



OPEN

## EGR1 and RXRA transcription factors link TGF- $\beta$ pathway and CCL2 expression in triple negative breast cancer cells

Alisa M. Gorbacheva<sup>1</sup>, Aksinya N. Uvarova<sup>1</sup>, Alina S. Ustiugova<sup>1</sup>, Arindam Bhattacharyya<sup>2</sup>, Kirill V. Korneev<sup>1</sup>, Dmitry V. Kuprash<sup>1,3</sup> & Nikita A. Mitkin<sup>1</sup>✉

Transforming growth factor beta (TGF- $\beta$ ) is the main cytokine responsible for the induction of the epithelial-mesenchymal transition of breast cancer cells, which is a hallmark of tumor transformation to the metastatic phenotype. Recently, research demonstrated that the chemokine CCL2 gene expression level directly correlates with the TGF- $\beta$  activity in breast cancer patients. CCL2 attracts tumor-associated macrophages and is, therefore, considered as an important inducer of breast cancer progression; however, the precise mechanisms underlying its regulation by TGF- $\beta$  are unknown. Here, we studied the behavior of the CCL2 gene in MDA-MB-231 and HCC1937 breast cancer cells representing mesenchymal-like phenotype activated by TGF- $\beta$ . Using bioinformatics, deletion screening and point mutagenesis, we identified binding sites in the CCL2 promoter and candidate transcription factors responsible for its regulation by TGF- $\beta$ . Among these factors, only the knock-down of EGR1 and RXRA made CCL2 promoter activity independent of TGF- $\beta$ . These factors also demonstrated binding to the CCL2 promoter in a TGF- $\beta$ -dependent manner in a chromatin immunoprecipitation assay, and point mutations in the EGR1 and RXRA binding sites totally abolished the effect of TGF- $\beta$ . Our results highlight the key role of EGR1 and RXRA transcription factors in the regulation of CCL2 gene in response to TGF- $\beta$  pathway.

### Abbreviations

ACTB	Beta-actin
CCL2	C-C motif ligand 2
cDNA	Complementary DNA
CXCR4	C-X-C chemokine receptor type 4
EGR1	Early growth response protein 1
EGR2	Early growth response protein 2
IgG	Immunoglobulin G
IL11	Interleukin 11
MMLV RT	Moloney murine leukemia virus reverse transcriptase
M-MULV	Moloney murine leukemia virus
PAX6	Paired box protein 6
PCNA	Proliferating cell nuclear antigen
RARA	Retinoic acid receptor alpha
RXRA	Retinoid X receptor alpha
SEM	Standard error of the mean
siRNA	Small interfering RNA
SMAD3	Mothers against decapentaplegic homolog 3
SP1	Specificity protein 1
STR analysis	Short tandem repeat analysis

<sup>1</sup>Center for Precision Genome Editing and Genetic Technologies for Biomedicine, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow 119991, Russia. <sup>2</sup>Immunology Laboratory, Department of Zoology, University of Calcutta, 35, Ballygunge Circular Road, Kolkata, West Bengal 700019, India. <sup>3</sup>Biological Faculty, Lomonosov Moscow State University, Moscow 119234, Russia. ✉email: mitkin.n.a@gmail.com

Transforming growth factor beta (TGF- $\beta$ ) is a cytokine that acts as a tumor suppressor in normal conditions and in the early stages of breast cancer (BC) development<sup>1</sup>. BC progression leads to the modification of TGF- $\beta$ -dependent signaling pathways, which results in the altered response of a cancer cells to TGF- $\beta$ , the elevated production of TGF- $\beta$ , and modification of the tumor microenvironment<sup>2</sup>. In particular, enhanced TGF- $\beta$  production is associated with the differentiation of myeloid-derived suppressor cells and regulatory T-helpers, which provoke tumor development by the inhibition of local inflammatory processes<sup>3</sup>. One of the key pathogenic functions of TGF- $\beta$  is the induction of the epithelial-mesenchymal transition that results in the loss of the epithelial phenotype of the tumor and its transition to the aggressive metastatic stage<sup>4</sup>. This process is accompanied by activation of the range of genes responsible for the invasion and migration of cancer cells. A series of studies demonstrated that TGF- $\beta$  was able to activate the expression of chemokines that enhance the metastatic potential of tumor cells<sup>5</sup>. Specifically, TGF- $\beta$  induced the expression levels of CXCL1—attracts myeloid cells in the primary tumor site<sup>6</sup>, CXCL5—attracts neutrophils<sup>7</sup>, and CCL20—stimulates proliferative and migrative activity of BC cells<sup>8</sup>.

However, until recently, there was no information regarding the effect of TGF- $\beta$  on CCL2 chemokine expression in BC cells, and the association of increased TGF- $\beta$ /Smad3 expression with CCL2 induction in smooth muscle cells is poorly translated to cancer cells<sup>9</sup>. CCL2 (classically named MCP-1, monocyte chemoattractant protein 1), is known to be a powerful monocyte chemoattractant and to strongly induce metastasis formation by BC cells. In particular, lung metastases are associated with the involvement of inflammatory monocytes expressing the CCR2 receptor, which differentiate into metastasis-associated macrophages that prepare the microenvironment for cancer cells extravasation, as well as stimulate tumor growth at the site of metastasis<sup>10</sup>. CCL2 has also been shown to play a key role in attracting tumor-associated macrophages (TAM) to the primary tumor site<sup>11</sup>. TAM contributes to tumor progression by inducing angiogenesis, intensifying the epithelial-mesenchymal transition, and stimulating the intravasation of cancer cells<sup>12</sup>. In general, the elevated expression of CCL2 by BC cells is associated with an aggressively metastasizing tumor phenotype and with poor prognoses in patients<sup>13</sup>.

Recent studies demonstrated that the CCL2 gene expression level directly correlates with the TGF- $\beta$  activity in breast cancer patients. These data are supported by the fact that, in the 4T1-BALB/c BC allograft mouse model, the inhibition of TGF- $\beta$  causes a significant reduction in the levels of CCL2 expression in the primary tumor site, which correlates with lower levels of lung metastases<sup>14</sup>. In addition, it has been reported that TGF- $\beta$  effects on CCL2 expression strongly depend on BC cell phenotype and activation is typical for advanced-stage BC cells<sup>14</sup>. In the current study, we aimed to identify the mechanism underlying regulation of CCL2 gene expression by TGF- $\beta$  in advanced-stage triple-negative BC cells with mesenchymal-like phenotype.

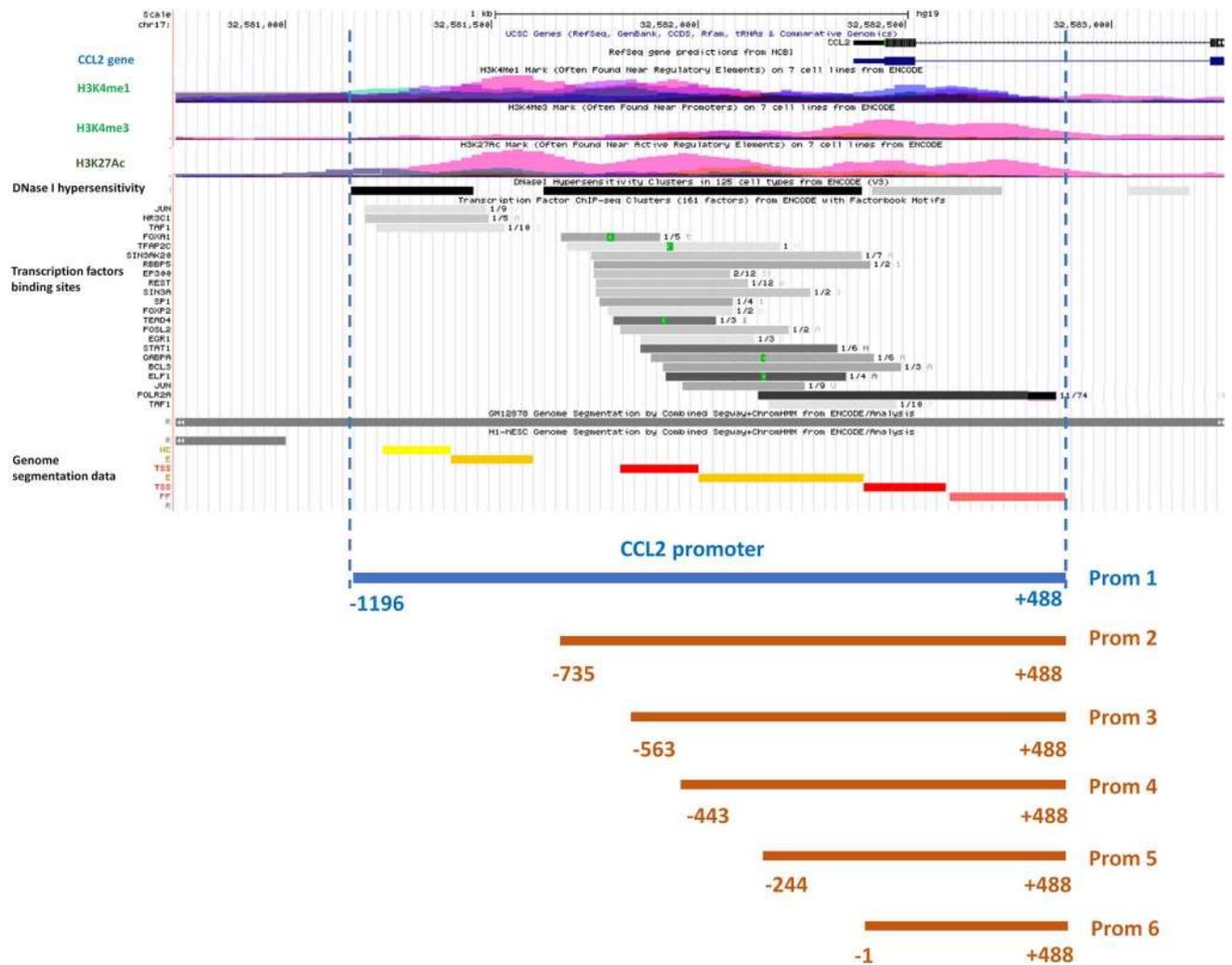
## Results

**Characterization of potential CCL2 promoter region.** The CCL2 promoter region was not described and characterized earlier; therefore, we performed its identification using the following epigenetic data represented in the UCSC Genome Browser (<http://genome.ucsc.edu/>, Human Feb. 2009 GRCh37/hg19 Assembly) as described earlier<sup>15</sup>: DNase-I hypersensitivity clusters, histone modifications (H3K4me3, H3K4me1, and H3K27ac), the localization of transcription factor (TF) binding sites, and information on chromatin segmentation that combines data on different epigenetic parameters<sup>16</sup>. Based on these data, we identified the region spanning from -1196 to +488 bp from the CCL2 transcription start site (TSS) as the potential promoter (Fig. 1).

**TGF- $\beta$ 1 induces CCL2 gene expression and CCL2 promoter activity in advanced stage triple-negative breast cancer cell lines.** In the current study, we used MDA-MB-231 and HCC1937 triple negative human breast cancer cell lines that are reported to overexpress CCL2<sup>17</sup> and demonstrate mesenchymal-like phenotype<sup>18,19</sup>. We verified that TGF- $\beta$  could induce CCL2 expression in triple-negative breast cancer cells with mesenchymal-like phenotype as it was indicated in previously published data<sup>14</sup>. The incubation of MDA-MB-231 and HCC1937 cells with 10 ng/mL of recombinant TGF- $\beta$ 1 protein (ab50036, Abcam, Cambridge, UK) for 24 h led to significant induction of expression levels of the IL-11 and CXCR4 genes (Fig. 2A), which are considered as the main TGF- $\beta$ -induced targets<sup>1</sup>. The incubation of MDA-MB-231 cells with TGF- $\beta$ 1 resulted in higher levels of the phospho-SMAD3 (Ser 204) and phospho-SMAD2 (Ser465/467) forms, the key participants of canonical TGF- $\beta$  intracellular cascade<sup>20</sup>. All of these TGF- $\beta$ -dependent effects correlated with the stimulation of both CCL2 mRNA (Fig. 2A) and protein (Fig. 2B) levels that indicated the response of CCL2 on TGF- $\beta$  stimulation.

In addition, incubation of MDA-MB-231 and HCC1937 breast cancer cells with TGF- $\beta$ 1 (10 ng/mL, for 24 h) led to a 1.5–twofold increase in the level of CCL2 promoter activity in the luciferase reporter system (Fig. 2C). These results highly correlate with the expression data and confirm the role of TGF- $\beta$  signaling in the regulation of the CCL2 promoter.

To perform a more comprehensive analysis of SMAD proteins in TGF- $\beta$ 1-dependent regulation of CCL2 promoter we measured the levels of several other SMAD family members, SMAD4, SMAD6 and SMAD7 (Fig. 2B). SMAD4 is known to form a complex with activated p-SMAD2/3 which acts as a direct transcriptional activator of TGF- $\beta$ -dependent target genes<sup>21,22</sup>. We detected no change in SMAD4 protein expression in response to TGF- $\beta$ 1 while the levels of p-SMAD2 and p-SMAD3 were induced. Apparently, SMAD4 activation is not required for the complex formation because the protein is already present in the cytoplasm and is recruited to the complex in a non-phosphorylated form<sup>21</sup>. Efficient complex formation is supported by the induction of IL11 and CXCR4 that have been reported to be direct p-SMAD2/3-SMAD4 target genes<sup>23</sup>. SMAD6 and SMAD7 have been shown to reduce the activity of p-SMAD2/3-SMAD4 complex acting in an antagonistic way<sup>22</sup> and are therefore considered as inhibitory members of the SMAD family. The unchanged SMAD6 and SMAD7 proteins levels support the activating nature of intracellular TGF- $\beta$  receptor signaling in MDA-MB-231 cells.

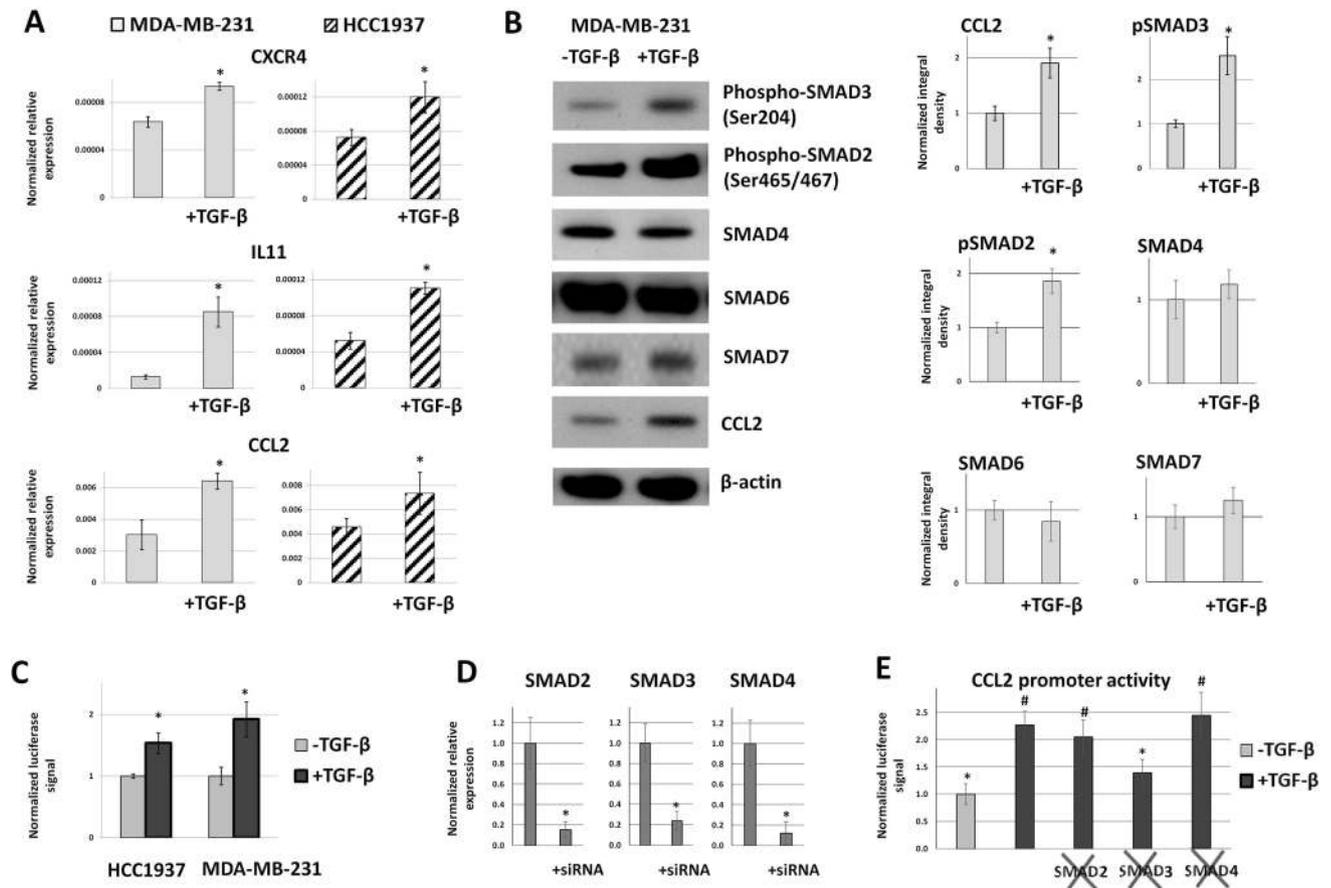


**Figure 1.** Illustration of the 5' area of the human CCL2 gene (UCSC Genome Browser, Human Feb. 2009 GRCh37/hg19 Assembly) with mapped potential CCL2 promoter variants used for deletion screening. The numbers indicate positions in response to the CCL2 transcription start site. *H3K4Me1* the track indicating areas of monomethylation of Histone H3 at lysine 4, *H3K4Me3* the track indicating areas of trimethylation of Histone H3 at lysine 4, *H3K27Ac* the track indicating areas of acetylation of Histone H3 at lysine 27, *DNase I hypersensitivity* the track indicating deoxyribonuclease I hypersensitivity clusters.

In order to check whether p-SMAD2/3-SMAD4 complex is able directly activate the CCL2 gene, we tested CCL2 promoter activity in TGF- $\beta$ -induced MDA-MB-231 cells in the presence of specific siRNAs against SMAD2, SMAD3 and SMAD4 (Fig. 4D,E). Only SMAD3 knockdown significantly abrogated CCL2 promoter activation after TGF- $\beta$  treatment indicating that other proteins of SMAD family do not interact directly with the promoter.

**- 244/- 1 region of the CCL2 promoter is responsible for its TGF- $\beta$ -dependent regulation.** We applied deletion screening of CCL2 promoter to identify its region responsible for TGF- $\beta$ 1-dependent activation. The activities of luciferase reporter constructs containing deletion variants of the CCL2 promoter were assessed in MDA-MB-231 breast cancer cells in the presence or absence of TGF- $\beta$ 1 (10 ng/mL, 24 h), (Fig. 3). All promoter regions except Prom 6 demonstrated the dependence on TGF- $\beta$ -induced activation. In the case of Prom 6, this effect was completely absent, although this region is essential for maintaining basal promoter activity. Comparing the character of influence of TGF- $\beta$ 1 on the activities of Prom 5 and Prom 6, the region is likely located within positions - 244 and - 1 in response to CCL2 transcription start site should be responsible for the dependence of CCL2 promoter on TGF- $\beta$ .

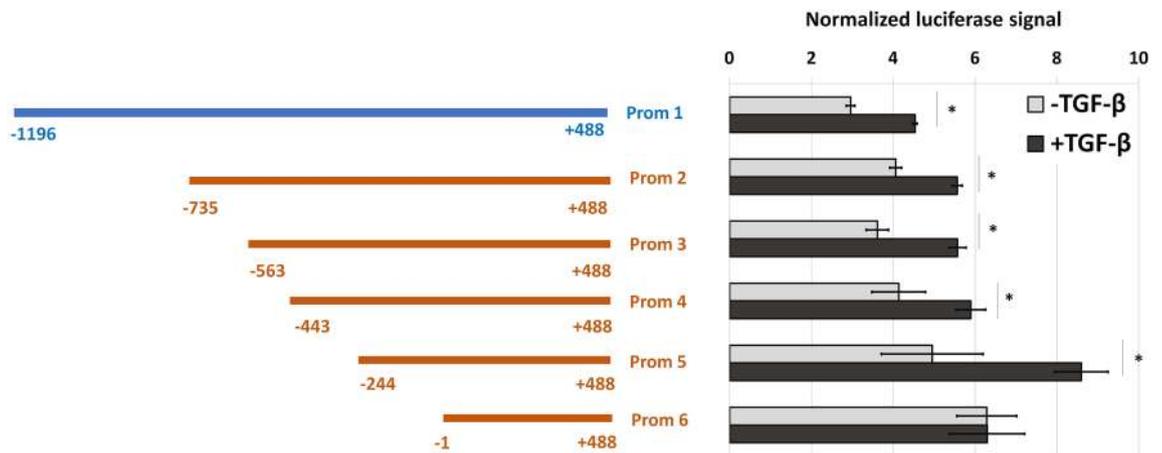
**EGR1 and RXRA transcription factors are essential for TGF- $\beta$ -dependent activation of the CCL2 promoter.** To identify potential transcription factors binding sites within the - 244/- 1 region of the CCL2 promoter we used the MoLoTool (<http://molotool.autosome.ru>) online service, which allows the search for transcription factor binding sites in a given sequence using curated models of these sites presented in the HOCOMOCO database (<http://hocomoco1.autosome.ru/>)<sup>24</sup>. For the first step, we formed a list of transcription



**Figure 2.** Transforming growth factor beta (TGF- $\beta$ 1)-induced activation of advanced stage triple-negative breast cancer cell lines enhanced CCL2 gene expression and CCL2 promoter activity. **(A)** The incubation of MDA-MB-231 and HCC1937 cells with 10 ng/mL TGF- $\beta$ 1 for 24 h led to significant induction of the expression levels of IL-11 and CXCR4 genes (the main TGF- $\beta$ -induced targets) and also resulted in elevated CCL2 mRNA levels. mRNA levels were measured by RT-PCR in real time. The results of five independent experiments are represented. \* $P < 0.01$ . **(B)** The incubation of MDA-MB-231 cells with TGF- $\beta$ 1 led to the induction of phospho-SMAD3 (Ser 204), phospho-SMAD2 (Ser465/467) and CCL2 protein levels with no influence on SMAD4, SMAD6 and SMAD7 protein levels. Representative western blots are shown (left panel). For chemiluminescent detection of the bands corresponding to SMAD, CCL2 and  $\beta$ -actin proteins we applied X-ray film exposure. The full-length blot images are represented in Supplementary Fig. S1. The experiment was repeated three times, the images were quantified using ImageJ software (right panel). The values of integral densities for CCL2, pSMAD3, pSMAD2, SMAD4, SMAD6 and SMAD7 are normalized on the respective  $\beta$ -actin bands. Normalized integral densities for CCL2, pSMAD3, pSMAD2, SMAD4, SMAD6 and SMAD7 in MDA-MB-231 cells without TGF- $\beta$ 1 activation are represented as 1. \* $P < 0.01$ . **(C)** TGF- $\beta$  induced CCL2 promoter activity in a luciferase reporter system in both MDA-MB-231 and HCC1937 cell lines. The data were obtained in five independent experiments and normalized to the Renilla luciferase activity. Normalized luciferase signals in MDA-MB-231 and HCC1937 cells without TGF- $\beta$ 1 activation are represented as 1. \* $P < 0.01$ . **(D)** The efficiency of knock-down of SMAD2, SMAD3 and SMAD4 genes in MDA-MB-231 cells. For each gene, real-time PCR signal was normalized to the sample without siRNA. The data are represented as the mean  $\pm$  SEM (five independent experiments). \* $P < 0.01$  in respect to control without siRNA. **(E)** TGF- $\beta$  induced CCL2 promoter activity in the presence of SMAD2-, SMAD3- or SMAD4-specific siRNAs in a luciferase reporter system in MDA-MB-231 cells (data from five independent experiments). Luciferase signal without TGF- $\beta$ 1 normalized to Renilla luciferase activity is represented as 1. \* $P < 0.01$  in respect to control with TGF- $\beta$ 1; # $P < 0.01$  in respect to control without TGF- $\beta$ 1.

factors with “strong” ( $-\log_{10}(P\text{-value}) > 5$ ) predicted sites within the 244/-1 region: MAZ, SP4, EGR2, THAP1, VEZF1, PAX6, SP1, RXRA, WT1, SP3, ZN768, BCL6, ZSC22, SP2, KLF3, RARA, ZBT14, and EGR1. In the next step, based on published data, we selected those factors from the list that are known to be involved in TGF- $\beta$ -induced cascades: EGR1<sup>25</sup>, EGR2<sup>26</sup>, PAX6<sup>27</sup>, RARA and RXRA<sup>28</sup>, and SP1<sup>29</sup>. All selected transcription factors demonstrated high expression levels in both MDA-MB-231 and HCC1937 cell lines according to the mRNA expression data represented in the Broad Institute Cancer Cell Line Encyclopedia (Supplementary Table S2), which was used as an additional selection criterion.

To reveal the roles of these factors in CCL2 promoter regulation, we estimated the activity of the full-size CCL2 promoter (Prom 1) and its Prom 5 and Prom 6 fragments in MDA-MB-231 cells in normal conditions and in the presence of TGF- $\beta$ 1 (10 ng/mL, 24 h) or/and specific siRNAs against EGR1, EGR2, PAX6, RARA, RXRA,



**Figure 3.** The  $-244/-1$  region of the CCL2 promoter is responsible for its TGF- $\beta$ 1-dependent regulation. Functional analysis of deletion variants of the CCL2 promoter. Left, schematic illustration of promoter variants with positions in response to the CCL2 transcription start site. Right, normalized luciferase signals in MDA-MB-231 breast cancer cells in normal conditions and after the incubation with TGF- $\beta$ 1 (10 ng/mL, 24 h). Note that shorter promoter fragments yield smaller plasmid sizes and better transfection efficiencies, resulting in a gradual increase of the overall luciferase signal. The data were obtained in five independent experiments and normalized to the Renilla luciferase activity. \* $P < 0.01$ .

and SP1 factors mRNAs. The efficiency of siRNAs in suppressing the expression of these factors was preliminarily estimated by qPCR in the MDA-MB-231 cell line and was not less than 80% (Fig. 4A). TGF- $\beta$ 1 induced expression of all of the studied transcription factors except for EGR1 and SP1<sup>30,31</sup>. siRNA knock-downs of all the factors abolished the activation of the corresponding genes and in the absence of TGF- $\beta$ 1 had no effects on the activity of CCL2 promoter variants (Fig. 4B, top panel). The Prom 6 fragment demonstrated no dependence on both TGF- $\beta$ 1 activation and the levels of any of the studied factors (Fig. 4B, bottom panel). The knock-down of any of EGR1 or RXRA completely abolished the effect of Prom1 and Prom 5 stimulation in response to TGF- $\beta$ 1. According to these results, EGR1 and RXRA could be considered as the main mediators of TGF- $\beta$ 1-dependent regulation of the CCL2 promoter.

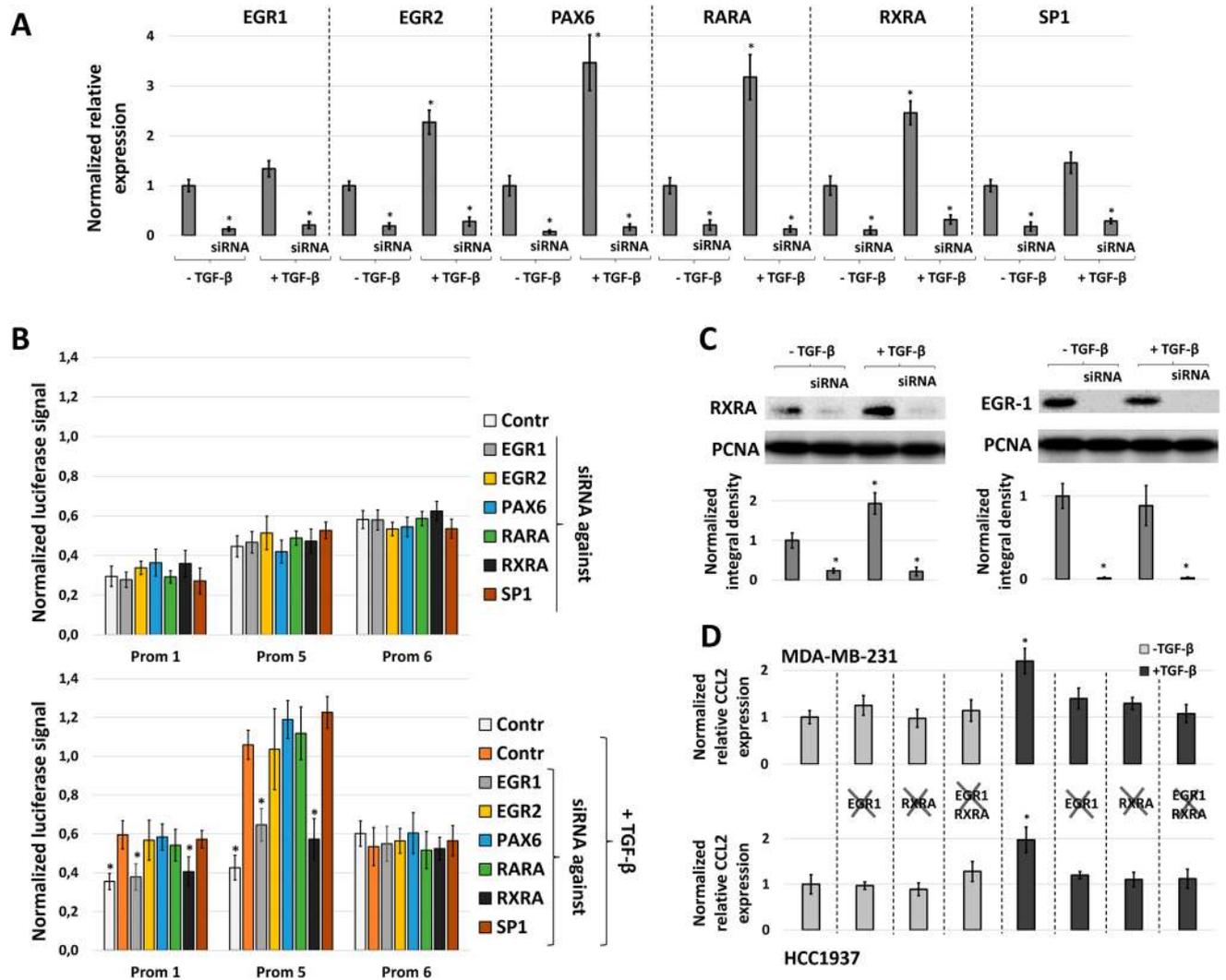
Transfection of MDA-MB-231 cells with EGR1- or RXRA-specific siRNAs led to reduced levels of the respective proteins in the nucleus (Fig. 4C) and these levels could not be induced by TGF- $\beta$ 1 (10 ng/mL, 24 h). In the absence of specific siRNAs, TGF- $\beta$ 1 treatment resulted in elevated RXRA protein levels in the nucleus and did not influence EGR1 protein level, in agreement with mRNA expression data (Fig. 4A). At the level of the CCL2 mRNA, either EGR1- or RXRA-specific siRNA abrogated the TGF- $\beta$ 1-triggered effect almost as completely as both siRNAs together, both in MDA-MB-231 and in HCC1937 cells (Fig. 4D).

### Both EGR1 and RXRA bind to the CCL2 promoter after TGF- $\beta$ treatment independently in a time-dependent manner.

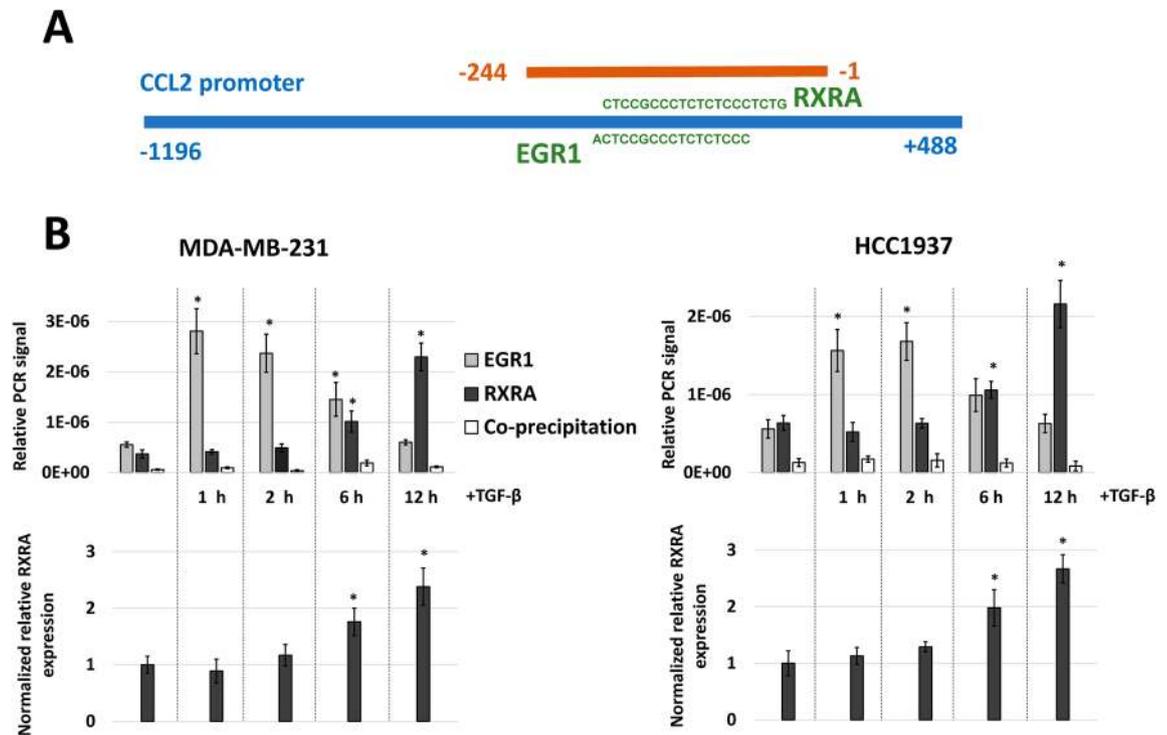
To estimate the influence of TGF- $\beta$  on the efficiency of EGR1 and RXRA binding to the CCL2 promoter in vivo, we used a classical chromatin immunoprecipitation (ChIP) assay. As the EGR1- and RXRA-binding sites are located in the same ( $-244/-1$ ) region of the promoter, we analyzed the amount of cross-linked DNA from MDA-MB-231 and HCC1937 cells precipitated with anti-EGR1 and anti-RXRA antibodies using the primers corresponding to this region (schematically illustrated in Fig. 5A). We performed ChIP analysis for both cell lines activated by TGF- $\beta$ 1 (10 ng/mL) for 1, 2, 6 and 12 h. In order to determine whether RXRA and EGR1 could cooperatively bind to the same region of the promoter we also performed co-immunoprecipitation i.e. after precipitation with anti-EGR1 antibodies an additional precipitation round was carried out with anti-RXRA antibodies. We also controlled RXRA expression level as this factor is induced by TGF- $\beta$ 1, unlike EGR1 mRNA and protein (Fig. 4A). EGR1 demonstrated maximal binding to CCL2 promoter already at 1 h of TGF- $\beta$ 1 treatment (Fig. 5B), in agreement with existing data on direct EGR1 activation by TGF- $\beta$ 1<sup>30</sup>. The strength of RXRA binding to CCL2 promoter also correlated with its expression level (Fig. 5B). These data indicate time-resolved binding of EGR1 and RXRA transcription factors to CCL2 promoter in response to TGF- $\beta$ , suggesting largely independent action of these transcription factors on CCL2 gene expression. Indeed, crosslinked chromatin fragments precipitated with anti-EGR1 antibodies could not be precipitated with anti-RXRA antibodies in co-immunoprecipitation assay. The strength of combined EGR1 and RXRA binding to CCL2 promoter correlates well with TGF- $\beta$ 1-induced CCL2 expression and promoter activity (Fig. 2) and provides a mechanistic explanation for transcriptional activation of the CCL2 gene by these two factors for an extended period of time.

### Point mutations in EGR1 and RXRA binding sites in CCL2 promoter completely abolished its activation by TGF- $\beta$ .

In order to verify the roles of the EGR1 and RXRA transcription factors in TGF- $\beta$ -dependent regulation of the CCL2 promoter, we generated a luciferase reporter construct containing the CCL2 promoter with point mutations in the EGR1 and RXRA sites. The EGR1- and RXRA-binding sites in CCL2



**Figure 4.** Knock-down of EGR1 and RXRA transcription factors resulted in the independence of the CCL2 promoter of TGF-β1 activation in MDA-MB-231 breast cancer cells. **(A)** The efficiency of knock-down of EGR1, EGR2, PAX6, RARA, RXRA, and SP1 genes after transfection of MDA-MB-231 cells with specific siRNAs in the presence or absence of TGF-β1 activation was tested by real-time PCR. For each gene, the normalization to the sample without siRNA and TGF-β1 was performed (the meanings for these samples are represented as 1). The real-time PCR data are represented as the mean ± SEM (five independent experiments). \* $P < 0.01$  in respect to control without siRNAs and TGF-β1. **(B)** Prom 1, Prom 5, and Prom 6 CCL2 promoter variants were tested using a luciferase reporter assay in MDA-MB-231 cells transfected with specific siRNAs in normal conditions (top panel) or in the presence of TGF-β1 (10 ng/mL, 24 h) (bottom panel). The data were obtained in seven independent experiments and normalized to the Renilla luciferase activity. \* $P < 0.01$  in respect to control with TGF-β1. **(C)** EGR1 and RXRA protein levels measured in nuclear extracts of MDA-MB-231 cells in the presence or in the absence of specific siRNAs and/or TGF-β1 (10 ng/mL, 24 h). The experiment was repeated for three times. Representative western blots are shown (upper panel). For the detection of RXRA, EGR1 and PCNA proteins the membranes were splitted horizontally into two pieces. The upper piece of each membrane was stained with anti-RXRA or anti-EGR1 primary antibodies, and the lower—with anti-PCNA antibodies. Chemiluminescent detection of the bands was performed using ChemiDoc Imager (Bio-Rad Laboratories, Hercules, CA, USA). Full-size images of splitted membranes are represented in Supplementary Fig. S2. Membrane images were quantified using ImageJ software (bottom panel). Integral densities for EGR1 and RXRA bands normalized to PCNA in MDA-MB-231 cells without TGF-β1 activation are represented as 1. \* $P < 0.01$ . **(D)** CCL2 mRNA levels were estimated by RT-PCR in real time. For each cell line, samples without siRNAs and TGF-β1 are represented as 1. The results of three independent experiments are shown. The full-length blot images are represented in Supplementary Fig. S2. \* $P < 0.01$  as compared to control without siRNAs and TGF-β1.



**Figure 5.** TGF- $\beta$ 1 induces EGR1 and RXRA binding to the CCL2 promoter in MDA-MB-231 and HCC1937 breast cancer cells in time-dependent manner. (A) Map of the CCL2 promoter, predicted EGR1-/RXRA-binding sites, and PCR product amplified in the chromatin immunoprecipitation (ChIP) assay. (B) Activation of MDA-MB-231 and HCC-1937 cells with TGF- $\beta$ 1 led to elevated time-dependent EGR1 and RXRA crosslinking to the amplicon containing corresponding consensus sites (top panel). As described Material and Methods, none of the specificity controls produced signals above background. TGF- $\beta$ 1 also induced RXRA expression as estimated by RT-PCR in real time (bottom panel). The RT-PCR data were normalized to the sample without TGF- $\beta$ 1 separately for each cell line. The results of three independent experiments are represented both for CHIP-assay and RT-PCR. \* $P < 0.01$  in response to control without TGF- $\beta$ 1.

promoter were overlapped, and their consensus sequences are highly similar; therefore, mutations in their key positions should damage both of them (schematically illustrated on the Fig. 6A).

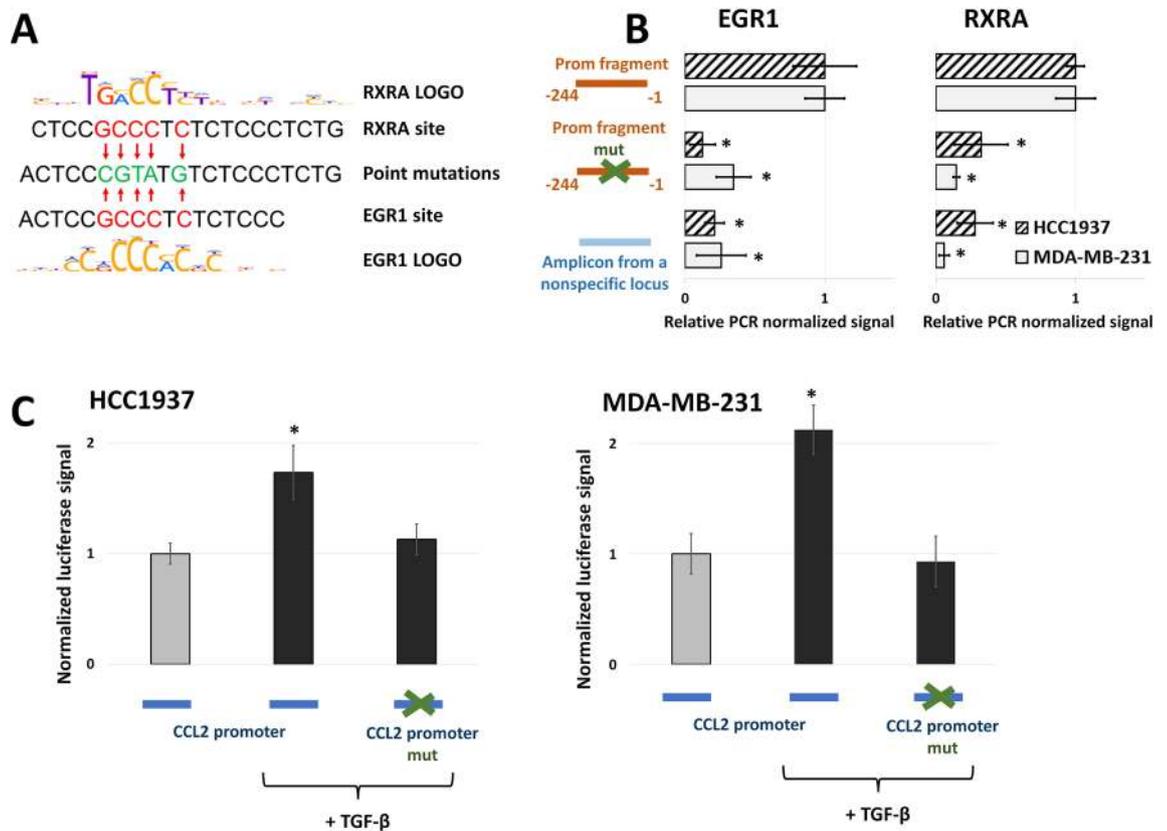
Preliminary, we applied a pull-down assay<sup>32</sup> to test the influence of these mutations on the efficiency of EGR1 and RXRA binding to the CCL2 promoter. For this reason, we generated two types of amplicons: the -244/-1 region of the CCL2 promoter and the same region with point mutations in the EGR1/RXRA sites (Fig. 6B). For precipitation, we used nuclear extracts from MDA-MB-231 and HCC1937 cells and anti-EGR1 or anti-RXRA antibodies. We concluded that the introduced mutations totally abrogated EGR1 and RXRA1 binding to the CCL2 promoter in both cell lines (Fig. 6B).

In the next step, we compared the activity of a full-size CCL2 promoter and its mutated variant in normal conditions and in the presence of TGF- $\beta$ 1 (10 ng/mL, 24 h) (Fig. 6C). Mutation of the EGR1/RXRA binding sites resulted in a full absence of the effect of TGF- $\beta$ 1 on CCL2 promoter activity in both the MDA-MB-231 and HCC1937 cell lines. This result agrees with both the pull-down and CHIP data and indicates the exclusive role of the EGR1/RXRA binding sites in the TGF- $\beta$ -dependent regulation of the CCL2 promoter.

## Discussion

A direct correlation between the CCL2 and TGF- $\beta$  levels is typical for progressive stages of breast cancer as was recently shown using both clinical samples and in vitro studies<sup>14</sup>. In vivo studies applying the 4T1-BALB/c BC mouse model demonstrated that the inhibition of TGF- $\beta$  causes a significant reduction in the levels of CCL2 expression in the primary tumor site, which correlates with lower levels of lung metastases<sup>14</sup>. We verified these data using MDA-MB-231 and HCC1937 triple negative human breast cancer cell lines, which are reported to overexpress CCL2<sup>17</sup>. These cells demonstrated a response to TGF- $\beta$ 1 that was detected by the induction of the expression levels of the main TGF- $\beta$  targets<sup>1</sup>, IL-11 and CXCR4 (Fig. 2A), and higher levels of phospho-SMAD3 (Ser 204) and phospho-SMAD2 (Ser465/467) (Fig. 2B), which are considered as the key participant of the TGF- $\beta$  intracellular cascade<sup>20</sup>. All of these TGF- $\beta$ -dependent effects correlated with stimulation of the CCL2 mRNA (Fig. 2A) and protein (Fig. 2B) levels that confirm the previously published data. Thereby, TGF- $\beta$  undoubtedly acts as an activator of CCL2 expression in progressive breast cancer; however, the molecular mechanism of this action remains unknown.

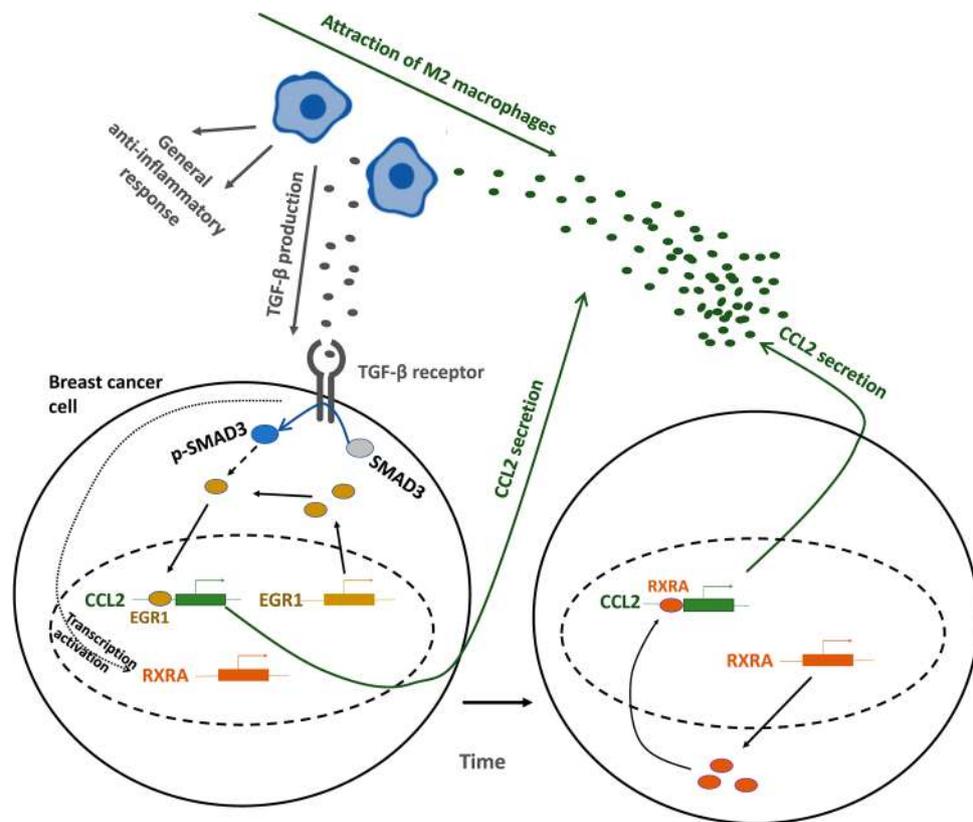
Using the current epigenetic data presented in the UCSC Genome Browser, we identified the region spanning from -1196 to +488 from the CCL2 transcription start site as a potential CCL2 promoter. TGF- $\beta$ 1-induced activation of MDA-MB-231 and HCC1937 breast cancer cells led to an increase in the activity of the CCL2 promoter



**Figure 6.** Point mutations in the EGR1 and RXRA binding sites in the CCL2 promoter completely abolished its activation by TGF- $\beta$ 1. (A) Position weight matrixes of the EGR1 and RXRA binding sites and the strategy of their point mutagenesis. Red arrows indicate nucleotides changes. (B) Point mutations of the EGR1/RXRA sites in the CCL2 promoter totally abrogated EGR1 and RXRA1 binding to the promoter. The efficiency of EGR1 and RXRA1 binding to different variants of the CCL2 promoter was estimated by pull-down assay with the use of nuclear extracts from MDA-MB-231 and HCC1937 breast cancer cells. The data were normalized to the normal -244/-1 amplicon after subtraction of the background signal (the signal for this sample is represented as 1 for each cell line). The results of five independent experiments are represented. \* $P < 0.01$ . (C) Mutation of the EGR1/RXRA binding sites resulted in a full absence of the effect of TGF- $\beta$ 1 (10 ng/mL, 24 h) on CCL2 promoter activity in both the MDA-MB-231 and HCC1937 cell lines as estimated using a luciferase reporter assay. The data were obtained in five independent experiments and normalized to the Renilla luciferase activity. Normalized luciferase signals of original CCL2 promoter without TGF- $\beta$ 1 activation are represented as 1 for each cell line. \* $P < 0.01$ .

in the luciferase reporter system (Fig. 2C), and this was fully provided by its -244/-1 region (Fig. 3). It should be noted that p-SMAD2/3-SMAD4 complex, which is considered as the canonical effector of intracellular signaling of TGF- $\beta$  receptor acting as a transcriptional activator<sup>21,22</sup>, is not involved in direct CCL2 promoter activation since only SMAD3 knockdown but not that of SMAD2 or SMAD4 influenced the level of CCL2 promoter activity after TGF- $\beta$ 1 treatment (Fig. 2D,E). Bioinformatical analysis of 244/-1 TGF- $\beta$ -dependent region made it possible to predict potential binding sites for the transcription factors. The extensive list of transcription factors has been shortened as a result of literature annotation. We were interested only in the factors whose activity is described as dependent on TGF- $\beta$ -signaling in cancer cells: EGR1<sup>25</sup>, EGR2<sup>26</sup>, PAX6<sup>27</sup>, RARA and RXRA<sup>28</sup>, and SP1<sup>29</sup>. Among these factors, the knock-down of only EGR1 and RXRA completely abolished the effect of CCL2 promoter activation in response to TGF- $\beta$ 1 (Fig. 4B). Interestingly, RARA knock-down did not have an effect on the promoter activity, indicating that RXRA probably binds CCL2 promoter as a homodimer. Additional experiments supported the role of EGR1 and RXRA in TGF- $\beta$ -dependent regulation of the CCL2 promoter as their knock-down resulted in the absence of TGF- $\beta$ -induced increase in CCL2 expression (Fig. 4D), both factors demonstrated binding to the promoter in a TGF- $\beta$ 1-dependent manner (Fig. 5), and point mutagenesis of their common binding site led to the loss of the ability of TGF- $\beta$ 1 to activate the CCL2 promoter (Fig. 6). Of note, EGR1 and RXRA bind to the CCL2 promoter independently but with different kinetics which may explain their essential and independent roles in TGF- $\beta$ -induced CCL2 promoter activation.

The described data indicate EGR1 and RXRA as the key link between the TGF- $\beta$ -induced pathways and CCL2 expression and allow us to propose a positive regulatory loop (Fig. 7). Research previously described that high CCL2 levels directly correlated with intense macrophage polarization to the M2 phenotype and secretion of CCL22, which attracts and accumulates Th2 cells<sup>14</sup>. The presence of M2 macrophages is associated with the general anti-inflammatory response<sup>33</sup> and with production of TGF- $\beta$ , which could cause further stimulation of



**Figure 7.** Schematic illustration of potential EGR1 and RXRA participation in TGF- $\beta$ -CCL2 positive regulatory loop.

CCL2 expression in breast cancer cells via EGR1 and RXRA. The participation of the EGR1 and RXRA transcription factors in such positive regulatory loops is in agreement with their anti-inflammatory functions. EGR1 is known to activate expression of a range of anti-inflammatory genes, including IL4<sup>34,35</sup>, and is also involved in the positive regulation of TGF- $\beta$  transcription<sup>36</sup>. In its turn, RXRA provides an anti-inflammatory effect by repressing transcription of IL-12 and TNF- $\alpha$  and induction of IL-10<sup>37</sup>. Therefore, activation of CCL2 expression could be an anti-inflammatory effect of these factors that was not previously described.

Special attention should be paid to the fact that both RXRA and EGR1 are independently essential for TGF- $\beta$ -dependent activation of the CCL2 gene promoter and subsequent induction of CCL2 expression. We did not set out to reveal the precise mechanism of participative action of these factors; however, we could propose a possible model (Fig. 7) based on the existing information and on our CHIP and gene expression data (Fig. 5). EGR1 is known to be a rapid-acting transcription factor that directly binds to the activated SMAD3, translocates to the nucleus, and performs transcriptional activation of the target genes<sup>30</sup>. The specific involvement of SMAD3 via interaction with EGR1 may explain why SMAD3 knockdown specifically abrogates the TGF- $\beta$ -induced activation of CCL2 promoter (Fig. 2E). Accordingly, we did not observe induction in EGR1 expression by TGF- $\beta$  (Fig. 4A,C), but registered rapid changes in its binding to CCL2 promoter (Fig. 5). In addition, EGR1 is reported to act as an epigenetic modulator in different tissues<sup>38,39</sup>, including breast cancer<sup>40</sup>, which results in the ability to influence the methylation context around its binding region and significantly change the affinity of other factors to this region<sup>40</sup>.

In addition to previously described anti-inflammatory functions of EGR1, its action could also provide some features of the pro-inflammatory response, including NF- $\kappa$ B activation<sup>41</sup>, and is inhibited by the secondary long-lasting processes triggered by TGF- $\beta$ -dependent signaling. For example, TGF- $\beta$  induces LXRA, RXRA, and RARA expression levels<sup>42</sup>. These factors are known to produce the overall anti-inflammatory response<sup>37,43</sup> and, particularly, reduce EGR1 expression levels<sup>44</sup>. These processes presumably lead to a decrease in EGR1 levels and to subsequent involvement of RXRA in the activation of the CCL2 promoter. Indeed, RXRA binding activity correlates well with its expression level (Fig. 5). In summary, rapid EGR1 action in response to TGF- $\beta$  could lead to both activation of the CCL2 promoter and modification of its epigenetic context but also is inhibited itself in a short time. RXRA, on the contrary, could bind to the prepared CCL2 promoter region and provide long-lasting activation effects. Thereby, both EGR1 and RXRA could be necessary for complex activation of the CCL2 promoter in response to TGF- $\beta$ .

We conclude that, in this study, we provided a possible mechanism that underlies TGF- $\beta$ -dependent regulation of the CCL2 gene promoter in triple-negative advanced breast cancer cells.

## Materials and methods

**Cell lines.** MDA-MB-231 and HCC1937 cell lines were maintained in a local stock of the Shemyakin–Ovchinnikov Institute of Bioorganic Chemistry RAS (Moscow, Russia) and kindly provided by Dr. D.S. Kravchenko. The stocks of MDA-MB-231 and HCC1937 cells were recently authenticated using commercially performed STR analysis (Gordiz, Moscow, Russia, data not shown). Both cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM, Paneco, Moscow, Russia) supplemented with 10% fetal bovine serum.

**Ethical approval.** Scientific Council of the Engelhardt Institute of Molecular Biology declared no ethical approval requirements for experiments performed on the commercially available cell lines used in the study.

**RNA isolation, reverse transcription, and real-time quantitative RT-PCR.** The total RNA was extracted from cells using ExtractRNA reagent (Evrogen, Moscow, Russia) according to the manufacturer's protocol. cDNA was generated using the M-MULV reverse transcriptase and oligo-dT primer from the MMLV RT kit (Evrogen, Moscow, Russia). Quantitative RT-PCR was performed using the protocol and PCR program described earlier<sup>45</sup> and specific primers designed to amplify intron-spanning fragments of human  $\beta$ -actin, CCL2, CXCR4, IL11, EGR1, EGR2, PAX6, RARA, RXRA, and SP1 genes (Supplementary Table S1). The results were calculated using  $\Delta$ Ct method, and all expression levels of target genes were normalized to  $\beta$ -actin.

**Western blot analysis.** The protocols for the total cell lysate preparation, electrophoresis, transfer to the nitrocellulose membrane, and band visualization were described earlier<sup>15</sup>. For analysis of EGR1 and RXRA protein levels we used nuclear extracts prepared according to the previously described protocol in the same manner as for the pull-down assay<sup>32</sup>. The membranes were pre-blocked using 3% BSA and incubated with anti-pSMAD3 (Ser204) antibodies (PA5-36877, Thermo Fisher Scientific, Waltham, MA, USA) at a 1:2000 dilution, anti-CCL2 antibodies (ab9669, Abcam, Cambridge, UK) at a 1:2000 dilution, anti-pSMAD2 (Ser465/467) antibodies (138D4, Cell Signaling Technology, Danvers, MA, USA) at a 1:2000 dilution, anti-SMAD4 antibodies (D3R4N, Cell Signaling Technology, Danvers, MA, USA) at a 1:3000 dilution, anti-SMAD6 antibodies (sc-25321, Santa Cruz Biotechnology, Santa Cruz, CA, USA) at a 1:2000 dilution, anti-SMAD7 antibodies (ab216428, Abcam, Cambridge, UK) at a 1:3000 dilution, anti-EGR1 antibodies (44d5, Cell Signaling Technology, Danvers, MA, USA) at a 1:3000 dilution, anti-RXRA antibodies (d6h10, Cell Signaling Technology, Danvers, MA, USA) at a 1:3000 dilution, anti-PCNA antibodies (D3H8P, Cell Signaling Technology, Danvers, MA, USA) at a 1:3000 dilution as a loading control for nuclear extracts and anti- $\beta$ -actin antibodies (ab8229, Abcam, Cambridge, UK) at a 1:3000 dilution as a loading control for whole-cell lysates. Gel images of three independent repeats were analyzed using ImageJ software to determine integral densities of the bands. The values of integral densities to  $\beta$ -actin for CCL2 and pSMAD3 were normalized and to PCNA for RXRA and EGR1.

**Molecular cloning and luciferase reporter constructs.** We amplified a potential full-size CCL2 promoter region (Prom 1) located within positions –1196 and +488 in response to the CCL2 transcription start site (TSS) by PCR using human genomic DNA isolated from MDA-MB-231 cells and specific primers containing HindIII and NcoI restriction sites (Supplementary Table S1). The Prom 1 region was cloned to a pGL3 base luciferase reporter vector (Promega, Madison, WI, USA). We assembled luciferase reporter plasmids containing deletion variants of CCL2 promoter: Prom 2 (–735/+488), Prom 3 (–563/+488), Prom 4 (–443/+488), Prom 5 (–244/+488), and Prom 6 (–1/+488), (Schematically illustrated on the Fig. 2); and a Prom 1 variant with point mutations in the predicted EGR1 and RXRA binding sites (Schematically illustrated in Fig. 6A).

**Luciferase reporter assay.** MDA-MB-231 and HCC1937 cells were electroporated with 10  $\mu$ g of purified plasmid DNA and 0.1  $\mu$ g of pRL-CMV Renilla luciferase control reporter vector (Promega, Madison, WI, USA) using the Neon Transfection System (Thermo Fisher Scientific, Waltham, MA, USA) and the following regimens: four 10 ms 1400 V pulses for MDA-MB-231, and three 10 ms 1550 V pulses for HCC1937. Firefly luciferase activity was measured as described earlier<sup>46</sup>, and normalized to the activity of Renilla luciferase in order to account for the variations in cell transfection efficiency.

**Knockdown of studied transcription factors and the proteins of SMAD family using siRNA.** We used previously published siRNAs against SMAD2, SMAD3<sup>47</sup> and SMAD4<sup>48</sup> proteins and EGR1, EGR2, PAX6, RARA, RXRA, and SP1 transcription factors<sup>49</sup>. The sequences of the siRNAs are represented in Supplementary Table S1. Commercially synthesized single-stranded RNAs (Syntol, Moscow, Russia) were annealed as previously described<sup>50</sup>. MDA-MB-231 and HCC1937 cells were transfected with siRNA duplexes (500 pmol of per 5 million cells) 48 h before transfection with the luciferase constructs and a further 200 pmol to extend the silencing effect.

**Chromatin immunoprecipitation (ChIP) assay.** We applied the cross-linking chromatin immunoprecipitation (X-ChIP) protocol as described earlier<sup>51</sup>. Briefly, after the lysis of  $2 \times 10^7$  of MDA-MB-231 or HCC1937 cells containing protein-DNA complexes preliminary cross-linked with 0.75% formaldehyde, these complexes were sonicated to obtain average DNA fragment sizes of 700 bp, incubated with antibodies to RXRA (d6h10, Cell Signaling Technology, Danvers, MA, USA) or EGR1 (44d5, Cell Signaling Technology, Danvers, MA, USA), and precipitated with pre-blocked protein A sepharose beads. After the elution of protein-DNA complexes and DNA purification, the target DNA was quantified by real-time PCR. We used four types of controls: a background control without lysate; a nonspecific precipitation control without antibodies; a control with isotype rabbit IgG

antibodies (DA1E, Cell Signaling Technology, Danvers, MA, USA); and an amplicon from a nonspecific locus containing no predicted RXRA- or EGR1-binding sites. The primer sequences are represented in Supplementary Table S1. None of the control precipitations produced signals above the background level.

**Pull-down assay.** We performed a pull-down assay according to our previously published original protocol<sup>32</sup>. Briefly, we amplified a 245-bp fragment of CCL2 promoter (−244/−1) and the same fragment with point mutations in EGR1/RXRA binding sites (Fig. 6B). An amplicon from a non-specific locus of the same length that did not contain any predicted EGR1/RXRA binding sites served as a negative control. The primer sequences are represented in Supplementary Table S1. The protocols of the isolation of nuclear extracts from MDA-MB-231 and HCC1937 cell lines, the immunoprecipitation of protein-DNA complexes, and quantification of bound DNA probes by real-time PCR were described earlier<sup>32</sup>. For precipitation, we used anti-RXRA (d6h10, Cell Signaling Technology, Danvers, MA, USA) and anti-EGR1 (44d5, Cell Signaling Technology, Danvers, MA, USA) antibodies. We used the following controls to calculate the background signal level: with an amplicon from a nonspecific locus; without nuclear extract; without antibodies; and isotype control with rabbit IgG (DA1E, Cell Signaling Technology, Danvers, MA, USA).

**Statistical analysis.** We used Microsoft Excel and Graphpad Prism software for our statistical analyses. Statistical significance was determined using a two-tailed unpaired Student's t-test. The data are represented as the mean ± SD.

Received: 21 December 2020; Accepted: 28 June 2021

Published online: 08 July 2021

## References

- Zarzynska, J. M. Two faces of TGF- $\beta$ 1 in breast cancer. *Mediat. Inflamm.* **2014**, 141747 (2014).
- Barcellos-Hoff, M. H. & Akhurst, R. J. Transforming growth factor- $\beta$  in breast cancer: too much, too late. *Breast Cancer Res.* **11**, 202 (2009).
- Pickup, M., Novitskiy, S. & Moses, H. L. The roles of TGF $\beta$  in the tumour microenvironment. *Nat. Rev. Cancer* **13**, 788–799 (2013).
- Wang, Y. & Zhou, B. P. Epithelial-mesenchymal transition in breast cancer progression and metastasis. *Chin. J. Cancer* **30**, 603–611 (2011).
- Bierie, B. *et al.* Abrogation of TGF- $\beta$  signaling enhances chemokine production and correlates with prognosis in human breast cancer. *J. Clin. Invest.* **119**, 1571–1582 (2009).
- Acharyya, S. *et al.* A CXCL1 paracrine network links cancer chemoresistance and metastasis. *Cell* **150**, 165–178 (2012).
- Hsu, Y.-L., Hou, M.-F., Kuo, P.-L., Huang, Y.-F. & Tsai, E.-M. Breast tumor-associated osteoblast-derived CXCL5 increases cancer progression by ERK/MSK1/Elk-1/Snail signaling pathway. *Oncogene* **32**, 4436–4447 (2013).
- Osula, K. O. & Sloane, B. F. Many roles of CCL20: Emphasis on breast cancer. *Postdoc J.* **2**, 7–16 (2014).
- Zhang, F. *et al.* Transforming growth factor- $\beta$  promotes recruitment of bone marrow cells and bone marrow-derived mesenchymal stem cells through stimulation of MCP-1 production in vascular smooth muscle cells. *J. Biol. Chem.* **284**, 17564–17574 (2009).
- Kitamura, T. *et al.* CCL2-induced chemokine cascade promotes breast cancer metastasis by enhancing retention of metastasis-associated macrophages. *J. Exp. Med.* **212**, 1043–1059 (2015).
- Chen, X. *et al.* CCL2/CCR2 regulates the tumor microenvironment in HER-2/neu-driven mammary carcinomas in mice. *PLoS ONE* **11**, e0165595 (2016).
- Williams, C. B., Yeh, E. S. & Soloff, A. C. Tumor-associated macrophages: Unwitting accomplices in breast cancer malignancy. *NPI Breast Cancer* **2**, 15025 (2016).
- Lim, S. Y., Yuzhalin, A. E., Gordon-Weeks, A. N. & Muschel, R. J. Targeting the CCL2-CCR2 signaling axis in cancer metastasis. *Oncotarget* **7**, 28697–28710 (2016).
- Mandal, P. K. *et al.* CCL2 conditionally determines CCL22-dependent Th2-accumulation during TGF- $\beta$ -induced breast cancer progression. *Immunobiology* **223**, 151–161 (2018).
- Mitkin, N. A. *et al.* Protective C allele of the single-nucleotide polymorphism rs1335532 is associated with strong binding of Ascl2 transcription factor and elevated CD58 expression in B-cells. *Biochim. Biophys. Acta Mol. Basis Dis.* **1864**, 3211–3220 (2018).
- Hoffman, M. M. *et al.* Unsupervised pattern discovery in human chromatin structure through genomic segmentation. *Nat. Methods* **9**, 473–476 (2012).
- Dutta, P., Sarkissyan, M., Paico, K., Wu, Y. & Vadgama, J. V. MCP-1 is overexpressed in triple-negative breast cancers and drives cancer invasiveness and metastasis. *Breast Cancer Res. Treat.* **170**, 477–486 (2018).
- Bai, F. *et al.* BRCA1 suppresses epithelial-to-mesenchymal transition and stem cell dedifferentiation during mammary and tumor development. *Cancer Res.* **74**, 6161–6172 (2014).
- Neve, R. M. *et al.* A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10**, 515–527 (2006).
- Wang, G., Matsuura, I., He, D. & Liu, F. Transforming growth factor- $\beta$ -inducible phosphorylation of Smad3. *J. Biol. Chem.* **284**, 9663–9673 (2009).
- Albers, R. E., Selesniemi, K., Natale, D. R. C. & Brown, T. L. TGF- $\beta$  induces Smad2 phosphorylation, are induction, and trophoblast differentiation. *Int. J. Stem Cells* **11**, 111–120 (2018).
- Derynck, R. & Zhang, Y. E. Smad-dependent and Smad-independent pathways in TGF- $\beta$  family signalling. *Nature* **425**, 577–584 (2003).
- Zhang, J. *et al.* The regulation of TGF- $\beta$ /SMAD signaling by protein deubiquitination. *Protein Cell* **5**, 503–517 (2014).
- Kulakovskiy, I. V. *et al.* HOCOMOCO: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).
- Chen, S.-J. *et al.* The early-immediate gene EGR-1 is induced by transforming growth factor- $\beta$  and mediates stimulation of collagen gene expression. *J. Biol. Chem.* **281**, 21183–21197 (2006).
- Gregory, K. J., Morin, S. M. & Schneider, S. S. Regulation of early growth response 2 expression by secreted frizzled related protein 1. *BMC Cancer* **17**, 473 (2017).
- Qian, Z. *et al.* Investigating the mechanism by which SMAD3 induces PAX6 transcription to promote the development of non-small cell lung cancer. *Respir. Res.* **19**, 262 (2018).

28. Pendaries, V., Verrecchia, F., Michel, S. & Mauviel, A. Retinoic acid receptors interfere with the TGF- $\beta$ /Smad signaling pathway in a ligand-specific manner. *Oncogene* **22**, 8212–8220 (2003).
29. Zhao, Y. *et al.* TGF- $\beta$  transactivates EGFR and facilitates breast cancer migration and invasion through canonical Smad3 and ERK/Sp1 signaling pathways. *Mol. Oncol.* **12**, 305–321 (2018).
30. Fortin, J. & Bernard, D. J. SMAD3 and EGR1 physically and functionally interact in promoter-specific fashion. *Cell. Signal.* **22**, 936–943 (2010).
31. Jungert, K. *et al.* Sp1 is required for transforming growth factor- $\beta$ -induced mesenchymal transition and migration in pancreatic cancer cells. *Cancer Res.* **67**, 1563–1570 (2007).
32. Mitkin, N. A., Korneev, K. V., Gorbacheva, A. M. & Kuprash, D. V. Relative efficiency of transcription factor binding to allelic variants of regulatory regions of human genes in immunoprecipitation and real-time PCR. *Mol. Biol.* **53**, 346–353 (2019).
33. Najafi, M. *et al.* Macrophage polarity in cancer: A review. *J. Cell. Biochem.* **120**, 2756–2765 (2019).
34. Lohoff, M. *et al.* Early growth response protein-1 (Egr-1) is preferentially expressed in T helper type 2 (Th2) cells and is involved in acute transcription of the Th2 cytokine interleukin-4. *J. Biol. Chem.* **285**, 1643–1652 (2010).
35. Veremeyko, T., Yung, A. W. Y., Anthony, D. C., Strelakova, T. & Ponomarev, E. D. Early growth response gene-2 is essential for M1 and M2 macrophage activation and plasticity by modulation of the transcription factor CEBP $\beta$ . *Front. Immunol.* **9**, 2515 (2018).
36. Baron, V., Adamson, E. D., Calogero, A., Ragona, G. & Mercola, D. The transcription factor Egr1 is a direct regulator of multiple tumor suppressors including TGF $\beta$ 1, PTEN, p53, and fibronectin. *Cancer Gene Ther.* **13**, 115–124 (2006).
37. Yamamoto, A., Kakuta, H., Miyachi, H. & Sugimoto, Y. Involvement of the retinoid X receptor ligand in the anti-inflammatory effect induced by peroxisome proliferator-activated receptor. *PPAR Res.* **2011**, 840194 (2011).
38. Sun, Z. *et al.* EGR1 recruits TET1 to shape the brain methylome during development and upon neuronal activity. *Nat. Commun.* **10**, 3892 (2019).
39. Spaapen, F. *et al.* The immediate early gene product EGR1 and polycomb group proteins interact in epigenetic programming during chondrogenesis. *PLoS ONE* **8**, e58083 (2013).
40. Zhou, X., Chen, Z. & Cai, X. Identification of epigenetic modulators in human breast cancer by integrated analysis of DNA methylation and RNA-Seq data. *Epigenetics* **13**, 473–489 (2018).
41. Shajahan-Haq, A. N. *et al.* EGR1 regulates cellular metabolism and survival in endocrine resistant breast cancer. *Oncotarget* **8**, 96865–96884 (2017).
42. Chen, Y., Takeshita, A., Ozaki, K., Kitano, S. & Hanazawa, S. Transcriptional regulation by transforming growth factor  $\beta$  of the expression of retinoic acid and retinoid X receptor genes in osteoblastic cells is mediated through AP-1. *J. Biol. Chem.* **271**, 31602–31606 (1996).
43. Xu, Q. & Kopp, J. B. Retinoid and TGF- $\beta$  families: Crosstalk in development, neoplasia, immunity, and tissue repair. *Semin. Nephrol.* **32**, 287–294 (2012).
44. Guillem-Llobat, P. & Íñiguez, M. A. Inhibition of lipopolysaccharide-induced gene expression by liver X receptor ligands in macrophages involves interference with early growth response factor 1. *Prostagland. Leukot. Essent. Fat. Acids* **96**, 37–49 (2015).
45. Mitkin, N. A., Muratova, A. M., Schwartz, A. M. & Kuprash, D. V. The A allele of the single-nucleotide polymorphism rs630923 creates a binding site for MEF2C resulting in reduced *cxc5* promoter activity in B-cell lymphoblastic cell lines. *Front. Immunol.* **7**, 515 (2016).
46. Gorbacheva, A., Korneev, K., Kuprash, D. & Mitkin, N. The risk G allele of the single-nucleotide polymorphism rs928413 creates a CREB1-binding site that activates IL33 promoter in lung epithelial cells. *Int. J. Mol. Sci.* **19**, 2911 (2018).
47. Phanish, M. K., Wahab, N. A., Colville-Nash, P., Hendry, B. M. & Dockrell, M. E. C. The differential role of Smad2 and Smad3 in the regulation of pro-fibrotic TGF $\beta$ 1 responses in human proximal-tubule epithelial cells. *Biochem. J.* **393**, 601–607 (2006).
48. Morishita, Y. *et al.* siRNAs targeted to Smad4 prevent renal fibrosis in vivo. *Sci. Rep.* **4**, 6424 (2014).
49. Hasson, S. A. *et al.* High-content genome-wide RNAi screens identify regulators of parkin upstream of mitophagy. *Nature* **504**, 291 (2013).
50. Mitkin, N. A. *et al.* p63 and p73 repress CXCR5 chemokine receptor gene expression in p53-deficient MCF-7 breast cancer cells during genotoxic stress. *Biochim. Biophys. Acta Gene Regul. Mech.* **1860**, 1169–1178 (2017).
51. Mitkin, N. A. *et al.* P53-dependent expression of CXCR5 chemokine receptor in MCF-7 breast cancer cells. *Sci. Rep.* <https://doi.org/10.1038/srep09330> (2015).

## Author contributions

A.G., N.M., A.B. and D.K. designed the study. A.G., N.M., A.N.U., A.S.U. and K.K. performed experiments. A.G., N.M. and D.K. wrote the manuscript. All authors analyzed data and contributed to the final version of the manuscript.

## Funding

Most of the promoter analysis (Figs. 1, 4) was supported by Grant 18-54-45020 from Russian Foundation for Basic Research. Gene expression studies (Fig. 2) were supported by Grant 19-74-00096 from Russian Science Foundation. Deletion screening studies (Fig. 3) were supported by Grant 19-14-00341 from Russian Science Foundation. Chromatin immunoprecipitation studies and point mutagenesis (Figs. 5, 6, 7) were supported by Grant 075-15-2019-1660 from the Ministry of Science and Higher Education of the Russian Federation.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-93561-6>.

**Correspondence** and requests for materials should be addressed to N.A.M.

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

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021