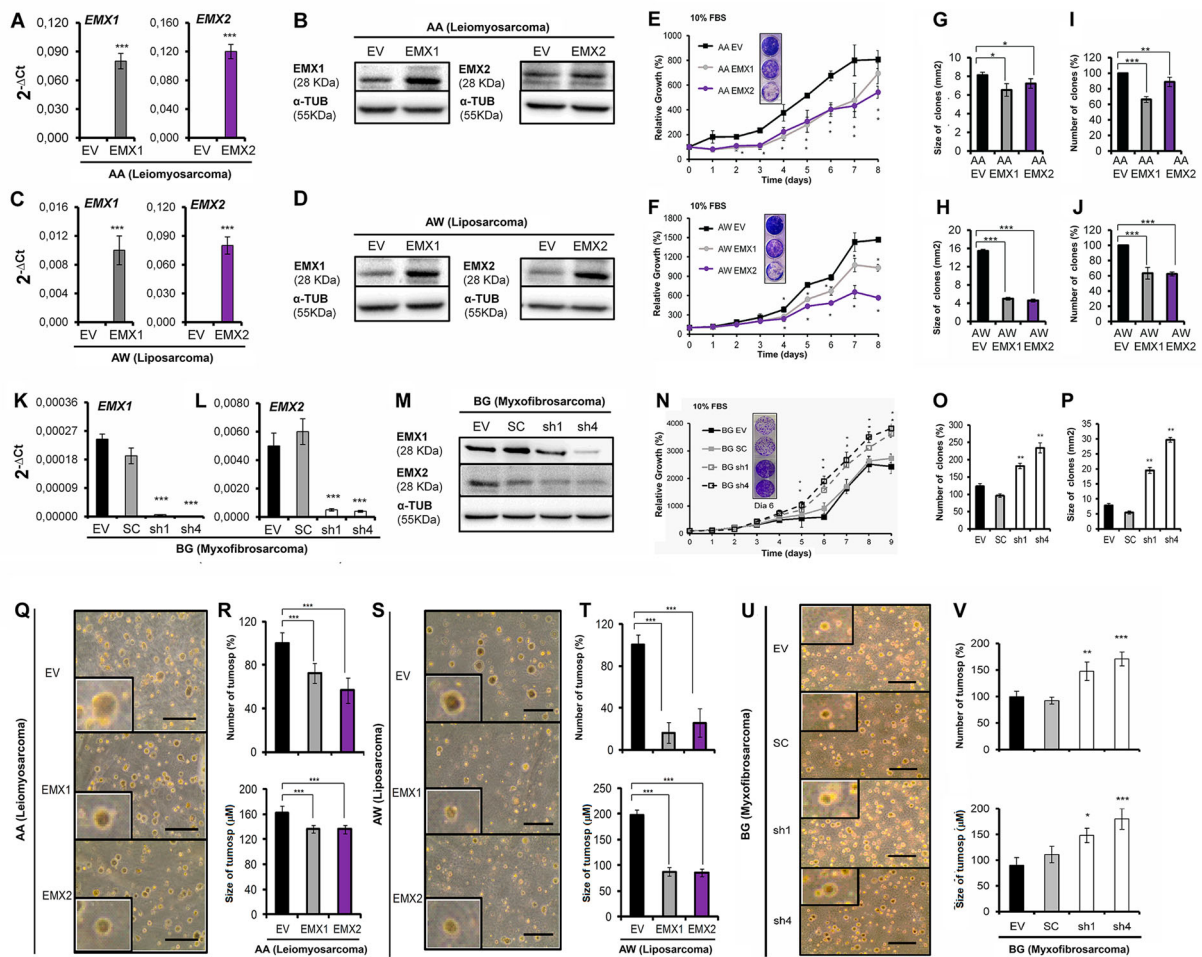
To explore the role of EMX, we first selected three different low-passage sarcoma cell lines from different tissues of origin; two sarcoma lines expressed low levels of EMX1 and EMX2 (Supplementary Fig. S1), as models to study the

effects of EMX overexpression: AA (leiomyosarcoma) and AW (liposarcoma). These cell lines were subject to ectopic expression of EMX1 or EMX2 cDNAs, whereas the parental control was transfected with an empty vector (pCMV6-EV). All models were validated both at the mRNA level using quantitative reverse-transcriptase PCR (qRT-PCR) (Fig. 1A, C) and at the protein level (Fig. 1B, D). Very low levels of EMX mRNA detected compared to detectable proteins suggests more stable protein and rapid mRNA processing. Overexpression of the EMX genes reduced proliferation (Fig. 1E, F) and the number and size of the colonies formed (Fig. 1G–J). We also selected a third line, with high expression of EMX1 and EMX2 (BG, mixoid fibrosarcoma, Supplementary Fig. S1) as the model to reduce the levels of both genes. The reduction of the levels of EMX1/EMX2 in BG was carried out with stable expression of two different short hairpin RNAs (shRNAs) against a common sequence of EMX1 and EMX2 (pRS-sh1 and pRS-sh4). Empty vectors (pRS-EV) and vectors with random shRNA sequence (pRS-SC) were used as controls. All models were validated both at the mRNA level (Fig. 1K, L) and at the protein level (Fig. 1M). For both shRNAs, we confirmed that both EMX genes were targeted, ensuring proper downregulation of EMX signaling. In this cell line, eliminating the EMX genes increased the proliferative capacity (Fig. 1N) and the capability to form colonies (Fig. 1O, P).

In all cases, we observed that high levels of EMX reduced the number and size of the colonies formed when grown in soft agar (Fig. 1Q–V). We also observed that in vivo, the xenograft tumors overexpressing EMX1 or EMX2 grew slower and were significantly smaller than the tumors formed by each of the parental lines of AA and AW (Fig. 2A, B, D, E). Furthermore, mice with xenografts from the EMX-overexpressing tumor cells survived longer (Fig. 2C, F). Similarly, with the BG line, tumors with reduced levels of EMX grew faster (Fig. 2G, H) and mice showed a marked reduction in survival compared to those with the control lines (Fig. 2I).

### The effect of the EMX1/EMX2 level on the stem cell phenotype

Next, we performed functional tests aiming to measure the stemness properties in cells with altered EMX levels. Initially, we measured tumorspheres described as enriched in multipotent progenitors<sup>35,36</sup> and have high levels of expression of tumor stem cell markers<sup>37,38</sup>. We measured stemness in two different ways: (1) planting 5000 cells and measuring primary and secondary tumorspheres; and (2) planting 1 single cell per well and leaving it to grow to reconstitute a cell population. Under EMX1 or EMX2 overexpression, a reduction in the number and size of primary and secondary tumorspheres was observed (Fig. 3A–F). Coincidentally, the downregulation of EMX genes promoted tumorsphere formation either in number or size (Fig. 3G–I). When the

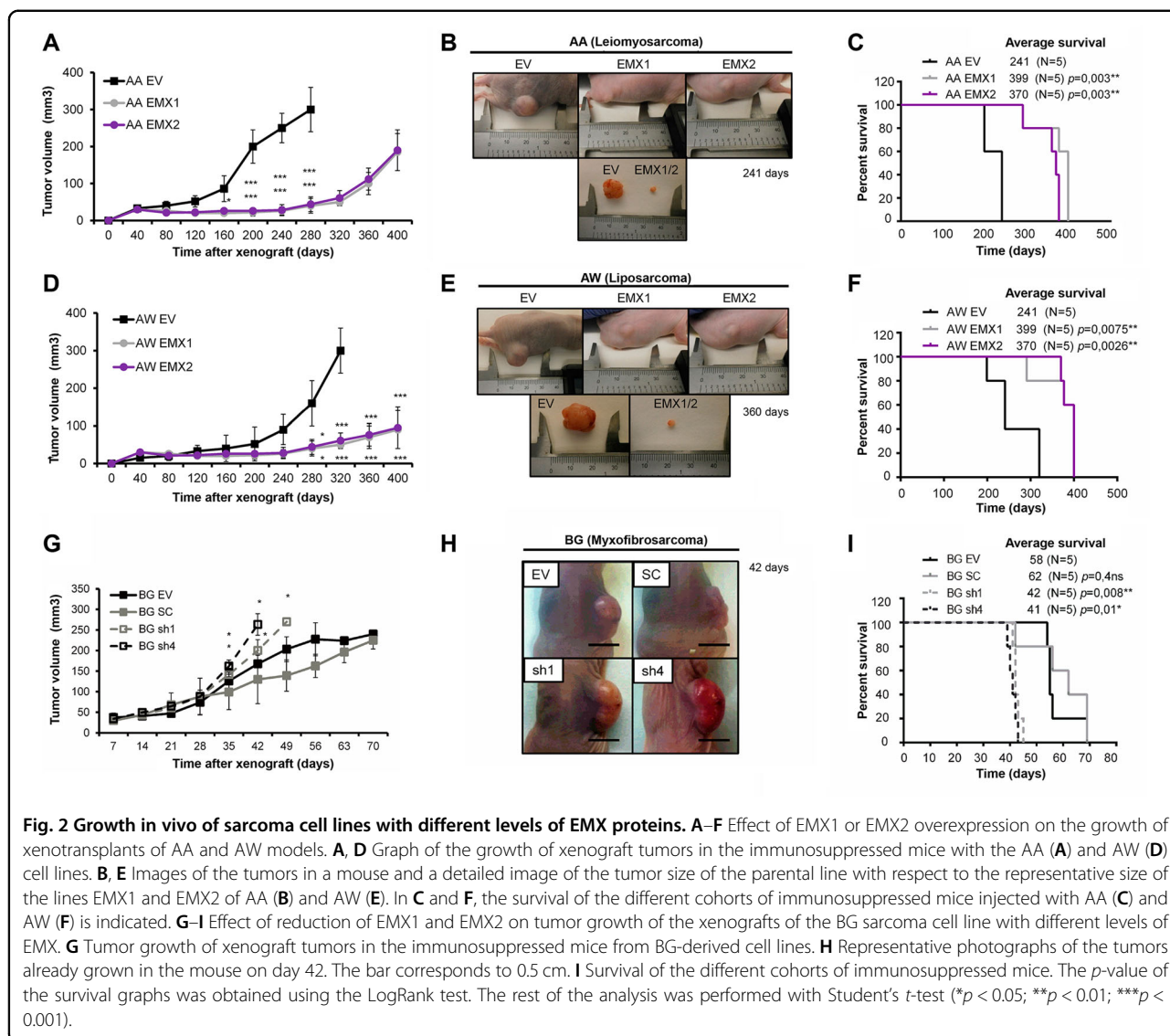


**Fig. 1** Effect of EMX1 and EMX2 expression levels in cell models expressing different levels of EMX proteins. Ectopically transfected empty vector (EV) or EMX1 or EMX2 cDNAs. In **A** and **C**, the expression levels were measured by qRT-PCR and in **B** and **D** by western blotting in the primary lines of AA sarcoma (**A, B**) and AW (**C, D**). **E–J** Effect of EMX1 and EMX2 overexpression on the proliferative and colony-forming capacities of sarcoma cell lines. **E, F** Growth curves of the overexpression of EMX1 in the AA (**E**) and AW (**F**) sarcoma lines. Inside, images of crystal violet-stained growth plate wells on the last day of the experiment. **G–J** The result of the clonability test assessing the effect of overexpression of EMX on the size and number of colonies in the AA (**G, I**) or AW (**H, J**) cell lines. **G, H** Graph of the size of the colonies. **I, J** Graph of the number of colonies. **K, L** Validation of cellular models of silencing of EMX1 and EMX2 expression levels by two independent shRNAs, sh1 and sh4, in the sarcoma line BG. **M** Protein levels of EMX1, EMX2, and the endogenous control  $\alpha$ -tubulin ( $\alpha$ -TUB) in the BG cell line expressing the shRNAs against EMX. **N–P** Effect of the reduction of EMX1 and EMX2 on the proliferative and colony-forming capacities in the sarcoma BG cell line. **N** Growth curves of the silencing models of EMX1 and EMX2 of the BG sarcoma line under the conditions of 10% FBS. Inside, images of crystal violet-stained growth plate wells on a given day of the experiment. **O, P** Clonability assay of the silencing models of the BG sarcoma line. **O** Graph of the number of colonies. **P** Size of the colonies. **Q–V** Effect of the EMX1 and EMX2 levels on anchor-independent growth. **Q, S, U** Representative images of the colonies embedded in the agar in the AA (**Q**), AW (**S**), or BG (**U**) lines with different levels of EMX protein. **R, T, and V** Percentage of the number of colonies and the mean size of the colonies for the AA (**R**), AW (**T**), or BG (**V**) cell lines with different levels of EMX protein. The horizontal bar corresponds to 500  $\mu$ m. In all cases, the mean of three independent experiments is represented in triplicate  $\pm$  SD. Statistical analysis was performed with Student's *t*-test (\**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001). In western blotting, a representative experiment of at least three independently performed experiments is shown. The endogenous control was  $\alpha$ -TUB ( $\alpha$ -tubulin).

tumorsphere formation test was performed with an isolated cell by flow cytometry, the results were similar; a reduction in size and tumor formation efficiency was observed with EMX1/EMX2 overexpression (Fig. 3J–O), while the

downregulation of EMX increased the efficiency and the size of the tumorspheres (Fig. 3P–R).

We analyzed the different clonal phenotypes, as depending on the phenotype the clones were classified



into (i) holoclones, enriched in cells with the capacity to regenerate the culture; (ii) paraclones, formed by differentiated cells that are not capable of reconstituting a culture; and (iii) meroclones, with intermediate characteristics<sup>39,40</sup>. We observed a reduction of the holoclones compared to the clones with the highest degree of differentiation (meroclones and paraclones) in cells ectopically expressing EMX1/EMX2 (Supplementary Figs. FS2 and FS3).

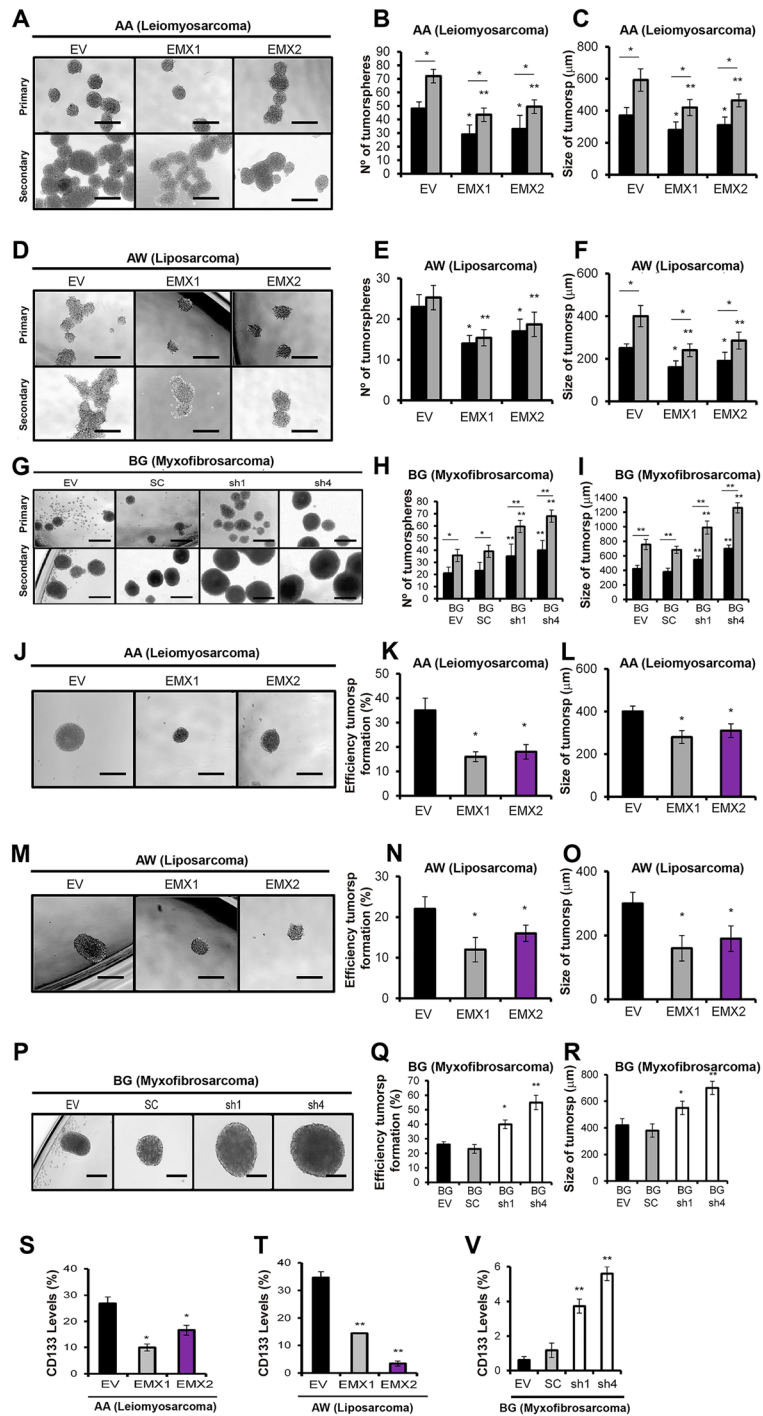
Finally, we determined the effect on the tumor cell pool by measuring a surface marker to identify cancer stem cells (CSCs) with CD133 for sarcoma (Fig. 3S–V). In exponentially growing cells, controls with vector only or cells ectopically expressing EMX1 or EMX2 were collected, and the content in the cells positive for CD133 was measured. We observed a marked decrease

in the CD133+ ratio, in the lines overexpressing the EMX genes (Fig. 3S, T). Cells with downregulated EMX genes increased the pool of CD133+ cells (Fig. 3V).

**EMX1/EMX2 repress genes related to stemness**

Next, we measured the expression of some transcription factors involved in signaling pathways that regulate the core of pluripotency of stem cell factors (OCT4, SOX2, KLF4, MYC, and NANOG), as well as the expression of BMI1, NES, and PROM1, which is responsible for the CD133 antigen<sup>41–45</sup>.

We measured the response of these transcripts in three different settings: whole in vitro culture (ET), cells growing in tumorspheres only (TO), or cells from the tumor growing in vivo in xenograft (XEN). We observed a significant and



**Fig. 3** Effect of different levels of EMX1 and EMX2 on tumorigenesis in sarcoma cell lines. **A, D** Micrographs of the results of the primary and secondary tumor formation assays in the AA (**A**) and AW (**D**) sarcoma lines with EMX1 or EMX2 overexpression or in the BG cell line with EMXs downregulated by shRNAs (**G**). **B, E, H** Average number of primary (black bars) and secondary (gray bars) tumorspheres in AA (**B**), AW (**E**), and BG (**H**). In **C, F, I**, the mean sizes of the primary (bars in black) and secondary (bars in gray) tumors of AA (**C**), AW (**F**), and BG (**I**) are shown. **J, M, P** Representative images of tumor formation from isolated cells in the AA (**J**), AW (**M**), and BG (**P**) cell lines with different levels of EMX proteins. Sphere-formation efficiency of single cells for AA (**K**), AW (**N**), or BG (**Q**). Size of tumorspheres formed from an isolated cell in AA (**L**), AW (**O**), or BG with different levels of EMX proteins. All bars correspond to 500 µm. The mean of three independent experiments is represented in triplicate ± SD. Statistical analysis was performed with Student's *t*-test (\**p* < 0.05; \*\**p* < 0.01; \*\*\**p* < 0.001). **S, T, V** Effect of EMX1 and EMX2 levels on the percentage of CD133+ cells in sarcoma cell lines with different levels of EMX proteins. Graphs of the percentage of expression of the CD133 marker (conjugated with PE: CD133-PE), measured by FACS in the different models of lines AA (**S**) and AW (**T**) or BG (**V**) with different levels of EMX proteins.

generalized reduction in the expression of stem cell genes under EMX1/EMX2 overexpression (Fig. 4A, B). The result presented greater differences in the expression levels in the tumorspheres (TO) with respect to the total extract of the cell line (ET) and the extract of the tumor generated in xenotransplants (XEN). In total, coincident results were obtained in the BG cell line, where reduction of EMX levels increased stemness markers (Fig. 4C).

In addition, we verified that there is a reduction in Prom1 transcription under high levels of EMX, confirming the reduction in CD133+ cells observed previously (Fig. 4A–C). A reduction in the expression of NESTIN (NES) was also observed in the EMX-positive cell population (Fig. 4A–C), reaffirming the markers NES and CD133 as markers enriched in stem cells in sarcoma and regulated by the expression of EMX. These results confirmed the negative relationship between tumor stem cell physiology and the presence of EMX1/EMX2. We observed similar correlation in bioinformatic analysis of EMX1/EMX2 in stem cell biology (Fig. 4D).

#### The absence of Emx1 or Emx2 increases the tumorigenic capacity of 3MC-induced sarcomas in murine KO models

To study the role of EMX genes in sarcomagenesis in vivo, we used a model for the specific tumor induction of sarcoma mediated by the carcinogen 3-methylcholanthrene (3MC). Thus, we established the experimental cohorts of 15 mice in both the *Emx1* and *Emx2* knockout (KO) mouse lines<sup>46,47</sup> and performed the 3MC carcinogen-mediated tumorigenic induction. We used two doses of intramuscular 3MC at 1 mg in the abductor muscle of the left paw of each mouse (Fig. 5A). We followed each animal for the first tumor signs and they were killed when the tumors reached a size of 1.7 cm<sup>3</sup>. Then, a complete autopsy was performed, observing the state of each organ and possible metastases. Tumors were observed to appear earlier with a statistically significant difference in the *Emx1* and *Emx2* KO (–/–) line in a gene dose-dependent manner (Fig. 5B, C). Mice null for *Emx* developed sarcoma earlier than heterozygous mice. In addition, the data on the appearance of the tumors correlated with the significant increase in tumor size (Fig. 5D, E) and a reduction in the survival of the *Emx*-null mice (Fig. 5F, G). To verify the physiological cause of reduced survival and increased tumor malignancy, we performed a histopathological evaluation of bone infiltration of the obtained tumor sections, comparing them with control muscle tissues. After analyzing the sections of all the cohorts of both murine lines, we classified the different degrees of infiltration of the femoral bone marrow into three groups as follows: (1) incipient bone infiltration, (2) patent bone infiltration, and (3) aggressive bone infiltration (Fig. 5H). We observed a

statistically significant increase in grade 3 aggressive bone infiltration in the mice with *Emx* downregulation (Fig. 5I, J), indicating greater malignancy and invasiveness of the tumor in the absence of the *Emx* genes.

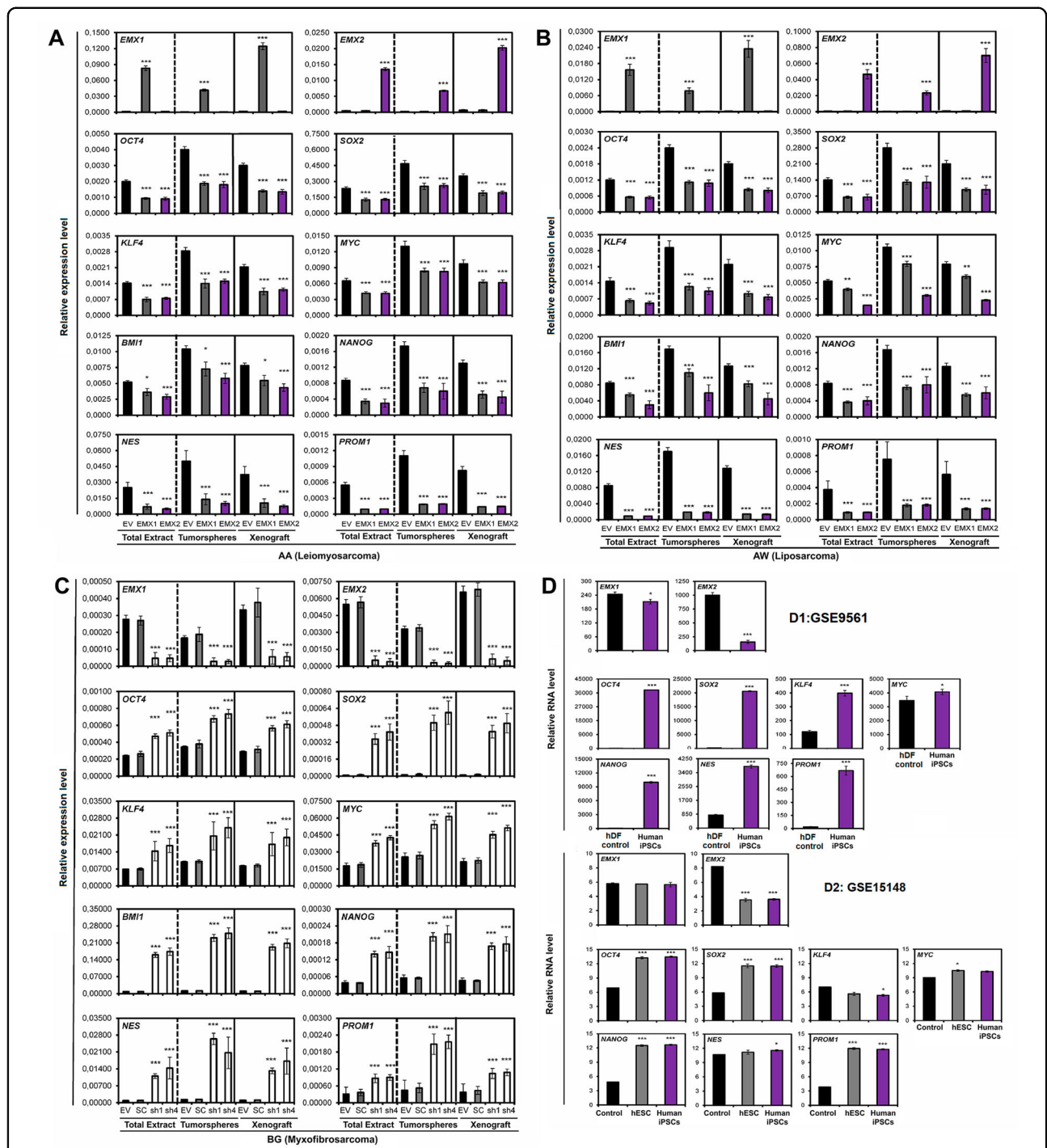
The absence of *Emx1* or *Emx2* increased inflammation in the 3MC-induced sarcomas (Supplementary Fig. FS4); we observed that all induced tumors in all genotypes showed similar KI67 immunostaining levels (60–65 ± 12%) compared to 8–9 ± 2% in non-tumoral tissue sections. This finding indicates the homogeneity of tumor induction in the model mediated by the carcinogen 3MC. We also observed a greater presence of both leukocytes measured by CD45 (Supplementary Fig. FS5) and macrophages measured by F4/80 (Supplementary Fig. FS5) in the hemizygous genotypes (+/–) and homozygous null genotypes (–/–) of both murine models.

#### The absence of Emx1 or Emx2 increases the phenotype of stem cells in the 3MC-induced sarcomas

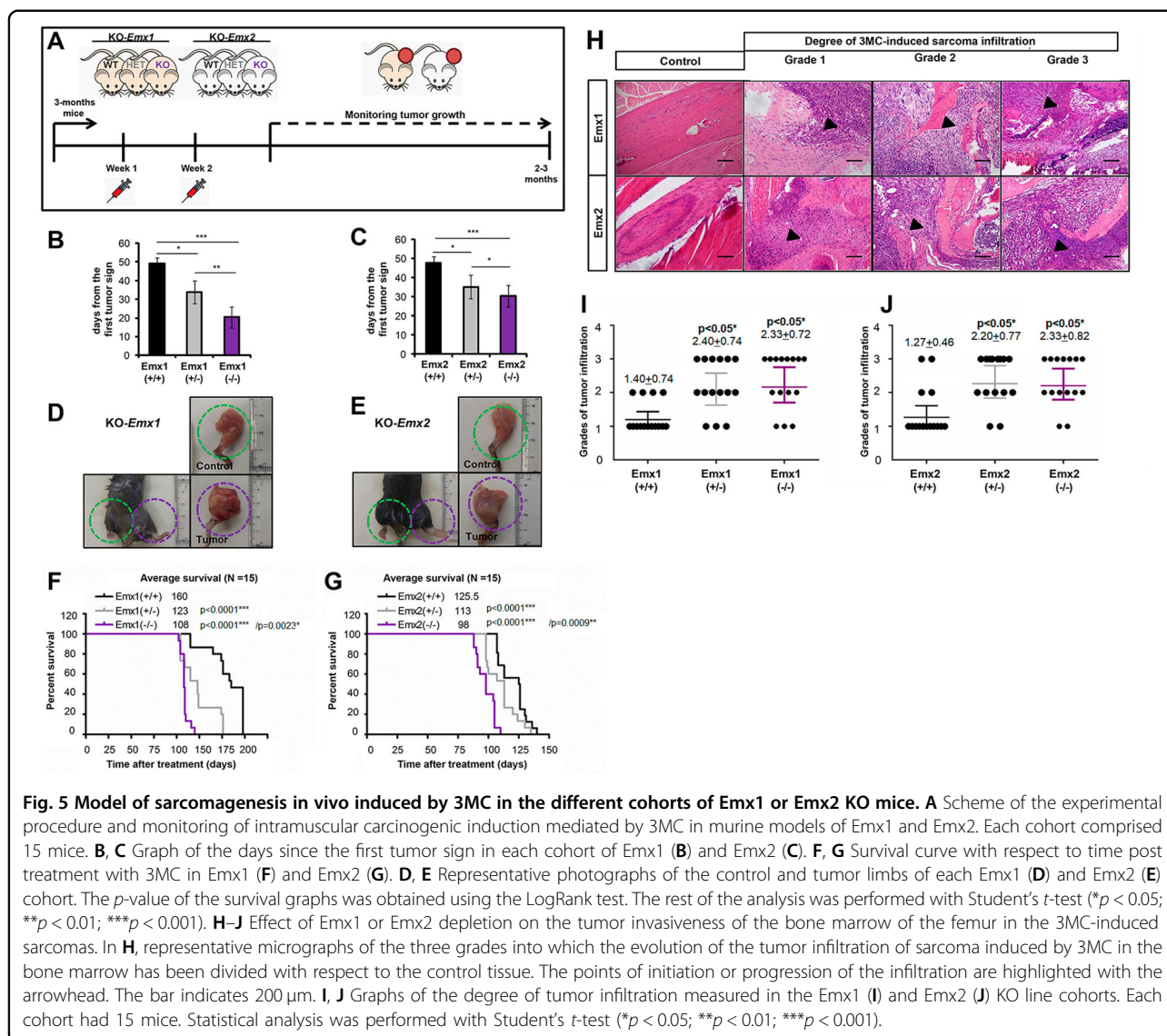
To correlate the results obtained in vitro and in vivo with respect to the relationship of the properties of stem cells with the EMX genes in the primary sarcoma lines, we studied the genes regulating the properties of stem cells in the two models. For that, we (1) measured protein levels by immunohistochemistry and (2) measured mRNA levels using qRT-PCR of *Nanog*, *Nes*, *Oct4*, *Sox2*, *Klf4*, and *Myc*, both in the induced sarcoma with respect to the control muscle tissues.

We observed 5–6 ± 1% *Emx*-positive nuclei in the control tissue and 21–39% ± 5% in the induced sarcoma tissue (Fig. 6A, B), indicating an increase in cells with expression of EMX in sarcoma, with no expression observed in the *Emx*-null compared to the wild genotype (Fig. 6A, B). This result negatively correlates with the protein expression of the neural marker NES, which has its highest level of expression in sarcoma tissues of the *Emx*-null mice. In contrast, NES presented very low levels in the induced sarcoma tissues of the wild-type (WT) mice, low levels in the WT genotype, and no levels in the control tissues of the hemizygous and recessive homozygote genotypes (Fig. 6C, D). These results together suggest a possible origin on neural crest-derived progenitors of these induced sarcomas in the EMX-downregulated mice.

Finally, a molecular characterization of the genes of the stem cells was performed, and the *Emx1* and *Emx2* genes were measured by qRT-PCR to characterize the two murine KO models, as well as the *Nanog*, *Nes*, *Oct4*, *Sox2*, *Klf4*, and *Myc* genes (Fig. 7). In the murine *Emx1* model, a reduction in *Emx1* levels was obtained in the heterozygous *Emx*(+/–) mice and no expression was found in the *Emx*(–/–)-null homozygotes (Fig. 7), with lower expression levels in induced sarcoma (T) than in control muscle tissue (C). We observed an increase in the *Nanog*,



**Fig. 4 Effect of EMX1 and EMX2 levels on the expression of genes related to the phenotype of stem cells.** Quantification of the relative mRNA levels of the *EMX1*, *EMX2*, *OCT4*, *SOX2*, *KLF4*, *MYC*, *BMI1*, *NANOG*, *NES*, and *PROM1* genes by qRT-PCR in the different cell lines with different levels of EMX proteins. The graphs show the expression levels ( $2^{-\Delta C_t}$ ) of the different indicated genes for the total extract of the cell lines (ET), the tumorspheres (TO), and the xenograft tumors (XEN). **A** Line AA, **B** line AW, and **C** line BG. The average of a minimum of three independent experiments is represented in triplicate  $\pm$  SD. Statistical analysis was performed with Student's *t*-test (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ). **D** Transcriptional analysis of the genes *EMX1* and *EMX2*, and those related to the properties of stem cells in public databases. **D1** Analysis of the Yamanaka database (GSE9561), where the relative RNA levels of *EMX1* and *EMX2*, and the genes related to the properties of stem cells (*OCT4*, *SOX2*, *KLF4*, and *MYC*) were measured. The database showed the transcriptome of induced pluripotent stem cells (iPSCs) generated from human dermal fibroblasts (hDFs). **D2** Analysis of the Thomson database (GSE15148), where the relative RNA levels of *EMX1* and *EMX2*, and stem cell genes were measured in human iPSCs generated from fibroblasts of the foreskin, used as a control and in human embryonic stem cells (hESCs). Statistical analysis was performed with Student's *t*-test (\* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ).



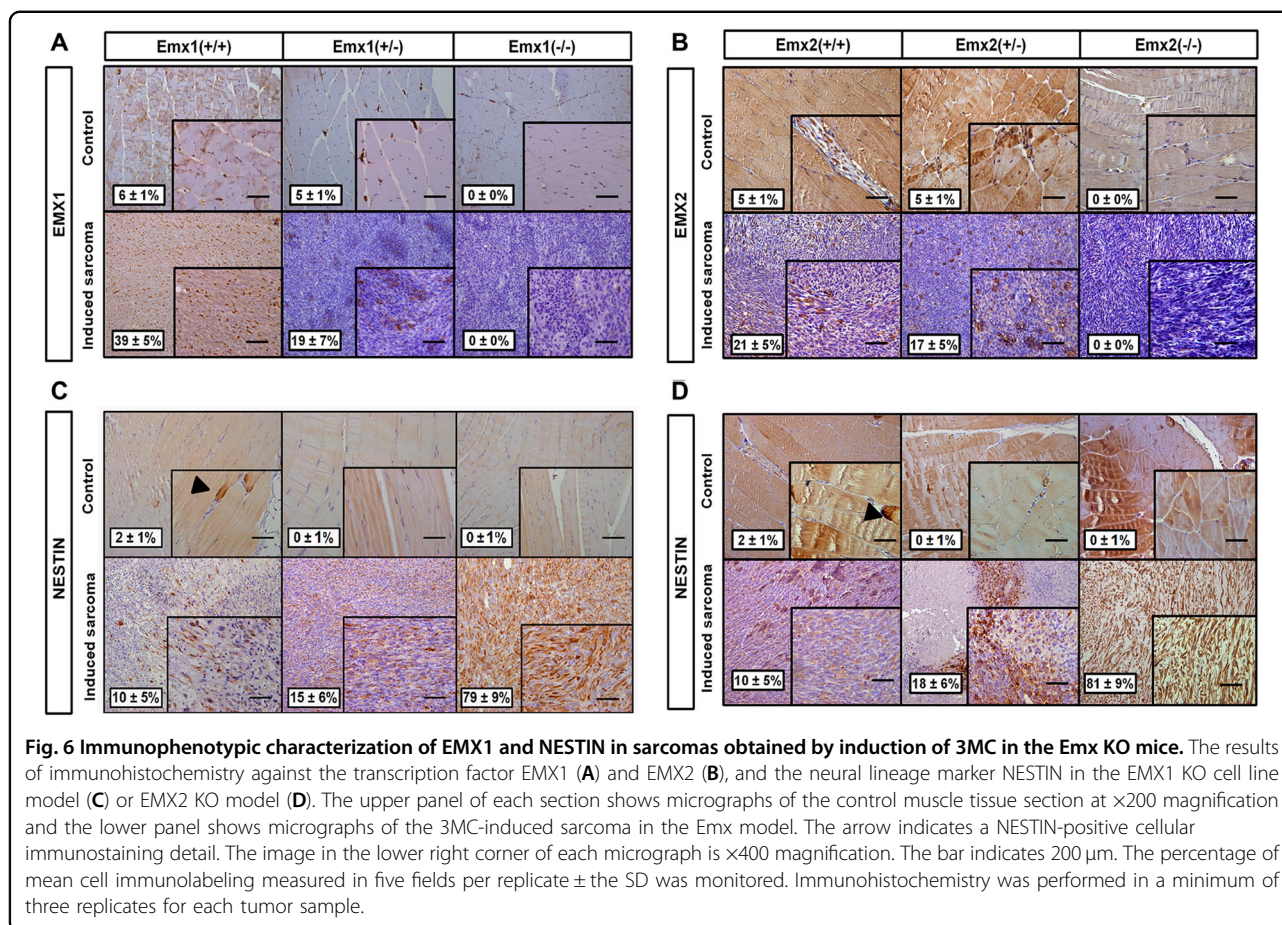
*Nes*, *Oct4*, *Sox2*, *Klf4*, and *Myc* mRNA levels in the mice with a functional allele of *Emx*(+/-) and especially in the null *Emx*(-/-), with higher levels in the induced sarcoma compared to the control muscle tissue (Fig. 7).

These results indicate that the reduction of Emx1/Emx2 allows the expression of genes related to the properties of stem cells (*Nanog*, *Nes*, *Oct4*, *Sox2*, *Klf4*, and *Myc*), confirming this overexpression in the case of induced sarcomas and in the null murine models of both genes. This result indicates that *Emx* genes are negatively correlated with genes related to stem cell properties in vivo. Analysis of overall survival in sarcoma databases with respect to EMX1 and EMX2 levels suggest that, in general terms, a reduction in disease-free survival or/and overall survival was observed in the group of patients with low levels of EMX1 and EMX2 (Supplementary Fig. FS6).

## Discussion

The transcription factors EMX1 and EMX2 act as tumor suppressors in sarcomas in vitro and in vivo by decreasing the expression of genes regulating the properties of stem cells and the stem cell phenotype. The tumor suppressor function of EMX found in this work is consistent with those previously obtained from the overexpression of EMX2 in lung cancers<sup>22,48,49</sup>, colorectal metastasis<sup>28</sup>, gastric cancer<sup>25</sup>, endometrial cancer<sup>26,27</sup>, and glioblastoma<sup>29,30</sup>. We found that the overexpression of EMX1 or EMX2 arrests CSCs, explaining the reduction in the regeneration of the culture or in the number and size of tumorspheres and tumorigenicity in vivo. This reduction in the CSC population dependent on EMX levels may be linked to reports associating a better prognosis of these tumors with EMX2 overexpression. The results indicated a reduction in the





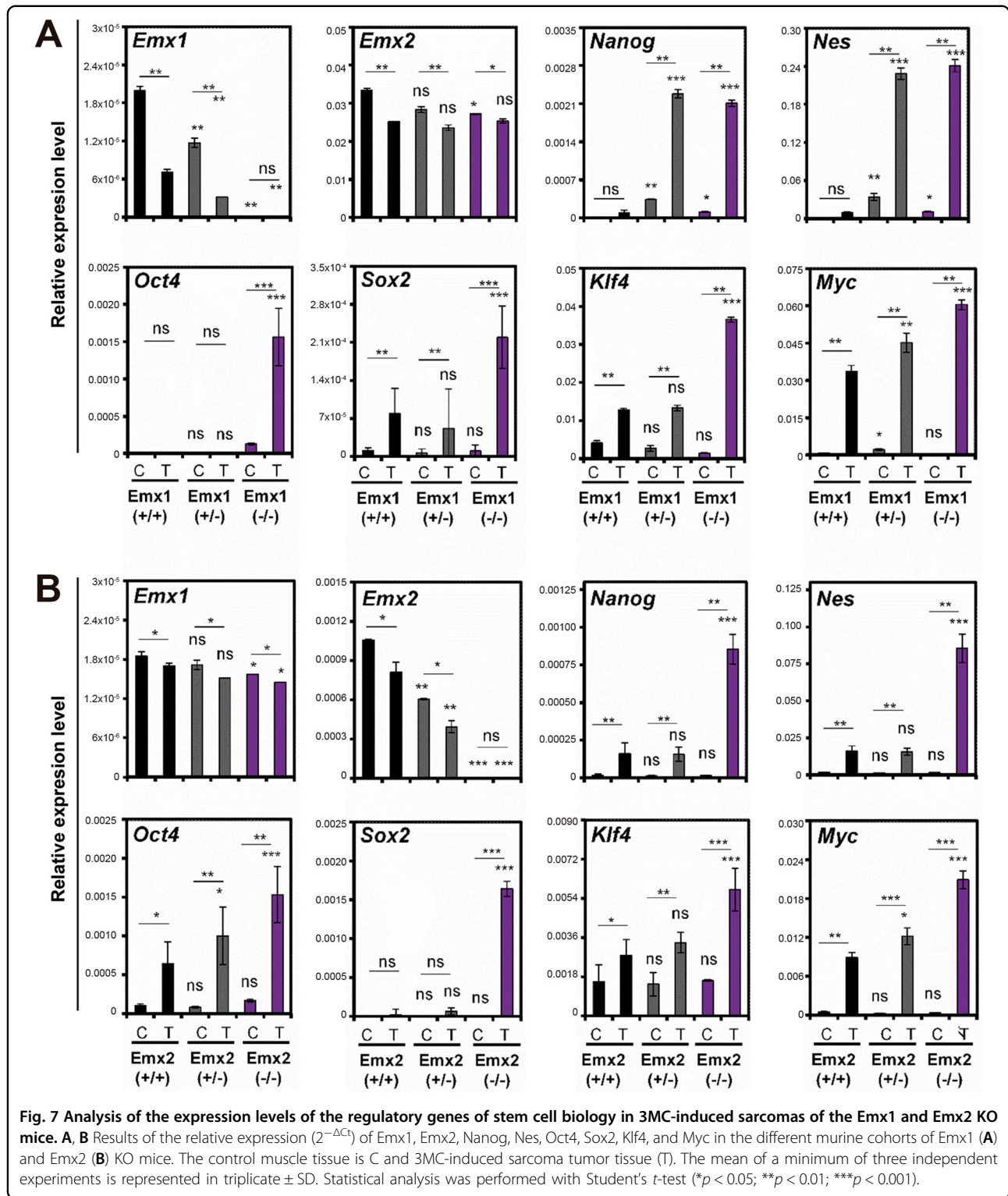
number, size, and efficiency of tumor formation, as well as a negative correlation of the CD133+ fraction with the EMX1 and EMX2 levels.

In addition, the negative correlation of stem cell genes (*OCT4*, *SOX2*, *KLF4*, *MYC*, *NANOG*, *NES*, and *PROM1*) with EMX1/EMX2 was verified in culture in vivo and in the in silico transcriptional analysis of the databases of embryonic and human induced pluripotent progenitors. Taken together, these results indicate that the EMX genes retain their ontogenetic function of blocking stemness. In the neural embryonic context, cell proliferation is reduced by directly blocking Sox2<sup>50,51</sup> and other precursors (such as Sox3 and Sox11) related to neural development. This process allows the creation of a gene expression profile that regulates the differentiation and migration of neural precursor cells in neurospheres cultured in vitro<sup>40</sup>.

Despite the different molecular processes by which sarcomagenesis is reported to occur, EMX1 and EMX2 negatively contribute to sarcomagenesis in vivo by regulating stemness. The 3MC carcinogen-induced sarcomas have a more aggressive phenotype, increased infiltration into the femoral bone marrow, and reduced

survival in the Emx1/Emx2 KO mice. This increased malignancy may be due to increased expression of stem cell genes in both the precursors of Emx1/Emx2 KO mice, as mature tissue does not have EMX expression (Fig. 6). In addition, the Emx1/Emx2 KO mice that are induced with sarcoma have decreased survival. These results can be explained if the induced sarcoma occurs as a consequence of the tumor transformation of the stem cells, which, in the absence of EMX proteins, facilitates the process of tumor transformation. One of the most representative neural proteins of stem cell regulation is NES<sup>52,53</sup>. A marked increase in NES was obtained in the tumors of both KO mice. These data correlate with our data for EMX1/EMX2 depletion found in the CD133+ fraction of human sarcoma primary lines. In other works, the presence of NES was associated with increased tumor malignancy and tumor stem cell populations in the tumor<sup>53</sup>.

Comparing the control muscle tissues and the induced sarcoma, we observed the highest expression of the stem cell genes (*Oct4*, *Sox2*, *Klf4*, *Myc*, and *Nanog*) and aggressivity in the sarcomas induced in the Emx1 or Emx2 KO mice. These results indicate that sarcomagenesis occurs



mainly in those populations or cell progenitors of the muscle tissue of KO animals that overexpress stem cell genes, such as NES. This overexpression of pluripotency genes in combination with the 3MC carcinogen enables

sarcomas to develop and tumor malignancy to increase, as demonstrated in other studies<sup>54–58</sup>. Therefore, all these data explain the negative correlation between the EMX genes and the properties of CSCs.

In embryonic studies of *Emx1* and *Emx2* in the development of neural progenitors in mice, expression was found in the dorsal region of the neural tube, allowing us to constitute an expression gradient that affects not only the neural tube but also the neural crest by direct proximity of the gradient<sup>40,50,51,59–61</sup>. Thus, as the process of ontogenetic development in the neural tube and neural crest progresses with the successive stages of cell differentiation, the levels of *Emx* expression gradually decrease, leading to the differentiation of the different tissue subtypes<sup>51,60</sup>. Therefore, in the progressive reduction of the expression of the *EMX* genes, it is expected that the expression of the stem cell genes will be allowed, as well as proliferative regulators. Thus, at this level of differentiation with loss of *EMX*, other signaling pathways or genes could participate in ontogenetic processes. *EMX* genes are essential in the early stages of development by inducing cell differentiation of different neural precursors.

The *EMX1/EMX2* genes are expressed in neural precursors (neuro-ectodermal), inducing cell differentiation, proliferation, and migration. After the development processes in which the *EMX* genes are believed to induce cell differentiation in early precursors, their hypermethylation and silencing in mature cells take place but are maintained in stem cells regulating their stem proliferative arrest. *EMX* lost by methylation in stem cells may be considered the first “hit” that allows the activation of certain genes in stem cells and further carcinogenic insult will promote tumorigenesis.

All together, our results suggest that sarcoma self-renewing populations originate from neural or mesodermal precursors where *EMX1* and/or *EMX2* expression is silenced. The processes of cellular differentiation during ontogenetic development produce the silencing of *EMX* genes by hypermethylation of their promoters in differentiating, arrested cells, whereas in individual stem cells from neural crest-derived progenitors, *EMX* levels are maintained at high levels, controlling proliferation.

We propose that the tumor suppressive effect of *EMX* might be extrapolated to the heterogeneity of sarcomas, although experiments in many other different sarcoma types should be performed to corroborate this hypothesis. This is because their behavior as tumor suppressors is robust in several different primary sarcoma lines (leiomyosarcoma, liposarcoma, and myxoid fibrosarcoma), xenografts thereof, a murine model of sarcomagenesis in vivo, and sarcoma tumors from the databases of analyzed patients. Sarcomas are a very heterogeneous group and it has been reported that they do not have a single embryological origin; most seem to derive from mesodermal precursors<sup>1,62–64</sup> and some from precursors of the neural crest<sup>3–5,65</sup>. Furthermore, head and neck sarcomas that derive exclusively from cartilage, bone, muscle, and adipose tissue have been proposed to originate from the

neural crest<sup>4</sup>. Furthermore, we proposed that *EMX1/EMX2* transcription factors could block the expression of essential stem cell genes during sarcomagenesis.

These findings may suggest why the cell population of neural-like and mesenchymal-like precursors with the enhanced multipotent properties are those with null expression of *EMX1/EMX2* and high levels of *NES*. Our model adapts to the existing developmental model, adding the data that certain populations of CSCs positive for the analyzed stem cell markers and dependent on *EMX* are the ones that can give rise to many sarcomas, regardless of the embryological origin, of the histology of the tumor, and the place of appearance. However, much more research is needed to confirm this point.

Until now, we have considered *EMX1* and *EMX2* as two genes with the same function based on the results obtained in our work. It should be noted that the scientific literature in general only mentions *EMX2*, either ontogenetically or oncologically. However, it is expected that *EMX1* and *EMX2* could have different functions, although they may be partially or completely redundant during sarcomagenesis.

In summary, our work showed that *EMX1* and *EMX2* act as tumor suppressors by suppressing the activity of stem cell regulatory genes (*OCT4*, *KLF4*, *MYC*, *SOX2*, *NANOG*, *NES*, and *PROM1*) in sarcoma. It is proposed that there is no or low expression of *EMX1* and *EMX2* in CSCs embedded in the tissue of origin of the sarcoma, allowing self-renewal of the tumor. Following overexpression of *EMX* transcription factors, there is a reduction in tumorigenesis and the capacities for self-renewal and maintenance of CSCs in primary sarcoma lines and in xenografts thereof in vivo. Murine KO models of *Emx1* or *Emx2* have been used to validate the initiation and development of sarcomagenesis in sarcomas induced with wild or hemizygous levels of *Emx1/Emx2*. The induced sarcoma of the KO group is more aggressive, more infiltrative, and has a greater capacity for tumor self-renewal than that of the WT group. Based on this evidence, *EMX* genes are proposed as possible new predictive biomarkers of response in sarcoma, where they play a fundamental role as tumor suppressor genes regulating various cellular genes.

## Materials and methods

### Transfections and plasmids

Subconfluent cells were transfected with TransIT-X2 reagent (Mirus) according to the manufacturer's instructions. At 48 h, the cells were seeded in 10 cm plates with media containing the appropriate selection drug (100–450 µg/ml G418, 0.25–0.4 µg/ml puromycin, or 0.25–0.4 µg/ml blasticidine). The cells were transfected with the plasmids from Table 1.

**Table 1 Plasmids used in this work.**

Plasmid	Insert sequence	R. eukar.	Origin
pRS-SC sh	-	Puro	Origene (TR313216)
pRS-sh1-EMX1	5'-GCTTCAATTAAGCCACAGTGTCTCCGAG-3'	Puro	Origene (TI363329)
pRS-sh4-EMX1	5'-GGCAGTCTCAGCCTCTCCGAGACGCAGGT-3'	Puro	Origene (TI363332)
pRS-sh1-EMX2+EMX1	5'-TCAAGCCATTACCAGGCTTCGGAGGAAG-3'	Puro	Origene (TI352857)
pRS-sh4-EMX2+EMX1	5'-CGGTGGAGAATCGCCACCAAGCAGGCGAG-3'	Puro	Origene (TI352860)
pCMV6-EV	-	G418	Origene (PS100001)
pCMV6-EMX1	EMX1 ADNc	G418	Origene (RC208006)
pCMV6-EMX2	EMX2 ADNc	G418	Origene (RC222758)

A. resistance, *Amp* ampicillin, *Blast* blasticidin, *EV* empty vector, *G418* geneticin, *Kan* kanamycin, *LAB* generated in the laboratory, *pRS* pRetroSuer, *Puro* puromycin, *SC* random sequence shDNA (scrambled shDNA), *T* transfection.

**Table 2 The cell lines used in this study were obtained from the ECACC commercial repository or generated in the laboratory as previously reported<sup>67–69</sup>.**

Cell line	Species	Tissue	Culture media	Growth	Origin
A-673	Human	Ewing's sarcoma	RPMI	Adherent	ATCC
SAOS-2	Human	Osteosarcoma	DMEM	Adherent	ATCC
SW-872	Human	Liposarcoma	DMEM	Adherent	ATCC
IMR-90	Human	Fibroblast*	DMEM	Adherent	ATCC
WI-38	Human	Fibroblast*	DMEM	Adherent	ATCC
CNIO-AA	Human	Leiomyosarcoma	F10	Adherent	LAB
CNIO-AW	Human	Liposarcoma	F10	Adherent	LAB
CNIO-AX	Human	Liposarcoma	F10	Adherent	LAB
CNIO-AZ	Human	Fibrous sarcoma	F10	Adherent	LAB
CNIO-BC	Human	MPNST	F10	Adherent	LAB
CNIO-BD	Human	Ewing's sarcoma	F10	Adherent	LAB
CNIO-BG	Human	Mixoid fibrosarcoma	F10	Adherent	LAB
CNIO-BO	Human	Mixoid fibrosarcoma	F10	Adherent	LAB
CNIO-BP	Human	Osteosarcoma	F10	Adherent	LAB
CNIO-CE	Human	Rabdomyosarcoma	F10	Adherent	LAB
CNIO-DA	Human	Mixoid fibrosarcoma	F10	Adherent	LAB
CNIO-DD	Human	Mixoid fibrosarcoma	F10	Adherent	LAB

\*Primary fibroblast.

### Cell culture

The cell lines were obtained from the European Collection of Authenticated Cell Cultures (ECACC) commercial repository or generated in the laboratory as previously reported<sup>66–68</sup>. No further authentication was conducted by the authors. The cells were negative for mycoplasma. The cell lines were maintained in Dulbecco's modified Eagle's medium, F10, or RPMI (AQmedia; Sigma) as indicated in Table 2, supplemented with 10%

fetal bovine serum (FBS) (Gibco), penicillin, streptomycin, and fungizone (Sigma).

### Transfection using the commercial TransIt-X2 kit (Mirus)

Cells were seeded in six-well plates the day before, so that they were at 80% confluence at the time of transfection. The procedure consists of mixing 2.5 µg of plasmid DNA and 250 µl of medium (amounts to transfect cells in one well). Then, 7.5 µl of TransIT-X2, a non-liposomal polymer that

coats the DNA favoring its transport into the cell, is added, mixed with the pipette, and incubated for 30 min at room temperature. Subsequently, the mixture is added dropwise on the cells, which were incubated for 48 h at 37 °C to then start the selection process.

After 24–48 h of incubation after transfection or infection, the cells were trypsinized, seeded at low density from a 3 cm plate to a 10 cm plate in the case of transfected cells, and the selection with the required antibiotic was started. The selection was maintained for the time necessary for the untransfected cells to die and subcultures were obtained on the plates of the transfected cells. Once the selection was completed, for maintenance the concentration of the antibiotic was reduced by half.

#### **For either EMX1 or EMX2 overexpression models with pCMV6 plasmids:**

AA cell line: 1 mg/ml G418 (half concentrated for maintenance: 0.5 mg/ml);

AW cell line: 1 mg/ml G418 (half concentrated for maintenance: 0.5 mg/ml);

Ham's F10 medium supplemented with 10% FBS and 1 mg/ml G418 for selection, and after selection 0.5 mg/ml for maintaining stable cell line.

Selection time: ~10–15 days (refreshing the culture medium every 2 days). Taking into account that EMX1 and EMX2 behave as tumor suppressors, clones that overexpress either EMX1 or EMX2 can be progressively lost, so that quantitative PCR (qPCR) measurements are periodically (once every 1–2 weeks) performed to see expression levels, otherwise the transfection was repeated.

#### **For knockdown BG model with pRS plasmids:**

BG cell line: 1 µg/ml G418 (half concentrated for maintenance: 0.5 µg/ml);

Ham's F10 medium supplemented with 10% FBS and 1 µg/ml puromycin for selection for maintaining stable cell line.

Selection time: ~8–12 days (refreshing the culture medium every 2 days). Taking into account that knocking down EMX levels promotes a proliferative phenotype, this model tends not to be lost; in any case, qPCR measurements are periodically performed (once every 1–2 weeks) to see expression levels, otherwise the transfection was repeated.

In all cases 1/2 dosis of the antibiotic for selection was used for continuous culture (maintenance), to ensure the maintenance of the transfection.

#### **Tumorigenesis assays**

Proliferation assays, clonability assays, and growth assays in soft agar, and cell migration and invasion assays were performed as indicated in refs. <sup>69,70</sup>.

#### **Tumorsphere assay**

A total of 5000 cells were seeded in triplicate in 24-well Ultra-Low Attachment Plates (Costar) containing 1 ml of MammoCult basal medium (Stem Cell Technologies) supplied with 10% MammoCult proliferation supplement, 4 µg/ml heparin, 0.48 µg/ml hydrocortisone, penicillin, and streptomycin. After 5–10 days, depending on the cell line (AA in 5–7 days, AW in 6–10 days, BG in 7–8 days), the number of primary tumorspheres formed was measured using an inverted microscope (Olympus IX-71).

Subsequently, the secondary tumorspheres were generated by collecting the medium with the tumorspheres, centrifuging for 5 min at 900 r.p.m., trypsinizing for 5 min, and centrifuging again at 900 r.p.m. for 5 min. Finally, all the cells trypsinized were resuspended in fresh, complete, supplemented MammoCult medium and 1 ml/well was seeded in a 24-well, low-stick plate. They were incubated for more 4–5 days and the number and size of the tumorspheres were measured again and quantified.

#### **Single-cell tumorsphere assay**

Single cells were individually seeded through cell sorting with a FACS Jazz flow cytometer (BD Biosciences) in 96-well Ultra-Low Attachment Plates containing 1 ml of MammoCult basal medium (Stem Cell Technologies) supplied with 10% MammoCult proliferation supplement, 4 µg/ml of heparin, 0.48 µg/ml of hydrocortisone, penicillin, and streptomycin. After 30 days, the number of individual primary tumorspheres formed was measured using an inverted microscope (Olympus IX-71).

#### **Analysis and classification of clonal phenotypes**

The cells were seeded at low density, 100 or 1000 depending on the cell line, in triplicate in 10 cm plates and cultured under standard conditions. The medium was changed twice a week. After 10–15 days, once the clones were formed, the morphology of each clone was observed under an inverted microscope (Olympus CKX41). Finally, each type of clone was counted based on the following classification: holoclone (more compact and rounded clone with a higher percentage of CSCs), meroclone (clone with a more irregular cell arrangement at the edges and a lower proportion of CSCs), or paraclone (greater separation of cells from each other, due to a greater degree of differentiation and lower percentage of CSCs)<sup>71,72</sup>.

#### **Marker analysis by flow cytometry**

The previously trypsinized suspension of  $1 \times 10^6$  cells was resuspended in 125 µl of phosphate-buffered saline (PBS) with 2% FBS and 5 mM EDTA. This suspension was then blocked by adding 12.5 µl of blocking agent (Miltenyi Biotec) and incubated for 10 min on ice. The anti-CD133 antibody conjugated to fluorochrome phycoerythrin (MACS) was then added at a 1 : 25 dilution and incubated

**Table 3 Probes for Q-PCR used in this study.**

HumanGene	Probe (human)	Mouse gene	Probe (mouse)
<i>GAPDH</i>	Hs03929097_g1	<i>Gadph</i>	Mm99999915_g1
<i>EMX1</i>	Hs00417957_m1	<i>Emx1</i>	Mm01182609_m1
<i>EMX2</i>	Hs00244574_m1	<i>Emx2</i>	Mm00550241_m1
<i>CD133</i>	Hs01009257_m1		
<i>NES</i>	Hs04187831_g1	<i>Nes</i>	Mm00450205_m1
<i>NANOG</i>	Hs04260366_g1	<i>Nanog</i>	Mm02019550_s1
<i>OCT4</i>	Hs00999632_g1	<i>Oct4</i>	Mm03053917_g1
<i>SOX2</i>	Hs01053049_s1	<i>Sox2</i>	Mm03053810_s1
<i>KLF4</i>	Hs00358836_m1	<i>Klf4</i>	Mm00516104_m1
<i>MYC*</i>	Hs00153408_m1	<i>c-Myc</i>	Mm00487804_m1
<i>BMI1</i>	Hs00995536_m1		

\*Primary fibroblast.

for 30 min in ice and darkness. A tube with  $1 \times 10^6$  cells was left unlabeled with the antibody to use as a negative control for staining. After the end of the incubation period, the cells were washed two times with PBS, with 2% FBS and 5 mM EDTA. Subsequently, the cells were centrifuged for 5 min at a speed of 1000 r.p.m. and resuspended in 500  $\mu$ l of PBS with 2% FBS and 5 mM EDTA. Finally, the cell suspension was examined in a Canto II analytical cytometer (BD Biosciences), separating in certain cases two populations positive and negative for the CD133 marker.

#### Reverse-transcriptase quantitative PCR

Total RNA from cell lines was extracted and purified using the miRNeasy kit (Qiagen) and reverse transcription was performed with 3  $\mu$ g of mRNA using the High Capacity cDNA Reverse Transcription kit (Life Technologies) according to the manufacturer's instructions. We used the following probes (Table 3):

#### Protein isolation and western blot analysis

Western blottings were performed as previously described elsewhere<sup>69–72</sup>. Membranes were incubated with the following primary antibodies (Table 4):

The proteins were detected using an ECL detection system (Amersham Biosciences) and a Bio-Rad Chemidoc Touch. Bio-Rad's Image Lab 5.1 program was used to quantify the protein bands. Each band was selected individually and the value was reported as a loading control on the same membrane, which was usually immunodetection for  $\alpha$ -tubulin.

#### Xenografts in nude mice

Tumorigenicity was assayed by the subcutaneous injection of  $4 \times 10^6$  cells into the right flanks of 4-week-

old female athymic nude mice. Cells were suspended in 50  $\mu$ l of Matrigel (Corning) prior to the injection. Animals were examined weekly and after 150–180 days, depending on the cell lines, the mice were killed and the tumors were extracted and conserved at  $-80^\circ\text{C}$ . Tumorsphere tumorigenicity was measured by seeding  $10^4$  cells as described in the tumorsphere assay section. After 5 days, tumorspheres were disaggregated with trypsin, resuspended in 50  $\mu$ l of Matrigel, and injected into the right flanks of 4-week-old female athymic nude mice. Animals were examined weekly and after 85–100 days, depending on the cell lines, mice were killed and tumors were extracted and conserved at  $-80^\circ\text{C}$ . Tumor volume ( $\text{mm}^3$ ) was measured using calipers. All animal experiments were performed according to the experimental protocol approved by the IBIS and HUVR Institutional Animal Care and Use Committee (0309-N-15).

#### Mouse EMX KO lines used and maintenance of the experimental colonies

The KO mice were provided by the Japanese Natural Science Research Center RIKEN (<http://www.riken.jp/en/>). They were generated from TT2 embryonic stem cells derived from an F1 embryo mixed between C57BL/6J and CBA/J. The mice of the KO-Emx1 line were kept in a mixed C57BL/6J  $\times$  CBA/J genetic background and the KO-Emx2 line in a pure C57BL/6J genetic background. The KO lines for Emx1 and Emx2 were maintained and crossed in heterozygosity to obtain the different genotypes of Emx1 or Emx2 in the same litter (Emx1/Emx2 (+/+), Emx1/Emx2 (+/-), and Emx1/Emx2 (-/-)), and thus reduce possible genetic variability. All mice were kept according to the standards established in the IBiS animal facility (based on the provisions of Royal Decree 53/2013) and were killed by CO<sub>2</sub> inhalation when any significant sign of disease was detected to avoid the suffering of the animal.

#### Carcinogenic tests

Treatment with 3MC was carried out to induce sarcomas in the abductor muscle of the murine model. For this, a solution of the carcinogen 3MC (Sigma-Merck) in sesame oil (Sigma-Merck) was prepared at a concentration of 10 mg/ml. Three-month-old mice were treated, inoculating a weekly dose of 1 mg of 3MC (100  $\mu$ l), performing two doses intramuscularly in the abductor muscle of the right hind leg. The mice were examined once a week and were killed when the tumor reached a size of 1.7  $\text{cm}^3$  or showed signs of disease.

#### Necropsy

After the killing of each animal, a complete necropsy of each animal was performed and samples of the tumor and affected tissues were collected for histological and molecular analysis. The tumor samples were divided into three

**Table 4 Antibodies used in this study.**

Antibody	Company		Dilution		
			WB	IHQ	
				DIL	DA
EMX1 (rabbit polyclonal)	Abcam	ab136102	1:1000	1:100	EDTA
EMX2 (mouse polyclonal)	Abcam	ab171818	1:1000	-	-
EMX2 (rabbit polyclonal)	Invitrogen	PA5-34415	-	1:500	TC
$\alpha$ -tubulin (mouse monoclonal)	Sigma-Merck	T9026	1:10,000	-	-
KI67 (rabbit monoclonal)	MAD	000310-QD	-	1:250	TC
NESTIN (rat polyclonal)	SCB	SC-101541	-	1:100	TC
Rabbit anti-mouse HRP (polyclonal) <sup>a</sup>	Abcam	ab97046	1:5000	-	-
Goat anti-rabbit HRP (polyclonal) <sup>a</sup>	Abcam	ab97051	1:5000	-	-
Goat anti-rabbit HRP (polyclonal) <sup>a</sup>	JAC	111-035-003	-	1:400	-
Goat anti-rat HRP (polyclonal) <sup>a</sup>	JAC	112-035-003	-	1:400	-
Rabbit anti-goat HRP (polyclonal) <sup>a</sup>	Abcam	ab97100	-	1:400	-

CST Cell Signaling Technologies, DA antigen unmasking, DIL dilution, EDTA buffer EDTA, IHQ immunohistochemistry, JAC Jackson Immuno Research, MAD Master Diagnostica, TC citrate buffer, SCB Santa Cruz Biotechnology, WB western blotting.

<sup>a</sup>Secondary antibodies.

parts: one for histology and the others for RNA or protein analysis. Histology samples were fixed in 25% formalin for 24 h. After a dehydration process with ethanol at different concentrations and xylol, they were introduced into paraffin blocks at 65 °C, obtaining blocks from which 2  $\mu$ m sections were made in an automatic microtome, which were stained with hematoxylin and eosin or immunostained with different antibodies, as described in section 7.5. The two pieces of the tumor destined for RNA and protein analysis were frozen in dry ice in cryotubes and were stored at -80 °C until pulverization with a mortar and liquid nitrogen.

### Immunohistochemistry

Samples were extracted from the different tissues of the mice and fixed in buffered 25% formalin for 24 h. After a dehydration process with ethanol at different concentrations and xylol, the samples were introduced in paraffin at 65 °C. We obtain 2  $\mu$ m sections in an automatic microtome from each paraffin block. For staining with the antibodies KI67, CD45, F4/80, EMX2, and NES, the samples were rehydrated by the usual procedure and antigenic recovery was performed with citrate buffer at pH 6.5 (Dako/Agilent) heating in a pressure cooker for 2 min. In the case of the EMX1 antibody, antigenic recovery was carried out with EDTA buffer at pH 9 (Dako/Agilent). All primary antibodies were incubated overnight (12 h) in a humid chamber and at 4 °C, following the dilutions (Table 4). Then, the corresponding secondary antibody was added 1 h and a half

at 1:400 in a humid chamber and at room temperature. The samples were washed with TBS + 0.1% Tween 20. After the endogenous peroxidase blocked with hydrogen peroxide and methanol (Dako/Agilent), 1% blocking solution was added (Sigma-Merck) and development was performed with the peroxidase system using diaminobenzidine plus (Dako/Agilent) as substrate. Counterstaining was performed with Harris hematoxylin. Finally, the sample was completely dehydrated with xylol and mounting was performed with Pertex. Photos were taken with an Olympus BX-61 microscope. The evaluation of the samples was carried out by quantitative microscopic analysis. Percentage values of immunostaining were assigned with respect to five fields.

### Analysis in public databases

The expression levels of EMX1/EMX2 and genes related to stem cell properties in public databases of human embryonic stem cells and human induced pluripotent stem cells were examined:

- Yamanaka, GEO ID: GSE9561.
- Thomson, GEO ID: GSE15148.

### Statistical analysis

We used the GraphPad Prism 6 software program for all statistical analyses of the experiments performed with the cell lines and the in vivo experiments. To analyze the differences in all the functional or transcriptomic tests

carried out with the lines transfected with the cDNA of EMX1, EMX2, shRNAs, and the lines transfected with the different control vectors, we used Student's *t*-test for unpaired samples or Student's *t*-test with Welch's correction. To determine the statistical significance of the survival graphs, we used a  $\chi^2$ -test (LogRank test). *P*-values < 0.05 were considered statistically significant and were represented according to the following classification: \**p* < 0.05, \*\**p* < 0.01, and \*\*\**p* < 0.001.

#### Acknowledgements

The KO mice were provided by the Japanese natural science research center RIKEN (<http://www.riken.jp/en/>). We thank RIKEN for their help. We thank Sandra Muñoz-Galvan for help with figures and critical reading of this manuscript.

#### Author details

<sup>1</sup>Instituto de Biomedicina de Sevilla (IBIS), Hospital Universitario Virgen del Rocío, Universidad de Sevilla, Consejo Superior de Investigaciones Científicas, Sevilla, Spain. <sup>2</sup>CIBER de Cáncer, IS Carlos III, Madrid, Spain. <sup>3</sup>Department of Cell Growth and Differentiation, Center for iPS Cell Research and Application, Kyoto University, Kyoto, Japan

#### Author contributions

A.C. conceived and designed this study. M.P.J.-G., D.O.-A., and A.L.-C. designed and performed the experiments. M.P.J.-G. and A.C. analyzed and interpreted the data. A.C. drafted the manuscript. All authors revised and edited the manuscript.

#### Funding

This work was supported by grants from the Ministerio de Ciencia, Innovación y Universidades (MCIU) Plan Estatal de I + D + I 2018, a la Agencia Estatal de Investigación (AEI) y al Fondo Europeo de Desarrollo Regional (MCIU/AEI/FEDER, UE): RTI2018-097455-B-I00; from AEI-MICIU/FEDER (RED2018-102723-T); from CIBER de Cáncer (CB16/12/00275), co-funded by FEDER from Regional Development European Funds (European Union); and from Consejería de Salud (PI-0397-2017) and Consejería of Economía, Conocimiento, Empresas y Universidad of the Junta de Andalucía (P18-RT-2501). Also special thanks to the Fundación AECC for supporting this work. M.P.J.-G. was funded by a FPU contract from the Ministerio de Educación (FPU13/00426).

#### Data availability

No datasets were generated during the current study. The datasets analyzed during the current study are publicly available in the different repositories.

#### Ethics statement

All methods were performed in accordance with the relevant guidelines and regulations of the Institute for Biomedical Research of Seville (IBIS) and University Hospital Virgen del Rocío (HUVR). Animal experiments were performed according to the experimental protocol approved by HUVR Animals Ethics (CEI 0309-N-15). No human trials were performed.

#### Conflict of interest

The authors declare no competing interests.

#### Publisher's note

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

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41419-021-03801-w>.

Received: 10 March 2021 Revised: 3 May 2021 Accepted: 3 May 2021  
Published online: 20 May 2021

#### References

- Mackall, C. L., Meltzer, P. S. & Helman, L. J. Focus on sarcomas. *Cancer Cell* **2**, 175–178 (2002).
- Jo, V. Y. & Fletcher, C. D. WHO classification of soft tissue tumours: an update based on the 2013 (4th) edition. *Pathology* **46**, 95–104 (2014).
- Dupin, E. & Sommer, L. Neural crest progenitors and stem cells: from early development to adulthood. *Dev. Biol.* **366**, 83–95 (2012).
- Maguire, L. H., Thomas, A. R. & Goldstein, A. M. Tumors of the neural crest: common themes in development and cancer. *Dev. Dyn.* **244**, 311–322 (2015).
- Yamada, K. et al. EWS/ATF1 expression induces sarcomas from neural crest-derived cells in mice. *J. Clin. Invest.* **123**, 600–610 (2013).
- Shinagawa, K. et al. The characteristics of human cranial bone marrow mesenchymal stem cells. *Neurosci. Lett.* **606**, 161–166 (2015).
- Mehrotra, P., Tseropoulos, G., Bronner, M. E. & Andreadis, S. T. Adult tissue-derived neural crest-like stem cells: Sources, regulatory networks, and trans-lational potential. *Stem Cells Transl. Med.* **9**, 328–341 (2020).
- Piacentino, M. L., Li, Y. & Bronner, M. E. Epithelial-to-mesenchymal transition and different migration strategies as viewed from the neural crest. *Curr. Opin. Cell Biol.* **66**, 43–50 (2020).
- Hovland, A. S., Rothstein, M. & Simoes-Costa, M. Network architecture and regulatory logic in neural crest development. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **12**, e1468 (2020).
- Sauka-Spengler, T. & Bronner-Fraser, M. A gene regulatory network orchestrates neural crest formation. *Nat. Rev. Mol. Cell Biol.* **9**, 557–568 (2008).
- Ziermann, J. M., Diogo, R. & Noden, D. M. Neural crest and the patterning of vertebrate craniofacial muscles. *Genesis* **56**, e23097 (2018).
- Leonard, C. E. & Taneyhill, L. A. The road best traveled: Neural crest migration upon the extracellular matrix. *Semin. Cell Dev. Biol.* **100**, 177–185 (2020).
- von Levetzow, C. et al. Modeling initiation of Ewing sarcoma in human neural crest cells. *PLoS ONE* **6**, e19305 (2011).
- Mii, Y. et al. Neural crest origin of clear cell sarcoma of tendons and aponeuroses. Ultrastructural and enzyme cytochemical study of human and nude mouse-transplanted tumours. *Virchows Arch. A Pathol. Anat. Histopathol.* **415**, 51–60 (1989).
- Mathew, T. Evidence supporting neural crest origin of an alveolar soft part sarcoma: an ultrastructural study. *Cancer* **50**, 507–514 (1982).
- Ishibe, T., Nakayama, T., Aoyama, T., Nakamura, T. & Toguchida, J. Neuronal differentiation of synovial sarcoma and its therapeutic application. *Clin. Orthop. Relat. Res.* **466**, 2147–2155 (2008).
- Cecchi, C. & Boncinelli, E. Emx homeogenes and mouse brain development. *Trends Neurosci.* **23**, 347–352 (2000).
- Sen, S., Reichert, H. & VijayRaghavan, K. Conserved roles of ems/Emx and otd/Otx genes in olfactory and visual system development in *Drosophila* and mouse. *Open Biol.* **3**, 120177 (2013).
- Piper, M. et al. Emx and Nfi genes regulate cortical development and axon guidance in the telencephalon. *Novartis Found. Symp.* **288**, 230–242 (2007). discussion 42–5, 76–81.
- Matsuo, I. et al. Otx and Emx functions in patterning of the vertebrate rostral head. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 545–553 (1997).
- Tole, S., Goudreau, G., Assimacopoulos, S. & Grove, E. A. Emx2 is required for growth of the hippocampus but not for hippocampal field specification. *J. Neurosci.* **20**, 2618–2625 (2000).
- Okamoto, J. et al. Downregulation of EMX2 is associated with clinical outcomes in lung adenocarcinoma patients. *Clin. Lung Cancer* **12**, 237–244 (2011).
- Kim, M. S., An, C. H., Yoo, N. J. & Lee, S. H. Rare somatic mutation and loss of expression of EMX2 gene in common solid cancers. *APMIS* **119**, 733–734 (2011).
- Giroux Leprieur, E. et al. The homeobox gene EMX2 is a prognostic and predictive marker in malignant pleural mesothelioma. *Lung Cancer* **85**, 465–471 (2014).
- Li, J. et al. Adenoviral delivery of the EMX2 gene suppresses growth in human gastric cancer. *PLoS ONE* **7**, e45970 (2012).
- Peiffer-Schneider, S. et al. Mapping an endometrial cancer tumor suppressor gene at 10q25 and development of a bacterial clone contig for the consensus deletion interval. *Genomics* **52**, 9–16 (1998).



27. Daftary, G. S. & Taylor, H. S. EMX2 gene expression in the female reproductive tract and aberrant expression in the endometrium of patients with endometriosis. *J. Clin. Endocrinol. Metab.* **89**, 2390–2396 (2004).
28. Aykut, B. et al. EMX2 gene expression predicts liver metastasis and survival in colorectal cancer. *BMC Cancer* **17**, 555 (2017).
29. Monnier, A. et al. The expression of EMX2 lead to cell cycle arrest in glioblastoma cell line. *BMC Cancer* **18**, 1213 (2018).
30. Falcone, C., Daga, A., Leanza, G. & Mallamaci, A. Emx2 as a novel tool to suppress glioblastoma. *Oncotarget* **7**, 41005–41016 (2016).
31. Bordogna, W. et al. EMX homeobox genes regulate microphthalmia and alter melanocyte biology. *Exp. Cell Res.* **311**, 27–38 (2005).
32. Zimmerman, L. et al. Independent regulatory elements in the nestin gene direct transgene expression to neural stem cells or muscle precursors. *Neuron* **12**, 11–24 (1994).
33. Hamasaki, T., Leingartner, A., Ringstedt, T. & O'Leary, D. D. EMX2 regulates sizes and positioning of the primary sensory and motor areas in neocortex by direct specification of cortical progenitors. *Neuron* **43**, 359–372 (2004).
34. Gangemi, R. M. et al. Emx2 in adult neural precursor cells. *Mech. Dev.* **109**, 323–329 (2001).
35. Konrad, C. V., Murali, R., Varghese, B. A. & Nair, R. The role of cancer stem cells in tumor heterogeneity and resistance to therapy. *Can. J. Physiol. Pharmacol.* **95**, 1–15 (2017).
36. Orecchioni, S. & Bertolini, F. Characterization of cancer stem cells. *Methods Mol. Biol.* **1464**, 49–62 (2016).
37. Garcia-Heredia, J. M. & Carnero, A. Role of mitochondria in cancer stem cell resistance. *Cells* **9**, 1693 (2020).
38. Munoz-Galvan, S. & Carnero, A. Targeting cancer stem cells to overcome therapy resistance in ovarian cancer. *Cells* **9**, 1402 (2020).
39. Locke, M., Heywood, M., Fawell, S. & Mackenzie, I. C. Retention of intrinsic stem cell hierarchies in carcinoma-derived cell lines. *Cancer Res.* **65**, 8944–8950 (2005).
40. Beaver, C. M., Ahmed, A. & Masters, J. R. Clonogenicity: holoclones and meroclones contain stem cells. *PLoS ONE* **9**, e89834 (2014).
41. Klemba, A. et al. Surface markers of cancer stem-like cells of ovarian cancer and their clinical relevance. *Contemp. Oncol.* **22**, 48–55 (2018).
42. Werner, S., Stenzl, A., Pantel, K. & Todenhofer, T. Expression of epithelial mesenchymal transition and cancer stem cell markers in circulating tumor cells. *Adv. Exp. Med. Biol.* **994**, 205–228 (2017).
43. de Jong, J. & Looijenga, L. H. Stem cell marker OCT3/4 in tumor biology and germ cell tumor diagnostics: history and future. *Crit. Rev. Oncog.* **12**, 171–203 (2006).
44. Saunders, A., Faiola, F. & Wang, J. Concise review: pursuing self-renewal and pluripotency with the stem cell factor Nanog. *Stem Cells* **31**, 1227–1236 (2013).
45. Hirai, H., Karian, P. & Kikyo, N. Regulation of embryonic stem cell self-renewal and pluripotency by leukaemia inhibitory factor. *Biochem. J.* **438**, 11–23 (2011).
46. Baquet, Z. C., Gorski, J. A. & Jones, K. R. Early striatal dendrite deficits followed by neuron loss with advanced age in the absence of anterograde cortical brain-derived neurotrophic factor. *J. Neurosci.* **24**, 4250–4258 (2004).
47. Kusaka, M. et al. Abnormal epithelial cell polarity and ectopic epidermal growth factor receptor (EGFR) expression induced in Emx2 KO embryonic gonads. *Endocrinology* **151**, 5893–5904 (2010).
48. Yue, D. et al. EMX2 is a predictive marker for adjuvant chemotherapy in lung squamous cell carcinomas. *PLoS ONE* **10**, e0132134 (2015).
49. Okamoto, J. et al. EMX2 is epigenetically silenced and suppresses growth in human lung cancer. *Oncogene* **29**, 5969–5975 (2010).
50. Mariani, J. et al. Emx2 is a dose-dependent negative regulator of Sox2 telencephalic enhancers. *Nucleic Acids Res.* **40**, 6461–6476 (2012).
51. Galli, R. et al. Emx2 regulates the proliferation of stem cells of the adult mammalian central nervous system. *Development* **129**, 1633–1644 (2002).
52. Krupkova, O. Jr. et al. Analysis of nuclear nestin localization in cell lines derived from neurogenic tumors. *Tumour Biol.* **32**, 631–639 (2011).
53. Krupkova, O. Jr., Loja, T., Zambo, I. & Veselska, R. Nestin expression in human tumors and tumor cell lines. *Neoplasma* **57**, 291–298 (2010).
54. Garcia-Heredia, J. M. & Carnero, A. The cargo protein MAP17 (PDZK1IP1) regulates the immune microenvironment. *Oncotarget* **8**, 98580–98597 (2017).
55. Garcia-Heredia, J. M. & Carnero, A. Dr. Jekyll and Mr. Hyde: MAP17's up-regulation, a crosspoint in cancer and inflammatory diseases. *Mol. Cancer* **17**, 80 (2018).
56. Jimenez-Garcia, M. P. et al. Inflammation and stem markers association to PIM1/PIM2 kinase-induced tumors in breast and uterus. *Oncotarget* **8**, 58872–58886 (2017).
57. Narlik-Grassow, M. et al. The essential role of PIM kinases in sarcoma growth and bone invasion. *Carcinogenesis* **33**, 1479–1486 (2012).
58. Narlik-Grassow, M. et al. Conditional transgenic expression of PIM1 kinase in prostate induces inflammation-dependent neoplasia. *PLoS ONE* **8**, e60277 (2013).
59. Falcone, C. & Mallamaci, A. Tuning of neocortical astrogenesis rates by Emx2 in neural stem cells. *Neural Regen. Res.* **10**, 550–551 (2015).
60. Gangemi, R. M. et al. Effects of Emx2 inactivation on the gene expression profile of neural precursors. *Eur. J. Neurosci.* **23**, 325–334 (2006).
61. Shinozaki, K., Yoshida, M., Nakamura, M., Aizawa, S. & Suda, Y. Emx1 and Emx2 cooperate in initial phase of archipallium development. *Mech. Dev.* **121**, 475–489 (2004).
62. Taylor, B. S. et al. Advances in sarcoma genomics and new therapeutic targets. *Nat. Rev. Cancer* **11**, 541–557 (2011).
63. Gibault, L. et al. New insights in sarcoma oncogenesis: a comprehensive analysis of a large series of 160 soft tissue sarcomas with complex genomics. *J. Pathol.* **223**, 64–71 (2011).
64. Chibon, F. et al. Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat. Med.* **16**, 781–787 (2010).
65. Dupin, E. & Coelho-Aguiar, J. M. Isolation and differentiation properties of neural crest stem cells. *Cytom. A* **83**, 38–47 (2013).
66. Moneo, V. et al. Levels of active tyrosine kinase receptor determine the tumor response to Zolypsis. *BMC Cancer* **14**, 281 (2014).
67. Perez, M., Munoz-Galvan, S., Jimenez-Garcia, M. P., Marin, J. J. & Carnero, A. Efficacy of CDK4 inhibition against sarcomas depends on their levels of CDK4 and p16ink4 mRNA. *Oncotarget* **6**, 40557–40574 (2015).
68. Perez, M. et al. Efficacy of bortezomib in sarcomas with high levels of MAP17 (PDZK1IP1). *Oncotarget* **7**, 67033–67046 (2016).
69. Ferrer, I. et al. Loss of the tumor suppressor spinophilin (PPP1R9B) increases the cancer stem cell population in breast tumors. *Oncogene* **35**, 2777–2788 (2016).
70. Felipe-Abrio, B., Verdugo-Sivianes E. M., Saez, C. & Carnero, A. Loss of MYBBP1A Induces Cancer Stem Cell Activity in Renal Cancer. *Cancers (Basel)* **11**, 235 (2019).
71. Lucena-Cacace, A., Otero-Albiol, D., Jimenez-Garcia, M. P., Munoz-Galvan, S. & Carnero, A. NAMPT is a potent oncogene in colon cancer progression that modulates cancer stem cell properties and resistance to therapy through Sirt1 and PARP. *Clin. Cancer Res.* **24**, 1202–1215 (2018).
72. Garcia-Heredia, J. M., Lucena-Cacace, A., Verdugo-Sivianes, E. M., Perez, M. & Carnero, A. The cargo protein MAP17 (PDZK1IP1) regulates the cancer stem cell pool activating the Notch pathway by abducting NUMB. *Clin. Cancer Res.* **23**, 3871–3883 (2017).