**Investigating Faster-Z in *Gallus gallus* and Taeniopygia guttata**

**Abstract**

Evolution predicts that Z chromosome genes evolve faster than autosomal ones, as natural selection and genetic drift are stronger in Z-linked genes. Here we investigated the faster-Z hypothesis (the faster evolution of genes in Z-chromosome) in chicken, Gallus gallus domesticus (GGA) ,and zebrafinch, Taeniopygia guttata (TGU), alongside autosomal control genes. Additionally, we used the temperature-dependent sex determination of two crocodilian species, Chinese alligator (Alligator sinensis) and gharial (Gavialis gangeticus) to work as an additional control, as they lack sex chromosomes altogether. We used codeml package from PAML to estimate the ratio of nonsynonymous to synonymous substitutions (dN/dS) for sex-linked genes alongside autosomal genes. We used 509 Z-linked genes in zebrafinch and chicken and 6907 autosomal genes. The dN/dS ratio was significantly higher for Z-linked (0.1886) than for genes linked to similarly sized autosomes 1-10 (0.1599; P=0.0045). Using Maximum likelihood, we found no faster-Z in chicken, alligator and gharial. These results suggest that evolution proceeds more quickly on some but not all Z chromosomes.

Sex chromosomes are significantly different from autosomes and these differences likely influence the rate of divergence of sex-linked genes. Sex chromosomes have reduced hemizygosity and recombination; but more importantly, they have a unique form of inheritance different from autosome, making them a good option to study different evolutionary phenomenon. The sex chromosomes usually have a differing effective population size than autosomes, since in a species with a 1:1 sex ratio, the population of X and Z chromosome is ¾ the size of autosomes and the population of W and Y is ¼ the size of autosome population. Reduced recombination in X and Z can contribute to the chromosomal degeneration, since without recombination, deleterious mutations are more likely to become fixed; whereas, in W and Y, the reduced population makes the effect of genetic drift stronger, allowing mildly deleterious to become fixed there as well. Furthermore, the reduced hemizygosity in heterogametic sex opens recessive mutation to the action of selection. Faster rates of divergence of coding sequences have been seen in X and Z chromosomes in many species including chicken, zebrafinch, *D. melanogaster*, *Mus castaneus* and *Homo sapiens*.

A 2007 paper titled “Fast-X on the Z: rapid evolution of sex-linked genes in birds.” by Mank et. Al discovered faster-Z in *Gallus gallus* (GGA) and Taeniopygia guttata (TGU). The paper evaluated the divergence of coding sequences in the Z chromosome and compared it to coding sequeces in autosomes. They compared it to CDSs in chromosome 1-10 as they have comparable sizes. The calculated the rate of synonymous and non-synonymous substitution per substitution site. Synonymous changes are the nucleotide substitutions that do not change the amino acid coded by the codon while non-synonymous substitutions change the amino acid. The ratio of the non-synonymous substitution rate per substitution site (dN) and the synonymous substitution rate per substitution site (dS) is referred to as dN/dS or omega (ω). dN/dS can be used as a metric of faster-Z evolution when averaged over a large genomic distance.

The authors used PAML (Phylogenetic Analysis by Maximum Likelihood) to calculate pairwise dN/dS between same gene CDSs in GGA and TGU where they found CDSs in the Z chromosome to have a significantly higher dN/dS compared to the autosomes. This paper established faster-Z in birds and has been the basis for many different publications regarding evolution of sex chromosome.

The pairwise dN/dS calculation is a rudimentary option in software that offers many advanced and complex analysis options. For my project I will be investigating this same phenomenon in PAML employing the more advanced maximum likelihood method of calculation of omega. This advancement in newer versions of PAML allows us to use different branch models according to different hypotheses. The TGA and GGA genomes from NCBI are improved compared to what they were in 2007. Additionally, I will include a temperature dependent sex determination (TSD) control for the analysis. TSD species are organisms whose sex determination does not depend on sex chromosome. The sex of the offspring is determined by the temperature of the nest in species like some turtles, alligator, crocodiles, lizards etc. These species lack sex chromosomes and thus all the genes present in their genome are in autosomal condition. This will offer a good contrast. I will be using two crocodilian TSD species with similar divergence time as TGU and GGA, *Alligator senensis* and *Gavialis gangeticus*. They lack sex chromosome and could be valuable as controls in the analyses. All of these species have genomes available in NCBI.



**Method**

We collected CDSs from the genome of zebrafinch, chicken, Chinese alligator and gharial from NCBI. ~~We used CDS of chicken Z chromosome and chromosome 1-10 against the other species and used those to prepare Alignments in PRANK~~. We aligned the CDS regions of chicken chromosomes Z and 1-10 against the CDS regions of the other species’ chromsomes, in PRANK. Each gene alignment was run through PAML individually with the free model to estimate the dN, dS and the dN/dS. The average dN/dS was calculated by dividing the sum of the number of substitutions over genes by the sum of the number of sites over genes. This circumvented the problem of infinite dN/dS when dS tends to zero. The data is also weighted for the length of the alignment of each gene. We then separately ran a pairwise model for avian-avian and crocodilian-crocodilian pairs of genes to see how they diverged from each other. We disregarded any dataset where dS was greater than 2. A permutation test using 10000 repetitions was used to assist the significance of the data along with bootstrap with 10000 repetitions to establish 95% confidence intervals. Most of this was done using R and PAML on the HPC clusters at Iowa State University.

**Results**

The result from the pairwise estimation of dN/dS between avian-only and crocodilian-only pairs of genes showed faster divergence in the avian pair than in the crocodilian pair. The dN/dS in avian pair was significant as shown in table 1.

|  |  |  |  |
| --- | --- | --- | --- |
| **Avian pair** | **Z-linked [95% CI]** | **Auto-linked [95% CI]** | **P-value** |
| dN | 0.0704 [0.0625-0.0781] | 0.0636 [0.0608- 0.0652] | 0.0962 |
| dS | 0.3857 [0.3524-0.3912] | 0.3857 [0.3759-0.3901] | 0.1523 |
| dN/dS | 0.1826 [0.1700-0.2083] | 0.1650 [0.1598-0.1692] | 0.0076\* |
| **Crocodilian pair** | **Z-linked [95% CI]** | **Auto-linked [95% CI]** | **P-value** |
| dN | 0.0287 [0.0238-0.0341] | 0.0308 [0.0291-0.0324] | 0.4941 |
| dS | 0.1340 [0.1122- 0.1577] | 0.1523 [0.1452-0.1597] | 1.0 |
| dN/dS | 0.2145 [0.1859-0.2447] | 0.2145 [0.1953-0.2091] | 0.3697 |

Table 1.0 . Pairwise mean dN, dS, and dN/dS for Z-linked and autosomal chicken–zebra finch orthologs and crocodilian orthologs.

The result from the maximum likelihood estimation of dN/dS for each species using free model in CODEML from PAML showed faster divergence in only zebrafinch and not chicken. The crocodilian species did not show faster evolution in the genes z-linked in chicken. The maximum likelihood estimations are show in table 2.

|  |  |  |  |
| --- | --- | --- | --- |
| **Chicken** | **Z-linked [95% CI]** | **Auto-linked [95% CI]** | **P-value** |
| dN | 0.0311 [0.0277-0.0346] | 0.0317 [0.0302-0.0332] | 0.8211 |
| dS | 0.1912 [0.1703-0.2237] | 0.1902 [0.1855-0.1950] | 0.9093 |
| dN/dS | 0.1625 [0.1418-0.1845] | 0.1666 [0.1611-0.1721] | 0.7026 |
| **Zebrafinch** | **Z-linked [95% CI]** | **Auto-linked [95% CI]** | **P-value** |
| dN | 0.0324 [0.0275-0.0316] | 0.0284 [0.1549-0.1649] | 0.04\* |
| dS | 0.1717 [0.1641-0.1756] | 0.1776 [0.1741-0.1810] | 0.4941 |
| dN/dS | 0.1887 [0.1695-0.2080] | 0.1599 [0.1549-0.1649] | 0.0045\* |
| **Alligator sinenesis** | **Z-linked [95% CI]** | **Auto-linked [95% CI]** | **P-value** |
| dN | 0.0125 [0.0102-0.0152] | 0.0104 [0.0129-0.0145] | 0.4671 |
| dS | 0.0562 [0.0553-0.0591] | 0.0639[0.0639-0.0709] | 0.0831 |
| dN/dS | 0.2230 [0.1918-0.2588] | 0.2019 [0.1952-0.2088] | 0.2088 |
| **Gavialis gangeticus** | **Z-linked [95% CI]** | **Auto-linked [95% CI]** | **P-value** |
| dN | 0.0090 [0.0081-0.0011] | 0.0104 [0.0098-0.0110] | 0.4818 |
| dS | 0.0449 [0.0388-0.0523] | 0.0515 [0.0494-0.0539] | 0.1089 |
| dN/dS | 0.2153 [0.1820-0.2505] | 0.2019 [0.1945-0.2115] | 0.3705 |

Table 2 . Maximum likelihood mean dN, dS, and dN/dS for Z-linked and autosomal for each species

**Discussions**

Here we see from the pairwise analysis that the genes in avian Z chromosome pairs have significantly higher dN/dS than austosomal genes on the similarly sized chromosomes 1-10. The same genes did not show significantly higher dN/dS in TSD species (crocodilians). When we apply maximum likelihood to the analysis we see that the faster-Z in the avian pairs relative to the crocodilian pairs was due to solely or primarily to individual differences between zebrafish and chicken; the dN/dS of Z-linked genes in chicken was not significantly greater than that of autosomal genes, but it was significant in zebrafinch. The crocodilian species did not show greater dN/dS in either of the analysis which supports the hypothesis that genes in the Z chromosome could evolve faster, both compared to autosomal genes within its species and also compared to equivalent sex-determining genes in the autosomal context of crocodilian species.

**References**

Mank, J. E., Axelsson, E., & Ellegren, H. (2007). Fast-X on the Z: rapid evolution of sex-linked genes in birds. *Genome Research*, *17*(5), 618-624.

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution*, *24*(8), 1586-1591.

Meisel, R. P., & Connallon, T. (2013). The faster-X effect: integrating theory and data. *Trends in Genetics*, *29*(9), 537-544.

Mank, J. E., Vicoso, B., Berlin, S., & Charlesworth, B. (2010). Effective population size and the Faster‐X Effect: empirical results and their interpretation. *Evolution*, *64*(3), 663-674.

Charlesworth, B., Coyne, J. A., & Barton, N. H. (1987). The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, *130*(1), 113-146.

Bachtrog, D., Kirkpatrick, M., Mank, J. E., McDaniel, S. F., Pires, J. C., Rice, W., & Valenzuela, N. (2011). Are all sex chromosomes created equal?. *TRENDS in Genetics*, *27*(9), 350-357.

Mank, J. E., Vicoso, B., Berlin, S., & Charlesworth, B. (2010). Effective population size and the Faster‐X Effect: empirical results and their interpretation. *Evolution*, *64*(3), 663-674.

Mank, J. E. (2012). Small but mighty: the evolutionary dynamics of W and Y sex chromosomes. *Chromosome Research*, *20*(1), 21-33.

Literman, R., Burrett, A., Bista, B., & Valenzuela, N. (2018). Putative Independent Evolutionary Reversals from Genotypic to Temperature-Dependent Sex Determination are Associated with Accelerated Evolution of Sex-Determining Genes in Turtles. *Journal of molecular evolution*, *86*(1), 11-26.