Published in final edited form as:
Gene. 2015 January 25; 555(2): 194–202. doi:10.1016/j.gene.2014.11.003.

PAX5 IS THE TRANSCRIPTIONAL ACTIVATOR OF MUCOLIPIN-2 (*MCOLN2*) GENE

Jessica A. Valadez[1] and **Math P. Cuajungco**[1,2]

[1]Department of Biological Science, and Center for Applied Biotechnology Studies, California State University Fullerton, CA, USA 92831

[2]Mental Health Research Institute, Melbourne Brain Centre, Parkville, Victoria, Australia 3052

Abstract

Transient Receptor Potential Mucolipin (TRPML) proteins belong to the TRP superfamily of non-selective cation channels. The TRPML1, -2, and -3 proteins are encoded by *Mucolipin (MCOLN)-1*, *-2* and *-3* genes, respectively. TRPML1 has been associated with Mucolipidosis type IV (MLIV), while no disease phenotype has been linked with TRPML2 or -3 protein. The TRPML proteins share high sequence similarities, form hetero-tetramers, and serve in membrane trafficking, autophagy, and metal homeostasis. Previous studies suggest that TRPML2 serves a role in the immune system; however, the evidence is mostly indirect. We hypothesize that if TRPML2 is involved in immune function its expression would be likely regulated by an immune-associated transcription factor protein. Thus, we set out to identify the core promoter region and the transcription factor responsible for *MCOLN2* gene expression. Using dual-luciferase assay and over-expression analyses, we reveal for the first time that B-cell lineage specific activator protein (BSAP), also known as paired box 5 (PAX5), controls *MCOLN2* expression. Specifically, heterologous expression of PAX5 in HEK-293 cells significantly increased endogenous *MCOLN2* transcript and TRPML2 protein levels, while RNA interference targeting endogenous PAX5 reduced its effect. Site-directed mutagenesis studies showed that the core promoter and PAX5 binding region to be between -79 and -60 base pairs upstream of the transcriptional start site. Thus, our findings add to a growing list of evidence for TRPML2's possible involvement in the immune system. The knowledge gained from this study could be used to further characterize the role of TRPML2 in B-cell development and function.

Keywords

BSAP; Mucolipidosis IV; endosomes; lysosomes; B-cell; immune system; TRPML2

Corresponding Author: Math P. Cuajungco, PhD, Dept. of Biological Science, California State University Fullerton, 800 N. State College Blvd., Fullerton, CA 92831 USA, Tel: + 1-657-278-8522, Fax: + 1-657-278-3426, mcuajungco@fullerton.edu.

The authors declare no conflict of interest.

INTRODUCTION

Transient Receptor Potential Mucolipin (TRPML) proteins belong to the TRP superfamily of non-selective cation channels. TRPMLs consist TRPML1, -2, and -3 proteins that are encoded by *Mucolipin (MCOLN)-1*, *-2* and *-3* genes, respectively. They contain a highly distinctive, large extracellular loop, and share sequence homology within the transmembrane (TM) domains of group 1 TRPs (Montell, 2005). TRPML1 was first identified by virtue of its causative effect on a neurodegenerative disease known as Mucolipidosis type IV (MLIV) (Bargal et al., 2000; Bassi et al., 2000; Sun et al., 2000). Positional cloning analysis subsequently identified the TRPML2 and TRPML3 proteins, which are both located in chromosome 1, and their mouse counterpart on chromosome 3 (Di Palma et al., 2002). Di Palma et al. (2002) reported that a point mutation in the *MCOLN3* gene is responsible for the auditory and vestibular deficits in the varitint waddler mouse. Interestingly, no human disease phenotype has been associated with *MCOLN2* or *-3* genes. Tissue distribution analyses show that TRPML1 is ubiquitously expressed in human or rodent tissues, but *MCOLN2* or *-3* transcripts have tissue-specific expression (Samie et al., 2009; Grimm et al., 2010). One interesting feature of TRPML proteins is their ability to interact with each other, and form hetero-tetramers (Venkatachalam et al., 2006; Curcio-Morelli et al., 2010; Zeevi et al., 2010). Furthermore, a closer inspection of their putative channel pore between TM5 and TM6 regions reveals 89% amino acid similarity (Zeevi et al., 2010), but that TRPML2 has a more similar channel characteristic with TRPML1 than TRPML3 (Samie et al., 2009). At the cellular level, the distribution of all three TRPML proteins shows mostly lysosomal localization. However, TRPML2 has been detected in long tubular recycling endosomes and late endosomes, while TRPML3 has been detected on the plasma membrane and late endosomes (Venkatachalam et al., 2006; Grimm et al., 2007; Karacsonyi et al., 2007; Grimm et al., 2010; Flores and Garcia-Anoveros, 2011). Despite recent advances in the field, the cellular function of the TRPML proteins remains obscure. Nevertheless, there are indications that these proteins are involved in endosomal-lysosomal biogenesis, vesicle fusion, membrane trafficking, organelle acidification, autophagy, exocytosis, and metal homeostasis (LaPlante et al., 2006; Song et al., 2006; Karacsonyi et al., 2007; Dong et al., 2008; Miedel et al., 2008; Vergarajauregui et al., 2008; Martina et al., 2009; Eichelsdoerfer et al., 2010; Samie et al., 2013; Cuajungco et al., 2014).

We are interested in TRPML2 because it has the potential to substitute for the missing function of TRPML1 in MLIV, and that it appears to play an important role in the immune system. Indeed, its subfamily member, TRPML1, has also been implicated in immune cell function, particularly in major histocompatibility complex (MHC)-II processing and trafficking in the lysosomes of macrophages (Thompson et al., 2007) and for toll-like receptor type 7-mediated nucleic acid (single-stranded RNA) processing and trafficking in the lysosomes of dendritic cells (Li et al., 2014). For TRPML2, we previously observed high tissue-specific expression levels of mouse *Mcoln2* and human *MCOLN2* transcripts in thymus and spleen, where B-cell lymphocytes mature (Samie et al., 2009; Grimm et al., 2010). Likewise, *MCOLN2* transcripts are detected in different stages of B-cell maturation, demonstrating a potential role in B-cell lymphocyte development and function (Lindvall et al., 2005). Note that stimulation of primary splenic B-cells with immunoglobulin or phorbol

12-myristate 13-acetate plus ionomycin (which activates protein kinase C) has been shown to up-regulate *MCOLN2* transcript levels (Lindvall et al., 2005). A similar outcome is observed in B-cells with dysfunctional Bruton's tyrosine kinase (BTK) protein; however, the amount of detectable *MCOLN2* mRNA is markedly lower in the BTK-deficient B-cells compared to wild-type B-cells. This result suggests that TRPML2 protein expression might be partly dependent on the BTK pathway. TRPML2 is also detected in "specialized" lysosomal compartments of chicken DT-40 B-cell lymphocyte lines (Song et al., 2006), which further signifies a role for this protein in antigen processing upon receptor activation and endocytosis. Interestingly, TRPML2 co-localizes with MHC-I and cluster of differentiation (CD)-59 within long tubular recycling endosomes (Karacsonyi et al., 2007), but it is not known if TRPML2 physically interacts with CD59 or MHC-I. It was also suggested that TRPML2 mediates membrane sorting through the ADP ribosylation factor 6 (ARF6) pathway (Karacsonyi et al., 2007) – a regulator of clathrin-independent endocytic membrane trafficking and actin remodeling in cells (Donaldson, 2003). This supposition is due to the observation that heterologously overexpressed TRPML2 in HeLa cells increases ARF6 protein activity by sevenfold, which subsequently increases GPI-AP (*i.e.* CD59) internalization, vesicle fusion, and formation of large vacuoles (Karacsonyi et al., 2007). While these evidences appear indirect for a role of TRPML2 in the immune system, these published observations are just too many to be considered insignificant. We thus hypothesize that if TRPML2 is involved in immune function its expression would be likely controlled by an immune-regulated or immune-associated transcription factor protein as well. We set out to determine the identity of the transcription factor and core promoter sequence involved in TRPML2 protein expression.

In the present study, we show for the first time that B-cell lineage specific activator protein (BSAP), also known as paired box 5 (PAX5), is the transcription factor that controls the expression of *MCOLN2* gene. We also mapped and identified the core promoter sequence of the *MCOLN2* gene between −79 and −60 base pairs upstream of the transcriptional start site (TSS) where a consensus PAX5-binding sequence is situated. Therefore, these observations provide additional supporting evidence that TRPML2 very likely plays a role in immune cell development and function. The knowledge gained from this study could be instrumental to further characterize the cellular function of TRPML2 and is the first step to investigate the possibility of using TRPML2 as a potential therapy to compensate for the loss of TRPML1 function in MLIV disease.

MATERIALS AND METHODS

Cell culture

Human embryonic kidney (HEK)-293 and HGT-1 gastric adenocarcinoma cells were grown and maintained in a standard humidified 37°C incubator, with 5% CO_2. The cell lines were cultured in Dulbecco's Modification of Eagle's Medium with 4.5 g/L glucose, L-glutamine, and sodium pyruvate (Mediatech, Manassas VA). The medium was supplemented with 10% fetal bovine serum (FBS; Life Technologies, Carlsbad CA) and 1% penicillin-streptomycin solution (Mediatech).

Standard and Real-time Quantitative PCR (QPCR)

Both standard and real-time quantitative PCR (QPCR) reactions were performed as previously described (Cuajungco et al., 2012) with minor modifications appropriate for this study. The primer sets used for each specific experiment are outlined in Table 1. For the real-time QPCR, we followed the *Minimum Information for Publication of Quantitative Real-Time PCR Experiments* (MIQE) guidelines (Bustin et al., 2009). All treatment samples were run in triplicates using the Sensimix™ Plus SYBR supermix (Bioline USA, Taunton MA). We used human *18S* ribosomal RNA (rRNA) as the reference gene to normalize the quantitative cycle (Cq) values obtained from each experiment due to its robust expression across many tissues and cell lines (Cuajungco et al., 2003). We used the CFX96 thermocycler (Bio-Rad, Hercules, CA) with the following thermocycler parameters: 10 min at 95 °C, and 40 cycles for 15 sec at 95 °C, 30 sec at 56 °C, and 30 sec at 72 °C. All QPCR reactions were done with at least three independent trials, and included a standard curve using pooled cDNA samples. The QPCR trials had an average correlation coefficient (R^2) value (mean ± SEM): $R^2 = 0.993 \pm 0.001$. The QPCR efficiencies for each primer set were (mean ± SEM): *MCOLN2* = 102.5 ± 2.1 %, *PAX5* = 99.3 ± 2.9 %, and *18S* rRNA = 97.6 ± 3.5 %. The QPCR data were analyzed using the Livak method ($2^{-\Delta\Delta Cq}$) and represented as normalized relative expression levels calibrated to untreated control (value = 1). We used the GraphPad Prism software to test for statistical significance by running an analysis of variance (ANOVA) followed by Bonferroni's multiple comparisons *post-hoc* test. The significance level was set at p-value < 0.05.

Dual Luciferase Reporter (DLR) Constructs and Experimental Assay

We engineered a dual luciferase reporter (DLR) pGL4.10 *Luc2-hRluc-Neo* vector by using a pGL4.10 *Luc2* vector (Promega, Madison WI) and inserting a constitutively active *Renilla-Neomycin* gene fusion (*hRluc-Neo*) containing both an SV40 early enhancer/promoter and a synthetic poly-A signal. This was done by PCR-amplifying the *hRluc-Neo* fragment using the pmiR Glo vector (Promega) as a template and ligating the product downstream of the *Luc2* gene (using the *BamH I* restriction enzyme site). The engineered pGL4.10 *Luc2-hRluc-Neo* vector was used in DLR assay using the Dual-Glo reagent (Promega). The engineered DLR construct allowed us to normalize the luciferase activity without the potentially confounding effects or issues of transfection efficiency in double-plasmid delivery into the cells.

We used the University of California Santa Cruz (UCSC) Genome Browser to identify the putative promoter region of the *MCOLN2* gene. We initially PCR cloned a 1,583 base pair (bp)-long genomic DNA region into the pGL4.10 *Luc2-hRluc-Neo* DLR vector using the 5' and 3' *Sfi I* restriction sites on the vector (see Table 1 for primer information). We then produced several DLR constructs containing shorter genomic DNA fragments using PCR deletion mutagenesis. We sent all DLR constructs for sequencing (Retrogen, San Diego CA) to verify sequence integrity before using them in the luciferase study.

We used HEK-293 in all DLR assays, because our previous study showed that *MCOLN2* is present and constitutively expressed in these cells (Samie et al., 2009; Grimm et al., 2010), which precludes the use of an exogenous transcriptional activator to detect promoter

construct activity in these cells. In subsequent experiments, we also used HGT-1 cells as additional verification of *MCOLN2* QPCR data obtained from HEK-293 cells. In general, the cells were cultured in a 96-well plate and transfected with a standard plasmid concentration of 0.2 μg using TurboFect™ reagent according to the manufacturer's recommendations (Thermo Scientific, Pittsburgh PA). The treated cells were analyzed for *Luciferase* and *Renilla* activities using the Dual-Glo Luciferase Assay kit (Promega, Madison WI) at 48 h post-transfection using a BioTek Synergy 2 microplate reader. Untransfected cells were used as negative control, while cell-free wells containing media only were used as blank control.

Expression Constructs of Candidate Transcription Factors

Sequence analysis of the 103 construct using the JASPAR (http://jaspar.binf.ku.dk) and TRANSFAC (http://www.gene-regulation.com) databases revealed the following candidate TF proteins and used in this study: 1) CCCTC-binding factor zinc finger protein (*CTCF*; NCBI Gene ID 10664); 2) Kruppel-like factor 4 (*KLF4*; NCBI Gene ID 9314); 3) myeloid zinc finger 1 (*MZF1*; NCBI Gene ID 7593); 4) nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 (*NF-κB1*: NCBI Gene ID 4790); and 5) B-cell specific lineage activator protein (BSAP) or paired box 5 (*PAX5*; NCBI Gene ID 5079).

All expression constructs used in the study were driven by the strong CMV promoter and contained a fluorescent protein fusion to monitor the efficiency of gene expression. The pmCherry and pmOrange empty vectors were a generous gift from Dr. Roger Tsien (UC San Diego). We PCR cloned NF-κB1 into the pmOrange vector, while MZF1and PAX5 into the pmCherry vector using a human cDNA library (Clontech, Mountain View CA). KLF4 was PCR cloned into the pmCherry vector using the FU-tetO-hKLF4 plasmid (Addgene plasmid ID 19777) as a template. The CTCF expression vector (pMY-mutCTCF-T2A-mOrange) was a kind gift from Drs. Viktor Lobanenkov and Steevenson Nelson (NIH). The expression constructs containing the cloned genes were sequence-verified prior to use.

To study if the candidate TFs are able to induce the transcription of endogenous *MCOLN2* mRNA levels, we transfected each TF construct (4 μg) into HEK-293 cells with TurboFect reagent. Non-transfected and empty vector transfected cells were used as controls. The transfection efficiency was monitored using fluorescence microscopy 24 h post-treatment. After a 48-h incubation period, total RNA was extracted using TRI Reagent (Molecular Research Center, Cincinnati OH) and reverse transcribed (RT) with Maxima H^- RT according to the manufacturer's recommendation (Thermo Scientific). Each RT reaction contained 2 μg of total RNA. The presence of cDNAs was confirmed by a standard PCR reaction using the housekeeping gene, human glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*). The amount of *MCOLN2* transcript was then analyzed using real-time QPCR as described earlier.

Western blot analyses were performed to determine the effects of candidate TF proteins on the endogenous TRPML2 protein levels. HEK-293 cells were transfected as described above and the cells were lysed 48 h post-transfection. The cell lysates where run on SDS-PAGE and the endogenous TRPML2 proteins were probed with anti-TRPML2 polyclonal antibody (pAb; 1:1000; Novus Biologicals, Littleton CO). The blot was viewed with a secondary anti-

rabbit 800CW IRdye (LICOR, Lincoln NE) and scanned using the LICOR Odyssey Sa™ imaging system.

RNA interference of endogenous PAX5 expression

We commercially purchased four short-hairpin RNA (shRNA) vectors that target PAX5 (Origene Technologies, Bethesda MD). We used HEK-293 cells to test the effectiveness of the shRNA vectors, and we identified only one out of four shRNA vectors to be effective in knocking down *PAX5* mRNA levels. This particular shRNA was then used in the study. The RNAi experiments were analyzed using real-time QPCR and Western blot techniques. For Western blot, anti-β-ACTIN monoclonal antibody (mAb; 1:1000; Rockland Immunochemicals, Limerick PA) was used as loading control, while anti-TRPML2 pAb described earlier was used to detect endogenous TRPML2 protein. The blot was viewed with a secondary anti-rabbit 800CW and anti-mouse 680LT IRdyes, and imaged using the LICOR Odyssey Sa™ imaging system.

RESULTS

Mapping of the MCOLN2 promoter region and identification of the core promoter sequence

We performed *in silico* studies using the UCSC Genome Browser and the NCBI's Epigenomics website (http://www.ncbi.nlm.nih.gov/epigenomics/) to map the genomic DNA of *MCOLN2* gene and the nearby gene, *MCOLN3*. Our analysis showed that the two genes are separated by 20,970 bp of genomic DNA (Fig. 1A). Furthermore, the database showed the presence of CpG island that spans 1,182 bp and ends at −780 bp upstream of the TSS, while another CpG island is present from −963 to −1593 bp upstream of the TSS (Fig. 1A).

Cross-referencing the UCSC Genome Browser and NCBI's RefSeq NM_153259.2 showed that the longest DLR construct we initially cloned included 123 bp of the 5'-untranslated region (UTR) of the gene. To accurately map the *MCOLN2* promoter region, we designated all DLR constructs with reference to the genomic sequence location of the TSS. The longest construct (1,583 bp) we originally made was then designated as 1,460 upstream of the TSS (Fig. 1B), while the shorter fragments that were subsequently created using PCR deletion mutagenesis are denoted as DLR constructs: 181 (Fig. 1C), 103 (Fig. 1D), and 5'-UTR+12 (Fig. 1E). The relevant genomic DNA sequences cloned into the DLR constructs were submitted to Genbank (Accession Nos. KM882898, KM979446, KM995813).

We fortuitously identified a 34-bp sequence that was deleted while PCR cloning the 103 bp DLR construct – we refer to this clone as 103(del34) (Fig. 1F). The 34-bp sequence consists of G and C nucleotides, which resembled multiple GC boxes bearing the consensus sequence "GGGCGG", which are often found in mammalian promoter regions (Suzuki et al., 2001). Further analysis of the putative promoter region of *MCOLN2* cloned into the 103 construct showed that it does not contain a TATA box. Nonetheless, our bioinformatics and experimental analyses of the 103 bp DNA sequence initially suggested that it might be the putative core promoter. To test this hypothesis, we used the 103 DLR construct to delete a 53-bp intervening sequence between −45 bp upstream of the TSS and +9 bp of *MCOLN2*'s

5’-UTR. This approach essentially placed the 34-bp sequence upstream of the 5’UTR+12 DNA sequence, which we called as the 5’UTR+12(ins34) construct (Fig. 1G).

MCOLN2 genomic DNA region between −79 and −52 bp upstream of the TSS is required and necessary for promoter activation

The dual-luciferase reporter assays revealed that the minimal core promoter region of *MCOLN2* is located within the 1460, 181, and 103 DNA segments, but not within 5’-UTR +12 fragment as evidenced by the luciferase activities of each DLR construct (Fig. 2A). Interestingly, a significantly higher luciferase signal was observed for the 103 construct, which further suggested that the core promoter is likely located within this DNA segment.

We initially did not use the 103(del34) construct in the dual-luciferase experimental trials, because we wanted to first establish that all sequence-intact promoter constructs are being expressed and that they have luciferase activity when assayed. After obtaining the results shown in Figure 2A, we then performed additional experiments comparing the 103(del34) mutant along with the 103 and 5’-UTR+12 constructs for comparison. We found that the 103(del34) mutant clone significantly lost its luciferase activity compared to the full 103 construct (Fig. 2B). The magnitude of the reduction in luciferase activity of the 103(del34) mutant was slightly higher than the luciferase activity of the 5’-UTR+12 construct, but it was not significantly different. Sequence analysis of the 103(del34) deletion mutant showed that the 34-bp region (Fig. 2C) contains three GC boxes.

To confirm our observation, we performed site-directed mutagenesis (SDM) experiments on the 103 DLR construct. Specifically, we targeted the DNA region containing the GGG (−71 to −69) motif into AAA, referred to as 103(3) point mutant (see Fig. 2C for the map). The 103(3) point mutant DLR construct exhibited a significant reduction in luciferase activity when compared to the intact 103 construct (Fig. 2D). In contrast, this mutant had significantly higher luciferase activity upon comparison with the 5’-UTR+12 construct. We then created a double trinucleotide point mutant DLR construct by mutating the CCC (−54 to −52) motif into TTT (Fig. 2C) within the original 103(3) point mutant construct. This double point mutation was subsequently called 103(6) point mutant. Analysis of the luciferase activity of the 103(6) point mutant showed significant reduction of promoter activation when compared with the intact 103 construct. The 103(6) point mutant also showed a reduction of luciferase activity relative to the 103(3) point mutant (Fig 2D). Interestingly, the magnitude of the luciferase activity obtained from the 103(6) point mutant was very similar to the 103(del34) mutant construct, which then narrowed down the length of the candidate core promoter sequence between −79 and −52 upstream of the TSS (Fig. 2C). Since the promoter activity of the 5’-UTR+12 construct was virtually non-existent (Figs. 2A-2B), we hypothesized that inserting this 34-bp DNA sequence into this construct could activate luciferase activity. We designated this insertion mutant as 5’-UTR+12(ins34) construct (Fig. 1G). We assayed the luciferase activity of the 5’-UTR+12(ins34) clone and found that it did increase the relative luciferase activity, but the magnitude of the response was not significant (Fig 2D). Nevertheless, the overall data showed that the 34 bp DNA sequence was very likely to be the binding site of an unknown transcription factor regulating *MCOLN2* gene expression.

PAX5 is the transcriptional activator of the MCOLN2 gene and the core promoter is at −79 and −60 bp upstream of the TSS

To elucidate the transcription factor for TRPML2 given the results we gathered from the DLR assays, we first performed *in silico* analyses of the 103 bp DNA segment using the JASPAR and TRANSFAC databases. The experimental data we obtained from the DLR assays (Figs. 2A-2B) initially restricted our search for the core promoter sequence between −79 and −46 upstream of the TSS, while deletion and SDM experiments further refined the location to −79 and −52 upstream of the TSS (Fig. 2D). Furthermore, the information gathered from testing the 103(del34) mutant clone has allowed us to further limit the list of candidate TF proteins obtained from the JASPAR and TRANSFAC databases. In this study, we identified five candidate TFs, namely, CTCF, KLF4, MZF1, NF-κB1, and PAX5 (Fig. 3A). The candidate TFs were cloned into a pmCherry or pmOrange fluorescent-tagged mammalian expression vector. Each TF protein was tested on its ability to increase the mRNA levels of endogenous *MCOLN2* gene using HEK-293. We found that heterologous expression of PAX5 in HEK-293 cells significantly increased the endogenous *MCOLN2* transcripts by 2.25 fold compared with untreated and empty vector controls (Fig. 3B). CTCF showed a modest effect on increasing *MCOLN2* transcripts relative to untreated and empty vector controls, but the data were not significant. Similarly, the other candidate TF proteins analyzed did not show any significant effect on endogenous *MCOLN2* transcript levels. To further corroborate the positive effect of PAX5 on endogenous *MCOLN2* transcripts, we used another cell line (HGT-1 adenocarcinoma) to heterologously express PAX5. We found that over-expression of PAX5 in HGT-1 cells also significantly increased the *MCOLN2* transcripts when compared with controls (data not shown), which further supported the data obtained using HEK-293 cells. We then performed a Western blot to determine if the increase in *MCOLN2* transcripts paralleled an elevation of TRPML2 protein, since changes in mRNA levels do not necessarily reflect protein levels for various reasons. We observed that PAX5 over-expression in HEK-293 cells significantly increased the endogenous TRPML2 protein levels compared with control treatments and other candidate TF proteins (Fig. 3C). We subsequently analyzed and quantified the integrated density values (IDV) of each band in all trials, which showed a significant increase in the relative amount of TRPML2 proteins in the presence of heterologously expressed PAX5 protein (Fig. 3D). This result further confined the core promoter to the PAX5 binding region of −79 to −60 bp upstream of the TSS (see Fig. 3A for a map).

To further confirm the effect of PAX5 on TRPML2, we used RNA interference (RNAi) targeting the *PAX5* gene. Four PAX5-shRNA vectors were commercially purchased, but upon testing, we identified only one PAX5-shRNA that significantly decreased the endogenous levels of *PAX5* transcripts present in HEK-293 cells (Fig. 4A). We then used this particular PAX5-shRNA to analyze its effect on regulating *MCOLN2* mRNA and TRPML2 protein levels. We found that knocking down endogenous PAX5 significantly decreased *MCOLN2* transcript levels of (Fig. 4B), and reduced TRPML2 protein levels (Fig. 4C). Note, however, that the effect of such indirect knock down on TRPML2 protein was less prominent compared to the *MCOLN2* mRNA.

Finally, we performed additional bioinformatics analysis to verify if our data were supported by other studies available in UCSC Genome Browser and NCBI databases. We found that the "Transcription Factor ChIP-Seq" data from the Encyclopedia of DNA Elements (ENCODE) provided evidence for both CTCF and PAX5 binding, but not for KLF4, MZF1, and NF-κB1 (not shown). We also checked whether the 34 bp sequence (Fig. 1F), which includes the PAX5 binding motif, 5'-GGCGGGGCGGGGCGGGGCGGGG-3', is highly conserved among mammalian species using NCBI's BLAST website. We found a bacterial artificial chromosome (BAC) clone from chromosome 7 of *Pan troglodytes* (Genbank Accession No. AC183601.2), which showed conservation of the 34 bp region along with downstream DNA sequence that includes the 5'-UTR of *MCOLN2* gene. Interestingly, the *Pan troglodytes* sequence had an additional "CGGGGC" motif at the 3' end of the PAX5 binding motif shown above. Meanwhile, our analysis of the mouse chromosome 3 genome, where *Mcoln2* is located, did not show consensus sequences for Pax5 or Ctcf protein. We did, however, notice the presence of four GC boxes within −120 bp upstream of the TSS, two of which were 30 bp away from the TSS of the *Mcoln2* gene (not shown).

In summary, the observational and experimental data performed in the current study showed quite convincingly that PAX5 is the transcription factor responsible for *MCOLN2* gene expression.

DISCUSSION

Promoters play an essential role in understanding the transcriptional mechanisms of genes. We show for the first time that B-cell lineage specific activator protein (BSAP), also known as PAX5, is the transcription factor that regulates the expression of TRPML2 protein. Our study also reveals that the core promoter sequence of *MCOLN2* gene contains GC boxes located between −79 and −46 base pairs upstream of the transcriptional start site.

Mapping of the MCOLN2 promoter region and identification of the core promoter sequence

The presence of a CpG island that spans −780 bp upstream of the TSS (Fig. 1B) was very informative, since it explains the constitutive expression of *MCOLN2* in the HEK-293 and HGT-1 cell lines we used in the study. This is because CpG islands located within promoter regions appear to create a more conducive chromatin state that favors transcription or allows gene expression silencing through intensive CpG methylation (Deaton and Bird, 2011). Interestingly, a previous report been suggested that CpG islands within promoter regions might be typically correlated with a diverse TSS and the lack TATA boxes (Deaton and Bird, 2011). Furthermore, certain genes that have GC boxes are associated with TATA-less boxes (Blake et al., 1990). Indeed, we observed that the *MCOLN2* promoter exhibits these aforementioned features whereby it several GC boxes, lacks a TATA box, and has a CpG island within its TSS.

Our discovery of the core promoter sequence was made easier by the significantly higher luciferase activity bestowed upon by the 103 DLR construct, as well as the serendipitous deletion of the 34-bp region in one of the clones created for the 103 construct. We are unable to explain why the luciferase activities of both the 1460 and 181 constructs were considerably less compared to that of the 103 construct. We can only speculate that these

DLR constructs were less efficient in activating the core promoter due to their longer nucleotide length, whereas the 103 construct was more efficient because it is shorter. Alternatively, there could be a repressive molecule that recognizes DNA sequences upstream of the 103-bp sequence, which resulted in less promoter activation. The findings of Cooper et al. (2006) possibly support this reasoning because some negative regulatory elements of certain genes have been detected between −1000 bp and −500 bp upstream of the TSS (Cooper et al., 2006). Future research using deletion mutagenesis strategy of both the 181 and 1460 DLR constructs could reveal if this is really the case.

MCOLN2 genomic DNA region between −79 and −52 bp upstream of the TSS is required and necessary for promoter activation

The study by Cooper et al. (2006) that some human gene promoters lie between −300 to −50 bp of the TSS gave us an *a priori* expectation that the *MCOLN2* core promoter would be identified within the 103-bp DNA segment cloned into the DLR construct. This belief was reinforced by our findings that the 34-bp candidate promoter region has GC boxes and that MCOLN2 might be one of a few genes classified as TATAA-less promoter (Blake et al., 1990). Either way, the effect of the trinucleotide GGG to AAA point mutations was remarkable, but the additional CCC to TTT point mutations downstream of this site produced an even more interesting result, because the luciferase activity of this double trinucleotide point mutant is comparable to that of the 34-bp deletion mutant. This result further validated and even confined the core promoter region to within −79 and −52 upstream of the TSS. Meanwhile, the modest effect of 5'-UTR+12(ins34) construct on increasing luciferase activity suggests that the 34-bp DNA segment alone is not sufficient to activate *MCOLN2* transcription (recall that the 5'-UTR+12 construct has completely loss its promoter activity). This finding somehow makes sense because the absence of 53-bp intervening sequence between the 34-bp region and the first 12 nucleotides of the 5'-UTR may be a mitigating factor in that the −1 to −51 bp region contains the binding sequence for other *cis*-regulatory elements and RNA polymerase II protein. Notwithstanding, our findings from various experiments confirmed that the core promoter of *MCOLN2* gene lies within the GC boxes located within −79 and −52 bp upstream of the TSS (28 nucleotides in total).

PAX5 is the transcriptional activator of the MCOLN2 gene and the core promoter is at −79 and −60 bp upstream of the TSS

Our bioinformatics analysis of the candidate TF proteins regulating the *MCOLN2* gene in reference to the 34-bp DNA sequence has helped us deduce which TF likely binds to the core promoter. The JASPAR and TRANSFAC databases gave us a list of candidate TF proteins that potentially bind to the GC boxes located within the 34-bp region. We narrowed down the list to five (CTCF, KLF4, MZF1, NF-κB1 and PAX5). Our results showed that heterologous over-expression of PAX5 significantly increases the amounts of *MCOLN2* transcript and TRPML2 protein (Fig. 3), suggesting that it is the transcriptional activator of *MCOLN2*. More importantly, RNAi of PAX5 significantly reduced the levels of constitutively expressed *MCOLN2* transcripts in HEK-293 cells, and Western blot analysis revealed a similar effect, albeit subtle (Fig. 4). Although knocking down PAX markedly reduced the expression of *MCOLN2* transcript and TRPML2 protein levels, the constitutive levels we detected post-RNAi treatment suggest that both transcript and protein levels are

fairly stable. It would be interesting to see in future study if the levels *MCOLN2* transcript and TRPML2 protein virtually disappear when PAX5 is completely knocked out in cells. Finally, the revelation that the ENCODE's Transcription Factor Chip-Seq data supported our experimental findings indicate that PAX5 is indeed responsible for mediating *MCOLN2* expression.

Note that CTCF appeared to have a modest effect in elevating the relative mRNA levels of *MCOLN2*, but no overt elevation in the levels of TRPML2 protein was observed. Furthermore, we would like to point out that the CCC to TTT point mutations located at position −54 to −52 (Fig. 2C) is situated where CTCF putatively binds (Fig. 3A). Recall that the 103(6) DLR construct with double trinucleotide point mutation had a lower luciferase activity than the 103(3) DLR construct with single trinucleotide point mutation (Fig. 2D). Moreover, our analysis of the ENCODE database shows that CTCF also binds to the *MCOLN2* promoter. Based on these apparent circumstantial evidences, it is still possible that CTCF may have a contributory role in modulating *MCOLN2* expression in addition to PAX5. An alternative explanation why the TTT point mutations located at position −54 to −52 had reduced luciferase activity is that this motif is the site of other *cis*-regulatory elements that regulate *MCOLN2* expression and not necessarily part of the CTCF binding region. Notwithstanding, the existence of other *cis*-regulatory elements and transcription factors influencing the constitutive expression of *MCOLN2* could also explain why the decrease in levels of *MCOLN2* transcript and TRPML2 protein levels upon PAX5-shRNA treatment was not as dramatic as the reduction of PAX5 expression.

We were quite intrigued to see that four of the five candidate transcription factors (CTCF, MZF1, NF-κB1 and PAX5/BSAP) have been implicated in immune cell function and development, inflammatory response, or immune-related pathology (Nutt et al., 1999; Wu et al., 2006; Cobaleda et al., 2007; Gil et al., 2007; Powell et al., 2013; Collins et al., 2014). For example, CTCF has been shown to regulate MHC class II expression in addition to other TF proteins (Ottaviani et al., 2012). MZF1 AND NF-κB1 have been shown to involved in early myeloid lineage differentiation (Powell et al., 2013), while PAX5 has been found to be required for normal B-cell development (Nutt et al., 1999). *MCOLN2* has been linked in the differentiation and maturation of B cells (Lindvall et al., 2005), so it is logical quite possible that one or more transcription factors that regulate immune activities could also up-regulate the expression of *MCOLN2* gene. Our data, however, indicate that only PAX5 significantly increases the levels of TRPML2 protein. Even though CTCF serves a role in immune cells, and MZF1 is critical in myeloid cell development, or that NF-κB1 is involved in pro-inflammatory responses in immune cells, our findings show that none of these TF proteins significantly activated the *MCOLN2* gene promoter, suggesting the possibility that none of them is directly involved in TRPML2 protein expression.

Concluding Remarks

Our interest in TRPML2 ion channel function lies in two fronts. First, previous reports have indicated a putative role for this protein in the immune system, specifically in B-cell lymphocytes. Second, the high degree of similarity in the structure and ion channel characteristics of TRPML2 with TRPML1, make TRPML2 an ideal protein to substitute for

the loss of TRPML1 in MLIV. This is because a previous study in *C. elegans* revealed that gene complementation of the non-functional CUP-5 protein (a TRPML1 ortholog) using either human TRPML1 or TRPML3 rescues the disease phenotype in worms (Treusch et al., 2004). Thus, based on the physical and functional redundancy between TRPML1 and TRPML2, we proposed that TRPML2 is the best candidate for a proof-of-principle gene complementation approach to substitute for the missing TRPML1 function in MLIV disease.

In conclusion, understanding the transcriptional and translational regulation of TRPML2 is an important first step to further characterize and dissect its biological significance in cells, particularly in B-cell lymphocyte development and function.

ACKNOWLEDGEMENTS

We are very grateful to Saima Ali, Vinod Valluri, Dewayne Campbell, Joshua Silva, and Ania Habibi for providing technical support. We thank Drs. Viktor Lobanenkov (NIH/NIAID) and Steevenson Nelson (NIH/NIAMS) for providing the CTCF expression clones, Dr. Ehud Goldin (NIH/NHGRI) for giving us a frozen stock of HGT-1 gastric adenocarcinoma cells, and Dr. Roger Tsien (UC San Diego) for providing the pmCherry and pmOrange vectors. We also thank Dr. Sean Murray for reading and critiquing the manuscript. This work was funded by grants to MPC from NIH R15-NS070774, NSF MCB-0920127, and CSUF Intramural Grants program. JV is a top awardee at the 28th Annual CSU Statewide Student Research Competition and the 2014 CSUF Student Research Competition.

REFERENCES

Bargal R, Avidan N, Ben-Asher E, Olender Z, Zeigler M, Frumkin A, Raas-Rothschild A, Glusman G, Lancet D, Bach G. Identification of the gene causing mucolipidosis type IV. Nat Genet. 2000; 26:118–123. [PubMed: 10973263]

Bassi MT, Manzoni M, Monti E, Pizzo MT, Ballabio A, Borsani G. Cloning of the gene encoding a novel integral membrane protein, mucolipidin-and identification of the two major founder mutations causing mucolipidosis type IV. Am J Hum Genet. 2000; 67:1110–1120. [PubMed: 11013137]

Blake MC, Jambou RC, Swick AG, Kahn JW, Azizkhan JC. Transcriptional initiation is controlled by upstream GC-box interactions in a TATAA-less promoter. Mol Cell Biol. 1990; 10:6632–6641. [PubMed: 2247077]

Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem. 2009; 55:611–622. [PubMed: 19246619]

Cobaleda C, Schebesta A, Delogu A, Busslinger M. Pax5: the guardian of B cell identity and function. Nature immunology. 2007; 8:463–470. [PubMed: 17440452]

Collins PE, Kiely PA, Carmody RJ. Inhibition of transcription by B cell Leukemia 3 (Bcl- 3) protein requires interaction with nuclear factor kappaB (NF-kappaB) p50. J Biol Chem. 2014; 289:7059–7067. [PubMed: 24459141]

Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM. Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome. Genome Res. 2006; 16:1–10. [PubMed: 16344566]

Cuajungco MP, Basilio LC, Silva J, Hart T, Tringali J, Chen CC, Biel M, Grimm C. Cellular Zinc Levels Are Modulated by TRPML1-TMEM163 Interaction. Traffic. 2014

Cuajungco MP, Leyne M, Mull J, Gill SP, Lu W, Zagzag D, Axelrod FB, Maayan C, Gusella JF, Slaugenhaupt SA. Tissue-specific reduction in splicing efficiency of IKBKAP due to the major mutation associated with familial dysautonomia. Am J Hum Genet. 2003; 72:749–758. [PubMed: 12577200]

Cuajungco MP, Podevin W, Valluri VK, Bui Q, Nguyen VH, Taylor K. Abnormal accumulation of human transmembrane (TMEM)-176A and 176B proteins is associated with cancer pathology. Acta histochemica. 2012; 114:705–712. [PubMed: 22244448]

Curcio-Morelli C, Zhang P, Venugopal B, Charles FA, Browning MF, Cantiello HF, Slaugenhaupt SA. Functional multimerization of mucolipin channel proteins. Journal of cellular physiology. 2010; 222:328–335. [PubMed: 19885840]

Deaton AM, Bird A. CpG islands and the regulation of transcription. Genes & development. 2011; 25:1010–1022. [PubMed: 21576262]

Di Palma F, Belyantseva IA, Kim HJ, Vogt TF, Kachar B, Noben-Trauth K. Mutations in Mcoln3 associated with deafness and pigmentation defects in varitint-waddler (Va) mice. Proc Natl Acad Sci U S A. 2002; 99:14994–14999. [PubMed: 12403827]

Donaldson JG. Multiple roles for Arf6: sorting, structuring, and signaling at the plasma membrane. J Biol Chem. 2003; 278:41573–41576. [PubMed: 12912991]

Dong XP, Cheng X, Mills E, Delling M, Wang F, Kurz T, Xu H. The type IV mucolipidosis-associated protein TRPML1 is an endolysosomal iron release channel. Nature. 2008; 455:992–996. [PubMed: 18794901]

Eichelsdoerfer JL, Evans JA, Slaugenhaupt SA, Cuajungco MP. Zinc dyshomeostasis is linked with the loss of mucolipidosis IV-associated TRPML1 ion channel. J Biol Chem. 2010; 285:34304–34308. [PubMed: 20864526]

Flores EN, Garcia-Anoveros J. TRPML2 and the evolution of mucolipins. Advances in experimental medicine and biology. 2011; 704:221–228. [PubMed: 21290298]

Gil A, Maria Aguilera C, Gil-Campos M, Canete R. Altered signalling and gene expression associated with the immune system and the inflammatory response in obesity. The British journal of nutrition. 2007; 98(Suppl 1):S121–S126. [PubMed: 17922949]

Grimm C, Cuajungco MP, van Aken AF, Schnee M, Jors S, Kros CJ, Ricci AJ, Heller S. A helix-breaking mutation in TRPML3 leads to constitutive activity underlying deafness in the varitint-waddler mouse. Proc Natl Acad Sci U S A. 2007; 104:19583–19588. [PubMed: 18048323]

Grimm C, Jors S, Saldanha SA, Obukhov AG, Pan B, Oshima K, Cuajungco MP, Chase P, Hodder P, Heller S. Small molecule activators of TRPML3. Chemistry & biology. 2010; 17:135–148. [PubMed: 20189104]

Karacsonyi C, Miguel AS, Puertollano R. Mucolipin-2 localizes to the Arf6-associated pathway and regulates recycling of GPI-APs. Traffic. 2007; 8:1404–1414. [PubMed: 17662026]

LaPlante JM, Sun M, Falardeau J, Dai D, Brown EM, Slaugenhaupt SA, Vassilev PM. Lysosomal exocytosis is impaired in mucolipidosis type IV. Mol Genet Metab. 2006; 89:339–348. [PubMed: 16914343]

Li X, Saitoh SI, Shibata T, Tanimura N, Fukui R, Miyake K. Mucolipin 1 positively regulates TLR7 responses in dendritic cells by facilitating RNA transportation to lysosomes. International immunology. 2014

Lindvall JM, Blomberg KE, Wennborg A, Smith CI. Differential expression and molecular characterisation of Lmo7, Myo1e, Sash1, and Mcoln2 genes in Btk-defective B-cells. Cellular immunology. 2005; 235:46–55. [PubMed: 16137664]

Martina JA, Lelouvier B, Puertollano R. The calcium channel mucolipin-3 is a novel regulator of trafficking along the endosomal pathway. Traffic. 2009; 10:1143–1156. [PubMed: 19497048]

Miedel MT, Rbaibi Y, Guerriero CJ, Colletti G, Weixel KM, Weisz OA, Kiselyov K. Membrane traffic and turnover in TRP-ML1-deficient cells: a revised model for mucolipidosis type IV pathogenesis. J Exp Med. 2008; 205:1477–1490. [PubMed: 18504305]

Montell C. The TRP superfamily of cation channels. Sci STKE. 2005; 2005:re3. [PubMed: 15728426]

Nutt SL, Heavey B, Rolink AG, Busslinger M. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. Nature. 1999; 401:556–562. [PubMed: 10524622]

Ottaviani D, Lever E, Mao S, Christova R, Ogunkolade BW, Jones TA, Szary J, Aarum J, Mumin MA, Pieri CA, Krawetz SA, Sheer D. CTCF binds to sites in the major histocompatibility complex that are rapidly reconfigured in response to interferongamma. Nucleic acids research. 2012; 40:5262–5270. [PubMed: 22367884]

Powell ND, Sloan EK, Bailey MT, Arevalo JM, Miller GE, Chen E, Kobor MS, Reader BF, Sheridan JF, Cole SW. Social stress up-regulates inflammatory gene expression in the leukocyte transcriptome via beta-adrenergic induction of myelopoiesis. Proc Natl Acad Sci U S A. 2013; 110:16574–16579. [PubMed: 24062448]

Samie M, Wang X, Zhang X, Goschka A, Li X, Cheng X, Gregg E, Azar M, Zhuo Y, Garrity AG, Gao Q, Slaugenhaupt S, Pickel J, Zolov SN, Weisman LS, Lenk GM, Titus S, Bryant-Genevier M, Southall N, Juan M, Ferrer M, Xu H. A TRP channel in the lysosome regulates large particle phagocytosis via focal exocytosis. Dev Cell. 2013; 26:511–524. [PubMed: 23993788]

Samie MA, Grimm C, Evans JA, Curcio-Morelli C, Heller S, Slaugenhaupt SA, Cuajungco MP. The tissue-specific expression of TRPML2 (MCOLN-2) gene is influenced by the presence of TRPML1. Pflugers Arch. 2009; 459:79–91. [PubMed: 19763610]

Song Y, Dayalu R, Matthews SA, Scharenberg AM. TRPML cation channels regulate the specialized lysosomal compartment of vertebrate B-lymphocytes. Eur J Cell Biol. 2006; 85:1253–1264. [PubMed: 17050035]

Sun M, Goldin E, Stahl S, Falardeau JL, Kennedy JC, Acierno JS Jr, Bove C, Kaneski CR, Nagle J, Bromley MC, Colman M, Schiffmann R, Slaugenhaupt SA. Mucolipidosis type IV is caused by mutations in a gene encoding a novel transient receptor potential channel. Hum Mol Genet. 2000; 9:2471–2478. [PubMed: 11030752]

Suzuki Y, Tsunoda T, Sese J, Taira H, Mizushima-Sugano J, Hata H, Ota T, Isogai T, Tanaka T, Nakamura Y, Suyama A, Sakaki Y, Morishita S, Okubo K, Sugano S. Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Res. 2001; 11:677–684. [PubMed: 11337467]

Thompson EG, Schaheen L, Dang H, Fares H. Lysosomal trafficking functions of mucolipin-1 in murine macrophages. BMC Cell Biol. 2007; 8:54. [PubMed: 18154673]

Treusch S, Knuth S, Slaugenhaupt SA, Goldin E, Grant BD, Fares H. Caenorhabditis elegans functional orthologue of human protein h-mucolipin-1 is required for lysosome biogenesis. Proc Natl Acad Sci U S A. 2004; 101:4483–4488. [PubMed: 15070744]

Venkatachalam K, Hofmann T, Montell C. Lysosomal localization of TRPML3 depends on TRPML2 and the mucolipidosis-associated protein TRPML1. J Biol Chem. 2006; 281:17517–17527. [PubMed: 16606612]

Vergarajauregui S, Connelly PS, Daniels MP, Puertollano R. Autophagic dysfunction in mucolipidosis type IV patients. Hum Mol Genet. 2008; 17:2723–2737. [PubMed: 18550655]

Wu D, Li T, Lu Z, Dai W, Xu M, Lu L. Effect of CTCF-binding motif on regulation of PAX6 transcription. Invest Ophthalmol Vis Sci. 2006; 47:2422–2429. [PubMed: 16723452]

Zeevi DA, Lev S, Frumkin A, Minke B, Bach G. Heteromultimeric TRPML channel assemblies play a crucial role in the regulation of cell viability models and starvationinduced autophagy. J Cell Sci. 2010; 123:3112–3124. [PubMed: 20736310]

Highlights

- The putative *MCOLN2* promoter has 3 GC boxes, a CpG island, and lacks a TATA box
- Luciferase assay reveals candidate promoter at −79 to −52 bp upstream of the TSS
- Bioinformatics analysis of promoter sequence shows 5 potential TF proteins
- Experimental data show that PAX5 is the transcriptional activator of *MCOLN2* gene
- The PAX5 binding region is located at −79 to −60 bp upstream of the TSS

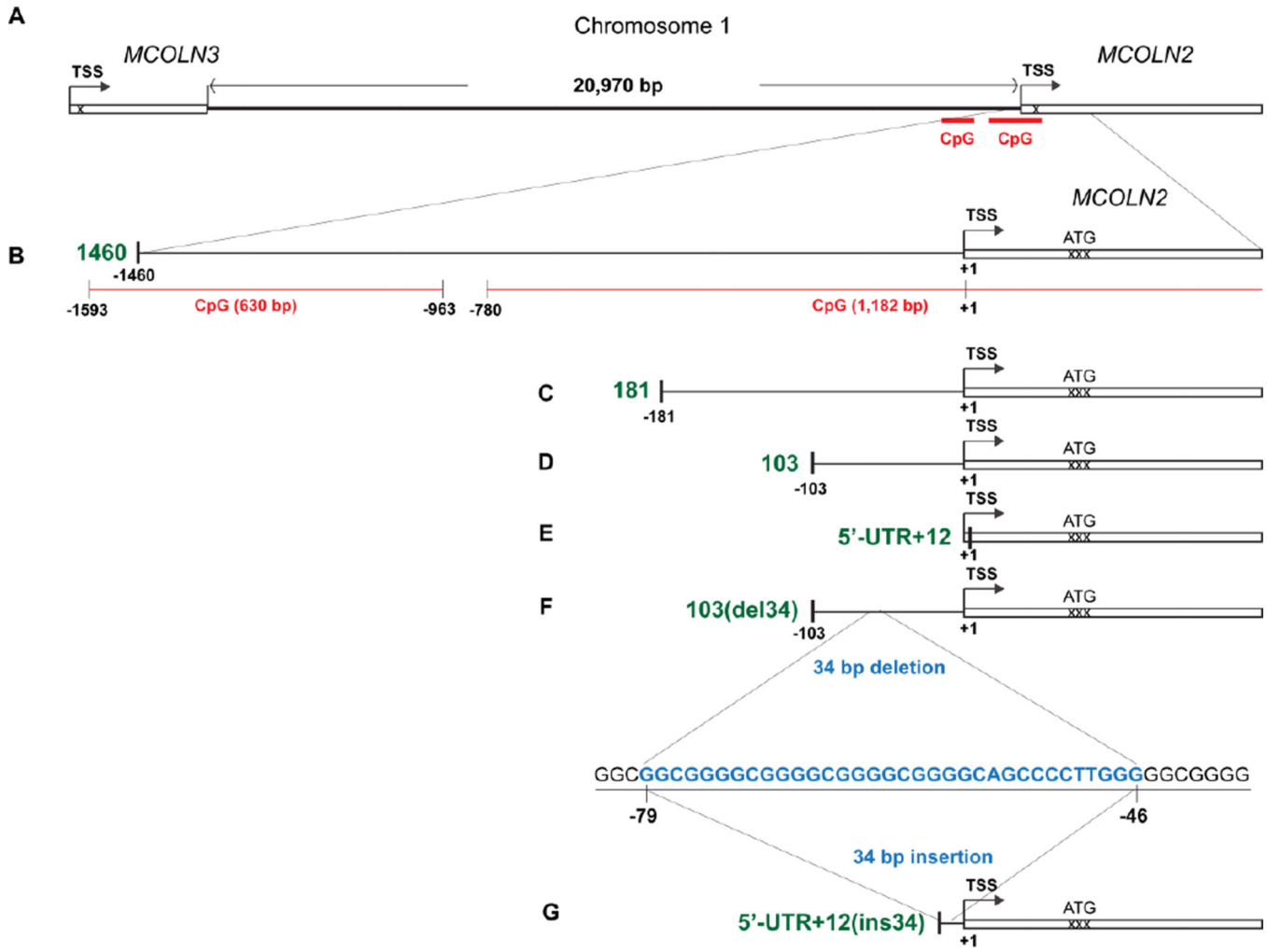

Figure 1. Schematic diagram of the genomic map of *MCOLN2* and dual-luciferase reporter (DLR) constructs

A) Chromosome 1 map of the *MCOLN2* and *MCOLN3* genes. Two CpG islands are shown to flank −780 bp upstream of the TSS (1,182 bp in length) and −963 bp upstream of the TSS. The DLR constructs containing DNA segments upstream of the TSS are shown in **B**) 1,460 bp, **C**) 181 bp, **D**) 103 bp. **E**) The 5'-UTR+12 DLR construct contains 114 bp of DNA sequence beginning from +12 of the UTR. **F**) The 103(del34) DLR construct is missing 34 bp of DNA sequence located at −79 to −46 upstream of the TSS. **G**) The 5'-UTR+12(ins34) DLR construct is the same as in E, but it contains the 34 bp DNA sequence deleted in F.

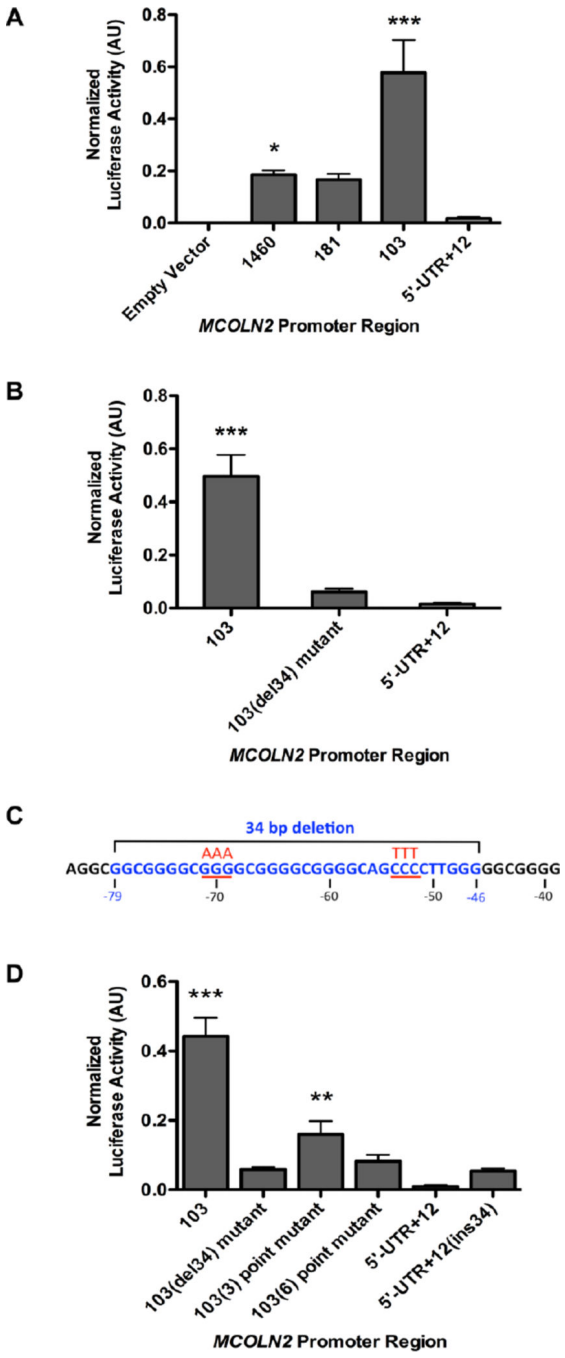

Figure 2. Dual luciferase reporter assay of various DNA DLR constructs containing the putative core promoter region of the *MCOLN2* gene

A) Luciferase activities of DLR constructs transfected into HEK-293 cells showed that the 1460, 181, and 103 vectors, but not the 5'-UTR+12 vector contain the putative core promoter of *MCOLN2*. **B**) The significant luciferase activity of the 103 DLR construct confined the location of the putative core promoter within this region. **C**) A schematic map of the 34 bp sequence (*blue text*) that was accidentally deleted during the PCR cloning of the 103 construct, denoted as 103(del34) mutant. The map also shows the location of the two

SDM constructs used in subsequent experiments, namely, 103(3) and 103(6) point mutants. **D**) DLR assays of the intact 103, 103(del34) mutant, 103(3) point mutant, and 103(6) point mutant showed that the DNA segment between −79 and −52 bp is the likely core promoter region of the *MCOLN2* gene. The presence of the 34 bp sequence in the 5'-UTR+12(ins34) DLR construct modestly increased its luciferase activity when compared with the 5'-UTR +12 clone, albeit not significantly. All luminescence values were normalized using the *Renilla* luciferase activity. (***$p < 0.001$, **$p < 0.01$; *$p < 0.05$; ANOVA followed by Bonferroni's multiple comparison *post-hoc* test, N > 3 independent trials).

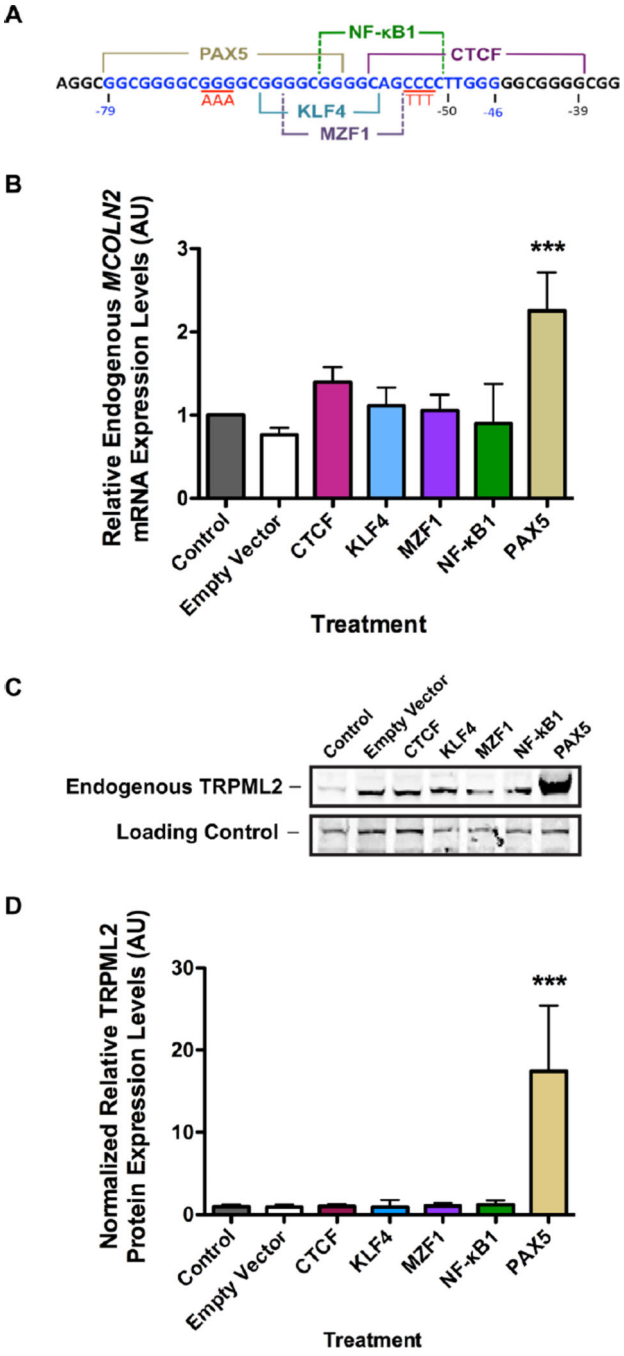

Figure 3. Analyses of endogenous *MCOLN2* transcript and TRPML2 protein upon heterologous expression of candidate transcription factors in HEK-293 cells

A) *In silico* analysis revealed five transcription factors that are predicted to bind within the TRPML2 promoter sequence identified by DLR assay. These candidate TF proteins were identified by searching and cross-referencing the JASPAR and TRANSFAC databases. **B**) Real-time QPCR analysis of HEK-293 cells transiently transfected with the corresponding TF proteins. Endogenous *MCOLN2* mRNA levels were quantified 48 h post-transfection. The QPCR data were analyzed as described in the *Materials and Methods* section using *18S*

rRNA as the normalizer, while the untreated control was used as the calibrator (value = 1). The data show that endogenous *MCOLN2* transcripts increased by as much as 2.25 fold in the presence of PAX5 (***p-value < 0.001; ANOVA followed by Bonferroni's multiple comparison *post-hoc* test, N > 3 independent trials). **C**) Representative Western blot image of endogenous TRPML2 protein upon transient transfection of candidate TF proteins. The top panel is TRPML2 blotted with anti-TRPML2 polyclonal antibody (pAb). The lower panel is the loading control taken from the same blot where the bands in each individual lane represented non-specific binding of the anti-TRPML2 pAb. The blot showed that PAX5 markedly increased endogenous TRPML2 protein levels compared to control and other TF proteins. **D**) Semi-quantitative analysis of endogenous TRPML2 protein bands upon heterologous expression of candidate TF proteins using integrated density values obtained from NIH ImageJ analysis. PAX5 significantly increased the TRPML2 protein levels compared with control and other transcription factors (***p-value < 0.001; ANOVA followed by Bonferroni's multiple comparison *post-hoc* test, N > 3 independent trials).

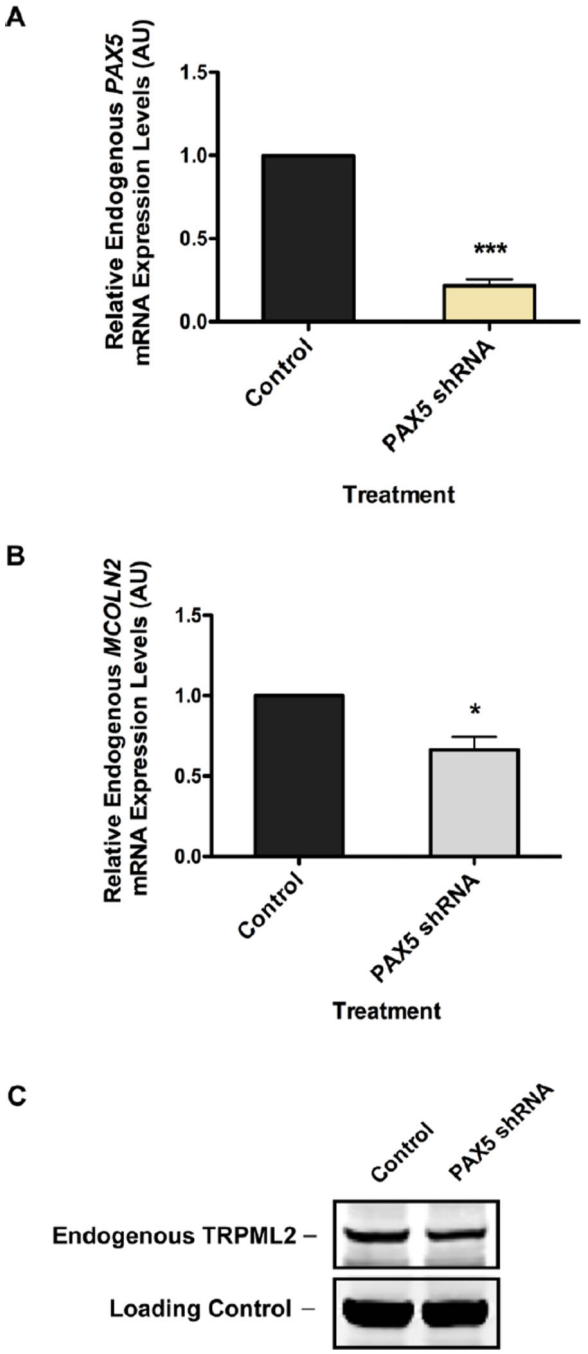

Figure 4. Analyses of endogenous *PAX5* and *MCOLN2* transcript levels, and TRPML2 protein levels upon RNA interference (RNAi) of PAX5 in HEK-293 cells

Real-time QPCR analyses of (**A**) *PAX5* and (**B**) *MCOLN2* mRNA levels upon RNAi-induced knockdown of PAX5 expression in HEK-293 cells. Endogenous *PAX5* and *MCOLN2* transcript levels were quantified 48 h post-transfection of short-hairpin RNA targeting endogenous PAX5 protein levels. The QPCR data were analyzed as described in the *Materials and Methods* section using *18S* rRNA as the normalizer, while the untreated control was used as the calibrator (value = 1). Endogenous *PAX5* and *MCOLN2* transcripts

significantly decreased in HEK-293 cells upon treatment of the PAX5-shRNA (***p-value < 0.001; *p < 0.05; Students t-test, paired, two-tailed, N > 3 independent trials). **C**) Representative Western blot image of endogenous TRPML2 protein expression level (control) and upon targeted knock down of PAX5 protein levels (PAX5-shRNA). The top panel is TRPML2 blotted with anti-TRPML2 pAb, while the lower panel is the loading control blotted with anti-β- ACTIN monoclonal Ab. Immunoblotting was performed as described in the *Materials and Methods* section.

TABLE 1

List of standard PCR and real-time quantitative PCR primers used in this study.

Name	Primer Sequence
MCOLN2-F1 DLR promoter 1460 bp	Fwd: 5'-**GGC CTA ACT GGC C**TT GCT GCT GCT ACC TCG TGA AAA G-3' (amplicon size = 1583 bp)
MCOLN2-F2* DLR promoter 181 bp	Fwd: 5'-**GGC CTA ACT GGC C**GC CGA CCC CCT TCT CCG GGA-3' (amplicon size = 304 bp)
MCOLN2-F3* DLR promoter 103 bp	Fwd: 5'-**GGC CTA ACT GGC C**AA GCT CCG GTC CCC GGT AGG A-3' (amplicon size = 226 bp)
MCOLN2-F4* DLR 5'-UTR+12 bp	Fwd; 5'-**GGC CTA ACT GGC C**CG CCT TTC AGC CGC CCG AGA A-3' (amplicon size = 114 bp)
MCOLN2-R* DLR 5'-UTR Universal	Rev: 5'-**GGC CTC GGC GGC C**TT CCC TTC TCT TAC CCT TTC T-3'
MCOLN2 Standard PCR	Fwd: 5'-ATG GCC CGG CAG CCT TAT CGT-3' Rev: 5'-TTA GCT AAT AGG TAT CAG TGA TCA TC-3'
GAPDH Standard PCR**	Fwd: 5'-ATG GGG AAG GTG AAG GTC G-3' Rev: 5'-AGT GGT CGT TGA GGG CAA T-3'
MCOLN2 Real-time QPCR	Fwd: 5'-CTG TGG CTG GAT TGT CTT AGG-3' Rev: 5'-CTG GAT TTG GGC AAA GGT TG-3'
PAX5 Real-time QPCR	Fwd: 5'-TGT TTG AGA GGC AGC ACT AC-3' Rev: 5'-CAT GGC TGA ATA CTC TGT GGT-3'
18S rRNA Real-time QPCR**	Fwd: 5'-GCC CGA AGC GTT TAC TTT G-3' Rev: 5'-CCC TCT TAA TCA TGG CCT CAG-3'

* The forward primers F2–F4 were used for PCR deletion strategy of the −1460 bp DLR construct, which gave DLR constructs containing smaller fragments of the putative promoter region and parts of the 5'-UTR. The universal reverse primer is located on the 5'-UTR of *MCOLN2* (+99 to +123), in which it was used to create the initial −1460 bp and shorter DLR constructs. The amplicons were cloned using the *Sfi I* restriction site (indicated as bolded text on the primer sequence).

** Cuajungco et al. (2012) previously reported these primer sets.