1080–1094 *Nucleic Acids Research, 2016, Vol. 44, No. 3* *Published online 12 October 2015*
doi: 10.1093/nar/gkv1032

DTIE, a novel core promoter element that directs start site selection in TATA-less genes

Nadav Marbach-Bar[1,†], Anat Bahat[1,†], Shaked Ashkenazi[1], Michal Golan-Mashiach[1], Ora Haimov[1], Shwu-Yuan Wu[2], Cheng-Ming Chiang[2], Anna Puzio-Kuter[3], Kim M. Hirshfield[3], Arnold J. Levine[3] and Rivka Dikstein[1,*]

[1]Dept. of Biological Chemistry, The Weizmann Institute of Science, Rehovot 76100, Israel, [2]Simmons Comprehensive Cancer Center, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA and [3]Cancer Institute of New Jersey, 195 Little Albany Street, New Brunswick, NJ 08903, USA

Received May 04, 2015; Revised September 09, 2015; Accepted September 26, 2015

ABSTRACT

The transcription start site (TSS) determines the length and composition of the 5′ UTR and therefore can have a profound effect on translation. Yet, little is known about the mechanism underlying start site selection, particularly from promoters lacking conventional core elements such as TATA-box and Initiator. Here we report a novel mechanism of start site selection in the TATA- and Initiator-less promoter of miR-22, through a strictly localized downstream element termed DTIE and an upstream distal element. Changing the distance between them reduced promoter strength, altered TSS selection and diminished Pol II recruitment. Biochemical assays suggest that DTIE does not serve as a docking site for TFIID, the major core promoter-binding factor. TFIID is recruited to the promoter through DTIE but is dispensable for TSS selection. We determined DTIE consensus and found it to be remarkably prevalent, present at the same TSS downstream location in ≈20.8% of human promoters, the vast majority of which are TATA-less. Analysis of DTIE in the tumor suppressor p53 confirmed a similar function. Our findings reveal a novel mechanism of transcription initiation from TATA-less promoters.

INTRODUCTION

The site of transcription initiation is critical for productive gene expression as it determines the length and composition of the 5′ UTR of mRNAs, which can have profound effects on translation efficiency (1). A major element controlling transcription start site (TSS) selection of RNA polymerase II (Pol II)-transcribed genes is the core promoter (for review see (2–4)). The core promoter consists of regulatory sequences around the TSS that bind and recruit the general transcription machinery (5). In addition to its role in TSS determination, the core promoter influences transcription initiation rates (6–8) and integrates the signals transmitted by enhancer-bound transcription factors (3). Initiation sites directed by mammalian promoters appear either as a narrow cluster of nucleotides or as broadly dispersed sites (9). Those directing focused TSSs typically have a TATA-box and/or Initiator (Inr) elements or combination of these basic elements with others. For example the BRE acts only in conjunction with the TATA-box, and the DPE and MTE are strictly dependent on the Inr. The second class, with the generic name TATA-less promoters, is largely uncharacterized even though it constitutes a majority among all the promoters (10–13). Several functional studies led to identification of core elements, such as XCPE1/2 and sINR, in TATA-less promoters (14–16), but these are present in only a small fraction of human genes. Thus, our current knowledge about the structure and function of core elements governing the large number of TATA-less genes is limited.

Attempts to characterize the core promoter region using bioinformatics analyses of mammalian promoters have pointed to several elements enriched in the core promoter region of TATA-less genes (10,13,16,17), but experimental evidence that these motifs function as core elements is for the most part missing. Furthermore, commonly used motif-identifying computational programs that work by extracting over-represented 'words' in a list of sequences are limited in their predictive power. For example they tend to ignore short or divergent words that have a lower statistical score but can nevertheless be functional, or they cannot efficiently identify composite elements consisting of two words that are separated by a gap of an unknown length. This is why these programs failed to identify the Inr and the DPE elements in mammalian promoters (our unpublished observations). Identification of new core elements, apparently, still requires the use of conventional molecular tools.

*To whom correspondence should be addressed. Tel: +972 8 9342117; Fax: +972 8 9344118; Email: rivka.dikstein@weizmann.ac.il
†These authors contributed equally to the paper as first authors.

The miR-22 promoter is of considerable interest since it directs a strictly localized TSS in the absence of a TATA-box or an Inr. We therefore investigated this promoter as a prototype of TATA-less and Inr-less class, and report the identification of a novel and highly prevalent downstream core element that we termed DTIE (Downstream Transcription Initiation Element). DTIE has a strict location and it cooperates with an upstream element for precise TSS positioning and promoter strength. DTIE indirectly recruits the general transcription factor TFIID, which we found to be dispensable for TSS selection. DTIE is highly prevalent specifically in TATA-less genes. One of these is the TATA- and Inr-less promoter of p53 in which we found DTIE to be similarly functional. Interestingly, a rare polymorphism in a conserved position of p53 DTIE caused a modest, but significant reduction in promoter activity. Our findings revealed that DTIE directs TSS selection in a subset of TATA- and Inr-less genes associated with cancer.

MATERIALS AND METHODS

Cells and transfection

HEK293T cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum. Transfections in HEK293T cells were performed using the standard $CaPO_4$ method. For reporter assays, subconfluent cells were transfected in a 24-well plate using 25 ng of the reporter plasmid, 5 ng RSV-Renilla and 10 ng CMV-GFP. 24 h after transfection cells were harvested and their luciferase and renilla activities were measured. For primer extension and northern blot assays, 150 mm dishes were transfected with 10 μg of the reporter plasmids, and 400 ng CMV-GFP, and cells were harvested 48 h later.

RNA analyses

Primer extension—Total RNA was prepared using the TRIzol reagent (MRC inc.), according to the manufacturer's instructions. RNA preparations were treated with RQ1 DNAse I (Promega) to avoid contamination with genomic DNA. Primers complementary to the luciferase or pEGFP transcribed sequences were radiolabeled using [^{32}P]γ-ATP and PNK (Fermentas) for 1 h at 37°C, followed by ethanol precipitation. The radiolabeled primers were annealed to the RNA as follows: 5 min at 85°C then a short spin and incubation for 1 h at 50°C. Reverse transcription was done by adding AMV reverse transcriptase (Promega), 33 μM dNTPs and actinomicyn D (Sigma), followed by ethanol precipitation. The labeled cDNA was resolved on a urea-polyacrylamide gel, and then visualized with a PhosphoImager (Fuji, BAS 2500). The sequencing reaction shown is from pEGFP.

Northern blot of small RNAs—5–10 μg total RNA was loaded onto a 15% acrylamide gel, and run at 180V for 90 min, in 1X TBE. RNA was then transferred from the gel to a GeneScreen-Plus membrane (NEN), at 200 mÅ for 2 h, in 0.5X TBE. Next, the RNA was crosslinked to the membrane with UV irradiation. Pre-hybridization was carried out at 42°C for 2 h in hybridization buffer (5X SSC, 20 mM Na_2HPO_4 pH 7.2, 7% SDS and 2X Denhardt's solution), after which the ^{32}P-labeled probe (miR-22 and internal control) was added. Hybridization was carried out for 16–24 h at 42°C. The membrane was then washed three times at 42°C in a washing solution (3X SSC, 25 mM NaH_2PO_4 (pH 7.5), 5% SDS and 10X Denhardt's solution). Hybridization products were visualized using phosphoimager (Fuji, BAS 2500).

5′-Rapid Amplification of cDNA Ends (5′-RACE)—first strand cDNA was synthesized from 1.5 μg RNA, using reverse transcriptase (SuperScript II, Invitrogen). The cDNA was purified using QIAquick polymerase chain reaction (PCR) purification kit (Qiagen) and a poly-G tail was added to the cDNA 3′-end using terminal deoxynucleotidyl transferase (Promega) according to manufacture instructions. The cDNA was purified and PCR was performed using Phusion (NEB), with 5 μl of cDNA and 10 pmol of a reverse primer nested to the primers used for reverse transcription and a forward primer $D(C)_{17}$. A fraction of the PCR product was resolved on a 6% polyacrylamide gel.

Plasmid construction

The miR-22 promoter from −487 to +55 relative to the TSS was previously described (18). 3′ end dissections of miR-22 promoter were constructed by PCR with a common forward primer and different reverse primers. Their sequences appear in Supplementary Table S1. The replacement mutations in miR-22 positions +8 to +31 were created by PCR using a reverse primer bearing the mutation to amplify the fragment between 487 bp upstream to the TSS up to 42 bp downstream to the TSS (Supplementary Table S1). This fragment was then inserted into the pGL2-basic plasmid through SmaI and HindIII restriction sites. To generate the linker mutant the construct bearing the dissection up to position +8 bp relative to the TSS was used as a template. A fragment containing the sequence between positions +8 and +42 bp relative to the TSS was prepared by annealing of primers and phosphorylation, and then insertion into the HindIII site, causing a destruction of the site that created the linker sequence. The miR-22 expression mini-gene is described elsewhere (6). To construct the plasmids for the site specific stable cell lines three fragments of miR-22 promoter (WT, ΔDTIE and ΔE1) were cloned into the pCDNA/FRT vector by replacing the original CMV promoter with a PCR fragment consisting of the miR-22 promoter variant and the luciferase gene using RF cloning method (19,20). The p53 promoter (from −150 to +40 relative to the TSS, genomic coordinates chr17:7,687,449–7,687,639 of the GRCh38/hg38 assembly) was amplified by genomic PCR and cloned into the promoter-less pGL2-basic via SmaI and HindIII sites. To create the mutations, a primer bearing the mutation was used for the PCR amplification (Supplementary Table S1), and the fragment containing the mutation was inserted into the pGL2-basic in a similar manner. All constructs were verified by sequencing.

Generation of miR-22 promoter stable cell lines

Stable clones were generated as described (21). Parental 293T Flp-In™ cells (Invitrogen) were maintained in

Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 10% FBS, 1% penicillin/streptomycin, 1% L-glutamine and 5% Amphotericin B (Sigma). The Flp-In expression vectors described above were co-transfected together with pOG44 recombinase (Invitrogen), which mediates the integration into the FRT site of the Flp-In™ cells, in a 1:9 ratio in 10 cm dish. 48 h later hygromycin (100 μg/ml) was added and cells were grown until discrete foci of hygromycin-resistant cells were evident (≈3 weeks of selection). Individual colonies were transferred to 24-well plate using 3MM discs that were previously dipped in trypsin. Stable colonies were then analyzed for Luciferase activity.

Electrophoretic mobility shift assay

DNA oligonucleotides containing the DTIE sequence were fluorescently labeled on the 5′-end with Cy5, or HEX (Integrated DNA Technologies, Inc). The oligonucleotides were annealed in 20 μl in a concentration of 10 pmol/μl and used as probes to the reaction. The binding reactions containing 100 ng of poly(dI-dC) and 10 μg of HEC-1B nuclear extract prepared as described (22), with binding buffer consisting of 25 mM HEPES (pH 7.9), 50 mM KCl, 10 μM Zn, 1 mM DTT and 10% glycerol. The reaction mix was incubated on ice for 10 min after which 20 fmol probe was added for an additional 20 min. Competitor double-stranded DNAs were added prior to the addition of the probe. The muted double-stranded DNA sequences have the same sticky ends. The reactions were separated by native electrophoresis at 4°C in a 6.5% polyacryamide gel with 1× TBE buffer at 150 V. The gel visualized with the Typhoon 9400 instrument (Amersham Biosciencs).

Human TFIID was purified from a HeLa-derived cell line expressing FLAG-tagged TBP as described (23). For EMSA with miR-22 promoter and purified TFIID, we used the mini-gene constructs (6) as templates, and amplified the promoter region using miR-22 -65 forward, and luc +75 reverse primers. The PCR product was cut using SacII and the 165 bp product was extracted from the gel, end-labeled with [γ-^{32}P]ATP (Izotop) using PNK (Fermentas), and cleaned with the PCR purification kit (Qiagen). The binding reactions contained 250 fmol probe, 0.04 μg poly(dG-dC) (Sigma), 5 μg BSA, 5 mM $MgCl_2$, 5 mM β-ME and 10 ng of purified TFIID, with D100. The reaction mix was incubated for 30 min at room temp after which it was loaded and run for 6 h at 45 V on a 4% polyacrylamide gel in Tris-boric acid running buffer. The gel was dried and visualized using PhosphoImager (PMI™ Personal Molecular Imager, Bio-Rad).

Chromatin immunoprecipitation

HEK293T cells that were transfected with miR-22 promoter constructs (2 μg) in 100 mm plates were cross-linked with 1% formaldehyde for 10 min at room temperature. Chromatin extraction and immuneprecipitations were carried out as previously described (24). DNA samples were analyzed by qPCR in an ABI 7300 Real Time PCR system using Power SYBR PCR reaction mix (ABI). Primer sequences are shown in the primer list (Supplementary Table S1).

Knockdown of TBP and TAF1

HEK293T cells were seeded on a 6-well plate and transfected with 100 nM Dharmacon siGENOME SMARTpool siRNA against TBP and TAF1 (Thermo Scientific) using DharmaFECT1 transfection reagent. The Dharmacon ON-TARGET*plus* Non-targeting siRNA #3 was used as a negative control. 48 h after the initial transfection, cells were transfected again with a mixture of the siRNA and the reporter plasmids. Cells were harvested 24 h after the second transfection for protein and RNA analyses.

Bioinformatics analysis

Human TSSs (30 nt range) and flanking sequences (−40 to +40) were retrieved from FANTOM5 (9) (http://fantom.gsc.riken.jp/5/sstar/Data_source) and UCSC genome browser (https://genome.ucsc.edu/), respectively, and analyzed for the consensus sequences of DTIE (GBBRDNHGG), TATA-box (TATAWA) and Inr (YYANWYY) by using 'DNA pattern' tool of the Regulatory Sequence Analysis Tools (http://rsat.ulb.ac.be/rsat/), analyzing only direct strand and allowing one substitution. For each element only sequences starting from positions indicated in the results were counted. Gene ontology analysis was done using a web-based tool of the Gene Ontology Consortium (http://geneontology.org/).

Analysis of p53 SNP rs17884410 in breast cancer patients

The p53 SNP rs17884410 (C/T) was analyzed using genomic DNA isolated from peripheral blood from a breast cancer cohort consisting of consecutively enrolled, consenting patients and healthy volunteers, invited to participate in the prospective study from 2004-present at The Cancer Institute of New Jersey (CINJ). The cohort included the following races: Caucasian (77%), African American (6%), Asian (6%), Hispanic (6%), other (5%). Ductal (84%), lobular (11%) and other (5%) subtypes of breast cancer were represented. The genomic DNA was prepared using a spin column-based method according to the manufacturer's protocol (QIAGEN). Genotyping for the p53 SNP was performed using Taqman assays on the ABI 7900HT Fast Real-Time PCR System (Applied Biosystems). Briefly, reactions were performed using 5–10 ng genomic DNA in 10 μl volume. PCR cycling conditions were 50°C for 2 min, 95°C for 10 min, followed by 40 cycles of 92°C for 15 s and 60°C for 1 min. Non-template controls (NTC) were used in every reaction to control for possible contaminations. Control plasmids for the assay were generated by cloning out a 480 bp region of the p53 gene from normal genomic DNA that carried the wild-type homozygous (TT) genotype. Mutagenesis using the Agilent Technologies Quick Change Lightning Site Directed Mutagenesis kit generated the homozygous (CC) genotype plasmid.

RESULTS

Identification of a downstream transcription initiation element in the miR-22 promoter

The human miR-22 is a Pol II-transcribed gene. Its promoter lacks TATA-box and Inr, but it nevertheless directs

a strictly localized TSS, which is unusual among TATA-less promoters (18). We therefore searched for a regulatory element involved in transcription initiation of this promoter. We had already mapped the minimal region of the miR-22 gene displaying full promoter activity to positions −487 to +55 relative to the TSS and identified 2 upstream enhancer elements necessary for full promoter activity (18). To determine whether the region downstream of the TSS also bears initiation regulatory activity we constructed a series of progressive deletions from the 3′-end of miR-22 promoter, and placed them upstream of a luciferase gene. The constructs were transfected into cells and 48 h later total RNA was prepared and analyzed by primer-extension using a ^{32}P-labeled primer complementary to the beginning of the luciferase mRNA (Figure 1A). In addition luciferase activity was measured 24 h after transfection (Figure 1C). Figure 1A shows that transcription driven by the miR-22 promoter (−487 to +42) initiates from a major TSS at exactly the same location as in the endogenous gene (18). No significant differences in promoter activity were observed upon deletions of sequences up to position +31, but deletion of an additional 6, 9 and 23 nt caused significant decrease in the signal of the major TSS, while the upstream minor TSSs (sequence shown in Figure 1B) were enhanced (Figure 1A). We also monitored the luciferase activity, which records the transcription from all the TSSs, and found good correlation with the primer extension results (Figure 1C), suggesting that TSS selection and promoter strength are coupled. These findings indicate that the downstream region contains a regulatory element that is important for promoter strength and transcription initiation. As the sequence of this region does not resemble known regulatory motifs we termed it DTIE for Downstream Transcription Initiation Element.

To examine the functional importance of DTIE for the miRNA expression we used a previously described minigene (6) consisting of the promoter (wild-type or DTIE mutant), the first exon, part of the intron and the second exon from which the miR-22 is processed (Figure 1D). These constructs were transfected into cells together with a plasmid directing expression of a 42 nucleotide small RNA under the control of the H1 promoter which served as internal control for transfection efficiency. 48 h later the levels of the miRNA and the control small RNA were determined by northern blot. As shown in Figure 1D, DTIE deletion caused a significant reduction in miR-22 levels, indicating that this element is also required for the proper expression of miR-22.

To define the boundaries of DTIE we constructed a series of mutations in the context of the fully active promoter (−487 to +42), by substituting three nucleotides at a time, from position +11 to +31. Since this DNA segment is rich in G and C, each triad was replaced with ATT nucleotides (Figure 2A). One exception is Mut4 in which the original T was replaced by a G. These constructs were transfected into cells and luciferase activity was measured. Mutations 1, 3 and 4 reduced luciferase activity, while mutations 2, 5, 6 and 7 had no effect (Figure 2B). To assess the effect of the mutations on mRNA levels and TSS location we performed a primer extension assay. TSS location was not changed by any of the mutations, however mutations 1 and 3 caused significant reduction in mRNA levels (Figure 2C, and Supplementary Figure S1). Mutation 4, which reduced luciferase levels, showed no reduction in mRNA levels. This incompatibility is most likely a consequence of introduction of an upstream ATG, that is not in-frame with the luciferase initiation codon, thereby reducing translation efficiency. Mut1 and Mut3 had a modest effect relative to the deletion of the entire motif, therefore we constructed a double mutant (Mut1+3) and found that its effect on the major TSS is greater than each one alone (Supplementary Figure S2). Interestingly, the intensity of the minor upstream TSS is increased in the double mutant (Supplementary Figure S2), as also noted upon deletion of DTIE (Figure 1A), masking the effect of these mutations on the luciferase activity (Supplementary Figure S2B). These findings suggest that DTIE lies between positions +23 and +31 relative to the TSS, in which two sub-regions, +23 to +25 and +29 to +31, are the most important for its function. By combining the mutagenesis data with the conservation of this sequence among all miR-22 orthologs (Figure 2D) we defined its consensus as G(G/C)G(A/G)(A/G/T)N(A/C/T)GG or GSGRDNHGG. To examine the sequence requirement of DTIE in more detail, each position in the two sub-regions (23–25 and 29–32) was further substituted by nucleotides that do not constitute the consensus. Analysis of the luciferase activity of these mutants confirmed the requirement of the sub-domains and expanded the consensus to GBBRDNHGG (Figure 2E).

DTIE cooperates with an upstream element to direct transcription initiation

To determine whether the location of DTIE relative to the TSS is important for its function, we first introduced 5 and 10 nt linkers between the TSS and DTIE at the +8 position that is beyond the upstream boundary of the motif (Figure 3A). The 5 nt linker introduces ≈a half helical turn, therefore is expected to alter the spatial arrangement, whereas the 10 nt primarily changes the distance of DTIE relative to the rest of the promoter. WT- or linker-bearing promoter constructs were transfected into cells and luciferase activity was determined. Addition of both downstream linkers resulted in decreased promoter activity (Figure 3A). We then conducted a 5′ RACE assay to determine the effect of the linkers on the TSS location. With the 5 nt linker initiation from the original TSS was diminished to almost undetectable levels (Figure 3B, red arrowhead), but new TSSs appeared downstream to the original one (Figure 3B, black arrowheads). Moreover the linker also caused significant enhancement of the minor upstream TSS, so the promoter now resembles TATA-less promoters that drive multiple dispersed TSSs (Figure 3B). A similar enhancement of the minor upstream TSSs was also seen upon DTIE deletion or Mut1+3 (Figure 1A and Supplementary Figure S2). The effect of the 5 nt linker on the TSS was verified by primer extension (Supplementary Figure S3). With the 10 nt linker the TSSs were almost undetectable (Figure 3B). To define the maximal distance of DTIE relative to the TSS for strict TSS selection we also inserted 2, 3 and 4 nt at the +8 position. Promoter activity was gradually decreased in proportion to the length of the linker, with the 4 nt linker hav-

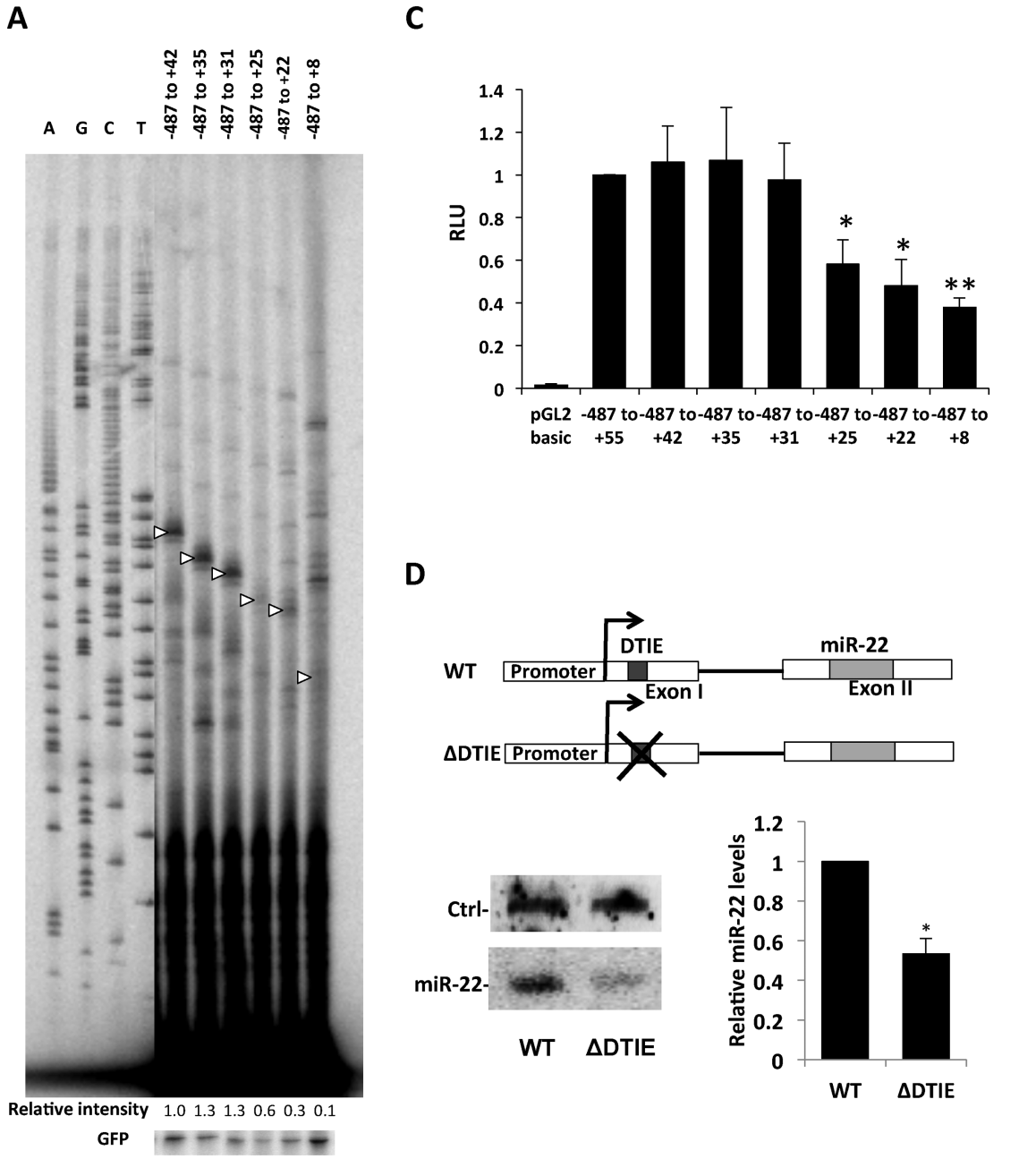
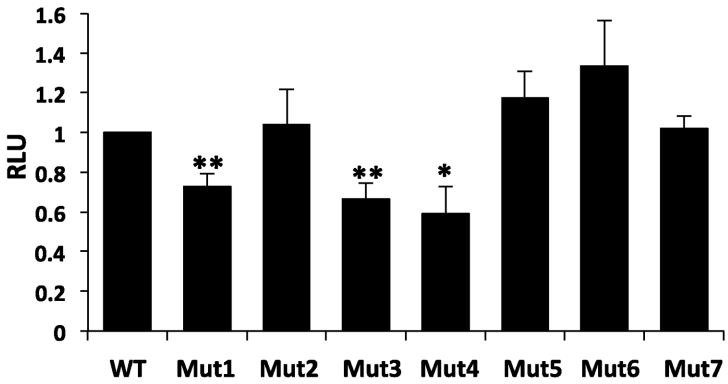


Figure 1. Identification of a downstream element (DTIE) that influences promoter strength and TSS location. (**A**) A series of successive 3′ end deletions generated in the miR-22 promoter that was cloned upstream to luciferase reporter gene. These constructs were transfected into cells together with CMV-GFP. Primer extension of luciferase and GFP was performed to analyze mRNA levels and TSS location. White arrowheads point to the expected TSS location. Quantification of the intensity of the major TSS of miR-22 relative to that of the GFP TSS is shown at the bottom. The gel is a representative of two independent experiments. The sequencing lanes were spliced to bring them close to the primer extension lanes. (**B**) The sequence of the miR-22 proximal promoter region with the major TSS marked in bold and the minor altered TSSs double-underlined. (**C**) Luciferase assay was used to analyze promoter strength. Luciferase activity throughout the paper was normalized to the co-transfected RSV-renilla activity. The promoter fragment between −487 and +55 relative to the TSS, which was previously shown to bear full promoter activity, was set 1. The results represent average ± SE of at least 4 independent transfection experiments. * and ** denote $P < 0.05$ and 0.005, respectively. Statistical significance was calculated by Student t-test throughout the paper. (**D**) A miR-22 mini-gene under the regulation of the miR-22 intact promoter (−487 to +55), or a DTIE deleted version (ΔDTIE, −487 to +8) was transfected into cells together with a plasmid directing a 42 nt small RNA that serves as internal control. 48 h after transfection total RNA was prepared and the level of miR-22 produced by both promoters and the control RNA were determined by northern blot of small RNAs. The graph shows the average of 3 independent experiments quantified by densitometry, in which the WT was set to 1. * denotes $P < 0.05$.

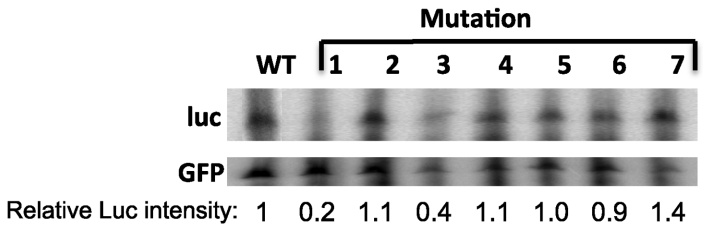
A

```
          +1          +11                         +31          +42
WT:       AAGAGACAGCGCCGCCGGCCGTGGGGAGCGGACGCAGTGATT
Mut1:     AAGAGACAGCGCCGCCGGCCGTGGGGAGATTACGCAGTGATT
Mut2:     AAGAGACAGCGCCGCCGGCCGTGGGATTCGGACGCAGTGATT
Mut3:     AAGAGACAGCGCCGCCGGCCGTATTGAGCGGACGCAGTGATT
Mut4:     AAGAGACAGCGCCGCCGGCATGGGGGAGCGGACGCAGTGATT
Mut5:     AAGAGACAGCGCCGCCATTCGTGGGGAGCGGACGCAGTGATT
Mut6:     AAGAGACAGCGCCATTGGCCGTGGGGAGCGGACGCAGTGATT
Mut7:     AAGAGACAGCATTGCCGGCCGTGGGGAGCGGACGCAGTGATT
```

B



C



D

```
          +1                                    +23      +31
Human     GCCGAGAAGAGACAGCGCCGCCGGCCGTGGGGAGCGGACGC
Mouse     GCTGGGAAGAGACAGAGCGGTCGGCCGTGCGGACAGGTCGC
Rat       GCTGGGAAGAGACAGAGCGGTCGGCCGTGCGGAGCGGTCGC
Dog       GCAGGGAAGACACAGCGCCGCCGGCCGTGGGGGATGGACGC
Opossum   GCTGGGGAGAGACGGTGGCGCCGGCCGAGGGGAGCGGACGC
Chicken   GCTGGGGAGAGACGGTGGCGCCGGCCGAGGGGAGCGGACGC
```

E

Mutated nucleotide:

	+23	+24	+25	+29	+30	+31
A	0.6 ± 0.0	0.1 ± 0.1	0.2 ± 0.1	-	-	0.4 ± 0.1
T	0.6 ± 0.0	0.9 ± 0.1	1 ±0 .1	-	0.2 ± 0.1	0.6 ± 0.1
C	0.3 ± 0.1	-	-	-	0.2 ± 0.1	0.7 ± 0.1
G	-	-	-	0.6 ± 0.1	-	-

Consensus:
GBBRDNHGG

Figure 2. Determination of the DTIE boundaries. (**A**) A scheme showing a series of 3 nucleotides substitution mutants (bold letters) spanning +31 to +11 relative to the TSS. (**B**) The mutants shown in A were transfected into cells and their effect on luciferase activity was determined. The data represent the average ± SE of 4 independent experiments in which * and ** denote $P < 0.05$ and 0.005, respectively. (**C**) Analysis of the mutants shown in A using primer extension analysis of luciferase and GFP reporter genes as indicated. The full gel is shown in Supplementary Figure S1. Lane 7 was spliced to bring it closer to lane 6. The gel is a representative of two independent experiments. (**D**) Alignment of miR-22 TSS downstream sequences between different species. DTIE sequence is boxed. (**E**) Nucleotide substitution of the indicated positions of DTIE and their transcriptional activity relative to the WT promoter are shown (average of 3 independent experiments ± SE). The DTIE consensus sequence derived from this analysis is shown on the right.

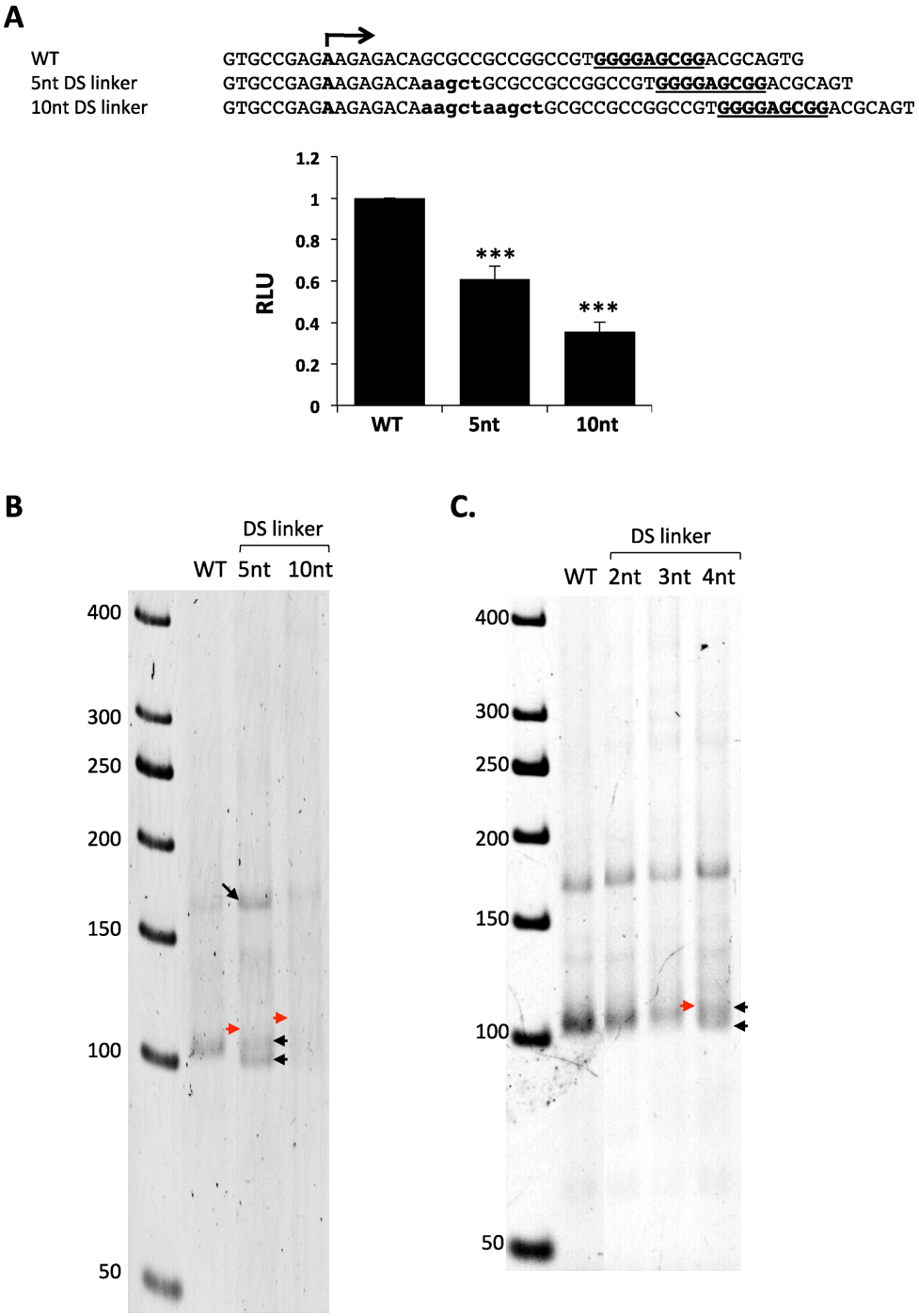


Figure 3. DTIE strict location is important for TSS selection. (**A**) A scheme showing the 5 and 10 nt linkers that were inserted between the DTIE and the TSS. The wild-type (WT) and the linker mutants were transfected into cells and promoter strength was analyzed by luciferase assay. The results represent the average ± SE of 7 (5 nt linker) and 3 (10 nt linker) independent experiments. *** denote $P < 0.0001$. (**B**) 5′ RACE analysis of the WT and the linker mutants. The red arrowhead point to the expected location of the TSS and the black arrowheads to the new TSSs. The marker lane was spliced to bring it closer to the other lanes. The gel is a representative of two independent experiments. (**C**) 5′ RACE analysis of the WT and 2, 3 and 4 nt linkers as described above. The WT lane was spliced to bring it closer to the other lanes.

ing the most damaging effect (Supplementary Figure S4). 5′ RACE analysis shows that addition of up to 3 nt linker retained the major TSS (Figure 3C) while the 4 nt linker affected the TSS similar to the 5 nt described above. We conclude that the native location of DTIE is critical for the specific start site selection as well as promoter strength.

The dramatic change in the level and position of the TSS upon alteration of DTIE location suggests that its specific arrangement and distance relative to another, upstream element is important for start site selection. To test whether this element is located upstream of the TSS we inserted a 5 nt linker at an upstream −8 position, a region that we validated, using bioinformatics, not to be part of a transcription factor-binding site. Like the downstream linker, the upstream linker decreases promoter strength (Figure 4A) and dramatically changes the TSS selection (Figure 4B). In attempting to identify the putative cooperating upstream element, we considered the previous characterization of the miR-22 promoter (18), in which two upstream enhancer elements had been revealed: the first is located in the region spanning positions −304 to −75 (E1) and the second lying between −75 and −65 (E2) relative to the TSS (Figure 4C). To determine which of these elements acts together with DTIE we examined the effect of DTIE mutation in the context of E1 and E2. As expected, removal of DTIE when both E2 and E1 are present (−487 to +8) caused significant reduction in reporter gene activity (Figure 4C). When E1 is deleted and the promoter activity is directed only by E2 (−75 to +55), the deletion of DTIE (−75 to +8) made no difference in its activity (Figure 4C). The reporter activity of the E2 alone is significantly reduced but is still more than 14-fold higher than the promoter-less construct (pGL2 basic column). These results suggest that DTIE is functional in the presence of E1 but not E2. Primer extension assay revealed that deletion of E1 was sufficient to eliminate the major TSS (Figure 4D), reminiscent of the effect of DTIE mutation. Thus DTIE appears to cooperate with a specific, relatively distal element to direct transcription initiation from a specific site. We next wished to examine whether DTIE is also functional in a chromatin context. For this purpose we generated stable cell lines in which the WT, DTIE- and E1-deleted promoters were integrated as a single copy at a specific site using the Flp-In-293 cell system (see Experimental Procedure for details). From each promoter variant we analyzed the luciferase activity of 6 independent colonies (Figure 4E) and also determined the TSS using 5′ RACE (Figure 4F). These measurements clearly show that both DTIE and E1 are essential for promoter activity also in a chromatin context.

To further map the DTIE upstream cooperating element (DUCE), the E1 region was further dissected (–227 to +55; −160 to +55 and −88 to +55) and the mutant constructs were transfected and analyzed by 5′ RACE. The results revealed that the TSS was retained with a promoter fragment spanning −160 to +55 but was lost with −88 to +55, suggesting that it lies between −160 and −88 (Figure 5B). Inspecting the sequence of this region revealed 3 Sp1 binding sites at −161, −123 and −98 (Figure 5A). Mutation in each of these sites did not change the TSS (Figure 5C, left). Likewise double Sp1 mutants did not alter the TSS position but did cause a reduction in its intensity (Figure 5C, right). From these results we concluded that the DUCE is probably not Sp1. To narrow down DUCE location, we introduced 5 nt linkers at −199, −137, −111 and −74 positions of the E1. The linker scanning analysis suggests that DUCE lies between −137 and −111 (Figure 5D). Analysis of the −137 to −111 sequence with a transcription factor search program did not reveal a binding site of a known transcription factor except for Sp1.

TFIID is dispensable for TSS selection of miR-22 promoter

Core promoter elements, such as TATA-box, Inr, DPE and DCE act through binding the general transcription factor TFIID to promote the assembly of the pre-initiation complex (4,25). We therefore examined whether the miR-22 promoter binds TFIID through the DTIE. For this purpose TFIID was purified from HeLa cell line expressing FLAG-tagged TBP (Supplementary Figure S5A) as previously described (23) and used for *in vitro* binding reactions with a ^{32}P-labeled miR-22 promoter fragment (from −65 to +55). The complexes were then resolved by Mg^{++}-containing native gel electrophoresis, which was shown to be most suitable for the analysis of TFIID-promoter complexes (26,27). As a positive control we used the same DNA fragment but with a TATA-box inserted at position –30 relative to the TSS. As can be seen in Figure 6A, the native miR-22 promoter failed to form a complex with TFIID (lanes 3 and 4) while the TATA-box-containing miR-22 promoter did (lanes 1 and 2), suggesting that DTIE on its own cannot recruit TFIID. In certain promoters TFIID forms contacts not only with the TATA-box but also with sequences downstream of the TATA-box up to +35 position. We therefore examined whether DTIE contributes to TFIID association with the TATA-box-containing promoter by analyzing the DTIE deleted promoter. As shown in Figure 6A (lanes 5 and 6 compared to lanes 1 and 2) the level of TFIID association with the promoter was unchanged upon DTIE deletion, indicating that DTIE is neither primary nor secondary docking element for TFIID.

While TFIID does not bind DTIE directly we examined whether TFIID would still be required for DTIE activity. For this purpose we transfected into cells the wild type promoter, DTIE mutant Mut1 which diminished transcription from the major TSS (see Figure 2), and the 5 nt linker mutant that derestricted the TSS (Figure 3). Using chromatin immunoprecipitation assay (ChIP) we analyzed the occupancy of these promoter variants by Pol II, and the TFIID subunits TBP and TAF1. TAF1 is TFIID specific whereas TBP is present in other transcription regulatory complexes (28,29). As can be seen in Figure 6B, Pol II, TBP and TAF1 are all enriched on the wild-type promoter. With DTIE Mut1 the levels of Pol II, TBP and TAF1 are significantly reduced. With the linker mutant only Pol II levels are reduced, whereas TBP and TAF1 remain almost unaffected. The *in vitro* binding (Figure 6A) and the ChIP assays (Figure 6B) together suggest that TFIID is recruited indirectly and is dependent more on the sequence of DTIE than its specific location.

To examine whether TFIID is involved in the strict TSS selection we analyzed the effect of TBP and TAF1 depletion on the position of the TSS using siRNA. Cells were trans-

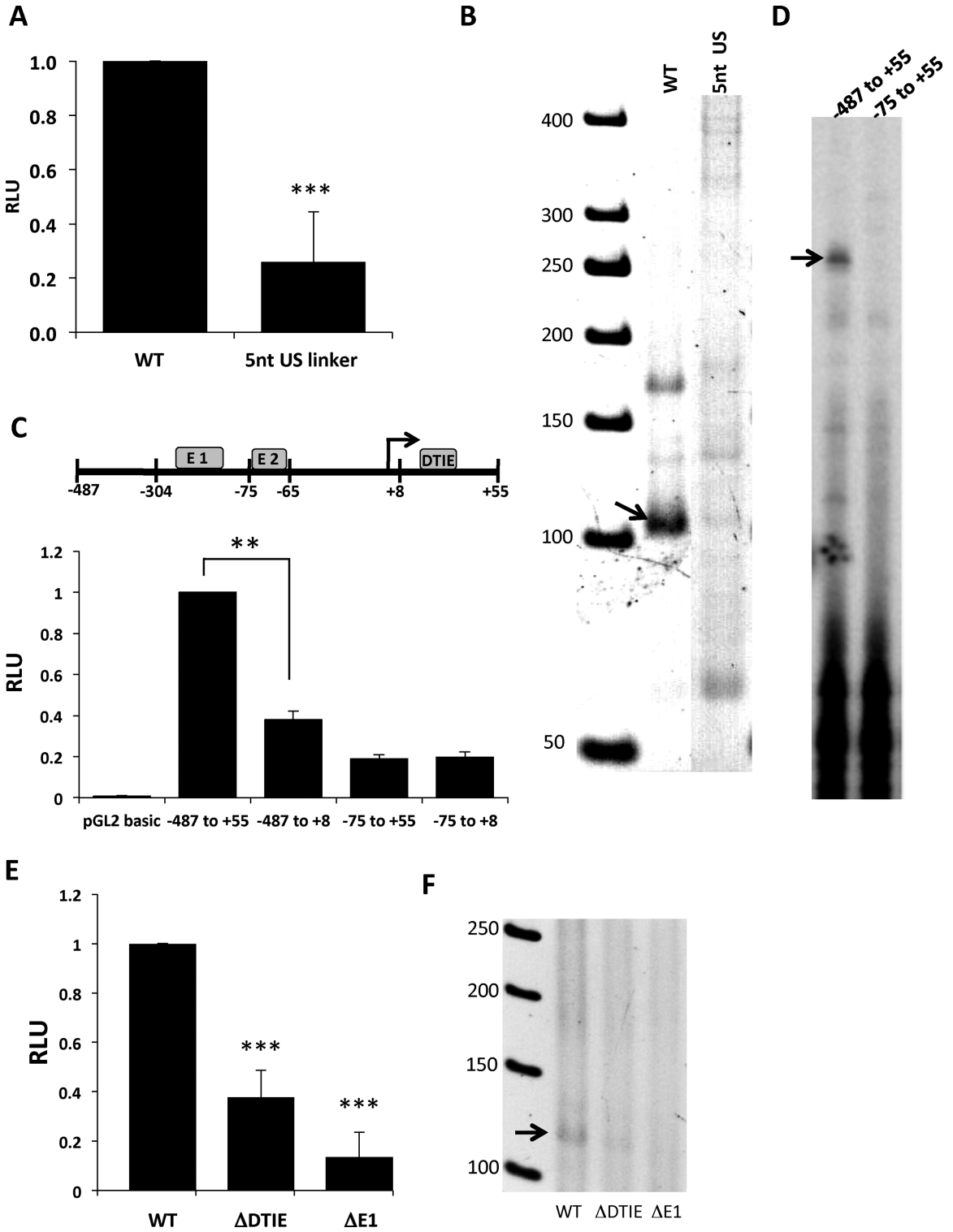


Figure 4. DTIE cooperates with an upstream, strictly localized element. (**A**) A 5 nt linker was inserted at the -8 position relative to the TSS, and the wild-type (WT) and linker mutant were analyzed by luciferase assay. The results represent the average ± SE of 5 independent experiments. *** denotes $P < 0.0001$. (**B**) 5′ RACE analysis of the WT and the upstream linker mutant described in A. The marker lane was spliced to bring it closer to the other lanes. The gel is a representative of two independent experiments. (**C**) The effect of DTIE deletion on the full promoter (containing E1 and E2 enhancers) and on a promoter construct lacking E1. The results represent the average ± SE of 8 independent experiments. ** denotes $P = 0.0001$. (**D**) Primer extension analysis of the wild type (−487 to +55) and the construct lacking E1 (−75 to +55). (**E** and **F**) DTIE and E1 are functional in a chromatin context. WT, DTIE and E1 deleted promoter-luciferase genes (ΔDTIE and ΔE1, respectively) were cloned in pcDNA5/FRT by replacing the CMV promoter. These constructs were transfected into Flp-In-293 cells to generate site-specific stable cell lines. Luciferase activity (E) and 5′ RACE analysis (F) of each construct was determined and normalized to total protein or RNA levels, respectively. The results of the luciferase activity represent the average ± SE of 6 clones from each construct. *** denote $P < 0.0001$.

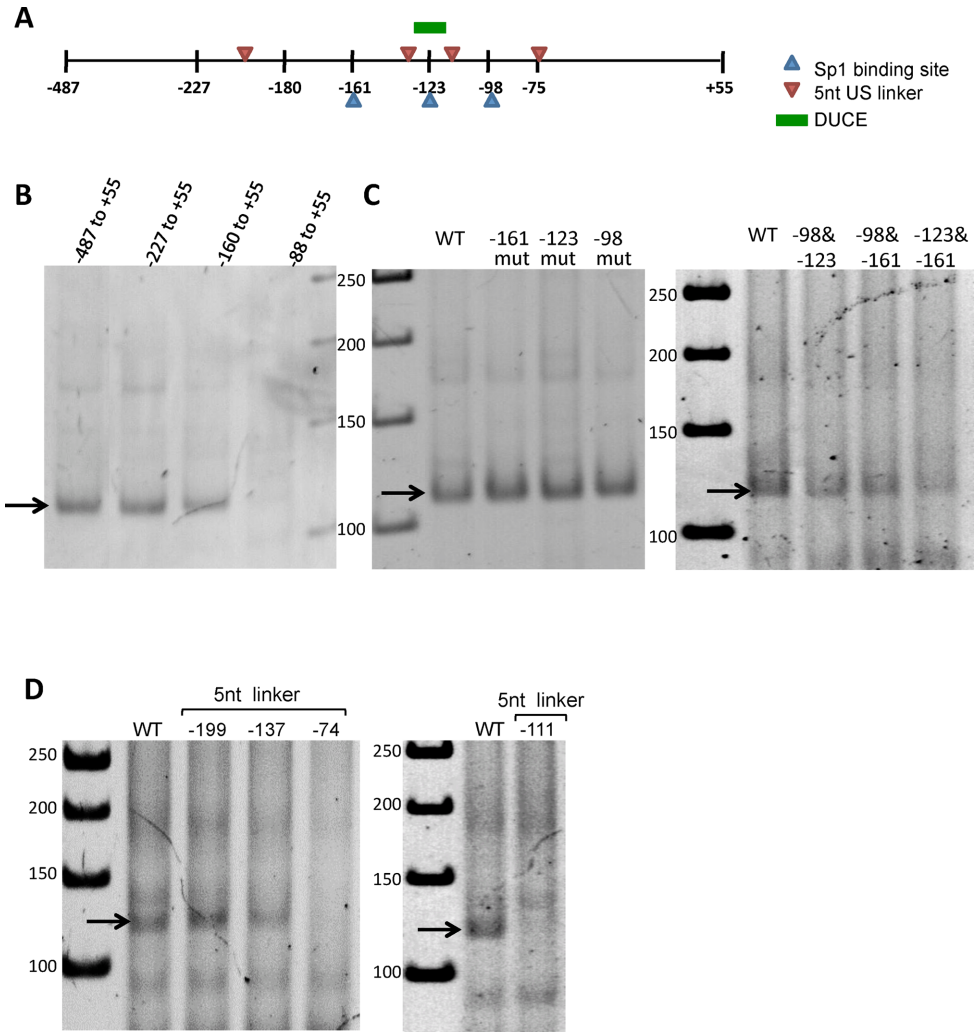


Figure 5. Mapping of DTIE upstream cooperating element (DUCE). (**A**) A scheme of the miR-22 promoter and the positions of the Sp1 binding sites, 5 nt linker insertions and the DUCE region. (**B**) The effect of 5′ end deletion mutants of the miR-22 promoter, as indicated on the top, on TSS position using 5′ RACE. (**C**) The impact of single and double mutations in the 3 Sp1 binding sites on the major TSS of the miR-22 promoter. (**D**) 5′ RACE analysis of 5 nt linkers inserted at the indicated positions. The gels are representatives of two independent experiments.

fected with siRNA and after 48 h transfected again with siRNA and with miR-22-Luc promoter. Immunoblot of the transfected cells showed efficient KD of both TAF1 and TBP (Figure 6C, left). RNA was extracted and subjected to TSS determination using 5′ RACE. The results revealed that the positions of the major and minor TSSs were not affected by TBP and TAF1 depletion (Figure 6C, right) indicating that TFIID is probably not required for the TSS selection directed by DTIE.

In a search for a specific DTIE-binding protein we used DNA binding assays with nuclear extract. A double stranded and labeled DNA fragment bearing DTIE and flanking sequences was incubated with nuclear extract and the DNA-protein complexes were analyzed by native gel electrophoresis. The DTIE-containing probe generated 3 major protein-DNA complexes (Figure 6D). The specificity of these complexes was analyzed by competition assays using excess unlabeled wild-type or DTIE-mutant oligonucleotides or a DNA fragment with the GC-rich Sp1 sequence (Figure 6D). The results revealed competition for complex B by the wild-type but not by the DTIE-mutant or the Sp1 oligos, indicating that complex B seems to be DTIE specific. We added antibodies against TBP and several TBP-associated factors (TAF1, 3, 4, 5, 6, 9, 10 and 12) to the DNA binding reactions, but none of them caused a shift in the migration of the complex (data not shown), meaning that the complex detected in the nuclear extract is distinct from TFIID.

DTIE is highly prevalent and functional in protein coding genes

Next we set out to examine the potential functional significance of DTIE in protein coding genes. For this purpose we determined the prevalence of DTIE among human promoters, which we retrieved from the recently published TSS data of the FANTOM5 project (see Experimental Procedures for details). The sequences were searched for the presence DTIE consensus sequence defined above (with up to 1

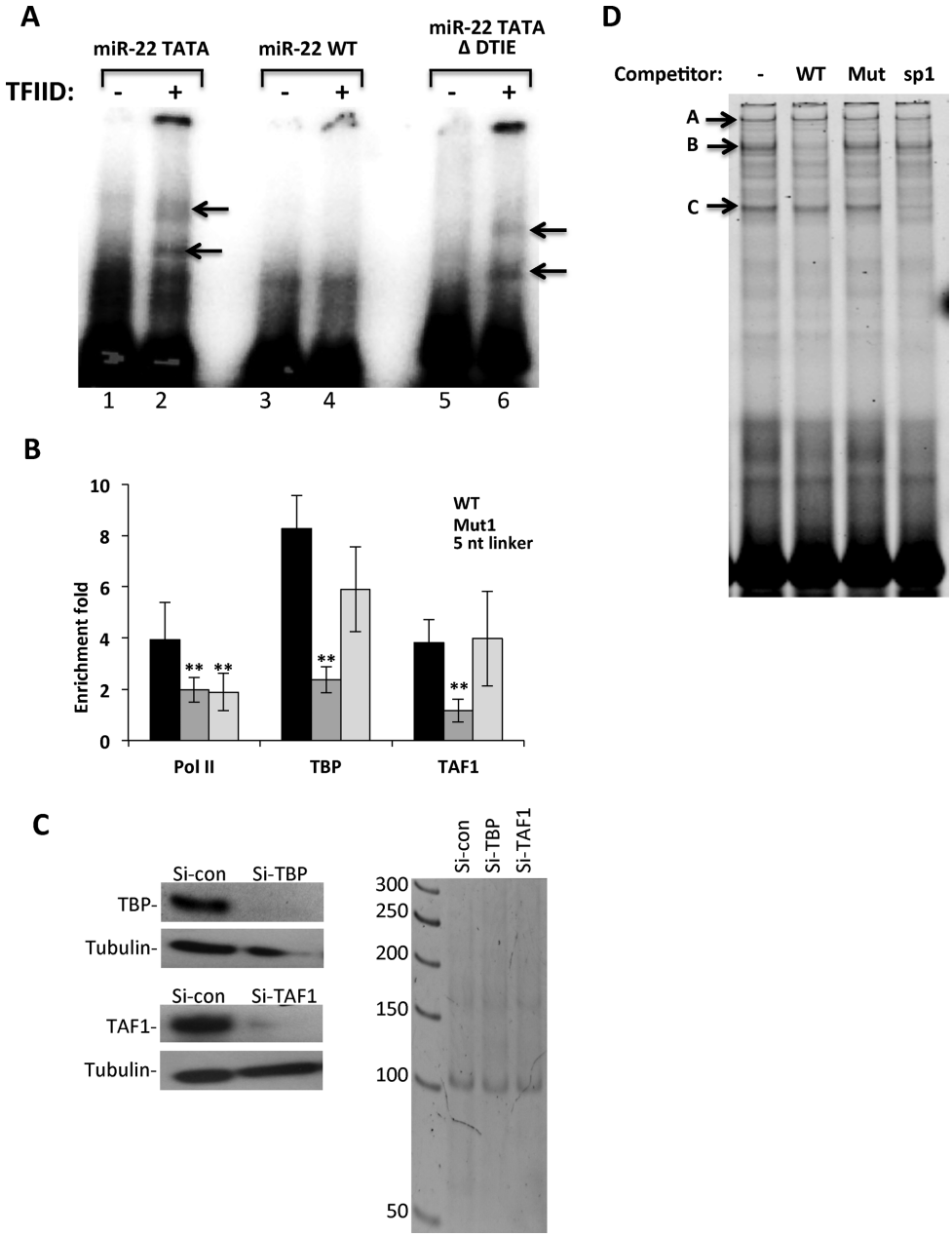


Figure 6. TFIID is indirectly recruited by DTIE. (**A**) DNA fragments containing either the wild-type miR-22 promoter (lanes 3–4), or a mutant bearing a TATA-box (lanes 1–2) or a mutant bearing a TATA-box but without DTIE (lanes 5–6), were end-labeled and used for *in vitro* binding assay with 10 ng of purified TFIID (+ lanes). The DNA-TFIID complexes were separated on Mg-containing native gel. The gel is a representative of 3 independent experiments. (**B**) HEK293T cells, that were transfected with WT, Mut 1 or 5 nt linker miR-22 promoter variants (described in Figures 2 and 3), were subjected to chromatin immunoprecipitation (ChIP) using Pol II, TBP, TAF1 and a non-relevant control antibodies. The immunoprecipitated chromatin was analyzed by qPCR and normalized to the input DNA and to the level of the control antibody. The graphs represent the average ± SE of 3–5 independent experiments. The asterisks indicate statistically significant difference ($P < 0.005$) relative to the control. (**C**) HEK293T cells were transfected with control, TBP or TAF1-specific siRNA and 48 h later cells were transfected again with the siRNA and miR-22 promoter-luciferase reporter. After 24 h the KD efficiency was verified by immunoblot (left) and the TSS position was determined by 5′ RACE (right). The gel is a representative of two independent experiments. (**D**) EMSA using a double stranded oligonucleotide containing DTIE sequence (DTIE WT) in the presence of 10 μg nuclear extract. WT, DTIE mut (1+3, shown in Figure 2A) and Sp1 double stranded oligonucleotides were used as competitors (25-fold excess) as indicated on the top of the lanes. The gel is a representative of two independent experiments.

Table 1. The prevalence of DTIE in human promoters and the co-occurrence of other core and proximal elements with DTIE

	Prevalence in all promoters (%)	Prevalence in DTIE promoters (%)
DTIE	20.8	100
TATA	15	1.6
Initiator	12.1	12.5

mismatch), specifically in positions +21 to +25 relative to the TSS. The prevalence of DTIE was found to be 20.8% (Table 1). A similar analysis of TATA box (TATAWA, with up to 1 mismatch) in positions −25 to −35 relative to the TSS and the Inr (YYANWYY, with up to 1 mismatch) at −2 relative to the TSS, revealed a 15% and 12.1% frequencies, respectively (Table 1). We next determined the frequency of the TATA-box and the Inr in DTIE-containing promoters. While the frequency of the Inr among DTIE promoters is similar to its general frequency (12.5% versus 12.1%), the TATA-box was found to be substantially under-represented in DTIE genes (1.6% versus 15%) (Table 1), suggesting a selection against co-occurrence of DTIE and TATA-box in the same promoter. Gene ontology analysis revealed transcription factor activity as the most significant enriched term (2.3-fold enrichment, $P = 6.14e\text{-}9$).

Among DTIE genes we found many which are involved in tumorogenesis such as the tumor suppressors/oncogenes p53, RB1, CDK inhibitors, E2Fs, etc (Supplementary Table S2). To examine the importance of DTIE in a cancer-associated gene we analyzed the function of this element in the promoter of the tumor suppressor p53. The p53 promoter has been studied in detail but in none of these studies was the DTIE sequence analyzed (for example see (30–33)). We therefore cloned the human p53 promoter and introduced mutations in the putative DTIE. Our findings revealed that DTIE is clearly required for the full activity of the p53 promoter as mutations spanning the 5′ and the 3′ parts of the element significantly reduced luciferase expression (Figure 7A). Comparing the conservation of this motif among several mammalian species, we found that all fit DTIE subdomains consensus (Supplementary Figure S6).

A rare polymorphism rs17884410 is found exactly within p53 DTIE. This polymorphism changes the seventh position of DTIE from A to G, which is not compatible with DTIE consensus. Moreover, an A in this position is highly conserved among p53 mammalian orthologs. These interesting data prompted us to examine the impact of this variation on p53 promoter activity, by substituting the A with a G and analyzing the corresponding luciferase activity. As shown in Figure 7A this change caused a modest, but significant reduction in p53 promoter activity. Considering cell fate decisions are highly sensitive to small changes in the expression levels of p53 (34–39), we analyzed 1659 breast cancer samples for this SNP but none contained this variation. The potential of this polymorphism to be a risk factor for cancer or responsiveness to drugs remains to be seen.

To determine the importance of DTIE for strict TSS selection we first validated the TSS location in the endogenous p53 gene using 5′ RACE (Figure 7B). The results confirm the presence of one major TSS at exactly the expected location. We next examined the impact of DTIE deletion on the position of the TSS using 5′ RACE following transfection of p53 promoter driven reporter gene. The results revealed that in the absence of DTIE the major TSS is much weaker (Figure 7C, red arrowhead) and additional start sites become apparent (black arrowhead), reminiscent of the effect of DTIE deletion seen in the miR-22 gene (Figure 1A). We also examined whether DTIE strict location is important for its function in the p53 promoter by introducing a 5 nt linker either at the +8 or −8 position, relative to the TSS. In the case of the downstream linker we found that the major TSS was eliminated (Figure 7D) while a dramatic change in the number and position of the TSSs was seen in the case of the upstream linker (Figure 7E). These findings indicate that the specific location of p53 DTIE relative to an upstream element is required for strict start site selection as it does in miR-22.

DISCUSSION

Selection of transcription initiation site can be critical for the subsequent mRNA translation process. Nevertheless very little is known about this stage, in particular in TATA-less and Inr-less promoters, which constitute a large fraction of Pol II genes. In the present study we used the miR-22 and the p53 promoters as models to investigate the mechanism of TSS selection in the TATA- and Inr-less promoter class. Our study revealed a novel and highly prevalent downstream regulatory element that we named DTIE. The features of DTIE fit most of the core promoter criteria: (i) it is located near the TSS in a region expected to encompass core elements; (ii) DTIE position is very strict, a property that is specific to core elements and shared by the well-known TATA, Inr and DPE elements. This feature differs from 'regular' proximal element (such as Sp1) that is dispersed within a certain range. (iii) DTIE is critical for promoter strength and TSS selection, which is also in common with other core element. (iv) the prevalence of DTIE is very high, reminiscent of a basic element. On the other hand DTIE is unique in the sense that it cooperates with an upstream, strictly localized element, to facilitate initiation from a specific site.

TFIID, the major core promoter-binding factor is recruited to the promoter via DTIE and most likely through its binding factor. This is evident from the observation that TFIID association with the promoter and the integrity of the DTIE sequence are coupled. However, based on the DNA binding assays, TFIID does not seem to interact directly with this element. We compared DTIE consensus to other downstream core elements known to bind TFIID such as DPE, MTE and DCE and found that DTIE is distinct in sequence and location from DPE and DCE (40,41) but resembles the 'bridge' organization of MTE (42) as both have downstream location and consist of two sub-regions. However the mechanisms by which the *Drosophila* DPE and MTE act are quite different from DTIE. Unlike DTIE, DPE and MTE bind TFIID directly (43,44). In addition, DPE and MTE activities are strictly dependent on the Initiator, which is dispensable for DTIE.

Our findings suggest that the arrangement of DTIE relative to another element within enhancer I (E1) of the miR-

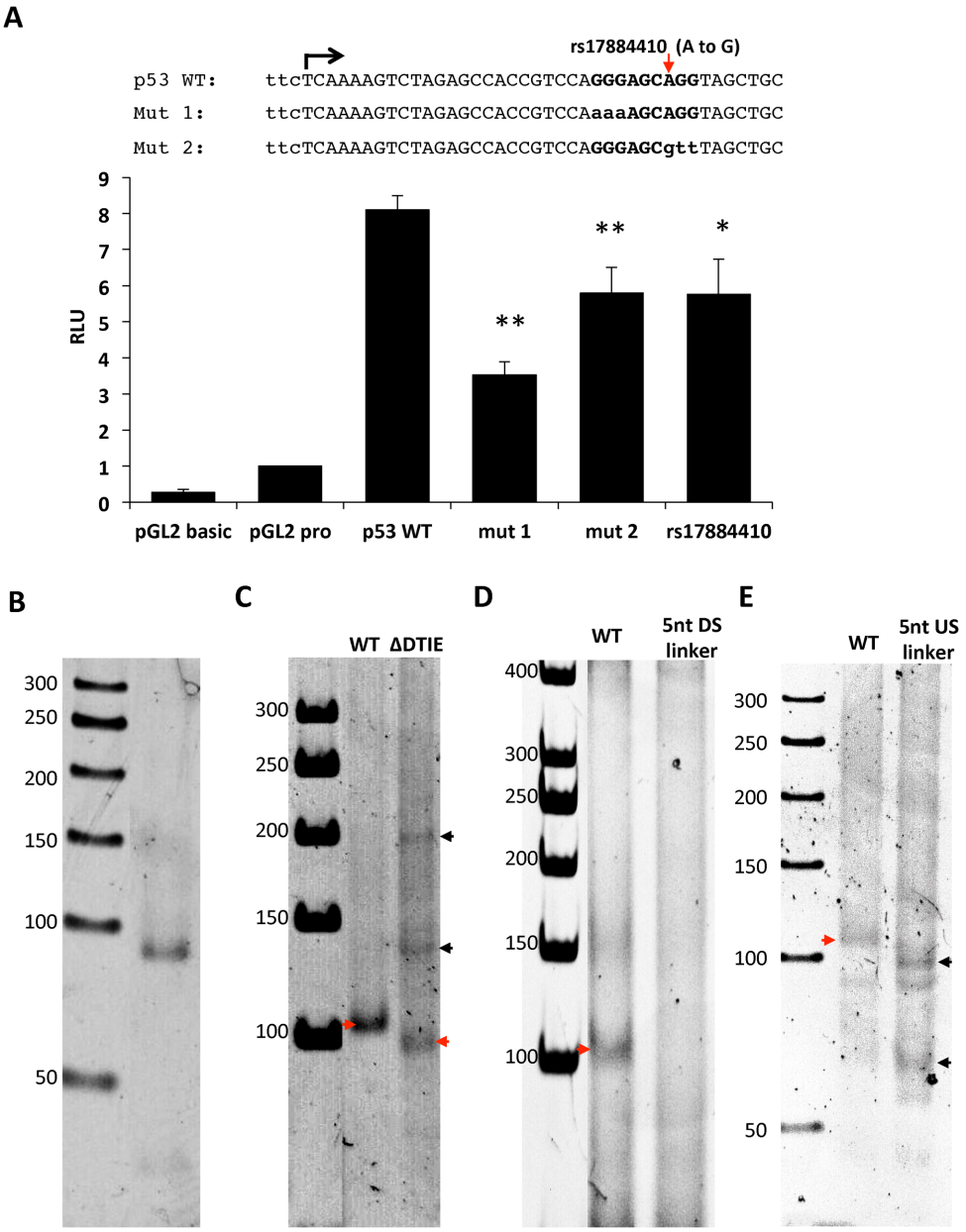


Figure 7. DTIE is functional in p53 promoter. (**A**) The sequence of the p53 TSS downstream region is shown on the top. DTIE is in bold letters and mutated nucleotides are indicated by lowercase letters. The position of rs17884410 SNP is indicated by arrow. Wild-type, DTIE promoter mutants and a point mutation mimicking the SNP were transfected into HEK293T cells and relative luciferase activity was determined 24 h later. The promoter-less and promoter-containing pGL2 basic and pGL2-pro, respectively were used as controls and pGL2 promoter activity was set to 1. The results represent the average ± SE of 8 independent experiments. * and ** indicate a statistically significant difference $P < 0.05$ and $P < 0.01$ respectively. (**B**) Analysis of the endogenous p53 TSS using 5′ RACE. (**C**) 5′ RACE analysis of the p53 promoter in which the 9 nucleotides of DTIE were deleted. (**D**) 5′ RACE analysis of the p53 promoter in which a 5 nt linker was inserted at the +8 position. (**E**) 5′ RACE analysis of the p53 promoter in which a 5 nt linker was inserted at the -8 position. The marker lane was spliced to bring it closer to the other lanes in B–D. The gels of this figure are representatives of 2–3 independent experiments.

22 promoter, is critical for its function. The dependency of start site selection on a regulatory element that is beyond the region considered core promoter, is unique. This element, that we termed DUCE, is located between −111 and –137 relative to the TSS and is important for strict TSS selection as well as promoter strength. Furthermore, the region containing DUCE does not resemble the binding site of any known transcription factor.

DTIE prevalence among human promoters, at +21 to +25 downstream location, is remarkably high, exceeding TATA-box frequency, indicating that this element and its binding factor are of paramount importance. This is underscored by the presence of DTIE in many central cancer-related genes such as p53, RB1, PTEN and miR-22 itself. In one particular case of p53, a rare polymorphism within p53 DTIE was found to cause a modest decrease in p53 pro-

moter strength. As maintenance of genome stability by p53 is sensitive to small changes in p53 protein levels (34–39), it is possible that this genetic alteration may be associated with increased cancer susceptibility.

In summary the characterization of DTIE shows that in spite of significant progress in computational genomics tools, the identification of a highly abundant proximal promoter element such as DTIE was possible only through conventional molecular approaches. While this highlights the relevance of molecular tools for discovery of novel and biologically significant regulatory elements, this study should also promote the development of improved computational means for prediction of regulatory elements. Future studies should address the mechanism of transcription initiation driven by these elements and the potential regulatory role they play.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We would like to thank Drs Michael Eberle and Luana McAuliffe from Ilumina Inc. and Young J. Kim from University of Science & Technology, Korea for their information regarding rs17884410. We also thank Drs Moshe Oren (Weizmann Institute), Igor Ulitzky (Weizmann Institute) and Tamar Juven-Gershon (Bar Ilan University) for their help and Dr Sandra Moshonov for editing.

FUNDING

Israel Science Foundation [1168/13]; Weizmann-Wolfson MicroRNA Research Center [to R.D.]; NIH [CA103867]; CPRIT [RP110471, RP140367]; Welch Foundation [I-1805 to C.-M.C.]. R.D. is the incumbent of the Ruth and Leonard Simon Chair of Cancer Research. Funding for open access charge: ISF.
Conflict of interest statement. None declared.

REFERENCES

1. Dikstein,R. (2012) Transcription and translation in a package deal: the TISU paradigm. *Gene*, **491**, 1–4.
2. Dikstein,R. (2011) The unexpected traits associated with core promoter elements. *Transcription*, **2**, 201–206.
3. Juven-Gershon,T., Hsu,J.Y. and Kadonaga,J.T. (2006) Perspectives on the RNA polymerase II core promoter. *Biochem. Soc. Trans.*, **34**, 1047–1050.
4. Smale,S.T. and Kadonaga,J.T. (2003) The RNA polymerase II core promoter. *Annu. Rev. Biochem.*, **72**, 449–479.
5. Thomas,M.C. and Chiang,C.M. (2006) The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.*, **41**, 105–178.
6. Marbach-Bar,N., Ben-Noon,A., Ashkenazi,S., Harush,A.T., Avnit-Sagi,T., Walker,M.D. and Dikstein,R. (2013) Disparity between microRNA levels and promoter strength is associated with initiation rate and Pol II pausing. *Nat. Commun.*, **4**, 1–12.
7. Morachis,J.M., Murawsky,C.M. and Emerson,B.M. (2010) Regulation of the p53 transcriptional response by structurally diverse core promoters. *Genes Dev.*, **24**, 135–147.
8. Yean,D. and Gralla,J. (1997) Transcription reinitiation rate: a special role for the TATA box. *Mol. Cell. Biol.*, **17**, 3809–3816.
9. Consortium,F., Pmi,R., Clst,T., Forrest,A.R., Kawaji,H., Rehli,M., Baillie,J.K., de Hoon,M.J., Haberle,V., Lassman,T. *et al.* (2014) A promoter-level mammalian expression atlas. *Nature*, **507**, 462–470.
10. Gershenzon,N.I. and Ioshikhes,I.P. (2005) Synergy of human Pol II core promoter elements revealed by statistical sequence analysis. *Bioinformatics*, **21**, 1295–1300.
11. Kim,T.H., Barrera,L.O., Zheng,M., Qu,C., Singer,M.A., Richmond,T.A., Wu,Y., Green,R.D. and Ren,B. (2005) A high-resolution map of active promoters in the human genome. *Nature*, **436**, 876–880.
12. Moshonov,S., Elfakess,R., Golan-Mashiach,M., Sinvani,H. and Dikstein,R. (2008) Links between core promoter and basic gene features influence gene expression. *BMC Genomics*, **9**, 1–10.
13. Yang,C., Bolotin,E., Jiang,T., Sladek,F.M. and Martinez,E. (2007) Prevalence of the initiator over the TATA box in human and yeast genes and identification of DNA motifs enriched in human TATA-less core promoters. *Gene*, **389**, 52–65.
14. Anish,R., Hossain,M.B., Jacobson,R.H. and Takada,S. (2009) Characterization of transcription from TATA-less promoters: identification of a new core promoter element XCPE2 and analysis of factor requirements. *PloS One*, **4**, e5103.
15. Tokusumi,Y., Ma,Y., Song,X., Jacobson,R.H. and Takada,S. (2007) The new core promoter element XCPE1 (X Core Promoter Element 1) directs activator-, mediator-, and TATA-binding protein-dependent but TFIID-independent RNA polymerase II transcription from TATA-less promoters. *Mol. Cell. Biol.*, **27**, 1844–1858.
16. Yarden,G., Elfakess,R., Gazit,K. and Dikstein,R. (2009) Characterization of sINR, a strict version of the Initiator core promoter element. *Nucleic Acids Res.*, **37**, 4234–4246.
17. Elfakess,R. and Dikstein,R. (2008) A translation initiation element specific to mRNAs with very short 5′ UTR that also regulates transcription. *PloS One*, **3**, e3094.
18. Bar,N. and Dikstein,R. (2010) miR-22 forms a regulatory loop in PTEN/AKT pathway and modulates signaling kinetics. *PloS One*, **5**, e10859.
19. Erijman,A., Dantes,A., Bernheim,R., Shifman,J.M. and Peleg,Y. (2011) Transfer-PCR (TPCR): a highway for DNA cloning and protein engineering. *J. Struct. Biol.*, **175**, 171–177.
20. Unger,T., Jacobovitch,Y., Dantes,A., Bernheim,R. and Peleg,Y. (2010) Applications of the Restriction Free (RF) cloning procedure for molecular manipulations and protein expression. *J. Struct. Biol.*, **172**, 34–44.
21. Yunger,S., Rosenfeld,L., Garini,Y. and Shav-Tal,Y. (2013) Quantifying the transcriptional output of single alleles in single living mammalian cells. *Nat. Protoc.*, **8**, 393–408.
22. Golan-Mashiach,M., Grunspan,M., Emmanuel,R., Gibbs-Bar,L., Dikstein,R. and Shapiro,E. (2012) Identification of CTCF as a master regulator of the clustered protocadherin genes. *Nucleic Acids Res.*, **40**, 3378–3391.
23. Chiang,C.M., Ge,H., Wang,Z., Hoffmann,A. and Roeder,R.G. (1993) Unique TATA-binding protein-containing complexes and cofactors involved in transcription by RNA polymerases II and III. *EMBO J.*, **12**, 2749–2762.
24. Amir-Zilberstein,L., Ainbinder,E., Toube,L., Yamaguchi,Y., Handa,H. and Dikstein,R. (2007) Differential regulation of NF-kappaB by elongation factors is determined by core promoter type. *Mol. Cell. Biol.*, **27**, 5246–5259.
25. Lee,D.H., Gershenzon,N., Gupta,M., Ioshikhes,I.P., Reinberg,D. and Lewis,B.A. (2005) Functional characterization of core promoter elements: the downstream core element is recognized by TAF1. *Mol. Cell. Biol.*, **25**, 9674–9686.
26. Carey,M.F., Peterson,C.L. and Smale,S.T. (2010) Magnesium-agarose electrophoretic mobility shift assay (EMSA) of transcription factor IID binding to DNA. *Cold Spring Harb. Protoc.*, 1223–1227.
27. Zerby,D. and Lieberman,P.M. (1997) Functional analysis of TFIID-activator interaction by magnesium-agarose gel electrophoresis. *Methods*, **12**, 217–223.
28. Cler,E., Papai,G., Schultz,P. and Davidson,I. (2009) Recent advances in understanding the structure and function of general transcription factor TFIID. *Cellular Mol. Life Sci.*, **66**, 2123–2134.
29. Matangkasombut,O., Auty,R. and Buratowski,S. (2004) Structure and function of the TFIID complex. *Adv. Protein Chem.*, **67**, 67–92.
30. Tuck,S.P. and Crawford,L. (1989) Characterization of the human p53 gene promoter. *Mol. Cell. Biol.*, **9**, 2163–2172.

31. Ronen,D., Rotter,V. and Reisman,D. (1991) Expression from the murine p53 promoter is mediated by factor binding to a downstream helix-loop-helix recognition motif. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 4128–4132.
32. Reisman,D., Elkind,N.B., Roy,B., Beamon,J. and Rotter,V. (1993) c-Myc trans-activates the p53 promoter through a required downstream CACGTG motif. *Cell Growth Differ.*, **4**, 57–65.
33. Deffie,A., Wu,H., Reinke,V. and Lozano,G. (1993) The tumor suppressor p53 regulates its own transcription. *Mol. Cell. Biol.*, **13**, 3415–3423.
34. Chen,X., Ko,L.J., Jayaraman,L. and Prives,C. (1996) p53 levels, functional domains, and DNA damage determine the extent of the apoptotic response of tumor cells. *Genes Dev.*, **10**, 2438–2451.
35. Murray-Zmijewski,F., Slee,E.A. and Lu,X. (2008) A complex barcode underlies the heterogeneous response of p53 to stress. *Nat. Rev. Mol. Cell Biol.*, **9**, 702–712.
36. Oren,M. (2003) Decision making by p53: life, death and cancer. *Cell Death Differ.*, **10**, 431–442.
37. Vousden,K.H. and Lu,X. (2002) Live or let die: the cell's response to p53. *Nat. Rev. Cancer*, **2**, 594–604.
38. Kracikova,M., Akiri,G., George,A., Sachidanandam,R. and Aaronson,S.A. (2013) A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. *Cell Death Differ.*, **20**, 576–588.
39. Levine,A.J. and Oren,M. (2009) The first 30 years of p53: growing ever more complex. *Nat. Rev. Cancer*, **9**, 749–758.
40. Burke,T.W., Willy,P.J., Kutach,A.K., Butler,J.E. and Kadonaga,J.T. (1998) The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold Spring Harb. Symp. Quant. Biol.*, **63**, 75–82.
41. Lewis,B.A., Kim,T.K. and Orkin,S.H. (2000) A downstream element in the human beta-globin promoter: evidence of extended sequence-specific transcription factor IID contacts. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 7172–7177.
42. Theisen,J.W., Lim,C.Y. and Kadonaga,J.T. (2010) Three key subregions contribute to the function of the downstream RNA polymerase II core promoter. *Mol. Cell. Biol.*, **30**, 3471–3479.
43. Burke,T.W. and Kadonaga,J.T. (1997) The downstream core promoter element, DPE, is conserved from Drosophila to humans and is recognized by TAFII60 of Drosophila. *Genes Dev.*, **11**, 3020–3031.
44. Lim,C.Y., Santoso,B., Boulay,T., Dong,E., Ohler,U. and Kadonaga,J.T. (2004) The MTE, a new core promoter element for transcription by RNA polymerase II. *Genes Dev.*, **18**, 1606–1617.