



The 5'ends of TFEB in melanoma

Katrín Möller



**Líf- og umhverfisvísindadeild
Háskóli Íslands
2014**

The 5`ends of TFEB in melanoma

Katrín Möller

12 eininga ritgerð sem er hluti af
Baccalaureus Scientiarum gráðu í Sameindalíffræði

Leiðbeinandi
Eiríkur Steingrímsson

Líf- og umhverfisvísindadeild
Verkfræði- og náttúruvísindasvið
Háskóli Íslands
Reykjavík, maí 2014

The 5'ends of TFEB in melanoma

12 eininga ritgerð sem er hluti af *Baccalaureus Scientiarum* gráðu í Sameindalíffræði

Höfundarréttur © 2014 Katrín Möller

Öll réttindi áskilin

Líf- og umhverfisvísindadeild

Verkfræði- og náttúruvísindasvið

Háskóli Íslands

Sturlugata 7

101 Reykjavík

Sími: 525 4600

Skráningarupplýsingar:

Katrín Möller, 2014, *The 5'ends of TFEB*, BS ritgerð, Líf- og umhverfisvísindadeild, Háskóli Íslands, 43 bls.

Prentun: Háskólaprent ehf.

Reykjavík, maí 2014

Útdráttur

Leysikornið er eins konar endurvinnslustöð frumunnar. Það gegnir lykilhlutverki í viðbrögðum við næringarskort og hreinsar upp gölluð prótein eða gömul og úr sér gengin frumulíffæri með sjálfsáti. Niðurbrotsefnin eru síðan nýtt sem orkugjafar eða sem byggingaeiningar frumunnar. Lykilstjórnprótein leysikornsins og sjálfsáts er umritunarpátturinn TFEB. Próteinkomplex á yfirborði leysikornsins sendir boð um næringarástand frumunnar til kjarnans í gegnum TFEB sem virkjar umritun ýmissa gena sem koma við sögu í viðbragði við sverti og uppsöfnun á gölluðum frumefnum. Undir venjulegum kringumstæðum er TFEB staðsett í umfrymi en er flutt inn í kjarna við hungur eða stress. Náskyldur umritunarpáttur, MITF, er lykilstjórnprótein litkorna í litfrumum, en litkorn eru leysikorns-skyld frumulíffæri sem framleiða litarefni. Rannsóknir hafa sýnt að MITF er eitt þeirra æxlisgena sem koma við sögu í sortuæxlum. Staðsetningu MITF er stýrt á sambærilegan hátt og TFEB en í sortuæxlum finnst MITF nánast eingöngu í kjarna. Það ísóform MITF sem er ríkjandi í litfrumum og sortuæxlum, MITF-M, vantar amínó-endann sem geymir boð um að fara aftur í umfrymi. Nýlegar ónæmislitánir á TFEB í sortuæxlum, sem gerðar voru á tilraunastofu Eiríks Steingrímssonar, hafa sýnt að TFEB er aðallega staðsett í kjarna í sortuæxlum, líkt og MITF. Því var framkvæmd tilraun til að athuga 5'-enda TFEB umrita í sortuæxlum. Tilraunin leiddi í ljós að mörg mismunandi TFEB umrit eru til staðar í sortuæxlisfrumum og mörg þeirra vantar 5'-endann og þar af leiðandi einnig merkið um staðsetningu próteinsins í umfrymi.

Abstract

The lysosome is the recycling center of the cell, where intracellular and extracellular material is degraded and reused for energy production or recycled for growth. Intracellular material is degraded through autophagy. The lysosome and autophagy are key components of the cells responses to nutrient deprivation and cellular stress. A protein complex at the lysosomal membrane signals the nutrient status of the cell to the nucleus via the transcription factor TFEB. Under normal conditions TFEB is located in the cytoplasm but during starvation it is translocated to the nucleus where it acts as an activator for lysosomal and autophagic genes, and other responses to starvation and cellular stress. MITF, a close relative to TFEB, is the master regulator of melanosomes in melanocytes. Melanosomes are lysosome-related organelles that produce melanin and studies have shown that MITF is one of the key oncogenes in melanoma. The location of MITF is regulated in a similar manner to TFEB but in melanocytes, MITF is mainly located in the nucleus. The dominant MITF isoform in melanocytes, MITF-M, is missing the N-terminus found in most other isoforms and this N-terminus has a cytoplasmic localization signal. Immunostaining experiments, performed in Steingrímson's lab, show TFEB mainly located to the nucleus in melanoma cells. This indicates that TFEB isoforms may also lack the N-terminus in melanoma cells in a similar manner to MITF. The aim of the study presented in this thesis was to observe the different 5'ends of TFEB in melanoma cells. The results show many different transcript variants with different 5'ends in melanoma cells and all of the variants observed were missing the N-terminal cytoplasmic localization signal.

Table of contents

Figures.....	vii
Tables.....	viii
Abbreviations.....	ix
Acknowledgements.....	x
1 Introduction	2
1.1 Lysosomes.....	2
1.2 Autophagy	3
1.3 TFEB	5
1.4 MITF and Melanoma	6
2 Aim of this project	8
3 Materials and methods.....	9
3.1 Bioinformatics.....	9
3.2 Cell culturing.....	9
3.2.1 The 501 cell culture	9
3.2.2 The Skmel28 cell culture	10
3.3 RNA extraction and purification	10
3.3.1 RNA isolation	10
3.3.2 DNase treatment	11
3.3.3 RNA purification	11
3.3.4 mRNA isolation	12
3.3.5 cDNA synthesis	12
3.4 5'RACE.....	13
3.4.1 Primer designing	13
3.4.2 The 5'RACE procedure	13
3.4.3 Polymerase chain reaction on 5'RACE products	15
3.4.4 Gel Electrophoresis on 5'RACE products.....	15
3.5 Cloning.....	16
3.5.1 DNA extraction from gel.....	16
3.5.2 Digestion and ligation.....	17
3.5.3 Transformation	17
3.5.4 Selection and culturing	17
3.5.5 Isolation of plasmid	18
3.6 Sequencing	18
3.7 PCR with testing primers	19
3.7.1 Primer designed	19
3.7.2 PCR on purified mRNA	20
3.7.3 Gel Electrophoresis on testing primer products	20

4 Results.....	22
4.1 Many TFEB transcript variants exist in melanoma	22
4.2 Many 5'ends of TFEB discovered	23
4.3 Sequencing revealed the existence of many transcript variants.....	25
4.4 The identified intronic sequences all have a translation start site in-frame with exon 4.....	28
4.5 Confirmation of the presence of intronic sequences	28
5 Discussions.....	30
References	31
Supplementary figures.....	33
Appendix	34

Figures

Figure 1: The main functions of the lysosome.....	2
Figure 2: The formation of the autophagosome and autolysosome.	4
Figure 3: mTOR keeps TFEB and related proteins inactive by phosphorylation.	5
Figure 4: The location of the gene specific primers.....	13
Figure 5: Location of the testing primer pairs.....	19
Figure 6: All transcript variants found in Ensembl.	22
Figure 7: All transcript variants found in melanoma samples from the Cancer Genome Atlas.....	23
Figure 8: The final results of the 5'RACE.	24
Figure 9: The 5'ends of the transcript variants in 501mel cells blated against the human genome.	26
Figure 10: The 5'ends of the transcript variants in Skmel28 cells blated against the human genome.	27
Figure 11: Gel electrophoresis of the PCR products of DNase treated RNA.	29

Tables

Table 1: Information on the gel pieces cut for DNA extraction.....	16
Table 2: Information on colonies grown from selection plates.....	17
Table 3: Information on testing primers.....	20

Abbreviations

ATP	Adenosine triphosphate
BLAST	Basic local alignment tool
BLAT	Blast like alignment tool
CLEAR	Coordinate lysosomal expression and regulation
CMA	Chaperone mediated autophagy
DMEM	Dulbecco's Modified Eagle Medium
FT	Flow through
HLH	Helix-loop-helix
LYNUS	Lysosomal nutrient sensing machinery
MITF	Microphthalmia-associated transcription factor
mTOR	Mammalian target of rapamycin
mTORC1	mTOR complex 1
N-terminal	Amino-terminal
PCR	Polymerase chain reaction
RACE	Rapid amplification of cDNA ends.
Rag GTPase	Rag guanosine triphosphatase
ROS	Reactive oxygen species
-RT	without reverse transcription
RT	Reverse transcription
rt	Room temperature
TdT	Terminal deoxynucleotide transferase
TFE3	Transcription factor E3
TFEB	Transcription factor EB
TFEC	Transcription factor EC
v-ATPase	H ⁺ -adenosine triphosphatase

Acknowledgements

I would like to thank Eiríkur Steingrímsson for the opportunity to be a part of his research team. It has been very interesting and educational. I would also like to thank Kristín Bergsteinsdóttir and Margrét H. Ögmundsdóttir for their patience, help and advice with this project. Steingrímson's lab members have also been very helpful and kind to assist.

1 Introduction

1.1 Lysosomes

The lysosome is the recycling station of the cell. Both intracellular and extracellular waste is transported to the lysosome where it is either degraded or recycled. But degradation and recycling are not the only functions of the lysosome, as it is also involved in other fundamental processes such as lipid metabolism, plasma membrane repair, secretion and signaling.

Lysosomes are vacuolar organelles with an acidic lumen separated from the cytoplasm by a single-lipid bilayer membrane. The acidic lumen is filled with hydrolases which break down substrates specifically marked for degradation. These substrates can be both extracellular and intracellular material but the pathways by which they reach the lysosome vary. Extracellular material is taken up by the endocytic pathway whereas intracellular material reaches the lysosome through autophagy. Intracellular material can either be proteins or old and malfunctioning organelles and the degraded parts can be reused for repair, secretion, building new organelles or for energy production. The secreted particles are transported out of the cell through lysosomal exocytosis and are used for different functions such as pigmentation, signaling, immune responses and cell protection.

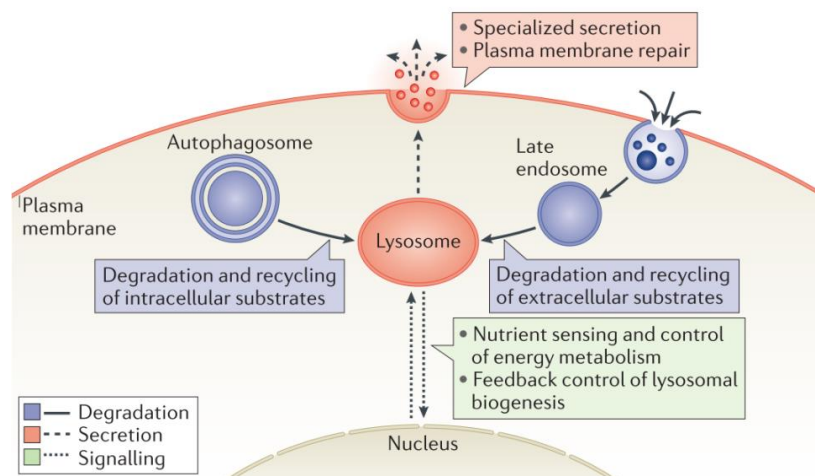


Figure 1: The main functions of the lysosome. *The lysosomal functions can be divided in three: degradation, secretion and signalling. External material reaches the lysosome through endocytosis, cytosolic material reaches it through autophagy and the degraded particles can be secreted through exocytosis. Proteins on the lysosomal surface sense the nutrient status of the lysosome and send signals to the nucleus (Settembre et al., 2013a)*

There are about 60 different hydrolases in the lysosomal lumen which are active at acidic pH, including protein families like sulphatases, peptidases, glycosidases, phosphatases, lipases and nucleases. The inside of the lysosomal membrane is covered with a thick polysaccharide layer, called glycocalyx, to protect the membrane from being degraded by these hydrolases. Many proteins are on the outside of the membrane and participate in the function and control of the lysosome. For example there are ion channels that help keep the lumen at acidic pH, transport proteins that can transport substrates marked for degradation into the lumen, trafficking and fusion machinery proteins that take part in the fusion of the lysosome with other vacuoles and nutrient sensing protein complexes that send signals to the nucleus about the nutrient status of the cell (Settembre et al., 2013a). An important regulatory complex is also located on the lysosomal surface, called the LYNUS complex (Lysosomal Nutrient Sensing Machinery). It is made of the mammalian target of rapamycin (mTOR) complex 1 (mTORC1), Rag guanosine triphosphatases (Rag GTPases) and vacuolar H⁺-adenosine triphosphatases (v-ATPases). Amino acids in the cytoplasm activate the Rag GTPases which, together with the v-ATPases, help locating mTORC1 on the lysosomal surface. mTORC1 is a protein kinase complex that is activated at the lysosomal surface and regulates lysosomal biogenesis and autophagy in response to nutrient deficiency and lysosomal stress (Zoncu et al., 2011).

1.2 Autophagy

Cells use autophagy for lysosomal degradation of intracellular material, such as protein, lipids or old and malfunctioning organelles. It is a part of normal development and function of the cell but it can be activated during stress conditions such as starvation (Chen and Klionsky, 2011). During these conditions anabolic activity is decreased and intracellular catabolism is increased to generate energy for cell survival (Settembre and Ballabio, 2011). Autophagy is divided into three classes, chaperone-mediated autophagy (CMA), microautophagy and macroautophagy. During CMA a chaperone protein binds to its target protein and to the lysosomal surface and then translocates the protein into the lysosome where degradation takes place. Microautophagy is the least understood of the three and translocates substrates from the cytosol into the lysosome through direct invagination. Macroautophagy is the best characterized type of autophagy and will here be referred to as autophagy. A phagophore forms and surrounds the cytoplasmic proteins or organelles, closes and forms an autophagosome, a double-membrane vesicle which then fuses with the lysosome to form the autolysosome.

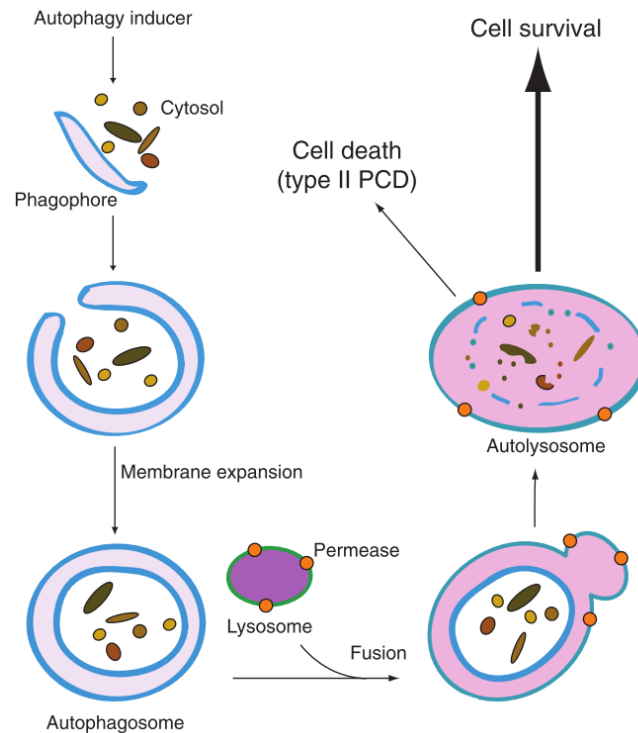


Figure 2: The formation of the autophagosome and autolysosome. A phagophore forms around substrates or organelles marked for degradation. When it closes completely it is called autophagosome. The autophagosome fuses with the lysosome where the lysosomal hydrolases degrade the contents of the autophagosome (Chen and Klionsky, 2011)

Inside the autolysosome the cytoplasmic proteins or organelles are degraded by the lysosomal hydrolases (Chen and Klionsky, 2011). The autophagic process generates amino acids and fatty acids for protein synthesis and β -oxidation, thereby generating new building material and energy in the form of ATP for cell survival (Settembre and Ballabio, 2011). But autophagy also breaks down malfunctioning organelles and misfolded proteins. Specific autophagy of mitochondria, called mitophagy, is a potential mechanism to reduce ROS (reactive oxygen species) and oxidative stress, by eliminating damaged mitochondria. This could be an important mechanism in tumor suppression by preventing tissue damage and cancer. This mechanism could also be tumor promoting, improving tumor cell survival in hypoxic regions by providing energy and cell maintenance material (reviewed in White, 2012).

1.3 TFEB

TFEB (transcription factor EB) is a key regulator of autophagy and lysosomal biogenesis. It is a member of the MITF (microphthalmia-associated transcription factor) subfamily the basic helix-loop-helix leucine zipper (bHLH-Zip) transcription factor family. This subfamily also includes MITF, a key regulator of melanogenesis, TFEC and TFE3 (Rehli et al., 1999; Sardiello et al., 2009). Members of this family share a DNA binding domain, a HLH and a Zip domain and they regulate gene expression through binding to DNA as a homodimer or as a heterodimer with another family member (Steingrímsson et al., 2004).

TFEB is normally located in the cytoplasm where it is phosphorylated by mTORC1 and bound to the 14-3-3 proteins, thereby kept inactive (Roczniak-ferguson et al., 2012). During starvation or lysosomal stress mTORC1 is inactivated and TFEB does thereby not get phosphorylated by the kinase. This results in the translocation of TFEB to the nucleus where it acts as a master regulator of the CLEAR (Coordinated lysosomal expression and regulation) network (Settembre et al., 2012).

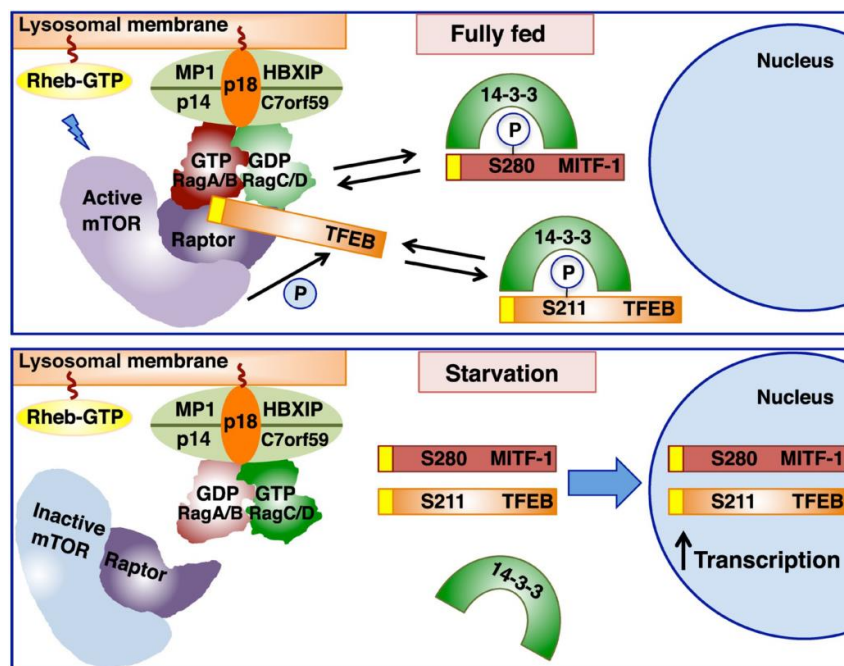


Figure 3: mTOR keeps TFEB and related proteins inactive by phosphorylation. Under normal conditions TFEB is kept in the cytoplasm through phosphorylation of serine 211. The 14-3-3 proteins bind to this phosphorylated serine 211 and thereby inhibit the translocation of TFEB to the nucleus. During starvation mTOR is inactivated. TFEB dephosphorylates and is able to translocate to the nucleus where it controls the transcription of genes involved in starvation responses. (Martina and Puertollano, 2013)

Members of this network are involved in various lysosomal mechanisms, such as lipid metabolism, protein degradation, autophagy and other metabolic responses to starvation as well as in exo- and endocytosis, phagocytosis and immune responses. TFEB recognizes and binds to E-box DNA sites at the promoters of the CLEAR network genes and thereby promotes their transcription (Palmieri et al., 2011). The *TFEB* gene has CLEAR sites in its own promoter region so TFEB can activate the expression of the *TFEB* gene and thereby induce and sustain the starvation response through an autoregulatory-feedback loop (Settembre et al., 2013b). When the nutrient level in the cell is back to normal or the lysosomal stress has been reduced it is important that TFEB is recruited back into the cytoplasm. Recent studies have shown that the Rag GTPases control this by interacting with the first 30 amino acid residues of TFEB, more precisely S3R4 and Q10L11, thereby promoting the recruitment of TFEB to the lysosomes, where they can be phosphorylated again by the active mTORC1. These studies also showed that some MITF isoforms are regulated in a similar manner, through interactions between Rag GTPases and the N-terminal amino acids of MITF (Martina and Puertollano, 2013).

1.4 MITF and Melanoma

Melanocytes originate from the neural crest cells as melanoblasts which migrate into the ectoderm and differentiate into mature melanocytes (Cheli et al., 2010). Melanocytes are dendritic, pigment-producing cells that are found in hair follicles, the epidermis, the inner ear, the choroid of the eye, the Harderian gland and the heart (reviewed in Steingrímsson et al., 2004). Their most studied role is melanin production for skin and eye pigmentation. Skin pigmentation is an important protection against UV radiation, thus preventing DNA damage and cancer. Melanocytes produce and transfer melanosomes, full of melanin, to neighboring keratinocytes through the dendrite tips of the melanocytes. Melanosomes are lysosome-related organelles that contain numerous proteins involved in melanogenesis. Three of these are *Tyrosinase*, *Tyrp1* and *DCT*, enzymes that convert tyrosine to melanin pigment.

During development, MITF is a master regulator of melanocyte growth, function and differentiation. It regulates the expression of *Tyrosinase*, *Tyrp1* and *DCT* as well as multiple other genes, many of which are involved in melanin synthesis and melanogenesis (Cheli et al., 2010). But MITF also plays an important role in melanoma, as mutations, overexpression and gene amplification have been detected in human melanoma. It is therefore proposed to be an

oncogene and is a very sensitive marker for melanoma (Garraway and Sellers, 2006; King et al., 1999).

MITF has several different isoforms, where MITF-A and MITF-H are expressed in multiple tissues but the MITF-M form is specific to melanocytes and melanoma (Yasumoto et al., 1998). The N-terminus of MITF-A and MITF-H have a high degree of homology to the N-terminus of TFEB but the N-terminus of MITF-M has no homology with TFEB. Interestingly, the MITF-M form accumulates in the nucleus of melanocytes (Grill et al., 2013; Martina and Puertollano, 2013).

2 Aim of this project

TFEB is a key regulator of autophagy, which is an important recycling mechanism in the cell. Autophagy is thought to be both tumor repressing and tumor promoting so it is important to know the controlling mechanisms of autophagy. The basal level of autophagy is high in melanoma cells but the underlying mechanisms of autophagy control in these cells are unclear.

A recent immunostaining of TFEB in our lab using a mouse monoclonal TFEB antibody showed that TFEB to be primarily nuclear in melanoma cells. This might suggest that the cytoplasmic localization signal at the N-terminus of TFEB isoforms are mutated or missing in melanoma cells, resulting in a similar localization as the lineage specific oncogene MITF-M.

The aim of this project was to map which N-terminal isoforms of TFEB are present in melanoma cells. Two different human melanoma cell lines were used in this study, 501mel and Skmel28. RNA was isolated from the two cell lines and the 5'ends of TFEB transcripts reverse transcribed. The ends were amplified and separated on gel to see if there were many 5'end variants in the melanoma cells.

The ends were extracted from the gel and ligated into a vector and the vectors transformed into *E.coli* for amplification. The amplified plasmids were purified and the TFEB inserts were sequenced for exact analysis of which 5'ends were found in melanoma cells.

3 Materials and methods

3.1 Bioinformatics

The aim of the first part of this project was to get information about the currently known isoforms of TFEB in melanoma and healthy cells. All TFEB isoforms found on ENSEMBL were mapped as well as the TFEB isoforms from melanoma cell lines found in the Cancer Genome Atlas, using the UCSC genome browser (picture).

- For sequence alignment, the Clustal Omega program was used [<http://www.ebi.ac.uk/Tools/msa/clustalo/>].
- For translation of one or more reading frames the EMBOSS sixpack program was used [https://www.ebi.ac.uk/Tools/st/emboss_sixpack/].
- To search for a sequence in the human genome, the Basic Local Alignment Tool (BLAST) from NCBI was used [<http://blast.ncbi.nlm.nih.gov/Blast.cgi>].
- For aligning one or more sequences visually to the human genome the Blast Like Alignment Tool (BLAT) from the UCSC genome browser was used [<http://genome.ucsc.edu/cgi-bin/hgBlat?command=start>].
- To search for possible domains in new sequences the NCBI Conserved Domain Search tool was used [<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>].

3.2 Cell culturing

Two melanoma cell lines were cultured for use in this study, 501mel and Skmel28. Both cell lines were cultured in 25 mL flasks in RPMI 1640 medium from GIBCO with 10% FBS and 0,5% PenStrep at 37°C until at least 80% confluent.

3.2.1 The 501 cell culture

The 501mel cell line was originally a gift from Ruth Halaban (Halaban et al., 1997). For the 5'RACE the 501mel cells were cultured by Kristín Bergsteinsdóttir on RPMI 1640 (+ L-Glutamine, + 25mM HEPES) medium from GIBCO, with 10% FBS and 0,5% PenStrep, in T25 flasks at 37°C and 5% CO₂. When they were 90-100% confluent they were used for RNA extraction.

For the mRNA isolation the 501mel cells were cultured by Kimberly Anderson on RPMI 1640 medium, with 10% FBS, in T25 flasks at 37°C and 5% CO₂. When they were 80-90% confluent they were used for RNA extraction

3.2.2 The Skmel28 cell culture

The human melanoma cell line Skmel28 (ATCC® HTB-72™) was purchased from ATCC. For the 5' RACE a P(5) lineage of this cell line was retrieved from liquid nitrogen and thawed at body temperature. DMEM (+4,5g/L D-Glucose, + NEAA, - L-Glutamine, - Pyruvate) medium from GIBCO with 10% FBS and 0,5% PenStrep was added to the cell culture and the tube was centrifuged at 1000 rpm for 5 min. The medium was removed and the pellet resuspended in new medium. The cells were cultured in T25 flask at 37°C and 5% CO₂. The medium was changed after 1 and 3 days and on day 4 they had gained 60-70% confluency. The cells were then split into two flasks. The medium was removed and EDTA/Trypsin solution added and removed from the flask. EDTA/Trypsin solution was added a second time and the culture incubated at 37°C and 5% CO₂ for 2-3 min. DMEM medium was added and the cells moved to a 15 mL Falcon tube. The tube was centrifuged at 1000 rpm for 5 min. Almost all the medium was removed and new medium added to the tube and two clean flasks. The cells were resuspended in the new medium and divided into the two new flasks. They were then grown at 37°C and 5% CO₂. After one day the medium was removed and new DMEM medium was added to one flask and RPMI 1640 medium added to the other one. After two days in culture the DMEM grown cells were about 50-60% confluent but the RPMI 1640 grown cells were about 70-80% confluent. The RPMI 1640 grown cells were therefore used for RNA extraction.

For the mRNA isolation the Skmel28 cells were cultured by Kimberly on RPMI 1640 medium in T25 flasks at 37°C and 5% CO₂. When they were 60-70% confluent they were used for RNA extraction

3.3 RNA extraction and purification

Total RNA was extracted from cells using TRIzol (Ambion) and re-purified with the RNeasy Mini kit from Qiagen.

3.3.1 RNA isolation

TRIzol protocol from Ambion was used for this procedure and the 501mel and Skmel28 cell cultures. The medium was removed from the cultures and 2 mL TRIzol added. The cells were lysed by pipetting them up and down several times. The lysate from each cell culture was

divided into two 1,7 mL Nuclease free Eppendorf tubes from Ambion and incubated at room temperature (rt) for 5 min. Chloroform was added, the tubes shaken vigorously for 15 seconds and incubated at rt for 3 min. The lysate was then centrifuged at 12.000 g for 15 min at 4°C to separate the lysate into different phases. The clear upper aqueous layer, containing RNA, was pipetted into new Nuclease free Eppendorf tubes and 100% isopropanol added. The new tubes were incubated at rt for 10 min and then centrifuged at 12.000 g for 10 min at 4°C. The isopropanol was removed and the RNA pellet washed two times with 75% ethanol, centrifuging at 7.500 g for 5 min in each wash. After the ethanol was removed the tubes were centrifuged again at 7.500 g for 30 sek and the last ethanol removed. The tubes were opened and the RNA pellet dried for 10 min. RNase free water from Ambion was used to resuspend the RNA. The resuspended RNA was kept on ice for 10 min and then the concentration measured using Nanodrop. The isolated RNA was stored at -80°C and one of the two tubes for each cell culture was purified.

3.3.2 DNase treatment

For mRNA isolation the total RNA samples were DNase treated before purification. The Qiagen RNase free DNase kit was used for the procedure. RNase free water was added to the resuspended RNA so the total reaction volume would become 100 µl. Then RDD buffer (optimized digestion buffer) and DNase I were added and the reaction incubated at 24°C for 10 minutes.

3.3.3 RNA purification

The isolated total RNA from the two cell cultures was purified using RNeasy mini kit from Qiagen. The isolated RNA was diluted so the starting concentration was 2-300 ng/µl. RLT buffer (lysing buffer) was added to the RNA and mixed by pipetting. 100% ethanol was added and the RNA solution pipetted a few times. The solution was then added to the RNeasy spin column in a collection tube and centrifuged at 8000g for 15 seconds. The flow through (FT) was discarded. RPE buffer (wash buffer) was added to the spin column and centrifuged again at 8000g for 15 seconds. The FT was discarded and RPE buffer added again to the column. The column was centrifuged at 8000g again but now for 2 min. The FT and collection tube were discarded, the spin column placed in a new collection tube and centrifuged open at 8000g for 5 min. The FT and collection tube were discarded and the spin column placed in a 1,7 ml RNase free Eppendorf tube. RNase free water was added to the spin column and incubated at rt for 1 min. The spin column in the Eppendorf tube was centrifuged at 10.000g for 1 min. The spin column was discarded and the concentration and quality of the eluted RNA were measured using a Nanodrop spectrophotometer and a Bioanalyzer.

3.3.4 mRNA isolation

mRNA isolation was done on purified, DNase treated RNA from 501mel and Skmel28 cells. The Qiagen Oligotex mRNA mini kit was used for this procedure.

OEB buffer (elution buffer) was kept at 70°C until used and OBB buffer (binding buffer) and the oligotex suspension were kept at 37°C until used. RNase free water was added to the purified RNA until final volume was 250µl. For the binding of RNA to the oligotex, 37°C OBB buffer was added as well as the 37°C oligotex suspension. The RNA-oligotex mixture was incubated at 70°C for 3 minutes and kept at rt for 10 minutes. Now the mixture was centrifuged at 16.000g for 2 minutes and all but 50 µl of the supernatant removed. OW2 buffer (wash buffer) was added to the RNA-oligotex pellet and the pellet resuspended by pipetting. The resuspension was then added to the spin column and the column centrifuged at 16.000g for 1 min. The column was moved to a new 1,7 RNase free collection tube, OW2 buffer added again and the column centrifuged at 16.000g for 1 minute. The column was moved to a new collection tube and 70°C hot OEB elution buffer added to the column. The elution buffer was pipetted up and down a few times before the column was centrifuged at 16.000g for 1 min. The elution step was repeated with new OEB buffer for maximum recovery. The spin column was discarded and the concentration and quality of the mRNA were measured using a Nanodrop spectrometer and a Bioanalyzer.

3.3.5 cDNA synthesis

The purified mRNA from 501mel and Skmel28 cells was reverse transcribed into cDNA using the High-Capacity cDNA Reverse Transcription Kit from Applied Biosystems.

The reaction for each cell line prepared in two 1,7 ml nuclease free Eppendorf tubes from Ambion. In order to do this, 2 µl 10X RT buffer, 2 µl RT random primers (10 mM), 0,8 µl dNTP mix (10 mM), 1 µl RNase inhibitor and 13,2 µl mRNA were mixed in each Eppendorf tube and 1 µl MultiScribe Reverse Transcriptase (50 U/µl) added to one Eppendorf tube for each cell line. For control without reverse transcription (-RT) 1 µl Nuclease free water from Ambion was added to the other tube for each cell line. The reactions were then incubated at 25°C for 10 minutes, at 37°C for 120 minutes and at 85°C for 5 minutes. The – RT controls were not made at the same time as the cDNA synthesis of the mRNA. The newly synthesized cDNA was stored at -80°C.

3.4 5'RACE

3.4.1 Primer designing

Three specific primers had to be designed for the 5'RACE system to amplify the 5' end of the different TFEB mRNA isoforms in melanoma cells. The first one is used by the reverse transcriptase, Superscript II, and has to pair with the mRNA sequence downstream of the region to be reverse transcribed and amplified using the 5'RACE system. This primer was called gene specific primer 1 (GSP1). Because the aim was to study the „cytoplasmic localization sequence“ which lies at the 5' end of exon 2 in TFEB, this primer was designed to pair with the second exon downstream from this sequence, in exon 4 (Figure 4). GSP1 has the sequence 5'ATTCAGGATTGATGTAG 3' and $T_m = 49,1^{\circ}\text{C}$. The second primer, GSP2, was designed as a nested primer to the GSP1, that is, it pairs to the newly made cDNA upstream of GSP1, also in exon 4. GSP2 was then used in the PCR amplification of the tailed cDNA. GSP2 has the sequence 5'ACATCGTCCAGACGCATAAT 3' and $T_m = 62^{\circ}\text{C}$. The third primer, GSP3, was designed as a nested primer to GSP2, also in exon 4, but with two restriction sites, for SacII and XbaI, and three random nucleotides at the 5' end of the primer. The Universal Anchor Primer (UAP) that is included in the 5'RACE kit, has restriction sites as well, for SalI and SpeI, so this way the PCR product could be cloned into a vector for further analysis. GSP3 has the sequence 5'ACGTCTAGACCGCGGGTTGTCAATGACATCATCCA 3'.

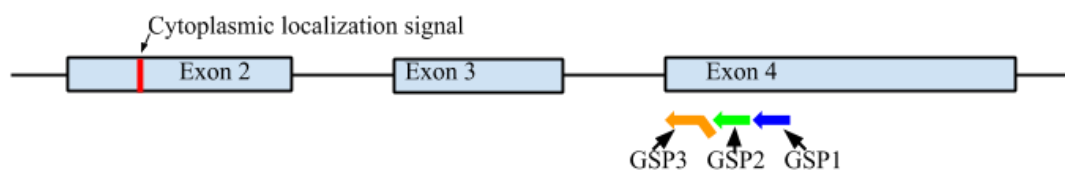


Figure 4: The location of the gene specific primers. *GSP2 was designed as a nested primer to GSP1 and GSP3 as a nested primer to GSP2. GSP3 has a flanking tail with restriction sites for XbaI and SacII. All primers pair to TFEB exon 4.*

3.4.2 The 5'RACE procedure

5'RACE System for Rapid Amplification of cDNA Ends, version 2.0, from Life Technologies, was used for this procedure. The procedure was first performed on control RNA and DNA provided with the kit, to verify that it worked. The results from the control

procedure were good so 5'RACE was performed on the purified RNA samples from the Skmel28 and 501mel cells.

The RNA samples and DEPC-treated water were put in RNase free 1,7 ml Eppendorf tubes as well as GSP1. The samples were centrifuged at 6000g for 30 seconds and incubated at 70°C for 10 min. The samples were then put on ice for 1 min and centrifuged again at 6000g for 30 seconds. 10X PCR buffer, MgCl₂ (25 mM), dNTP mix (10mM) and DTT (0,1mM) were added to the samples and the mixture was centrifuged at 6000g for 30 seconds. The samples were then incubated for 2 minutes at 42°C. Then the Superscript II reverse transcriptase was added to the samples and pipetted up and down a few times. The samples were incubated at 42°C for 50 minutes for the Superscript to work. The cDNA samples were then incubated at 70°C for 15 minutes and centrifuged at 6000g for 30 seconds. Next, RNase mix was used to remove the RNA template along with all the other RNA. The cDNA samples were kept at 37°C while RNase mix was added, then incubated at 37°C for 30 min and centrifuged at 6000g for 30 seconds. 2µl of the RNase treated cDNA were put in TE buffer for possible troubleshooting (A).

A S.N.A.P column system was used to purify the cDNA. Binding solution was added to each Eppendorf tube and the cDNA samples were transferred to the S.N.A.P column. The column was centrifuged at 13.000g for 20 seconds and the FT saved for later (FT-sample). The column was washed four times with 1x Wash buffer, centrifuging at 13.000g for 20 seconds and discarding the FT after each wash. The column was washed again, now two times with 70% ethanol, centrifuging at 13.000g for 20 seconds and discarding the FT after each wash. The column was centrifuged again at 13.000g, now for 1 minute and the column moved to a new collection tube. The cDNA was eluted from the column with 65°C hot DEPC-treated water and centrifuged at 13.000g for 20 seconds. 5µl of the eluted cDNA were put in TE buffer for possible troubleshooting (B).

Next, a poly-C tail was added to the 5' ends of the cDNA using a Terminal deoxynucleotide transferase (TdT). For each sample, 501mel cDNA, Skmel28 cDNA and control cDNA, two 1,7 mL RNase free eppendorf tubes were used for the TdT tailing. To each tube DEPC-water, 5X tailing buffer, dCTP (2mM) and 10µl of cDNA sample were added and the tubes incubated at 94°C for 2,5 min. The mixture was kept on ice for 1 minute and centrifuged at 6000g for 30 seconds. The mixture was then kept on ice. To one of the two tubes for each sample we added 1 µl of TdT. The reaction was incubated at 37°C for 10 minutes and then at 65°C for 10 minutes. The tailed cDNA samples were kept at -20°C overnight.

3.4.3 Polymerase chain reaction on 5'RACE products

Polymerase chain reaction (PCR) was used to amplify the tailed product of the 5'RACE system. HotStart Taq polymerase from Qiagen was used for the procedure.

The tailed cDNA samples and the other PCR reagents were thawed on ice. The cDNA samples that included the TdT reaction were used for the PCR. Three 50 µl reaction for each of the 501mel and Skmel28 samples were prepared using sterile deionized water, 10X PCR buffer, 25 mM MgCl₂ (final concentration 2,0 mM), 10 mM dNTP mix, 10 mM GSP2, 10 mM Abridged Anchor Primer (AAP), that pairs to the oligoT tail, and tailed cDNA. The reaction mixtures were centrifuged lightly and HotStart Taq Polymerase from Qiagen added and mixed. The reactions were performed for 15 minutes at 95°C to activate the polymerase, then 35 PCR cycles of [1 minute at 94°C, 1 minute at either 57°C, 59°C or 61°C for each sample and 2 minutes at 72°C] were performed, then 7 minutes at 72°C and finally the reaction was kept at 5°C.

A second PCR was done with GSP3 to get a more specific result from the 5'RACE system. A 50 µl reaction was prepared for each amplified cDNA sample that used 57°C and 59°C for annealing temperature, using the same amounts and material as in the first one, except now using GSP3 instead of GSP2 and the Universal Amplification Primer (UAP) instead of AAP. The reaction was performed for 15 minutes at 95°C to activate the polymerase, then 30 PCR cycles of [30 seconds at 94°C, 30 seconds at 58°C and 1 minute at 72°C] were performed, then 7 minutes at 72°C and finally the reaction was kept at 4°C.

3.4.4 Gel Electrophoresis on 5'RACE products

To analyze the products from each PCR, 2% agarose gel was used. To make the agarose gel, 1g agarose was mixed with 50 ml of TAE buffer and heated in a microwave until the agarose had dissolved completely. We then added 2,5 µl ethidium bromide to the agarose mixture poured it into a mold with a comb and left for cooling for about half an hour. The comb was removed and the gel placed in an electrophoresis apparatus filled with TAE buffer with the same concentration of ethidium bromide as the gel. For the first PCR 5 µl of each PCR reaction sample were mixed with 1,25 µl 6X loading dye from Fermentas. Each sample was then loaded on the agarose gel. The O'RangeRuler 100 bp DNA ladder from Thermo Scientific was used as a marker. The samples were run at 80 V for 50 minutes and a picture under UV light taken.

3.5 Cloning

The isolated isoforms of TFEB were ligated into a pBluescript vector and then transformed into High Efficiency competent *E.coli* from New England Biolabs.

3.5.1 DNA extraction from gel

The bands that appeared on the gel from the nested PCR on the 5'RACE products were cut of the gel and the DNA was isolated. The remaining products from the nested PCR were run on a gel for the extraction. This time 20 µl of each sample were mixed with 2,2 µl of 6X loading dye and loaded onto a 2% agarose gel. The samples were run at 60V for 85 minutes. The gel was kept under UV light while cutting. For the 501mel cells the PCR products that had an annealing temperature at 57°C during the first PCR were cut of the gel. The bands were not very clearly separated so the 200 bp band was cut out of the gel in one piece, the 3-400 bp bands were cut in one gel piece and the bands bigger than 400 bp were cut in one gel piece. A similar pattern of bands was observed and they were cut in a similar way as the 501mel cells.

The Nucleospin, Gel and PCR cleanup kits from Macherey-Nagel was used for the cleanup procedure. Each gel piece was put in a 1,7 ml Eppendorf tube and weighed. Proper amount of NTI buffer was added, see table 1.

Table 1: *The weight of the gel pieces and the volume of NTI buffer added*

Sample	Weight (mg)	NTI volume (µl)
501mel – 200bp	200	600
501mel – 3-400 bp	210	630
501mel – >400 bp	340	1020
Skmel28 – 200 bp	230	690
Skmel28 – 3-400 bp	210	630
Skmel28 – >400 bp	350	1050

The samples were incubated at 50°C for 15 minutes and the tubes were shaken every once in a while. We loaded 600 µl of each dissolved gel-cDNA mixture on separate spin columns and centrifuged at 11.000g for 30 seconds. The FT was discarded and the loading step repeated until all the mixture from each sample was on the column. Now the columns were washed two times with NT3 buffer, centrifuging at 11.000g for 30 seconds and discarding the FT after each wash. The columns were then centrifuged at 11.000g for 1 minute. 15 µl of 70°C hot NE elution buffer were added to the column and centrifuged at 100g for 1 minute. The column was then centrifuged again at 11.000g for 1 minute and the eluate was loaded on the same column again for maximum recovery. The column was centrifuged again at 11.000g for 1 minute. The cDNA concentration was measured using Nanodrop spectrometry.

3.5.2 Digestion and ligation

The restriction enzymes, SacII and SalI from New England Biolabs, were used to cut both the PCR products and a pBluescript vector. A T4 DNA ligase from Thermo Scientific was used to ligate them together. The restriction enzymes use different buffers so sequential digestion had to be set up. First the enzyme using lower salt concentration, SacII, was used to digest the isolated cDNAa fragment in NEB 2.1 buffer. The reaction incubated at 37°C for 1 hour. Then Tris-HC and NaCl were added to the reaction so the concentration of Tris-HCl was 50 mM and NaCl was 100 mM. The other restriction enzyme, SalI, was added and the reaction incubated at 37°C for 1 hour. The same procedure was done on an empty pBluescript vector. Before the ligation, the digested cDNA fragments and pBluescript vectors were purified using ethanol precipitation. A 20 µl reaction was prepared using 0,015 pmol of cut cDNA fragments and 0,01 pmol cut pBluescript vector, 10X T4 DNA ligase buffer, T4 DNA ligase and Nuclease free water from Ambion. The reaction was incubated at 16°C overnight.

3.5.3 Transformation

The *E.coli* cells were thawed and half of each tube used for each TFEB fragment sample. We added 100-150 pg (or 1 µl) of each fragment to 27 µl of *E.coli* in a 1,7 ml eppendorf tube. The culture was kept on ice for 30 minutes and then heat shocked at 42°C for 30 seconds. The culture was then kept on ice for 5 minutes and NEB SOC media added. The culture was incubated at 37°C for 1 hour in a shaker. Each culture was spread in different volumes, 40 µl or 200 µl (for the 501mel 200bp culture 40 µl and 400 µl were used), on two selection plates using LB media with 100 µg/ml ampicillin and X-gal and IPTG spread over. The plates were incubated overnight at 37°C.

3.5.4 Selection and culturing

White colonies from the plates were selected and transferred to 15 ml Falcon tubes with 3 ml SOC media (Super optimal broth (SOB) + 10mM MgCl₂, 10 mM MgSO₄ and 2 mM glucose) with 100 µg/ml Ampicillin. In total, 10 colonies were cultured from the 501mel cells and 12 colonies from the Skmel28 (see Table 2). One falcon tube with 3 ml SOC media was used as a negative control

Table 2: Colonies selected and which selection plate they were picked from

Sample	Number of colonies	Names of selections	From selection plate
501mel – 200bp	2	A1, A2	400 µl
501mel – 3-400 bp	3	B1, B2, B3	200 µl
501mel – >400 bp	5	C1, C2, C3, C4, C5	40 µl
Skmel28 – 200 bp	3	D1, D2, D3	200 µl
Skmel28 – 3-400 bp	3	E1, E2, E3	200 µl
Skmel28 – >400 bp	5	F1, F2, F3, F4, F5	200 µl
Skmel28 – >400 bp	1	F6	40 µl

The falcon tubes with the medium and selected colonies were incubated at 30°C for 3 hours in a shaker and then at 37°C overnight in a shaker.

For possible later analysis, 250 µl of each overnight culture and 250 µl 50% glycerol were mixed in a 3 ml tube and stored at -80°C

3.5.5 Isolation of plasmid

The plasmids were isolated from the *E.coli* by using the Nucleospin plasmid miniprep kit from Macherey-Nagel. The Falcon tubes were centrifuged at 3000 rpm for 10 minutes at 4°C and the supernatant discarded. Resuspension buffer A1 was added and the pellet resuspended by pipetting. Lysis buffer A2 was added, the resuspension moved to a 1,7 ml Eppendorf tube and then incubated at rt for 6 minutes. Neutralization buffer A3 was then added and the tubes centrifuged two times at 11.000g for 5 min. 600µl of the supernatant were loaded on the spin column and centrifuged at 11.000g for 1 minute. The FT was discarded and AW wash buffer added to the column. The column was centrifuged at 11.000g for 1 minute and washed again, this time with A4 wash buffer, and centrifuged at 11.000g for 1 minute, discarding the FT after each wash. The column was then centrifuged for 11.000g for 2 minutes and placed in a 1,7 ml Eppendorf tube. Elution buffer was added to the column and incubated for 2 minutes at RT. The columns were centrifuged at 11.000g for 1 minute and the column discarded. The plasmid concentrations were measured using Nanodrop spectrophotometer. The plasmids were named by the same system as the colonies picked from the selection plates and then stored at -20°C

3.6 Sequencing

The inserts of the isolated pBluescript plasmids were amplified and sequenced using BigDye® Terminator v1.1 Cycle Sequencing kit and ABI PRISM® 3100 Genetic Analyzer.

For each plasmid a 10µl reaction was prepared in a PCR strip using 2µl BigDye 1.1 Ready Reaction Premix, 1µl BigDye sequencing buffer, 1µl T7 primer (10mM), 5µl nuclease free water from Ambion and 1µl purified plasmid. Each reaction was centrifuged briefly. The PCR program used for the reaction was 96°C for 1 min, 25 circles of [96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes] then 60°C for 4 minutes and finally the reaction was kept at 4°C.

The amplified inserts were precipitated by adding 125 mM EDTA to the reaction products and moving each of them to a 1,7 ml Eppendorf tube. 96% ethanol was added and incubated at RT for 15 minutes. The reaction was then centrifuged at 3000g for 30 minutes at 4°C. The supernatant was discarded and the pellet washed with 70% ethanol and centrifuged at 1700g for 15 minutes. The supernatant was discarded and the pellet incubated at 37°C for 10-15 minutes for drying. The dry pellets were kept in a cooler overnight.

For sequencing, the pellets were dissolved in 15 µl formamide. The resuspension was moved to a sequencing plate and the plate centrifuged at 1000g for 10 seconds at 6°C. The sequencing plate was placed in the Genetic Analyzer and sequenced.

3.7 PCR with testing primers

To confirm that a transcript variant exists that includes an intron, and that it is not a result of DNA contamination, a PCR was done on DNase treated cDNA from 501mel cells and Skmel28 cells. Primer pairs, spanning the intron-exon junction and some known exon-exon junctions, were designed for this.

3.7.1 Primer designed

Testing primers were designed as seen in figure 5 and table 3.

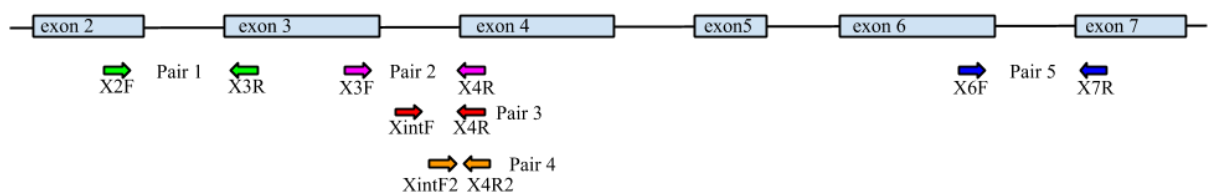


Figure 5: Location of the testing primer pairs. *The primers were designed to amplify a sequence from one exon to the next to ensure the specific amplification of mRNA derived cDNA. Primer pair 1 covered a sequence from TFEB exon 2 to exon 3. Primer pair 2 covered a sequence from exon 3 to exon 4. Primer pairs 3 and 4 covered a sequence from the intron between exon 3 and 4, with two product lengths. Primer pair 5 covered a sequence from exon 6 to exon 7.*

Table 3: *Names, sequences, T_m and product length of testing primers*

Primer pair number	Primer pair covering exons	Name of primer	Sequences [5' -> 3']	T _m [°C]	Product length [bp]
1	2 -> 3	X2F	CCATCAATACCCCCGTCCAC	60,18	252
		X3R	CAGCGGAGGAGGACAGCA	61,39	
2	3 -> 4	X3F	GGAGTACCTGTCCGAGACCT	60,03	231
		X4R	ACATCGTCCAGACGCATAAT	57,11	
3	Intron -> 4	XintF	CCAAGCCAGAGACAGAAGTGA	59,65	295
		X4R	ACATCGTCCAGACGCATAAT	57,11	
4	Intron -> 4	XintF2	CCTCCCCAGATGGACTTTTC	57,57	85
		X4R2	TGATGTAGCCAAGGACATCGT	59,17	
5	6 -> 7	X6F	AAGGAGCGGCAGAAGAAAGA	59,31	81
		X7R	CCAACCTCCTTGATGCGGTCA	60,32	

The primer pairs covering the TFEB intron-exon 4 junction will only give a product if the intron exists in a TFEB transcript in these cell lines. The other primer pairs are used as controls, as exons 2, 3, 4, 6 and 7 are used in most transcript variants.

3.7.2 PCR on purified mRNA

The cDNA made from DNase treated mRNA was used for the amplification. Five reactions were prepared in PCR strips for each cell line and two extra for the –RT controls. For each reaction 12,6 µl Nuclease free water from Ambion, 2 µl 10X PCR buffer, 2 µl MgCl₂ (25 mM), 2 µl cDNA, 0,4 µl dNTP mix (10 mM), 0,4 µl of a forward primer (10 mM) and 0,4 µl of a reverse primer (10 mM) (see table 3 for each primer pair) were . For the –RT control, the –RT mRNA was used instead of cDNA. Finally 0,2 µl HotStart Taq polymerase from Qiagen was added and the reactions vortexed and centrifuged briefly. The PCR program used for the reactions was: 95°C for 15 minutes, 31 PCR cycles of [94°C for 30 seconds, annealing temperature for 30 seconds, 72°C for 1 minute], 72°C for 7 minutes and 4°C until reactions were collected. The annealing temperature started at 57°C but after circle 2 it was raised by 0,2°C after each circle, the final circle ending in 62°C.

3.7.3 Gel Electrophoresis on testing primer products

To analyse the products from each PCR, 2% agarose gel was used. To make the agarose gel, 2g agarose was mixed with 100 ml of TAE buffer and heated in a microwave until the agarose had dissolved completely. We then added 5 µl ethidium bromide and the mixture was poured into a mold with a comb and left for cooling for about half an hour. The comb was removed

and the gel placed in an electrophoresis apparatus filled with TAE buffer with the same concentration of ethidium bromide. Each PCR reaction sample was mixed with 4 μ l 6X loading dye from Fermentas. Each dyed sample was then loaded on the agarose gel, as well as a GeneRuler DNA ladder mix and a GeneRuler Ultra Low Range DNA ladder from Thermo Scientific. The samples were run at 80 V for 60 minutes and a picture under UV light taken.

4 Results

4.1 Many TFEB transcript variants exist in melanoma

Data retrieved from Ensembl and UCSC revealed many different transcript variants of TFEB, almost all of which have different 5'ends. All the transcript variants but one, found in these databases, included the cytoplasmic localization signal as can be seen in Figures 6 (Ensembl) and 7 (UCSC).

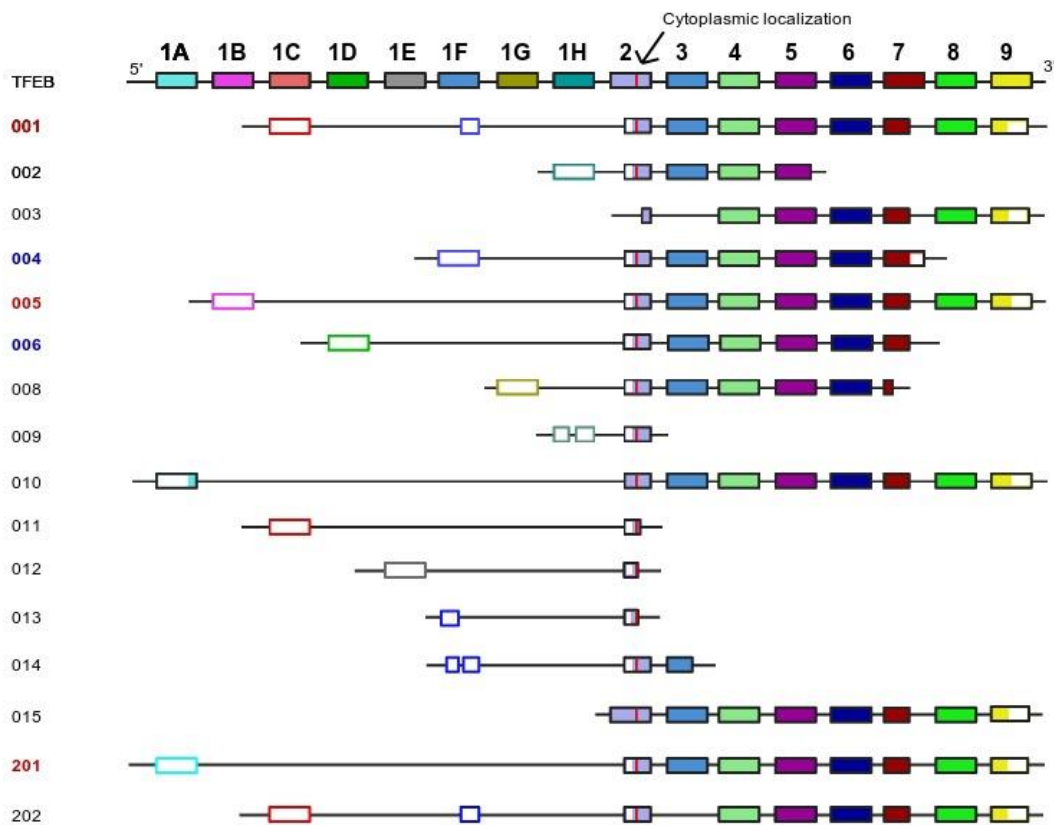


Figure 6: All transcript variants found on Ensembl. The first lane shows all the TFEB exons and their numbers. The different transcript variants are numbered on the left side of the picture, the red numbers show the transcript variants that are translated into isoform 1, the blue numbers show the transcript variants that are translated into isoform 2. Boxes filled with colour are translated, the empty boxes are untranslated. The red stripe in exon 2 shows the location of the S3R4 and Q10L11 amino acids, or the cytoplasmic localization signal.

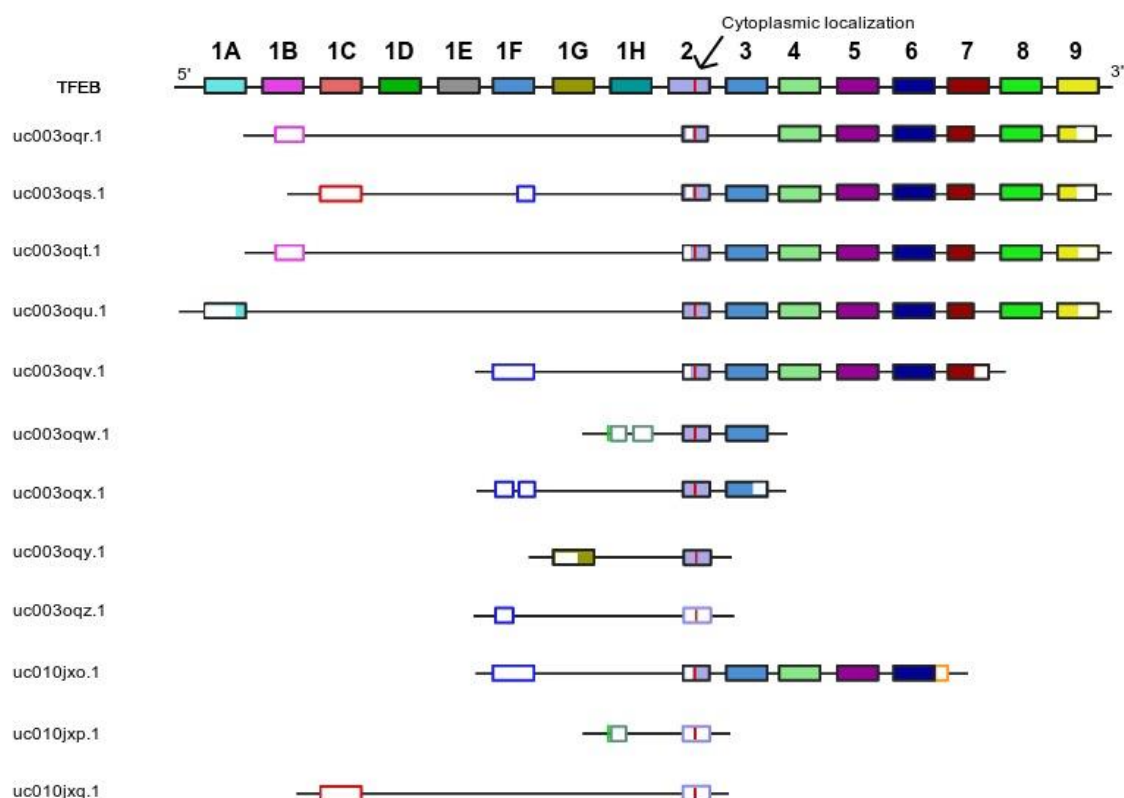


Figure 7: All transcript variants found in melanoma samples from the Cancer Genome Atlas. The top line shows all the TFEB exons and their numbers. The different transcript variants are numbered on the left side of the picture. Boxes filled with colour are translated, the empty boxes are untranslated. Transcript variants number 2-5 are also found in the Ensembl database, the other transcripts are not.

Three of the transcript variants found at the Ensembl website are translated into TFEB isoform 1, the 476 amino acid version, and two of the variants are translated into TFEB isoform 2, the 391 amino acid version, missing a part of the carboxyl end found in isoform 1. Thus it is clear that many different transcript variants and isoforms of TFEB are found, both in normal cells and melanoma.

4.2 Many 5'ends of TFEB discovered

In order to identify different isoforms of TFEB in melanoma cells, a 5'RACE procedure was performed on purified RNA from both 501Mel cell line and SKMel28 cell line. Three gene specific primers (GSPs) for TFEB were used for the procedure. To optimize the PCR for best analysis of the different 5'ends of TFEB, different salt concentration, different annealing temperature and different total PCR time were used and the results read from agarose gels.

The first PCR on the 5'ends using GSP2 resulted in one vague band at around 700 bp for the 501mel sample but nothing for the Skmel28 sample or the control sample

(Supplementary figure 1). After some adjustments to the PCR protocol, using different annealing temperatures, salt concentration and total PCR time than the original protocol recommended we got the right band for the control sample (not shown). A second PCR amplification using the nested GSP3 was done and clearer results were obtained from the cell samples. The PCR results show a band at just below 200 bp for both cell lines, two bands at around 300 bp and a ladder of bands at 400 bp and above, as can be seen in Figure 8. This indicates many different 5'ends of TFEB existing in the melanoma cell lines.

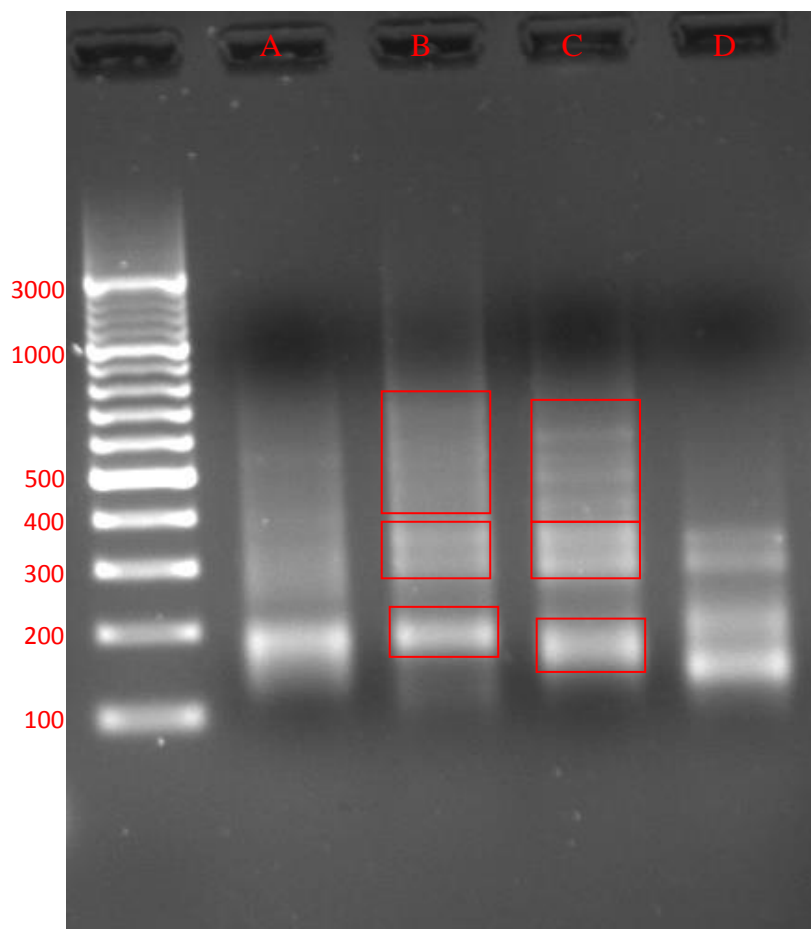


Figure 8: The final results of the 5'RACE. The nested GSP3 primer was used for a second amplification of the first products from the 5'RACE. The first well to the left has a O'RangeRuler 100 bp DNA ladder, A) has 501mel 5'RACE products from the reaction using 57°C as annealing temperature, B) has 501mel 5'RACE products from the reaction using 59°C as annealing temperature, C) has Skmel28 5'RACE products from the reaction using 57°C as annealing temperature and D) has Skmel28 5'RACE products from the reaction using 59°C as annealing temperature. The red boxes show how the bands were excised from the gel for isolation.

4.3 Sequencing revealed the existence of many transcript variants

The PCR products were cut from the agarose gel and purified. The gel in Figure 8 represents the best separation of bands that was obtained. Because of how many different bands were present on the gel, the area around the 200 bp band was cut out of the gel, a portion of the gel containing the 300-400 bp bands was cut out in a piece and the ladder of bands above 400 bp were cut together in one piece. This was done for both cell lines. The gel purification had very low efficiency, only around 10 ng/μl were retrieved for each gel piece. As the GSP3 had a restriction site for SacII and the UAP primer from the 5'RACE kit had a restriction site for SalI, these two enzymes were used to digest the isolated DNA fragments as well as a pBluescript vector, and the fragments ligated into the vector with T4 DNA ligase. These vectors were then transformed into high competent *E.coli* cells for amplification. The vectors were isolated from the *E.coli* cells and the inserts sequenced. The sequencing results were aligned against the human genome using BLAT to find out which parts the sequences covered.

The sequencing results showed many different TFEB transcript variants for both cell lines (Figure 9 and 10). Nine of the sequenced clones from the 501mel cells covered parts of known TFEB exons, mostly exons 2 and 3, but one had part of the intron between exons 3 and 4. For the Skmel28 cells, nine of the sequenced clones covered parts of known exons, very similar to what is seen in the 501mel cells, but 3 of the sequences had part of the intron between exons 3 and 4. Interestingly, none of the sequenced fragments covered the cytoplasmic localization site of the protein.

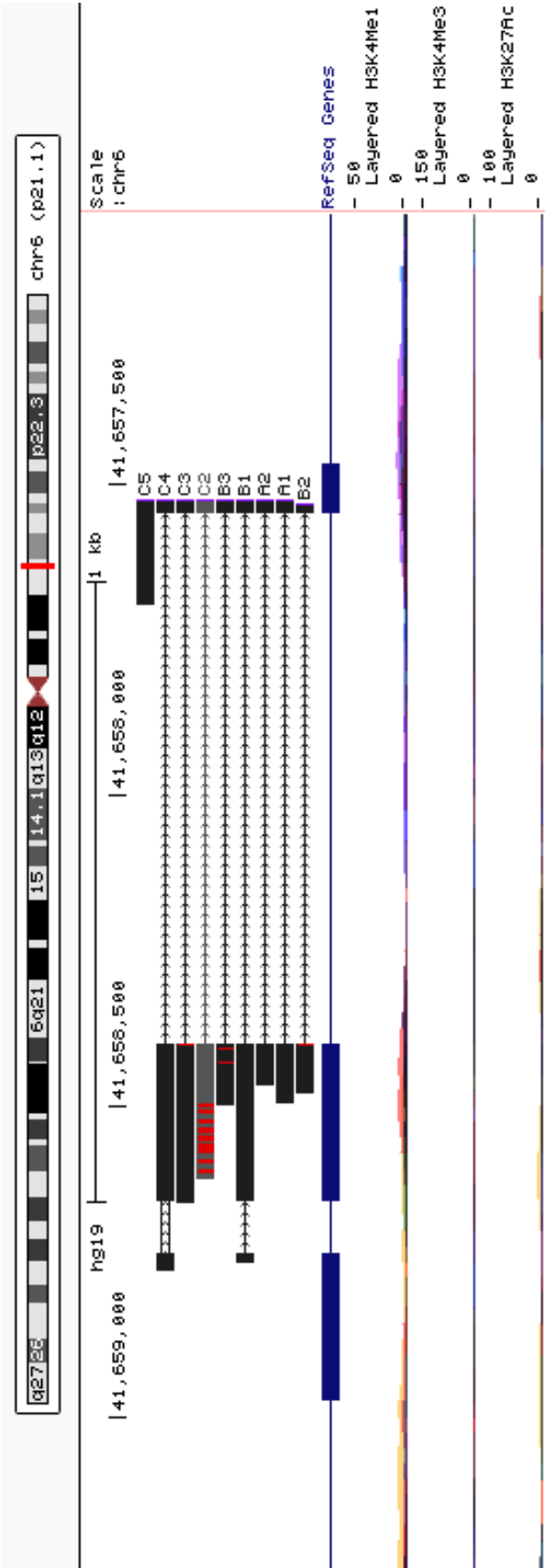


Figure 9: The 5'ends of the transcript variants in 501mel cells blated against the human genome.
The picture shows the location of the TFEB gene at the p arm, location 21.1 on chromosome 6. The RefSeq gene shows the reference gene of TFEB and above that all the transcript variants sequenced from 501 mel cells. Their names are indicated on the right site of each transcript. The exon farthest to the right is exon 4 and exons 3 and 2 can be seen to the left. The three tracks at the bottom show H3K4Me1 marks, often found near regulatory elements, H2K4M3 marks, often found near promoters and H3K27Ac marks, often found near active regulatory elements.

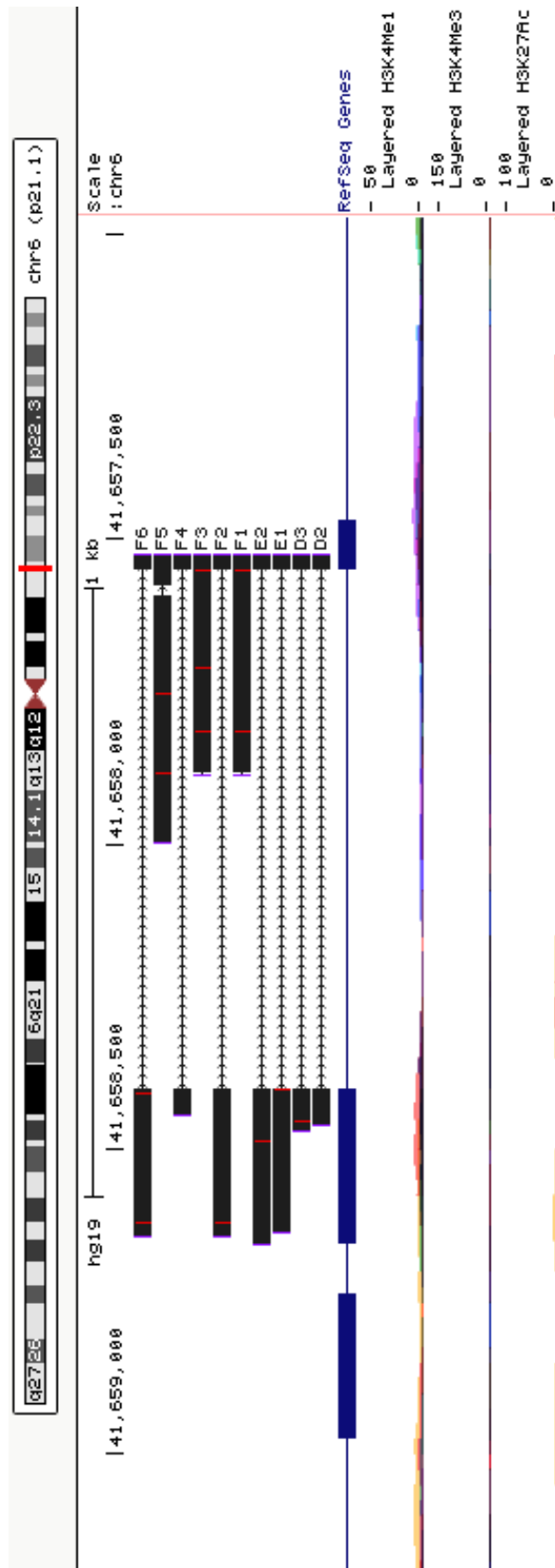


Figure 10: The 5'ends of the transcript variants in Skmel28 cells blated against the human genome.
The top of the picture shows the location of the TFEB gene at the p arm, location 21.1 on chromosome 6. The RefSeq gene shows the reference gene of TFEB and above that all the transcript variants sequenced from 501 mel cells. Their names are indicated on the right site of each transcript. The exon farthest to the right is exon 4 and exon 3 and 2 can be seen to the left. The three tracks at the bottom show H3K4Me1 marks, often found near regulatory elements, H2K4M3 marks, often found near promoters and H3K27Ac marks, often found near active regulatory elements.

4.4 The identified intronic sequences all have a translation start site in-frame with exon 4

To see if the transcripts variant starting in the intron between exon 3 and 4 had a possible translation start site in-frame with exon 4, the sequences including the intron were translated in all three possible forward reading frames. To find out a possible translation start site in the reading frame that was in-frame with the translation of exon 4, we looked for an ATG site. All of the transcripts had one or more ATG sites in frame with exon 4 but none of them had a full KOZAC sequence (5'A/GNNAUGG 3'). When the intron was examined using UCSC Genome Browser, no predicted enhancer or promoter markers were found (Figures 9 and 10). When the intronic sequence of the F5 clone was run in the NCBI Conserved Domain Search a GVQW domain was found. This is a highly conserved GVQW domain, rich in proline and cysteine. The domains function is not known but might be an interacting domain or a binding domain.

4.5 Confirmation of the presence of intronic sequences

To make sure that the new transcript variants starting within an intron were not a result of DNA contamination or half spliced transcripts, total RNA was isolated again from 501mel cells and Skmel28 cells. The total RNA was treated with DNase and then purified. The mRNA was then isolated specifically. The final concentration of mRNA was very low, only around 12 ng/μl, which is normal as mRNA is only 1-5% of the cell's total RNA. The mRNA was reverse transcribed to cDNA and PCR was done on the cDNA from each cell line using 5 different primer sets to test if the intron between exon 3 and 4 is truly a part of a transcript (Figure 5 and Table 3).

The results from the PCR are shown in Figure 11. It is clear that all the primer sets gave a band of the expected size and the –RT control shows no DNA contamination. Thus, part of the intron exists in some of the TFEB transcripts and their transcription might possibly start there. This suggests that a novel transcript of TFEB might exist which starts in the intron and leads to a functional protein lacking the cytoplasmic localization signal.

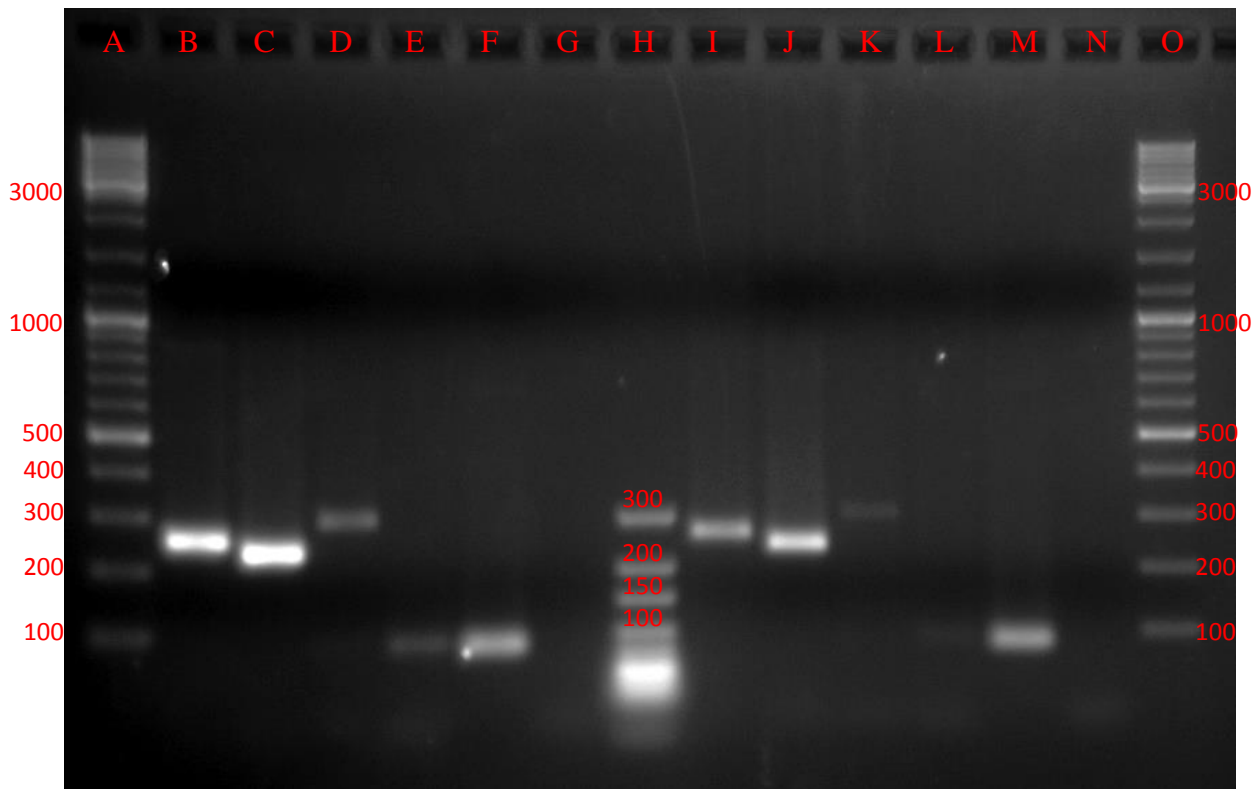


Figure 11: Gel electrophoresis of the PCR products of DNase treated RNA. In well A) GeneRuler DNA ladder mix, B) 501mel PCR product from exon 2->3, expected band size of 252 bp, C) 501mel PCR product from exon 3->4, expected band size of 231 bp, D) 501mel PCR product from intron ->exon 4, expected band size of 295 bp, E) 501mel PCR product from intron -> exon 4, expected band size of 85 bp, F) 501mel PCR product from exon 6->7, expected band size of 81 bp, G) -RT control for 501mel samples, H) GeneRuler Ultra low range DNA ladder, I) Skmel28 PCR product from exon 2->3, expected band size of 252 bp J) Skmel28 PCR product from exon 3->4, expected band size of 231 bp K) Skmel28 PCR product from intron ->exon 4, expected band size of 295 bp, L) Skmel28 PCR product from intron -> exon 4, expected band size of 85 bp, M) Skmel28 PCR product from exon 6->7, expected band size of 81 bp, N) -RT control for 501mel samples, O) GeneRuler DNA ladder mix. Clear bands can be seen in well D, E and K and vaguely in well L, indicating that the transcript variant starting in an intron exists in both cell lines.

5 Discussions

This study shows that many isoforms of TFEB exist in human melanoma cells. The N-terminus of the different isoforms varies and this study suggests that more transcript variants exist than already published in open access databanks. TFEB is located in the cytoplasm under normal conditions but under starvation or stress conditions it is translocated to the nucleus. When the cellular conditions are back to normal the cytoplasmic recruitment of TFEB is important and is regulated by the interaction between Rag GTPases and the N-terminus of TFEB. In a recent study in Steingrímsson's lab, TFEB staining using mouse monoclonal antibody (MyBiosource) showed TFEB mostly nuclear in 501mel cells and ASPC1 cells (a pancreatic cell line), a similar result as has already been shown for MITF-M in melanoma .

Many different TFEB transcript variants exist in both normal cells and melanoma cells and most of the published TFEB transcripts contain the cytoplasmic localization signal at the N-terminus. In this study many 5'ends of TFEB were found in both 501mel and Skmel28 cells, all different from the transcripts found in Ensembl and UCSC databases. A possible new start-of-translation site was discovered, in both 501mel and Skmel28 cell lines, starting in an intron. These results suggest that TFEB exists in two forms like MITF, one that can be held in the cytoplasm and one that is lacking the N-terminus and stays nuclear. This might explain the mostly nuclear localization of TFEB in melanoma cells, seen with the mouse antibody, as the transcripts found in this study all lack the N-terminal region that includes the cytoplasmic localization signal.

There is a possibility that some TFEB transcripts may have been lost during the extraction procedure. A different method or kit might be used for the DNA extraction from the gel for better recovery if the experiment were to be repeated. For further research it would be an interesting approach to quantify the different transcript variants found in both cell lines used in this study, using Absolute Quantification qPCR.

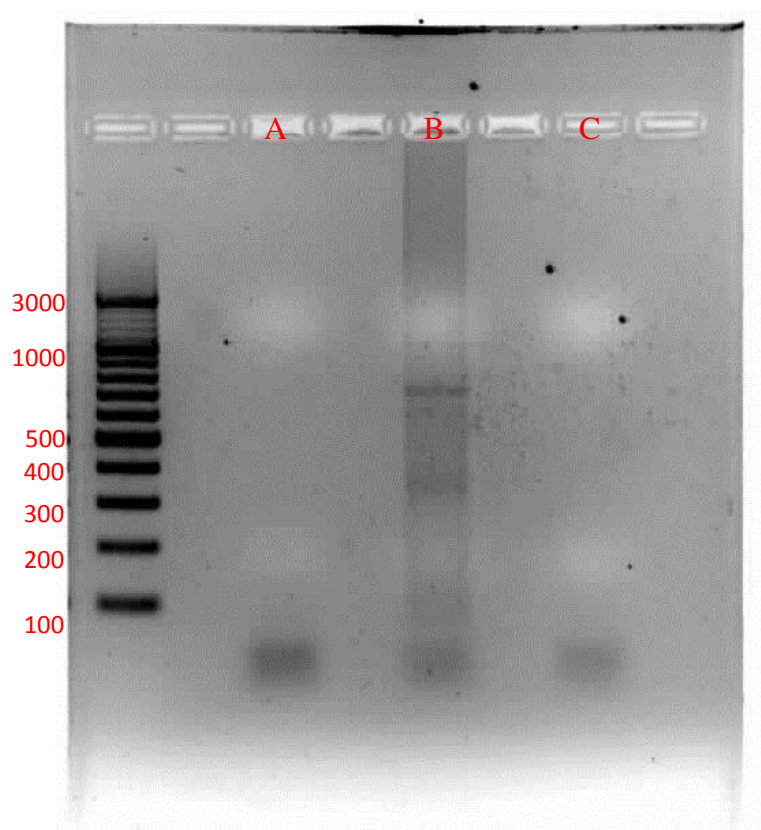
It is important to map and determine the control of TFEB and MITF on autophagy as the autophagic control of these proteins might be different in melanoma than in normal cells. Results from other studies done in Steingrímsson's lab show interactions between MITF and TFEB and that both proteins take part in regulating mTORC1 and autophagy. This study of the 5'ends of TFEB is thus an important contribution to revealing the role of TFEB in autophagic control in melanoma.

References

- Cheli, Y., Ohanna, M., Ballotti, R., and Bertolotto, C. (2010). Fifteen-year quest for microphthalmia-associated transcription factor target genes. *Pigment Cell Melanoma Res.* 23, 27–40.
- Chen, Y., and Klionsky, D.J. (2011). The regulation of autophagy - unanswered questions. *J. Cell Sci.* 124, 161–170.
- Garraway, L.A., and Sellers, W.R. (2006). Lineage dependency and lineage-survival oncogenes in human cancer. *Nat. Rev. Cancer* 6, 593–602.
- Grill, C., Steingrímsson, E., Bergsteinsdóttir, K., Ögmundsdóttir, M.H., Pogenberg, V., Schepsky, A., Wilmanns, M., and Pingault, V. (2013). MITF mutations associated with pigment deficiency syndromes and melanoma have different effects on protein function. *Hum. Mol. Genet.* 22, 1–11.
- Halaban, R., Cheng, E., Zhang, Y., Moellmann, G., Hanlon, D., Michalak, M., Setaluri, V., and Hebert, D.N. (1997). Aberrant retention of tyrosinase in the endoplasmic reticulum mediates accelerated degradation of the enzyme and contributes to the dedifferentiated phenotype of amelanotic melanoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 94, 6210–6215.
- King, R., Weilbaecher, K.N., McGill, G., Cooley, E., Mihm, M., and Fisher, D.E. (1999). Microphthalmia Transcription Factor A Sensitive and Specific Melanocyte Marker for Melanoma Diagnosis. *Am. J. Pathol.* 155, 731–738.
- Martina, J. a, and Puertollano, R. (2013). Rag GTPases mediate amino acid-dependent recruitment of TFEB and MITF to lysosomes. *J. Cell Biol.* 200, 475–491.
- Palmieri, M., Impey, S., Kang, H., di Ronza, A., Pelz, C., Sardiello, M., and Ballabio, A. (2011). Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways. *Hum. Mol. Genet.* 20, 3852–3866.
- Rehli, M., Elzen, N. Den, Cassady, A., Ostrowski, M.C., and Hume, D.A. (1999). Cloning and characterization of the murine genes for bHLH-ZIP transcription factors TFEC and TFEB reveal a common gene organization for all MiT subfamily members. *Genomics* 03, 111–120.
- Roczniak-ferguson, A., Petit, C.S., Froehlich, F., Qian, S., Ky, J., Angarola, B., Walther, T.C., and Ferguson, S.M. (2012). The Transcription Factor TFEB Links mTORC1 Signaling to Transcriptional Control of Lysosome Homeostasis. *Sci. Signal.* 5.

- Sardiello, M., Palmieri, M., di Ronza, A., Medina, D.L., Valenza, M., Gennarino, V.A., Di Malta, C., Donaudy, F., Embrione, V., Polishchuk, R.S., et al. (2009). A gene network regulating lysosomal biogenesis and function. *Science* 325, 473–477.
- Settembre, C., and Ballabio, A. (2011). TFEB regulates autophagy. *Autophagy* 7, 1379–1381.
- Settembre, C., Zoncu, R., Medina, D.L., Vetrini, F., Erdin, S., Erdin, S., Huynh, T., Ferron, M., Karsenty, G., Vellard, M.C., et al. (2012). A lysosome-to-nucleus signalling mechanism senses and regulates the lysosome via mTOR and TFEB. *EMBO J.* 31, 1095–1108.
- Settembre, C., Fraldi, A., Medina, D.L., and Ballabio, A. (2013a). Signals from the lysosome: a control centre for cellular clearance and energy metabolism. *Nat. Rev. Mol. Cell Biol.* 14, 283–296.
- Settembre, C., De Cegli, R., Mansueto, G., Saha, P.K., Vetrini, F., Visvikis, O., Huynh, T., Carissimo, A., Palmer, D., Klisch, T.J., et al. (2013b). TFEB controls cellular lipid metabolism through a starvation-induced autoregulatory loop. *Nat. Cell Biol.* 15, 647–658.
- Steingrímsson, E., Copeland, N.G., and Jenkins, N. a (2004). Melanocytes and the microphthalmia transcription factor network. *Annu. Rev. Genet.* 38, 365–411.
- White, E. (2012). Deconvoluting the context-dependent role for autophagy in cancer. *Nat. Rev. Cancer* 12, 401–410.
- Yasumoto, K., Amae, S., Udon, T., Fuse, N., Takeda, K., and Shibahara, S. (1998). A big gene linked to small eyes encodes multiple Mitf isoforms: many promoters make light work. *Pigment Cell Res.* 11, 329–336.
- Zoncu, R., Bar-Peled, L., Efeyan, A., Wang, S., Sancak, Y., and Sabatini, D.M. (2011). mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the Vacuolar H⁺-ATPase. *Science* (80-.). 334, 678–683.

Supplementary figures



Supplementary Figure 1: The first PCR on the 5'RACE products. *In well A) 5'RACE products from Skmel28 cells, B) 5'RACE products from 501mel cells and C) 5'RACE products from control sample. No band appeared for the control sample so adjustments were required for the PCR procedure.*

Appendix

Composition of buffers and solutions

The buffers that are not listed were provided with the appropriate kits and did not have information about their composition.

Oligotex mRNA Mini kit from Qiagen:

OEB buffer	5 mM Tris·Cl, pH 7.5
OBB buffer	20 mM Tris·Cl, pH 7.5, 1 M NaCl, 2 mM EDTA, 0.2% SDS
OW2 buffer	10 mM Tris·Cl, pH 7.5, 150 mM NaCl, 1 mM EDTA
Oligotex suspension	10% (w/v) suspension (= 1 mg/10 µl) Oligotex particles in 10 mM Tris·Cl, pH 7.5, 500 mM NaCl, 1 mM EDTA, 0.1% SDS and 0.1% NaN ₃ .

5'RACE System for Rapid Amplification of cDNA Ends, version 2.0 from Life Technologies:

10X PCR buffer	200 mM Tris-HCl (pH 8.4), 500 mM KCl
5X tailing buffer	50 mM Tris-HCl (pH 8.4), 125 mM KCl, 7.5 mM MgCl ₂
Binding solution	6 M sodium iodide

Restriction enzyme buffer from New England Biolabs:

NEB 2.1	50mM NaCl, 10mM Tris-HCl, 10mM MgCl ₂ , 100µg/ml BSA, pH 7.9 at 25°C
---------	---

Ligation buffer from Thermo Scientific:

10X T4 DNA ligation buffer	400 mM Tris-HCl, 100 mM MgCl ₂ , 100 mM DTT, 5 mM ATP, pH 7.8 at 25°C
----------------------------	--