

Rapid and recurrent evolution of MUC 7 repeats within and across primate species

Emilie Redwood, Duo Xu, Stefan Ruhl, Omer Gokcumen



Abstract

MUC7 codes for one of the few intrinsic salivary proteins and the genetic variation affecting this gene has been associated with asthma. The gene is also unusual as it contains exonic repeats, which harbor functionally important O-glycosylation sites. Here, I conducted bioinformatic analysis of human and nonhuman primate genomes, revealing unusually rapid and recurrent nucleotide and copy number evolution of these repeats throughout primate species. In addition, I conducted a phylogenetic analysis of this variation to predict the evolutionary chronology of individual repeats. Collectively, my results are concordant with diversifying selection acting on MUC7, likely as a response to diverse pathogenic pressures.

Introduction

- Mucins are glycoproteins which are the major components of the mucus layer covering epithelial surfaces.
- It's functions include maintaining oral health by promoting the clearance of bacteria.
- MUC7 has been linked to potential protection against atopic asthma in humans.

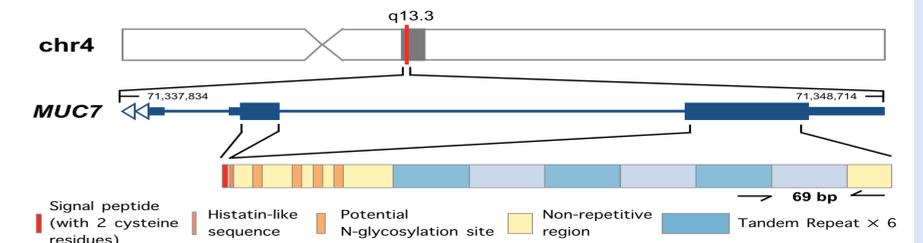


Figure 1. Genomic location, organization and exonic copy number variation of MUC7 gene in human reference genome (GRCh37/hg19).

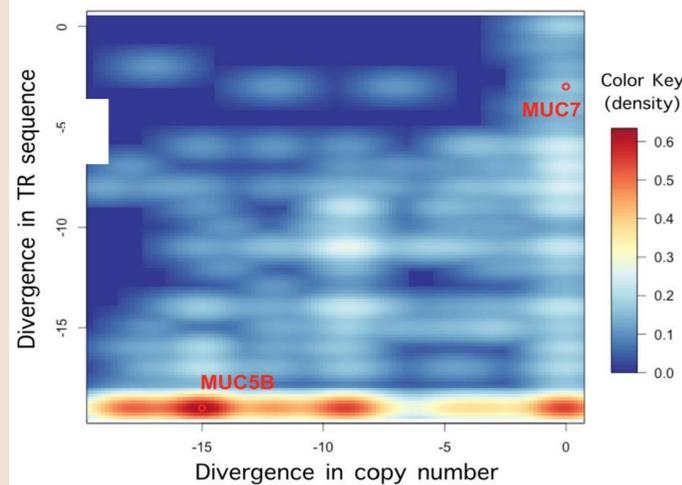


Figure 2. Divergence for both copy number variation and repeat sequence among all proteins with exonic tandem repeats across mammals.

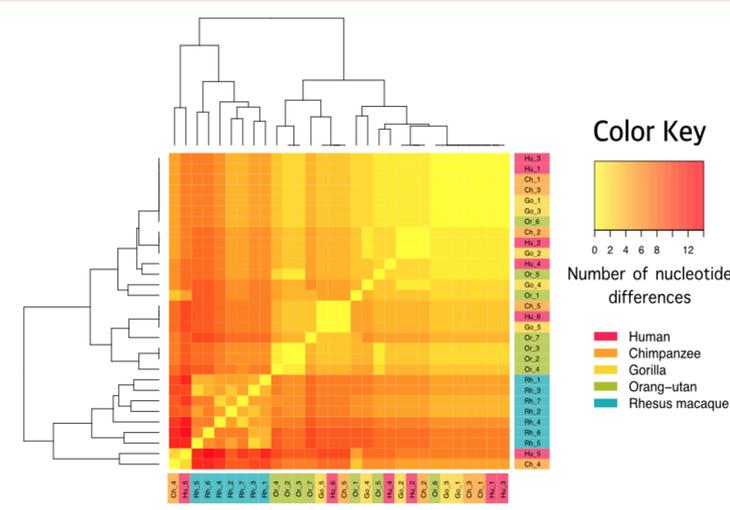


Figure 3. Pairwise nucleotide differences between each repeat sequence among primates.

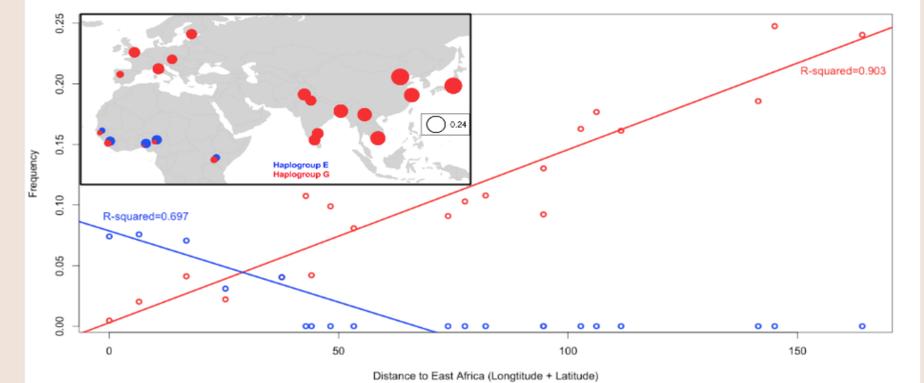


Figure 6. Worldwide distribution of both human 5 copy haplogroups and the correlation with their geographical distance to East Africa.

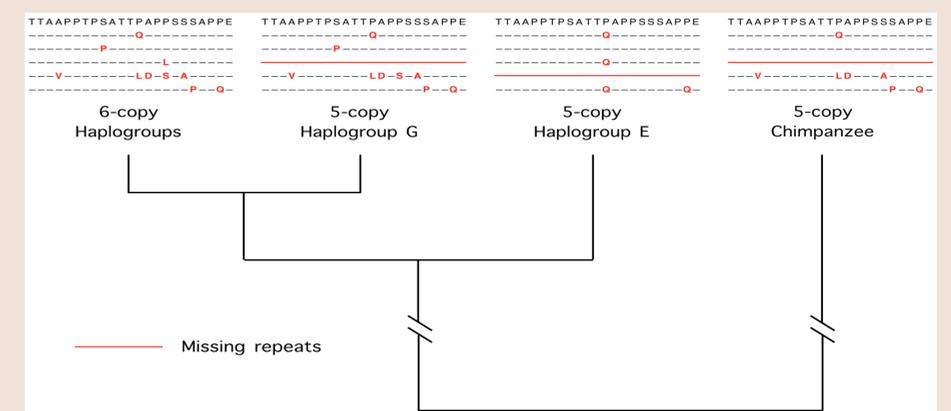


Figure 7. Repeat patterns in different haplogroups.

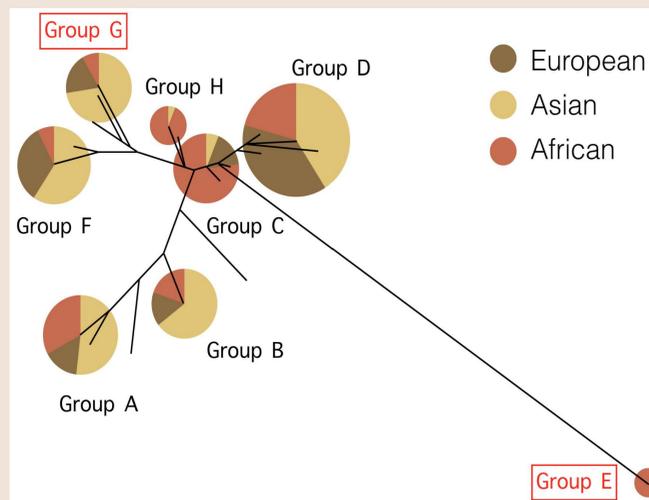


Figure 4. Distribution of human populations in each haplogroup. Groups G and E are 5 copy haplotypes.

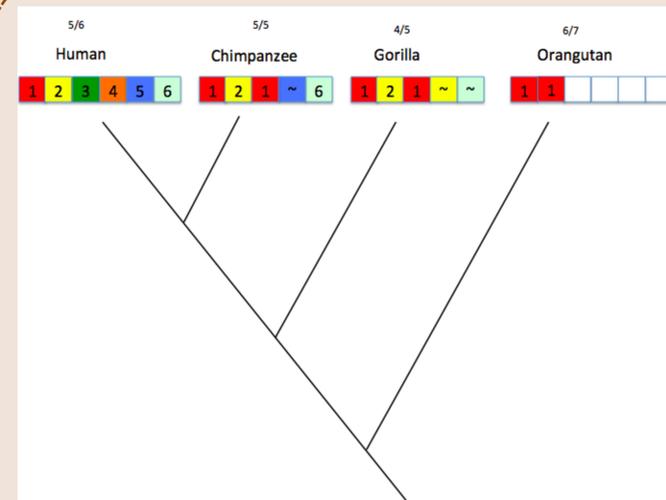


Figure 5. Human centric analysis using reference sequence data to show map of potential repeat order, ancestral copies and evolution.

Conclusions and Future Directions

- The data show recurrent and rapidly evolving repeats.
- Recurrent variants shape the global distribution of variation in humans.
- The variation in copy number and haplotype may have a distinct functional impact. For example, they were recently found to affect the type and profile of oral microbes present.

Acknowledgements: The Gokcumen lab including Colin Flanagan, Sara Fox, Onta Lin, Easwarkhanth Muthukrishnan and Omer Gokcumen, also Clara Redwood.