

SCIENTIFIC REPORTS



OPEN

HMGA1P7-pseudogene regulates *H19* and *Igf2* expression by a competitive endogenous RNA mechanism

Received: 21 April 2016
Accepted: 01 November 2016
Published: 22 November 2016

Marco De Martino¹, Floriana Forzati¹, Marianna Marfella¹, Simona Pellicchia¹, Claudio Arra², Luigi Terracciano³, Alfredo Fusco¹ & Francesco Esposito¹

Recent studies have revealed that pseudogene transcripts can function as competing endogenous RNAs, and thereby can also contribute to cancer when dysregulated. We have recently identified two pseudogenes, *HMGA1P6* and *HMGA1P7* for the *HMGA1* gene whose overexpression has a critical role in cancer progression. These pseudogenes work as competitive endogenous RNA decoys for *HMGA1* and other cancer related genes suggesting their role in carcinogenesis. Looking for new *HMGA1* pseudogene ceRNAs, we performed RNA sequencing technology on mouse embryonic fibroblasts deriving from transgenic mice overexpressing *HMGA1P7*. Here, we report that *HMGA1P7* mRNA sustains the *H19* and *Igf2* overexpression by acting as miRNA decoy. Lastly, the expression of *HMGA1P7* was significantly correlated with *H19* and *IGF2* levels in human breast cancer thereby suggesting a role for *HMGA1P7* deregulation in this neoplasia.

Pseudogenes are a subclass of long non coding RNA (lncRNA) sharing high sequence identity with protein-coding parental counterparts. As stated by the GENCODE pseudogene annotations (v.17), there are almost 15,000 human pseudogenes¹. They frequently possess features, such as premature stop codons, deletions/insertions, or frame shift mutations, that impede them to produce functional proteins. There are three groups of pseudogenes: processed, duplicated, and unitary^{1,2}. Processed pseudogenes do not have introns and are thought to arise from reverse transcription of mRNA followed by reinsertion into the genome^{1,2}. Duplicated pseudogenes contain introns and sometimes even upstream regulatory elements since they are produced by gene duplication. For each pseudogene belonging to these two classes there is an associated protein-coding gene that is highly similar in sequence^{1,2}. The last type of pseudogenes are the unitary ones, which take place when protein-coding genes accumulate mutations and lose their coding potential^{1,2}. Consequently, unitary pseudogenes do not have parental genes. From the time of their discovery in 1977, pseudogenes have commonly been thought as “biologically inconsequential” and non-functional³. However, recent studies have unveiled different mechanisms by which pseudogenes control gene expression such as the generation of siRNAs^{4,5}, competition for RNA-binding proteins or the translation apparatus^{6–8}, and engagement of proteins by pseudogene antisense RNAs to corresponding sites in the parental gene to modify chromatin transcription and remodeling^{9,10}. The latest function identified for pseudogenes is post-transcriptional regulation of mRNA levels by competing for microRNAs (miRNAs). Indeed, processed pseudogenes maintain 5' and 3' untranslated region (UTR) sequences of their parental genes¹¹. Given that miRNAs inhibit target gene expression by binding to the 3' UTR, pseudogenes can be targeted by miRNAs that modify the expression of coding genes. Definitely, pseudogene transcripts exert regulatory control of their parental gene expression levels by competing for the same miRNAs¹².

We have recently characterized two processed pseudogenes, *HMGA1P6* and *HMGA1P7*, for the *HMGA1* gene that codes for the HMGA1a and HMGA1b proteins highly overexpressed in most of the human malignancies^{13,14}.

¹Istituto di Endocrinologia ed Oncologia Sperimentale del CNR c/o Dipartimento di Medicina Molecolare e Biotecnologie Mediche, Scuola di Medicina e Chirurgia di Napoli, Università degli Studi di Napoli “Federico II”, via Pansini 5, 80131 Naples, Italy. ²Istituto Nazionale dei Tumori, Fondazione Pascale, via Mariano Semmola, 80131 Naples, Italy. ³Institute of Pathology, Molecular Pathology Division, University of Basel, Schonbeinstrasse 40, 4003 Basel, Switzerland. Correspondence and requests for materials should be addressed to A.F. (email: alfusco@unina.it) or F.E. (email: francesco.esposito2@unina.it)

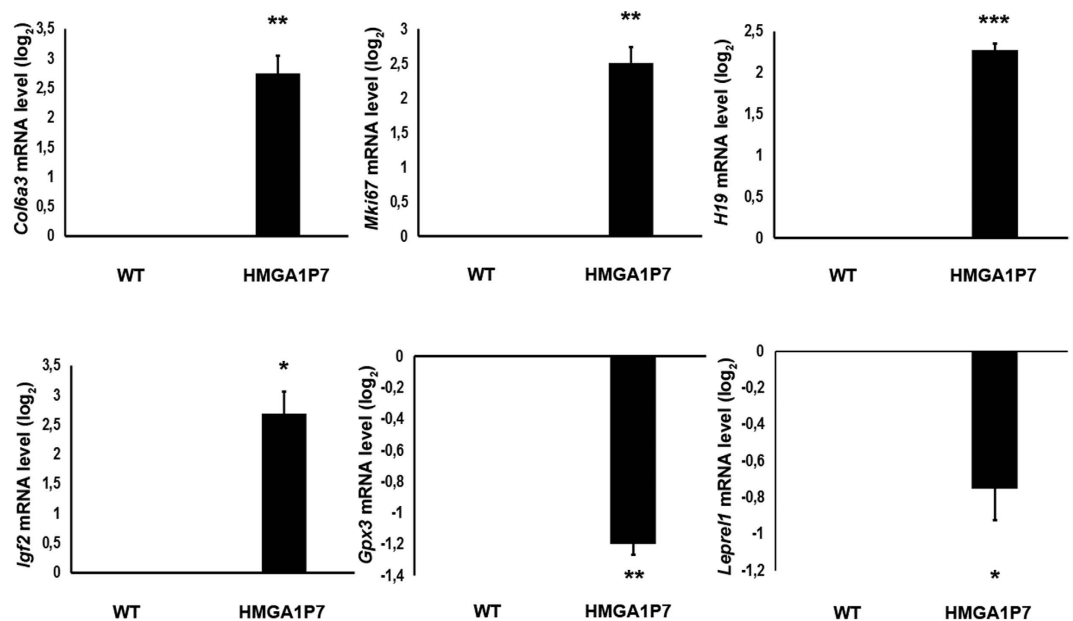


Figure 1. Validation of RNA-seq analyses on *HMGA1P7* MEFs. qRT-PCR analysis of selected deregulated genes from RNA-seq performed on WT and *HMGA1P7* transgenic MEFs. The results are reported as the mean of values. The error bars represent mean \pm SE; * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ (t test).

Moreover, it has been previously demonstrated an association between *HMGA1* overexpression and a poor patient survival¹³, and that their overexpression is even required for cell transformation^{15,16}, and is able to induce benign and malignant neoplasias in mice¹³. *HMGA1P6* and *HMGA1P7* pseudogenes, present only in human genome, have preserved seed matches for miRNAs targeting the *HMGA1* oncogene. *HMGA1* pseudogenes (*HMGA1Ps*) overexpression, working as competitive endogenous RNA (ceRNA), increases *HMGA1* protein levels by blocking the suppression of *HMGA1* protein synthesis exerted by miRNAs^{17–20}. *HMGA1Ps* have also oncogenic activity by suppressing apoptosis and promoting cell proliferation and migration^{17–20}. Moreover, we have previously show that *HMGA1Ps* are overexpressed in anaplastic thyroid carcinomas but not in the differentiated ones, indicating a critical role of them in cancer progression¹⁷. Since the *HMGA1Ps* contain several seed sequences for miRNAs, their overexpression derepresses the expression of different cancer-related genes, as already demonstrated for *HMGA2*, *VEGF*, *EZH2*^{17–20}. Therefore, the aim of this study has been to find novel ceRNA interactors differentially expressed in *HMGA1P7* transgenic mouse embryonic fibroblasts (MEFs) with respect to the wild-type (WT) ones, which do not express *HMGA1* pseudogenes, using a RNA sequencing (RNA-seq) approach. By this analysis, we found a set of mRNAs up- or down-regulated in *HMGA1P7* overexpressing MEFs in comparison with WT cells. Among them, we focused our attention on two of the most overexpressed and *HMGA1P7* miRNA-sharing genes: *H19* and *insulin-like growing factor 2* (*Igf2*).

H19 and *IGF2* genes are closely linked, showing highly similar patterns of gene expression, but they are reciprocally imprinted. In fact, *H19* is expressed solely from the maternally inherited chromosome, whereas *IGF2* expression is from the paternal chromosome. In particular, the noncoding *H19* has a critical role in genomic imprinting during cell growth and development²¹. The loss of imprinting results in misexpression of *H19* and was detected in many tumors including hepatocellular²², bladder²³, gastric^{24,25} and colon²⁶ cancer. *IGF2* codes for a mitogenic growth factor that is active in early development and has a critical role in embryonic and fetal growth²⁷. Increased expression of *IGF2* is a common feature of both pediatric and adult malignancies²⁷, and mounting evidence implicates *IGF2* as a major factor contributing to oncogenesis^{27–29}. Here, we report that *HMGA1P7* mRNA induces the *H19* and *Igf2* overexpression by acting as miRNA decoy.

Results

RNA-seq on *HMGA1P7* transgenic MEFs. To identify the genes regulated by *HMGA1P7* expression, we analyzed the whole transcriptome of WT and *HMGA1P7* transgenic MEFs by RNA-seq analyses. To this aim, the entire population of RNA transcripts extracted from WT and *HMGA1P7*-MEFs were sequenced. The genome-wide RNA expression profiles studies reveal that about one hundred fifty transcripts (32 upregulated and 116 downregulated) were regulated by *HMGA1P7* expression with a significant fold-change variation (FDR adjusted p-value of 0.05). To validate the results obtained by RNA-seq, we analyzed the expression of some upregulated genes such as *Collagen Type VI Alpha 3* (*Col6a3*), *Marker of Proliferation Ki-67* (*Mki67*), *H19*, *Igf2* and downregulated genes such as *Glutathione Peroxidase 3* (*Gpx3*), *Leprecan-Like 1* (*Leprel1*) by Real-time PCR (qRT-PCR). As shown in Fig. 1, the quantitative qRT-PCR analyses confirmed the data obtained from the RNA-seq analyses. Interestingly, these genes have been related to several human cancers (colon, gastric, liver, breast and hematological cancers), and are considered possible therapeutic targets^{30–35}.

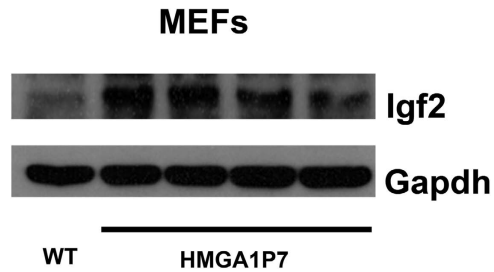


Figure 2. Igf2 is upregulated in *HMGA1P7* MEFs. Western blot analysis of Igf2 from WT and *HMGA1P7* transgenic MEFs.

Among the differentially expressed mRNAs found in MEFs overexpressing *HMGA1P7*, we focused our attention on *H19* and *Igf2* since they, other than to be involved in carcinogenesis, showed the highest fold change among the upregulated genes, and are also targeted by several miRNAs that are able to bind to the *HMGA1P7* mRNA. Western blot analysis for Igf2 confirmed the qRT-PCR data (Fig. 2). Moreover, qRT-PCR and Western blot analysis showed that *H19* and *Igf2* were also upregulated in heart, spleen and kidney from *HMGA1P7* adult transgenic mice (Fig. 3A,B and C). As expected from previous results, qRT-PCR shows upregulation of *H19* and *Igf2* following *HMGA1P7* pseudogene overexpression in NIH3T3 cells (Fig. 3D). Western blot confirms *Igf2* upregulation also at protein level in tissues from *HMGA1P7* and in the *HMGA1P7*-transfected NIH3T3 cells (Fig. 3C and D).

Taken together, these results strongly support the hypothesis that *HMGA1P7* could act as ceRNA for *H19* and *Igf2*.

***HMGA1P7* act as decoy for *H19* and *Igf2* targeting miRNAs.** To test whether the effect of the *HMGA1P7* pseudogene on *H19* and *Igf2* expression is dependent on sharing targeting-miRNAs, we evaluated the ability of *HMGA1P7*-targeting miRNAs¹⁷ to bind to *H19* and *Igf2*. To this aim, we transfected miR-15, miR-16, miR-214 and miR-761 (already reported to target *HMGA1P7*)¹⁷ into NIH3T3 cells, and analyzed *H19* and *Igf2* mRNA levels by qRT-PCR. As presented in Fig. 4A, the transfection of the *HMGA1P7*-targeting miRNAs yield a significant reduction of *H19* and *Igf2* mRNA levels. Western blot confirms Igf2 downregulation also at protein level following the transfection of the *HMGA1P7*-targeting miRNAs (Fig. 4A). To define whether the *HMGA1P7*-targeting miRNAs straightly interacted with *Igf2* mRNA, we cloned the *Igf2* 3' UTR downstream of the luciferase open reading frame. This reporter vector was transfected into NIH3T3 cells together with miRNA precursors and a control non-targeting scrambled oligonucleotide. The luciferase signal was considerably lower after transfection with miR-15, miR-16, miR-214 and miR-761 in comparison with the cells transfected with the scrambled oligonucleotide (Fig. 4B). The overexpression of *H19* and *Igf2* induced by upregulation of *HMGA1P7* was depleted in Dicer-knockdown cells (Fig. 4C) then supporting the hypothesis that *HMGA1P7*, *H19* and *Igf2* follow the same miRNA-mediated post-transcriptional regulation. In fact, silencing of Dicer, the enzyme that leads miRNA maturation process, results in reduced levels of mature miRNAs compared to control. Moreover, to verify whether *H19* and *Igf2* can act as ceRNA each-other, we transfected siRNA-*Igf2* into NIH3T3 cell line in combination or not with Anti miR-16, which is able to block miR-16 repression on *HMGA1P7*, *Igf2* and *H19*, and a siRNA-control, then evaluating the *H19* mRNA levels. As proposed by our model, siRNA-*Igf2* transfection induces a significant *H19* downregulation, that is reverted by the transfection with the Anti miR-16 oligonucleotide, suggesting that both *H19* and *Igf2* transcripts can talk each-other through miRNAs mediation (Fig. 4D). These data are consistent with the hypothesis that *HMGA1P7* requires mature miRNAs to regulate *H19* and *Igf2* levels.

***HMGA1P7*, *H19* and *IGF2* expression positively correlates in human breast cancer.** Then, we investigated whether *HMGA1P7* functions as ceRNA through, or partially through *H19*, *IGF2* and *HMGA1* in breast cancer human cells. As expected, we found upregulation of *H19*, *IGF2* and *HMGA1* following *HMGA1P7* overexpression in MCF7 cells (human breast adenocarcinoma cell line) (Fig. 5A). Moreover, MCF7-*HMGA1P7* cells grow faster than the control transfected cells as consequence of *HMGA1P7* ceRNA pathway activation (Fig. 5B).

To confirm whether *HMGA1P7* works as miRNA sponge for the regulation of *H19* and *IGF2* expression levels also in human cancer, we evaluated the expression of *H19*, *IGF2* and *HMGA1P7* in a panel of breast carcinoma samples by qRT-PCR, since *H19* and *IGF2* have been reported to be overexpressed in this type of tumor^{36,37}. As shown in Fig. 5C, *HMGA1P7* was overexpressed in most of the carcinoma samples as well as *H19* and *IGF2*. Moreover, the direct correlation between *HMGA1P7* and *H19* expression (Spearman $r = 0,8656$; $p < 0,001$) and between *HMGA1P7* and *IGF2* expression (Spearman $r = 0,7958$; $p < 0,001$) underlines that these genes are co-regulated (Fig. 5D). Altogether, these results strongly support the idea that *HMGA1P7* could act as ceRNAs in human breast cancer and represent a novel potential mechanism accounting for *H19* and *IGF2* upregulation in these tumors.

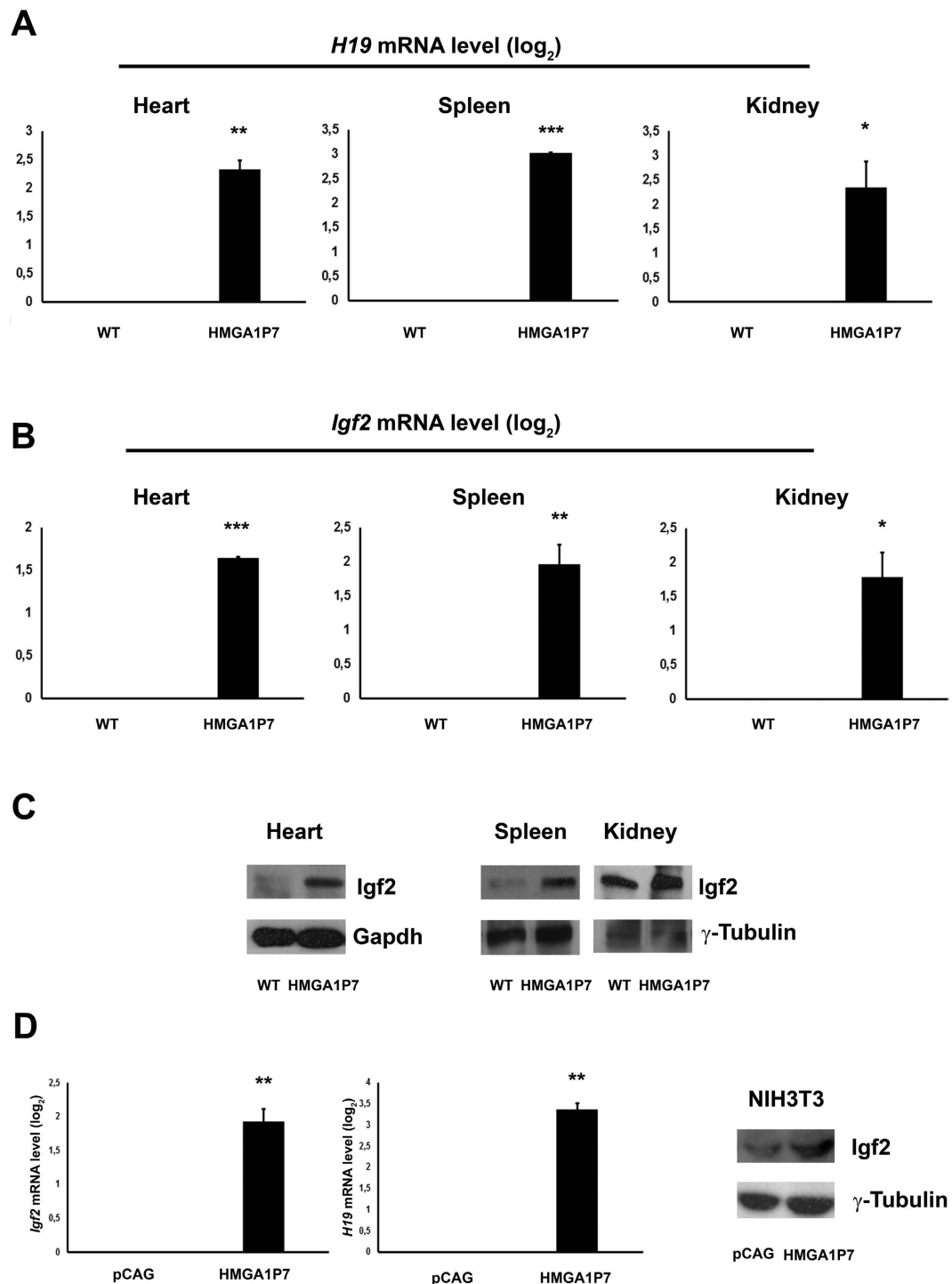


Figure 3. *H19* and *Igf2* are positively regulated by *HMGA1P7*. (A) qRT-PCR analysis of *H19* from hearts, spleens and kidneys of WT and *HMGA1P7* transgenic mice. (B) qRT-PCR analysis of *Igf2* from hearts, spleens and kidneys of WT and *HMGA1P7* transgenic mice. (C) Western blot analysis of *Igf2* from heart, spleen and kidney of WT and *HMGA1P7* transgenic mice. (D) Left Panel, qRT-PCR analysis of *H19* and *Igf2* from control and *HMGA1P7* overexpressing NIH3T3 cells. Right Panel, Western blot analysis of *Igf2* from control and *HMGA1P7* overexpressing NIH3T3 cells. The results are reported as the mean of values. The error bars represent mean \pm SE; * $P < 0.05$ ** $P < 0.01$ *** $P < 0.001$ (t test).

Discussion

lncRNAs are involved in regulating the complexity of biological processes with specific regulatory mechanisms, thereby, attracting considerable research interest³⁸. We have previously isolated and characterized two pseudogenes, *HMGA1P6* and *HMGA1P7*, for the *HMGA1* gene and demonstrated that they act as decoys for

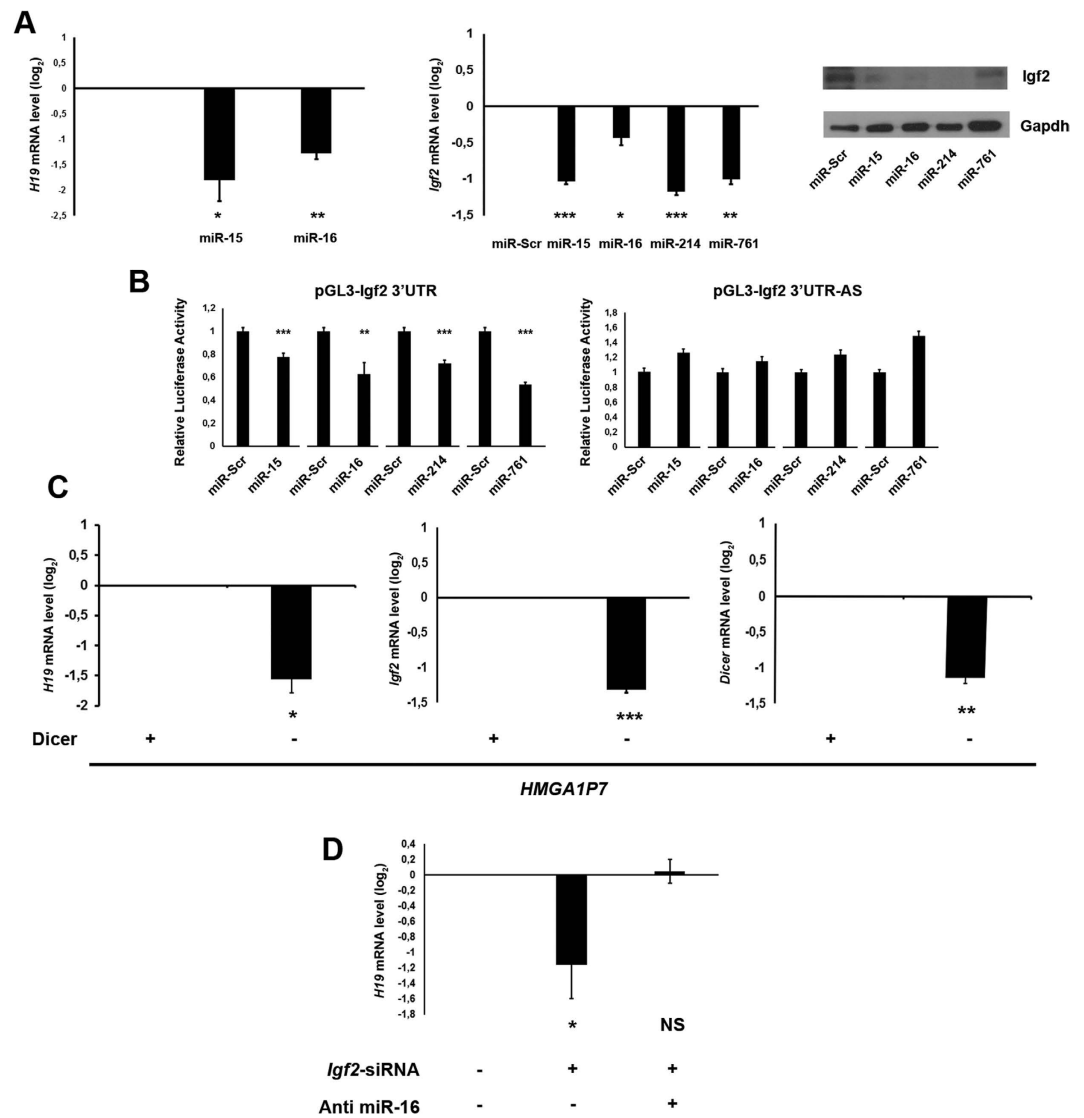


Figure 4. *H19* and *Igf2* are targeted by *HMGA1P7*-targeting miRNAs. (A) qRT-PCR analysis of *H19* (left Panel) and *Igf2* (middle Panel) mRNA from the NIH3T3 cells transfected with scrambled-oligonucleotide, miR-15, miR-16, miR-214 and miR-761. Right Panel, Western blot analysis of *Igf2* in the same samples as in the middle panel. (B) *Igf2* was cloned into the pGL3 control vector. Relative luciferase activity in HEK293 cells transiently transfected with miR-15, miR-16, miR-214, miR-761 and a control scrambled oligonucleotide. (C) *H19* and *Igf2* mRNA levels 24 h after the transfection of *HMGA1P7* in scrambled oligonucleotide or siRNA-Dicer NIH3T3 transfected cells. (D) qRT-PCR analysis of *H19* mRNA levels from the NIH3T3 cells transfected with siRNA-control, siRNA-*Igf2* alone or in combination with the Anti miR-16 oligonucleotide.

HMGA1-targeting miRNAs¹⁷. In fact, their overexpression enhances *HMGA1* protein levels whereas their knocking down results in the reduction of *HMGA1* mRNA and protein amounts. Moreover, their decoy activity protected the synthesis of other proteins involved in carcinogenesis^{17,18}. In this study, we used RNA-seq technology to identify additional mRNAs differentially expressed in MEFs transgenic for *HMGA1P7*.

We found that the expression of several genes were influenced by *HMGA1P7* including also genes involved in cancer progression such as *Col6a3*, *Mki67*, *H19*, *Igf2*, *Gpx3* and *Leprel1*^{30–37}. Indeed, oncomine analyses and tissue-microarray immunohistochemistry showed overexpression of *COL6A3* in colorectal carcinomas that was significantly and directly correlated with Dukes stage, T stage, stage, recurrence and smoking status and then with a poor prognosis³⁰. The *Mki-67* protein (also known as *Ki67*) is a cellular marker for proliferation. *Ki-67* protein is expressed during all active phases of the cell cycle (G_1 , S, G_2 , and mitosis), but is absent from resting cells (G_0)³². *GPX3* gene codes for the Glutathione peroxidase 3, also known as plasma glutathione peroxidase (GPX-P), the variations in activity of GPX1, GPX2, and GPX3 isoforms may be associated with the development of cancers, for example, prostate cancer or even colon cancer³⁹. *Leprecan-like 1* is a potential tumor suppressor gene since it has been demonstrated to be downregulated in the hepatocarcinoma tissues and its overexpression inhibits cancer cell proliferation and colony formation through regulation of the cell cycle by downregulation of cyclins⁴⁰.

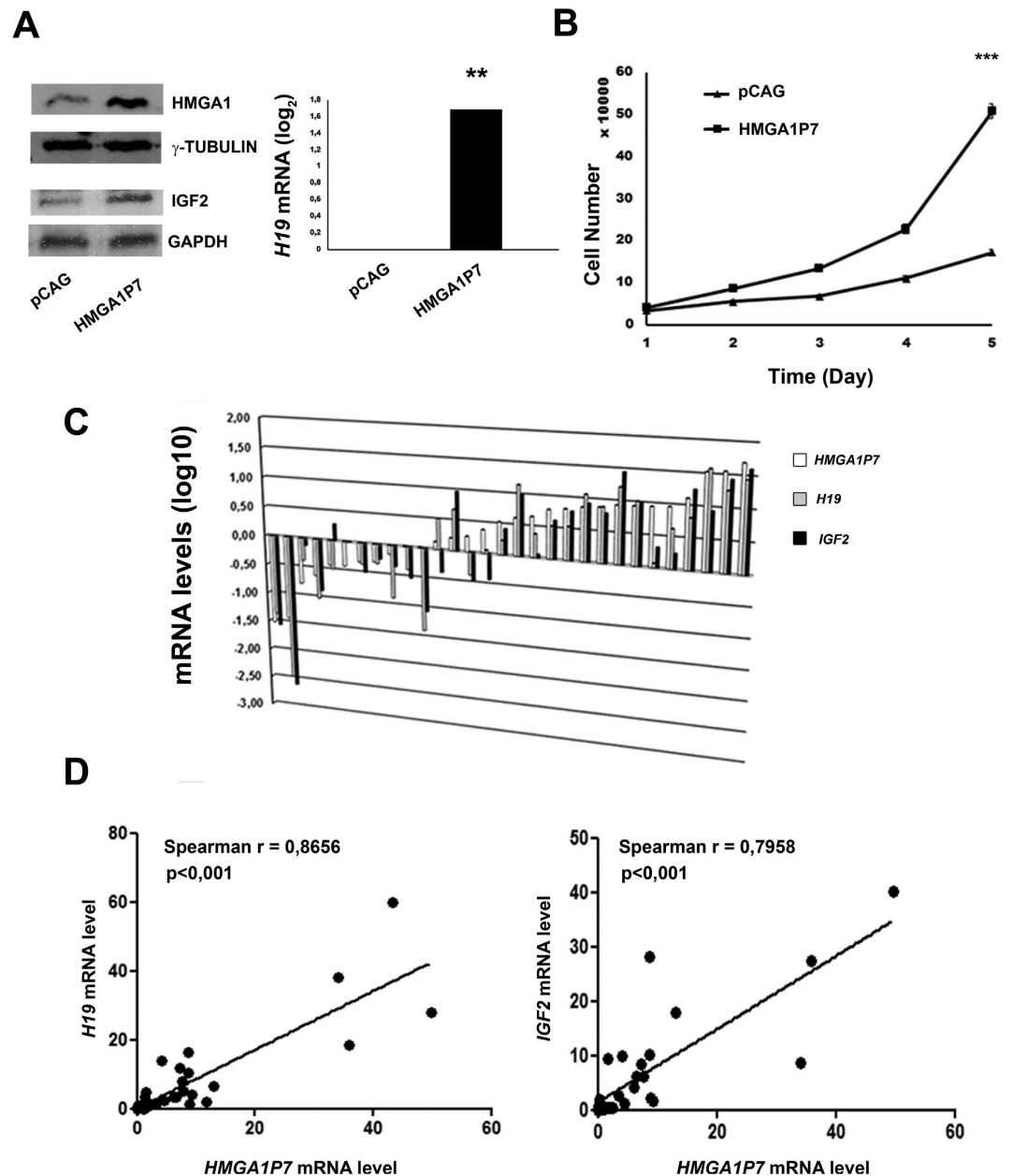


Figure 5. *H19* and *IGF2* expression positively correlates with *HMGA1P7* in breast cancer. (A) Left Panel, Western blot analysis of HMGA1 and IGF2 protein levels in control and *HMGA1P7*-overexpressing MCF7 cells. Right Panel, qRT-PCR analysis of *H19* expression of control and *HMGA1P7*-overexpressing MCF7 cells. (B) MCF7 cell proliferation of control and *HMGA1P7*-overexpressing cells (C) qRT-PCR analysis in tumor and normal breast tissues. The fold change indicates the relative change in expression levels between tumor samples and normal samples, assuming that the value of normal sample is equal to 1. (D) Correlation analysis of *HMGA1P7* versus *H19* and *HMGA1P7* versus *IGF2* are shown. The Spearman's rank correlation coefficient is shown.

Deregulation of *H19* noncoding gene was found in many tumors such as hepatocellular and bladder cancer^{22,23}. Finally, *IGF2* overexpression is widely reported in pediatric and adult tumors²⁷, and several studies involve *IGF2* as a key factor leading to cancerogenesis^{27–29}.

Among the most deregulated genes, we selected and studied *H19* non-coding gene and *Igf2*, that share several miRNAs with *HMGA1P7*. Here, we report the ceRNA relationship between *HMGA1P7*, *H19* and *Igf2*. We demonstrate that *HMGA1P7* overexpression increases *H19* and *Igf2* levels inhibiting their mRNA suppression by miRNAs that target *HMGA1P7* gene, namely, miR-15, miR-16, miR-214, and miR-761. Interestingly, preliminary results show an analogous ceRNA connection between *H19*, *Igf2* and *HMGA1P6*, supporting the oncogenic role of the both *HMGA1* pseudogenes.

Finally, we show that expression of *HMGA1P7* significantly correlates with *H19* and *IGF2* levels in human breast cancer, suggesting the upregulation of *HMGA1P7* may increase *H19* and *IGF2* expression by a ceRNA mechanism then contributing to cancer progression. Interestingly, the oncogenic role of *HMGA1P7* is also supported by the development of malignant hematological neoplasias in *HMGA1P7* transgenic mice (manuscript in preparation).

Then, the data reported here confirm the oncogenic role of the *HMGA1P7* pseudogene that is exerted by the increased expression through a ceRNA mechanism of *HMGA1* and other cancer-related genes. Future studies need, however, to characterize other genes regulated by the *HMGA1* pseudogenes and thereby better define the mechanisms by which they can contribute to cancer progression.

Materials and Methods

Cell culture and transfections. MEFs and MCF7 were cultured in DMEM supplemented with 10% foetal calf serum (Thermo Fisher Scientific Inc). NIH3T3 cells were maintained in DMEM supplemented with 10% calf serum (Thermo Fisher Scientific Inc), glutamine and antibiotics. MycoAlert (Lonza) was regularly used to test that cells were not infected by mycoplasma. Lipofectamine plus reagent was used to transfect the cells (Thermo Fisher Scientific Inc) according to the manufacturer's instructions. The transfected cells were selected in a medium containing geneticin (Sigma). Transfection efficiency was tested for each experiment by assessing GFP signal. To inhibit *Dicer* and *Igf2* expression, small interfering RNAs and corresponding scramble small interfering RNAs were designed and used as suggested by the manufacturer (RIBOXX).

RNA-sequencing. RNA samples were initially checked for quality and quantity using a Bioanalyzer with the total RNA Pico chip (Agilent Technologies, Inc) and a Qubit® with RNA Assay Kit (Thermo Fisher Scientific Inc) respectively.

Spike-In Mix 1 and Spike-In Mix 2, each containing the full complement of 92 polyadenylated transcripts from the ERCC plasmid reference library, were added to samples.

mRNA was selected from total RNA preparation using MicroPolyA Purist kit (Ambion, Inc).

SOLiD™ Total RNA-Seq Kit (Life Technologies Corporation) was used to convert RNA transcripts into a cDNA library, starting from low input amounts of poly(A) RNA, for analysis on the 5500 Genetic Analysis System. First of all, mRNA was incubated in a thermal cycler at 95 °C for 10 minutes to fragment the RNA by chemical hydrolysis. The RNA digested was hybridized and ligated with Solid specific adaptors. Two rounds of size selection using Agencourt® AMPure® XP Reagent were performed to increase the percentage of library inserts that were in the desired size range, >150 bp. The purified DNA was amplified and barcoded by 18 PCR cycles to enable sequencing of all the samples in a single multiplexed SOLiD System sequencing run.

The yield and size distribution of the amplified DNA libraries was assessed running the samples on an Agilent® 2100 Bioanalyzer™ Instrument with the DNA HS Kit following the manufacturer's instructions and Qubit® dsDNA HS kit.

Ninety nanograms of each library was pooled together and one E80 emulsion was prepared following SOLiD® EZ Bead™ System. About 400 millions of pooled templated beads were deposited on a 4 lanes of a 6 lanes slide and the sequencing was performed up to a read length of 50 bp, based on 5500 Genetic Analysis System Run sequencer protocol.

Bioinformatic Analysis. Four samples were analysed: two from WT and two from *HMGA1P7* transgenic MEFs. The comparison performed was WT versus transgenic, two biological replicates for condition.

Sequencing reads in SOLiD "xsq" format were mapped against the reference genome (UCSC GRC38/mm10); reference gene structure was Refseq from the refGene.txt file of the UCSC genome browser FTP site; the mapping software was the Whole Transcriptome Analysis module from the Lifescope 2.5.1 Genomic Analysis Software analysis suite from Applied Biosystems/ThermoFisher Scientific.

A filter file, containing 6415 sequences (sequencing adaptors; barcodes; tRNAs; rRNAs; rRNA fragments; repetitive sequences; ERCC RNA sequences) was used (1) to filter the transcripts for non-significant reads and (2) to quantify the absolute expression using the External RNA Controls Consortium (ERCC) RNA Spike-In Mix.

The genome-mapped reads were then correlated with Refseq genes and the resulting gene-associated read counts were analysed with a Genomni proprietary procedure based on the Bioconductor library edgeR⁴¹. The chosen limit for evaluating differential expression was 5 counts per millions in at least half of the examined samples. The normalization procedure used was the standard for edgeR (TMM). Genes were called differentially expressed when the comparison was evaluated with a FDR < 0.05. Absolute gene expression was evaluated from the read counts in RPKM (Reads per kilo base per million mapped reads).

Primary gene annotation was performed using the Bioconductor libraries biomaRt and GOstats, while functional clustering of the genes was performed using the DAVID functional annotation web site (<https://david.ncifcrf.gov/>).

Mouse embryo and tissue samples. The use of mouse embryos and tissues and the experiments performed in this study were approved by the Ministero della Salute; the methods and experiments were carried out in accordance with the approved guidelines by the Ministero della Salute.

Human breast tissue samples. Normal and neoplastic human breast tissues were obtained from surgical specimens and immediately frozen in liquid nitrogen. 32 breast samples were collected at the Institute of Pathology, University of Basel, Switzerland. The tumor samples were frozen until required for RNA extraction. The use of human tissues and the experiments performed in this study were approved by the Institute of Pathology, Molecular Pathology Division, University of Basel; the methods and experiments were carried out

in accordance with the approved guidelines by the University of Basel. We declare that informed consent for the scientific use of biological material was obtained from all patients.

RNA extraction and quantitative reverse transcription PCR. Total RNA was extracted from cells or tissues with TRIsure (Aurogene) according to the manufacturer's instructions. For mRNA transcripts detection, we reverse transcribed total RNA from samples by using the QuantiTect Reverse Transcription Kit (Qiagen), and then Real-time PCR was performed by using Power SYBR Green PCR Master Mix (Bio-Rad) and the following primers:

HMGA1P7-Fw 5'-gctccttctcggtcctc-3'
HMGA1P7-Rev 5'-gcttggcctctttatgg-3'
mIgf2 Fw 5'-cctccttacccaactcaggt-3'
mIgf2 Rv 5'-aagagatgagaagcaccacac-3'
mh19 Fw 5'-atgtcttcatttccctatagcc-3'
mh19 Rv 5'-gtcatcctcgcttcagtg-3'
mG6pd-Fw 5'-cagcggcaactaaactcaga-3'
mG6pd-Rev 5'-ttcctcaggatcccacac-3'
mCol6a3 Fw 5'-ggaggtgtacaggaagttccac-3'
mCol6a3 Rev 5'-gactgagccgtcaaagagga-3'
mMki67 Fw 5'-gctgtcctcaagacaatcatca-3'
mMki67 Rev 5'-ggcgttatcccaggagact-3'
mGpx3 Fw 5'-gtgaacggggagaaagagc-3'
mGpx3 Rev 5'-tgagcccaggagtctgc-3'
mLeprel1 Fw 5'-tggaccctttaccgagaa-3'
mLeprel1 Rev 5'-tgatccaagatggcaatcac-3'
hActin Fw 5'-ccaaccgagagaagatga-3'
hActin Rv 5'-ccagaggcgtacaggatag-3'
hH19 Fw 5'-ttactcctccacggagtcg-3'
hH19 Rv 5'-gagctgggtagcaccatttc-3'
hIGF2 Fw 5'-gctggcagaggagtcc-3'
hIGF2 Rv 5'-gggattcccattggtgtct-3'

The $2^{-\Delta\Delta CT}$ formula was used to calculate the differential gene expression, and described elsewhere⁴².

Plasmid and miRNA oligonucleotides. For transfection of miRNA oligonucleotides, cells were transfected with 50 nmol/ml of miRNA precursors or with a control no-targeting scrambled oligonucleotides (Thermo Fisher Scientific Inc) using siPORT neoFX Transfection Agent (Thermo Fisher Scientific Inc). For transfection of Anti miR-16 oligonucleotides, cells were transfected with 50 nmol/ml of Anti miR-16 or with a control no-targeting scrambled oligonucleotides (Thermo Fisher Scientific Inc).

For *Igf2* luciferase reporter construct (pGL3-*Igf2*), the miRNA seed sequence containing fragment of *Igf2* gene (ENSMUST00000000033) was amplified by using the primers:

Igf2 Fw 5'-aatttctagaccaaaatctcacttttccc-3'
Igf2 Rev 5'-aatttctagagatggccataggtgtgctc-3'.

The amplified fragment was cloned into pGL3-Control luciferase reporter vector (Promega).

All the generated vectors were confirmed by sequencing. The Renilla luciferase vector (pRL-CMV), for transient transfection efficiency, was purchased from Promega.

Protein extraction, western blotting and antibodies. Protein extraction and Western blotting were performed as previously described^{43,44}. The primary antibodies used were anti-IGF2 (#32592) from Sabbiotech; anti-GAPDH (sc-32233) and anti- γ -Tubulin (sc-17787) from Santa Cruz Biotechnology. Blots were visualized by using the Western blotting detection reagents (Thermo Fisher Scientific Inc).

Dual-luciferase reporter assay. For dual-luciferase reporter assay, 3×10^5 NIH3T3 cells were co-transfected in 6-well plates with the pGL3-*Igf2* or the pGL3-*H19* luciferase reporter vectors, together with the Renilla luciferase plasmid and miRNA precursors or a control no-targeting scrambled oligonucleotides (Thermo Fisher Scientific Inc), using siPORT neoFX Transfection Agent (Thermo Fisher Scientific Inc). The pRL-TK control vector expressing Renilla luciferase (Promega) was used for normalization of cell number and transfection efficiency. Luciferase activity was measured 48 hours after transfection using the Dual-Luciferase Reporter Assay System (Promega) with a Lumat LB 9507 apparatus (Berthold Technologies).

Growth curve assay. For each experimental point 3×10^4 cells were plated in a 60 mm plate. Cells were counted in triplicate for 5 days with Burkert hemocytometer chamber.

Statistical analysis. Data were analyzed using a two-sided unpaired t test (GraphPad Prism, GraphPad Software, Inc.). Values of $P < 0.05$ were considered statistically significant. Regression analysis, correlation coefficients and statistical analysis were generated using GraphPad Prism, GraphPad Software, Inc.

References

1. Welch, J. D., Baran-Gale, J., Perou, C. M., Sethupathy, P. & Prins, J. F. Pseudogenes transcribed in breast invasive carcinoma show subtype-specific expression and ceRNA potential. *BMC Genomics*. **16**, 113 (2015).
2. Zhang, Z. D., Frankish, A., Hunt, T., Harrow, J. & Gerstein, M. Identification and analysis of unitary pseudogenes: historic and contemporary gene losses in humans and other primates. *Genome Biol.* **11**, R26 (2010).
3. Jacq, C., Miller, J. R. & Brownlee, G. G. A pseudogene structure in 5S DNA of *Xenopus laevis*. *Cell*. **12**, 109–120 (1977).

4. Tam, O. H. *et al.* Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*. **453**, 534–538 (2008).
5. Watanabe, T. *et al.* Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*. **453**, 539–543 (2008).
6. Bier, A. *et al.* Connexin43 pseudogene in breast cancer cells offers a novel therapeutic target. *Mol Cancer Ther.* **8**, 786–793 (2009).
7. Chiefari, E. *et al.* Pseudogene-mediated posttranscriptional silencing of HMGA1 can result in insulin resistance and type 2 diabetes. *Nat Commun.* **1**, 40 (2010).
8. 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 (2011).
9. 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 (2010).
10. Johnsson, P. *et al.* A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nat Struct Mol Biol.* **20**, 440–446 (2013).
11. Polisenio, L. *et al.* A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*. **465**, 1033–1038 (2010).
12. Arvey, A., Larsson, E., Sander, C., Leslie, C. S. & Marks, D. S. Target mRNA abundance dilutes microRNA and siRNA activity. *Mol Syst Biol.* **6**, 363 (2010).
13. Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat Rev Cancer.* **7**, 899–910 (2007).
14. Esposito, F., Tornincasa, M., Chieffi, P., De Martino, I., Pierantoni, G. M. & Fusco, A. High-Mobility Group A1 Proteins Regulate p53-Mediated Transcription of *Bcl-2* Gene. *Cancer Res.* **70**, 5379–88 (2010).
15. Berlingieri, M. T. *et al.* Inhibition of HMGI-C protein synthesis suppresses retrovirally induced neoplastic transformation of rat thyroid cells. *Mol Cell Biol.* **15**, 1545–1553 (1995).
16. Berlingieri, M. T., Pierantoni, G. M., Giancotti, V., Santoro, M. & Fusco, A. Thyroid cell transformation requires the expression of the HMGI proteins. *Oncogene*. **21**, 2971–2980 (2002).
17. Esposito, F. *et al.* HMGA1 pseudogenes as candidate proto-oncogenic competitive endogenous RNAs. *Oncotarget.* **5**, 8341–8354 (2014).
18. Esposito, F., De Martino, M., Forzati, F. & Fusco, A. HMGA1-pseudogene overexpression contributes to cancer progression. *Cell Cycle.* **13**, 3636–3639 (2014).
19. Esposito, F. *et al.* A HMGA1-pseudogene expression is induced in human pituitary tumors. *Cell Cycle.* **14**, 1471–1475 (2015).
20. De Martino, M., Forzati, F., Arra, C., Fusco, A. & Esposito, F. HMGA1-pseudogenes and cancer. *Oncotarget.* [Epub ahead of print] (2016).
21. Gabory, A., Jammes, H. & Dandolo, L. The H19 locus: role of an imprinted non-coding RNA in growth and development. *Bioessays.* **32**, 473–480 (2010).
22. Imad, J. *et al.* The H19 Non-Coding RNA Is Essential for Human Tumor Growth. *PLoS ONE.* **2**, e845 (2007).
23. Liu, C. *et al.* H19-derived miR-675 contributes to bladder cancer cell proliferation by regulating p53 activation. *Tumour Biol.* **37**, 263–270 (2015).
24. Zhou, X., Yin, C., Dang, Y., Ye, F. & Zhang, G. Identification of the long non-coding RNA H19 in plasma as a novel biomarker for diagnosis of gastric cancer. *Sci Rep.* **5**, 11516 (2015).
25. Chen, J. S., Zhang, X. Q., Lv, J. M., Li, Y. & Liu, T. X. H19 serves as a diagnostic biomarker and up-regulation of H19 expression contributes to poor prognosis in patients with gastric cancer. *Neoplasma.* **63**, 223–230 (2016).
26. Liang, W. C. *et al.* The lncRNA H19 promotes epithelial to mesenchymal transition by functioning as miRNA sponges in colorectal cancer. *Oncotarget.* **6**, 22513–22525 (2015).
27. Foulstone, E. *et al.* Insulin-like growth factor ligands, receptors, and binding proteins in cancer. *J Pathol.* **205**, 145–153 (2005).
28. Nielsen, H. M. *et al.* Copy number variations alter methylation and parallel IGF2 overexpression in adrenal tumors. *Endocr Relat Cancer.* **22**, 953–967 (2015).
29. Brouwer-Visser, J. & Huang, G. S. IGF2 signaling and regulation in cancer. *Cytokine Growth Factor Rev.* **26**, 371–377 (2015).
30. Qiao, J. *et al.* Stroma derived COL6A3 is a potential prognosis marker of colorectal carcinoma revealed by quantitative proteomics. *Oncotarget.* **6**, 29929–29946 (2015).
31. Xie, X., Liu, X., Zhang, Q. & Yu, J. Overexpression of collagen VI $\alpha 3$ in gastric cancer. *Oncol Lett.* **7**, 1537–1543 (2014).
32. Li, L. T., Jiang, G., Chen, Q. & Zheng, J. N. Ki67 is a promising molecular target in the diagnosis of cancer. *Mol Med Rep.* **11**, 1566–1572 (2015).
33. Zhou, J. D. *et al.* GPX3 hypermethylation serves as an independent prognostic biomarker in non-M3 acute myeloid leukemia. *Am J Cancer Res.* **5**, 2047–2055 (2015).
34. Qi, X. *et al.* Clinical significance and therapeutic value of glutathione peroxidase 3 (GPx3) in hepatocellular carcinoma. *Oncotarget.* **5**, 11103–11120 (2014).
35. Shah, R. *et al.* The prolyl 3-hydroxylases P3H2 and P3H3 are novel targets for epigenetic silencing in breast cancer. *Br J Cancer.* **100**, 1687–1696 (2009).
36. Shetty, P. J. *et al.* Regulation of IGF2 transcript and protein expression by altered methylation in breast cancer. *J Cancer Res Clin Oncol.* **137**, 339–345 (2011).
37. Sun, H. *et al.* H19 lncRNA mediates 17 β -estradiol-induced cell proliferation in MCF-7 breast cancer cells. *Oncol Rep.* **33**, 3045–3052 (2015).
38. Lieberman, J. *et al.* Noncoding RNAs and cancer. *Cell.* **153**, 9–10 (2013).
39. Zmorzyński, S., Świdarska-Kołac, G., Koczkodaj, D. & Filip, A. A. Significance of Polymorphisms and Expression of Enzyme-Encoding Genes Related to Glutathione in Hematopoietic Cancers and Solid Tumors. *Biomed Res Int.* **2015**, 853573 (2015).
40. Wang, J. *et al.* LEPREL1 Expression in Human Hepatocellular Carcinoma and Its Suppressor Role on Cell Proliferation. *Gastroenterol Res Pract.* **2013**, 109759 (2013).
41. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research.* **40**, 4288–4297 (2012).
42. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* **25**, 402–408 (2001).
43. Esposito, F. *et al.* Aurora B expression in post-puberal testicular germ cell tumours. *J Cell Physiol.* **221**, 435–439 (2009).
44. Esposito, F. *et al.* The high-mobility group A1-estrogen receptor β nuclear interaction is impaired in human testicular seminomas. *J Cell Physiol.* **227**, 3749–3755 (2012).

Acknowledgements

This study has been supported by grants from: PNR-CNR Aging Program 2012–2014, POR Campania FSE 2007–2013 (CREMe), CNR Flagship Projects (Epigenomics-EPIGEN), PON 01-02782 (Nuove strategie nanotecnologiche per la messa a punto di farmaci e presidi diagnostici diretti verso cellule cancerose circolanti), Associazione Italiana per la Ricerca sul Cancro (AIRC IG 11477).

Author Contributions

F.E. and A.F. conceived and designed the study. M.D.M. with help from F.E., F.F., M.M. and S.P. performed all experiments, and F.E., M.D.M., C.A., L.T. and A.F. analyzed most data. F.E. and A.F. wrote the manuscript with contributions from all authors.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: De Martino, M. *et al.* *HMGA1P7*-pseudogene regulates *H19* and *Igf2* expression by a competitive endogenous RNA mechanism. *Sci. Rep.* **6**, 37622; doi: 10.1038/srep37622 (2016).

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



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016