A. Letra[1,2]*[†], M. Zhao[1†], R.M. Silva[1,2],
A.R. Vieira[3], and J.T. Hecht[1,2]

[1]Craniofacial Research Center, University of Texas Health Science Center at Houston School of Dentistry, Houston, TX, USA; [2]Department of Pediatrics, Pediatric Research Center, University of Texas Health Science Center Medical School at Houston, Houston, TX, USA; and [3]Departments of Oral Biology and Pediatric Dentistry, University of Pittsburgh School of Dental Medicine, Pittsburgh, PA, USA; [†]joint first authors; *corresponding author, ariadne.m.letra@uth.tmc.edu

Functional Significance of *MMP3* and *TIMP2* Polymorphisms in Cleft Lip/Palate
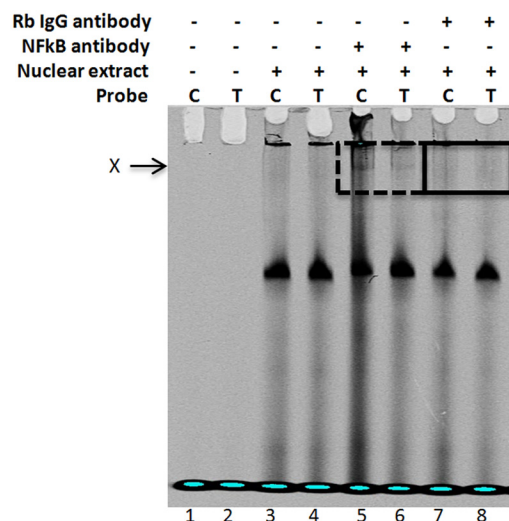
APPENDIX

Appendix Figure. Electrophoretic mobility shift assays and supershift assays of *TIMP2* rs8179096. Double-stranded, infrared-labeled C and T probes incubated in the absence (lanes 1, 2) or presence of cell nuclear extract (lanes 3-8). NFκB antibody was added in lanes 5 and 6 for supershift assay (in dotted frame). Rabbit IgG antibody was added in lanes 7 and 8 as isotype control.

MATERIALS & METHODS

Reporter Constructs

Constructs and reagents were purchased from SwitchGear Genomics (Menlo Park, CA, USA). A reporter construct showing 1.057 kilobases of the *MMP3* gene promoter containing the rs522616 A/G variant was generated for the A allele and used for site-directed mutagenesis to obtain the G allele (constructs *MMP3*_A and *MMP3*_G). Next, reporter constructs containing the rs522616 A/G and rs3025058 5A/6A variant alleles combined were generated by cloning the promoter segments from -2074 kilobases upstream to the transcription start site, into a

pLightSwitch_Prom vector (SwitchGear Genomics, Carlsbad, CA, USA). Constructs were generated with the ancestral alleles; alternate alleles were obtained by site-directed mutagenesis (constructs *MMP3*_5A_A, *MMP3*_5A_G, *MMP3*_6A_A, and *MMP3*_6A_G). Reporter constructs for the *TIMP2* rs8179096 C/T variant were generated showing a 0.979-kilobase pair of the gene promoter (constructs *TIMP2*_C and *TIMP2*_T). Insert sequences are presented below. Cloning efficiency for all constructs was verified by direct sequencing. GenBank sequences for *MMP3* (NG_012100.1) and *TIMP2* (NC_000017.10) genes were used as references.

Reporter Constructs Insert Sequences and Respective Blasting Results, *per* the NCBI Database

Sequence of inserts for MMP3 rs3025058 5A[6A]/rs522616 A[G] **(base changes in bold):** TGACTCTTCCTGGGAC TTGGGAAACATCTTTCAGGGGAAGTCGTCGAA GCTGTTTTAAAATATAGCAAACTTTTGTATTTAGTTCAG GAACAGCATGGCCCATTTTGCCAATCACATCTTAACAG TTGGAAAAGCAAACATATTATCTATCAGGCTTTCCTCT AAACTTTAAATATGTTTTATAAGTTATAACTCCAGAGAA AATTTACAAAGGATAAACCTTAATATAGAAGGAATTA GAGCTGCCACAGCTTCTACACTTTTAACCTCTCAATAT TTTATCTGTTGGGCTCCACTGTTTCTTCCTGGAATTCAC ATCACTGCCACCACTCTGTTCTCCTTGTCCTCATATCA ATGTGGCCAAATATTTTCCCTGTATTTCAATCAGGACAA GACATGG**TTTTT[T]**CCCCCCATCAAAGGAATGGAGAA CCATAGAATACTAGTTTTAAAATGTCTTTAGGCCAGGT GCCGTGACCCATGTCTGTAATCCTAGCACTTTGAGAGG TTGAGGCAGGAGAATCACTTGAGCCCAGAGCTCGAA ACCAGCCTGGGCAACATAGTGAAACCTCTGTCTCTATT TTTTAAATAAAATTTGAAAAAGTCTTTAGACATAATCTA GTCTAAAAATGAAGGCTTAAATGTGATGTATAGCCCCC TGCCAAGTGGCTATCAACTGTGTGGGCATCTTCAGTCA TAGGGATCTTATTGCCACAGAGAAATCCCTTTAAACTT ATTGGGTAAAATCTCTCCAATGTTTATTAAGAAACACA CAAAAAATAAAGCAAAGAAGAAAATGCAAAGAGTT ATAAATGAGAGGAAGCAAAATGGGCACTTATTAAAGG TCTAATAAATGCACATTTGTATCCATCATTCTACTGAGT

TCTTACTCCCAAGATGTTCTTCCCTTTAGCAAACAAAT
AAGCAAGTCAGCAAAGAAAGAAAGAACAAACAAAAT
GTGGTGATCAGGGAAGCATTGAGGAGATGGATGGTG
GCAGGTGGCAAGAGGACTATAAAAGTTTTACAAAATG
TCTTCCTCTGAATATGTTTAGAGTCTTGCATTCAAGCAT
TTATTATACACCAATAATGTGAGCAACACTTTACTTGAC
AAAGAAACAGAAAAGAAAGGAAAGGAAGAAAACAG
AAGAGCATGAAGAGAAAATTTAGGATGGATTCTGTTC
TTCAACTTCAAAGCATCTGCTAATTTGAATTTAGGGAG
GAGGGGAAAAGGTTGAAAGAGAATAAGACATGTGTA
GAAGACAAGGACAGAGAGAATTTCAGTCCGGTAAGC
AATGTAATTCATTTCA**A[G]**TTCTACAACTATTTATGGAG
CAGCTACGTGGGCCCATCACCCATTAATAAATTGGTTA
CAGAATTAAAACCAACCCAAAGGGAATATACTTCCTT
CTTTTTCACAGACCCTCTTTGTTCTATTCTGCCCATGA
GGTTTTCCTCCTCAAGAACCAGCAAATCCAACGACAG
TCAATAGCAGGCATTACAAATCAGATTCAGAAAAA
TAAATCACCCCTTCTAAATTTCTTCTAGATATTATCTT
TTATGTTTTGAGTATAATTGTATATAGTATAGACTATAGC
TATGTATGTACACTTTCCACTTACATCTTTTATTTGCTTT
TATAATGTGTTTCTTAAAATAAAACTGCTTTTAGAAGTT
CTGCACAATTCTGATTTTTACCAAGTCAACCTACTTCTT
CTCTCAAAAGGACAAACATAAATTGTCTAGTGAATTCC
AGTCAATTTTTCCAGAAGAAAAAAATGCTCCAGTTTT
CTCCTCTACCAAGACAGGAAGCACTTCCTGGAGATTA
ATCACTGTGTTGCCTTGCAAAATTGGGAAGGTTGAGA
GAAATTAGTAAAGTAGGTTGTATCATCCTACTTTGAATT
TGGAATGTTTGGAAATGGTCCTGCTGCCATTTGGATGA
AAGCAAGGATGAGTCAAGCTGCGGGTGATCCAAACAA
ACACTGTCACTCTTTAAAAGCTGCGCTCCCGAGGTTG
GACCTACAAGGAGGCAGGCAAGACAGCAAGGCATAG
AGACAACATAGAGCTAAGTAAAGCCAGTGGAA

Result of Inserts MMP3 rs3025058 5A[6A]/rs522616 A[G] Blasted to the Human Genome, *per* the NCBI Database
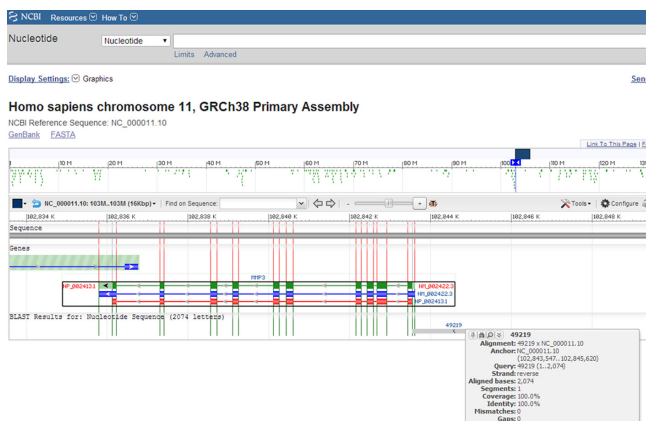
Sequence of Inserts for MMP3 rs522616 A[G] (base changes in bold):

GCAGGTGGCAAGAGGACTATAAAAGTTTTACAAAATG
TCTTCCTCTGAATATGTTTAGAGTCTTGCATTCAAGCAT
TTATTATACACCAATAATGTGAGCAACACTTTACTTGA

CAAAGAAACAGAAAAGAAAGGAAAGGAAGAAAACA
GAAGAGCATGAAGAGAAAATTTAGGATGGATTCTGTT
CTTCAACTTCAAAGCATCTGCTAATTTGAATTTAGGGA
GGAGGGGAAAAGGTTGAAAGAGAATAAGACATGTGT
AGAAGACAAGGACAGAGAGAATTTCAGTCCGGTAAG
CAATGTAATTCATTTCA**A[G]**TTCTACAACTATTTATGGA
GCAGCTACGTGGGCCCATCACCCATTAATAAATTGGTT
ACAGAATTAAAACCAACCCAAAGGGAATATACTTCCTT
CTTTTTCACAGACCCTCTTTGTTCTATTCTGCCCATGAG
GTTTTCCTCCTCAAGAACCAGCAAATCCAACGACAGT
CAATAGCAGGCATTACAAATCAGATTCAGAAAAATAAA
TCACCCCTTCTAAATTTCTTCTAGATATTATCTTTTATGT
TTTGAGTATAATTGTATATAGTATAGACTATAGCATGTA
TGTACACTTTCCACTTACATCTTTTATTTGCTTTTATAAT
GTGTTTCTTAAAATAAAACTGCTTTTAGAAGTTCTGCA
CAATTCTGATTTTTACCAAGTCAACCTACTTCTTCTCTC
AAAAGGACAAACATAAATTGTCTAGTGAATTCCAGTC
AATTTTTCCAGAAGAAAAAAATGCTCCAGTTTTCTCC
TCTACCAAGACAGGAAGCACTTCCTGGAGATTAATCA
CTGTGTTGCCTTGCAAAATTGGGAAGGTTGAGAGAAA
TTAGTAAAGTAGGTTGTATCATCCTACTTTGAATTTGGA
ATGTTTGGAAATGGTCCTGCTGCCATTTGGATGAAAGC
AAGGATGAGTCAAGCTGCGGGTGATCCAAACAAACAC
TGTCACTCTTTAAAAGCTGCGCTCCCGAGGTTGGACCT
ACAAGGAGGCAGGCAAGACAGCAAGGCATAGAGACA

## Result of Inserts MMP3 rs522616 A[G] Blasted to the Human Genome, *per* the NCBI Database
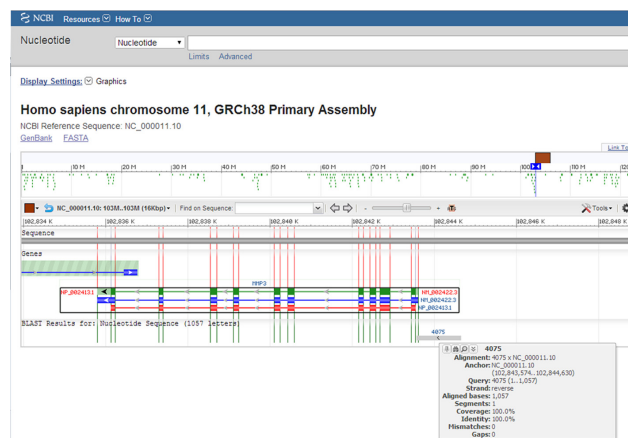
Sequence of Inserts for TIMP2 rs8179096 C[T] (base changes in bold):

TCACCACCCTACACATGTTCATCCCTTCATATGCCTGGG
TCTTTCCTGGAACACAAAACTGTTTTGGGAAAAGTCC
TTCGGCTCCCACATTTGTGGACTAAGAGAGGAAACTG
TGAGCGGAACCCAGCCAATGCCTCTGCTGCGATCCTA
CTGGCTCCTGGGCGCCTGGGCCCACCCCCGTCTCTTGT
TGGCTGGTCAAAAATATGGCCAGTTTATATAAAATCCT
GTTTTGTTCACAGTAACCACACCCCCCACCCCCCAACT
AAACTGGCCAGGCGCACTTAAAATTCTAAGGCCTCCA
TTTGAAAAAGGGATCCTGTCAGTTTCTCAATAGGCCA
CCCGCCCACAGAAACGGGGAGGTGGCGACAGGGAAC
GGCCCCTGCTCCAAAGGACACCCCTTGGCTCGCCCCG

```
AGGCTGGGCTCGAAGGGACCCCGGGGTGGCGGGGGA
CGGAGCAGCGTAGCCCTCCAGAGTCGAGCTGAAGGG
GAAAGGGTAGCGGGTGGGTCGCCTGGTGTCCTGGAAG
AACGGGGCGCGAGTCCCCGCGCTGAGTCAGGGACC
CCGGGCGCAGAAGGCCACGCAGCGGGGACCGGGGTC
GGGGGGCTGGGGGCGTCCGGGGACCCCCGCGCGGGT
GCGGGTCGCGGGCGCCAGGTGGTGC[T]GGGAAGCCC
CCGACGTGCCAGGCCGGGCACAACAAAAGCGCGGGC
TGGGGGGAGGCGCGGGCGGAGGGGGAGGAGGGGGC
TGCTGGGGAGCGCCCAGAGCCTGCATTGGCCGCCAGCC
ACCGGGAGGAGGAGCAGAAAATCCTCCGAGCGCAATA
AAACTGCGGCCCGGCCCAAGCCCGCAGCAAACACATC
CGTAGAAGGCAGCGCGGCCGCCGAGAACCGCAGCGC
CGCTCGCCCGCCGCCCCCCACCCCGCCGCCCCGCCCG
GCGAATTGCGCCCCGCGCCCCTCCCCTCGCGCCCCCG
AGACAAAGAGGAGAGAAAGTTTGCG
```

Results of Inserts TIMP2 rs8179096 C[T] Blasted to the Human Genome, *per* the NCBI Database
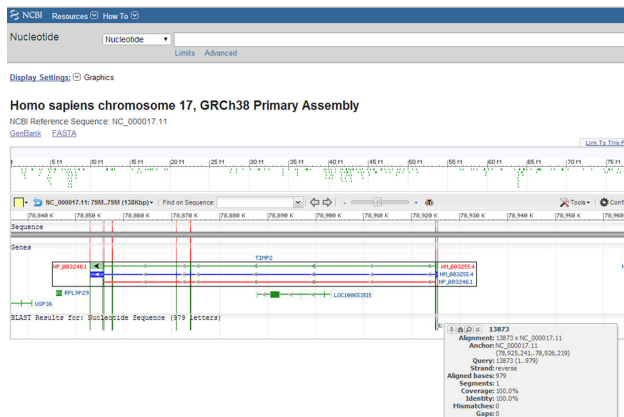
Appendix Table. Oligonucleotide Sequences for EMSAs

Oligonucleotide	Sequence
MMP3 rs522616 A	TTCATTTCA[A]TTCTACAA
MMP3 rs522616 G	TTCATTTCA[G]TTCTACAA
TIMP2 rs8179096 C	CAGGTGGTG[C]GGGAAGCC
TIMP2 rs8179096 T	CAGGTGGTG[T]GGGAAGCC
NFκB consensus oligonucleotide	AGTTGAG**GGGACTTTCCC**AGGC*
Mutated NFκB consensus oligonucleotide	AGTTGAG**GGCACTTTCCC**AGGC*

*Bolded nucleotides represent binding site to NFκB protein.

EMSA for Verification of NFκB Binding to TIMP2 rs8179096

To further confirm that C/T probe binds to NFκB protein, we performed additional EMSA experiments using wild-type or mutant NFκB consensus-binding oligonucleotide. Double-stranded, infrared-labeled C and T probes were incubated in the absence (lanes 1, 2) or presence of cell nuclear extract (lanes 3-12). One- or two-fold molar excess (in comparison with probe concentration) of NFκB consensus-unlabeled oligonucleotide and mutant NFκB consensus-unlabeled oligonucleotide was added in the reaction as competitors.

Proteins Identified by Mass Spectrometry Analyses for MMP3 rs522616 A Allele

The following figures reflect the proteins identified as exclusively bound to rs522616 A, showing over 95% probability levels (in green). The last 2 columns (blank) show that these 34 proteins were not found bound to the G allele or to a negative control band.