



OPEN

# HMGA1-pseudogene7 transgenic mice develop B cell lymphomas

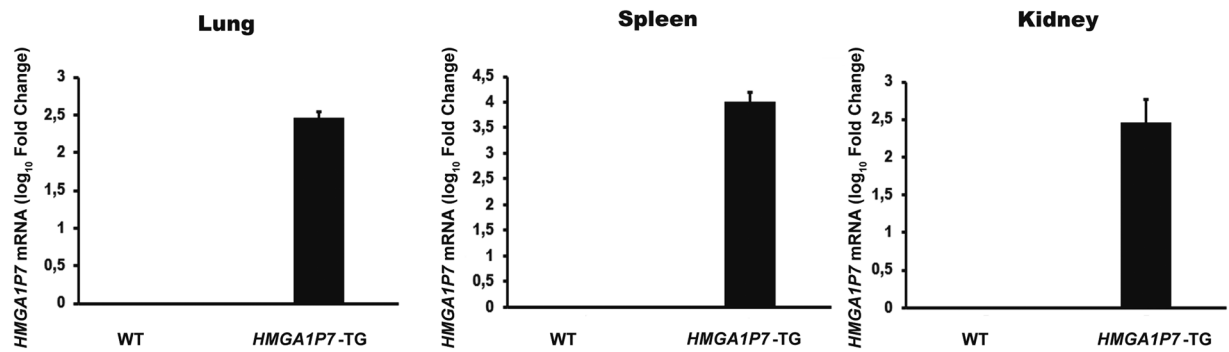
Marco De Martino<sup>1</sup>, Davide De Biase<sup>2</sup>, Floriana Forzati<sup>1</sup>, Sara Carmela Credendino<sup>1</sup>, Giuseppe Palma<sup>3</sup>, Antonio Barbieri<sup>3</sup>, Claudio Arra<sup>3</sup>, Orlando Paciello<sup>2</sup>, Eugenio Gaudio<sup>4</sup>, Maurilio Ponzoni<sup>5</sup>, Gabriella De Vita<sup>1</sup>, Paolo Chieffi<sup>6</sup>, Francesco Bertoni<sup>4</sup>, Alfredo Fusco<sup>1</sup>✉ & Francesco Esposito<sup>1</sup>✉

We have recently identified and characterized two pseudogenes (*HMGA1P6* and *HMGA1P7*) of the *HMGA1* gene, which has a critical role in malignant cell transformation and cancer progression. *HMGA1P6* and *HMGA1P7* act as microRNA decoy for *HMGA1* and other cancer-related genes upregulating their protein levels. We have previously shown that they are upregulated in several human carcinomas, and their expression positively correlates with a poor prognosis and an advanced cancer stage. To evaluate *in vivo* oncogenic activity of *HMGA1* pseudogenes, we have generated a *HMGA1P7* transgenic mouse line overexpressing this pseudogene. By a mean age of 12 months, about 50% of the transgenic mice developed splenomegaly and accumulation of lymphoid cells in several body compartments. For these mice FACS and immunohistochemical analyses suggested the diagnosis of B-cell lymphoma that was further supported by clonality analyses and RNA expression profile of the pathological tissues of the *HMGA1P7* transgenic tissues. Therefore, these results clearly demonstrate the oncogenic activity of *HMGA1* pseudogenes *in vivo*.

Many evidences indicate that long non-coding RNAs (lncRNAs) are key modulators of different biological phenomena. Given this scenario, it is predictable that deregulated expression and aberrant role of lncRNAs are involved in the development of several diseases including cancer<sup>1</sup>. Among lncRNAs, pseudogenes, a subgroup of genes that arises from protein-coding genes that have lost the capacity to produce proteins, have been considered for long time as non-functional genomic junk<sup>1</sup>. However, recent studies have unveiled important functions of pseudogenes in the regulation of the expression of the parental genes. Indeed, the majority of the identified pseudogenes has high sequence homology with their protein-coding parental counterparts, enabling them to take part in post-transcriptional control of their parental genes. The regulation of parental gene relies on several mechanisms: (i) the generation of endogenous short interfering RNAs (siRNAs)<sup>2,3</sup>; (ii) the engagement of regulatory proteins on the parental gene by pseudogene RNAs to control gene expression and chromatin remodelling<sup>4,5</sup>; (iii) the ability of the pseudogenes to compete with the parental genes for RNA-binding proteins and the translation machinery<sup>6-8</sup>; (iiii) the ability of pseudogenes to compete with their parental genes for a common pool of shared microRNAs (miRNAs)<sup>9</sup> through the high sequence homology of the 3' Untranslated region (UTR), thus regulating each other expression as competitive endogenous RNAs (ceRNAs)<sup>10</sup>.

The HMGA protein family includes the HMGA1a, HMGA1b and HMGA2 members<sup>11</sup>. The first two are coded for by the same gene through an alternative splicing. They have no transcriptional activity *per se*, but, modifying the chromatin architecture, they are able to positively or negatively regulate the expression of several genes, particularly those involved in cancer progression<sup>11,12</sup>. Consistently, these proteins are expressed at very low levels in normal adult tissues, but are abundant in almost all the human malignant neoplasms<sup>11</sup>, and their expression significantly correlates with the capability of cancer cells to metastasize and a patient poor prognosis<sup>13-15</sup>. Moreover, *in vitro* and *in vivo* models support a causal role of the HMGA proteins in cell transformation and cancer development<sup>11,16,17</sup>.

<sup>1</sup>Istituto di Endocrinologia ed Oncologia Sperimentale - CNR c/o Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Università degli Studi di Napoli "Federico II", Naples, Italy. <sup>2</sup>Department of Veterinary Medicine and Animal Production, University of Naples Federico II, Naples, Italy. <sup>3</sup>S.S.D. Sperimentazione Animale, Istituto Nazionale Tumori, IRCCS, Fondazione Pascale, Naples, Italy. <sup>4</sup>Institute of Oncology Research, Faculty of Biomedical Sciences, USI, Bellinzona, Switzerland. <sup>5</sup>Vita-Salute San Raffaele University & Pathology Unit, IRCCS San Raffaele Scientific Institute, Milan, Italy. <sup>6</sup>Department of Psychology, University of Campania "L. Vanvitelli", Caserta, Italy. ✉e-mail: [alfusco@unina.it](mailto:alfusco@unina.it); [francesco.esposito2@unina.it](mailto:francesco.esposito2@unina.it)



**Figure 1.** Analysis of *HMGA1P7* expression in transgenic mice qRT-PCR analysis of total RNA from lungs, spleens and kidneys of WT (n = 3) and *HMGA1P7* (n = 3) transgenic mice. The error bars represent mean  $\pm$  SD.

We have recently identified two human *HMGA1* processed pseudogenes (*HMGA1P6* and *HMGA1P7*) that are not present in mouse genome. *HMGA1P6* and *HMGA1P7* can compete with *HMGA1* for miRNA binding, leading to the upregulation of *HMGA1* cellular levels, thereby enhancing the expression of cell malignant features<sup>18–23</sup>. The overexpression of these *HMGA1* pseudogenes (*HMGA1Ps*) also increases the levels of *HMGA2* and other cancer-related genes, such as *EZH2* and *VEGF*, by inhibiting the suppression of their synthesis<sup>18</sup>. Noteworthy, *HMGA1Ps* were found overexpressed in several human cancer types supporting their involvement in carcinogenesis<sup>18,20–23</sup>. To investigate the role of *HMGA1* pseudogenes overexpression *in vivo*, we generated transgenic mouse model overexpressing *HMGA1P7* (*HMGA1P7-TG*)<sup>18,22–24</sup>. Mouse Embryonic Fibroblasts (MEFs) derived from *HMGA1* pseudogene transgenic mice showed a higher growth rate and a later onset of senescence than the wild-type (WT) counterpart<sup>18</sup>.

Here, we report that *HMGA1* pseudogene transgenic mice develop haematological neoplasia characterized by monoclonal B-cell populations, most of them diagnosed as large B-cell lymphoma. These results validate the oncogenic role of the *HMGA1* pseudogenes<sup>18</sup>.

## Results

***HMGA1P7* transgenic mice develop lymphoproliferative lesions.** Transgenic mice carrying the *HMGA1P7* gene were generated by the injection of the transgene into C57BL/6N derived zygotes and, then transferred into pseudo-pregnant as previously described<sup>18</sup>. The expression of the *HMGA1P7* was assessed in lungs, spleens and kidneys explanted from transgenic mice (Fig. 1).

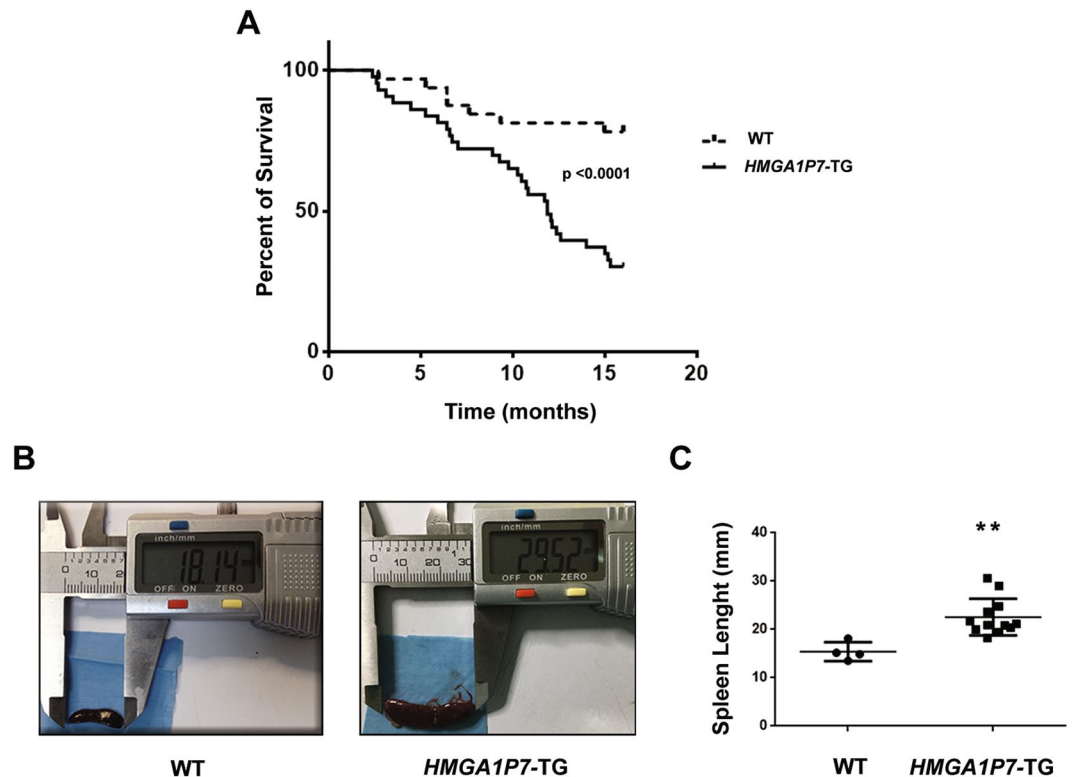
Interestingly, *HMGA1P7* mice showed significant increased mortality with respect to the WT mice (Gehan Breslow Wilcoxon test,  $p < 0.0001$ ) with a mean age of death of about 52 weeks (Fig. 2A). About 50% of 12 months-old *HMGA1P7* transgenic mice displayed splenomegaly at necropsy, whereas WT mice showed no relevant alteration in splenic size or weight (Fig. 2B,C). Histological sections of the *HMGA1P7-TG* spleens showed a clear distinction between the red and the white pulp. In the red pulp multiple foci of extramedullary haematopoiesis, as well as hemosiderin-laden macrophages were frequently observed (Fig. 3A,III). White pulp showed a moderate expansion with some confluent areas and partial loss of normal structures and germinal centers. In some mice, higher magnification showed a diffuse, monotonous lymphoid population composed of medium-to-large rounded cells with scant cytoplasm, round to oval nuclei and single or multiple, prominent nucleoli often adherent to the nuclear membrane (Fig. 3A,IV). Mitotic activity was medium to high ( $< 10 \times 10$  HPF). Intriguingly, histopathological analyses revealed monotonous lymphoid cells infiltrating liver ( $\approx 25\%$ ), kidneys ( $\approx 25\%$ ), lung ( $\approx 30\%$ ), and pancreas ( $\approx 20\%$ ) (Fig. 3B). Immunohistochemical analysis of lymphoid component displayed a predominant CD45/B220-positive population intermingled with few, scattered CD3-positive cells (Fig. 3C). Based on morphology and immunophenotype, a diagnosis of large B-cell lymphoma with immunoblastic features was made (human counterpart: DLBCL, immunoblastic variant)<sup>25</sup>.

Furthermore, FACS analysis of lymphocytes isolated from WT or pathological spleens using the CD3, CD19 and NK anti-mouse antibodies confirmed the immunohistochemical data. CD19 population resulted almost doubled, while CD3 population was decreased in *HMGA1P7-TG* mouse spleens in comparison with WT animals (Fig. 4A).

To investigate the clonal status of the accumulation of the CD19 positive population in *HMGA1P7-TG* mice, genomic DNAs from TG and WT spleens were analysed. As shown in Fig. 4B, only one dominant PCR product was generated by the amplification of the DNA extracted from the transgenic spleens, whereas DNA derived from a WT spleen yielded three prominent PCR products of 1.0, 0.7 and 0.12 kb, corresponding to DJH2, DJH3 and DJH4 Immunoglobulin (Ig) gene rearrangements, respectively<sup>26</sup>.

Taken together, these results indicate that *HMGA1P7-TG* mice lymphoid expansion was monoclonal, therefore further supporting the diagnosis of B-cell lymphoma.

**Identification of the genes modulated by *HMGA1P7* expression in pathological spleens.** Since *HMGA1* did not result upregulated by *HMGA1P7* overexpression in the analyzed pathological spleens and other mouse tissues (Fig. 5), we compared the transcriptome of spleens derived from *HMGA1P7* transgenic mice (n = 2) versus that of WT spleens (n = 2) by RNA-Seq analyses, in order to better understand the mechanisms



**Figure 2.** *HMGA1P7* expression *in vivo* induces splenomegaly and premature death (A) Survival curve of WT (n = 30) and *HMGA1P7* (n = 40) transgenic mice. The survival rate of WT mice was significantly higher than *HMGA1P7* transgenic ones (Gehan Breslow Wilcoxon test,  $p < 0.0001$ ). (B) Representative images of spleens from WT and *HMGA1P7* transgenic mice. (C) Spleens from *HMGA1P7* (n = 12) transgenic mice were larger than spleens from WT (n = 4) (Mann-Whitney Test,  $**p < 0.0011$ ). The error bars represent mean  $\pm$  SD.

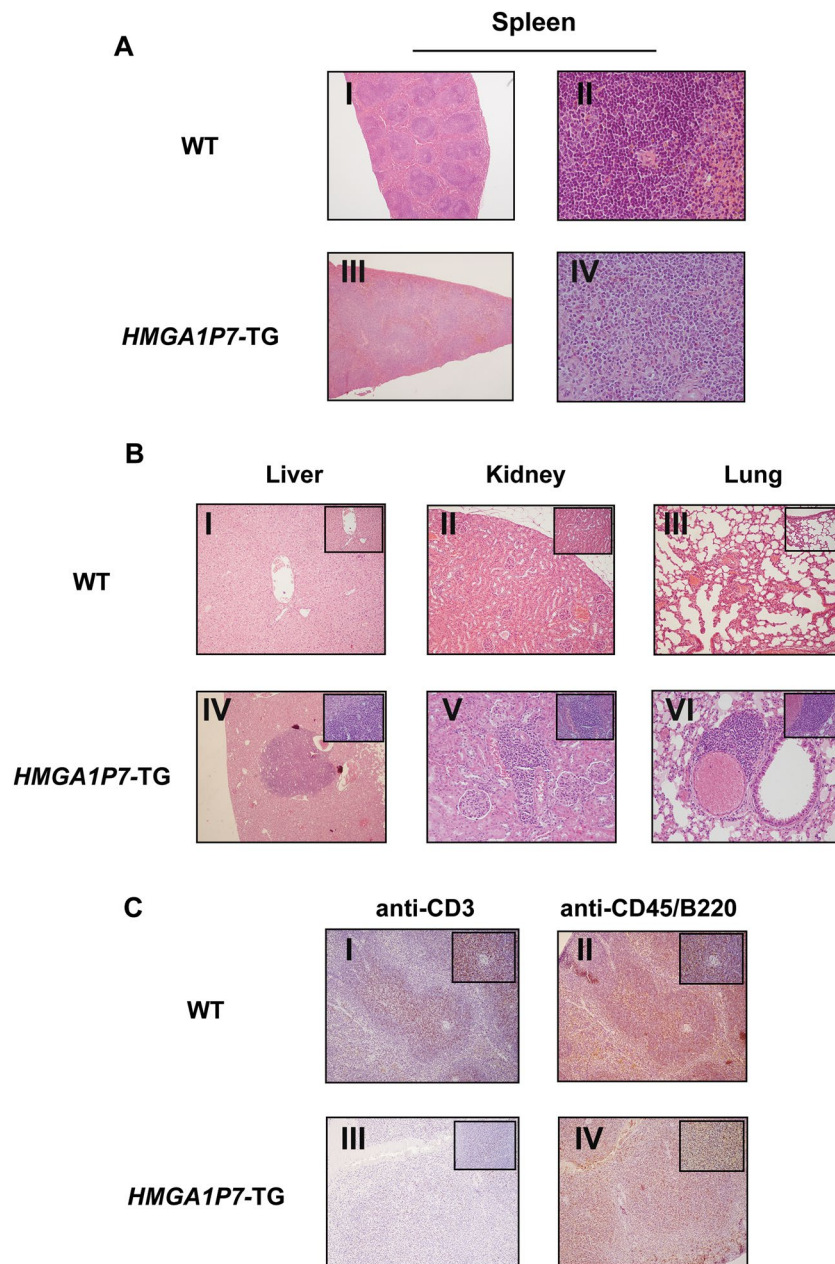
leading to lymphoid cell proliferation in transgenic mice. The upregulated transcripts included genes involved in inflammation (*Ccl24*, *Il1a*, *Rgs16*, *Ccl5*)<sup>27–30</sup>, in the NF $\kappa$ B pathway and in IL6/JAK/STAT3 and MTOR signalling, in oxidative phosphorylation (*Uqcrc1*, *Ndufa1*, *Cox5a*, *Atp5d*)<sup>31–34</sup> and targets of MYC, E2F, STAT3, API1, ATF3. In addition, the spleens of transgenic mice presented a gene expression signature compatible with an induction of senescence (*Il13ra2*, *Il1a*, *Mmp3*, *Il1b*)<sup>35–38</sup> and immune escape (*Pvrl2*, *Il10*, *Cd160*, *Ido1*)<sup>39–42</sup>. Enrichment of genes downregulated by B cell receptor (BCR) inhibitors in diffuse large B cell lymphomas (DLBCL) was also unveiled by this analysis. Moreover, among the genes that showed a decreased expression in the transgenic spleens we found (i) transcripts down-regulated in post-germinal center (GC) BCL6 dependent B cell lymphomas when compared to MYC driven pre-GC lymphomas (*Cnot6l*, *Sh3kbp1*)<sup>43,44</sup>, (ii) genes repressed by BLIMP1 (*Stat6*, *Zfp361l*)<sup>45,46</sup>; (iii) genes present in the GC B-cell type (GCB) DLBCL signature (*Dtx1*, *Cux1*, *Sh3pxd2a*, *Klhl6*)<sup>47–49</sup> (Fig. 6).

Then, we validated the results obtained by RNA-Seq analyses, testing the expression of a panel of deregulated mRNAs in spleens from *HMGA1P7* by qRT-PCR (Fig. 7). Among the upregulated genes we chose CCAAT/enhancer-binding protein delta (*Cebpd*), chemokine (C-C motif) ligand 24 (*Ccl24*), Bcl-2-like 1 (*Bcl2l1*), *Fos*, Interleukin 1 Alpha (*Il1a*), BTB and CNC homolog 2 (*Bach2*), one of the downregulated genes. Next, the increased expression levels of *Cebpd*, *Bcl2l1* and *Fos* were also confirmed by western blot analyses (Fig. 7). Finally, to demonstrate that *HMGA1P7* acts through a ceRNA mechanism on the genes deregulated in pathological spleens (Fig. 8A), we inserted downstream of the luciferase open reading frame the 3'-UTRs of these genes. These reporter vectors were transfected into NIH3T3 cells overexpressing or not *HMGA1P7*. As expected, the luciferase activity was markedly increased in the cells that overexpressed *HMGA1P7* (Fig. 8B), confirming the ceRNA action induced by *HMGA1P7* on these new targets.

Therefore, on the basis of the FACS and immunohistochemical data combined with the RNA-Seq analyses we can assess that the lymphoproliferation in the *HMGA1P7* transgenic mice shares transcriptome features with DLBCL of the non-GCB type<sup>43–49</sup>.

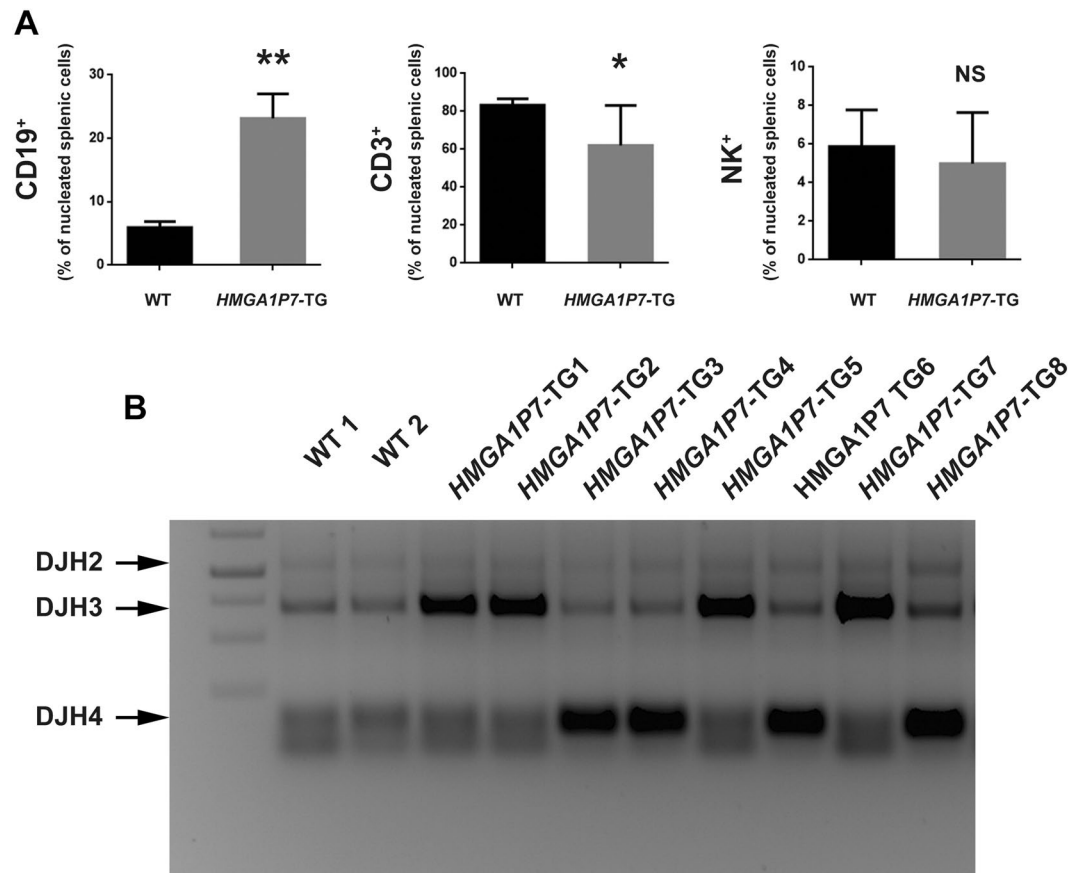
## Discussion

We have previously reported that the overexpression of *HMGA1Ps* accelerates cell proliferation, by enhancing the G1-S transition, increases cell migration ability, likely raising the levels of HMGA1 and other oncogenic proteins such as HMGA2 and EZH2<sup>18</sup>. Moreover, the MEFs obtained from *HMGA1Ps* transgenic mice showed a reduced proliferation time and senescence in comparison with the WT MEFs<sup>18</sup>.



**Figure 3.** *HMGA1P7* transgenic mice show a lymphoid malignancy (A) (I and II) Spleen from WT mouse showing normal morphology. (III) Representative image of immunoblastic lymphoma from a *HMGA1P7*-TG mouse. Expansion and merging of adjacent white pulp areas with loss of normal structures and germinal centers. (IV) A higher magnification shows a monotonous lymphoid population composed of fairly large rounded cells with scant cytoplasm and round to oval nuclei with finely dispersed chromatin and inconspicuous nucleoli. Haematoxylin and eosin (Original Magnification 4X for I and III panels, 40X for II and IV panels). (B) Representative images of liver (I), kidney (II) and lung (III) from a WT mouse with no pathological alteration. Monotonous lymphoid cells infiltrating the liver (IV), the kidney (V), the lung (VI) of *HMGA1P7*-TG mice. At higher magnification (inset), lymphoid population was composed of fairly large rounded cells with scant cytoplasm and round to oval nuclei with finely dispersed chromatin and inconspicuous nucleoli. Haematoxylin and eosin (Original Magnification 10X, inset 40X). (C) (I and II) immunohistochemistry of a WT mouse spleen showing normal distribution of CD3 immunolabelled T-cells and CD45/B220 immunolabelled B-cells. (III, IV) Immunohistochemical analysis of *HMGA1P7*-TG mouse spleen revealed a predominant CD45/B220 neoplastic lymphoid population with fewer scattered CD3-immunolabelled cells (Original Magnification 10X, inset 40X).

Therefore, the aim of this study was to better characterize the transgenic mice overexpressing the *HMGA1P7* to possibly validate its oncogenic activity *in vivo*. The analysis of *HMGA1P7* transgenic mice at 12 months of age, shows that about 50% of these mice developed a pathology characterized by splenomegaly and invasion



**Figure 4.** Mice overexpressing *HMGA1P7* develop monoclonal expansion of the CD19 positive population. **(A)** FACS analysis of splenic cells isolated from WT (n = 8) and *HMGA1P7* (n = 14) transgenic mice using CD19, CD3, and NK1.1 anti-mouse antibodies. The results are reported as the mean of values. The error bars represent mean  $\pm$  SD; \* $P < 0.05$  \*\* $P < 0.01$  (t test). **(B)** Genomic DNA isolated from the spleens of two WT mice and eight *HMGA1P7*-TG derived-lymphomas were amplified with DSF and JH4 primers. Three expected DJH bands with the spleen DNA were indicated. Within all eight tumor samples, DJH3 or DJH4 bands were predominantly detected.

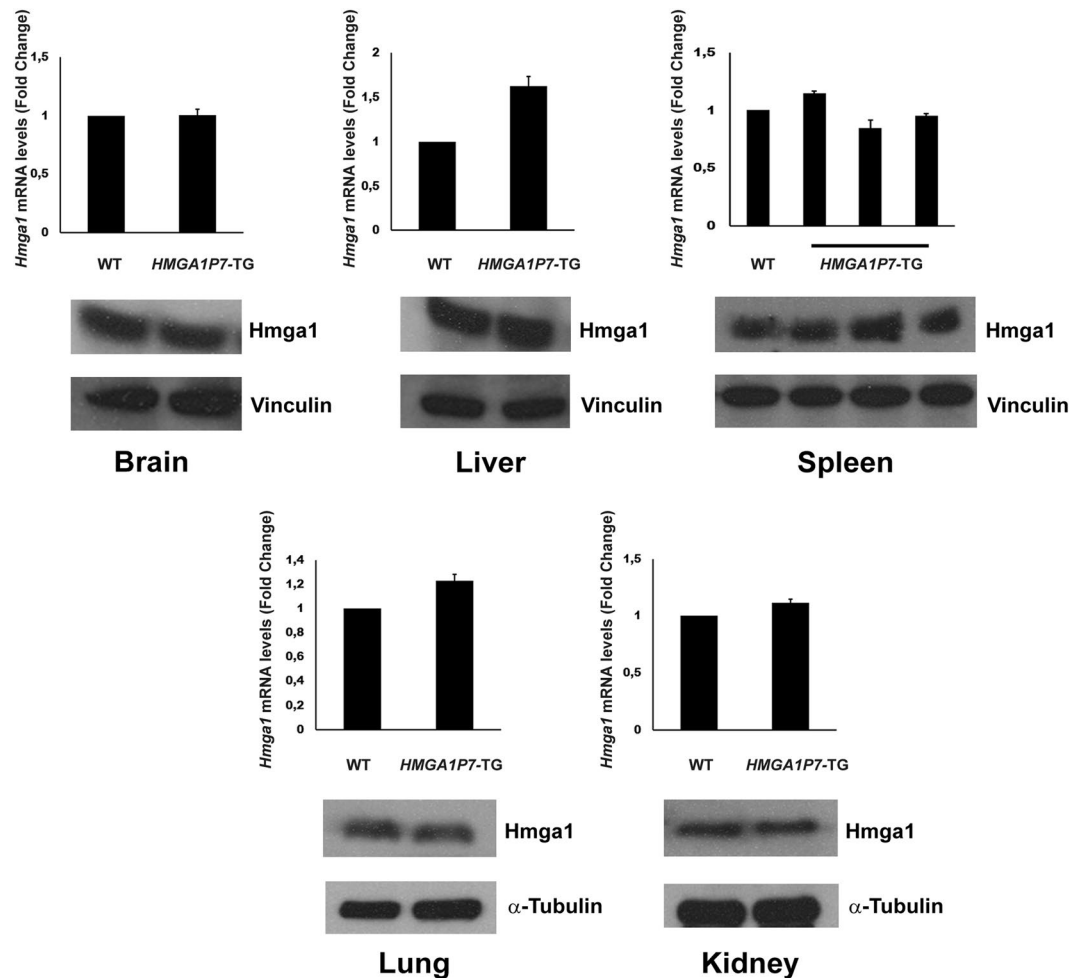
of lymphoid cells in different anatomical districts. The pathological spleens showed a diffuse and monotonous lymphoid population effacing the splenic parenchyma with the loss of typical structures and germinal centres. Neoplastic lymphoid cells were medium to large, rounded, with scant cytoplasm and round to ovular nuclei with single or multiple prominent nucleoli.

By immunohistochemistry and FACS analyses, we found that the neoplastic cells were respectively positive for CD45/B220 and CD19 proposing a B cell phenotype of the lymphoid cells. Clonality assay on pathological spleens evidenced the clonal expansion of CD19-positive lymphoid population supporting a diagnosis of B cell lymphomas for these lesions. Interestingly, RNA-Seq analyses performed on spleens derived from WT and *HMGA1P7* mice revealed a deregulation of several genes, likely due to *HMGA1P7*-ceRNA activity. The deregulated genes were involved in inflammation pathways such as NF $\kappa$ B pathway, the IL6/JAK/STAT3 and MTOR signalling, the oxidative phosphorylation, and targets of MYC, E2F, STAT3, AP1, ATF3. Moreover, the spleens from *HMGA1P7* mice had a gene expression signature compatible with an induction of senescence and immune escape (*Il13ra2*, *Il1a*, *Mmp3*, *Il1b*, *Pvrl2*, *Il10*, *Cd160*, *Ido1*)<sup>35–42</sup>.

Noteworthy, the genes suppressed by BCR inhibitors in DLBCL were found significantly enriched in the pathological tissues of *HMGA1P7* mice. In particular, the downregulated genes were enriched of transcripts decreased in post-GC BCL6 dependent B cell lymphomas and present in the GCB DLBCL signature. Therefore, the transcriptome study of the lymphoproliferative lesions in the *HMGA1P7* transgenic mice unveils a pathology compatible with DLBCL of the non-GCB type.

Consistently with the ability of the *HMGA1Ps* to regulate gene transcription by a ceRNA mechanism, bioinformatic analyses demonstrate that several upregulated genes emerged from RNA-Seq data shared the same microRNA Responsive Elements with *HMGA1P7* (i.e. *Cebpd*, *Ccl24*, *Bcl2l1*, *Fos*, *Il1a*).

Surprisingly, *HMGA1* did not result upregulated by *HMGA1P7* overexpression in the analysed pathological spleens, suggesting that pseudogene-induced lymphomas were based on other molecular targets already described<sup>27–49</sup>. Moreover, we did not find any change in *HMGA1* expression levels during spleen development of *HMGA1P7*-transgenic mice (data not shown). However, we cannot exclude the possibility of increased *HMGA1* protein levels in a limited cell compartment in the initial steps of lymphomagenesis.



**Figure 5.** HMGA1 is not upregulated in *HMGA1P7-TG* explanted-organs. qRT-PCR and Western Blot analyses of HMGA1 expression levels in WT (n = 3) and *HMGA1P7* (n = 3) transgenic brain, liver, spleen, lung and kidney organs.

Altogether the data presented here show that deregulated expression of *HMGA1P7* pseudogene has oncogenic role also *in vivo*, thus representing a new class of genes involved in cancer pathology as their upregulation occurs frequently in multiple human cancers<sup>50</sup>. An oncogenic role for pseudogenes has been already reported. Indeed, mice engineered to overexpress the full-length murine *B-Raf* pseudogene *Braf-rs1* develop an aggressive malignancy resembling human diffuse large B cell lymphoma by ceRNA mechanism that elevates *BRAF* expression<sup>50</sup>.

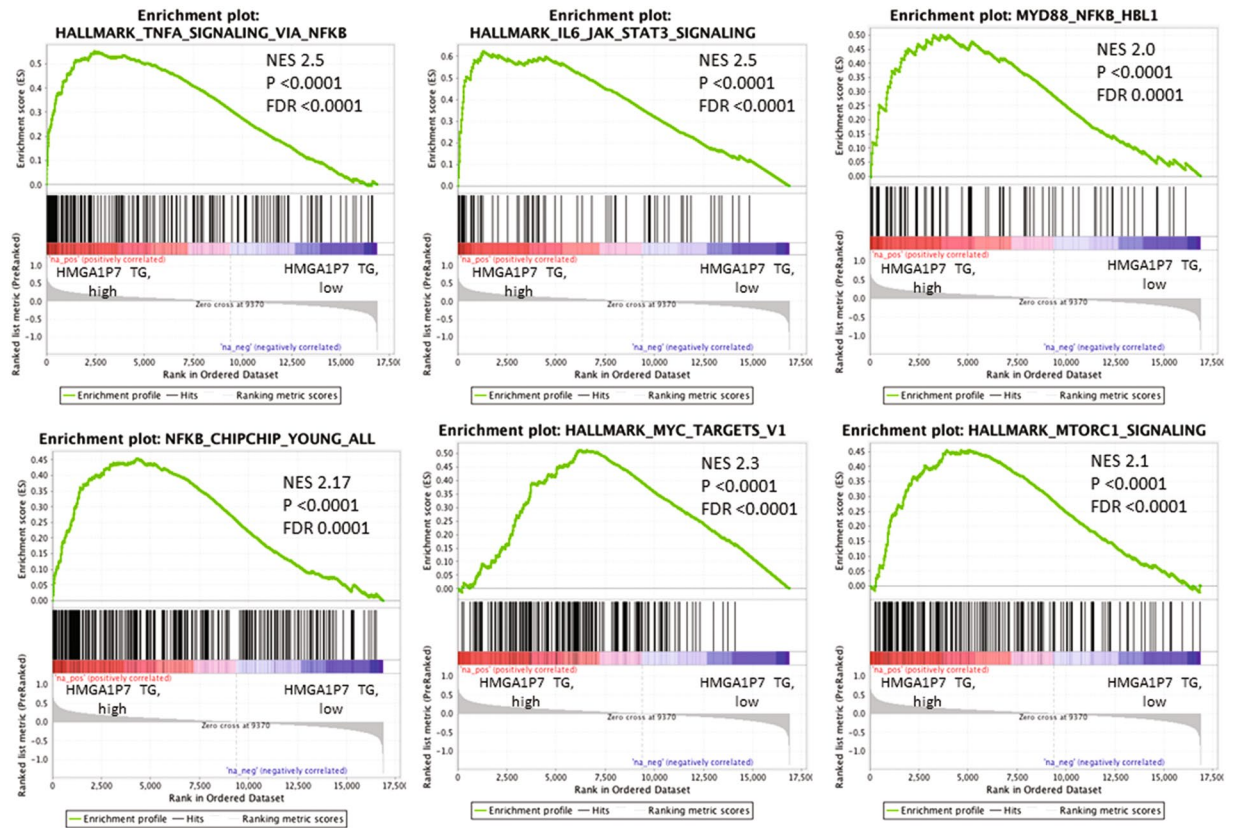
Noteworthy, preliminary studies on a mouse strain overexpressing *HMGA1P6* pseudogene show that several mice develop a lymphoid pathology characterized by splenomegaly that resembles that found in *HMGA1P7-TG* mice.

Therefore, our mouse model confirms the oncogenic potential of pseudogenes and provides compelling support for a causal link between altered pseudogene expression and cancer, mediated by ceRNA mechanism. Studies are in progress to evaluate the expression levels of *HMGA1* pseudogenes in human lymphomas. Preliminary results indicate *HMGA1P1* overexpression that could contribute to lymphomagenesis by a similar ceRNA mechanism.

## Materials and Methods

**Transgenic mice.** *HMGA1P7* overexpressing mice have been previously generated and described<sup>18</sup>. Animals were housed in IRCCS “G. Pascale” animal facility as previously reported<sup>51</sup> (project “Ruolo degli pseudogeni di HMGA1 nel cancro” Cod. 893/2013 approved by Italian Health Ministry on 13/05/2013). The experimental protocols were in compliance with the European Communities Council Directive (63/2010/EEC).

**Cell culture and transfections.** NIH3T3 cells were maintained in DMEM supplemented with 10% calf serum (ThermoFisher, Waltham, MA, USA), glutamine and antibiotics. MycoAlert (Basel, Switzerland) was regularly used to test that cells were not infected with mycoplasma. Cell transfection protocol was reported elsewhere<sup>19</sup>.



**Figure 6.** Transcriptome of *HMGAI1P7*-TG derived-lymphomas. Representative GSEA plots illustrating the transcriptional expression signature enrichment in genes upregulated in transgenic spleens compared to normal spleens. Green line, enrichment score; bars in the middle portion of the plots show where the members of the gene set appear in the ranked list of genes. Positive or negative ranking metric indicate respectively correlation or inverse correlation with the profile. NES, normalised enrichment score. P, p-value. FDR, false discovery rate.

## Histology and immunohistochemistry

Light microscopy was performed as previously described<sup>52</sup>. Definition and classification of lymphoid disease were based on criteria reported elsewhere<sup>25</sup>.

For immunohistochemistry (IHC), 4- $\mu$ m-thick sections were processed as previously described<sup>24,53</sup>. Primary antibodies included rabbit monoclonal to CD3 (ab16669, Abcam, Cambridge, UK) diluted 1:200 and rabbit polyclonal to CD45/B220 as a marker for B cells<sup>54</sup> (ab10558 Abcam, Cambridge, UK) diluted 1:1000.

## RNA extraction and qRT-PCR

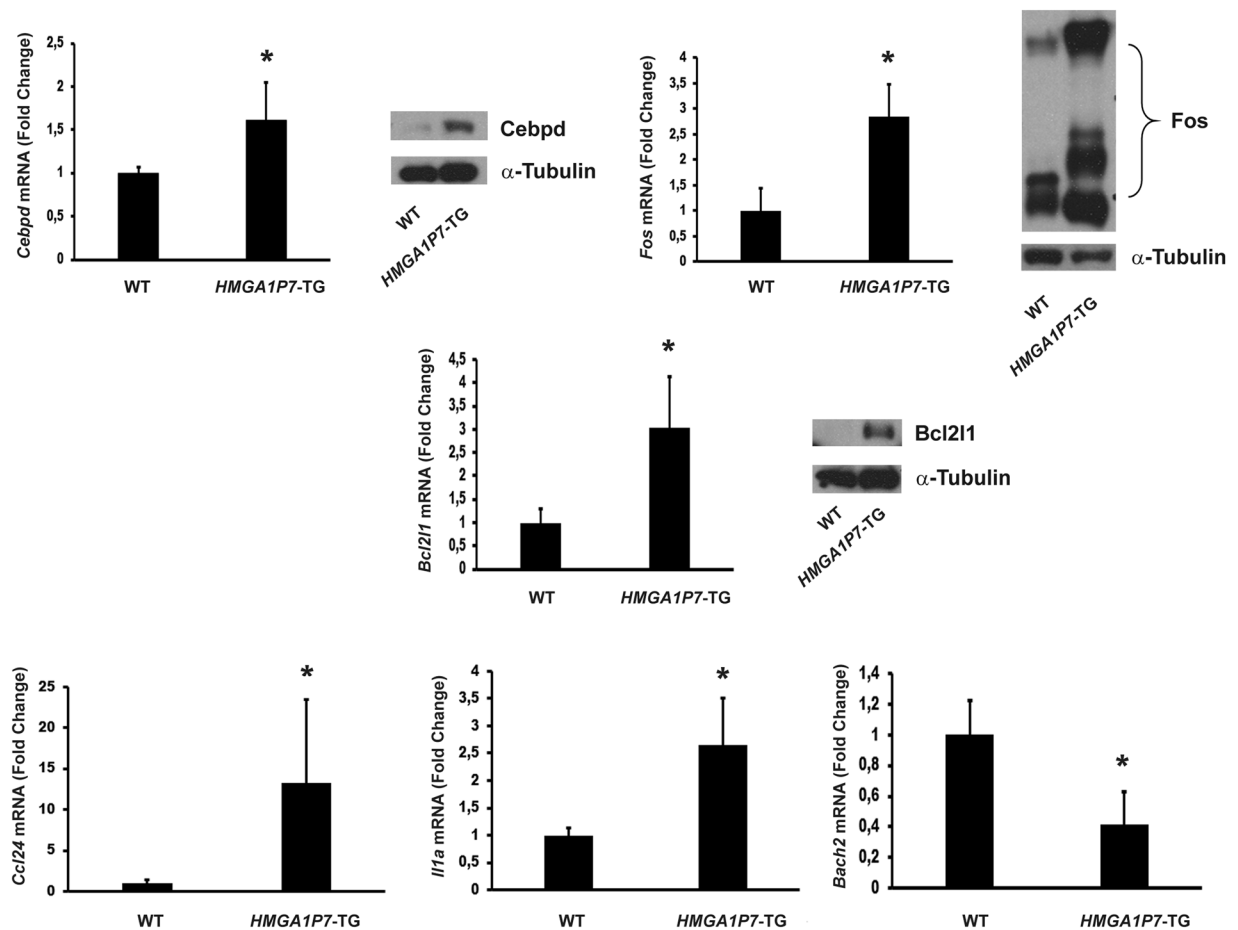
RNA extraction, RNA reverse transcription and Real-time PCR was performed as described elsewhere<sup>55</sup>. The following primers were used:

*HMGAI1P7*-Fw 5'-gctcctctcggtcctc-3'/*HMGAI1P7*-Rev 5'-gcttggcctctttatgg-3'  
*Hmga1*-Fw 5'-ggcagaccaagaactgg-3'/*Hmga1*-Rev 5'-ggcactgagtggtgat-3'  
*Cebpd*-Fw  $\lambda$ 5'-cttttagtggtgccgaag-3'/*Cebpd*-Rev 5'-ggcaacgaggaatcaagttt-3'  
*Ccl24*-Fw 5'-gcagcatctgtccaagg-3'/*Ccl24*-Rev 5'-gcagcttggggtcagtaca-3'  
*G6pd*-Fw 5'-cagcggcaactaactcaga-3'/*G6pd*-Rev 5'-ttcctcaggtaccacac-3'  
*Bcl2l1*-Fw 5'-tgaccacctagagccttga-3'/*Bcl2l1*-Rev 5'-gctcattgttcccgtaga-3'  
*Fos*-Fw 5'-gggacgccttctactacc-3'/*Fos*-Rev 5'-agatctgcaaaagtccg-3'  
*Il1a*-Fw 5'-ttggttaatgacctgcaaca-3'/*Il1a*-Rev 5'-gagcgtcacgaacagttg-3'  
*Bach2*-Fw 5'-gcagacagtgagctgtcc-3'/*Bach2*-Rev 5'-gttcttgggaaggtctgtga-3'

## Flow cytometric analysis (FACS)

For FACS analyses, spleens were collected from WT and transgenic mice, hard-pressed through a stainless-steel mesh, resuspended in PBS and then in Red Blood Lysing Buffer (Sigma-Aldrich, Saint Louis, MI, US) for 3 min. After two washes in PBS, lymphocytes ( $5 \times 10^5$ ) were set in 96-well round-bottom dishes.

Monoclonal antibodies utilized: NK-FITC (clone # PK146), CD3 APC-H7 (clone # 560176), CD19 PE/cy7 (clone # HIB19). All antibodies were from BioLegend (San Diego, CA, US).



**Figure 7.** Validation of RNA-Seq analyses on *HMGA1P7* spleens. qRT-PCR and Western Blot analyses of selected deregulated genes performed on WT (n = 4) and *HMGA1P7* (n = 4) transgenic spleens. The results are reported as the mean of values. The error bars represent mean  $\pm$  SD; \* $P < 0.05$  (Mann-Whitney Test).

### Analysis of the clonality of lymphomas

Genomic DNA was extracted from fresh spleens through Phenol/Chloroform/Isoamyl Alcohol Extraction (ThermoFisher, Waltham, MA, USA). The obtained DNAs were utilized as PCR templates with DSF and JH4 primers that recognize mouse DNA DJ rearrangement<sup>26</sup>.

DSF primer: 5'-AGGGATCCTTGTAAGGGATCTACTACTGTG-3';  
JH4 primer: 5'-AAAGACCTCCAGAGGCCATTCTTACC-3'.

### RNA-Seq analyses

Genomix4life S.R.L. (Baronissi, Salerno, Italy) performed the next generation sequencing analysis, including samples quality control and Bioinformatics analysis. Following the producer's guidelines, indexed libraries were obtained from 500 ng/ea RNA through TruSeq Stranded total RNA Sample Prep Kit (Illumina, San Diego, CA, USA). The libraries quantification was performed by the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) and Qubit fluorometer (ThermoFisher, Waltham, MA, USA), then combined in order that every index-tagged sample was in equimolar amounts, with 2nM pooled samples final concentration. Illumina HiSeq. 2500 System (Illumina, San Diego, CA, USA) sequenced the pooled samples with a format of  $2 \times 100$  paired-end at 8 pmol final concentration.

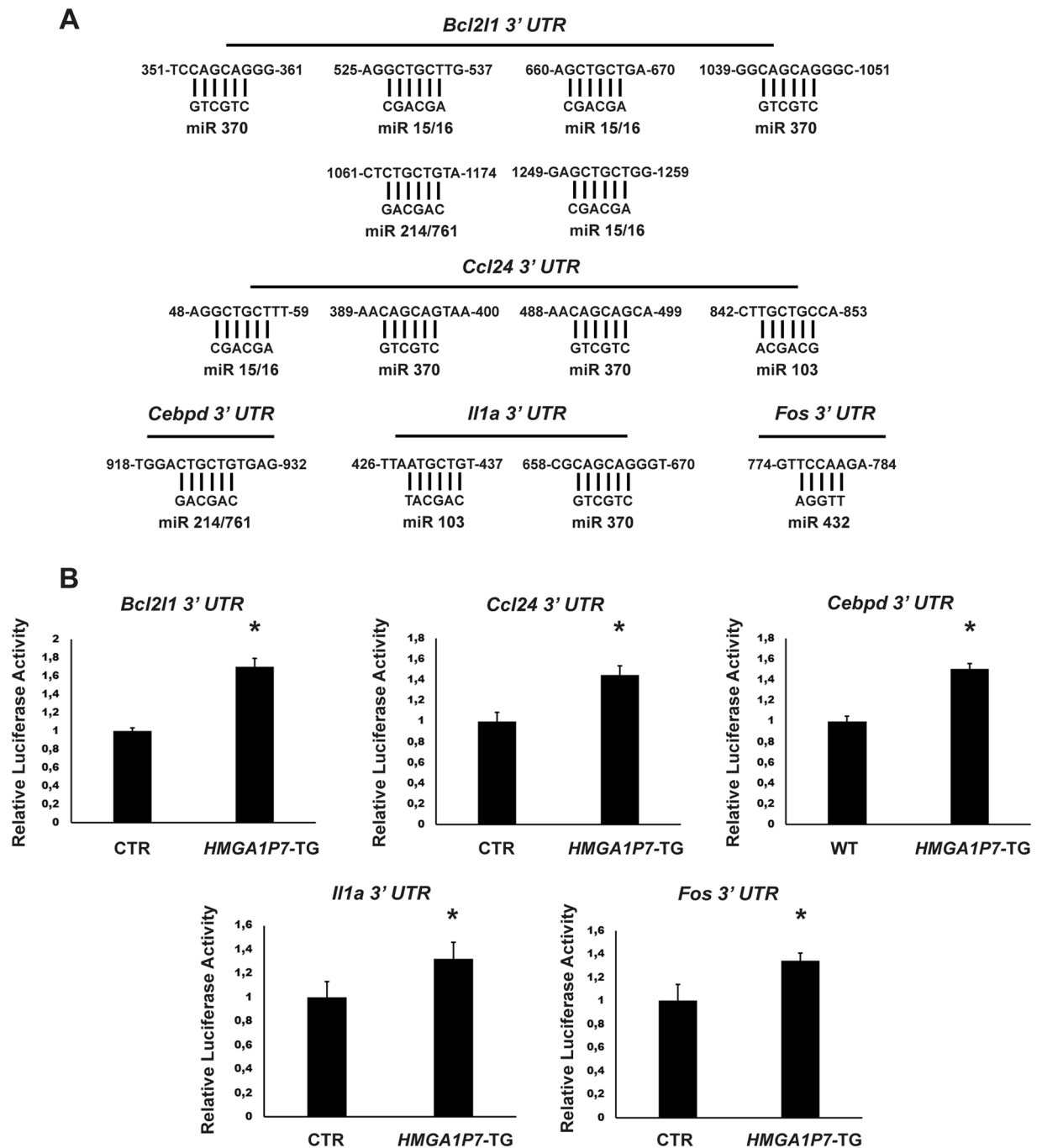
FastQC tool<sup>56</sup> was utilized for the quality control analysis of the generated raw sequence files (.fastq files). Cutadapt was used in order to eliminate the adapter sequences. Paired-end reads were mapped using STAR (version 2.5.2b)<sup>57</sup> on reference genome assembly mm10 acquired from Ensembl<sup>58,59</sup>. The quantification of transcripts expressed for each replicate of the sequenced samples was performed using HTSeq-Count algorithm<sup>60</sup>. The differential expression analysis was performed through DESeq. 2<sup>61</sup>.

Gene Set Enrichment Analysis (GSEA) was used for functional annotation on pre-ranked lists using the MSigDB 5.2<sup>62</sup>, the SignatureDB collection<sup>63</sup> and genesets obtained from different publications<sup>64,65</sup>, applying false discovery rate (FDR) values  $< 0.05$  as threshold.

### Luciferase assay

Dual-luciferase reporter assays were performed as previously described<sup>18</sup>.





**Figure 8.** Deregulated genes from RNA-Seq are modulated by *HMGA1P7* through a ceRNA mechanism. (A) 3'UTRs of *Bcl2l1*, *Ccl24*, *Cebpd*, *Il1a* and *Fos* share several microRNA Responsive Elements with *HMGA1P7*. (B) The *Bcl2l1*, *Ccl24*, *Cebpd*, *Il1a* and *Fos* 3'UTRs were cloned into the pMIR control vector and then transfected in NIH3T3 cells stably overexpressing the control vector or *HMGA1P7*. The results are reported as the mean of values. Error bars represent mean  $\pm$  SD; n = 4. \*P < 0.05 (Mann-Whitney Test).

### Plasmids

For the mCebpd (NM\_007679), mCcl24 (NM\_019577), mBcl2l1 (NM\_009743), mFos (NM\_010234) and mIl1a (NM\_010554) 3'UTR luciferase reporter constructs, the 3'UTR sequences were amplified by using the following primers:

mCebpd 3'utr-Fw 5'-gcagagctcagaattctgccttctactaagatactggttg-3'  
mCebpd 3'utr-Rv 5'-gcatcgcttgaattcttagtcttctgggagctgcc-3'  
mCcl24 3'utr-Fw 5'-gcagagctcagaattcccgcctctctctgtccc-3'  
mCcl24 3'utr-Rv 5'-gcatcgcttgaattcacatcctggcagcaagagg-3'  
mBcl2l1 3'utr-Fw 5'-gcagagctcagaattcgagcctctcgggaatgcttttc-3'

mBcl2l1 3'utr-Rv 5'-gcatcgcttgaattccgcacagcaagccagcag-3'  
 mFos 3'utr-Fw 5'-gcagagctcagaattcgaatgtctgacattacagttttc-3'  
 mFos 3'utr-Rv 5'-gcatcgcttgaattctcaacttaaatgcttttattgac-3'  
 mIl1a 3'utr-Fw 5'-gcagagctcagaattcctcaaatgacagttgagtagga-3'  
 mIl1a 3'utr-Rv 5'-gcatcgcttgaattcaggagactacatctaactgaccac-3'

The amplified fragments were cloned into pMirTarget vector (OriGene, Rockville, MD, USA) using In-Fusion HD Cloning kit (Takara Bio, Mountain View, CA, USA). *HMGA1P7* overexpressing vector was previously described<sup>18</sup>.

## Western blot

Western blot analyses were performed as previously described<sup>66</sup>. The primary antibodies used were: anti-Cebpδ #7077 (ProSci, Poway, CA, USA); anti-Fos sc-166940 (Santa Cruz Biotechnology, Dallas, TX, USA); anti-Bcl2l1 #2762 (Cell Signaling, Danvers, MA, USA). Antibody against HMGA1 protein was described elsewhere<sup>67</sup>.

## Statistical analysis

Two-sided unpaired Student's *t* tests and Mann-Whitney tests were utilized to analyse data (GraphPad Prism, GraphPad Software, Inc.). *P* < 0.05 values were taken into account as statistically significant. The mean values ± s.d were obtained from three or more separate experiments. GraphPad Prism, GraphPad Software, Inc. was used to obtain regression analyses and correlation coefficients.

Received: 8 September 2019; Accepted: 22 March 2020;

Published: 27 April 2020

## References

- Poliseno, L. Pseudogenes: newly discovered players in human cancer. *Sci. Signal.* **5**, re5, <https://doi.org/10.1126/scisignal.2002858> (2012).
- Tam, O. H. *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* **453**, 534–538, <https://doi.org/10.1038/nature06904> (2008).
- Watanabe, T. *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* **453**, 539–543, <https://doi.org/10.1038/nature06908> (2008).
- Hawkins, P. G. & Morris, K. V. Transcriptional regulation of Oct4 by a long non-coding RNA antisense to Oct4-pseudogene 5. *Transcription* **1**, 165–175, <https://doi.org/10.4161/trns.1.3.13332> (2010).
- Johnsson, P., Lipovich, L., Grander, D. & Morris, K. V. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim. Biophys. Acta* **1840**, 1063–1071, <https://doi.org/10.1016/j.bbagen.2013.10.035> (2014).
- Bier, A. *et al.* Connexin43 pseudogene in breast cancer cells offers a novel therapeutic target. *Mol. Cancer Ther.* **8**, 786–793, <https://doi.org/10.1158/1535-7163.MCT-08-0930> (2009).
- Chiefari, E. *et al.* Pseudogene-mediated posttranscriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes. *Nat. Commun.* **1**, 40, <https://doi.org/10.1038/ncomms1040> (2010).
- Han, Y. J., Ma, S. F., Yourek, G., Park, Y. D. & Garcia, J. G. A transcribed pseudogene of MYLK promotes cell proliferation. *FASEB J.* **25**, 2305–2312, <https://doi.org/10.1096/fj.10-177808> (2011).
- Poliseno, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* **465**, 1033–1038, <https://doi.org/10.1038/nature09144> (2010).
- Salmena, L., Poliseno, L., Tay, Y., Kats, L. & Pandolfi, P. P. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* **146**, 353–358, <https://doi.org/10.1016/j.cell.2011.07.014> (2011).
- Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat. Rev. Cancer* **7**, 899–910, <https://doi.org/10.1038/nrc2271> (2007).
- Sgarra, R. *et al.* High Mobility Group A (HMGA) proteins: Molecular instigators of breast cancer onset and progression. *Biochim. Biophys. Acta Rev. Cancer* **1869**, 216–229, <https://doi.org/10.1016/j.bbcan.2018.03.001> (2018).
- Pegoraro, S. *et al.* A novel HMGA1-CCNE2-YAP axis regulates breast cancer aggressiveness. *Oncotarget* **6**, 19087–19101, <https://doi.org/10.18632/oncotarget.4236> (2015).
- Pegoraro, S. *et al.* HMGA1 promotes metastatic processes in basal-like breast cancer regulating EMT and stemness. *Oncotarget* **4**, 1293–1308, <https://doi.org/10.18632/oncotarget.1136> (2013).
- Gorbounov, M. *et al.* High mobility group A1 (HMGA1) protein and gene expression correlate with ER-negativity and poor outcomes in breast cancer. *Breast Cancer Res. Treat.* **179**, 25–35, <https://doi.org/10.1007/s10549-019-05419-1> (2020).
- Roy, S. *et al.* HMGA1 overexpression correlates with relapse in childhood B-lineage acute lymphoblastic leukemia. *Leuk. Lymphoma* **54**, 2565–2567, <https://doi.org/10.3109/10428194.2013.782610> (2013).
- Xu, Y. *et al.* The HMG-I oncogene causes highly penetrant, aggressive lymphoid malignancy in transgenic mice and is overexpressed in human leukemia. *Cancer Res.* **64**, 3371–3375, <https://doi.org/10.1158/0008-5472.CAN-04-0044> (2004).
- Esposito, F. *et al.* HMGA1 pseudogenes as candidate proto-oncogenic competitive endogenous RNAs. *Oncotarget* **5**, 8341–8354, <https://doi.org/10.18632/oncotarget.2202> (2014).
- Esposito, F., De Martino, M., Forzati, F. & Fusco, A. HMGA1-pseudogene overexpression contributes to cancer progression. *Cell Cycle* **13**, 3636–3639, <https://doi.org/10.4161/15384101.2014.974440> (2014).
- Esposito, F. *et al.* HMGA1-pseudogene expression is induced in human pituitary tumors. *Cell Cycle* **14**, 1471–1475, <https://doi.org/10.1080/15384101.2015.1021520> (2015).
- De Martino, M., Forzati, F., Arra, C., Fusco, A. & Esposito, F. HMGA1-pseudogenes and cancer. *Oncotarget* **7**, 28724–28735, <https://doi.org/10.18632/oncotarget.7427> (2016).
- De Martino, M. *et al.* HMGA1P7-pseudogene regulates H19 and Igf2 expression by a competitive endogenous RNA mechanism. *Sci. Rep.* **6**, 37622, <https://doi.org/10.1038/srep37622> (2016).
- De Martino, M. *et al.* The HMGA1 Pseudogene 7 Induces miR-483 and miR-675 Upregulation by Activating Egr1 through a ceRNA Mechanism. *Genes (Basel)* **8**, <https://doi.org/10.3390/genes8110330> (2017).
- De Biase, D. *et al.* Characterization of inflammatory infiltrate of ulcerative dermatitis in C57BL/6NcrJ-Tg(HMGA1P6)1Pg mice. *Lab Anim.* **23677218815718**, <https://doi.org/10.1177/0023677218815718> (2018).
- Morse, H. C. III. *et al.* Bethesda proposals for classification of lymphoid neoplasms in mice. *Blood* **100**, 246–258 (2002).
- Chang, Y., Paige, C. J. & Wu, G. E. Enumeration and characterization of DJH structures in mouse fetal liver. *EMBO J.* **11**, 1891–1899 (1992).

27. Menzies-Gow, A. *et al.* Eotaxin (CCL11) and eotaxin-2 (CCL24) induce recruitment of eosinophils, basophils, neutrophils, and macrophages as well as features of early- and late-phase allergic reactions following cutaneous injection in human atopic and nonatopic volunteers. *J. Immunol.* **169**, 2712–2718 (2002).
28. Calvano, S. E. *et al.* A network-based analysis of systemic inflammation in humans. *Nature* **437**, 1032–1037, <https://doi.org/10.1038/nature03985> (2005).
29. Lippert, E. *et al.* Role of regulator of G protein signaling 16 in inflammation-induced T lymphocyte migration and activation. *J. Immunol.* **171**, 1542–1555 (2003).
30. Medzhitov, R. & Horng, T. Transcriptional control of the inflammatory response. *Nat. Rev. Immunol.* **9**, 692–703, <https://doi.org/10.1038/nri2634> (2009).
31. Kunej, T. *et al.* Functional UQCRC1 polymorphisms affect promoter activity and body lipid accumulation. *Obesity* **15**, 2896–2901, <https://doi.org/10.1038/oby.2007.344> (2007).
32. Aguilera-Aguirre, L. *et al.* Mitochondrial dysfunction increases allergic airway inflammation. *J. Immunol.* **183**, 5379–5387, <https://doi.org/10.4049/jimmunol.0900228> (2009).
33. Gong, Y. Y. *et al.* Hypermethylation of Cox5a promoter is associated with mitochondrial dysfunction in skeletal muscle of high fat diet-induced insulin resistant rats. *PLoS One* **9**, e113784, <https://doi.org/10.1371/journal.pone.0113784> (2014).
34. Liu, H. *et al.* Identification of a new Mpl-interacting protein, Atp5d. *Mol. Cell Biochem.* **391**, 85–94, <https://doi.org/10.1007/s11010-014-1990-x> (2014).
35. Sattiraju, A. *et al.* IL13RA2 targeted alpha particle therapy against glioblastomas. *Oncotarget* **8**, 42997–43007, <https://doi.org/10.18632/oncotarget.17792> (2017).
36. Laberge, R. M. *et al.* MTOR regulates the pro-tumorigenic senescence-associated secretory phenotype by promoting IL1A translation. *Nat. Cell Biol.* **17**, 1049–1061, <https://doi.org/10.1038/ncb3195> (2015).
37. Young, A. R. & Narita, M. SASP reflects senescence. *EMBO Rep.* **10**, 228–230, <https://doi.org/10.1038/embor.2009.22> (2009).
38. Salama, R., Sadaie, M., Hoare, M. & Narita, M. Cellular senescence and its effector programs. *Genes. Dev.* **28**, 99–114, <https://doi.org/10.1101/gad.235184.113> (2014).
39. Stamm, H. *et al.* Immune checkpoints PVR and PVRL2 are prognostic markers in AML and their blockade represents a new therapeutic option. *Oncogene* **37**, 5269–5280, <https://doi.org/10.1038/s41388-018-0288-y> (2018).
40. Couper, K. N., Blount, D. G. & Riley, E. M. IL-10: the master regulator of immunity to infection. *J. Immunol.* **180**, 5771–5777 (2008).
41. Cai, G. & Freeman, G. J. The CD160, BTLA, LIGHT/HVEM pathway: a bidirectional switch regulating T-cell activation. *Immunol. Rev.* **229**, 244–258, <https://doi.org/10.1111/j.1600-065X.2009.00783.x> (2009).
42. Mbongue, J. C. *et al.* The Role of Indoleamine 2, 3-Dioxygenase in Immune Suppression and Autoimmunity. *Vaccines* **3**, 703–729, <https://doi.org/10.3390/vaccines3030703> (2015).
43. Luo, X., Shi, F., Qiu, H., Tong, Y. & Gao, X. Identification of potential key genes associated with diffuse large B-cell lymphoma based on microarray gene expression profiling. *Neoplasma* **64**, 824–833, [https://doi.org/10.4149/neo\\_2017\\_603](https://doi.org/10.4149/neo_2017_603) (2017).
44. Chen, J. *et al.* Dysregulated CXCR4 expression promotes lymphoma cell survival and independently predicts disease progression in germinal center B-cell-like diffuse large B-cell lymphoma. *Oncotarget* **6**, 5597–5614, <https://doi.org/10.18632/oncotarget.3343> (2015).
45. Wang, L. *et al.* Blimp-1 induced by IL-4 plays a critical role in suppressing IL-2 production in activated CD4 T cells. *J. Immunol.* **181**, 5249–5256 (2008).
46. Nasir, A. *et al.* ZFP36L1 negatively regulates plasmacytoid differentiation of BCL1 cells by targeting BLIMP1 mRNA. *PLoS One* **7**, e52187, <https://doi.org/10.1371/journal.pone.0052187> (2012).
47. Gupta-Rossi, N. *et al.* Specific over-expression of deltex and a new Kelch-like protein in human germinal center B cells. *Mol. Immunol.* **39**, 791–799 (2003).
48. McNerney, M. E. *et al.* CUX1 is a haploinsufficient tumor suppressor gene on chromosome 7 frequently inactivated in acute myeloid leukemia. *Blood* **121**, 975–983, <https://doi.org/10.1182/blood-2012-04-426965> (2013).
49. Xu-Monette, Z. Y. *et al.* Clinical and Biologic Significance of MYC Genetic Mutations in De Novo Diffuse Large B-cell Lymphoma. *Clin. Cancer Res.* **22**, 3593–3605, <https://doi.org/10.1158/1078-0432.CCR-15-2296> (2016).
50. Karreth, F. A. *et al.* The BRAF pseudogene functions as a competitive endogenous RNA and induces lymphoma *in vivo*. *Cell* **161**, 319–332, <https://doi.org/10.1016/j.cell.2015.02.043> (2015).
51. Fedele, M. *et al.* HMGA2 cooperates with either p27(kip1) deficiency or Cdk4(R24C) mutation in pituitary tumorigenesis. *Cell Cycle* **17**, 580–588, <https://doi.org/10.1080/15384101.2017.1403682> (2018).
52. Fedele, M. *et al.* Overexpression of the HMGA2 gene in transgenic mice leads to the onset of pituitary adenomas. *Oncogene* **21**, 3190–3198, <https://doi.org/10.1038/sj.onc.1205428> (2002).
53. Palumbo, A. Jr. *et al.* UBE2C is overexpressed in ESCC tissues and its abrogation attenuates the malignant phenotype of ESCC cell lines. *Oncotarget* **7**, 65876–65887, <https://doi.org/10.18632/oncotarget.11674> (2016).
54. Bleesing, J. J. & Fleisher, T. A. Human B cells express a CD45 isoform that is similar to murine B220 and is downregulated with acquisition of the memory B-cell marker CD27. *Cytometry B Clin. Cytom.* **51**, 1–8, <https://doi.org/10.1002/cyto.b.10007> (2003).
55. Forzati, F. *et al.* miR-155 is positively regulated by CBX7 in mouse embryonic fibroblasts and colon carcinomas, and targets the KRAS oncogene. *BMC Cancer* **17**, 170, <https://doi.org/10.1186/s12885-017-3158-z> (2017). <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
56. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21, <https://doi.org/10.1093/bioinformatics/bts635> (2013).
57. Hubbard, T. *et al.* The Ensembl genome database project. *Nucleic Acids Res.* **30**, 38–41 (2002).
58. [http://www.ensembl.org/Mus\\_musculus/Info/Index](http://www.ensembl.org/Mus_musculus/Info/Index).
59. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169, <https://doi.org/10.1093/bioinformatics/btu638> (2015).
60. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq. 2. *Genome Biol.* **15**, 550, <https://doi.org/10.1186/s13059-014-0550-8> (2014).
61. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550, <https://doi.org/10.1073/pnas.0506580102> (2005).
62. Shaffer, A. L. *et al.* A library of gene expression signatures to illuminate normal and pathological lymphoid biology. *Immunol. Rev.* **210**, 67–85, <https://doi.org/10.1111/j.0105-2896.2006.00373.x> (2006).
63. Tarantelli, C. *et al.* PQR309 Is a Novel Dual PI3K/mTOR Inhibitor with Preclinical Antitumor Activity in Lymphomas as a Single Agent and in Combination Therapy. *Clin. Cancer Res.* **24**, 120–129, <https://doi.org/10.1158/1078-0432.CCR-17-1041> (2018).
64. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCB gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
65. Nicolau-Neto, P. *et al.* UBE2C Is a Transcriptional Target of the Cell Cycle Regulator FOXM1. *Genes (Basel)* **9**, <https://doi.org/10.3390/genes9040188> (2018).
66. Puca, F. *et al.* HMGA1 negatively regulates NUMB expression at transcriptional and post transcriptional level in glioblastoma stem cells. *Cell Cycle* **18**, 1446–1457, <https://doi.org/10.1080/15384101.2019.1618541> (2019).

## Acknowledgements

This work was supported by grants from: PNR-CNR Aging Program 2012–2014, CNR Flagship Projects (Epigenomics-EPIGEN), Associazione Italiana per la Ricerca sul Cancro (AIRC IG 11477). M.D.M is recipient of a Fellowship from “Programma Valere Plus”, University of Campania “L. Vanvitelli”, Caserta, Italy. The Authors sincerely thank Mrs. Mariarosaria Montagna for her research support.

## Author contributions

M.D.M., F.F., A.F. and F.E. made substantial contributions to conception and design of the study, as well as analysis and interpretation of data and drafting of the article; S.C.C., D.D.B., O.P. and M.P. significantly contributed to the pathological evaluation of specimens and revision of the paper; M.D.M. realized FACS analyses; E.G. contributed to data interpretation; F.B. performed the bioinformatic analyses of the RNA-Seq data, giving substantial contributions to the interpretation of data and drafting of the article; G.P., C.A., A.B., G.D.V. and P.C. contributed to conception of the study and critically revised the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to A.F. or F.E.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons 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 Author(s) 2020