

Abstract

In humans and most mammals the amelogenin gene has been found embedded in the first intron of ARHGAP6 -about 40 Kb upstream of a conserved 160 bp exon (human exon 2). Amelogenin is not very conserved between species and is usually found embedded or upstream of ARHGAP6 in humans and other mammals. Recent sequencing of the anole genome allowed for analysis of alternative first exon-promoter regions.

Background

- The amelogenin protein is responsible for amelogenesis.
- ARHGAP6 is a rho GTPase activating protein responsible for actin remodeling and consist of thirteen exons (humans).
- Amelogenin is a gene that has been found nested or embedded inside of the ARHGAP6 gene with reverse transcriptional orientation in humans and other mammals.
- ARHGAP6 in humans has 10 alternative mRNA transcripts, 8 of which include the region containing AMEL.
- ARHGAP6 exon 2 is highly conserved among species, however exon 1 is not!!
- The human gene contains several different exon 1 variants.

Hypothesis

The amelogenin gene is embedded in the ARHGAP6 gene in the *Anolis carolinensis* genome.

Methods

Promoter and TSS Identification

Eukaryotic Promoter Database was used to locate promoter regions, and promoter sequences were identified for each. (TATA, CCAAT, GGGCC, ATTGG). Human ARHGAP6 variant transcripts were downloaded from ACEVIEW.

Sequence Alignments

Sequences were aligned using NCBI BLAST. Human ARHGAP6 variants to scaffold and exon 1 and exon 2 alignments were done in blastn. Transcript alignments were compared to scaffold, exon 1 and 2 alignments were done in megablast.

Mapping & Charting

Alignments for exon by exon comparisons were charted for analysis.

Results

Anole unplaced scaffold GL 343702.1

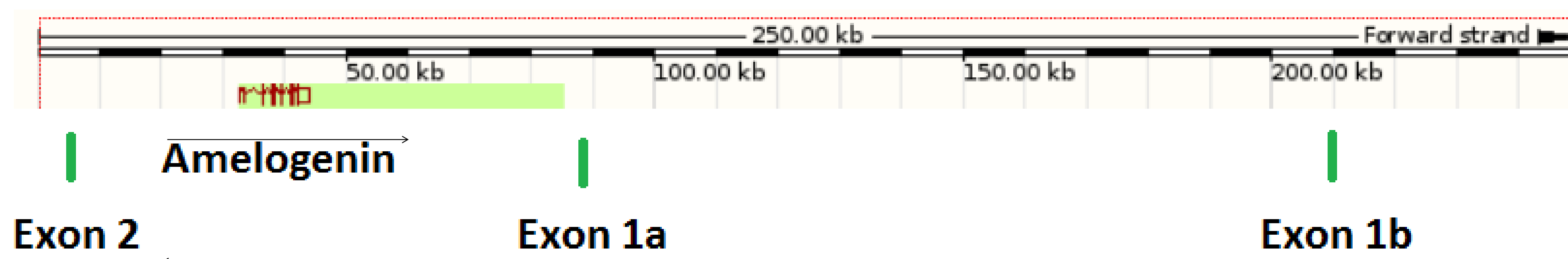


Figure 1. Anole unplaced scaffold

Exon 2- identified in both transcripts
Exon 1a- transcript GAFZ01196792.1
Exon 1b- transcript GAFZ01196793.1

Alternative Transcriptional Start Sites and associated Promoters Human ARHGAP6

Transcript name	Length (bp)/ mRNA matching the genome	Length (aa)/ Best predicted protein	Exon 1	Location start	Location end	Exon 2	Location start	Location end	Intron 1 size	Intron start	Intron end	Total # of exons
Promoter region				11,684,321	11,683,722							
aAug10	5,118	974	1,461	11,683,821	11,682,361	160	11,272,827	11,272,668	409,533	11,682,360	11,272,828	13
dAug10	3,632	765	1,461	11,683,821	11,682,361	160	11,272,827	11,272,668	409,533	11,682,360	11,272,828	11
Promoter region				11,683,703	11,683,104							
Promoter region				11,683,448	11,682,849							
fAug10	6,299	658	843	11,683,203	11,682,361	160	11,272,827	11,272,668	409,533	11,682,360	11,272,828	14
Promoter region				11,446,393	11,445,794							
bAug10	3,885	794	232	11,445,899	11,445,668	160	11,272,827	11,272,668	172,840	11,445,667	11,272,828	13
Promoter region				11,370,156	11,369,557							
iAug10	1,795	467	151	11,369,656	11,369,506	160	11,272,827	11,272,668	96,678	11,369,505	11,272,828	12
Promoter region				11,309,098	11,308,499							
hAug10	2,267	601	96	11,308,598	11,308,503	160	11,272,827	11,272,668	35,675	11,308,502	11,272,828	11
Promoter region				11,284,624	11,284,025	Overlapping sequence						
cAug10	3,995	771	338	11,284,143	11,283,806	160	11,272,827	11,272,668	10,978	11,283,805	11,272,828	13
gAug10	2,104	604	134	11,284,123	11,283,990	160	11,272,827	11,272,668	11,162	11,283,989	11,272,828	14
Promoter region				11,200,694	11,200,095							

Table 1. Alternative transcriptional start sites and associated promoters Human ARHGAP6

Discussion

Currently in the annotated genome, the region containing ARHGAP6 appears to be divided between ch 3 and the unplaced scaffold which contains amelogenin.

ARHGAP6 has several identified promoters

BLAST comparisons: Promoter regions to transcripts higher similarities in the conserved exon 2 vs exon 1.

Based on alignments and comparative analysis amelogenin is embedded in the first intron of the ARHGAP6 gene in the *Anolis carolinensis*.

The unplaced scaffold GL 343702.1 should be annotated on ch 3.

Future Directions

- Align all human exons with transcripts.
- Correspond with The Broad Institute to get unplaced scaffold GL 343702.1 annotated.
- Further analysis of transcripts that were identified but don't contain exon 2.

References

- Gibson CW, Thomson NH, Abrams WR, Kirkham J. 2005. Nested genes: Biological implications and use of AFM for analysis. *Gene* 350(1):15-23.
- Zhou YL and Snead ML. 2000. Identification of CCAAT/enhancer-binding protein alpha as a transactivator of the mouse amelogenin gene. *J Biol Chem* 275(16):12273-80.
- Salido EC, Yen PH, Koprivnikar K, Yu LC, Shapiro LJ. 1992. The human enamel protein gene amelogenin is expressed from both the X and the Y chromosomes. *Am J Hum Genet* 50(2):303-16.
- Nguyen T. 2012. The Amelogenin Gene in *Anolis Carolinensis*, Research report prepared for Biology 4351, Texas Wesleyan University
- NCBI Database: <http://www.ncbi.nlm.nih.gov/gene/395>
Ensembl database: <http://www.ensembl.org/index.html>
EPD: http://epd.vital-it.ch/promoter_elements.php
Aceview: www.ncbi.nlm.nih.gov/IEB/Research/Acembly/

Acknowledgements: Thank you to Dr. Clark for all of your help this semester!!