Amino Acids

Involvement of non-coding RNAs and transcription factors in the induction of Transglutaminase isoforms by ATRA

--Manuscript Draft--

Manuscript Number:	AMAC-D-19-00182R1
Full Title:	Involvement of non-coding RNAs and transcription factors in the induction of Transglutaminase isoforms by ATRA
Article Type:	Original Article
Corresponding Author:	Nicoletta Bianchi, Ph.D. Universita degli Studi di Ferrara Dipartimento di Scienze Biomediche e Chirurgico Specialistiche ferrara, ferrara ITALY
Corresponding Author E-Mail:	nicoletta.bianchi@unife.it
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	Universita degli Studi di Ferrara Dipartimento di Scienze Biomediche e Chirurgico Specialistiche
Corresponding Author's Secondary Institution:	
First Author:	Ornella Franzese
First Author Secondary Information:	
Order of Authors:	Ornella Franzese
	Linda Minotti
	Gianluca Aguiari
	Fabio Corrà
	Carlo Cervellati
	Carlo Ferrari
	Stefano Volinia
	Carlo M Bergamini
	Nicoletta Bianchi, Ph.D.
Order of Authors Secondary Information:	
Funding Information:	Università degli Studi di Ferrara (FIR2017 (Unife)) — Prof. Carlo M Bergamini
	Università degli Studi di Ferrara (FAR2018) — Prof. Carlo M Bergamini
	Associazione Italiana per la Ricerca sul Cancro (AIRC (IG 17063)) — Prof. Stefano Volinia
Abstract:	The multifunctional protein Transglutaminase 2 is associated with cancer epithelial-mesenchymal transition, invasiveness, stemness and drugs resistance. Several variant isoforms and non-coding RNAs are present in cancer and we have now explored the expression of these transcripts of the TGM2 gene in cancer cell lines after induction with all-trans retinoic acid demonstrating the expression of truncated variants along with two long non-coding RNAs, one coded from the first intron and the Last Exon Variant constituted by a sequence corresponding to the last 3 exons and the 3'UTR. Analysis of ChIP-seq data from ENCODE project highlighted factors interacting with intronic sequences which could interfere with the progression of RNApol II at checkpoints during the elongation process. Using RNA immunoprecipitation we found some relevant transcription factors bound in an ATRA-dependent way, notably GATA3

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	mainly enriched to Last Exon Variant non-coding RNA. We also underline the involvement of NMD in the regulation of the ratio among these transcripts, as the prevalent recovering of Last Exon Variant to phUPF1-complexes, with decrease of the binding towards other selective targets. This study contributes to identify molecular mechanisms regulating the ratio among the variants and improves the knowledge about regulatory roles of the non-coding RNAs of the TGM2 gene.
Response to Reviewers:	Ref.: Ms. No. AMAC-D-19-00182 Involvement of non-coding RNAs and transcription factors in the induction of Transglutaminase isoforms by ATRA Ferrara July 19th, 2019 Dear Prof. Beninati, referring to the paper above (AMAC-D-19-00182) we submitted previously for publication on Amino Acids, please find enclosed an amended version we have modified according to the suggestions of the reviewers. Notably we have substituted the proposed summary to the original one; employed the term Transglutaminase type 2 throughout the text, modified the expression of the cell numbers in culture plates and of volumes, and avoided references to “we” and “our”, as suggested. With best regards Carlo Bergamini and Nicoletta Bianchi

Ref.: Ms. No. AMAC-D-19-00182

Involvement of non-coding RNAs and transcription factors in the induction of Transglutaminase isoforms by ATRA

Ferrara July 19th, 2019

Dear Prof. Beninati,

referring to the paper above (AMAC-D-19-00182) we submitted previously for publication on Amino Acids, please find enclosed an amended version we have modified according to the suggestions of the reviewers. Notably we have substituted the proposed summary to the original one; employed the term Transglutaminase type 2 throughout the text, modified the expression of the cell numbers in culture plates and of volumes, and avoided references to “we” and “our”, as suggested.

With best regards

Carlo Bergamini and Nicoletta Bianchi

Manuscript red marked version

Click here to access/download;Manuscript;Manuscript Franzese O. et al 19-07-2019 marked version.pdf

Click here to view linked References

Involvement of non-coding RNAs and transcription factors in the induction of Transglutaminase isoforms by ATRA.

Ornella Franzese[1], Linda Minotti[2], Gianluca Aguiari[3], Fabio Corrà[2], Carlo Cervellati[3], Carlo Ferrari[4], Stefano Volinia[2], Carlo M. Bergamini[3,*], Nicoletta Bianchi[3,*]

[1] Department of Systems Medicine, School of Medicine, University of Rome Tor Vergata, Rome, Italy
[2] Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, Ferrara, Italy
[3] Department of Biomedical Sciences and Specialist Surgery, University of Ferrara, Ferrara, Italy
[4] Department of Clinical and Molecular Sciences, Marche Polytechnic University, Ancona, Italy

*** Correspondence**:
Nicoletta Bianchi and Carlo M. Bergamini, Department of Biomedical Sciences and Specialist Surgery, University of Ferrara, Via Luigi Borsari, 46 - 44121 Ferrara, Italy
nicoletta.bianchi@unife.it; Tel.: +39-0532-455447;
bgc@unife.it; Tel.: +39-0532-455428

ORCID linked to this article:

Ornella Franzese: https://orcid.org/0000-0002-6396-4603
Linda Minotti: https://orcid.org/0000-0003-2116-8411
Gianluca Aguiari: https://orcid.org/0000-0002-0007-0805
Fabio Corrà: https://orcid.org/0000-0002-6803-0797
Carlo Cervellati: https://orcid.org/0000-0003-4777-6300
Carlo Ferrari: https://orcid.org/0000-0001-5112-7448
Stefano Volinia: https://orcid.org/0000-0003-0910-3893
Carlo M. Bergamini: https://orcid.org/0000-0002-9430-8625
Nicoletta Bianchi: https://orcid.org/0000-0001-9280-6017

Acknowledgements
The cell lines employed in this study were kind gifts from Professor Paola Secchiero (HL-60 and NB4 cells), Michael U. Martin (SK-N-SH cells) and Ralf Hass (SH-SY5Y cells) respectively. The authors thank Prof. Giovanni Bernacchia for having made available the instrument to perform qPCR reactions. We are grateful to Prof. Cristian Taccioli and Dr. Marco Ricci for their advice and valuable help to address the study of TG2-lncRNA molecules. Finally, we thank Co.Pe.Go. (Soc.Coop.O.P.) for liberal voluntary contribution to the present research.

Handling Editor: Simone Beninati

Abstract

The multifunctional protein Transglutaminase type 2, is associated with cancer epithelial mesenchymal transition, invasiveness, stemness and drugs resistance. Several variant isoforms and non-coding RNAs are present in cancer and this report explored the expression of these transcripts of the *TGM2* gene, after induction, in cancer cell lines, with *all-trans* retinoic acid. The expression of truncated variants along with two long non-coding RNAs, was demonstrated. One of these, coded from the first intron and the Last Exon Variant, constituted by a sequence corresponding to the last 3 exons and the 3'UTR. Analysis of ChIP-seq data, from ENCODE project, highlighted factors interacting with intronic sequences, which could interfere with the progression of RNApol II at checkpoints, during the elongation process. Some relevant transcription factors bound, in an ATRA-dependent way, were found by RNA immunoprecipitation, notably GATA3 mainly enriched to Last Exon Variant non-coding RNA. The involvement of NMD in the regulation of the ratio among these transcripts was observed, as the prevalent recovering of Last Exon Variant to phUPF1-complexes, with decrease of the binding towards other selective targets. This study contributes to identify molecular mechanisms regulating the ratio among the variants and improves the knowledge about regulatory roles of the non-coding RNAs of the *TGM2* gene.

Keywords: Transglutaminase type 2; transcriptional variants; non-coding RNA; retinoic acid; GATA3

ABBREVIATIONS

AAs	amino acids
Abs	antibodies
APL	acute promyelocytic leukemia
AS	antisense oligonucleotide
ATRA	*all-trans* retinoic acid
CTCF	CCCTC-binding factor
FAB	French–American–British
FBS	foetal bovine serum
GEO	Gene Expression Omnibus
HDACs	histone deacetylases
LEV	Last Exon Variant
lncRNA	long non-coding RNA
NI	normalized signal intensity
NMD	nonsense-mediated mRNA decay
PBS	phosphate-buffered saline
phUPF1	phosphorylated ATP-dependent RNA helicase upframeshift 1
RA	retinoic acid
RARα	retinoic acid receptor α
RAREs	Retinoic Acid Responsive Elements
RIP	RNA-immunoprecipitation
RRE	retinoid receptor element
RT-qPCR	reverse transcription and real time quantitative PCR
TGM2	Transglutaminase type 2 gene
TG2	Transglutaminase type 2

Introduction

Transglutaminase type 2 (TG2) has attracted much attention in experimental oncology because of its involvement in the control of cell cycle and tumour cell stemness along with its higher expression in secondary than in primary tumours. TG2 is a bifunctional enzyme involved in post-translational protein modification (transamidation) and in membrane signalling, with opposite effects on cell survival and death (Tee et al. 2010, Beninati et al. 2017) in relation to occurrence of apoptosis (Fésüs and Szondy 2005), Epithelial-Mesenchymal Transition, expansion of cancer stem cells and resistance to chemotherapy (Tee et al. 2010, Kerr et al. 2017).

In addition to the full-length enzyme, normal and pathological tissues express at variable levels transcriptional variants of the protein that are apparently tailored to carry out the transamidation and the signalling activities, which are performed by the N- and the C-terminal regions respectively. The forms described in literature include the full-length, the two *tTGv1* and *tTGv2* variants with altered structure at the C-terminus due to alternative splicing (Phatak et al. 2013) and the shorter deleted isoform TGH. All these isoforms are insensitive to regulation by GTP (Lai and Greenberg 2013). A much shorter form, TGH2, full-length TG2 promotes cell survival, while the altered isoforms lacking the 3'-terminal portion with de-regulated enzyme catalytic activity are associated with cell death, it follows that the different forms carry out different cellular functions (Király et al. 2011).

Gene dysregulation often leads to the accumulation of altered variants, which are often considered as tumor markers. Similarly, the treatments with anti-cancer drugs modulate the levels of protein isoforms, as occurs for Transglutaminase type 2 gene (*TGM2*) among other genes. Therefore, the study of their modification and the use of these experimental models is critical for basic studies aimed at identifying factors involved in the transcriptional control and providing new tools useful for the identification of useful clinical perspectives and the development of new molecular therapies.

The expression of the *TGM2* is induced by retinoids *in vitro* and *in vivo* in patients affected by leukemia (Benetti et al. 1996) and neuroblastoma, who are successfully treated with retinoic acid (RA) with a rapid increase of the protein (Tee et al. 2010, Masetti et al. 2012). This regulated expression is under the control of multiple Retinoid Response Elements (RRE) (Daniel et al. 2014).

So far, investigations concerning the effects of RA have been defective due to the expression of the different TG2 isoforms, a relevant issue in the light of their functional differences. The recognized association between *TGM2* expression and tumor status prompted us to investigate the differential expression of the enzyme in selected acute promyelocytic leukemia (APL) and neural tumor cell lines, in which *all-trans* retinoic acid (ATRA) acts as differentiating agent. The APL cell lines employed in the present investigation were the HL-60 cells, lineage-bipotent myeloblasts (French–American–British, FAB M2) which can be induced to differentiation into granulocytes by RA (Jensen et al. 2015), and the leukemic NB4 cells belonging to FAB M3. Both cell lines are sensitive to ATRA but throughout the engagement of different receptors and Retinoic Acid Responsive Elements (RAREs). The neuroblastoma lines employed were the SK-N-SH cells, isolated from a bone marrow biopsy of a patient with neuroblastoma, and its more aggressive subclone SH-SY5Y. The parental SK-N-SH cells display two morphologically distinct neuroblast-like and epithelial-like phenotypes, the former expressing enzymes typical of the catecholaminergic neurons that are instead absent in the latter. SK-N-SH cells are forced to differentiate towards a neuronal phenotype with extensive neurite outgrowth by treatment with ATRA, which is actually used in the adjuvant therapy of residual or relapsed neuroblastoma. ATRA affects differently *TGM2* expression in SH-SY5Y cells (Lee et al. 2012) and neuron-like phenotype is induced only in low serum condition, which may also promote cell survival through activation of the phosphatidylinositol 3-kinase/Akt signalling pathway and increase of anti-apoptotic Bcl-2 (Kovalevich and Langford 2013).

The expression of mRNAs coding for the full-length and for the variant isoforms, as well as that of the regulatory long non-coding RNAs (lncRNAs), including the TG2-lncRNA reported by Minotti et al. (2018), was analyzed in these cellular models, along with their mechanisms of action upon treatment with RA (Jensen et al. 2015, Redner et al. 2018). The present study highlighted also the expression of a novel RNA, which we named Last Exon Variant (LEV), reported in "NEDO human cDNA sequencing project" (GenBank: AK126508.1) on the Genome Browser database https://genome.ucsc.edu/). LEV is constituted by the sequences of the last 3 exons of *TGM2* gene including the complete 3'UTR region (Bianchi et al. 2018). To identify the factors engaged in the differential expression of the various *TGM2* transcripts under ATRA modulation, we have taken advantage of ChIP-seq data from ENCODE (Consortium on Genome Browser Human Genome hg19) and of results from original experiments performed in our laboratory, including RNA immunoprecipitation (RIP) and real time quantitative PCR (RT-qPCR) assays. We have also explored approaches to target selectively the expression of the lncRNAs by means of antisense molecules in tumor cell lines treated with ATRA.

Transcription variants differing in their altered 3' last exons likely represent a strategy set in motion by cells to escape the effects of different miRNAs (Sandberg et al. 2008). In addition, several recent studies indicate that lncRNAs drive many important cancer phenotypes through their interactions either with DNA, proteins or RNA, regulating gene expression at the transcriptional, post-transcriptional, and epigenetic levels. Moreover, together with miRNAs, these molecules can act as competing endogenous RNAs capable of controlling mRNA targets, thus playing an important role in cancer. Indeed, lncRNAs can be sequestered from miRNAs of other specific targets, although the mechanisms underlying these interactions have not been elucidated yet (Zhang et al. 2018).

Methods

Cell line cultures and treatments

All materials for cell culture were purchased from Sigma-Aldrich.

Human leukemia HL-60 and NB4 cells (a gift from Prof. Paola Secchiero, University of Ferrara, Italy) were cultured at 37°C in a 5% CO_2 humidified atmosphere in RPMI 1640 medium, supplemented with 10% foetal bovine serum (FBS), 50 U/mL penicillin and 50 µg/mL streptomycin (Minotti et al. 2018). Cells were seeded at $12x10^4$/mL and induced to differentiation by treatment with 1 and 3 µM ATRA for different time intervals (from 16 to 72h).

Neuroblastoma SK-N-SH cells (a gift from Prof. Michael U. Martin, Justus-Liebig-University Giessen, Germany) were cultured at 37°C in a 5% CO_2 humidified atmosphere in EMEM (EBSS), supplemented with 15% FBS, 2 mM glutamine, 50 U/mL penicillin and 50 µg/mL streptomycin. SH-SY5Y cells (provided by Prof. Ralf Hass, Hannover Medical School, Germany) were cultured using a mixture 1:1 of Eagle's Minimum Essential Medium and F12 Medium with 10% FBS. For sub-confluent cultures, cells (at 70-80% of confluence) were splited using 0.25% Trypsin and incubated for 5 min at 37°C and 5% CO_2, then seeded at $3x10^3$-10^4 cells/mL. Differentiation was induced by treatment with 3 and 10 µM ATRA for different time intervals depending on the growth rate of the cell lines (16 and 48h). Pellets were collected during exponential growth and total RNA was extracted from $5x10^6$ cells.

DNA antisense oligonucleotides (AS) to target the transcripts of TG2-lncRNA (5'-TAGGGATTCAGCTCCCACCGGGTC-3'), LEV (5'-CTCATGATGCAGAATCACC-3'), and a scramble oligonucleotide (5'-AATTCTCCGAACGTGTCACGT-3') were used at the final concentration of 1.5 µM to modulate the expression of 10^6 cells needed to perform enough reactions to analyse the different *TGM2* transcripts, adjusting as appropriated the protocol of siPORT™ NeoFX™ Transfection Agent (ThermoFisher Scientific, Life Technologies, Applied Biosystems, Italy).

Reverse transcription and real-time quantitative PCR reactions to quantify *TGM2* transcripts

Reverse transcription reactions were carried out using the TaqMan® Reverse Transcription Reagents kit using the protocol from ThermoFisher Scientific. PCR reactions were performed using PowerUp SYBR Green Master Mix (ThermoFisher Scientific, Life Technologies, Applied Biosystems, Italy). Primers to amplify the targets are reported in Table 1. The couple of primers F2TGH/FR2TGH and Fshort3ex/Rshort3ex were designed with the Primer Blast tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The amplification was performed using Thermal cycler CFX96 Touch™ Real-Time PCR Detection System (Bio-rad Laboratories S.r.l., Segrate, Milan, Italy).

Fold increase was determined by comparing the cycle threshold (CT) values of the amplified targets with that of the reference HPRT1 gene, which is reported as one of those with better stability (de Kok et al. 2005). We have calculated the negative exponent in formula $2^{-\Delta\Delta CT}$ ($\Delta\Delta CT$) to quantify fold changes in gene expression of treated samples with respect to control cells, as suitably specified in the legends (Schmittgen and Livak 2008). The t test was performed and p values were calculated as detailed in the legends.

Analysis of *TGM2* gene expression array using datasets of NB4 cells from GEO database

Analysis of *TGM2* gene expression was performed (as detailed in Table 2) on data from public Gene Expression Omnibus (GEO) considering 2 datasets of total RNA analyzed with HumanExon1_0ST Affymetrix on NB4 cell line (Luesink et al. 2009), which is composed of specific probes distributed along the whole genome to evaluate with high performance the expression of altered transcripts including intronic sequences. The probes considered for the *TGM2* gene are reported in Minotti et al. (2018). For each culture we compared signal Normalized Intensity (NI) from untreated and treated cells with ATRA.

GEO dataset and Chip-seq analysis

Chip-seq analysis (Table 2) was derived from datasets through Genome Browser tools (HudsonAlpha/Caltech ENCODE group) obtained from neuroblastoma SK-N-SH cells untreated or treated with 6 µM ATRA for 48h. The tracks of ChIP-seq represent a comprehensive set of human transcription factor binding sites based on experiments generated by a group in the ENCODE Project Consortium 2011 and 2012, which developed Integrative analysis of experiments performed in duplicate and scored against two controls displaying the regions of enrichment as peak calls.

Preparation of nuclear extracts and RNA immunoprecipitation

RIP has been developed to identify RNA species that bind to a protein of interest by a protocol including: (1) preparation of protein lysate from target cells, (2) immunoprecipitation of the protein with specific antibodies (Ab)s and extraction of protein-bound RNAs. We followed the protocol reported by Brun (2012) as modified according to Comandini et al. (2013). Native nuclei were obtained from $15x10^6$ cells washed three times in Phosphate-buffered saline (PBS), and resuspended in a solution of 2 mL PBS plus 2 mL nuclear isolation Buffer (1.28 M Sucrose, 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 4% (v/v) Triton X-100) and 6 mL H_2O treated with diethylpyrocarbonate. After 20 min incubation on ice native nuclei were collected by centrifugation at 2500 g for 15 min at 4°C, suspended in 200 µl freshly prepared ice-cold RIP buffer (25 mM Tris pH 7.5,150 mM KCl, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, plus HALT protease inhibitor cocktail 1x and RNase Inhibitor 20 U/mL) (Thermo Fisher Scientific, Italy) and sheared on ice by passage through a 27 gauge needle with 15–20 strokes. Nuclear membranes and debris were removed by centrifugation at 16000 g for 10 min at 4°C and supernatants of nuclear lysates were saved. Each extract was incubated at 4°C on gentle rotation overnight with 10 µg of Abs against RNApol II (Pierce ma146093, Anti-RNA Polymerase II CTD monoclonal mouse), Max (ThermoFisher Scientific pa529745, MAX Polyclonal Antibody), GATA3 (Abcam, [EPR16651]-ChIP grade ab199428), RXRα (Abcam, EPR7106, ab125001) and upframeshift 1 (phUPF1) (Ser1127) (EMD Millipore Corporation). Twenty µl of protein A/G PLUS conjugated agarose beads, pre-washed three times with 500 µl ice-cold RIP buffer, were incubated in rotation for 3 hours at 4°C with specific Ab. The supernatant was removed after centrifugation at 4000 g for 2 min followed by 3 washing steps with ice cold 500 µl RIP buffer for the extraction of the bound RNA. The incubation with Abs against RNApol II, was preceded by one step of 2h at 4°C in the presence of 1 µg of competitor oligonucleotide to reduce the background.

Purification of transcripts bound to immunoprecipitated proteins and data analysis

After IP reactions RNAs were isolated from the protein-A-G conjugated agarose complexes the beads by extraction with 1 mL TRIzol®, as described in the manufacturer's instructions. RNAs were precipitates as reported by Brun (2012) and suspended in 10 µl RNase free H_2O.

One microgram of total RNA was employed as starting material in 20 µl Reverse Transcription reaction. We employed TaqMan® Reverse Transcription Reagents kit, as above reported. RNA contents have been normalized using HPRT1 for setting of samples. Quantitative real-time PCRs were carried out in order to assay target RNAs using PowerUp SYBR Green Master Mix in 10 µl volume, according to the manufacturer's instructions.

CT values considered in our analysis were derived from arithmetic mean of 3 independent experiments. To analyse data of RIP experiments Fold Enrichment of RIP reactions was calculated by plugging the average CT values to respective Input RNA fractions (derived from amounts of samples before the RIP), also reported in specific user guide (https://www.sigmaaldrich.com). The quantification of RNA targets associated to the Ab complexes was performed as follows: i) normalization of each RIP RNA fractions' CT value to the Input RNA fraction CT value for the same qPCR Assay (ΔCT) considering also dilution factor: **ΔCT [normalized RIP]** = CT [RIP] – (CT [Input] – Log2 Input Dilution Factor), where Input fraction is 1% (log2 of 100); ii) the binding of Ab to protein A/G PLUS conjugated agarose beads (after and pre-RIP) was used to adjust the data for Non-Specific (NS) background: **ΔCT [normalized NS]**; iii) finally we have calculated: **ΔΔCT [RIP/NS]** = ΔCT [normalized RIP] – ΔCT [normalized NS]; iv) for quantification of the **Fold Enrichment**= $2^{-\Delta\Delta CT\ [RIP/NS]}$.

Western blotting

Cells were collected by centrifugation at 800xg for 10 min and washed twice in PBS buffer containing a cocktail of protease inhibitors (Sigma-Aldrich). Cells were lysed for 30 min by using a single detergent lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 1% v/v Triton X-100) with protease inhibitors. Nuclear extracts were prepared by removing cell debris by centrifugation at 10000 g for 10 min.

Twenty five µg of total protein or nuclear extract were solubilized in SDS loading buffer (60 mM Tris-HCl, pH 6.8, 1.8% SDS, 6% (v/v) glycerol, 1% β-mercaptoethanol, 0.002 (w/v) bromophenol blue) and analyzed by 12.5% SDS-PAGE. Proteins were transferred on nitrocellulose filters (EuroClone) in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol, pH 8.3) for 2h by using an elettroblotting chamber (Bio-Rad). Nitrocellulose filters, blocked in a solution containing 5% non-fat dry milk in PBS buffer 0.05% Tween 20®, were treated overnight at 4°C with specific Abs. The Anti-TG2 mouse monoclonal Abs A036 and A037 (Zedira GmbH, Darmstadt, Germany), recognize respectively the epitope K^{527}YLLNLNL^{534} of the beta-Barrel-1 Domain, present in both TG2 and TGH isoforms, and the epitope V^{604}AEVSLQNP^{612} present in beta-Barrel-2-Domain, but not in TGH isoform. For the analysis of transcription factors we employed Abs against Max (ThermoFisher Scientific pa529745, MAX Polyclonal Ab), GATA3 (Abcam, [EPR16651]-ChIP grade ab199428), RXRα (Abcam, EPR7106, ab125001). Protein loads were compared using an Ab against β-actin. After hybridization, the nitrocellulose membranes were washed three times in PBS buffer with 0.05% Tween20® and incubated with the HRP-conjugated secondary Abs. Further, three washes in PBS 0.05% with Tween 20®

were performed and the protein bands were visualised through autoradiography by using the enhanced chemiluminescence system (EuroClone) (de Stephanis et al. 2018).

Results

Expression of transcriptional variants of *TGM2* upon ATRA treatment

The induction of *TGM2* expression by ATRA by RT-qPCR was investigated in HL-60 cells. The transcripts for canonical full-length TG2, truncated TGH2 and TGH isoforms, along with alternative spliced variants *tTGv1* and *tTGv2* (Phatak et al. 2013), TG2-lncRNA and LEV were quantified by RT-qPCR using the primers reported in Table 1. Incubation with 3 µM ATRA was performed for time intervals ranging between 16 and 72h (Breitman et al. 1980). A drug-induced transcriptional up-regulation of *TGM2* variants as compared with untreated cells was observed at the same time of culture. However, each variant and regulatory RNA showed different profiles as described in Fig. 1, showing different levels of expression appreciable at the specific time points. Therefore, LEV was induced more intensively than TG2-lncRNA by ATRA before 36h of culture with slight reversal of the trend after this time. After 48h of induction the highest TG2-lncRNA levels declined, whereas full-length TG2 further increased. In contrast, the highest levels of TGH2 and TGH variants were reached at 24h of treatment and decreased over extended time intervals, while the strong initial increase of the spliced *tTGv1* and *tTGv2* variants showed fluctuating trends.

Experiments were carried out also on leukemic NB4 cells, in which *TGM2* is among the highest induced genes with appreciable levels of the protein also in untreated cells. The gene expression profiling of *TGM2* from the GSE18397 total RNA dataset of GEO (Table 2) was performed by GeneChip Human Exon 1.0 ST Array using probes recognizing both exonic and intronic sequences belonging to the transcriptional variants allow to discern differences across single variants based on the use of specific intronic sets of probes (Minotti et al. 2018). NB4 cells cultured for 48, 72 and 96h in the absence or presence of 1 µM ATRA displayed particularly high levels of both the TG2-lncRNA (probe 3905186) (Minotti et al. 2018) and the LEV transcript (probe 3905153) when cultured in the presence of the drug. ATRA treatment increased the overall expression of *TGM2* (Fig. 2), including the transcripts derived from altered splicing and polyadenylation (Benedetti et al. 1996), as shown by comparing the signals of the probe 3905145 (that represents the average intensity of signals from probes hybridizing all exons of full-length TG2 mRNA) and those recognising the intronic sequences present in the altered variants (3905171 and 3905172, selective for TGH2, 3905161 for TGH, 3905147 and 3905148 for *tTGv1* and *tTGv2* mRNAs). All these signals increased during treatment with ATRA, in contrast to the 3905144 probe specific for an intronic sequence that remained substantially untranscribed. The probe 3905153, specific for LEV, was increased after 48h of ATRA treatment, but decreased during persistent exposure to the drug (at 72h) before the decline of TG2-lncRNA observed at 96h (Fig. 2). Therefore, ATRA stimulates the expression of all *TGM2* transcripts in APL cells, despite the cell line differences possibly arising because the gene is silenced in untreated HL-60 but not in control NB4 cells.

The analysis was further extended to brain tumor cell lines to investigate whether ATRA affects the *TGM2* gene also in this model, by considering SK-N-SH cells. We tested the effects of 3 and 10 µM ATRA at 16h and 48h incubation times. The results reported in Table 3 show that the expression of all *TGM2* transcripts (including TG2-lncRNA and LEV) increased with respect to the untreated control at early culture time (16h) while 3 or 10 µM ATRA reduced the expression especially of the alternatively spliced *tTGv1* and *tTGv2* transcripts with the protracted exposure (48h). Only the expression of TG2-lncRNA was not reduced after 48h of treatment at high concentration (10 µM). Furthermore, values reported in Table 3 showed that the decline in expression is more pronounced in the 3' portion of the gene. SH-SY5Y cells did not reveal a consistent ATRA-mediated induction of *TGM2* (Table 3), probably because the gene was already actively transcribed in untreated SH-SY5Y cells.

GATA3 drives ATRA-dependent expression of LEV, a regulatory ncRNA influencing the levels of the other *TGM2* transcripts

Several studies have analyzed the binding of transcription factors to *TGM2* promoter and other regulatory regions, but only a few reports describe the interactions between different proteins along the gene, including the intronic regions (de Santiago et al. 2017). The involvement of specific transcription factors was detected in the regulation of *TGM2* expression after ATRA stimulation using ChIP-seq information from ENCODE (Table 2). Fig. 3 shows peak calls (regions of enrichment) for the binding of regulatory factors to the *TGM2* gene in addition to RNApol II. The sites of binding are mainly concentrated at the promoter, but some of them are also present at intronic positions critical for the generation of transcription variants. In untreated SK-N-SH cells, ChIP-seq highlights the binding of RXRα complexes to RAREs in the introns 1, 3, 6 and 10 of *TGM2*, where they may potentially interact or interfere with other regulatory factors listed and summarized in Fig. 3. They include activators, such as MAX, and co-repressors recruiting Histone deacetylases (HDACs), along with the insulator CCCTC-binding factor (CTCF), RAD21, or YY1. The pattern of enrichment of these latter factors in ChIP-seq is modified in SK-N-SH cells treated with ATRA (Fig. 3, bottom panel).

It is relevant to call attention to GATA3 because it binds to intron 10 of *TGM2* (downstream the altered polyadenylation site of TGH and really upstream the LEV starting site) with higher affinity than to other regions (Fig. 3, top panel).

Since GATA3 and other transcription factors are associated with putative regions for modulation of *TGM2* expression in neuronal cells, and considering their involvement in the expression of the *TGM2* locus, the association of these proteins with mRNAs and regulatory RNAs has been investigated by RNA immunoprecipitation in native nuclear extracts of NB4 cells (Fig. 4A). RIP experiment performed using an Ab against RNApol II, showed that the treatment with 1 µM ATRA led to an enrichment of complexes associated with LEV in particular, and, in contrast, to a decreased of those with full-length TG2 and TGH transcripts. RIP performed using Abs against the Max factor allowed only the recovery of the regulatory TG2-lncRNA and LEV. Finally, the use of Abs against RXRα enriched the complexes with TG2-lncRNA, TGH and LEV transcripts, in contrast with the decline of full-length TG2 associated to GATA3, which appeared to bind mainly LEV after ATRA administration (Fig. 4A). It must be underlined that the TGH isoform (Fig. 4B) and LEV ncRNA were observed only upon ATRA treatment. Moreover, while RXRα was negatively modulated by ATRA, as showed by Western blots of Fig. 4B, GATA3 was equally detected in nuclear extracts of treated and untreated cells.

Despite the possibility that LEV could derive from the processing of primary unspliced RNA, the ENCODE data show that the region upstream LEV start site includes sequences recognized by RNApol II and other factors. In the sequence upstream of LEV a H3K27Ac mark is present with feature of regions undergoing high transcriptional activity (as shown in Fig. 3), as elsewhere within *TGM2*, or close to the region of TG2-lncRNA (Minotti et al. 2018, Creyghton et al. 2010). RT-qPCR amplification using oligo-dT primers proved that the LEV transcript is polyadenylated (data not shown). In addition, no short peptides of 150 AAs corresponding to possible LEV transcripts was detected, using an Ab against the epitope V^{604}AEVSLQNP^{612} of the C-terminus of the protein (Ab A037).

These observations demonstrate specific transcription factors binding to ncRNAs of *TGM2*. Since ncRNAs and altered gene transcripts are susceptible to degradation by nonsense-mediated mRNA decay (NMD), RIP assays were performed using an against the phosphorylated ATP-dependent RNA helicase phUPF1, to distinguish cellular targets of NMD processing. In this context the regulatory ncRNAs of *TGM2* could contribute to the accumulation of altered variants, by interfering with their degradation systems. In the presence of ATRA, the levels of full-length TG2 and TGH declined in the complexes with phUPF1, whereas LEV was transcribed and appeared to stringly bind phUPF1 (Fig. 4A).

Since LEV is mainly constituted by the 3'UTR of TGM2 and hypothesizing its protective role against the degradation of the main transcripts, the effects of a AS-LEV adjacent to the start site within intron 10 were studied in HL-60 cells in the presence of ATRA. In these cells the transcription factors involved are expressed at different levels than in NB4 cells, as shown by the western blot (Fig. 5A) and this is mirrored by a different time course of transcripts accumulation, notably the full-length TG2 (Fig. 5B). The treatment of HL-60 cells with AS-LEV or scrambled AS-neg oligonucleotide for 24h in the presence of 1 µM ATRA (antisense oligonucleotides were added 8h after ATRA addition), induced a significant decrease of TG2-lncRNA correlated to LEV transcript after 24h (Fig. 5C).

Predicted structure and regulatory properties of TG2-lncRNA

The sequence corresponding to the TG2-lncRNA was analyzed using Vienna (http://rna.tbi.univie.ac.at/), which has predicted the secondary RNA structure as reported in Supplementary Fig. 1. RNA fold was defined using two different algorithms, one considering Minor Free-Energy, the other based on Centroid approach. An extended region (in red, Supplementary Fig. 1) showed high probability of pairing with each analysis method. Additional analysis in Genome Browser revealed that this transcript is constituted by 2 exons, the first containing an Inverted Repeat IR0 motif likely binding retinoic acid receptor α (RARα) and RARγ (Jolma et al. 2013). This consensus sequence can be recognized also by PML-RARα and its heterodimers having low restriction sequence selectivity (Hosoda et al. 2015). The existence of rs2076380 SNP in this site is also reported (Supplementary Table 1), witnessing a potential regulative role (Jolma et al. 2013). The TG2-lncRNA sequence contains additional elements for binding to transcription factors, which are modified after ATRA treatment, as evidenced by enriched ChIP-sequencing on SK-N-SH cells reported in Fig. 3. The probability of interaction between the TG2-lncRNA and these proteins has been investigated using RPISeq web server (http://pridb.gdcb.iastate.edu/RPISeq/) selecting two prediction methods, the Random Forest and the Support Vector Machine classifier. The data obtained indicated that the probabilities of interaction are respectively 0.85 and 0.6 with GATA3, 0.85 and 0.76 with RXRα, 0.85 and 0.84 with specific RARα, but were also elevated with YY1, 0.80 and 0.85, with CTCF, 0.75 and 0.79, and with RAD21, 0.85 and 0.84 (>0.5 positive threshold). These values are in agreement with our data obtained from RIP experiments, in which TG2-lncRNA associated with RNApol II, RXRα and GATA3 overall after ATRA treatment (Fig. 4), although RARα could be a better target than RXRα.

Notably, 77 nucleotides in the second exon of the TG2-lncRNA (CAAGGTTTATGAGATTTGTTTTCCCCCAAGACC CCCACCCATCTGTCACCCTGACCTGCCCTGGTCCTCTGCTCCA) are identical to another sequence in the intron spacing the two exons, close to the splicing site, and characterized by consensus sequence for TFIID, RXRα and YY1.

Other portions of its long sequence predict a role in interacting with nucleic acid. Indeed, it contains SINE motifs (Short Interspersed Nuclear Elements, which include ALU motifs) a Repeat Mirc-family located in the first exon and frequently recognized by YY1 polycomb proteins, acting as transcriptional repressor interacting with histone deacetylases/

acetyltransferases (Usmanova et al. 2008, Smit et al. 1996-2010, RepeatMasker Open-3.0). These regions are disseminated with consensus sequences for both RAR/RXR and other factors capable of interacting with transcriptional factors, some of which are also present on the gene promoter, in the introns along the gene and just upstream of the LEV sequence. Further, a LINE motif (Long Interspersed Nuclear Elements) characterizes a portion of the second exon with feature of L1MC5 (Jurka 2000), which is complementary to intron 10 and possibly affects the starting region of LEV.

With the aim to investigate a potential activity of the TG2-lncRNA in the control of gene expression we used an antisense molecule (AS-lncRNA, see Material and Methods), targeting the TG2-lncRNA in NB4 cells stimulated with ATRA, in order to have a detectable but not too high expression model, as described for HL-60 cells. Under these conditions AS-lncRNA did not reduce the levels of TG2-lncRNA, but promoted up-regulation of full-length TG2 and LEV mRNA after 24h of treatment and restrained further TG2-lncRNA increase after 48h (Fig. 6, as referred to activity of an AS-neg molecule used as negative control).

Discussion

Transglutaminase type 2 is a complex protein with multiple catalytic activities that, under the control of allosteric effectors and cell reducing potential, drive opposite actions on cell death and survival (Griffin et al. 2002). The existence of variant isoforms preferentially displaying one of the activities (protein crosslinkage or membrane signalling) adds further complexity (Lai and Greenberg 2013). One of our recent analysis of gene expression has highlighted intricate aspects also at this level, since untranslated RNAs with regulatory functions can interfere with the above mentioned processes (Bianchi et al. 2018). In the present report, we investigated the expression of *TGM2* gene in cancer cell lines responsive to ATRA in terms of variants, kinetics, regulatory/functional roles of the ncRNAs, and involvement of transcription factors. Full-length TG2, two variants alternatively spliced at the GTP binding region, the TGH truncated form (lacking almost completely the forth domain) and a very short variant lacking domains 3 and 4, beside the two additional regulatory ncRNAs were examined.

Some of the analyzed cell lines, notably HL-60 with silenced *TGM2*, required ATRA to induce all transcripts while others, like NB4, transcribed the gene in the absence of the drug which is essential for the expression of TGH and of the regulatory ncRNAs. In neuroblastoma SK-N-SH cells that also express *TGM2* in the absence of ATRA, the induction was detected early after treatment with a decline at high concentrations or after prolonged exposure. Finally, ATRA did not produce any increase over basal conditions in SH-SY5Y cells, where the enzyme persistence has been associated with tumor invasiveness (Lee et al. 2012).

In the kinetics of changes triggered by ATRA, the induction of *TGM2* is a relatively early event tightly coupled with the synthesis of the TG2-lncRNA, potentially due to the direct effects of retinoid receptors at the level of the promoter and of the regulatory regions in the TG2-lncRNA sequence.

The induction of the altered transcripts occurred despite the decline of the RAR/RXR complexes, indicating that other factors can sustain the response. During protracted observations an appreciable decline of transcription was detected for LEV, of which induction was strongly determined by ATRA, occurring only after treatment and declining first among all *TGM2* transcripts. In this perspective, the switch-off of transcription involves LEV before other transcripts proceeding back from the 3' end of the gene, as it has been reported also for other genes undergoing temporal transcriptional control in the TNFα/NF-kB network (Wada et al. 2009). In SK-N-SH cells, ATRA apparently drives a "structural closure" of the locus inhibiting transcription, leading mostly to accumulation of TG2-lncRNA.

The gene sensitivity to retinoids is under the control of transcription factors with opposite roles of distinct RAR/RXR complexes (Girardi et al. 2019). Indeed, ATRA decreases RXRα but not RARα in HL-60 and in NB4 cells, in which however the PML-RARα fusion protein modulates retinoid effects (Jensen et al. 2015, Mikesch et al. 2010). In contrast, in neuronal SK-N-SH and SH-SY5Y cells, ATRA down-regulate both RARα and RXRα with concomitant up-regulation of RARβ (Girardi et al. 2019, Nagai et al. 2004). Since RXR content is different in the analyzed cell lines, it can be hypothesized that the different patterns of ATRA-mediated *TGM2* induction to the combined effects of regulatory RNAs and transcription factors on the reactivity of RREs. Notably RAREs arise through aggregation of hexameric units 5′-(A/G)G(G/T)TCA-3′, assembled in different ways as dimeric palindromes (Inverted or Everted Repeats) to form families of canonical and atypical RAREs, differently recognised by RAR, RXR and PML-RARα (Kamashev et al. 2004, Saeed et al. 2012). RAREs are located in the promoter, in intronic regions along the *TGM2* gene and in the TG2-lncRNA, likely in relation to pausing of transcriptional complexes during elongation (Wada et al. 2009), so that the different behaviours of these complexes might be the keys to understand the different expression of *TGM2* transcripts.

We interpreted our results in the frame of the model of Wada et al. (2009) suggesting that the progression of RNApol II is hindered by interposition of regulatory factors. The responsive motives occur at the promoter level, involving sites of RAR and RXR within transcriptional control sites of the ncRNAs and highly represented also at intronic regions. Thus, the rate of elongation would be slowed down by occupancy of these sites, leading either to a complete stop with production of truncated transcripts or to the selection of alternative sites of splicing and polyadenylation. The efficiency of transcription would require the removal of loops formed by repressors and regulatory complexes conditioning the stalling sites of the RNApol II complex along the template strand (Liu et al 2013). Interesting experiments by Daniel et al. shed light on the RXR-mediated response in murine macrophages defining active binding sites, transcriptional changes and

chromatin structure loops (Daniel et al. 2014). These regions interact with intergenic and intronic motives involving RXR and PU.1 on one side, CTCF and RAD21 on the other. These factors are critical to the genomic regulation because RAD21/CTCF binding sites, flanking the pre-initiation complex, are enrolled in the structural remodelling of the chromatin, and RXRs bind regions associated with PU.1, a repressor of *TGM2* promoter activity (Jensen et al. 2015). Upon RA activation, RXR complexes recruit p300 and leave PU.1 giving start to transcription and producing enhancer RNAs at the *TGM2* locus. These same regions can interact with other transcription factors and enhancer RNA molecules to modulate expression of *TGM2* (Sándor et al. 2016), as do the RAD21 and CTCF insulators (Wada et al. 2009). In untreated cells, RNApol II bound target genes at sites overlapping the active chromatin marker H3K4me3 and RAD21/CTCF sites, as major checkpoints downstream of the transcription start site which must be progressed through during elongation along the gene. CTCF forms complexes also with Sin3A to inhibit MYC activity by recruiting HDACs and interacting with YY1 (Zlatonova and Caiafa 2009), presenting distinctive functions either as transcriptional activator or repressor (He and Casaccia-Bonefil 2008, Kleiman et al. 2016). ChIP-seq data on SK-N-SH cells showed that the binding sites of factors listed in Fig. 3 disappear in the presence of ATRA, while the binding of RAD21 increases at the overlapping sites. After ATRA treatment, peaks of enrichment are evident for YY1, CTCF, RAD21 and p300 at the beginning of *TGM2* within the first intron through a mechanism of closure at the DNA sequence corresponding to the TG2-lncRNA.

We have analyzed the functional roles of TG2-lncRNA in NB4 cells by employing AS-lncRNA. It is fair to underscore that the silencing assay required ATRA treatment because only under this condition TG2-lncRNA is highly expressed. AS-lncRNA did not show a distinct antisense efficacy, because it increased, rather than reduced, the levels of transcripts. This unexpected result can possibly be explained either by the removal of an inhibitory effect or by a simple enhancement of expression due to the capacity of AS-lncRNA sequences to form duplexes with decoy activity. In any case, TG2-lncRNA was recovered in complexes with transcription factors by RIP experiments.

ChIP-seq data on untreated SK-N-SH cells highlighted also the binding of other important nuclear factors to intronic regions, where they may drive the generation of LEV. Among them GATA3 shows a large enrichment at intron 10, with features of promoter, upstream of the LEV transcriptional start. RIP experiments using Abs against GATA3 demonstrated its involvement in the transcriptional complex driving the overall expression of LEV under ATRA control. Interestingly, the appearance of LEV and TGH were associated with response to ATRA also in NB4 cells. It' is important to draw attention to the levels of GATA3 and RXRα: GATA3 was constant in nuclear extracts of NB4 cells treated with the drug, declined in SK-N-SH (Hosoda et al. 2015) and was highly expressed in SH-SY5Y cells (Peng et al. 2015), while RXRα declined during the first day of treatment in SH-SY5Y cells (Girardi et al. 2019), as well as decreased in NB4 cells after 48h of ATRA stimulation. This trend shows an association with *TGM2* expression.

We assume that the presence of specific factors at intronic checkpoints may facilitate the alternative use of poly(A) and splicing sites (Colgan and Manley 2018, Curran et al. 2018), giving rise to a heterogeneity of mRNAs, some of them with differences in the 3′ portion, while others with features of ncRNAs which can selectivity interfere with degradation processes (Smith and Baker 2015, Kurosaki et al. 2018). Usually, NMD proceeds to the quick degradation of abnormal RNA preventing the deleterious accumulation of truncated isoforms, which might potentially escape degradation by readdressing NMD to ncRNAs. In this context, RIP experiments using Abs against phUPF1, the component of NMD responsible for active degradation of targets (Luesink et al. 2009), demonstrated the recovery of LEV upon ATRA treatment with a simultaneous decrease of binding to other transcripts and accumulation of TG2 and TGH proteins in the cells, further controlled by TG2 ligands (Bergamini et al. 2007). Therefore, LEV might represent a strategy to elude the surveillance of NMD versus truncated variants, which are translated into specific proteins. Furthermore, the effects of antisense molecule against LEV suggests also a relationship with TG2-lncRNA. We did not uncover a peptide with a molecular weight consistent with the length of the LEV transcript, supporting the hypothesis that the polyadenylated LEV transcript, which we have herein described for the first time, plays a truly regulatory role. A hypothesis for a potential role of LEV as a mRNA-masking microRNA catcher at its 3'UTR, saving the full-length TG2 transcript from degradation by RNA-induced Silencing Complex, also deserves further studies.

Only recently it has been highlighted that degradation by NMD is able to generate RNA fragments complementary to genomic sequences which can interact with factors, as COMPASS complex and modulate gene expression by interfering with H3K4me3 mark regions (Wilkinson 2019), such as those present at intronic sequences flanking both TG2-lncRNA and LEV. From this point of view, both the NMD (Lareau et al. 2007) and the cooperation between transcriptional factors and regulatory RNAs appear crucial for the efficient transcription of *TGM2* and for the generation/accumulation of variants with dramatically different catalytic activities, in particular for the involvement of this enzyme in cancer and drug sensitivity (Kerr et al. 2017).

Author Contributions

Methodology and analysis, O.F., F.C., G.A. C.C., and N.B.; Biostatistical analysis: L.M.; devising of the study, C.M.B. and N.B.; writing, O.F., C.F., C.M.B. and N.B.; supervision, S.V., C.M.B., N.B.; funding acquisition, C.M.B. and S.V.

Funding

The research was supported by grants FIR2017 (Unife) and FAR2018 (Unife) to C.M.B. and by Associazione Italiana Ricerca sul Cancro AIRC (IG 17063) to S.V.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Electronic supplementary material

The online version of this article contains supplementary material, which is available to authorized users.

References

Benedetti, L, Grignani F, Scicchitano BM, Jetten AM, Diverio D, Lo Coco F, Avvisati G, Gambacorti-Passerini C, Adamo S, Levin AA, Pelicci PG, Nervi C (1996) Retinoid-induced differentiation of acute promyelocytic leukemia involves PML-RARalpha-mediated increase of type II transglutaminase. Blood 87:1939-1950. PMID: 8634442

Beninati S, Piacentini M, Bergamini CM (2017) Transglutaminase 2, a double face enzyme. Amino Acids 49:415-423. https://10.1007/s00726-017-2394-5

Bergamini CM (2007) Effects of ligands on the stability of tissue transglutaminase: studies in vitro suggest possible modulation by ligands of protein turn-over in vivo. Amino Acids 33, 415-421. https://10.1007/s00726-006-0457-0

Bianchi N, Beninati S, Bergamini CM (2018) Spotlight on the transglutaminase 2 gene: a focus on genomic and transcriptional aspects. Biochem J 475:1643-1667. https://10.1042/BCJ20170601

Breitman TR, Selonick SE, Collins SJ (1980) Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. Proc Natl Acad Sci USA 77:2936-2940. PMID: 6930676

Brun ES (2012) Thesis for the degree Master in Molecular Biotechnology MBI-3941, Dissertation, University of Tromsø, Norway.

Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. Genes Dev 11:2755-2766. https://10.1101/gad.11.21.2755

Comandini A, Naro C, Adamo R, Akbar AN, Lanna A, Bonmassar E, Franzese O (2013) Molecular mechanisms involved in HIV-1-Tat mediated inhibition of telomerase activity in human CD4(+) T lymphocytes. Mol Immunol 54:181-192. https://10.1016/j.molimm.2012.12.003

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA 107:21931-21936. https://10.1073/pnas.1016071107

Curran E, Wang H, Hinds TR, Zheng N, Wang EH (2018) Zinc knuckle of TAF1 is a DNA binding module critical for TFIID promoter occupancy. Sci Rep 8:4630-4643. https://10.1038/s41598-018-22879-5

Daniel B, Nagy G, Hah N, Horvath A, Czimmerer Z, Poliska S, Gyuris T, Keirsse J, Gysemans C, Van Ginderachter JA, Balint BL, Evans RM, Barta E, Nagy L (2014) The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages. Genes Dev 28:1562-1577. https://10.1101/gad.242685.114

de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, Swinkels DW, Span PN (2005) Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Lab Invest 85:154-159. https://.1038/labinvest.3700208

de Santiago I, Liu W, Yuan K, O'Reilly M, Chilamakuri CS, Ponder BA, Meyer KB, Markowetz F (2017) BaalChIP: Bayesian analysis of allele-specific transcription factor binding in cancer genomes. Genome Bio 18:39-56. https://10.1186/s13059-017-1165-7

de Stephanis L, Mangolini A, Servello M, Harris PC, Dell'Atti L, Pinton P, Aguiari G (2018) MicroRNA501-5p induces p53 proteasome degradation through the activation of the mTOR/MDM2 pathway in ADPKD cells. J Cell Physiol 233, 6911-6924. https://10.1002/jcp.26473

Fésüs L, Szondy Z (2005) Transglutaminase 2 in the balance of cell death and survival. FEBS Lett 579:3297-3302. https://10.1016/j.febslet.2005.03.063

Fraij BM, Gonzales RA (1996) A third human tissue transglutaminase homologue as a result of alternative gene transcripts. Biochim Biophys Acta 1306:63-74. https://10.1016/0167-4781(95)00219-7

Girardi CS, Rostirolla DC, Lini FJM, Brum PO, Delgado J, Ribeiro CT, Teixeira AA, Peixoto DO, Heimfarth L, Kunzler A, Moreira JCF, Gelain DP (2019) Nuclear RXRα and RXRβ receptors exert distinct and opposite effects on RA-mediated neuroblastoma differentiation. Biochim Biophys Acta Mol Cell Res 1866:317-328. https://10.1016/j.bbamcr.2018.11.014

Griffin M, Casadio R, Bergamini CM (2002) Transglutaminases: nature's biological glues. Biochem J 368:377-396. https://10.1042/BJ20021234

He Y, Casaccia-Bonnefil P (2008) The Yin and Yang of YY1 in the nervous system. J Neurochem 106:1493-1502. https://10.1111/j.1471-4159.2008.05486.x

Hosoda K, Sato M, Yanai K (2015) Identification and characterization of human genomic binding sites for Retinoic Acid Receptor/Retinoid X Receptor heterodimers. Advances in Biological Chemistry 5:58-72. https://10.4236/abc.2015.52006

Jensen HA, Yourish HB, Bunaciu RP, Varner JD, Yen A (2015) Induced myelomonocytic differentiation in leukemia cells is accompanied by noncanonical transcription factor expression. FEBS Open Bio 5:789-800. https://10.1016/j.fob.2015.09.008

Jolma A, Yan J, Whitington T, Toivonen J, Nitta KR, Rastas P, Morgunova E, Enge M, Taipale M, Wei G, Palin K, Vaquerizas JM, Vincentelli R, Luscombe NM, Hughes TR, Lemaire P, Ukkonen E, Kivioja T, Taipale J (2013) DNA-binding specificities of human transcription factors. Cell 152:327-339. 1 Doi: 0.1016/j.cell.2012.12.009

Jurka J (2000) Repbase Update: a database and an electronic journal of repetitive elements. Trends Genet 16:418-420. PMID: 10973072

Kamashev D, Vitoux D, De Thé H (2004) PML-RARA-RXR oligomers mediate retinoid and rexinoid/cAMP cross-talk in acute promyelocytic leukemia cell differentiation. J Exp Med 199:1163-1174. https://10.1084/jem.20032226

Kerr C, Szmacinski H, Fisher ML, Nance B, Lakowicz JR, Akbar A, Keillor JW, Lok Wong T, Godoy-Ruiz R, Toth EA, Weber DJ, Eckert RL (2017) Transamidase site-targeted agents alter the conformation of the transglutaminase cancer stem cell survival protein to reduce GTP binding activity and cancer stem cell survival. Oncogene 36:2981–2990. https://10.1038/onc.2016.452

Király R, Demény M, Fésüs L (2011) Protein transamidation by transglutaminase 2 in cells: a disputed Ca2+-dependent action of a multifunctional protein. FEBS J 278:4717-4739. https://10.1111/j.1742-4658.2011.08345.x

Kleiman E, Jia H, Loguercio S, Su AI, Feeney AJ (2016) YY1 plays an essential role at all stages of B-cell differentiation. Proc Natl Acad Sci USA 113:E3911-E3920. https://10.1073/pnas.1606297113

Kovalevich J, Langford D (2013) Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. Methods Mol Biol 1078:9-21. https://10.1007/978-1-62703-640-5_2

Kurosaki T, Hoque M, Maquat LE (2018) Identifying cellular Nonsense-Mediated mRNA Decay (NMD) targets: immunoprecipitation of phosphorylated UPF1 followed by RNA sequencing (p-UPF1 RIP-Seq). Methods Mol Biol 1720:175-186. https://10.1007/978-1-4939-7540-2_13

Lai TS, Greenberg CS (2013) TGM2 and implications for human disease: role of alternative splicing. Front Biosci (Landmark Ed) 18:504-519. PMID: 23276939

Lareau LF, Brooks AN, Soergel DAW, Meng Q, Brenner SE (2007) The Coupling of Alternative Splicing and Nonsense-Mediated mRNA Decay. Chapter 12. In: Benjamin J. Blencowe and Brenton R. Graveley (ed) Alternative Splicing in the Postgenomic Era. Landes Bioscience and Springer Science+Business Media. https://escholarship.org/content/qt6089j8mq/qt6089j8mq.pdf

Lee HJ, Park MK, Bae HC, Yoon HJ, Kim SY, Lee CH (2012) Transglutaminase-2 is involved in all-trans retinoic acid-induced invasion and matrix metalloproteinases expression of SH-SY5Y neuroblastoma cells via NF-κB pathway. Biomol Ther (Seoul) 20:286-292. https://10.4062/biomolther.2012.20.3.286

Liu X, Bushnell DA, Kornberg RD (2013) RNA Polymerase II Transcription: structure and mechanism. Biochim Biophys Acta 1829:2–8. https://10.1016/j.bbagrm.2012.09.003

Luesink M, Pennings JL, Wissink WM, Linssen PC, Muus P, Pfundt R, de Witte TJ van der Reijden BA, Jansen JH (2009) Chemokine induction by all-trans retinoic acid and arsenic trioxide in acute promyelocytic leukemia: triggering the differentiation syndrome. Blood 114:5512-5521. https://10.1182/blood-2009-02-204834

Masetti R, Biagi C, Zama D, Vendemini F, Martoni A, Morello W, Gasperini P, Pession A (2012) Retinoids in pediatric onco-hematology: the model of acute promyelocytic leukemia and neuroblastoma. Adv Ther 29:4747-4762. https://10.1007/s12325-012-0047-3

Mikesch JH, Gronemeyer H, So CW (2010) Discovery of novel transcriptional and epigenetic targets in APL by global ChIP analyses: emerging opportunity and challenge. Cancer Cell 17:112-114. https://10.1016/j.ccr.2010.01.012

Minotti L, Baldassari F, Galasso M, Volinia S, Bergamini CM, Bianchi N (2018) A long non-coding RNA inside the type 2 transglutaminase gene tightly correlates with the expression of its transcriptional variants. Amino Acids 50:421-438. https://10.1007/s00726-017-2528-9

Nagai J, Yazawa T, Okudela K, Kigasawa H, Kitamura H, Osaka H (2004) Retinoic acid induces neuroblastoma cell death by inhibiting proteasomal degradation of retinoic acid receptor alpha. Cancer Res 64:7910-7917. https://10.1158/0008-5472.CAN-04-1178

Peng H, Ke XX, Hu R, Yang L, Cui H, Wei Y (2015) Essential role of GATA3 in regulation of differentiation and cell proliferation in SK-N-SH neuroblastoma cells. Mol Med Rep 11:881-886. https://10.3892/mmr.2014.2809

Phatak VM, Croft SM, Rameshaiah Setty SG, Scarpellini A, Hughes DC, Rees R, McArdle S, Verderio EA (2013) Expression of transglutaminase-2 isoforms in normal human tissues and cancer cell lines: dysregulation of alternative splicing in cancer. Amino Acids 44:33-44. https://10.1007/s00726-011-1127-4

Redner RL, Beumer JH, Kropf P, Agha M, Boyiadzis M, Dorritie K, Farah RJZ, Hou A, Im SH, Lim A, Raptis A, Sehgal SM, Christner D, Normolle Johnson DE (2018) A phase-1 study of dasatinib plus all-trans retinoic acid in acute myeloid leukemia. Leuk Lymphoma 59:2595-2601. https://10.1080/10428194.2018.1443330

Saeed S, Logie C, Francoijs KJ, Frigè G, Romanenghi M, Nielsen FG, Raats L, Shahhoseini M, Huynen M, Altucci L, Minucci S, Martens JH, Stunnenber HG (2012) Chromatin accessibility, p300, and histone acetylation define PML-RARα and AML1-ETO binding sites in acute myeloid leukemia. Blood 120:3058-3068. https://10.1182/blood-2011-10-386086

Sandberg R, Neilson JR, Sarma A, Sharp PA & Burge CB (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320:643-647. https://10.1126/science.1155390

Sándor K, Daniel B, Kiss B, Kovács F, Szondy Z (2016) Transcriptional control of transglutaminase 2 expression in mouse apoptotic thymocytes. Biochim Biophys Acta 1859:964-974. https://10.1016/j.bbagrm.2016.05.011

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3:1101-1108. PMID: 18546601

Smit AFA, Hubley R, Green P (1996-2010) RepeatMasker Open-3.0. http://www.repeatmasker.org/faq.html

Smith JE, Baker KE (2015) Nonsense-mediated RNA decay--a switch and dial for regulating gene expression. Bioessays 37:612-623. https://10.1002/bies.201500007

Tee AE, Marshall GM, Liu PY, Xu N, Haber M, Norris MD, Iismaa SE, Liu T (2010) Opposing effects of two tissue transglutaminase protein isoforms in neuroblastoma cell differentiation. J Biol Chem 285:3561-3567. https://10.1074/jbc.M109.053041

Usmanova NM, Kazakov VI, Tomilin NV (2008) SINEs in mammalian genomes can serve as additional signals in formation of facultative heterochromatin. Cell and Tissue Biology 2:217–221. https://link.springer.com/article/10.1134/S1990519X08030012

Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S (2009) A wave of nascent transcription on activated human genes. Proc Natl Acad Sci USA 106:18357-18361. https://10.1073/pnas.0902573106

Wilkinson MF (2019) Genetic paradox explained by nonsense. Nature 568:179-180. https://10.1038/d41586-019-00823-5

Zlatanova J, Caiafa P (2009) CTCF and its protein partners: divide and rule? Cell Sci 122:1275-1284. https://10.1242/jcs.039990

Zhang G, Pian C, Chen Z, Zhang J, Xu M, Zhang L, Chen Y (2018) Identification of cancer-related miRNA-lncRNA biomarkers using a basic miRNA-lncRNA network. PLoS One eCollection 13(5):e0196681. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0196681

Figure legends

Fig. 1 a RT-qPCR analysis of the expression of regulatory TG2-lncRNA, LEV and full-length TG2, and **b** of truncated TGH2 and TGH, of alternatively spliced *tTGv1* and *tTGv2* transcripts in HL-60 cells treated with 3 µM ATRA. Total RNA was extracted from samples after 16, 24, 36, 48, 60 and 72h of culture, reverse transcribed and the quantification of targets was performed using HPRT (de Kok et al. 2005) as housekeeping gene. Fold changes were quantified using $2^{-\Delta\Delta CT}$ formula with respect to untreated cells analyzed at the same times of culture. The mean ± SD was determined from n=3 independent experiments. $p<0.05$ for each transcript analyzed with One-way ANOVA.

Fig. 2 Analysis of expression of *TGM2* variants in NB4 cells untreated (white circles) and treated with 1 µM ATRA (triangles) for 48h (black), 72h (grey) or 96h (white). The data were extracted from GSE18397 of GEO database (Luesink et al. 2009). Normalized Intensity of the samples was obtained using probe set from HumanExon1_0ST Affymetrix array distributed along the whole *TGM2* gene. The arrows indicate distinct variants. The probe 3905145 represents the mean of signals derived from all expressed gene sequences.

Fig. 3 ChIP-sequencing from ENCODE database of transcription factors and chromatin structural proteins binding the *TGM2* gene to regulatory intronic regions in untreated SK-N-SH cells (upper box) and treated with 6 µM ATRA (lower box). The positions of the binding factors are reported with respect to the gene, depicted on the top of the figure (canonical full-length TG2). The RNApol II sites and the H3K27Ac marks are also reported. Squares indicated the binding sequences recognized by factors within *TGM2*. Colors (blue, light green, dark green, red, orange, pink, violet, black, yellow, ocher, brown) indicate factors binding the same region in ChIP experiments, while shades intensity (light or dark blue) depends on their binding affinity.

Fig. 4 a RT-qPCR analysis of samples obtained from RIP experiments carried out on NB4 cells untreated or treated with 1 µM for 48h. RIP was performed with native extracts obtained from isolated nuclei using Abs against GATA3, RXRα, Max, RNApol II and phUPF1. Sybrgreen and specific primers were used to amplify TG2-lncRNA (oblique slashes), full-length TG2 (white), TGH mRNA (grey) and LEV transcript (black). Fold Enrichment = $2^{-\Delta\Delta Ct\ [RIP/NS]}$ was calculated (as detailed in Material and methods), where NS is Non-Specific Ab binding, as background and the $\log_{10}$ of the values was plotted. Bars on the right of the histogram plots represent the fold changes ratio of the samples treated with 1 µM ATRA /untreated, where (*) indicates $p<0.05$ (using unpaired t test). **b** Western blotting analysis was carried out on nuclear extracts or on total cellular extracts, under conditions of low expression of the analyzed factor, as indicated. Abs AbA037 and AbA036 were specific for full-length TG2 (as well as *tTGv1* and *tTGv2*), while AbA036 is specific for TGH. Abs directed against Actin was used as internal control to calculate the values of the ratio to specific Ab hybridization densitometry.

Fig. 5 a Evaluation of the content of transcription factors influencing *TGM2* gene expression in HL-60 cells in the absence or in the presence of ATRA 1 µM; the values of the ratio to specific Ab hybridization densitometry was reported. **b** The expression of TG2-lncRNA (oblique slashes), full-length TG2 (white), TGH (grey) and LEV transcripts (black) of HL-60 was compared to that of NB4 cells at the same ATRA concentration. **c** Analysis by RT-qPCR of the effects on *TGM2* transcripts of antisense oligonucleotide against LEV (AS-LEV) at 1.5 µM concentration for 24h in HL-60 ATRA-induced, employing scramble oligonucleotide (AS-neg) as negative control. Fold changes were calculated with respect to AS-neg using $2^{\wedge-\Delta\Delta CT}$ method. Histograms represent the mean ± SD of 3 independent experiments; (*) $p<0.05$ (using unpaired t test, two-tailed).

Fig. 6 a RT-qPCR analysis of the expression of TG2-lncRNA (oblique slashes), full-length (white), TGH (grey) and LEV transcripts (black) in NB4 cells treated with 1 µM ATRA and 1.5 µM of specific AS (AS-lncRNA) against TG2-lncRNA transcript. The treatment with a scramble oligonucleotide (AS-neg) in the presence of the same concentration of ATRA and representing negative control was used to quantify the fold increase in the samples. Fold changes were calculated using $2^{-\Delta\Delta CT}$ method. Histograms represent the mean ± SD of 3 independent experiments; (*) $p<0.05$ (using unpaired t test, two-tailed). **b** Cartoon illustrating the potential interference activity of TG2-lncRNA towards the *TGM2* gene expression. Critical binding sites for transcription factors could act like enhancer/decoy or scaffolding to re-direct them to other regions of the gene.

Supplementary Fig. 1 Drawing describing the secondary RNA structure of the TG2-lncRNA with base-paired regions generated by using Vienna web RNA service considering two approaches for the analysis: **a** to have the structure with less free-energy structure, **b** to employ a model based on centroids. In the colour bar, red corresponds to high and blue to a low probability that the structure is hired.

Manuscript
Click here to access/download;Manuscript;Manuscript Franzese O. et al 19-07-2019.pdf
Click here to view linked References

Involvement of non-coding RNAs and transcription factors in the induction of Transglutaminase isoforms by ATRA.

Ornella Franzese[1], Linda Minotti[2], Gianluca Aguiari[3], Fabio Corrà[2], Carlo Cervellati[3], Carlo Ferrari[4], Stefano Volinia[2], Carlo M. Bergamini[3,*], Nicoletta Bianchi[3,*]

[1] Department of Systems Medicine, School of Medicine, University of Rome Tor Vergata, Rome, Italy
[2] Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, Ferrara, Italy
[3] Department of Biomedical Sciences and Specialist Surgery, University of Ferrara, Ferrara, Italy
[4] Department of Clinical and Molecular Sciences, Marche Polytechnic University, Ancona, Italy

*** Correspondence**:
Nicoletta Bianchi and Carlo M. Bergamini, Department of Biomedical Sciences and Specialist Surgery, University of Ferrara, Via Luigi Borsari, 46 - 44121 Ferrara, Italy
nicoletta.bianchi@unife.it; Tel.: +39-0532-455447;
bgc@unife.it; Tel.: +39-0532-455428

ORCID linked to this article:

Ornella Franzese: https://orcid.org/0000-0002-6396-4603
Linda Minotti: https://orcid.org/0000-0003-2116-8411
Gianluca Aguiari: https://orcid.org/0000-0002-0007-0805
Fabio Corrà: https://orcid.org/0000-0002-6803-0797
Carlo Cervellati: https://orcid.org/0000-0003-4777-6300
Carlo Ferrari: https://orcid.org/0000-0001-5112-7448
Stefano Volinia: https://orcid.org/0000-0003-0910-3893
Carlo M. Bergamini: https://orcid.org/0000-0002-9430-8625
Nicoletta Bianchi: https://orcid.org/0000-0001-9280-6017

Acknowledgements
The cell lines employed in this study were kind gifts from Professor Paola Secchiero (HL-60 and NB4 cells), Michael U. Martin (SK-N-SH cells) and Ralf Hass (SH-SY5Y cells) respectively. The authors thank Prof. Giovanni Bernacchia for having made available the instrument to perform qPCR reactions. We are grateful to Prof. Cristian Taccioli and Dr. Marco Ricci for their advice and valuable help to address the study of TG2-lncRNA molecules. Finally, we thank Co.Pe.Go. (Soc.Coop.O.P.) for liberal voluntary contribution to the present research.

Handling Editor: Simone Beninati

Abstract

The multifunctional protein Transglutaminase type 2, is associated with cancer epithelial mesenchymal transition, invasiveness, stemness and drugs resistance. Several variant isoforms and non-coding RNAs are present in cancer and this report explored the expression of these transcripts of the *TGM2* gene, after induction, in cancer cell lines, with *all-trans* retinoic acid. The expression of truncated variants along with two long non-coding RNAs, was demonstrated. One of these, coded from the first intron and the Last Exon Variant, constituted by a sequence corresponding to the last 3 exons and the 3'UTR. Analysis of ChIP-seq data, from ENCODE project, highlighted factors interacting with intronic sequences, which could interfere with the progression of RNApol II at checkpoints, during the elongation process. Some relevant transcription factors bound, in an ATRA-dependent way, were found by RNA immunoprecipitation, notably GATA3 mainly enriched to Last Exon Variant non-coding RNA. The involvement of NMD in the regulation of the ratio among these transcripts was observed, as the prevalent recovering of Last Exon Variant to phUPF1-complexes, with decrease of the binding towards other selective targets. This study contributes to identify molecular mechanisms regulating the ratio among the variants and improves the knowledge about regulatory roles of the non-coding RNAs of the *TGM2* gene.

Keywords: Transglutaminase type 2; transcriptional variants; non-coding RNA; retinoic acid; GATA3

ABBREVIATIONS

AAs	amino acids
Abs	antibodies
APL	acute promyelocytic leukemia
AS	antisense oligonucleotide
ATRA	*all-trans* retinoic acid
CTCF	CCCTC-binding factor
FAB	French–American–British
FBS	foetal bovine serum
GEO	Gene Expression Omnibus
HDACs	histone deacetylases
LEV	Last Exon Variant
lncRNA	long non-coding RNA
NI	normalized signal intensity
NMD	nonsense-mediated mRNA decay
PBS	phosphate-buffered saline
phUPF1	phosphorylated ATP-dependent RNA helicase upframeshift 1
RA	retinoic acid
RARα	retinoic acid receptor α
RAREs	Retinoic Acid Responsive Elements
RIP	RNA-immunoprecipitation
RRE	retinoid receptor element
RT-qPCR	reverse transcription and real time quantitative PCR
TGM2	Transglutaminase type 2 gene
TG2	Transglutaminase type 2

Introduction

Transglutaminase type 2 (TG2) has attracted much attention in experimental oncology because of its involvement in the control of cell cycle and tumour cell stemness along with its higher expression in secondary than in primary tumours. TG2 is a bifunctional enzyme involved in post-translational protein modification (transamidation) and in membrane signalling, with opposite effects on cell survival and death (Tee et al. 2010, Beninati et al. 2017) in relation to occurrence of apoptosis (Fésüs and Szondy 2005), Epithelial-Mesenchymal Transition, expansion of cancer stem cells and resistance to chemotherapy (Tee et al. 2010, Kerr et al. 2017).

In addition to the full-length enzyme, normal and pathological tissues express at variable levels transcriptional variants of the protein that are apparently tailored to carry out the transamidation and the signalling activities, which are performed by the N- and the C-terminal regions respectively. The forms described in literature include the full-length, the two *tTGv1* and *tTGv2* variants with altered structure at the C-terminus due to alternative splicing (Phatak et al. 2013) and the shorter deleted isoform TGH. All these isoforms are insensitive to regulation by GTP (Lai and Greenberg 2013). A much shorter form, TGH2, full-length TG2 promotes cell survival, while the altered isoforms lacking the 3'-terminal portion with de-regulated enzyme catalytic activity are associated with cell death, it follows that the different forms carry out different cellular functions (Király et al. 2011).

Gene dysregulation often leads to the accumulation of altered variants, which are often considered as tumor markers. Similarly, the treatments with anti-cancer drugs modulate the levels of protein isoforms, as occurs for Transglutaminase type 2 gene (*TGM2*) among other genes. Therefore, the study of their modification and the use of these experimental models is critical for basic studies aimed at identifying factors involved in the transcriptional control and providing new tools useful for the identification of useful clinical perspectives and the development of new molecular therapies.

The expression of the *TGM2* is induced by retinoids *in vitro* and *in vivo* in patients affected by leukemia (Benetti et al. 1996) and neuroblastoma, who are successfully treated with retinoic acid (RA) with a rapid increase of the protein (Tee et al. 2010, Masetti et al. 2012). This regulated expression is under the control of multiple Retinoid Response Elements (RRE) (Daniel et al. 2014).

So far, investigations concerning the effects of RA have been defective due to the expression of the different TG2 isoforms, a relevant issue in the light of their functional differences. The recognized association between *TGM2* expression and tumor status prompted us to investigate the differential expression of the enzyme in selected acute promyelocytic leukemia (APL) and neural tumor cell lines, in which *all-trans* retinoic acid (ATRA) acts as differentiating agent. The APL cell lines employed in the present investigation were the HL-60 cells, lineage-bipotent myeloblasts (French–American–British, FAB M2) which can be induced to differentiation into granulocytes by RA (Jensen et al. 2015), and the leukemic NB4 cells belonging to FAB M3. Both cell lines are sensitive to ATRA but throughout the engagement of different receptors and Retinoic Acid Responsive Elements (RAREs). The neuroblastoma lines employed were the SK-N-SH cells, isolated from a bone marrow biopsy of a patient with neuroblastoma, and its more aggressive subclone SH-SY5Y. The parental SK-N-SH cells display two morphologically distinct neuroblast-like and epithelial-like phenotypes, the former expressing enzymes typical of the catecholaminergic neurons that are instead absent in the latter. SK-N-SH cells are forced to differentiate towards a neuronal phenotype with extensive neurite outgrowth by treatment with ATRA, which is actually used in the adjuvant therapy of residual or relapsed neuroblastoma. ATRA affects differently *TGM2* expression in SH-SY5Y cells (Lee et al. 2012) and neuron-like phenotype is induced only in low serum condition, which may also promote cell survival through activation of the phosphatidylinositol 3-kinase/Akt signalling pathway and increase of anti-apoptotic Bcl-2 (Kovalevich and Langford 2013).

The expression of mRNAs coding for the full-length and for the variant isoforms, as well as that of the regulatory long non-coding RNAs (lncRNAs), including the TG2-lncRNA reported by Minotti et al. (2018), was analyzed in these cellular models, along with their mechanisms of action upon treatment with RA (Jensen et al. 2015, Redner et al. 2018). The present study highlighted also the expression of a novel RNA, which we named Last Exon Variant (LEV), reported in "NEDO human cDNA sequencing project" (GenBank: AK126508.1) on the Genome Browser database https://genome.ucsc.edu/). LEV is constituted by the sequences of the last 3 exons of *TGM2* gene including the complete 3'UTR region (Bianchi et al. 2018). To identify the factors engaged in the differential expression of the various *TGM2* transcripts under ATRA modulation, we have taken advantage of ChIP-seq data from ENCODE (Consortium on Genome Browser Human Genome hg19) and of results from original experiments performed in our laboratory, including RNA immunoprecipitation (RIP) and real time quantitative PCR (RT-qPCR) assays. We have also explored approaches to target selectively the expression of the lncRNAs by means of antisense molecules in tumor cell lines treated with ATRA.

Transcription variants differing in their altered 3' last exons likely represent a strategy set in motion by cells to escape the effects of different miRNAs (Sandberg et al. 2008). In addition, several recent studies indicate that lncRNAs drive many important cancer phenotypes through their interactions either with DNA, proteins or RNA, regulating gene expression at the transcriptional, post-transcriptional, and epigenetic levels. Moreover, together with miRNAs, these molecules can act as competing endogenous RNAs capable of controlling mRNA targets, thus playing an important role in cancer. Indeed, lncRNAs can be sequestered from miRNAs of other specific targets, although the mechanisms underlying these interactions have not been elucidated yet (Zhang et al. 2018).

Methods

Cell line cultures and treatments

All materials for cell culture were purchased from Sigma-Aldrich.

Human leukemia HL-60 and NB4 cells (a gift from Prof. Paola Secchiero, University of Ferrara, Italy) were cultured at 37°C in a 5% CO_2 humidified atmosphere in RPMI 1640 medium, supplemented with 10% foetal bovine serum (FBS), 50 U/mL penicillin and 50 μg/mL streptomycin (Minotti et al. 2018). Cells were seeded at $12x10^4$/mL and induced to differentiation by treatment with 1 and 3 μM ATRA for different time intervals (from 16 to 72h).

Neuroblastoma SK-N-SH cells (a gift from Prof. Michael U. Martin, Justus-Liebig-University Giessen, Germany) were cultured at 37°C in a 5% CO_2 humidified atmosphere in EMEM (EBSS), supplemented with 15% FBS, 2 mM glutamine, 50 U/mL penicillin and 50 μg/mL streptomycin. SH-SY5Y cells (provided by Prof. Ralf Hass, Hannover Medical School, Germany) were cultured using a mixture 1:1 of Eagle's Minimum Essential Medium and F12 Medium with 10% FBS. For sub-confluent cultures, cells (at 70-80% of confluence) were splited using 0.25% Trypsin and incubated for 5 min at 37°C and 5% CO_2, then seeded at $3x10^3$-10^4 cells/mL. Differentiation was induced by treatment with 3 and 10 μM ATRA for different time intervals depending on the growth rate of the cell lines (16 and 48h). Pellets were collected during exponential growth and total RNA was extracted from $5x10^6$ cells.

DNA antisense oligonucleotides (AS) to target the transcripts of TG2-lncRNA (5'-TAGGGATTCAGCTCCCACCGGGTC-3'), LEV (5'-CTCATGATGCAGAATCACC-3'), and a scramble oligonucleotide (5'-AATTCTCCGAACGTGTCACGT-3') were used at the final concentration of 1.5 μM to modulate the expression of 10^6 cells needed to perform enough reactions to analyse the different *TGM2* transcripts, adjusting as appropriated the protocol of siPORT™ NeoFX™ Transfection Agent (ThermoFisher Scientific, Life Technologies, Applied Biosystems, Italy).

Reverse transcription and real-time quantitative PCR reactions to quantify *TGM2* transcripts

Reverse transcription reactions were carried out using the TaqMan® Reverse Transcription Reagents kit using the protocol from ThermoFisher Scientific. PCR reactions were performed using PowerUp SYBR Green Master Mix (ThermoFisher Scientific, Life Technologies, Applied Biosystems, Italy). Primers to amplify the targets are reported in Table 1. The couple of primers F2TGH/FR2TGH and Fshort3ex/Rshort3ex were designed with the Primer Blast tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The amplification was performed using Thermal cycler CFX96 Touch™ Real-Time PCR Detection System (Bio-rad Laboratories S.r.l., Segrate, Milan, Italy).

Fold increase was determined by comparing the cycle threshold (CT) values of the amplified targets with that of the reference HPRT1 gene, which is reported as one of those with better stability (de Kok et al. 2005). We have calculated the negative exponent in formula $2^{-\Delta\Delta CT}$ ($\Delta\Delta$CT) to quantify fold changes in gene expression of treated samples with respect to control cells, as suitably specified in the legends (Schmittgen and Livak 2008). The t test was performed and p values were calculated as detailed in the legends.

Analysis of *TGM2* gene expression array using datasets of NB4 cells from GEO database

Analysis of *TGM2* gene expression was performed (as detailed in Table 2) on data from public Gene Expression Omnibus (GEO) considering 2 datasets of total RNA analyzed with HumanExon1_0ST Affymetrix on NB4 cell line (Luesink et al. 2009), which is composed of specific probes distributed along the whole genome to evaluate with high performance the expression of altered transcripts including intronic sequences. The probes considered for the *TGM2* gene are reported in Minotti et al. (2018). For each culture we compared signal Normalized Intensity (NI) from untreated and treated cells with ATRA.

GEO dataset and Chip-seq analysis

Chip-seq analysis (Table 2) was derived from datasets through Genome Browser tools (HudsonAlpha/Caltech ENCODE group) obtained from neuroblastoma SK-N-SH cells untreated or treated with 6 μM ATRA for 48h. The tracks of ChIP-seq represent a comprehensive set of human transcription factor binding sites based on experiments generated by a group in the ENCODE Project Consortium 2011 and 2012, which developed Integrative analysis of experiments performed in duplicate and scored against two controls displaying the regions of enrichment as peak calls.

Preparation of nuclear extracts and RNA immunoprecipitation

RIP has been developed to identify RNA species that bind to a protein of interest by a protocol including: (1) preparation of protein lysate from target cells, (2) immunoprecipitation of the protein with specific antibodies (Ab)s and extraction of protein-bound RNAs. We followed the protocol reported by Brun (2012) as modified according to Comandini et al. (2013). Native nuclei were obtained from $15x10^6$ cells washed three times in Phosphate-buffered saline (PBS), and resuspended in a solution of 2 mL PBS plus 2 mL nuclear isolation Buffer (1.28 M Sucrose, 40 mM Tris-HCl pH 7.5, 20 mM $MgCl_2$, 4% (v/v) Triton X-100) and 6 mL H_2O treated with diethylpyrocarbonate. After 20 min incubation on ice native nuclei were collected by centrifugation at 2500 g for 15 min at 4°C, suspended in 200 µl freshly prepared ice-cold RIP buffer (25 mM Tris pH 7.5,150 mM KCl, 5 mM EDTA, 0.5 mM DTT, 0.5% NP40, plus HALT protease inhibitor cocktail 1x and RNase Inhibitor 20 U/mL) (Thermo Fisher Scientific, Italy) and sheared on ice by passage through a 27 gauge needle with 15–20 strokes. Nuclear membranes and debris were removed by centrifugation at 16000 g for 10 min at 4°C and supernatants of nuclear lysates were saved. Each extract was incubated at 4°C on gentle rotation overnight with 10 µg of Abs against RNApol II (Pierce ma146093, Anti-RNA Polymerase II CTD monoclonal mouse), Max (ThermoFisher Scientific pa529745, MAX Polyclonal Antibody), GATA3 (Abcam, [EPR16651]-ChIP grade ab199428), RXRα (Abcam, EPR7106, ab125001) and upframeshift 1 (phUPF1) (Ser1127) (EMD Millipore Corporation). Twenty µl of protein A/G PLUS conjugated agarose beads, pre-washed three times with 500 µl ice-cold RIP buffer, were incubated in rotation for 3 hours at 4°C with specific Ab. The supernatant was removed after centrifugation at 4000 g for 2 min followed by 3 washing steps with ice cold 500 µl RIP buffer for the extraction of the bound RNA. The incubation with Abs against RNApol II, was preceded by one step of 2h at 4°C in the presence of 1 µg of competitor oligonucleotide to reduce the background.

Purification of transcripts bound to immunoprecipitated proteins and data analysis

After IP reactions RNAs were isolated from the protein-A-G conjugated agarose complexes the beads by extraction with 1 mL TRIzol®, as described in the manufacturer's instructions. RNAs were precipitates as reported by Brun (2012) and suspended in 10 µl RNase free H_2O.

One microgram of total RNA was employed as starting material in 20 µl Reverse Transcription reaction. We employed TaqMan® Reverse Transcription Reagents kit, as above reported. RNA contents have been normalized using HPRT1 for setting of samples. Quantitative real-time PCRs were carried out in order to assay target RNAs using PowerUp SYBR Green Master Mix in 10 µl volume, according to the manufacturer's instructions.

CT values considered in our analysis were derived from arithmetic mean of 3 independent experiments. To analyse data of RIP experiments Fold Enrichment of RIP reactions was calculated by plugging the average CT values to respective Input RNA fractions (derived from amounts of samples before the RIP), also reported in specific user guide (https://www.sigmaaldrich.com). The quantification of RNA targets associated to the Ab complexes was performed as follows: i) normalization of each RIP RNA fractions' CT value to the Input RNA fraction CT value for the same qPCR Assay (ΔCT) considering also dilution factor: **ΔCT [normalized RIP]** = CT [RIP] – (CT [Input] – Log2 Input Dilution Factor), where Input fraction is 1% (log2 of 100); ii) the binding of Ab to protein A/G PLUS conjugated agarose beads (after and pre-RIP) was used to adjust the data for Non-Specific (NS) background: **ΔCT [normalized NS**]; iii) finally we have calculated: **ΔΔCT [RIP/NS]** = ΔCT [normalized RIP] – ΔCT [normalized NS]; iv) for quantification of the **Fold Enrichment**= $2^{-\Delta\Delta CT\,[RIP/NS]}$.

Western blotting

Cells were collected by centrifugation at 800xg for 10 min and washed twice in PBS buffer containing a cocktail of protease inhibitors (Sigma-Aldrich). Cells were lysed for 30 min by using a single detergent lysis buffer (10 mM Tris-HCl, pH 7.5, 10 mM NaCl, 3 mM $MgCl_2$, 1% v/v Triton X-100) with protease inhibitors. Nuclear extracts were prepared by removing cell debris by centrifugation at 10000 g for 10 min.

Twenty five µg of total protein or nuclear extract were solubilized in SDS loading buffer (60 mM Tris-HCl, pH 6.8, 1.8% SDS, 6% (v/v) glycerol, 1% β-mercaptoethanol, 0.002 (w/v) bromophenol blue) and analyzed by 12.5% SDS-PAGE. Proteins were transferred on nitrocellulose filters (EuroClone) in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol, pH 8.3) for 2h by using an elettroblotting chamber (Bio-Rad). Nitrocellulose filters, blocked in a solution containing 5% non-fat dry milk in PBS buffer 0.05% Tween 20®, were treated overnight at 4°C with specific Abs. The Anti-TG2 mouse monoclonal Abs A036 and A037 (Zedira GmbH, Darmstadt, Germany), recognize respectively the epitope K^{527}YLLNLNL^{534} of the beta-Barrel-1 Domain, present in both TG2 and TGH isoforms, and the epitope V^{604}AEVSLQNP^{612} present in beta-Barrel-2-Domain, but not in TGH isoform. For the analysis of transcription factors we employed Abs against Max (ThermoFisher Scientific pa529745, MAX Polyclonal Ab), GATA3 (Abcam, [EPR16651]-ChIP grade ab199428), RXRα (Abcam, EPR7106, ab125001). Protein loads were compared using an Ab against β-actin. After hybridization, the nitrocellulose membranes were washed three times in PBS buffer with 0.05% Tween20® and incubated with the HRP-conjugated secondary Abs. Further, three washes in PBS 0.05% with Tween 20®

were performed and the protein bands were visualised through autoradiography by using the enhanced chemiluminescence system (EuroClone) (de Stephanis et al. 2018).

Results

Expression of transcriptional variants of *TGM2* upon ATRA treatment

The induction of *TGM2* expression by ATRA by RT-qPCR was investigated in HL-60 cells. The transcripts for canonical full-length TG2, truncated TGH2 and TGH isoforms, along with alternative spliced variants *tTGv1* and *tTGv2* (Phatak et al. 2013), TG2-lncRNA and LEV were quantified by RT-qPCR using the primers reported in Table 1. Incubation with 3 μM ATRA was performed for time intervals ranging between 16 and 72h (Breitman et al. 1980). A drug-induced transcriptional up-regulation of *TGM2* variants as compared with untreated cells was observed at the same time of culture. However, each variant and regulatory RNA showed different profiles as described in Fig. 1, showing different levels of expression appreciable at the specific time points. Therefore, LEV was induced more intensively than TG2-lncRNA by ATRA before 36h of culture with slight reversal of the trend after this time. After 48h of induction the highest TG2-lncRNA levels declined, whereas full-length TG2 further increased. In contrast, the highest levels of TGH2 and TGH variants were reached at 24h of treatment and decreased over extended time intervals, while the strong initial increase of the spliced *tTGv1* and *tTGv2* variants showed fluctuating trends.

Experiments were carried out also on leukemic NB4 cells, in which *TGM2* is among the highest induced genes with appreciable levels of the protein also in untreated cells. The gene expression profiling of *TGM2* from the GSE18397 total RNA dataset of GEO (Table 2) was performed by GeneChip Human Exon 1.0 ST Array using probes recognizing both exonic and intronic sequences belonging to the transcriptional variants allow to discern differences across single variants based on the use of specific intronic sets of probes (Minotti et al. 2018). NB4 cells cultured for 48, 72 and 96h in the absence or presence of 1 μM ATRA displayed particularly high levels of both the TG2-lncRNA (probe 3905186) (Minotti et al. 2018) and the LEV transcript (probe 3905153) when cultured in the presence of the drug. ATRA treatment increased the overall expression of *TGM2* (Fig. 2), including the transcripts derived from altered splicing and polyadenylation (Benedetti et al. 1996), as shown by comparing the signals of the probe 3905145 (that represents the average intensity of signals from probes hybridizing all exons of full-length TG2 mRNA) and those recognising the intronic sequences present in the altered variants (3905171 and 3905172, selective for TGH2, 3905161 for TGH, 3905147 and 3905148 for *tTGv1* and *tTGv2* mRNAs). All these signals increased during treatment with ATRA, in contrast to the 3905144 probe specific for an intronic sequence that remained substantially untranscribed. The probe 3905153, specific for LEV, was increased after 48h of ATRA treatment, but decreased during persistent exposure to the drug (at 72h) before the decline of TG2-lncRNA observed at 96h (Fig. 2). Therefore, ATRA stimulates the expression of all *TGM2* transcripts in APL cells, despite the cell line differences possibly arising because the gene is silenced in untreated HL-60 but not in control NB4 cells.

The analysis was further extended to brain tumor cell lines to investigate whether ATRA affects the *TGM2* gene also in this model, by considering SK-N-SH cells. We tested the effects of 3 and 10 μM ATRA at 16h and 48h incubation times. The results reported in Table 3 show that the expression of all *TGM2* transcripts (including TG2-lncRNA and LEV) increased with respect to the untreated control at early culture time (16h) while 3 or 10 μM ATRA reduced the expression especially of the alternatively spliced *tTGv1* and *tTGv2* transcripts with the protracted exposure (48h). Only the expression of TG2-lncRNA was not reduced after 48h of treatment at high concentration (10 μM). Furthermore, values reported in Table 3 showed that the decline in expression is more pronounced in the 3' portion of the gene. SH-SY5Y cells did not reveal a consistent ATRA-mediated induction of *TGM2* (Table 3), probably because the gene was already actively transcribed in untreated SH-SY5Y cells.

GATA3 drives ATRA-dependent expression of LEV, a regulatory ncRNA influencing the levels of the other *TGM2* transcripts

Several studies have analyzed the binding of transcription factors to *TGM2* promoter and other regulatory regions, but only a few reports describe the interactions between different proteins along the gene, including the intronic regions (de Santiago et al. 2017). The involvement of specific transcription factors was detected in the regulation of *TGM2* expression after ATRA stimulation using ChIP-seq information from ENCODE (Table 2). Fig. 3 shows peak calls (regions of enrichment) for the binding of regulatory factors to the *TGM2* gene in addition to RNApol II. The sites of binding are mainly concentrated at the promoter, but some of them are also present at intronic positions critical for the generation of transcription variants. In untreated SK-N-SH cells, ChIP-seq highlights the binding of RXRα complexes to RAREs in the introns 1, 3, 6 and 10 of *TGM2*, where they may potentially interact or interfere with other regulatory factors listed and summarized in Fig. 3. They include activators, such as MAX, and co-repressors recruiting Histone deacetylases (HDACs), along with the insulator CCCTC-binding factor (CTCF), RAD21, or YY1. The pattern of enrichment of these latter factors in ChIP-seq is modified in SK-N-SH cells treated with ATRA (Fig. 3, bottom panel).

It is relevant to call attention to GATA3 because it binds to intron 10 of *TGM2* (downstream the altered polyadenylation site of TGH and really upstream the LEV starting site) with higher affinity than to other regions (Fig. 3, top panel).

Since GATA3 and other transcription factors are associated with putative regions for modulation of *TGM2* expression in neuronal cells, and considering their involvement in the expression of the *TGM2* locus, the association of these proteins with mRNAs and regulatory RNAs has been investigated by RNA immunoprecipitation in native nuclear extracts of NB4 cells (Fig. 4A). RIP experiment performed using an Ab against RNApol II, showed that the treatment with 1 µM ATRA led to an enrichment of complexes associated with LEV in particular, and, in contrast, to a decreased of those with full-length TG2 and TGH transcripts. RIP performed using Abs against the Max factor allowed only the recovery of the regulatory TG2-lncRNA and LEV. Finally, the use of Abs against RXRα enriched the complexes with TG2-lncRNA, TGH and LEV transcripts, in contrast with the decline of full-length TG2 associated to GATA3, which appeared to bind mainly LEV after ATRA administration (Fig. 4A). It must be underlined that the TGH isoform (Fig. 4B) and LEV ncRNA were observed only upon ATRA treatment. Moreover, while RXRα was negatively modulated by ATRA, as showed by Western blots of Fig. 4B, GATA3 was equally detected in nuclear extracts of treated and untreated cells.

Despite the possibility that LEV could derive from the processing of primary unspliced RNA, the ENCODE data show that the region upstream LEV start site includes sequences recognized by RNApol II and other factors. In the sequence upstream of LEV a H3K27Ac mark is present with feature of regions undergoing high transcriptional activity (as shown in Fig. 3), as elsewhere within *TGM2*, or close to the region of TG2-lncRNA (Minotti et al. 2018, Creyghton et al. 2010). RT-qPCR amplification using oligo-dT primers proved that the LEV transcript is polyadenylated (data not shown). In addition, no short peptides of 150 AAs corresponding to possible LEV transcripts was detected, using an Ab against the epitope V^{604}AEVSLQNP^{612} of the C-terminus of the protein (Ab A037).

These observations demonstrate specific transcription factors binding to ncRNAs of *TGM2*. Since ncRNAs and altered gene transcripts are susceptible to degradation by nonsense-mediated mRNA decay (NMD), RIP assays were performed using an against the phosphorylated ATP-dependent RNA helicase phUPF1, to distinguish cellular targets of NMD processing. In this context the regulatory ncRNAs of *TGM2* could contribute to the accumulation of altered variants, by interfering with their degradation systems. In the presence of ATRA, the levels of full-length TG2 and TGH declined in the complexes with phUPF1, whereas LEV was transcribed and appeared to stringly bind phUPF1 (Fig. 4A).

Since LEV is mainly constituted by the 3'UTR of TGM2 and hypothesizing its protective role against the degradation of the main transcripts, the effects of a AS-LEV adjacent to the start site within intron 10 were studied in HL-60 cells in the presence of ATRA. In these cells the transcription factors involved are expressed at different levels than in NB4 cells, as shown by the western blot (Fig. 5A) and this is mirrored by a different time course of transcripts accumulation, notably the full-length TG2 (Fig. 5B). The treatment of HL-60 cells with AS-LEV or scrambled AS-neg oligonucleotide for 24h in the presence of 1 µM ATRA (antisense oligonucleotides were added 8h after ATRA addition), induced a significant decrease of TG2-lncRNA correlated to LEV transcript after 24h (Fig. 5C).

Predicted structure and regulatory properties of TG2-lncRNA

The sequence corresponding to the TG2-lncRNA was analyzed using Vienna (http://rna.tbi.univie.ac.at/), which has predicted the secondary RNA structure as reported in Supplementary Fig. 1. RNA fold was defined using two different algorithms, one considering Minor Free-Energy, the other based on Centroid approach. An extended region (in red, Supplementary Fig. 1) showed high probability of pairing with each analysis method. Additional analysis in Genome Browser revealed that this transcript is constituted by 2 exons, the first containing an Inverted Repeat IR0 motif likely binding retinoic acid receptor α (RARα) and RARγ (Jolma et al. 2013). This consensus sequence can be recognized also by PML-RARα and its heterodimers having low restriction sequence selectivity (Hosoda et al. 2015). The existence of rs2076380 SNP in this site is also reported (Supplementary Table 1), witnessing a potential regulative role (Jolma et al. 2013). The TG2-lncRNA sequence contains additional elements for binding to transcription factors, which are modified after ATRA treatment, as evidenced by enriched ChIP-sequencing on SK-N-SH cells reported in Fig. 3. The probability of interaction between the TG2-lncRNA and these proteins has been investigated using RPISeq web server (http://pridb.gdcb.iastate.edu/RPISeq/) selecting two prediction methods, the Random Forest and the Support Vector Machine classifier. The data obtained indicated that the probabilities of interaction are respectively 0.85 and 0.6 with GATA3, 0.85 and 0.76 with RXRα, 0.85 and 0.84 with specific RARα, but were also elevated with YY1, 0.80 and 0.85, with CTCF, 0.75 and 0.79, and with RAD21, 0.85 and 0.84 (>0.5 positive threshold). These values are in agreement with our data obtained from RIP experiments, in which TG2-lncRNA associated with RNApol II, RXRα and GATA3 overall after ATRA treatment (Fig. 4), although RARα could be a better target than RXRα.

Notably, 77 nucleotides in the second exon of the TG2-lncRNA (CAAGGTTTATGAGATTTGTTTTCCCCCAAGACC CCCACCCATCTGTCACCCTGACCTGCCCTGGTCCTCTGCTCCA) are identical to another sequence in the intron spacing the two exons, close to the splicing site, and characterized by consensus sequence for TFIID, RXRα and YY1.

Other portions of its long sequence predict a role in interacting with nucleic acid. Indeed, it contains SINE motifs (Short Interspersed Nuclear Elements, which include ALU motifs) a Repeat Mirc-family located in the first exon and frequently recognized by YY1 polycomb proteins, acting as transcriptional repressor interacting with histone deacetylases/

acetyltransferases (Usmanova et al. 2008, Smit et al. 1996-2010, RepeatMasker Open-3.0). These regions are disseminated with consensus sequences for both RAR/RXR and other factors capable of interacting with transcriptional factors, some of which are also present on the gene promoter, in the introns along the gene and just upstream of the LEV sequence. Further, a LINE motif (Long Interspersed Nuclear Elements) characterizes a portion of the second exon with feature of L1MC5 (Jurka 2000), which is complementary to intron 10 and possibly affects the starting region of LEV.

With the aim to investigate a potential activity of the TG2-lncRNA in the control of gene expression we used an antisense molecule (AS-lncRNA, see Material and Methods), targeting the TG2-lncRNA in NB4 cells stimulated with ATRA, in order to have a detectable but not too high expression model, as described for HL-60 cells. Under these conditions AS-lncRNA did not reduce the levels of TG2-lncRNA, but promoted up-regulation of full-length TG2 and LEV mRNA after 24h of treatment and restrained further TG2-lncRNA increase after 48h (Fig. 6, as referred to activity of an AS-neg molecule used as negative control).

Discussion

Transglutaminase type 2 is a complex protein with multiple catalytic activities that, under the control of allosteric effectors and cell reducing potential, drive opposite actions on cell death and survival (Griffin et al. 2002). The existence of variant isoforms preferentially displaying one of the activities (protein crosslinkage or membrane signalling) adds further complexity (Lai and Greenberg 2013). One of our recent analysis of gene expression has highlighted intricate aspects also at this level, since untranslated RNAs with regulatory functions can interfere with the above mentioned processes (Bianchi et al. 2018). In the present report, we investigated the expression of *TGM2* gene in cancer cell lines responsive to ATRA in terms of variants, kinetics, regulatory/functional roles of the ncRNAs, and involvement of transcription factors. Full-length TG2, two variants alternatively spliced at the GTP binding region, the TGH truncated form (lacking almost completely the forth domain) and a very short variant lacking domains 3 and 4, beside the two additional regulatory ncRNAs were examined.

Some of the analyzed cell lines, notably HL-60 with silenced *TGM2*, required ATRA to induce all transcripts while others, like NB4, transcribed the gene in the absence of the drug which is essential for the expression of TGH and of the regulatory ncRNAs. In neuroblastoma SK-N-SH cells that also express *TGM2* in the absence of ATRA, the induction was detected early after treatment with a decline at high concentrations or after prolonged exposure. Finally, ATRA did not produce any increase over basal conditions in SH-SY5Y cells, where the enzyme persistence has been associated with tumor invasiveness (Lee et al. 2012).

In the kinetics of changes triggered by ATRA, the induction of *TGM2* is a relatively early event tightly coupled with the synthesis of the TG2-lncRNA, potentially due to the direct effects of retinoid receptors at the level of the promoter and of the regulatory regions in the TG2-lncRNA sequence.

The induction of the altered transcripts occurred despite the decline of the RAR/RXR complexes, indicating that other factors can sustain the response. During protracted observations an appreciable decline of transcription was detected for LEV, of which induction was strongly determined by ATRA, occurring only after treatment and declining first among all *TGM2* transcripts. In this perspective, the switch-off of transcription involves LEV before other transcripts proceeding back from the 3' end of the gene, as it has been reported also for other genes undergoing temporal transcriptional control in the TNFα/NF-kB network (Wada et al. 2009). In SK-N-SH cells, ATRA apparently drives a "structural closure" of the locus inhibiting transcription, leading mostly to accumulation of TG2-lncRNA.

The gene sensitivity to retinoids is under the control of transcription factors with opposite roles of distinct RAR/RXR complexes (Girardi et al. 2019). Indeed, ATRA decreases RXRα but not RARα in HL-60 and in NB4 cells, in which however the PML-RARα fusion protein modulates retinoid effects (Jensen et al. 2015, Mikesch et al. 2010). In contrast, in neuronal SK-N-SH and SH-SY5Y cells, ATRA down-regulate both RARα and RXRα with concomitant up-regulation of RARβ (Girardi et al. 2019, Nagai et al. 2004). Since RXR content is different in the analyzed cell lines, it can be hypothesized that the different patterns of ATRA-mediated *TGM2* induction to the combined effects of regulatory RNAs and transcription factors on the reactivity of RREs. Notably RAREs arise through aggregation of hexameric units 5′-(A/G)G(G/T)TCA-3′, assembled in different ways as dimeric palindromes (Inverted or Everted Repeats) to form families of canonical and atypical RAREs, differently recognised by RAR, RXR and PML-RARα (Kamashev et al. 2004, Saeed et al. 2012). RAREs are located in the promoter, in intronic regions along the *TGM2* gene and in the TG2-lncRNA, likely in relation to pausing of transcriptional complexes during elongation (Wada et al. 2009), so that the different behaviours of these complexes might be the keys to understand the different expression of *TGM2* transcripts.

We interpreted our results in the frame of the model of Wada et al. (2009) suggesting that the progression of RNApol II is hindered by interposition of regulatory factors. The responsive motives occur at the promoter level, involving sites of RAR and RXR within transcriptional control sites of the ncRNAs and highly represented also at intronic regions. Thus, the rate of elongation would be slowed down by occupancy of these sites, leading either to a complete stop with production of truncated transcripts or to the selection of alternative sites of splicing and polyadenylation. The efficiency of transcription would require the removal of loops formed by repressors and regulatory complexes conditioning the stalling sites of the RNApol II complex along the template strand (Liu et al 2013). Interesting experiments by Daniel et al. shed light on the RXR-mediated response in murine macrophages defining active binding sites, transcriptional changes and

chromatin structure loops (Daniel et al. 2014). These regions interact with intergenic and intronic motives involving RXR and PU.1 on one side, CTCF and RAD21 on the other. These factors are critical to the genomic regulation because RAD21/CTCF binding sites, flanking the pre-initiation complex, are enrolled in the structural remodelling of the chromatin, and RXRs bind regions associated with PU.1, a repressor of *TGM2* promoter activity (Jensen et al. 2015). Upon RA activation, RXR complexes recruit p300 and leave PU.1 giving start to transcription and producing enhancer RNAs at the *TGM2* locus. These same regions can interact with other transcription factors and enhancer RNA molecules to modulate expression of *TGM2* (Sándor et al. 2016), as do the RAD21 and CTCF insulators (Wada et al. 2009). In untreated cells, RNApol II bound target genes at sites overlapping the active chromatin marker H3K4me3 and RAD21/CTCF sites, as major checkpoints downstream of the transcription start site which must be progressed through during elongation along the gene. CTCF forms complexes also with Sin3A to inhibit MYC activity by recruiting HDACs and interacting with YY1 (Zlatonova and Caiafa 2009), presenting distinctive functions either as transcriptional activator or repressor (He and Casaccia-Bonefil 2008, Kleiman et al. 2016). ChIP-seq data on SK-N-SH cells showed that the binding sites of factors listed in Fig. 3 disappear in the presence of ATRA, while the binding of RAD21 increases at the overlapping sites. After ATRA treatment, peaks of enrichment are evident for YY1, CTCF, RAD21 and p300 at the beginning of *TGM2* within the first intron through a mechanism of closure at the DNA sequence corresponding to the TG2-lncRNA.

We have analyzed the functional roles of TG2-lncRNA in NB4 cells by employing AS-lncRNA. It is fair to underscore that the silencing assay required ATRA treatment because only under this condition TG2-lncRNA is highly expressed. AS-lncRNA did not show a distinct antisense efficacy, because it increased, rather than reduced, the levels of transcripts. This unexpected result can possibly be explained either by the removal of an inhibitory effect or by a simple enhancement of expression due to the capacity of AS-lncRNA sequences to form duplexes with decoy activity. In any case, TG2-lncRNA was recovered in complexes with transcription factors by RIP experiments.

ChIP-seq data on untreated SK-N-SH cells highlighted also the binding of other important nuclear factors to intronic regions, where they may drive the generation of LEV. Among them GATA3 shows a large enrichment at intron 10, with features of promoter, upstream of the LEV transcriptional start. RIP experiments using Abs against GATA3 demonstrated its involvement in the transcriptional complex driving the overall expression of LEV under ATRA control. Interestingly, the appearance of LEV and TGH were associated with response to ATRA also in NB4 cells. It' is important to draw attention to the levels of GATA3 and RXRα: GATA3 was constant in nuclear extracts of NB4 cells treated with the drug, declined in SK-N-SH (Hosoda et al. 2015) and was highly expressed in SH-SY5Y cells (Peng et al. 2015), while RXRα declined during the first day of treatment in SH-SY5Y cells (Girardi et al. 2019), as well as decreased in NB4 cells after 48h of ATRA stimulation. This trend shows an association with *TGM2* expression.

We assume that the presence of specific factors at intronic checkpoints may facilitate the alternative use of poly(A) and splicing sites (Colgan and Manley 2018, Curran et al. 2018), giving rise to a heterogeneity of mRNAs, some of them with differences in the 3′ portion, while others with features of ncRNAs which can selectivity interfere with degradation processes (Smith and Baker 2015, Kurosaki et al. 2018). Usually, NMD proceeds to the quick degradation of abnormal RNA preventing the deleterious accumulation of truncated isoforms, which might potentially escape degradation by readdressing NMD to ncRNAs. In this context, RIP experiments using Abs against phUPF1, the component of NMD responsible for active degradation of targets (Luesink et al. 2009), demonstrated the recovery of LEV upon ATRA treatment with a simultaneous decrease of binding to other transcripts and accumulation of TG2 and TGH proteins in the cells, further controlled by TG2 ligands (Bergamini et al. 2007). Therefore, LEV might represent a strategy to elude the surveillance of NMD versus truncated variants, which are translated into specific proteins. Furthermore, the effects of antisense molecule against LEV suggests also a relationship with TG2-lncRNA. We did not uncover a peptide with a molecular weight consistent with the length of the LEV transcript, supporting the hypothesis that the polyadenylated LEV transcript, which we have herein described for the first time, plays a truly regulatory role. A hypothesis for a potential role of LEV as a mRNA-masking microRNA catcher at its 3'UTR, saving the full-length TG2 transcript from degradation by RNA-induced Silencing Complex, also deserves further studies.

Only recently it has been highlighted that degradation by NMD is able to generate RNA fragments complementary to genomic sequences which can interact with factors, as COMPASS complex and modulate gene expression by interfering with H3K4me3 mark regions (Wilkinson 2019), such as those present at intronic sequences flanking both TG2-lncRNA and LEV. From this point of view, both the NMD (Lareau et al. 2007) and the cooperation between transcriptional factors and regulatory RNAs appear crucial for the efficient transcription of *TGM2* and for the generation/accumulation of variants with dramatically different catalytic activities, in particular for the involvement of this enzyme in cancer and drug sensitivity (Kerr et al. 2017).

Author Contributions

Methodology and analysis, O.F., F.C., G.A. C.C., and N.B.; Biostatistical analysis: L.M.; devising of the study, C.M.B. and N.B.; writing, O.F., C.F., C.M.B. and N.B.; supervision, S.V., C.M.B., N.B.; funding acquisition, C.M.B. and S.V.

Funding

The research was supported by grants FIR2017 (Unife) and FAR2018 (Unife) to C.M.B. and by Associazione Italiana Ricerca sul Cancro AIRC (IG 17063) to S.V.

Conflict of interest

The authors declare that they have no conflict of interest.

Ethical statement

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Electronic supplementary material

The online version of this article contains supplementary material, which is available to authorized users.

References

Benedetti, L, Grignani F, Scicchitano BM, Jetten AM, Diverio D, Lo Coco F, Avvisati G, Gambacorti-Passerini C, Adamo S, Levin AA, Pelicci PG, Nervi C (1996) Retinoid-induced differentiation of acute promyelocytic leukemia involves PML-RARalpha-mediated increase of type II transglutaminase. Blood 87:1939-1950. PMID: 8634442

Beninati S, Piacentini M, Bergamini CM (2017) Transglutaminase 2, a double face enzyme. Amino Acids 49:415-423. https://10.1007/s00726-017-2394-5

Bergamini CM (2007) Effects of ligands on the stability of tissue transglutaminase: studies in vitro suggest possible modulation by ligands of protein turn-over in vivo. Amino Acids 33, 415-421. https://10.1007/s00726-006-0457-0

Bianchi N, Beninati S, Bergamini CM (2018) Spotlight on the transglutaminase 2 gene: a focus on genomic and transcriptional aspects. Biochem J 475:1643-1667. https://10.1042/BCJ20170601

Breitman TR, Selonick SE, Collins SJ (1980) Induction of differentiation of the human promyelocytic leukemia cell line (HL-60) by retinoic acid. Proc Natl Acad Sci USA 77:2936-2940. PMID: 6930676

Brun ES (2012) Thesis for the degree Master in Molecular Biotechnology MBI-3941, Dissertation, University of Tromsø, Norway.

Colgan DF, Manley JL (1997) Mechanism and regulation of mRNA polyadenylation. Genes Dev 11:2755-2766. https://10.1101/gad.11.21.2755

Comandini A, Naro C, Adamo R, Akbar AN, Lanna A, Bonmassar E, Franzese O (2013) Molecular mechanisms involved in HIV-1-Tat mediated inhibition of telomerase activity in human CD4(+) T lymphocytes. Mol Immunol 54:181-192. https://10.1016/j.molimm.2012.12.003

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA, Jaenisch R (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci USA 107:21931-21936. https://10.1073/pnas.1016071107

Curran E, Wang H, Hinds TR, Zheng N, Wang EH (2018) Zinc knuckle of TAF1 is a DNA binding module critical for TFIID promoter occupancy. Sci Rep 8:4630-4643. https://10.1038/s41598-018-22879-5

Daniel B, Nagy G, Hah N, Horvath A, Czimmerer Z, Poliska S, Gyuris T, Keirsse J, Gysemans C, Van Ginderachter JA, Balint BL, Evans RM, Barta E, Nagy L (2014) The active enhancer network operated by liganded RXR supports angiogenic activity in macrophages. Genes Dev 28:1562-1577. https://10.1101/gad.242685.114

de Kok JB, Roelofs RW, Giesendorf BA, Pennings JL, Waas ET, Feuth T, Swinkels DW, Span PN (2005) Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes. Lab Invest 85:154-159. https://.1038/labinvest.3700208

de Santiago I, Liu W, Yuan K, O'Reilly M, Chilamakuri CS, Ponder BA, Meyer KB, Markowetz F (2017) BaalChIP: Bayesian analysis of allele-specific transcription factor binding in cancer genomes. Genome Bio 18:39-56. https://10.1186/s13059-017-1165-7

de Stephanis L, Mangolini A, Servello M, Harris PC, Dell'Atti L, Pinton P, Aguiari G (2018) MicroRNA501-5p induces p53 proteasome degradation through the activation of the mTOR/MDM2 pathway in ADPKD cells. J Cell Physiol 233, 6911-6924. https://10.1002/jcp.26473

Fésüs L, Szondy Z (2005) Transglutaminase 2 in the balance of cell death and survival. FEBS Lett 579:3297-3302. https://10.1016/j.febslet.2005.03.063

Fraij BM, Gonzales RA (1996) A third human tissue transglutaminase homologue as a result of alternative gene transcripts. Biochim Biophys Acta 1306:63-74. https://10.1016/0167-4781(95)00219-7

Girardi CS, Rostirolla DC, Lini FJM, Brum PO, Delgado J, Ribeiro CT, Teixeira AA, Peixoto DO, Heimfarth L, Kunzler A, Moreira JCF, Gelain DP (2019) Nuclear RXRα and RXRβ receptors exert distinct and opposite effects on RA-mediated neuroblastoma differentiation. Biochim Biophys Acta Mol Cell Res 1866:317-328. https://10.1016/j.bbamcr.2018.11.014

Griffin M, Casadio R, Bergamini CM (2002) Transglutaminases: nature's biological glues. Biochem J 368:377-396. https://10.1042/BJ20021234

He Y, Casaccia-Bonnefil P (2008) The Yin and Yang of YY1 in the nervous system. J Neurochem 106:1493-1502. https://10.1111/j.1471-4159.2008.05486.x

Hosoda K, Sato M, Yanai K (2015) Identification and characterization of human genomic binding sites for Retinoic Acid Receptor/Retinoid X Receptor heterodimers. Advances in Biological Chemistry 5:58-72. https://10.4236/abc.2015.52006

Jensen HA, Yourish HB, Bunaciu RP, Varner JD, Yen A (2015) Induced myelomonocytic differentiation in leukemia cells is accompanied by noncanonical transcription factor expression. FEBS Open Bio 5:789-800. https://10.1016/j.fob.2015.09.008

Jolma A, Yan J, Whitington T, Toivonen J, Nitta KR, Rastas P, Morgunova E, Enge M, Taipale M, Wei G, Palin K, Vaquerizas JM, Vincentelli R, Luscombe NM, Hughes TR, Lemaire P, Ukkonen E, Kivioja T, Taipale J (2013) DNA-binding specificities of human transcription factors. Cell 152:327-339. 1 Doi: 0.1016/j.cell.2012.12.009

Jurka J (2000) Repbase Update: a database and an electronic journal of repetitive elements. Trends Genet 16:418-420. PMID: 10973072

Kamashev D, Vitoux D, De Thé H (2004) PML-RARA-RXR oligomers mediate retinoid and rexinoid/cAMP cross-talk in acute promyelocytic leukemia cell differentiation. J Exp Med 199:1163-1174. https://10.1084/jem.20032226

Kerr C, Szmacinski H, Fisher ML, Nance B, Lakowicz JR, Akbar A, Keillor JW, Lok Wong T, Godoy-Ruiz R, Toth EA, Weber DJ, Eckert RL (2017) Transamidase site-targeted agents alter the conformation of the transglutaminase cancer stem cell survival protein to reduce GTP binding activity and cancer stem cell survival. Oncogene 36:2981–2990. https://10.1038/onc.2016.452

Király R, Demény M, Fésüs L (2011) Protein transamidation by transglutaminase 2 in cells: a disputed Ca2+-dependent action of a multifunctional protein. FEBS J 278:4717-4739. https://10.1111/j.1742-4658.2011.08345.x

Kleiman E, Jia H, Loguercio S, Su AI, Feeney AJ (2016) YY1 plays an essential role at all stages of B-cell differentiation. Proc Natl Acad Sci USA 113:E3911-E3920. https://10.1073/pnas.1606297113

Kovalevich J, Langford D (2013) Considerations for the use of SH-SY5Y neuroblastoma cells in neurobiology. Methods Mol Biol 1078:9-21. https://10.1007/978-1-62703-640-5_2

Kurosaki T, Hoque M, Maquat LE (2018) Identifying cellular Nonsense-Mediated mRNA Decay (NMD) targets: immunoprecipitation of phosphorylated UPF1 followed by RNA sequencing (p-UPF1 RIP-Seq). Methods Mol Biol 1720:175-186. https://10.1007/978-1-4939-7540-2_13

Lai TS, Greenberg CS (2013) TGM2 and implications for human disease: role of alternative splicing. Front Biosci (Landmark Ed) 18:504-519. PMID: 23276939

Lareau LF, Brooks AN, Soergel DAW, Meng Q, Brenner SE (2007) The Coupling of Alternative Splicing and Nonsense-Mediated mRNA Decay. Chapter 12. In: Benjamin J. Blencowe and Brenton R. Graveley (ed) Alternative Splicing in the Postgenomic Era. Landes Bioscience and Springer Science+Business Media. https://escholarship.org/content/qt6089j8mq/qt6089j8mq.pdf

Lee HJ, Park MK, Bae HC, Yoon HJ, Kim SY, Lee CH (2012) Transglutaminase-2 is involved in all-trans retinoic acid-induced invasion and matrix metalloproteinases expression of SH-SY5Y neuroblastoma cells via NF-κB pathway. Biomol Ther (Seoul) 20:286-292. https://10.4062/biomolther.2012.20.3.286

Liu X, Bushnell DA, Kornberg RD (2013) RNA Polymerase II Transcription: structure and mechanism. Biochim Biophys Acta 1829:2–8. https://10.1016/j.bbagrm.2012.09.003

Luesink M, Pennings JL, Wissink WM, Linssen PC, Muus P, Pfundt R, de Witte TJ van der Reijden BA, Jansen JH (2009) Chemokine induction by all-trans retinoic acid and arsenic trioxide in acute promyelocytic leukemia: triggering the differentiation syndrome. Blood 114:5512-5521. https://10.1182/blood-2009-02-204834

Masetti R, Biagi C, Zama D, Vendemini F, Martoni A, Morello W, Gasperini P, Pession A (2012) Retinoids in pediatric onco-hematology: the model of acute promyelocytic leukemia and neuroblastoma. Adv Ther 29:4747-4762. https://10.1007/s12325-012-0047-3

Mikesch JH, Gronemeyer H, So CW (2010) Discovery of novel transcriptional and epigenetic targets in APL by global ChIP analyses: emerging opportunity and challenge. Cancer Cell 17:112-114. https://10.1016/j.ccr.2010.01.012

Minotti L, Baldassari F, Galasso M, Volinia S, Bergamini CM, Bianchi N (2018) A long non-coding RNA inside the type 2 transglutaminase gene tightly correlates with the expression of its transcriptional variants. Amino Acids 50:421-438. https://10.1007/s00726-017-2528-9

Nagai J, Yazawa T, Okudela K, Kigasawa H, Kitamura H, Osaka H (2004) Retinoic acid induces neuroblastoma cell death by inhibiting proteasomal degradation of retinoic acid receptor alpha. Cancer Res 64:7910-7917. https://10.1158/0008-5472.CAN-04-1178

Peng H, Ke XX, Hu R, Yang L, Cui H, Wei Y (2015) Essential role of GATA3 in regulation of differentiation and cell proliferation in SK-N-SH neuroblastoma cells. Mol Med Rep 11:881-886. https://10.3892/mmr.2014.2809

Phatak VM, Croft SM, Rameshaiah Setty SG, Scarpellini A, Hughes DC, Rees R, McArdle S, Verderio EA (2013) Expression of transglutaminase-2 isoforms in normal human tissues and cancer cell lines: dysregulation of alternative splicing in cancer. Amino Acids 44:33-44. https://10.1007/s00726-011-1127-4

Redner RL, Beumer JH, Kropf P, Agha M, Boyiadzis M, Dorritie K, Farah RJZ, Hou A, Im SH, Lim A, Raptis A, Sehgal SM, Christner D, Normolle Johnson DE (2018) A phase-1 study of dasatinib plus all-trans retinoic acid in acute myeloid leukemia. Leuk Lymphoma 59:2595-2601. https://10.1080/10428194.2018.1443330

Saeed S, Logie C, Francoijs KJ, Frigè G, Romanenghi M, Nielsen FG, Raats L, Shahhoseini M, Huynen M, Altucci L, Minucci S, Martens JH, Stunnenber HG (2012) Chromatin accessibility, p300, and histone acetylation define PML-RARα and AML1-ETO binding sites in acute myeloid leukemia. Blood 120:3058-3068. https://10.1182/blood-2011-10-386086

Sandberg R, Neilson JR, Sarma A, Sharp PA & Burge CB (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320:643-647. https://10.1126/science.1155390

Sándor K, Daniel B, Kiss B, Kovács F, Szondy Z (2016) Transcriptional control of transglutaminase 2 expression in mouse apoptotic thymocytes. Biochim Biophys Acta 1859:964-974. https://10.1016/j.bbagrm.2016.05.011

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C(T) method. Nat Protoc 3:1101-1108. PMID: 18546601

Smit AFA, Hubley R, Green P (1996-2010) RepeatMasker Open-3.0. http://www.repeatmasker.org/faq.html

Smith JE, Baker KE (2015) Nonsense-mediated RNA decay--a switch and dial for regulating gene expression. Bioessays 37:612-623. https://10.1002/bies.201500007

Tee AE, Marshall GM, Liu PY, Xu N, Haber M, Norris MD, Iismaa SE, Liu T (2010) Opposing effects of two tissue transglutaminase protein isoforms in neuroblastoma cell differentiation. J Biol Chem 285:3561-3567. https://10.1074/jbc.M109.053041

Usmanova NM, Kazakov VI, Tomilin NV (2008) SINEs in mammalian genomes can serve as additional signals in formation of facultative heterochromatin. Cell and Tissue Biology 2:217–221. https://link.springer.com/article/10.1134/S1990519X08030012

Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S (2009) A wave of nascent transcription on activated human genes. Proc Natl Acad Sci USA 106:18357-18361. https://10.1073/pnas.0902573106

Wilkinson MF (2019) Genetic paradox explained by nonsense. Nature 568:179-180. https://10.1038/d41586-019-00823-5

Zlatanova J, Caiafa P (2009) CTCF and its protein partners: divide and rule? Cell Sci 122:1275-1284. https://10.1242/jcs.039990

Zhang G, Pian C, Chen Z, Zhang J, Xu M, Zhang L, Chen Y (2018) Identification of cancer-related miRNA-lncRNA biomarkers using a basic miRNA-lncRNA network. PLoS One eCollection 13(5):e0196681. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0196681

Figure legends

Fig. 1 a RT-qPCR analysis of the expression of regulatory TG2-lncRNA, LEV and full-length TG2, and **b** of truncated TGH2 and TGH, of alternatively spliced *tTGv1* and *tTGv2* transcripts in HL-60 cells treated with 3 µM ATRA. Total RNA was extracted from samples after 16, 24, 36, 48, 60 and 72h of culture, reverse transcribed and the quantification of targets was performed using HPRT (de Kok et al. 2005) as housekeeping gene. Fold changes were quantified using $2^{-\Delta\Delta CT}$ formula with respect to untreated cells analyzed at the same times of culture. The mean ± SD was determined from n=3 independent experiments. $p<0.05$ for each transcript analyzed with One-way ANOVA.

Fig. 2 Analysis of expression of *TGM2* variants in NB4 cells untreated (white circles) and treated with 1 µM ATRA (triangles) for 48h (black), 72h (grey) or 96h (white). The data were extracted from GSE18397 of GEO database (Luesink et al. 2009). Normalized Intensity of the samples was obtained using probe set from HumanExon1_0ST Affymetrix array distributed along the whole *TGM2* gene. The arrows indicate distinct variants. The probe 3905145 represents the mean of signals derived from all expressed gene sequences.

Fig. 3 ChIP-sequencing from ENCODE database of transcription factors and chromatin structural proteins binding the *TGM2* gene to regulatory intronic regions in untreated SK-N-SH cells (upper box) and treated with 6 µM ATRA (lower box). The positions of the binding factors are reported with respect to the gene, depicted on the top of the figure (canonical full-length TG2). The RNApol II sites and the H3K27Ac marks are also reported. Squares indicated the binding sequences recognized by factors within *TGM2*. Colors (blue, light green, dark green, red, orange, pink, violet, black, yellow, ocher, brown) indicate factors binding the same region in ChIP experiments, while shades intensity (light or dark blue) depends on their binding affinity.

Fig. 4 a RT-qPCR analysis of samples obtained from RIP experiments carried out on NB4 cells untreated or treated with 1 µM for 48h. RIP was performed with native extracts obtained from isolated nuclei using Abs against GATA3, RXRα, Max, RNApol II and phUPF1. Sybrgreen and specific primers were used to amplify TG2-lncRNA (oblique slashes), full-length TG2 (white), TGH mRNA (grey) and LEV transcript (black). Fold Enrichment = $2^{-\Delta\Delta Ct\,[RIP/NS]}$ was calculated (as detailed in Material and methods), where NS is Non-Specific Ab binding, as background and the $\log_{10}$ of the values was plotted. Bars on the right of the histogram plots represent the fold changes ratio of the samples treated with 1 µM ATRA /untreated, where (*) indicates $p<0.05$ (using unpaired t test). **b** Western blotting analysis was carried out on nuclear extracts or on total cellular extracts, under conditions of low expression of the analyzed factor, as indicated. Abs AbA037 and AbA036 were specific for full-length TG2 (as well as *tTGv1* and *tTGv2*), while AbA036 is specific for TGH. Abs directed against Actin was used as internal control to calculate the values of the ratio to specific Ab hybridization densitometry.

Fig. 5 a Evaluation of the content of transcription factors influencing *TGM2* gene expression in HL-60 cells in the absence or in the presence of ATRA 1 µM; the values of the ratio to specific Ab hybridization densitometry was reported. **b** The expression of TG2-lncRNA (oblique slashes), full-length TG2 (white), TGH (grey) and LEV transcripts (black) of HL-60 was compared to that of NB4 cells at the same ATRA concentration. **c** Analysis by RT-qPCR of the effects on *TGM2* transcripts of antisense oligonucleotide against LEV (AS-LEV) at 1.5 µM concentration for 24h in HL-60 ATRA-induced, employing scramble oligonucleotide (AS-neg) as negative control. Fold changes were calculated with respect to AS-neg using $2^{\wedge-\Delta\Delta CT}$ method. Histograms represent the mean ± SD of 3 independent experiments; (*) $p<0.05$ (using unpaired t test, two-tailed).

Fig. 6 a RT-qPCR analysis of the expression of TG2-lncRNA (oblique slashes), full-length (white), TGH (grey) and LEV transcripts (black) in NB4 cells treated with 1 µM ATRA and 1.5 µM of specific AS (AS-lncRNA) against TG2-lncRNA transcript. The treatment with a scramble oligonucleotide (AS-neg) in the presence of the same concentration of ATRA and representing negative control was used to quantify the fold increase in the samples. Fold changes were calculated using $2^{-\Delta\Delta CT}$ method. Histograms represent the mean ± SD of 3 independent experiments; (*) $p<0.05$ (using unpaired t test, two-tailed). **b** Cartoon illustrating the potential interference activity of TG2-lncRNA towards the *TGM2* gene expression. Critical binding sites for transcription factors could act like enhancer/decoy or scaffolding to re-direct them to other regions of the gene.

Supplementary Fig. 1 Drawing describing the secondary RNA structure of the TG2-lncRNA with base-paired regions generated by using Vienna web RNA service considering two approaches for the analysis: **a** to have the structure with less free-energy structure, **b** to employ a model based on centroids. In the colour bar, red corresponds to high and blue to a low probability that the structure is hired.

Click here to access/download;Figure;Figure 1. HL-60 + ATRA.tif

Figure 1

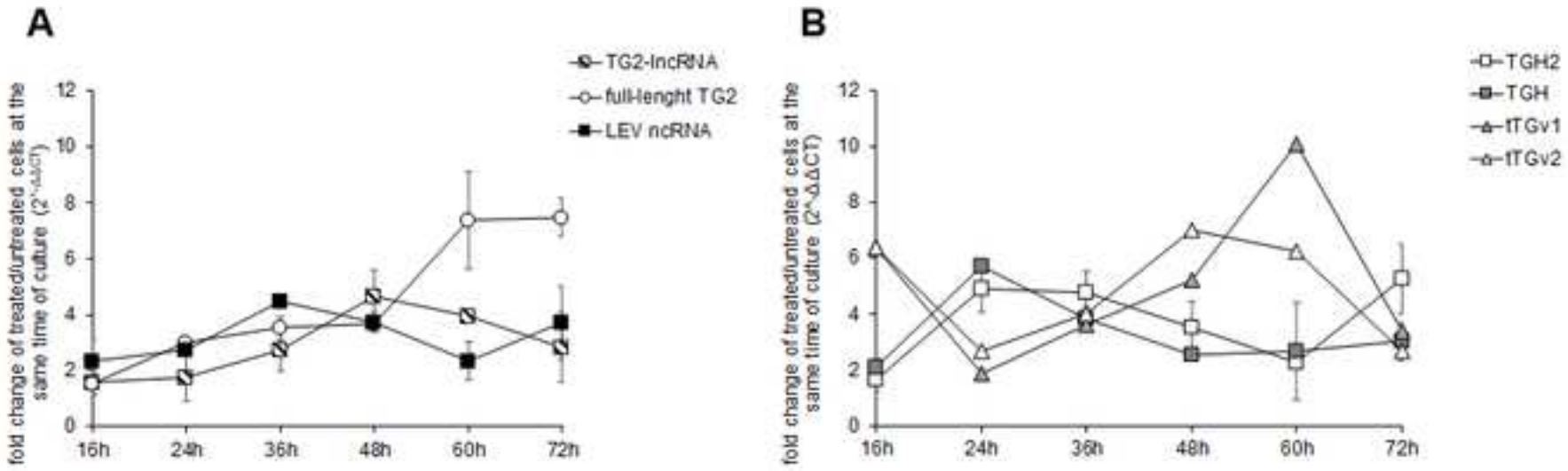
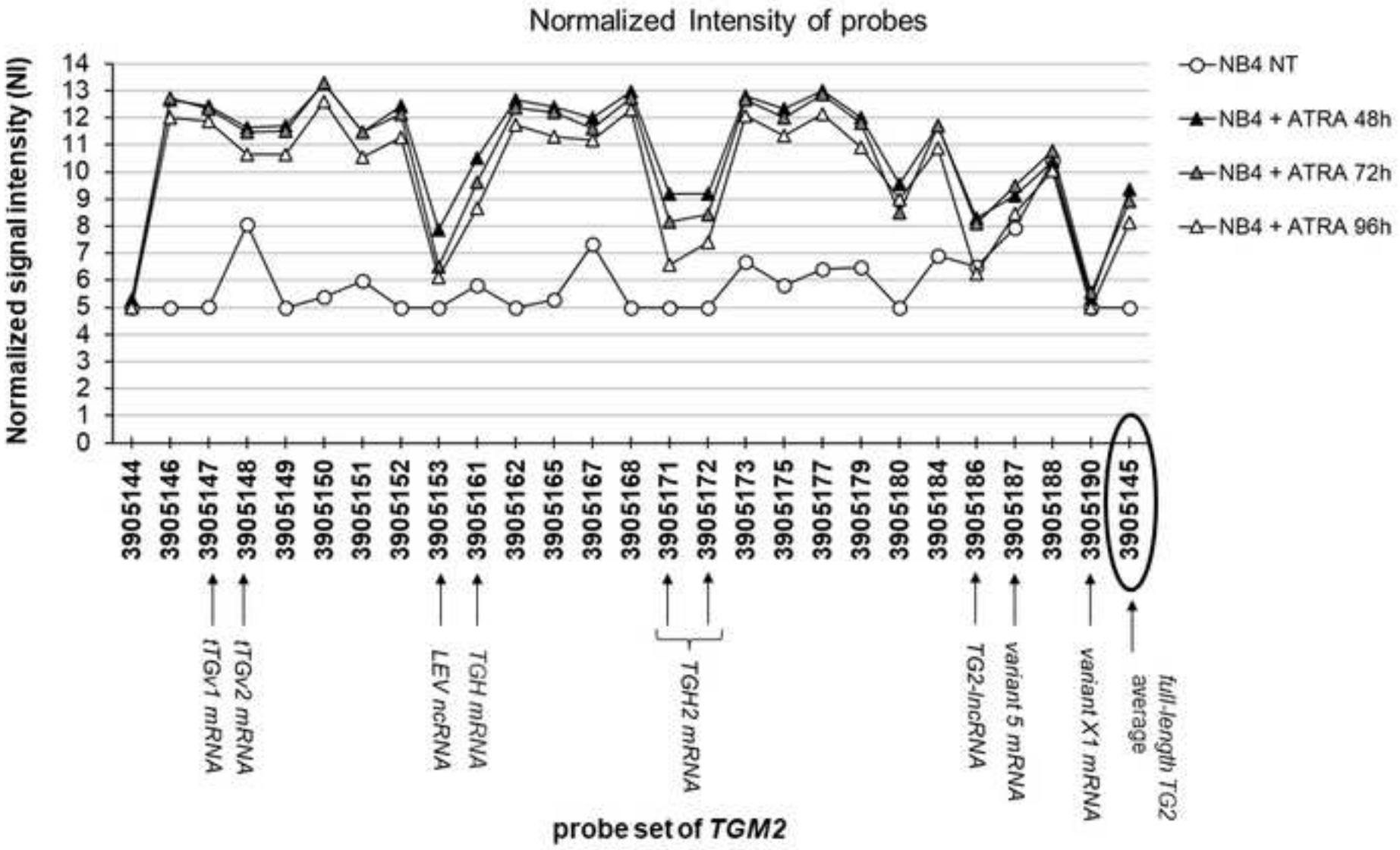


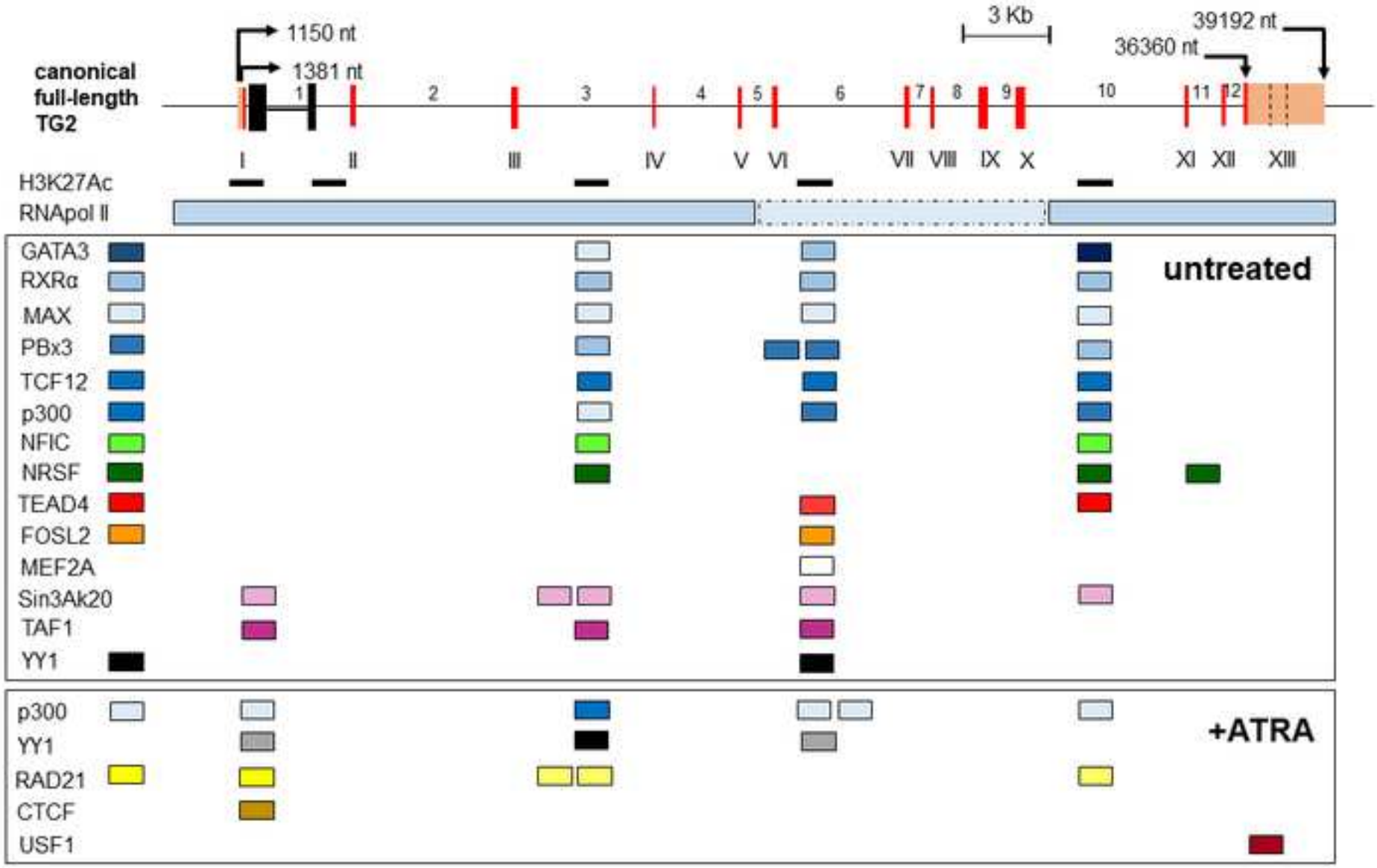
Figure 2

Click here to access/download;Figure;Figure 2 NB4 array.tif



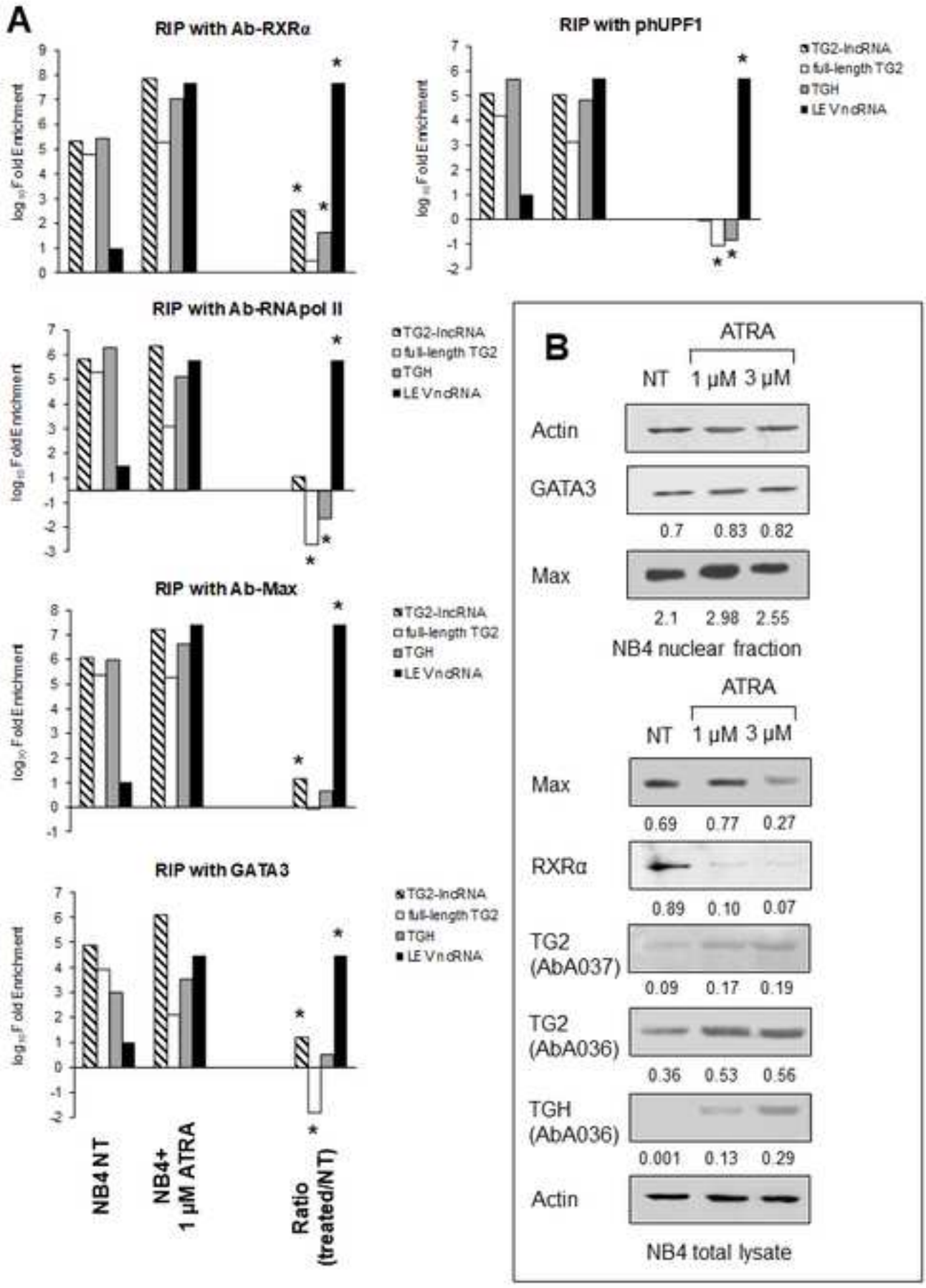
Click here to access/download;Figure;Figure 3 ChIP SK-N-SH.tif

Figure 3



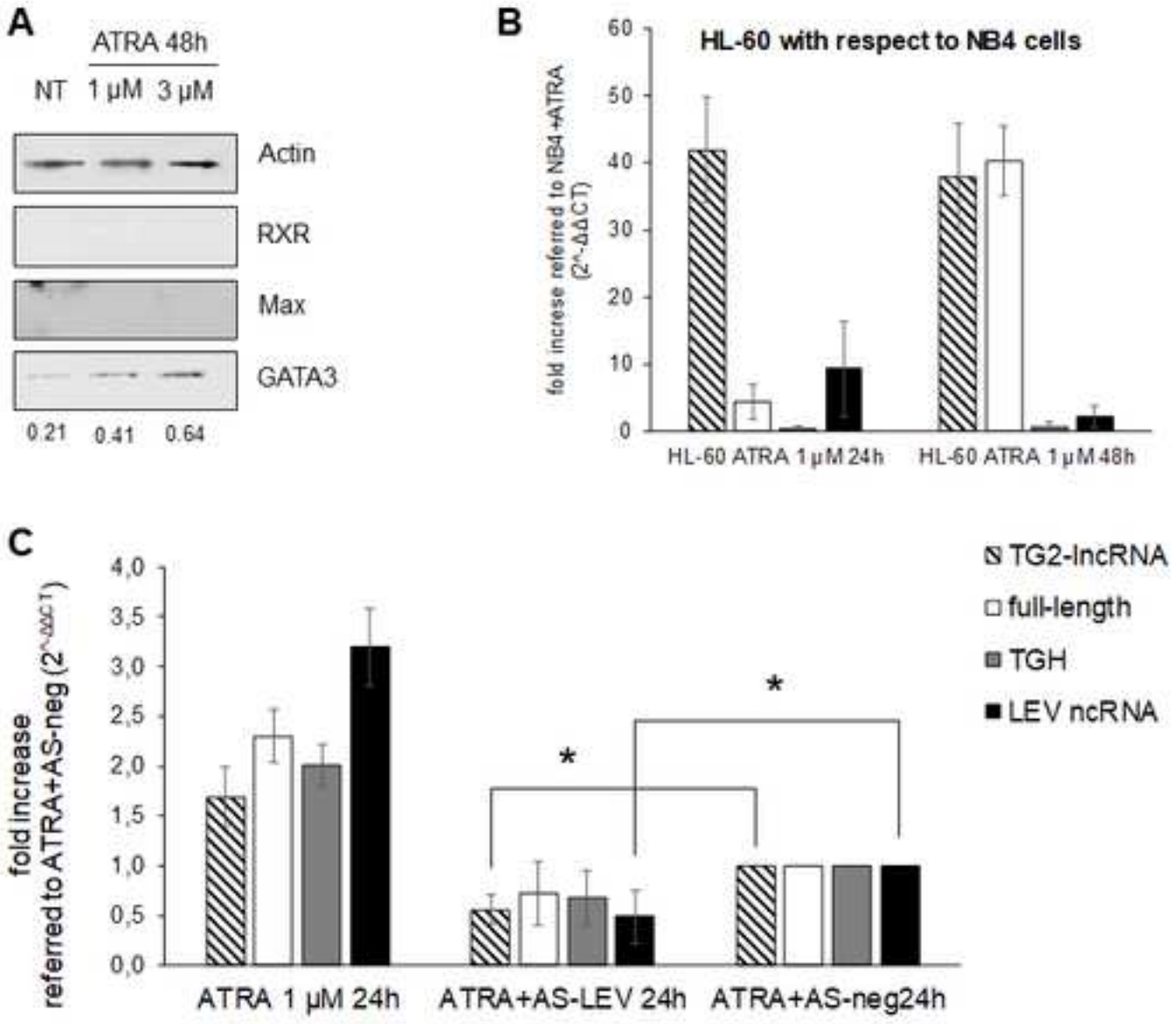
Click here to access/download;Figure;Figure 4 NB4 RIP.tif

Figure 4



Click here to access/download;Figure;Figure 5 HL-60 AS-LEV.tif

Figure 5



Click here to access/download;Figure;Figure 6 NB4 AS-lncRNA.tif

Figure 6

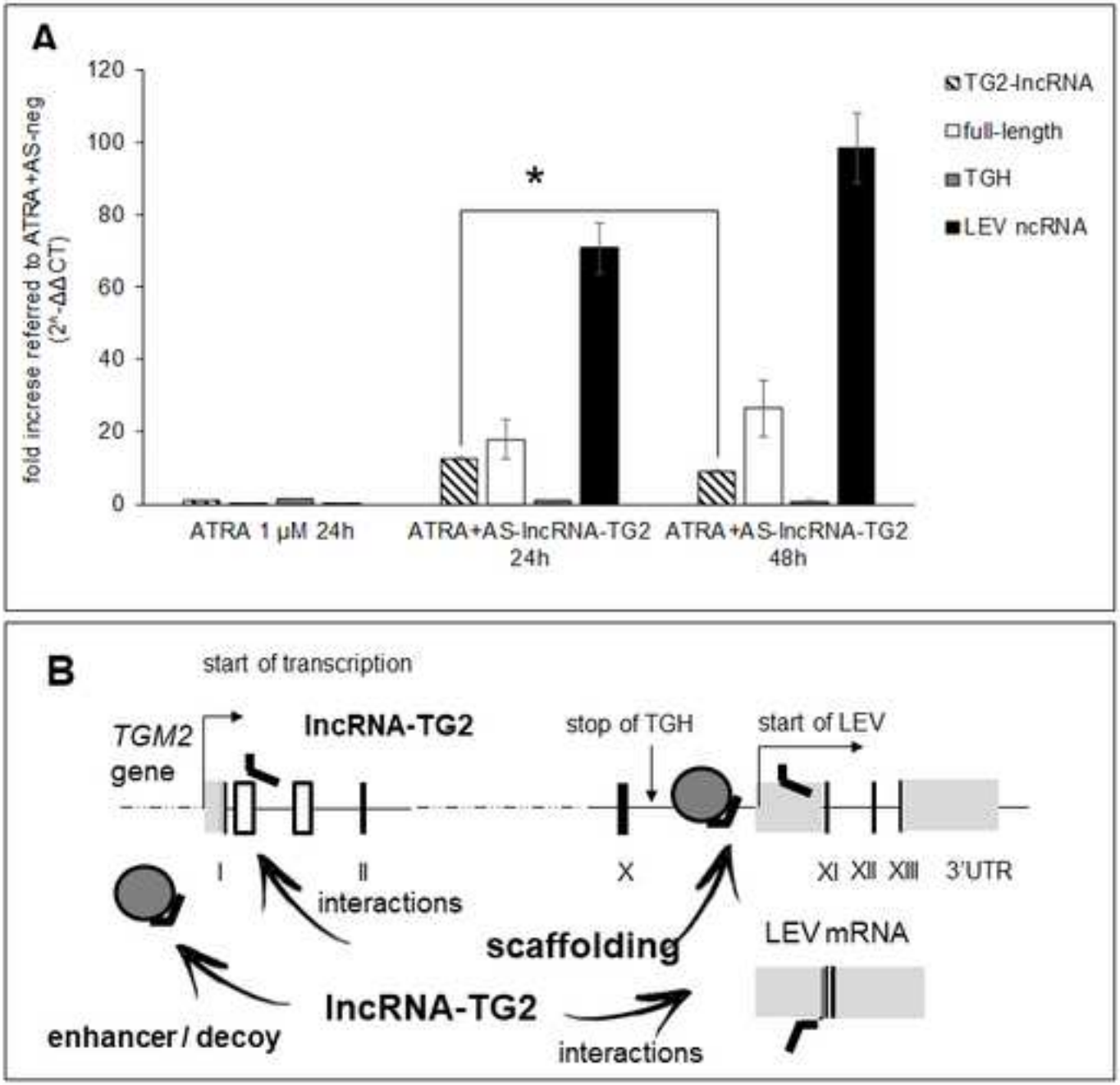


Table 1

Table 1 Description of primers. Sequence, localization, PCR product and conditions used for amplification reactions. The amplification program was 1 cycle 50°C 2 min, 1 cycle 95°C 5 min, and 40 cycles as specifically indicated

Primer name	Primer sequence	Exon recognized	Size of PCR product (bp)	PCR conditions
HPRT1F (Phatak et al. 2013)	5'-TGACACTGGCAAAACAATGCA-3'	exon VI	94	95°C 15''
HPRT1 R (Phatak et al. 2013)	5'-GGTCCTTTTCACCAGCAAGCT-3'	exon VII		60°C 1'
lncRNA F (Minotti et al. 2018)	5'-TCACGTTCTGCAACCCTAGAC-3'	exon I	117	95°C 15''
lncRNA R (Minotti et al. 2018)	5'- GGGGTCTTGGGGGAAAACAAA-3'	exon II		60°C 1'
FTG2ex11 (Phatak et al. 2013)	5'-CCTTACGGAGTCCAACCTCA-3'	exon XI	245	95°C 30''
RTG2ex12 (Phatak et al. 2013)	5'-CCGTCTTCTGCTCCTCAGTC-3'	exon XII		64°C 1'
FTG2ex11 (Phatak et al. 2013)	5'-CCTTACGGAGTCCAACCTCA-3'	exon XI	214	95°C 30''
RtTGv1 (Phatak et al. 2013)	5'-CTGGGATGTGGAGGTGCA-3'	alternative splicing 1 in exon XIII		64°C 1'
FTG2ex11 (Phatak et al. 2013)	5'-CCTTACGGAGTCCAACCTCA-3'	exon XI	214	95°C 30''
RtTGv2 (Phatak et al. 2013)	5'-CACTGGTGTGGAGGTGCAGC-3'	alternative splicing 2 in exon XIII		64°C 1'
F2TGH	5'-TACAAATACCCAGAGGGGTCC-3'	splicing site exon IX-X	338	95°C 30''
R2TGH	5'-ATGTTGTCAGTTGGCGGTCA-3'	intron X		62°C 1'
FTGH2 (Phatak et al. 2013)	5'-GGTGAGTGGCATGGTCAACT-3'	exon V	226	95°C 15''
RTGH2 (Phatak et al. 2013)	5'-AGGGCTCATGACCCACATC-3'	intron VI		65°C 1'
Fshort3ex	5'-TCATGAGCACCCTCCTTTCT-3'	intron IX	377	95°C 30''
Rshort3ex	5'-GGTCCATTCTCACCTTAACTTCCT-3'	exon XIII		62°C 1'

Table 2

Table 2 Samples from GEO database analyzed with HumanExon1_0ST Affymetrix and from ENCODE database of ChIP-seq analysis

Cell line and treatments	Description	Accession of dataset (GEO or ENCODE)	Number of samples	Reference
NB4 cells ± ATRA (1 μM)	Expression profiling of NB4 cells after treatment with ATRA	GSE18397	4 (NT, 48h, 72h, 96h)	(Luesink et al 2009)
SK-N-SH cells ± ATRA (6 μM)	ChIP-sequencing	wgEncodeEH003249 (GATA3) wgEncodeEH003302 (RXRA) wgEncodeEH003299 (MAX) wgEncodeEH003300 (PBX3) wgEncodeEH003227 (TCF12) wgEncodeEH003270 (p300) wgEncodeEH003237 (NFIC) wgEncodeEH002269 (NRSF) wgEncodeEH001547 (NRSF) wgEncodeEH003286 (TEAD4) wgEncodeEH003297 (FOSL2) wgEncodeEH003298 (MEF2A) wgEncodeEH002271 (SIN3A) wgEncodeEH002301 (TAF1) wgEncodeEH003228 (YY1) wgEncodeEH002270 (RNApol II) wgEncodeEH001616 (EP300) wgEncodeEH000439 (CTCF) wgEncodeEH001613 (CTCF) wgEncodeEH001614 (RAD21) wgEncodeEH001615 (YY1) wgEncodeEH001621 (USF1)	2 unreplicated (48h)	(Richard Myers, Hudson Alpha Institute for Biotechnology)

Table 3

Table 3 *TGM2* gene expression in SK-N-SH and SH-SY5Y cell lines exposed by 3 µM and 10 µM ATRA.
Fold change were quantified with respect to untreated control and using HPRT1 as housekeeping gene

Transcript	**SK-N-SH + 3 µM ATRA/NT**		**SK-N-SH + 10 µM ATRA/NT**	**SH-SY5Y + 3 µM ATRA/NT**		**SH-SY5Y + 10 µM ATRA/NT**
	16h	**48h**	**48h**	**16h**	**48h**	**48h**
lncRNA-TG2	4.74 ± 1.00	4.72 ± 1.08	8.00 ± 0.26	1.05 ± 0.43	1.37 ± 0.10	1.20 ± 0.40
TGH2 mRNA	5.52 ± 1.06	3.08 ± 0.47	2.19 ± 0.52	0.58 ± 0.01	0.69 ± 0.04	0.73 ± 0.01
TGH mRNA	6.63 ± 0.18	2.49 ± 1.39	2.39 ± 0.58	0.71 ± 0.27	0.62 ± 0.11	0.70 ± 0.22
LEV	4.45 ± 0.10	1.42 ± 0.35	0.64 ± 0.06	0.65 ± 0.15	0.64 ± 0.10	0.84 ± 0.06
tTGv1 mRNA	4.08 ± 1.05	0.54 ± 0.28	0.16 ± 0.05	1.11 ± 0.28	0.79 ± 0.04	1.78 ± 0.38
tTGv2 mRNA	2.72 ± 0.93	0.41 ± 0.18	0.13 ± 0.08	1.05 ± 0.04	0.56 ± 0.08	1.09 ± 0.14
TG2 mRNA	2.58 ± 0.36	0.71 ± 0.34	0.46 ± 0.13	1.21 ± 0.34	0.73 ± 0.10	1.89 ± 0.11

Supplementary table

Click here to access/download
Supplementary Material
Suppl. Table 1 SNP RAR-RXR.docx.pdf

Supplementary figure

Click here to access/download
Supplementary Material
Suppl. Figure 1. lncRNA-TG2.tif