Great ape-specific *ATF4* retrocopies may act to regulate wild type *ATF4* activity and viral infections

Hans M. Dalton, Katie G. Owings, Nels C. Elde, Clement Y. Chow
Department of Human Genetics, University of Utah School of Medicine, Salt Lake City, UT, United States of America

Multiple stress signals trigger the integrated stress response (ISR) - including endoplasmic reticulum stress and viral infection. Dysregulation of the ISR is implicated in diseases such as cancer and diabetes. The ISR activates kinases that phosphorylate eIF2alpha, causing strong inhibition of protein synthesis machinery. Despite this, certain genes contain specialized upstream Open Reading Frame (uORF) regulatory sequences in their 5'UTR region that allow, or even upregulate, translation during the ISR. One of the most upregulated uORF-containing transcripts is *ATF4* (**Fig. 1**), a conserved transcription factor that activates pro-survival or apoptotic genes. We have found conservation of multiple *ATF4* retrocopies - labeled in humans as pseudogenes *ATF4P1-P4* (**Fig. 2**), while *hATF4P3* and *P4* are ape-specific (**Fig. 3-5**).

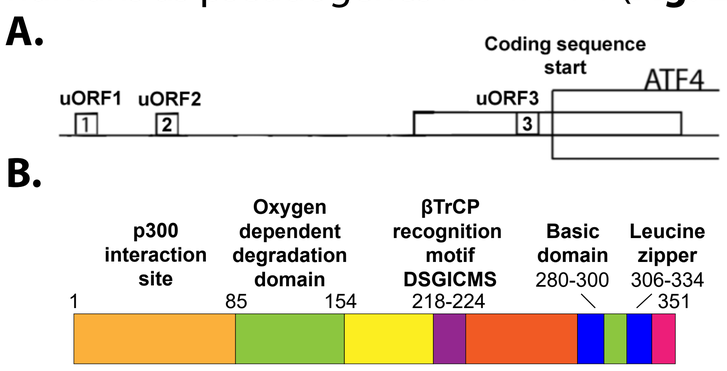

Figure 1 - Basics of ATF4 structure. A. Upstream Open Reading Frames (uORFs) are start codon-containing segments in the 5'UTR of the *ATF4* transcript. Normally, uORF3 inhibits *ATF4* translation; under stress, uORF3 is skipped and *ATF4* is translated. **B.** ATF4 is a leucine zipper transcription factor with two main degradation sites and a p300 interaction site that stabilizes and improves its transcriptional activity.

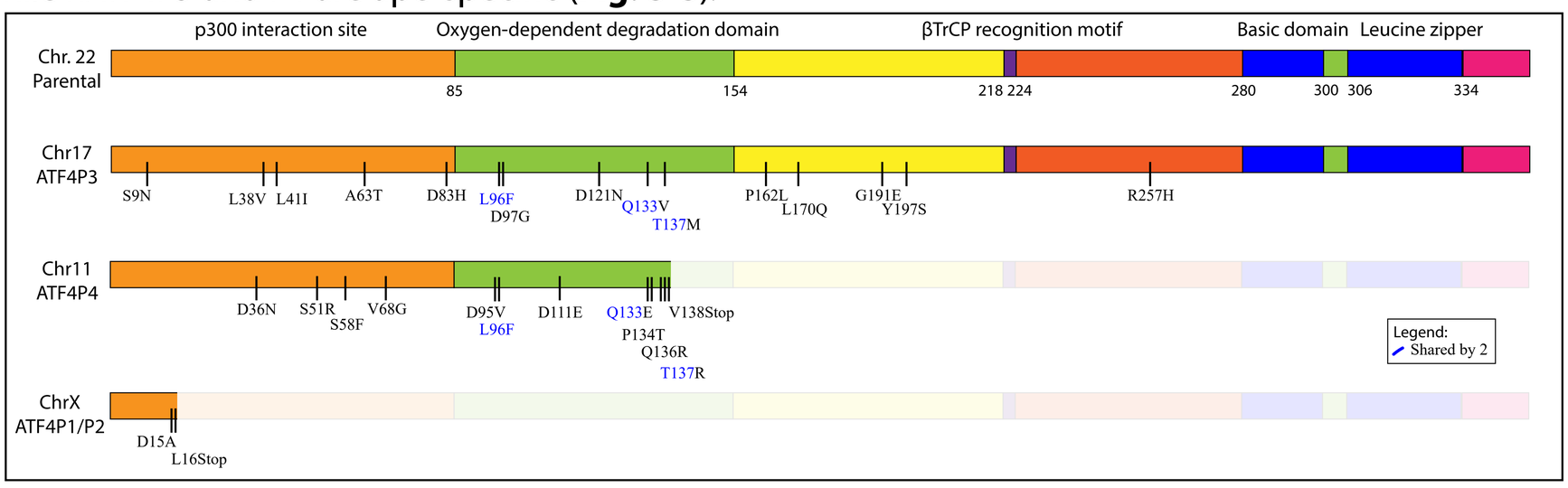

Figure 2 - Humans have 4 retrocopies of *ATF4*. Each has been previously identified as a pseudogene (P1-P4). *ATF4P3* contains no early stops (a "full length" copy). *ATF4P4* has an early stop midway through, upstream of the DNA binding domain. *ATF4P1* and *P2* are duplicates of each other with a very early stop. Amino acid changes from the parental copy are listed. Each retrocopy lacks introns (as expected), but all UTR's are intact.

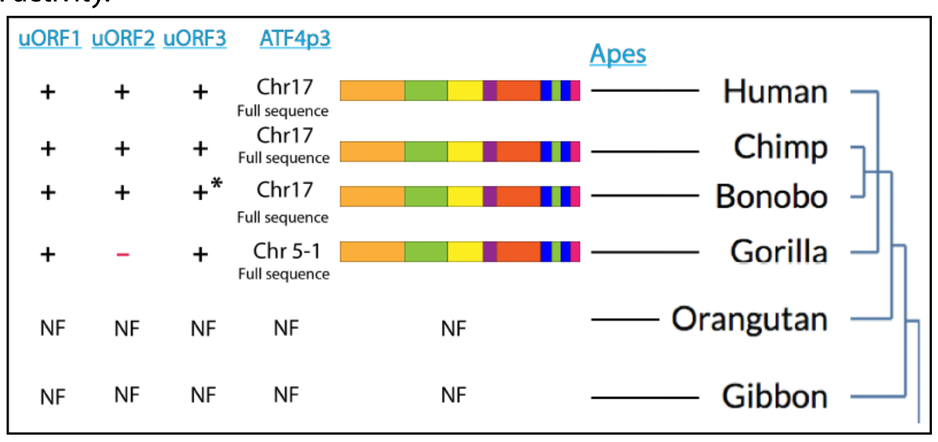

Figure 3 - Full-length copy. *hATF4P3* is a full length copy that maintains all domains and uORF regulatory regions (Gorilla losing 1 short uORF) with 95% amino acid (AA) identity with the parent *ATF4* gene. The 5% AA that are diverged from the parent gene are primarily in the N-terminal side, with no differences in the C-terminal DNA binding domain. * = new early stop, but still overlaps CDS.
Key: **+** = Maintained, **-** = Disrupted, NF = Not Found

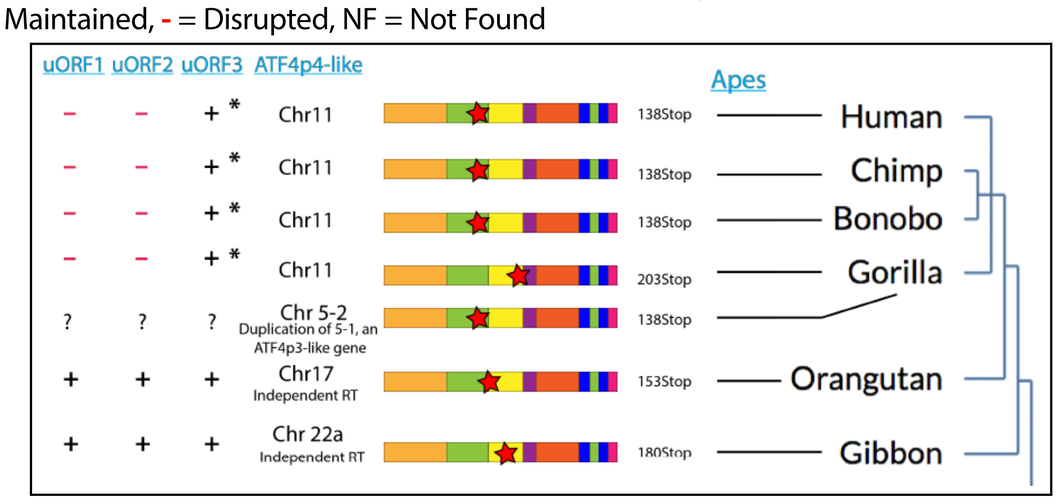

Figure 4 - Truncated copy. *hATF4P4* which has a truncation event removing its DNA binding domain, yet maintains 90% AA identity with the parent gene in the truncated half. *hATF4P4* is well conserved in humans, chimps, bonobos, and gorillas (Homininae) and ancestral to their most common ancestor. Strikingly, independent retrocopies with similar truncation events have occurred in orangutan, gibbon, and gorilla genomes (an extra gorilla copy). * = Interrupted by new uORF2 Stop codon in Homininae. Key: **+** = Maintained, **-** = Disrupted, ? = Unknown

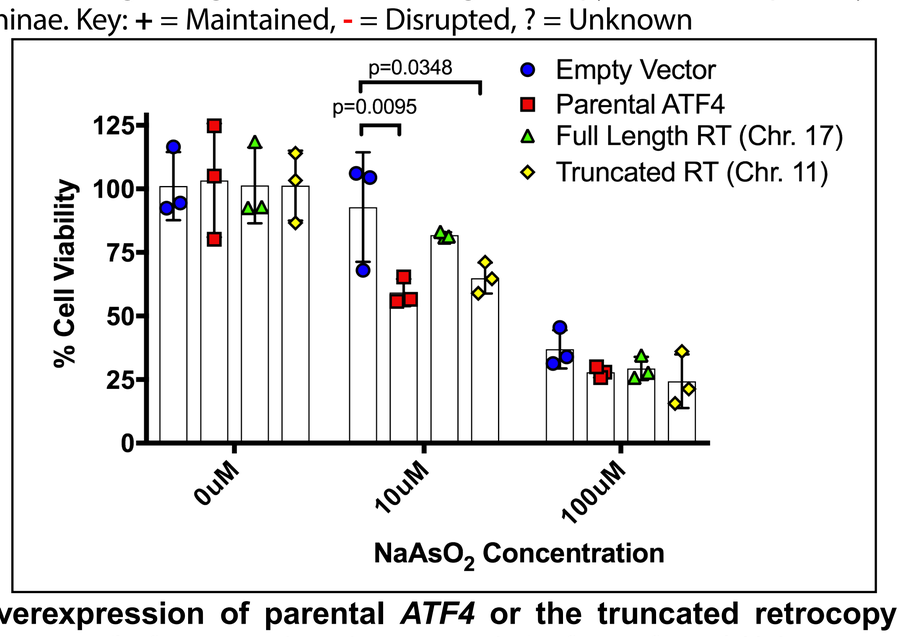

Figure 6 - Overexpression of parental *ATF4* or the truncated retrocopy reduces cell viability. Initial pools (before selecting single colonies) of transfected HeLa cells transfected with either empty vector, parental, full-length, or truncated ATF4 were subjected to 24hrs of Sodium Arsenite ($NaAsO_2$). The MTT assay was used to determine cell viability. At 10uM of Sodium Arsenite treatment, pools with increased parental or truncated ATF4 had reduced cell viability. Statistics: 2-way ANOVA, Dunnet's multiple comparisons test. Note: experiment done without secondary confirmation of transfection (primary being the selection media) - in progress.

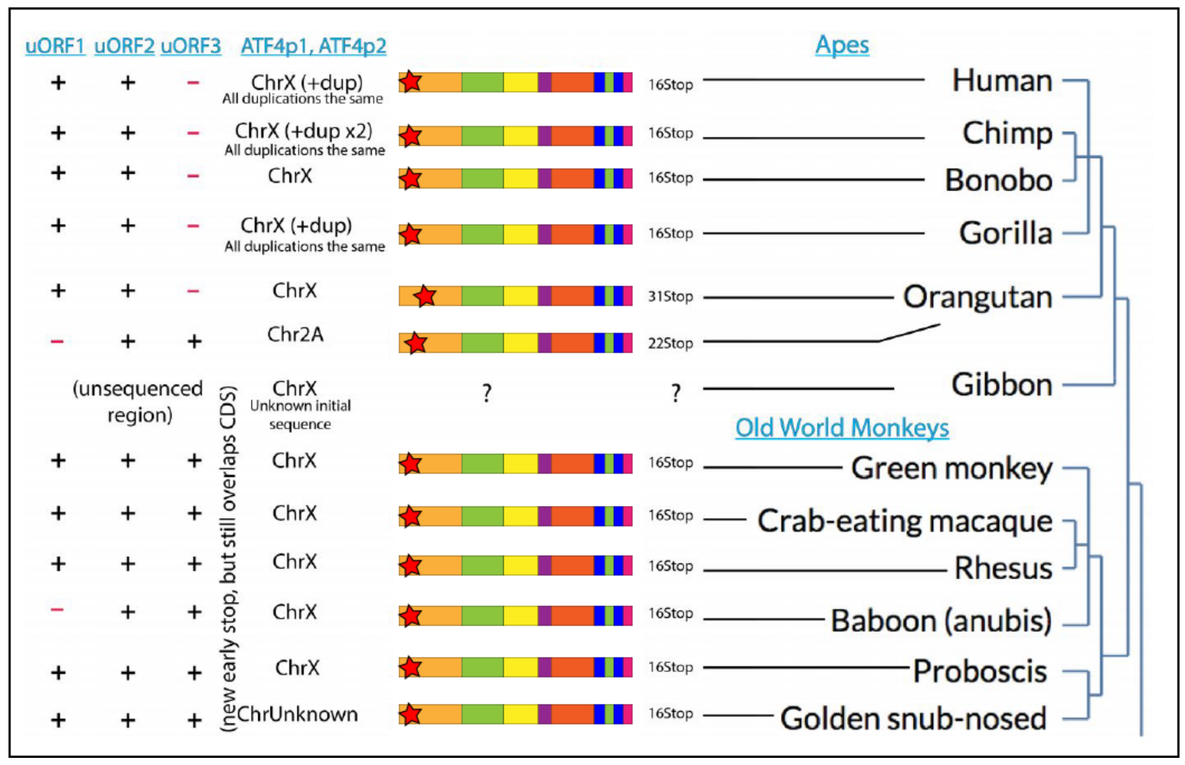

Figure 5 - Very truncated copy. *hATF4P1* and *P2* are short, truncated retrocopies on Chr. X that date to the common ancestor between apes and old world monkeys. Given the early stop, *P1* and *P2* are likely nonfunctional as proteins. Given the strong conservation, it suggests a possible importance of its well-conserved uORF region.
Key: **+** = Maintained, **-** = Disrupted, ? = Unknown

Future directions -

Strong conservation and independent events in different lineages suggests that each of these retrocopies has the potential to be functional. For example, all truncated *hATF4P4*-like retrocopies maintain the transcription-enhancing p300 protein binding domain. We hypothesize that *hATF4P4*-like retrocopies may regulate the parent *ATF4* gene by sequestering proteins like p300 away from ATF4. In addition, viruses, like HIV, can hijack important machinery during the ISR, suggesting an alternative hypothesis where hATF4P4 may buffer ATF4 to inhibit viral infection while not affecting *ATF4* transcription.

Ongoing experiments include:
- Cell culture work to determine functional consequences of translated *ATF4* retrocopies
- Substitution of retrocopy uORF regions with parental copies to determine effect on translation.
- Further elucidation of Gorilla *ATF4* retrocopy duplications and its potential meaning in *ATF4* evolution

Funding:

Dalton - NIH NIGMS 1 F32 GM136057

Chow - Mario Capecchi Endowed Chair in Genetics; Glenn Foundation for Medical Research; NIH NIGMS 1 R35 GM124780

Chow Lab
cchow@genetics.utah.edu
hdalton@genetics.utah.edu
www.ChowLab.org
twitter: @ChowLab