Published online April 26, 2004

*Nucleic Acids Research, 2004, Vol. 32, No. 8* ***2315–2322***
*DOI: 10.1093/nar/gkh571*

# Regulation of the MiTF/TFE bHLH-LZ transcription factors through restricted spatial expression and alternative splicing of functional domains

**Roland P. Kuiper*, Marga Schepens, José Thijssen, Eric F. P. M. Schoenmakers and Ad Geurts van Kessel**

Department of Human Genetics, University Medical Center Nijmegen, Nijmegen, The Netherlands

Received December 4, 2003; Revised January 19, 2004; Accepted April 6, 2004

## ABSTRACT

**The MiTF/TFE (MiT) family of basic helix–loop–helix leucine zipper transcription factors is composed of four closely related members, MiTF, TFE3, TFEB and TFEC, which can bind target DNA both as homo- or heterodimers. Using real-time RT–PCR, we have analyzed the relative expression levels of the four members in a broad range of human tissues, and found that their ratio of expression is tissue-dependent. We found that, similar to the *MiTF* gene, the genes for *TFEB* and *TFEC* contain multiple alternative first exons with restricted and differential tissue distributions. Seven alternative 5′ exons were identified in the *TFEB* gene, of which three displayed specific expression in placenta and brain, respectively. A novel *TFEC* transcript (*TFEC-C*) encodes an N-terminally truncated TFEC isoform lacking the acidic activation domain (AAD), and is exclusively expressed in kidney and small intestine. Furthermore, we observed that a considerable proportion of the *TFEC* transcripts splice out protein-coding exons, resulting in transcription factor isoforms lacking one or more functional domains, primarily the basic region and/or the AAD. These isoforms were always co-expressed with the intact transcription factors and may act as negative regulators of MiTF/TFE proteins. Our data reveal that multiple levels of regulation exist for the MiTF/TFE family of transcription factors, which indicates how these transcription factors may participate in various cellular processes in different tissues.**

## INTRODUCTION

The basic helix–loop–helix leucine zipper (bHLH-LZ) transcription factors MiTF, TFE3, TFEB and TFEC comprise a family (MiTF/TFE or MiT family) of closely related proteins that bind DNA as homo- and heterodimers (1). All possible combinations of DNA-binding MiTF/TFE dimers have been shown to exist *in vitro* (1–3).

The members of this family are believed to be implicated in pivotal developmental and cellular processes in various cell types. MiTF is involved in the maturation of melanocytes of neural crest origin, retinal pigment epithelium, and bone marrow-derived mast cells and osteoclasts (4); whereas TFEB is essential for placental vascularization (5). TFE3 plays a role in TGF-β-activated signal transduction (6,7) and B-cell activation (8), and cooperates with MiTF and TFEC in osteoclast development (9–11). Furthermore, chromosomal translocations involving the *TFE3* and *TFEB* genes have been implicated in subtypes of renal cell carcinomas in children and young adults. In addition, TFE3 rearrangements were found in alveolar soft part sarcomas (12–19).

The apparent distinctive cellular functions of MiTF/TFE family members in the various cell types require regulatory mechanisms that tightly control MiTF/TFE functioning. Such mechanisms may involve interactions with cell type-specific factors or specific protein modifications, but the relative levels of co-expression of the various transcription factors may be important as well. Indeed, the expression levels of each of the four MiTF/TFE family members appear to differ considerably between cell types (3,20,21), which points towards a model in which the ratio of MiTF/TFE expression dictates specific dimerization patterns and, subsequently, the expression of target genes. Although there are indications that heterodimeric interactions are not essential for proper MiTF/TFE functioning (11), MiTF–TFE3 and MiTF–TFEC heterodimers do exist *in vivo*, and are believed to be involved in the regulation of cell type-specific genes (10,21,22).

The *MiTF* gene is expressed in different isoforms that are under the control of distinct promoters (23,24). Currently, at least eight major MiTF isoforms have been identified that are differentially expressed in a variety of tissues, including melanocytes, heart and mast cells (4,21,24–28). These isoforms share the important functional domains of MiTF, including the transactivation domain, basic region, helix–loop–helix and leucine zipper, but differ in their N-termini. These N-termini may contribute to the cell type-specific properties of the various isoforms, for example by recruiting particular cofactors to the transcription apparatus. Thus far, little is known about the tissue distribution of the other three MiTF/TFE family members TFE3, TFEB and TFEC.

*To whom correspondence should be addressed. Tel: +31 24 3614017; Fax: +31 24 3540488; Email: r.kuiper@antrg.umcn.nl

 

In this study, we have determined the relative expression levels of MiTF, TFE3, TFEB and TFEC in a broad range of human tissues, and identified transcript variants. The ratio of expression of the four MiTF/TFE family members was found to be unique for each tissue. Furthermore, similar to MiTF, we found various splice forms of TFEB and TFEC with alternative first exons and very restricted expression patterns, suggesting regulation through alternative promoter usage.

## MATERIALS AND METHODS

### cDNA synthesis and RT–PCR

A panel of total RNA pools from 20 human tissues was obtained from BD Biosciences Clontech (Palo Alto, CA). Reverse transcription was performed using 2 µg of total RNA and Superscript II (Life Technologies, Grand Island, NY) according to the manufacturer's instructions. RT reactions were performed using random hexamers, except for the experiment dealing with alternative intron 7 splicing of TFEC transcripts, for which oligo-dT priming was used in order to minimize contamination with incompletely spliced nuclear mRNAs. Primers used for PCR were TFEB-1Af (5′-CAACAAGGGAAGGTGACATGAAAGG-3′), TFEB-1Bf (5′-CGGACAGATTGACCTTCAGAGG-3′), TFEB-1C1for (5′-ATGCCTGAGAGGGGGTCGGAG-3′), TFEB-1C2for (5′-GAGAGGAGGGAGATGAAAACACACC-3′), TFEB-1Df (5′-TTAGAGCTCGCTGGGGAAGTGCAG-3′), TFEB-1Ef (5′-TTGAGGTCCCTGCCTGCTAACCTGACG-3′), TFEB-1Ff (5′-CCTCCTTCCAGGGGGCTGCC-3′), TFEB-1Gf (5′-GGTGTGAGCTGACGAGAGCTTCAG-3′), TFEBexon2r (5′-AACCCTATGCGTGACGCCATGGTGG-3′), TFEC-1Af (5′-ACTTGTCCCAGCAAGACCACA-3′), TFEC-1Bf (5′-AAATCCACTCATTGCTGGTCC-3′), TFEC-1Cf (5′-GAGAAAACCATTGCTATTGTGAAGG-3′), TFECr (5′-TTGTAGCCACTTGATGTACTCCACT-3′), TFEC-6f (5′-CACTCTTATTCCAAAGTCTAATGATCCT-3′), TFECintron7r (5′-CCTTGAGGCAATGAAATCCATAGC-3′) and TFEC-8r (5′-TCTTTGTTCCTCTTTCAATGCAGC-3′).

### Real-time quantitative RT–PCR

Real-time RT–PCR on *MiTF/TFE* cDNAs was performed on a TaqMan ABI 7700 Sequence Detection System (PE Applied Biosystems) using heat-activated TaqDNA polymerase (Amplitaq Gold; PE Applied Biosystems) as previously described (18). The specificity of the primer/probe sets for each of the four *MiTF/TFE* members relative to the other three was confirmed using different dilutions of DNA constructs containing the bHLH-LZ-encoding regions of *TFE3*, *TFEB*, *TFEC* or *MiTF*. The four primer/probe sets recognized their specific *MiTF/TFE* templates with almost identical efficiencies. For quantitative analysis of the data, *TFEB* $C_T$-values were normalized to those of endogenous *GAPDH* (with use of standard TaqMan human *GAPDH* control reagents; PE Applied Biosystems) using the $\Delta\Delta C_T$ technique (29).

### Rapid amplification of 5′-cDNA ends (5′-RACE)

The 5′-RACE System (Invitrogen) was used according to the manufacturer's instructions. A *TFEB* gene-specific primer (5′-ATTTCAGGATTGATGTAGC-3′) was used for *TFEB* cDNA synthesis. The *TFEB*-specific primers (located in exon 2) were TFEB 5′R442 (5′-TATTGATGGCCGGGGTGGGCG-3′) and TFEB 5′R376 (5′-TAATGCATGACAGCCTGTTGCTGCATG-3′), of which the latter was used for the hemi-nested reaction. For the *TFEC* 5′-RACE we used the *TFEC* gene-specific primer (5′-TTTCAATGAGGTTGTGG-3′) for cDNA synthesis and primers TFEC 5′R547 (5′-ATTGGTAGACTACTTGGACAAGAAGCAC-3′) and TFEC 5′R494 (5′-AATTCCTTGTTCACCGCTATACACATCC-3′) for the PCR. The PCR profile was composed of five cycles of 30 s at 94°C, 1 min at 48°C, 30 s at 72°C; five cycles of 30 s at 94°C, 1 min at 52°C, 30 s at 72°C; 25 cycles of 30 s at 94°C, 1 min at 55°C, 30 s at 72°C. PCR products were separated on agarose gels, purified, subcloned into the pGemT vector (Promega, Leiden, The Netherlands), and subsequently sequenced using a Ready Reaction Dye Terminator Cycle sequencing kit (PE Applied Biosystems, Foster City, CA) and an ABI 3700 automated sequencer (PE Applied Biosystems).

### Transfection experiments and western blotting

Full-length RT–PCR products corresponding to the wild-type *TFEB* and *TFEC* transcripts were cloned into the eukaryotic expression vector pSG8-VSV (30). The *TFEC* cDNAs were cloned in-frame in front of a VSV-tag. Transfection and western blotting was performed as previously described (18). For immunodetection of TFEB, an affinity-purified rabbit polyclonal anti-TFEB-N antiserum was used (18). The VSV-tagged TFEC was detected using the mouse anti-VSV monoclonal antibody P5D4. Immunostaining was performed using chemiluminesence.

## RESULTS

### The MiTF/TFE transcription factor family shows a specific tissue distribution

*In vitro*, TFE3, TFEB, TFEC and MiTF form homo- and heterodimers in all possible combinations to bind DNA (1). Based on this observation, it can be hypothesized that differential tissue-specific dimerization may determine the repertoire of target genes that are expressed. In order to analyze the relative expression levels of the MiTF/TFE members in different tissues, we performed quantitative real-time RT–PCR analysis on a panel of 20 human tissues. We established primer sets located within the bHLH-encoding regions, capable of detecting all MiTF/TFE isoforms. As shown in Figure 1A, the expression ratios of the four *MiTF/TFE* members showed large variations between tissues and none of the members appeared to be co-regulated. In Figure 1B, the same set of data is represented as absolute levels, which allows comparison of the *MiTF/TFE* expression levels between tissues. Although each of the four members could be detected in all tissues examined, *TFEC* and *MiTF* showed more variation in expression between tissues (Fig. 1B). Highest expression levels of both *TFE3* and *TFEB* were found in placenta and lung, but their expression levels differed in adrenal gland (relatively high for *TFE3*), and prostate and spleen (higher levels of *TFEB*). *MiTF* was most prominent in the uterus. The *TFEC* gene was expressed at lower levels compared to the other three members, and was prominent in spleen, kidney, bone marrow and small intestine. These data

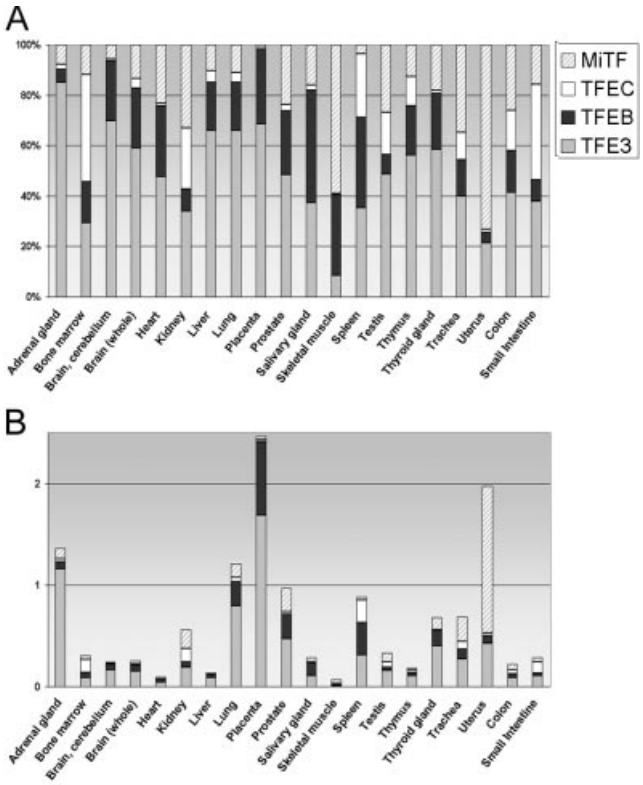


**Figure 1.** Bar graphs showing the relative expression levels of the MiTF/TFE transcription factors in twenty human tissues, as analyzed by real-time RT–PCR. (**A**) The sum of the expression levels of TFE3, TFEB, TFEC and MiTF is set to 100% to enable comparison of the MiTF/TFE expression ratios in each tissue. (**B**) The same data set depicted as absolute levels to illustrate the variation in MiTF/TFE expression between tissues. The data were based on two independent in duplo experiments.

indicate that expression of each of the four members must be regulated in a tissue-specific manner.

### The *TFEB* gene contains alternative 5′-non-coding exons

The temporal and spatial expression of the human *MiTF* gene is controlled by the differential usage of at least six distinct promoters, driving the expression of transcripts with alternative first exons (24–26,28). We investigated whether a similar mechanism might act in the other MiTF/TFE members. The *TFEB* gene, which resides on chromosome 6p21, is composed of nine exons, with a postulated initiation ATG preceded by a perfect ribosomal binding ('Kozak') sequence in exon 2 (18). The numerous TFEB-encoding ESTs that are present in the nucleotide sequence databases reveal the presence of at least three frequently occurring alternative 5′ exons (GenBank Accession nos M33782, NM007162 and AK095061, respectively). To validate whether more *TFEB* transcripts might exist that differ in their 5′-end, we performed 5′-RACE analysis on cDNA pools of various tissues, using TFEB-specific primers located in exon 2. Using this approach, we were able to identify *TFEB* transcripts containing seven distinct alternative 5′ exons, of which five were novel, named exons 1a to 1g1 (Table 1 and Fig. 2). The resulting TFEB transcripts were designated *TFEB-A* to *TFEB-G*, respectively. The *TFEB-B* (31) and *TFEB-C* transcripts were found in liver, kidney, spleen and brain, whereas the remaining five *TFEB* transcripts were found in only one of the tissues examined (Table 1). In addition, we observed three alternative splicing events in the 5′ region of the *TFEB-G* transcripts, generating transcripts containing either 1g1 alone, 1g1/1g2 or the complete exon 1g at their 5′-ends (Fig. 2B and Table 1). Similarly, we found alternative *TFEB-E* splice variants in the brain cDNA pool which contained exon 1c2 and, in some cases, the entire sequence from 1e to 1c2 (Fig. 2B). This latter longest *TFEB-E* variant corresponds to the previously identified full-length *TFEB* transcript AK095061. Another 5′-RACE product, obtained from skeletal muscle cDNA, appeared to be an alternatively spliced variant of *TFEC-C*, lacking exon 1c2 (Fig. 2B). Of all alternative 5′ exons, only exon 1a contains a putative in-frame ATG codon which, if active, would result in 14 extra N-terminal amino acids in TFEB-A. Therefore, most alternative 5′ exons are non-coding. Together, we conclude from these data that the *TFEB* gene contains multiple alternative first exons.

**Table 1.** Alternative 5′ exons of the *TFEB* gene and their occurrence in different tissues

| 5′ exon | Sequence at junction | | Number of 5′-RACE products/total | | | |
|---|---|---|---|---|---|---|
| | Exon | Intron | Liver $n = 5$ | Kidney $n = 11$ | Brain $n = 54$ | Spleen $n = 31$ |
| 1a | gctcaggctg | gtgagtgggg | | 8 | | |
| 1b | tcagagcgag | gtgagcggga | 4 | 2 | 15 | 14 |
| 1c2 | ctggtacatg | gtgagtacat | 1 | 1 | 24 | 11 |
| 1d | ggaagtgcag | gttaattact | | | 1 | |
| 1e | taacctgacg | gtatgctttc | | | 5 | |
| 1f | cctctcgaag | gtatgagagc | | | | 6 |
| 1g1 | agagcttcag | gtgggagcct | | | 3 | |
| 1g2 | ggtgaaagag | gtcagtggat | | | 6 | |

### Differential tissue distribution of alternative *TFEB* transcripts

To analyze the relative expression of the above *TFEB* variants in human tissues, we performed RT–PCR analysis on the human tissue panel using alternative exon 1-specific primers for the various transcripts. Six *TFEB* transcripts showed restricted and clearly different tissue distribution patterns (Fig. 2C). Only *TFEB-D* could not be amplified in any of the tissues, possibly due to very low expression levels. Despite the fact that the splice acceptor site of exon 1d perfectly matches the consensus criteria (Table 1), the *TFEB-D* transcript was found only once in our 5′-RACE analysis, and its existence thus still needs to be confirmed. *TFEB-B* and *TFEB-C* showed the broadest tissue distribution, with *TFEB-C* being detectable in all tissues but liver. Interestingly, *TFEB-A* was almost exclusively expressed in placenta, with only minor levels of expression in kidney, lung and prostate. Similarly, the *TFEB-E* and *TFEB-G* transcripts appeared to be brain-specific, whereas *TFEB-F* was highest in spleen (Fig. 2C).

The multiple RT–PCR products observed in the TFEB-E and TFEB-G panels were isolated and sequenced, and appeared to correspond to the splice variants identified with 5′-RACE analysis (Fig. 2B). Based on the intensities of these products in the two panels, these splice variants appear to be expressed in a tissue-specific manner. Despite their differen-

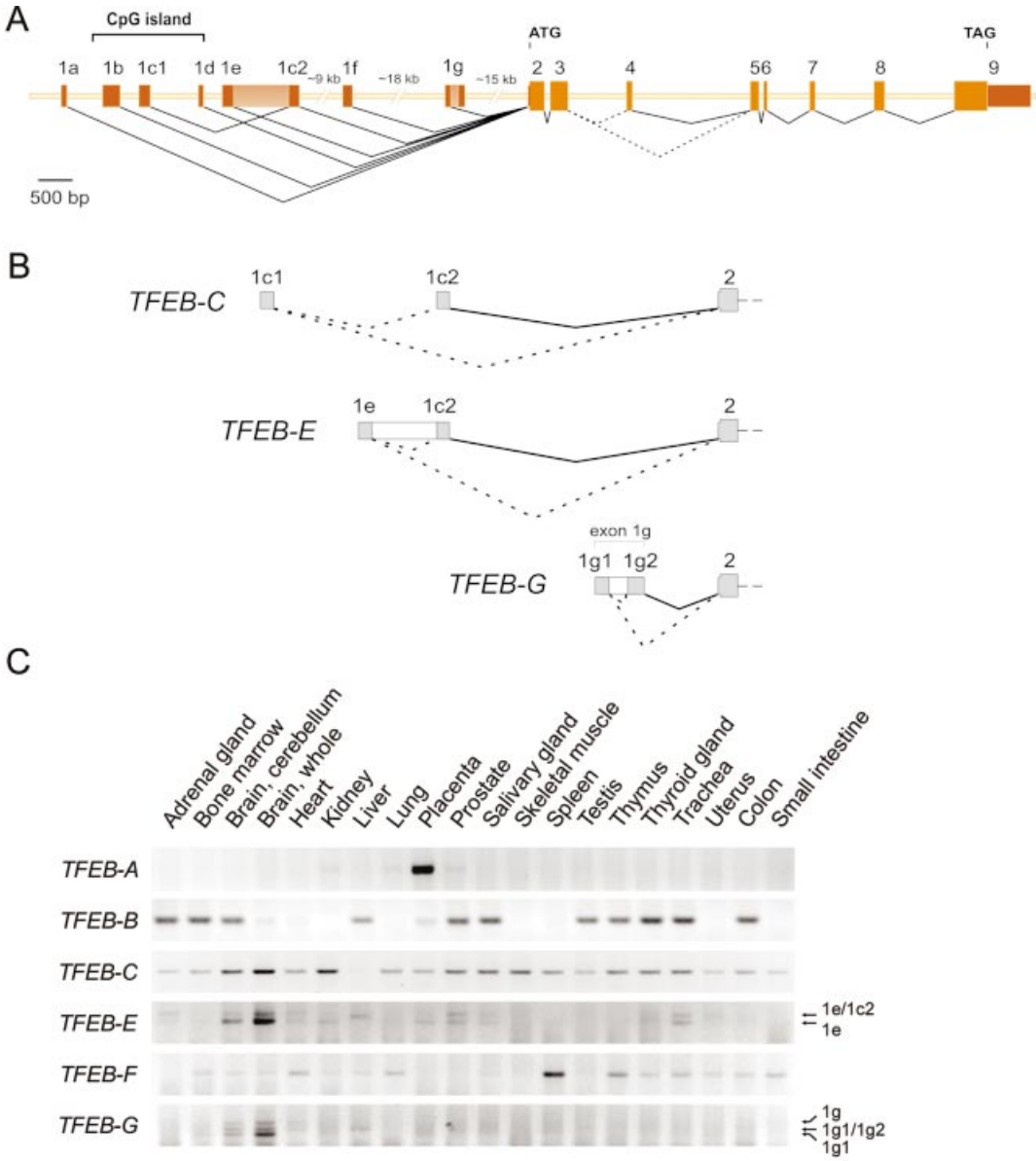


**Figure 2.** (**A**) The genomic organization of the human *TFEB* gene. The ATG start codon in exon 2 and TAG stop codon in exon 9 are indicated. Dotted lines depict alternative splicing. A 1600-bp CpG island colocalizes with exons 1b, 1c1 and part of exon 1d. (**B**) Alternative splice events in the 5′-regions of *TFEB-C*, *TFEB-E* and *TFEB-G*. Dotted lines depict alternative splicing. The largest of three *TFEB-G* derived products contained the complete exon 1g, including 1g1, 1g2 and the intermediate 'intronic' sequence. Accession numbers of the nine novel TFEB transcript fragments are AJ608786–AJ608794. (**C**) Tissue distribution of the human *TFEB* variants determined by RT–PCR analysis on 20 human tissues.

tial tissue distribution, none of these variants contain upstream ATG start codons, indicating that they all encode TFEB proteins with identical N-termini. Variations in the 5′-UTR sequences might, however, influence the efficiency of translation, as has been described for the cyclin-dependent kinase inhibitor p18INK4c and the nitric oxide synthase gene *NOS1* (32,33). We conclude that alternative transcripts may control the expression of TFEB in a tissue-specific manner.

### Differential tissue distribution of *TFEC* variants

In the *TFEC* gene (Fig. 3), at least two alternative transcription start sites appear to be present, resulting in transcripts here referred to as *TFEC-A* and *TFEC-B* (GenBank Accession nos BM807021 and NM012252, respectively). Exon 2 is shared by the two *TFEC* variants, and contains the ATG start codon. The *TFEC-A* variant has not been described before, and contains three exons preceding exon 2. Two additional in-frame ATG codons are present in exon 1a3 (Fig. 3A). Using 5′-RACE analysis on human kidney cDNA, we identified a third *TFEC* transcript, which initiates in an alternative 5′ exon positioned between exons 3 and 4 (exon 1c; Fig. 3A). RT–PCR analysis using a reverse primer in exon 1c in combination with a forward primer in any of the upstream exons never resulted in a PCR product, indicating that exon 1c is indeed the first exon of this transcript. *TFEC-C* thus lacks the protein-coding exons 2 and 3. In-frame ATG start codons were present in both exon

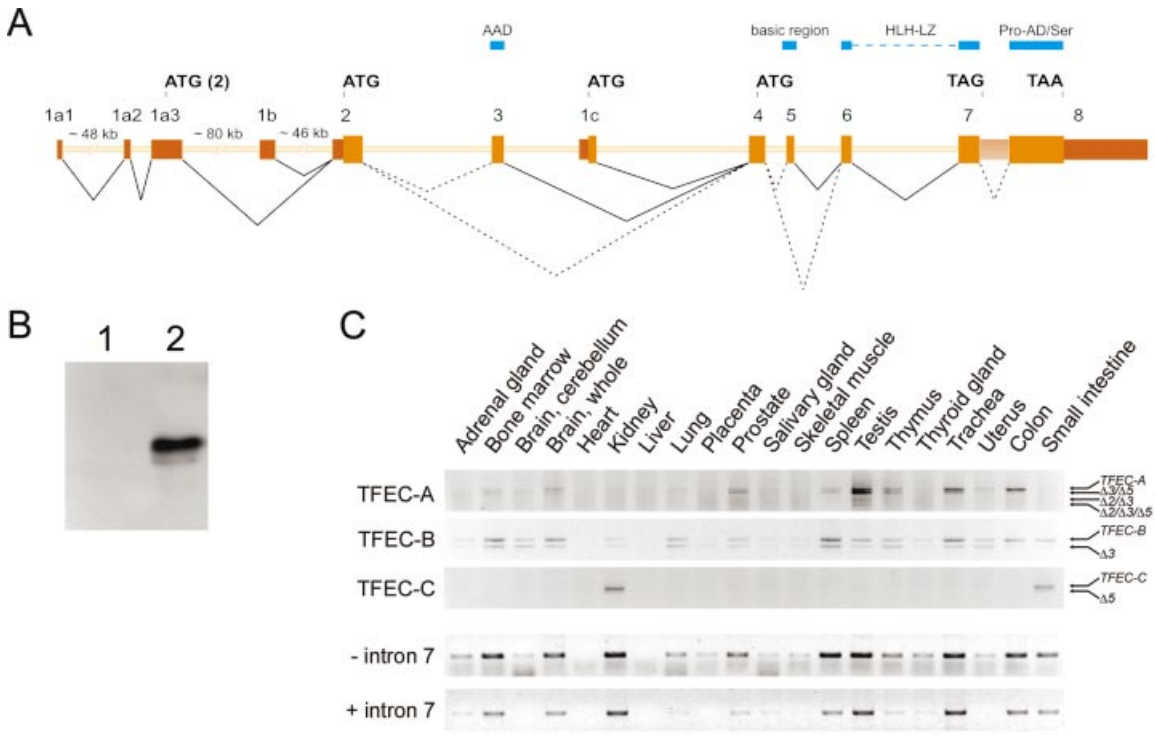


**Figure 3.** (**A**) Schematic representation of the human *TFEC* gene containing three alternative 5′ exons. For clarity reasons, exon sizes are depicted five times larger than the intron sizes. Relevant start and stop codons are indicated, as well as the regions encoding the AAD, the basic region, the bHLH-LZ and the proline-rich activation domain and serine-rich stretch (Pro-AD/Ser). Dotted lines depict alternative splicing. The accession number of the novel *TFEC-C* transcript fragment is AJ608795. (**B**) Western blot analysis of COS-1 cells transfected with empty vector (lane 1) or VSV-tagged *TFEC-C* cDNA (lane 2). (**C**) Tissue distribution of the *TFEC* transcripts determined by RT–PCR analysis. For each of the panels, the various PCR fragments were isolated and sequenced. Next to the full-length *TFEC* variants, alternatively spliced transcripts were amplified that lacked exon 2 (Δ2), exon 3 (Δ3) and/or exon 5 (Δ5). The lower two panels show RT–PCR products using a forward primer in exon 6, and reverse primers in either exon 8 or intron 7, which depict the presence of *TFEC* transcripts encoding the complete C-terminus of TFEC or a truncated C-terminus, respectively. In order to enrich for active cytoplasmic transcripts, we used oligo-dT priming in the reverse transcription reaction.

1c and exon 4. Subcloning of the *TFEC-C* cDNA in front of an in-frame VSV-tag, and subsequent transfection into COS-1 cells, revealed that the transcript indeed encodes a protein of the predicted size (~33 kDa; Fig. 3B). A *TFEC-C* transcript was also found in the mouse (Genbank Accession no. BB843583), and appeared to be highly conserved at both the nucleotide (83%) and amino acid (86%) level.

Similar to TFEB, the tissue distributions of *TFEC* transcripts was analyzed by RT–PCR (Fig. 3C). *TFEC-A* was predominantly expressed in testis, thymus, trachea, colon and prostate. *TFEC-B* showed a broader tissue distribution, and appeared to be absent only from heart and liver, whereas the newly identified *TFEC-C* transcript was exclusively expressed in kidney and small intestine.

Together, these findings suggest that next to the *MiTF* gene, the *TFEB* and *TFEC* genes also express transcripts with alternative 5′ exons that show restricted and clearly distinct tissue distributions.

Unlike *MiTF*, *TFEB* and *TFEC*, there are no *TFE3* transcripts known that would indicate the presence of alternative first exons. Furthermore, RT–PCR analysis on the human tissue panel, using a primer set that included the first exon and a primer set located within the bHLH domain, gave identical tissue distributions (data not shown). Therefore, we assume that expression of the *TFE3* gene is regulated by a single promoter.

## Alternative splicing of exons encoding functional domains

RT–PCRs with 5′ exon-specific primers generated multiple bands in all cases (Fig. 3C), which indicates that from each transcription start site various alternative transcripts can be generated. To reveal their identity, we have isolated and sequenced these different products. The largest RT–PCR product in the *TFEC-A* panel represents a transcript containing all protein-coding exons, whereas the lower band was found to lack the N-terminal acidic activation domain (AAD)-encoding exon 3 and the basic region-encoding exon 5 (Δ3/Δ5). In testis, the tissue that showed the highest expression of *TFEC-A*, two minor additional products could be isolated as well, and turned out to represent *TFEC* transcripts lacking exons 2 and 3 (Δ2/Δ3) and exons 2, 3 and 5 (Δ2/Δ3/Δ5), respectively. Interestingly, the relative expression ratios of these *TFEC-A* variants was different between the various tissues. Similarly, the doublet in the *TFEC-B* panel (Fig. 3C) appeared to represent transcripts that differ in the presence or absence of exon 3, and were previously presented as *TFEC-L* (long) and *TFEC-S* (short), respectively (34). Apparently, these two

forms are mostly co-expressed, although the ratio in each of the tissues examined was different. Finally, we also detected a weakly expressed, smaller variant of *TFEC-C*, which differed from the major form in that it again lacked the basic region-encoding exon 5 (Δ5).

Apart from exon 3 and exon 5, we found indications for another frequently occurring alternative splicing event in the *TFEC* gene. This alternative splicing event involves the C-terminus of TFEC, which became apparent when we used a reverse RT–PCR primer positioned in exon 8, the 3′-terminal exon of the gene. The C-terminal region of the protein is highly conserved within the MiTF/TFE family and contains two functional domains, namely a proline-rich activation domain (35), and a serine-rich region that contains an Rsk-1 phosphorylation site that is known to play a role in transactivation and degradation of MiTF (36). In most, if not all, *TFEC* expressing tissues, transcripts could be detected in which intron 7 is still present. Translation of such a transcript would result in a truncated TFEC protein that lacks the proline-rich activation domain and the serine-rich domain (Fig. 3C, two bottom panels). Although the RT–PCRs were performed under semi-quantitative conditions, it appears that the presence or absence of intron 7 in the *TFEC* transcripts again varies between tissues.

The expression of alternatively spliced variants of *TFEB* was less apparent than that of *TFEC*. The *TFEB* variants were mainly expressed as full-length transcripts that contained an alternative exon 1 followed by the protein-encoding exons 2 to 9. Nevertheless, we did find alternatively spliced *TFEB* transcripts in a limited set of tissues lacking exon 3 and exon 4, of which the latter encodes the AAD (data not shown).

In conclusion, we have established MiTF/TFE expression levels and ratios in 20 human tissues. These expression patterns appeared to be unique for each tissue. Furthermore, we identified (novel) alternative 5′ exons in the proximal regions of the *TFEB* and *TFEC* genes, and demonstrated that the expression of MiTF/TFE isoforms lacking functional domains is a characteristic for all members of the family, but is predominant for TFEC.

## DISCUSSION

We have determined the expression profile of the MiTF/TFE family of bHLH-LZ transcription factors in a large series of human tissues and identified alternative transcripts generated by the respective genes. We found that the ratio in which these four members are expressed differs extensively between the various tissues examined. The tight spatial regulation of MiTF/TFE expression appears to be mediated at least in part by alternative transcription start sites. Next to *MiTF* (4,21,24–28), both *TFEB* and *TFEC* were found to express a variety of transcripts that contain alternative 5′ (non-coding) exons and show very restricted differential tissue distributions. The use of alternative 5′ exons in a tissue-specific manner point towards the existence of alternative promoters in both genes. Indeed, through the characterization of an oligo-capped human cDNA library, the exact position of the transcription start site and thereby the position of a core-promoter region has been identified for the *TFEB-E* transcript (see http://dbtss.hgc.jp/). Furthermore, three more downstream located exons (1b, 1c1 and 1d) are positioned within a 1600-bp CpG island that is conserved between human and mouse. Such regions are known to colocalize with 60% of the human promoters (37). Although we cannot exclude that some of the *TFEB* variants we found are the result of alternative transcription start site usage within a single promoter region or alternative splicing in the 5′-UTR, our data strongly suggests the presence of alternative promoters in the *TFEB* and *TFEC* genes.

We identified seven alternative 5′ exons in the *TFEB* gene. Except for *TFEB-A*, all these 5′ exons appeared to be non-coding, suggesting that the respective transcripts encode the same TFEB protein. Together, the *TFEB* transcripts drive the ubiquitous expression of TFEB. However, each transcript was represented in a restricted set of tissues, and we found three novel *TFEB* transcripts that were predominantly expressed in one particular tissue. Of these three *TFEB* transcripts, *TFEB-E* and *TFEB-G* appeared to be specifically abundant in brain, whereas *TFEB-A* was expressed almost exclusively in placenta. Previous studies in transgenic mice already revealed that TFEB, but not the other MiTF/TFE members, plays an essential role in placental vascularization during the early stages of placental development (5,11). Our finding of a placenta-enriched *TFEB* variant points towards a similar role in the human placenta. This placental expression of TFEB appears to be controlled by a distinct promoter. A total of three alternative 5′ exons were found to be present in the *TFEC* gene, of which two (exons 1a and 1c) were novel. The three *TFEC* transcripts were all expressed in a limited set of tissues, with *TFEC-C* being exclusively found in kidney and small intestine. Despite this restricted expression of *TFEC* variants, *TFEC* was observed in all tissues examined, similar to *TFEB*. From these data, we conclude that next to the *MiTF* gene, the genes for *TFEB* and *TFEC* are also equipped with different alternative first exons and multiple promoter regions. Such a genomic organization allows a tightly regulated expression of the respective transcription factors in a cell type-specific manner, and may explain the observed tissue-specific ratios of the four MiTF/TFE members. These relative expression levels may affect MiTF/TFE homo- or heterodimerization and, thereby, influence the regulation of downstream target genes (18).

Another important mechanism for functional regulation may be the co-expression of MiTF/TFE members with isoforms lacking particular functional domains. Co-expression of such isoforms appears to occur for all MiTF/TFE proteins, but was most evident for TFEC. Multiple splice variants, lacking exons encoding a conserved mitogen-activated protein kinase (MAPK) phosphorylation site (exon 2), the N-terminal AAD (exon 3), the basic region (exon 5), and/or the C-terminal proline-rich activation domain and serine-rich stretch (exon 8), were co-expressed with *TFEC-A*, *TFEC-B* and *TFEC-C*. These domains are highly conserved in the MiTF/TFE family members and all play a role in the transactivating capacities of the proteins. Therefore, these isoforms may act as weaker transactivators or even repressors, as has been shown for TFE3 (35,38) and MiTF (23,36,39). Furthermore, splice variants lacking the basic region have been described for MiTF as well (23), and may act as negative regulators by forming non-DNA-binding homo- or heterodimers similar to what has been proposed for the HLH Id proteins (40,41). The N-terminally truncated TFEC-C isoform appears to be

expressed from an internal promoter, and also lacks the MAPK phosphorylation site and the AAD. Remarkably, this isoform was exclusively expressed in kidney and small intestine, which might indicate that it is involved in the regulation of processes that are shared by these two tissues, like for example (re)absorption of ions and/or organic compounds. The co-expression of these different TFEC isoforms may allow fine regulation of MiTF/TFE activities, and may be crucial for the proper spatio-temporal functioning of these transcription factors. Other well-characterized examples of genes that produce multiple isoforms with a regulatory role are *p63* and *p73*, two members of the *p53* gene family. Both genes produce splice variants as well as N-terminally truncated forms that originate from internal promoters (42,43). These isoforms act as dominant-negative regulators of p63, p73 and their close relative p53 by forming heteromeric complexes (43–45). The truncated isoforms were found to be vital for normal growth and development, and provide a regulatory system for modulating cell survival and cell death (42,46). Based on our findings, we hypothesize that the various MiTF/TFE isoforms lacking functional domains may form heterodimers with the full-length counterparts and, thereby, modulate their role in downstream target gene regulation, either by preventing DNA binding, changing the DNA-binding specificity, or altering the binding capacities of transcriptional cofactors. In particular, the *TFEC* gene, which generates the widest variety of functionally distinct isoforms, may play a prominent role in such regulatory mechanisms.

## ACKNOWLEDGEMENTS

The authors thank L. van der Logt, A. C. Poirters and E. M. Evers for technical assistance and advice. This work was supported by a grant from the Dutch Cancer Society (Koningin Wilhelmina Fonds), KUN 99–1912.

## REFERENCES

1. Hemesath,T.J., Steingrimsson,E., McGill,G., Hansen,M.J., Vaught,J., Hodgkinson,C.A., Arnheiter,H., Copeland,N.G., Jenkins,N.A. and Fisher,D.E. (1994) Microphthalmia, a critical factor in melanocyte development, defines a discrete transcription factor family. *Genes Dev.*, **8**, 2770–2780.
2. Fisher,D.E., Carr,C.S., Parent,L.A. and Sharp,P.A. (1991) TFEB has DNA-binding and oligomerization properties of a unique helix–loop–helix/leucine-zipper family. *Genes Dev.*, **5**, 2342–2352.
3. Zhao,G.Q., Zhao,Q., Zhou,X., Mattei,M.G. and de Crombrugghe,B. (1993) TFEC, a basic helix–loop–helix protein, forms heterodimers with TFE3 and inhibits TFE3-dependent transcription activation. *Mol. Cell. Biol.*, **13**, 4505–4512.
4. Hodgkinson,C.A., Moore,K.J., Nakayama,A., Steingrimsson,E., Copeland,N.G., Jenkins,N.A. and Arnheiter,H. (1993) Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix–loop–helix-zipper protein. *Cell*, **74**, 395–404.
5. Steingrimsson,E., Tessarollo,L., Reid,S.W., Jenkins,N.A. and Copeland,N.G. (1998) The bHLH-Zip transcription factor Tfeb is essential for placental vascularization. *Development*, **125**, 4607–4616.
6. Hua,X., Liu,X., Ansari,D.O. and Lodish,H.F. (1998) Synergistic cooperation of TFE3 and smad proteins in TGF-β-induced transcription of the plasminogen activator inhibitor-1 gene. *Genes Dev.*, **12**, 3084–3095.
7. Hua,X., Miller,Z.A., Benchabane,H., Wrana,J.L. and Lodish,H.F. (2000) Synergism between transcription factors TFE3 and Smad3 in transforming growth factor-beta-induced transcription of the Smad7 gene. *J. Biol. Chem.*, **275**, 33205–33208.
8. Merrell,K., Wells,S., Henderson,A., Gorman,J., Alt,F., Stall,A. and Calame,K. (1997) The absence of the transcription activator TFE3 impairs activation of B cells *in vivo*. *Mol. Cell. Biol.*, **17**, 3335–3344.
9. Weilbaecher,K.N., Motyckova,G., Huber,W.E., Takemoto,C.M., Hemesath,T.J., Xu,Y., Hershey,C.L., Dowland,N.R., Wells,A.G. and Fisher,D.E. (2001) Linkage of M-CSF signaling to Mitf, TFE3, and the osteoclast defect in Mitf(mi/mi) mice. *Mol. Cell*, **8**, 749–758.
10. Mansky,K.C., Sulzbacher,S., Purdom,G., Nelsen,L., Hume,D.A., Rehli,M. and Ostrowski,M.C. (2002) The microphthalmia transcription factor and the related helix–loop–helix zipper factors TFE-3 and TFE-C collaborate to activate the tartrate-resistant acid phosphatase promoter. *J. Leukoc. Biol.*, **71**, 304–310.
11. Steingrimsson,E., Tessarollo,L., Pathak,B., Hou,L., Arnheiter,H., Copeland,N.G. and Jenkins,N.A. (2002) Mitf and Tfe3, two members of the Mitf-Tfe family of bHLH-Zip transcription factors, have important but functionally redundant roles in osteoclast development. *Proc. Natl Acad. Sci. USA*, **99**, 4477–4482.
12. Weterman,M.A., Wilbrink,M. and Geurts van Kessel,A. (1996) Fusion of the transcription factor TFE3 gene to a novel gene, PRCC, in t(X;1)(p11;q21)-positive papillary renal cell carcinomas. *Proc. Natl Acad. Sci. USA*, **93**, 15294–15298.
13. Weterman,M.A., Wilbrink,M., Janssen,I., Janssen,H.A., van den Berg,E., Fisher,S.E., Craig,I. and Geurts van Kessel,A. (1996) Molecular cloning of the papillary renal cell carcinoma-associated translocation (X;1)(p11;q21) breakpoint. *Cytogenet. Cell. Genet*., **75**, 2–6.
14. Clark,J., Lu,Y.J., Sidhar,S.K., Parker,C., Gill,S., Smedley,D., Hamoudi,R., Linehan,W.M., Shipley,J. and Cooper,C.S. (1997) Fusion of splicing factor genes PSF and NonO (p54nrb) to the TFE3 gene in papillary renal cell carcinoma. *Oncogene*, **15**, 2233–2239.
15. Heimann,P., El Housni,H., Ogur,G., Weterman,M.A., Petty,E.M. and Vassart,G. (2001) Fusion of a novel gene, RCC17, to the TFE3 gene in t(X;17)(p11.2;q25.3)-bearing papillary renal cell carcinomas. *Cancer Res.*, **61**, 4130–4135.
16. Ladanyi,M., Lui,M.Y., Antonescu,C.R., Krause-Boehm,A., Meindl,A., Argani,P., Healey,J.H., Ueda,T., Yoshikawa,H., Meloni-Ehrig,A. *et al.* (2001) The der(17)t(X;17)(p11;q25) of human alveolar soft part sarcoma fuses the TFE3 transcription factor gene to ASPL, a novel gene at 17q25. *Oncogene*, **20**, 48–57.
17. Argani,P., Antonescu,C.R., Illei,P.B., Lui,M.Y., Timmons,C.F., Newbury,R., Reuter,V.E., Garvin,A.J., Perez-Atayde,A.R., Fletcher,J.A. *et al.* (2001) Primary renal neoplasms with the ASPL-TFE3 gene fusion of alveolar soft part sarcoma: a distinctive tumor entity previously included among renal cell carcinomas of children and adolescents. *Am. J. Pathol.*, **159**, 179–192.
18. Kuiper,R.P., Schepens,M., Thijssen,J., van Asseldonk,M., van den Berg,E., Bridge,J., Schuuring,E., Schoenmakers,H.F.P.M. and Geurts van Kessel,A. (2003) Upregulation of the transcription factor TFEB in t(6;11)(p21;q13)-positive renal cell carcinomas due to promoter substitution. *Hum. Mol. Genet.*, **12**, 1661–1669.
19. Davis,I.J., Hsi,B.L., Arroyo,J.D., Vargas,S.O., Yeh,Y.A., Motyckova,G., Valencia,P., Perez-Atayde,A.R., Argani,P., Ladanyi,M. *et al.* (2003) Cloning of an alpha-TFEB fusion in renal tumors harboring the t(6;11)(p21;q13) chromosome translocation. *Proc. Natl Acad. Sci. USA*, **100**, 6051–6056.
20. Rehli,M., Lichanska,A., Cassady,A.I., Ostrowski,M.C. and Hume,D.A. (1999) TFEC is a macrophage-restricted member of the microphthalmia-TFE subfamily of basic helix–loop–helix leucine zipper transcription factors. *J. Immunol.*, **162**, 1559–1565.
21. Takemoto,C.M., Yoon,Y.J. and Fisher,D.E. (2002) The identification and functional characterization of a novel mast cell isoform of the microphthalmia-associated transcription factor. *J. Biol. Chem.*, **277**, 30244–30252.
22. Weilbaecher,K.N., Hershey,C.L., Takemoto,C.M., Horstmann,M.A., Hemesath,T.J., Tashjian,A.H. and Fisher,D.E. (1998) Age-resolving osteopetrosis: a rat model implicating microphthalmia and the related transcription factor TFE3. *J. Exp. Med.*, **187**, 775–785.
23. Hallsson,J.H., Favor,J., Hodgkinson,C., Glaser,T., Lamoreux,M.L., Magnusdottir,R., Gunnarsson,G.J., Sweet,H.O., Copeland,N.G., Jenkins,N.A. *et al.* (2000) Genomic, transcriptional and mutational analysis of the mouse microphthalmia locus. *Genetics*, **155**, 291–300.
24. Udono,T., Yasumoto,K., Takeda,K., Amae,S., Watanabe,K., Saito,H., Fuse,N., Tachibana,M., Takahashi,K., Tamai,M. *et al.* (2000) Structural organization of the human microphthalmia-associated transcription factor

gene containing four alternative promoters. *Biochim. Biophys. Acta*, **1491**, 205–219.
25. Amae,S., Fuse,N., Yasumoto,K., Sato,S., Yajima,I., Yamamoto,H., Udono,T., Durlu,Y.K., Tamai,M., Takahashi,K. *et al.* (1998) Identification of a novel isoform of microphthalmia-associated transcription factor that is enriched in retinal pigment epithelium. *Biochem. Biophys. Res. Commun.*, **247**, 710–715.
26. Fuse,N., Yasumoto,K., Takeda,K., Amae,S., Yoshizawa,M., Udono,T., Takahashi,K., Tamai,M., Tomita,Y., Tachibana,M. *et al.* (1999) Molecular cloning of cDNA encoding a novel microphthalmia-associated transcription factor isoform with a distinct amino-terminus. *J. Biochem.*, **126**, 1043–1051.
27. Oboki,K., Morii,E., Kataoka,T.R., Jippo,T. and Kitamura,Y. (2002) Isoforms of mi transcription factor preferentially expressed in cultured mast cells of mice. *Biochem. Biophys. Res. Commun.*, **290**, 1250–1254.
28. Takeda,K., Yasumoto,K., Kawaguchi,N., Udono,T., Watanabe,K., Saito,H., Takahashi,K., Noda,M. and Shibahara,S. (2002) Mitf-D, a newly identified isoform, expressed in the retinal pigment epithelium and monocyte-lineage cells affected by Mitf mutations. *Biochim. Biophys. Acta*, **1574**, 15–23.
29. Fink,L., Seeger,W., Ermert,L., Hanze,J., Stahl,U., Grimminger,F., Kummer,W. and Bohle,R.M. (1998) Real-time quantitative RT–PCR after laser-assisted cell picking. *Nat. Med.*, **4**, 1329–1333.
30. Cuppen,E., Gerrits,H., Pepers,B., Wieringa,B. and Hendriks,W. (1998) PDZ motifs in PTP-BL and RIL bind to internal protein segments in the LIM domain protein RIL. *Mol. Biol. Cell*, **9**, 671–683.
31. Carr,C.S. and Sharp,P.A. (1990) A helix–loop–helix protein related to the immunoglobulin E box-binding proteins. *Mol. Cell. Biol.*, **10**, 4384–4388.
32. Phelps,D.E., Hsiao,K.M., Li,Y., Hu,N., Franklin,D.S., Westphal,E., Lee,E.Y. and Xiong,Y. (1998) Coupled transcriptional and translational control of cyclin-dependent kinase inhibitor p18INK4c expression during myogenesis. *Mol. Cell. Biol.*, **18**, 2334–2343.
33. Wang,Y., Newton,D.C., Robb,G.B., Kau,C.L., Miller,T.L., Cheung,A.H., Hall,A.V., Van Damme,S., Wilcox,J.N. and Marsden,P.A. (1999) RNA diversity has profound effects on the translation of neuronal nitric oxide synthase. *Proc. Natl Acad. Sci. USA*, **96**, 12150–12155.
34. Chung,M.C., Kim,H.K. and Kawamoto,S. (2001) TFEC can function as a transcriptional activator of the nonmuscle myosin II heavy chain-A gene in transfected cells. *Biochemistry*, **40**, 8887–8897.
35. Artandi,S.E., Merrell,K., Avitahl,N., Wong,K.K. and Calame,K. (1995) TFE3 contains two activation domains, one acidic and the other proline-rich, that synergistically activate transcription. *Nucleic Acids Res.*, **23**, 3865–3871.
36. Wu,M., Hemesath,T.J., Takemoto,C.M., Horstmann,M.A., Wells,A.G., Price,E.R., Fisher,D.Z. and Fisher,D.E. (2000) c-Kit triggers dual phosphorylations, which couple activation and degradation of the essential melanocyte factor Mi. *Genes Dev.*, **14**, 301–312.
37. Antequera,F. (2003) Structure, function and evolution of CpG island promoters. *Cell Mol. Life Sci.*, **60**, 1647–1658.
38. Roman,C., Cohn,L. and Calame,K. (1991) A dominant negative form of transcription activator mTFE3 created by differential splicing. *Science*, **254**, 94–97.
39. Hemesath,T.J., Price,E.R., Takemoto,C., Badalian,T. and Fisher,D.E. (1998) MAP kinase links the transcription factor Microphthalmia to c-Kit signalling in melanocytes. *Nature*, **391**, 298–301.
40. Benezra,R., Davis,R.L., Lockshon,D., Turner,D.L. and Weintraub,H. (1990) The protein Id: a negative regulator of helix–loop–helix DNA binding proteins. *Cell*, **61**, 49–59.
41. Norton,J.D., Deed,R.W., Craggs,G. and Sablitzky,F. (1998) Id helix–loop–helix proteins in cell growth and differentiation. *Trends Cell Biol.*, **8**, 58–65.
42. Pozniak,C.D., Radinovic,S., Yang,A., McKeon,F., Kaplan,D.R. and Miller,F.D. (2000) An anti-apoptotic role for the p53 family member, p73, during developmental neuron death. *Science*, **289**, 304–306.
43. Yang,A., Kaghad,M., Wang,Y., Gillett,E., Fleming,M.D., Dotsch,V., Andrews,N.C., Caput,D. and McKeon,F. (1998) p63, a p53 homolog at 3q27–29, encodes multiple products with transactivating, death-inducing, and dominant-negative activities. *Mol. Cell*, **2**, 305–316.
44. Nakagawa,T., Takahashi,M., Ozaki,T., Watanabe,K.K., Todo,S., Mizuguchi,H., Hayakawa,T. and Nakagawara,A. (2002) Autoinhibitory regulation of p73 by Delta Np73 to modulate cell survival and death through a p73-specific target element within the Delta Np73 promoter. *Mol. Cell. Biol.*, **22**, 2575–2585.
45. Fillippovich,I., Sorokina,N., Gatei,M., Haupt,Y., Hobson,K., Moallem,E., Spring,K., Mould,M., McGuckin,M.A., Lavin,M.F. *et al.* (2001) Transactivation-deficient p73alpha (p73Deltaexon2) inhibits apoptosis and competes with p53. *Oncogene*, **20**, 514–522.
46. Yang,A., Schweitzer,R., Sun,D., Kaghad,M., Walker,N., Bronson,R.T., Tabin,C., Sharpe,A., Caput,D., Crum,C. *et al.* (1999) p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. *Nature*, **398**, 714–718.