International Journal of Chemical and Biomedical Science
2015; 1(2): 38-43
Published online April 30, 2015 (http://www.aascit.org/journal/ijcbs)




**Keywords**
Zinc Finger Transcription Factor,
Dental Pulp Stem Cell,
DNA Sequencing

Received: March 30, 2015
Revised: April 18, 2015
Accepted: April 19, 2015

# Identification of a Predicted Zinc-Finger Transcription Factor (*windpipe* Homolog or WH1) in Dental Pulp Stem Cell mRNA

**Seyed Mohammadi[1], Sheila Heraypur[1], Karl Kingsley[2, *]**

[1]University of Nevada, Las Vegas, School of Dental Medicine, USA
[2]Department of Biomedical Sciences, School of Dental Medicine, University of Nevada, Las Vegas, Nevada, USA

## Email address
Karl.Kingsley@unlv.edu (K. Kingsley)

## Citation
Seyed Mohammadi, Sheila Heraypur, Karl Kingsley. Identification of a Predicted Zinc-Finger Transcription Factor (*windpipe* Homolog or WH1) in Dental Pulp Stem Cell mRNA. *International Journal of Chemical and Biomedical Science*. Vol. 1, No. 2, 2015, pp. 38-43.

## Abstract
The objective of this project was to find novel genes integral to head and neck development. An early developmental Drosophila gene, windpipe, was recently discovered to have a predicted human ortholog. Using the National Center for Biotechnology Information (NCBI) BLAST interface, this predicted gene (hmm30337) was identified from the transcript NM_178549. This corresponded to an 8565 base pair coding region containing three, distinct Open Reading Frames (ORFs), the first of which is a 1746 base pair transcript encoding the 582 amino acid (68.1 kDa) hypothetical protein (MGC42493) – a potential zinc-finger transcription factor (ZN678). Overlapping RT-PCR primers were designed to span this region. Total RNA derived from adult human tissue cell lines was isolated and mRNA expression screened. Negative results were obtained from all commercially available adult cell lines tested, which included breast, mammary gland, foreskin fibroblast, oral gingiva, colorectal, cervical and multiple dental pulp mesenchymal stem cell lines. However, positive results were obtained from two undifferentiated mesenchymal dental pulp stem cell lines, revealing mRNA expression corresponding with the 1746 bp of the ORF containing ZN678/MGC42493. Overlapping cDNA sequence analysis was used to identify a novel zinc-finger transcription factor transcribed in a small subset of undifferentiated dental pulp-derived mesenchymal stem cells.

## 1. Introduction

Most tissues and organs develop using a complex series of interactions between adjacent structures, such as epithelia and the underlying mesenchyme (1). Although much is already known about head and neck development and morphogenesis, this process is not completely understood and many of its mechanisms and pathways remain to be discovered (2). For example, although many of the genes and signaling molecules, as well as the transcription factors essential to these processes have been described – there are still several thousand predicted human genes that have been identified but have not yet had their potential function or role in any particular tissue described (3-5).

One early developmental gene expressed in the developing tracheal system of *Drosophila*, *windpipe* (*wnd*) was recently discovered to have several mammalian orthologs, including three in *Rattus norvegicus* and one in *Homo sapiens* (6). This early developmental gene encodes a transmembrane, leucine-rich repeat protein that appears

abruptly during embryogenesis to form a peptide with both a transmembrane and extracellular domain that may facilitate protein-protein interactions (7). It was hoped that a more thorough investigation of the human *wnd* ortholog might also reveal early developmental functions important to the development or morphogenesis of tissues within the head or neck.

Based upon this information, the overall purpose of this project was to screen for the human *wnd* ortholog, a predicted gene (hmm30337), identified from a predicted transcript or known mRNA (NM_178549) encoding a hypothetical protein from the mammalian gene collection (MGC042493). The working hypothesis for this project was that uncommitted, non-differentiated mesenchymal stem cells might express mRNA specific for this predicted gene product.

## 2. Methods

### 2.1. Genomic Search

Using the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST), nucleotide databases were searched using the nucleotide query from the cDNA of the *Drosophila wnd* gene CG3414; chromosomal location 2R 58D 2-3 (6,8) to identify potential orthologs. The genomic sequence of *Drosophila wnd* was entered in FASTA format and submitted for sequence comparison. Homologous sequence matches were then aligned with the query sequence.

### 2.2. Primer Design

Using NCBI-BLAST, overlapping reverse transcriptase polymerase chain reaction (RT-PCR) primers were designed to span the coding region of the 8.53 kb predicted gene from the human ortholog (hmm30337) from human chromosome 1 (1q42.13). Negative control primers were designed for the predicted non-coding (bp 0-177), as well as the predicted coding regions (bp 177 – 1923). These primers were synthesized by SeqWright (Houston, TX):

Negative control primer sets:

NM178549.31F: CGGCTATTTATCCCCAGCTGCGG ; bp 78

NM178549.31R: GTTTTCCGGCTTTCGGGGTGT

NM178549.32F: CCAGCTGGAGCTTTGGTCCCGTA; bp 51

NM178549.32R: AAATTTCGCTGGGCAGGGTCCAG

Experimental primer sets:

BC042500.1F: CAGCTGGAGCTTTGGTCCCG; bp 182

BC042500.1R: GAATGCCAGTAGTCCCGTTTTCCG

BC042500.2F: AACCTGGTCTCCCTGGGACTTCCAG; bp 410

BC042500.2R: AAGTGCAAATTGTCAAGGCCCCAG

### 2.3. RNA Screening

To screen for mRNA expression using these primers, the following human cell lines screened for this study were obtained from American Type Culture Collection (ATCC: Manassas, VA): Human Gingival Fibroblasts, HGF-1 (CRL-2014), Human Tongue SCC15 (CRL-1623), SCC25 (CRL-1628) and CAL27 (CRL-2095), Normal Breast Hs578 Bst (CRL-1424) and MCF10A (10317), Breast Cancer Hs578T (CRL-2329), MCF7 (HTB-22) and T-47D (CRL-2865), Cervical Adenocarcinoma CaSKi (CRL-1550) and GH354 (CRL-13003), Normal Foreskin Fibroblast Hs27 (CRL-1634), Normal Colorectal Hs722 (CRL-7456) and Colorectal Cancer Hs722T (CRL-7273). Total RNA from adult tissue cell lines was isolated and screened for mRNA expression specific to this ortholog, as previously described (6,9). In brief, RNA was isolated from 1.5 x $10^7$ cells of each cell line using ABgene Total RNA Isolation Reagent (Epsom, Surrey, UK) in accordance with the procedure recommended by the manufacturer (10).

RT-PCR was then performed on total RNA using the ABgene Reverse-iT One-Step RT-PCR kit (ReadyMix Version) and a mastercycler gradient thermocycler (Eppendorf: Hamburg, Germany) using one microgram of template (total) RNA for each screening reaction. The reverse transcription step ran for 30 minutes at 47C, followed by denaturation for two minutes at 94C. Fifty amplification cycles were run, consisting of 20 second denaturation at 94C, 30 seconds of annealing and 6.5 minutes of extension at 72C. Annealing temperatures were 56.3C for NM178549.31, 58.8C for 178549.32, 59.1C for BC042500.1 and 57.4C for BC042500.2 primers. Final extension was run for five minutes at 72C and reaction products separated by gel electrophoresis using Reliant 4% NuSieve 3:1 Plus Agarose gels (Lonza: Rockland, ME). Bands were visualized by UV illumination of ethidium-bromide stained gels and captured using a Kodak Gel Logic 100 Imaging System and 1D Image Analysis Software (Eastman Kodak: Rochester, NY).

### 2.4. Dental Pulp Stem Cells

In addition to commercial cell lines, dental pulp stem cells (DPSC) were also screened. The original protocol for the DPSCs titled “Isolation of Non-Embryonic Stem Cells from Dental Pulp” at the University of Nevada, Las Vegas – School of Dental Medicine was reviewed and approved by the Office of Research Integrity / Human Subjects (OPRS protocol 0907-3148) in February 2010 (9,11). All cell lines used were previously found to express intracellular (NANOG, Oct4, Sox2, Klf4, Rnf12, Bin1) and cell surface (CD44, CD133) markers that are used to identify and characterize DPSC isolates (9, 11,12). Total RNA was submitted to Functional Biosciences (Madison, WI) for generation of RT-PCR reaction products (cDNA) and sequencing. Final cDNA sequence was aligned using BLASTN (13), the NCBI nucleotide alignment tool that provides sequence homology and gap information between two sets of nucleotide sequences (predicted, actual).

### 2.5. Northern Blot

In brief, RNA was isolated from 1.5 x $10^7$ cells of each cell line using ABgene Total RNA Isolation Reagent (Epsom, Surrey, UK) in accordance with the procedure recommended by the manufacturer (10), as described above. Total RNA was subsequently separated on 1% agarose gels and transferred to positively charged nylon membranes from Millipore (Bedford, MA) according to the manufacturer's instructions (10-15V for 90 mins). Membranes were cross-linked to the RNA (60C for 1 hr) and probed with non-radioactive DIG-labeled probes generated from the BC042500.1 primer set (Roche Applied Science, Indianapolis, IN). Following hybridization, the membranes were washed at 37C using low stringency buffer solution (2 x SSC, 0.1% SDS) for ten minutes and imaged.

## 3. Results

Using NCBI-BLAST, genomic databases were searched using the *Drosophila wnd* sequence for the primary query (Figure 1). This search revealed several mammalian homologs, including three in rats (*R. novegicus;* NW_047658.2, NW_047774.2, and NW_047399.2) and one in humans (*H. sapiens;* NM_178549.3). This ortholog corresponded to an 8565 base pair (bp) coding region containing three, distinct Open Reading Frames (ORFs), the first of which is a 1746 bp transcript encoding the 582 amino acid (68.1 kDa) hypothetical protein (MGC042493) – a potential zinc-finger transcription factor (ZN678).

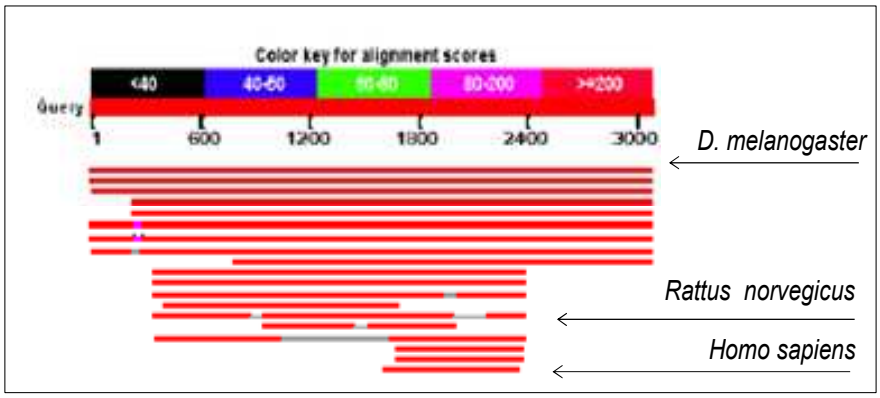


***Figure 1.*** *NCBI-BLAST search for Windpipe (wnd) gene orthologs. Several homologs were identified in D. melanogaster, as well as three predicted orthologs in R. norvegicus and one in H. sapiens.*

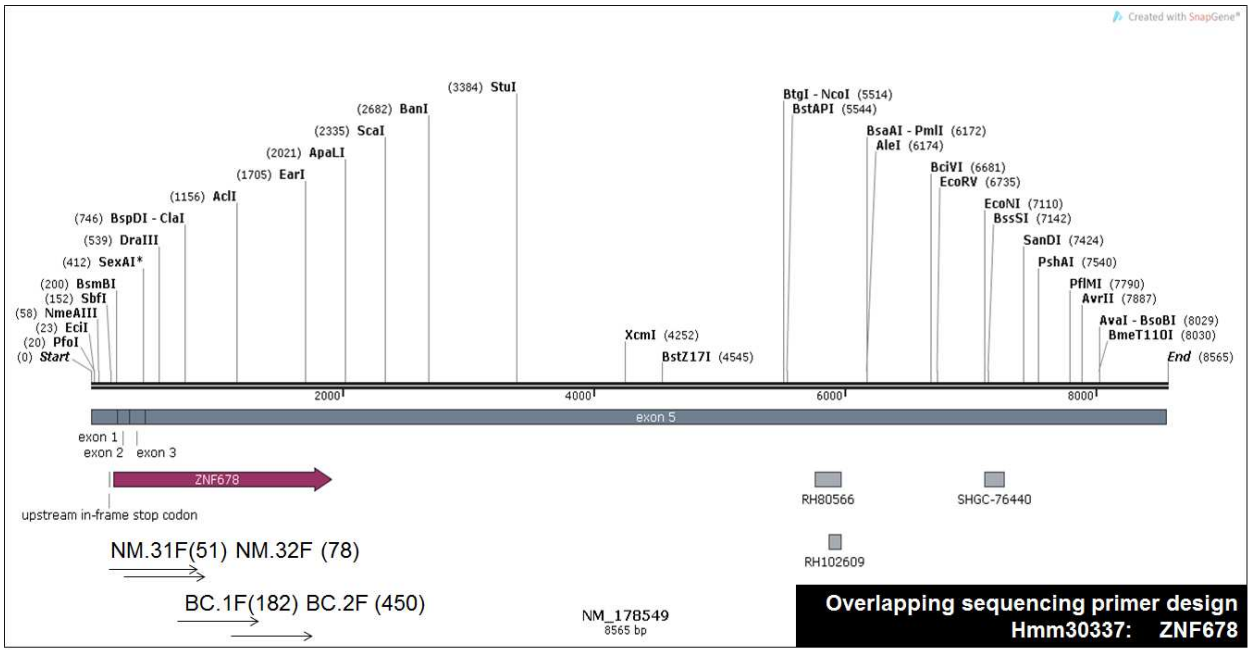


***Figure 2.*** *RNA screening for hmm30337 expression. Overlapping RT-PCR primers spanning the coding region of the 8.53 kb predicted gene were used to screen total RNA from adult tissue cell lines and dental pulp stem cells (DPSC).*

Using overlapping RT-PCR primers designed by NCBI-BLAST, spanning the coding region of the 8.53 kb predicted gene from the human ortholog (hmm30337), human cell lines were screened for mRNA expression. This screening revealed that no commercial cell lines available to this research group were found using this cDNA screening (Figure 2). This included oral gingiva, tongue, breast mammary and ductal tissues, cervical epithelium, foreskin fibroblasts, and

colorectal tissues. In addition, fifteen dental pulp stem cells (DPSC) were also screened, which revealed two positive results (BC042500.1) from DPSC-11836 and DPSC-11750 isolates. An analysis of the primer sites revealed these potential transcripts were generated from the experimental screening primers (BC04251110.1), corresponding a sequence site (bp 182) within the predicted coding region (bp 177-1943), however no results were obtained using the second experimental primer set (bp 410). No results were obtained from the negative control primer sets external to the predicted coding region (NM178549.31 bp 78; NM178549.32 bp 51).

Total RNA from DPSC-11836 and DPSC-11750, as well as experimental primer sets, were sent out for generation of RT-PCR reaction products (cDNA) and for subsequent sequencing (Figure 3). These results suggest the presence of mRNA (1746 bp) roughly corresponding with the predicted hmm30337 exon 3 at the predicted start (bp 177), which terminates at bp 1922. Although sequencing from the start codon (P1F, P1R) yielded a 552 bp sequence with low match (20/20; 523 unmatched), the final overlap (P3F, P3R) confirmed a 564 bp cDNA with 92% homology to the predicted sequence (459/500 matched). This provides some rationale for the negative result with the second experimental primer set, designed to anneal with bp 410. These results suggest an alternative DNA sequence is present within this coding region, whereas the original reported DNA sequence may be a cross-over or translocation event due to the high homology (99% homology with a zinc-finger pseudogene LOC100533641) on chromosome 12. Despite these sequence variations, the final sequence coding for a 582 amino acid peptide has high homology with the predicted zinc-finger transcription factor (ZN678).

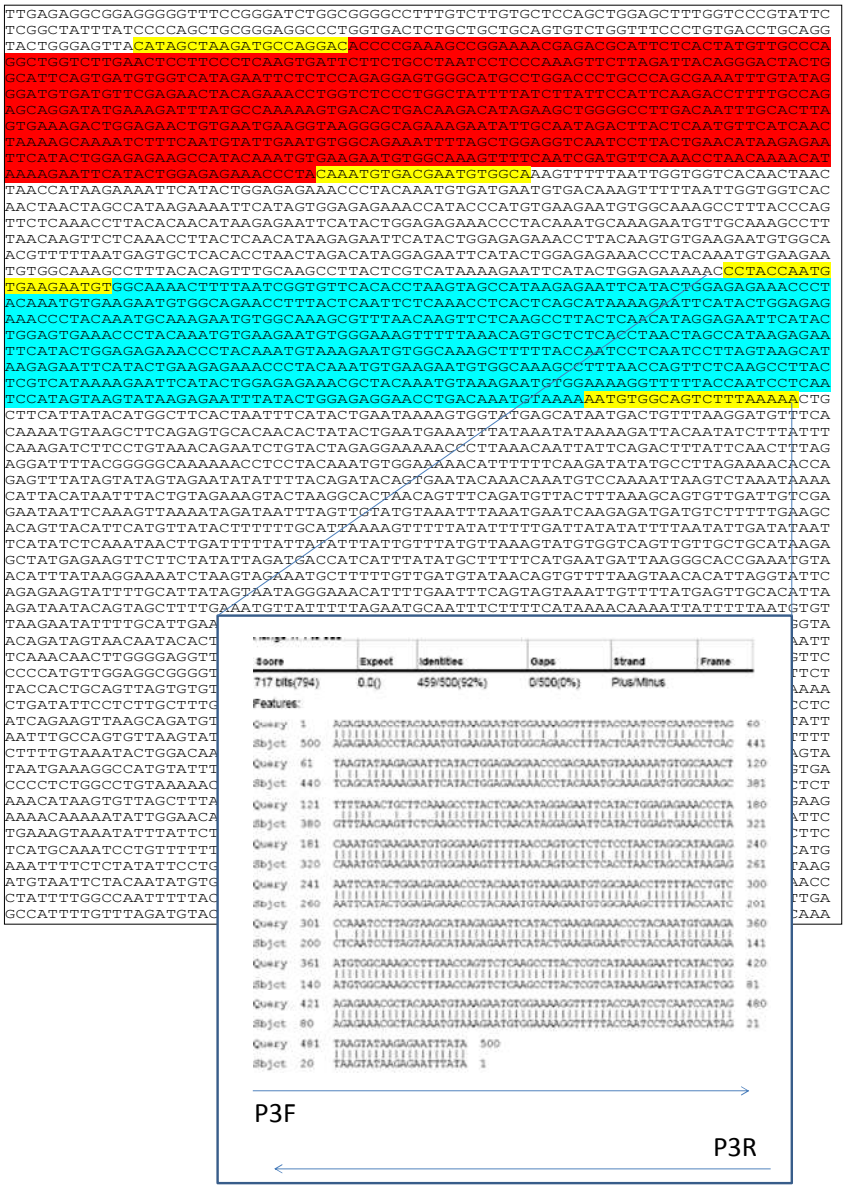


***Figure 3.*** *Sequence of predicted gene hmm30337. Total RNA from dental pulp stem cells (DPSCs) revealed a 1746 bp mRNA transcript (start bp 177, termination bp 1922) with areas of low and high homology to the predicted sequence. This transcript corresponds with a 582 amino acid zinc-finger transcription factor with high homology to the predicted ZNF678 gene product found on human chromosome 1 (1q42.13), as seen in Tables 1 and 2.*

***Table 1.*** *DNA sequence results Actual DNA sequence (1746 bp).*

ATGCCAGGACACCCCGAAAGCCGGAAAACGAGACGCATTCTCACTATGTTGCCCAGGCTGGTCTTGAACTCCTTCCCTCAAGTGATTCTTCTGC
CTAATCCTCCCAAAGTTCTTAGATTACAGGGACTACTGGCATTCAGTGATGTGGTCATAGAATTCTCTCCAGAGGAGTGGGCATGCCTGGACC
CTGCCCAGCGAAATTTGTATAGGGATGTGATGTTCGAGAACTACAGAAACCTGGTCTCCCTGGCTATTTTATCTTATTCCATTCAAGACCTTTT
GCCAGAGCAGGATATGAAAGATTTATGCCAAAAAGTGACACTGACAAGACATAGAAGCTGGGGCCTTGACAATTTGCACTTAGTGAAAGACT
GGAGAACTGTGAATGAAGGTAAGGGGCAGAAAGAATATTGCAATAGACTTACTCAATGTTCATCAACTAAAAGCAAAATCTTTCAATGTATT
GAATGTGGCAGAAATTTTAGCTGGAGGTCAATCCTTACTGAACATAAGAGAATTCATACTGGAGAGAAGCCATACAAATGTGAAGAATGTGG
CAAAGTTTTCAATCGATGTTCAAACCTAACAAAACATAAAAGAATTCATACTGGAGAGAAACCCTACAAATGTGACGAATGTGGCAAAGTTTT
TAATTGGTGGTCACAACTAACTAACCATAAGAAAATTCATACTGGAGAGAAACCCTACAAATGTGATGAATGTGACAAAGTTTTTAATTGGTG
GTCACAACTAACTAGCCATAAGAAAATTCATAGTGGAGAGAAACCATACCCATGTGAAGAATGTGGCAAAGCCTTTACCCAGTTCTCAAACCT
TACACAACATAAGAGAATTCATACTGGAGAGAAACCCTACAAATGCAAAGAATGTTGCAAAGCCTTTAACAAGTTCTCAAACCTTACTCAACA
TAAGAGAATTCATACTGGAGAGAAACCTTACAAGTGTGAAGAATGTGGCAACGTTTTTAATGAGTGCTCACACCTAACTAGACATAGGAGAA
TTCATACTGGAGAGAAACCCTACAAATGTGAAGAATGTGGCAAAGCCTTTACACAGTTTGCAAGCCTTACTCGTCATAAAAGAATTCATACTG
GAGAAAAACCCTACCAATGTGAAGAATGTGGCAAAACTTTTAATCGGTGTTCACACCTAAGTAGCCATAAGAGAATTCATACTGGAGAGAAA
CCCTACAAATGTGAAGAATGTGGCAGAACCTTTACTCAATTCTCAAACCTCACTCAGCATAAAAGAATTCATACTGGAGAGAAACCCTACAAA
TGCAAAGAATGTGGCAAAGCGTTTAACAAGTTCTCAAGCCTTACTCAACATAGGAGAATTCATACTGGAGTGAAACCCTACAAATGTGAAGA
ATGTGGGAAAGTTTTTAAACAGTGCTCTCACCTAACTAGCCATAAGAGAATTCATACTGGAGAGAAACCCTACAAATGTAAAGAATGTGGCA
AAGCTTTTTACCAATCCTCAATCCTTAGTAAGCATAAGAGAATTCATACTGAAGAGAAACCCTACAAATGTGAAGAATGTGGCAAAGCCTTTA
ACCAGTTCTCAAGCCTTACTCGTCATAAAAGAATTCATACTGGAGAGAAACGCTACAAATGTAAAGAATGTGGAAAAGGTTTTTACCAATCCT
CAATCCATAGTAAGTATAAGAGAATTTATACTGGAGAGGAACCTGACAAATGTAAAAAA TGTGGCAGTCTTTAAAAA

*83% match to Predicted DNA (BLASTN) with 0-1% gaps.

| Query (bp) | Identity | Gaps |
|---|---|---|
| 466-1649 | 989/1193 (83%) | 18/1193 (1%) |
| 634-1735 | 919/1110 (83%) | 16/1110 (1%) |
| 145-1483 | 856/1031 (83%) | 10/1031 (0%) |

***Table 2.*** *Amino acid sequence results.*

Predicted 582 amino acid (peptide) sequence (cDNA 1746 bp)

MPGHPESRKTRRILTMLPRLVLNSFPQVILLPNPPKVLRLQGLLAFSDVVIEFSPEEWACLDPAQRNLYRDVMFENYRNLVSLAILSYSIQDLLPEQDM
KDLCQKVTLTRHRSWGLDNLHLVKDWRTVNEGKGQKEYCNRLTQCSSTKSKIFQCIECGRNFSWRSILTEHKRIHTGEKPYKCEECGKVFNRCSNL
TKHKRIHTGEKPYKCDECGKVFNWWSQLTNHKKIHTGEKPYKCDECDKVFNWWSQLTSHKKIHSGEKPYPCEECGKAFTQFSNLTQHKRIHTGEKP
YKCKECCKAFNKFSNLTQHKRIHTGEKPYKCEECGNVFNECSHLTRHRRIHTGEKPYKCEECGKAFTQFASLTRHKRIHTGEKPYQCEECGKTFNRC
SHLSSHKRIHTGEKPYKCEECGRTFTQFSNLTQHKRIHTGEKPYKCKECGKAFNKFSSLTQHRRIHTGVKPYKCEECGKVFKQCSHLTSHKRIHTGEK
PYKCKECGKAFYQSSILSKHKRIHTEEKPYKCEECGKAFNQFSSLTRHKRIHTGEKRYKCKECGKGFYQSSIHSKYKRIYTGEEPDKCKKCGSL*K

Finally, to confirm these results, Northern blot analysis was performed on total RNA isolated from the DPSC isolates (Figure 4). These results demonstrate the presence of a transcript corresponding to an approximate size of 1.7 kb. No positive results were found in any other differentiated or cancerous cell lines evaluated.

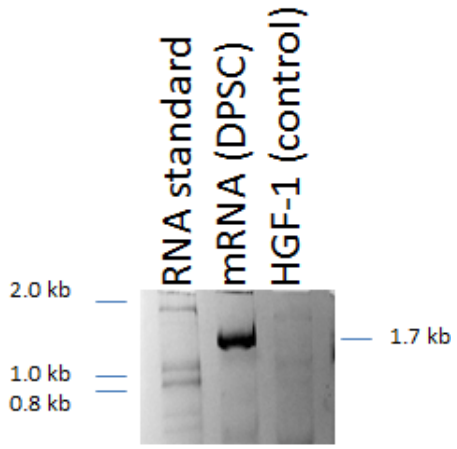


***Figure 4.*** *Northern blot analysis of hmm30337 mRNA expression. A transcript of approximately 1.5-1.7 kb was detected in RNA isolated from DPSC isolates. No other positive results were found among the differentiated or cancerous cell lines evaluated (HGF-1 normal oral gingiva shown).*

## 4. Discussion

Although many genes and pathways have been identified that facilitate the development of tissues within the head and neck, many unidentified pathways and genetic interactions have yet to be discovered (1-5). This project used an early development gene required for early tracheal development in fruit flies to search for potential correlates in other species (6,7), revealing that a previously predicted gene (hmm30337) may be expressed in a subset of uncommitted, undifferentiated dental pulp stem cells (DPSCs). Although the actual sequence obtained from cDNA analysis differed somewhat from the predicted sequence, the final confirmed sequence revealed a transcript coding for a zinc-finger transcription factor with high homology to the predicted protein (ZNF678) – although the final mRNA transcript of 1746 bp provides for a protein of only 582 amino acids, rather than 678.

Interestingly, the finding that one of the experimental primer sequences (BCC042500.2) did not produce any RT-PCR or Northern blot results provides some evidence to further support these findings. For example, the alternative sequence reported in this study significantly differs only within the first 650 base pairs, precisely where this primer set

would have annealed (bp 410). That the original predicted sequence bears homology to with *Homo sapiens* 12q BBAC RP11-58A17, a zinc finger protein 100 pseudogene (LOC1005336411) on chromosome 12 appears to suggest a possible cross-over or translocation event that could not have been predicted without direct sequence analysis of the potential transcript.

Although the function and characterization of this predicted protein has not yet been completed, the finding that only a small subset of DPSCs, but no differentiated (adult) cell lines, were found to express RNA corresponding with this predicted gene product further suggests that this transcription factor may be involved with a process of differentiation that may be observed in (but not limited to) DPSCs. As many zinc-finger transcription factors have already been demonstrated as essential for DPSC differentiation in vivo, these results suggest this protein may represent a novel transcription factor that may be involved in a subset of previously unidentified alternative pathways of wound repair or healing arising from DPSC residing in adult tissues (4-6).

## 5. Conclusions

The human genome project identified many possible gene products, which are yet to be confirmed. This project provides evidence that the predicted human gene (hmm30337), which would have encoded a 678 amino acid zinc-finger transcription factor (ZNF678) may actually be a 1746 bp mRNA sequence expressed in dental pulp stem cells (DPSCs) that encodes a 582 bp zinc-finger transcription factor (ZNF582) with high homology to the original, predicted sequence from chromosome 1 (1q42.13) named *windpipe* homolog 1 or WH1. Further studies will be needed to sequence and characterize this protein, as well as the potential functions within these cell types.

## Acknowledgments

The authors would like to thank Functional Biosciences for their technical assistance with this project. SM would like to thank the Graduate and Professional Student Association (GPSA) for partial funding of some aspects of this project.

## References

[1] Rué P, Martinez Arias A. Cell dynamics and gene expression control in tissue homeostasis and development. *Mol Syst Biol*. 2015, 11(2):792. Review. PMID: 25716053

[2] Tan J, Xu X, Lin J, Fan L, Zheng Y, Kuang W. Dental Stem Cell in Tooth Development and Advances of Adult Dental Stem Cell in Regenerative Therapies. *Curr Stem Cell Res Ther*. 2014 [Epub ahead of print] PMID: 25274563

[3] Zhang YD, Chen Z, Song YQ, Liu C, Chen YP. Making a tooth: growth factors, transcription factors, and stem cells. *Cell Res.* 2005, 15(5):301-16. PMID: 15916718

[4] Viale-Bouroncle S, Felthaus O, Schmalz G, Reichert TE, Morsczeck C. Transcription factors for dental stem cell differentiation. *Int J Oral Maxillofac Implants*. 2013, (6):e478-86. PMID: 24278957

[5] Strausberg RL et al.; Mammalian Gene Collection Program Team. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A*. 2002, 99(26):16899-903. PMID: 12477932

[6] Huff JL, Kingsley KL, Miller JM, Hoshizaki DK. Drosophila windpipe codes for a leucine-rich repeat protein expressed in the developing trachea. *Mech Dev.* 2002, 111(1-2):173-6. PMID: 11804792

[7] Kobe B, Deisenhofer J. The leucine-rich repeat: a versatile binding motif. *Trends Biochem Sci.* 1994, 19(10):415-21. Review. PMID: 7817399

[8] Adams et al. The genome sequence of Drosophila melanogaster. *Science*, 287 (2000), pp. 2185–2195. PMID: 10731132

[9] Loveland K, Young A, Khadiv M, Culpepper M, Kingsley K. Dental Pulp Stem Cell (DPSC) Pluripotency Enhanced by Transforming Growth Factor (TGF-β1) in Vitro may be Inhibited by Differentiation-Inducing Factors Laminin-5 and Dexamethasone. *International Journal of Biological Sciences and Applications.* 2014, 1(3): 55-61.

[10] Davis JE, Freel N, Findley A, Tomlin K, Howard KM, Seran CC, Cruz P, Kingsley K.A molecular survey of S. mutans and P. gingivalis oral microbial burden in human saliva using relative endpoint polymerase chain reaction (RE-PCR) within the population of a Nevada dental school revealed disparities among minorities. *BMC Oral Health*. 2012, 12:34. PMID:22925755

[11] Hung E, Lee S, Fitzgerald B, Hill CK, Kingsley K. Dental pulp-derived stem cell (DPSC) survival and viability may correlate with specific patient demographics. *Forum for Dental Student Research and Innovation* (*FDSRI*) Winter 2013, 1(3): 14-21.

[12] Bojic S, Volarevic V, Ljujic B, Stojkovic M. Dental stem cells--characteristics and potential. *Histol Histopathol.* 2014, 29(6):699-706. PMID:24446280

[13] Zhang Z, Schwartz S, Wagner L, Miller W. A greedy algorithm for aligning DNA sequence. *J Comput Biol* 2000, 7(1-2):203-14.