**Introduction**

Thermosensitivity is important for species exhibiting temperature dependent sex determination (TSD), and understanding its role in development is an important, unanswered question. RNA-seq experiments in *Chrysemys picta* (a TSD species) and *Apalone spinifera* (a genotypic SD species) have revealed interesting thermosensitivity patterns in some genes known to be related to urogenital networks in mammals (Radhakrishnan *et al*, 2017). TSD is believed to be the ancestral state in turtles, with GSD arising a number of times (and possibly reverting back to TSD) (Valenzuela and Adams, 2011).

Wt1 is a transcription factor involved in urogenital development (GCID: GC11M032365) and exhibits thermosensitivity in both *C. picta* and *A. spinifera* (Radhakrishnan *et al*, 2017), two distantly related *Cryptodiran* turtle species. In *C. picta* it has shown male-biased differential expression during the thermosensitive period (TSP) (developmental stages 19 and 22). The pattern is mixed in *A. spinifera* with male biased expression prior to and late in the TSP (developmental stages 15 and 22) but female biased expression earlier in the TSP (developmental stage 19) (Radhakrishnan *et al*, 2017). This raises the question of why a GSD species from an ancient GSD lineage would still exhibit thermosensitivity in genes proposed to be involved in sex determination? Is it simply residual? Does it serve some other function?

Sf1 is a transcription factor involved in sex determination (GCID: GC09M124481). Its pattern of thermosensitivity is more as one might expect between TSD and GSD turtles with thermosensitivity present in *C. picta* and absent in *A. spinifera*. It shows male-biased differential expression during the TSP (developmental stages 19 and 22) (Radhakrishnan *et al*, 2017). How will its evolutionary history compare to that of Wt1? Both are important for sex determination but one has retained thermosensitivity and another does not. Questions regarding the thermosensitivity patterns of these genes render it worth examining their evolutionary history using phylogenetic approaches.

Table 1. Species used in this study, abbreviations associated with their analysis (6-letters are used in analysis files only), and corresponding sex determination mechanisms. Outgroups are in bold.

|  |  |  |
| --- | --- | --- |
| Species Name | Abbreviations | Sex Determining Mechanism |
| *Chrysemys picta* | ChrPic, CPI | TSD |
| *Trachemys scripta* | TraScr, TSC | TSD |
| *Glyptemys insculpta* | GlyIns, GIN | GSD |
| *Staurotypus triporcatus* | StaTri, STR | GSD |
| *Chelonia mydas* | CheMyd, CMY | TSD |
| *Carettochelys insculpta* | CarIns, CIN | TSD |
| *Pelodiscus sinensis* | PelSin, PSI | GSD |
| *Apalone spinifera* | ApaSpi, ASP | GSD |
| *Podocnemis expansa* | PodExp, PEX | TSD |
| *Emydura macquarii* | EmyMac, EMA | GSD |
| ***Alligator mississippiensis*** | **AllMis, AMI** | **TSD** |
| ***Gallus gallus*** | **GalGal, GGA** | **GSD** |

**Hypotheses**

Given the conserved thermosensitivity patterns observed in Wt1 across *C. picta* and *A. spinifera*, it will exhibit similar evolutionary rates across GSD and TSD species.

Considering the loss of thermosensitivity observed in Sf1 in *A. spinifera* relative to *C. picta*, it is predicted it will exhibit faster evolution in GSD species relative to TSD species.

**Methods**

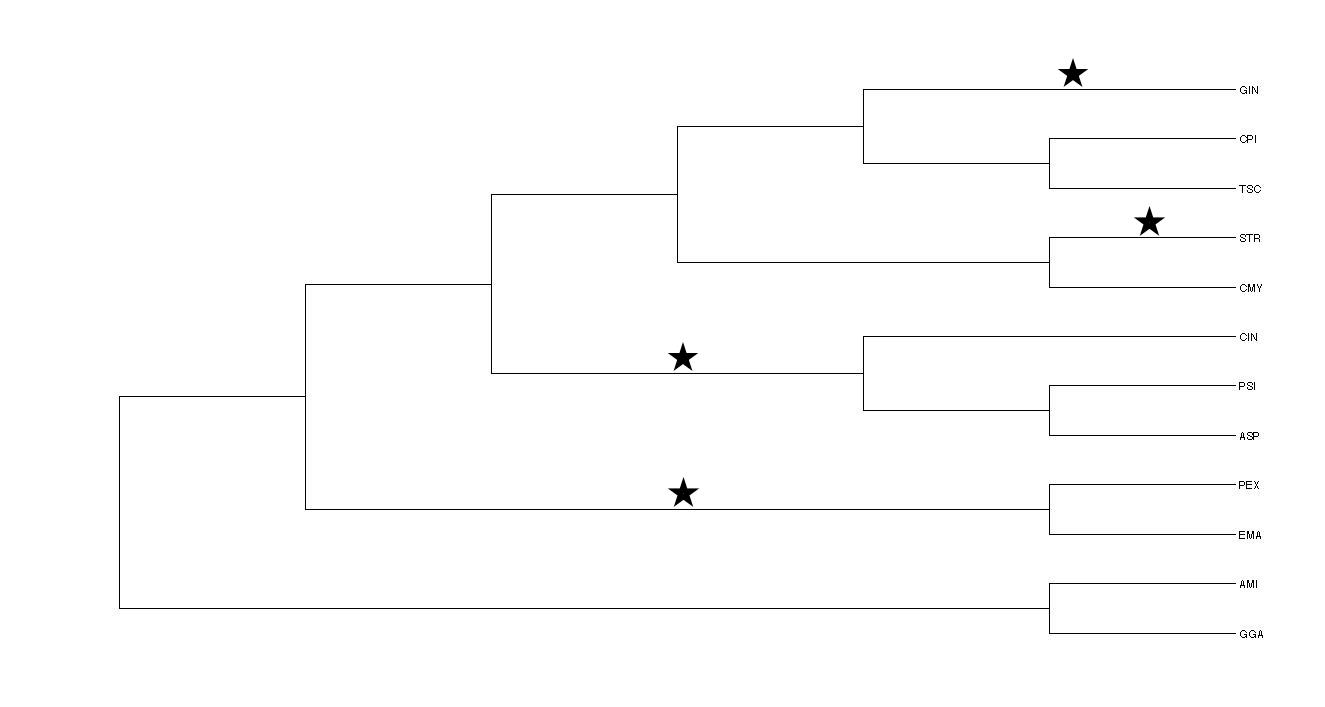


Figure 1. Species tree used in analysis of gene evolution rates (based on Literman et al, 2017). AMI and GGA represent the outgroup. Stars indicate branches where TSD 🡪 GSD transitions have been hypothesized to occur (Valenzuela and Adams 2011). See Table # for name abbreviations. Test groups were as follows: 1-GIN, CPI, TSC; 2-STR, CMY; 3-CIN, PSI, ASP; 4-PEX, EMA. Tree produced in Dendroscope (v3.5.9).

Coding sequences were obtained from a previous study (Literman *et al*, 2017) for ten turtle species and an alligator species for the genes SF1 (also known as NR5A1) and WT1 . Chicken coding sequences for these genes were obtained from the latest NCBI release (103). Separately, sequences were concatenated into a single file and aligned with MAFFT (v7.245) using the --auto parameter for the alignment and --phylipout parameter to produce phylip alignment files. Alignments were tested for differences in rates of evolution between TSD and GSD species for each gene, respectively. Using CODEML (PAML v4.9), log likelihoods were calculated for the hypothesis that GSD and TSD species would exhibit different rates of evolution for these genes (parameters: runmode=0, CodonFreq=2, model=2, kappa and omega were estimated).

Due to the broad distribution of the two sex determining mechanisms across the species tree for turtles (Fig. 1) (Valenzuela and Adams, 2011), species were broken up into four test groups to look for trends in evolutionary rate differences of each gene. This enabled more straightforward labeling of hypothesis trees in such a way that controlled for divergence times of the group. Test groups were taxonomically defined (Fig. 1 and Table 2).

In addition to the above analyses, the same kind of test was run comparing rates on the branches hypothesized to contain TSD 🡪 GSD transitions (Fig. 1) (Valenzuela and Adams 2011) to the null that all branches experienced the same rate. Finally, CPI and ASP branches were compared for differences in evolutionary rates due to their being species of interest. CPI and ASP represent two greatly diverged lineages of *Cryptodiran* turtles exhibiting different sex determination mechanisms (TSD and GSD, respectively).

Above analyses were evaluated for significance using likelihood ratio tests with one degree of freedom.

Table 2. Hypotheses tested for SF1 and WT1. Hash marks indicate how branches were labeled for hypotheses of rate variation. Branches were labeled back to the point of divergence for each group or comparison. Trees were rooted with A. mississippiensis and G. gallus.

|  |  |  |
| --- | --- | --- |
| Test | Alternative Hypothesis | Null Hypothesis |
| Group 1 | #1:CPI, TSC #2:GIN | #1:CPI, TSC, GIN |
| Group 2 | #1:CMY  #2:STR | #1:CMY, STR |
| Group 3 | #1:CIN  #2:PSI, ASP | #1:CIN, PSI, ASP |
| Group 4 | #1:PEX  #2:EMA | #1:PEX, EMA |
| TSD 🡪 GSD | Marked by stars in Fig. 1 | All branches the same |
| CPI vs. ASP | #1: CPI  #2: ASP | #1:CPI, ASP |

**Results**

Likelihood ratio tests (LRT) indicated that among the four subgroups tested, group 4 (PEX vs EMA) had greater maximum likelihood support (p<0.001)for the tree specifying different rates between GSD and TSD species for WT1 (Table 3). The omega value for the GSD branch (EMA) was 0.03633, for the TSD branch (PEX) was 0.32891, and for the remainder of the branches was 0.06428. Additionally the LRT indicated that group 2 (STR vs CMY) had greater maximum likelihood support (p<0.1) for the tree specifying different rates between GSD and TSD species for WT1 (Table 3). The omega value for the GSD branch (STR) was 0.06193, for the TSD branch (CMY) was 0.36941, and for the remainder of the branches was 0.08930. All other results were non-significant for both genes (Tables 3 and 4).

Table 3. Likelihood ratio tests (LRT) for test case hypotheses for the gene WT1. Italics indicates significance of LTR at the p<0.1 level. Bold indicates significance at the p<0.001 level. LnL are log likelihood values produced from CODEML as a part of the PAML program.

|  |  |  |  |
| --- | --- | --- | --- |
| WT1 | Alternate(LnL) | Null(LnL) | LRT |
| Group 1 | -3703.27162 | -3703.334128 | 0.125016 |
| Group 2 | -3702.398412 | -3704.295335 | *3.793846* |
| Group 3 | -3703.338517 | -3704.32346 | 1.969886 |
| Group 4 | -3685.700631 | -3692.447046 | **13.49283** |

Table 4. Likelihood ratio tests (LTR) for test case hypotheses for the gene SF1. LnL are log likelihood values produced from CODEML as a part of the PAML program.

|  |  |  |  |
| --- | --- | --- | --- |
| SF1 | Alternate(LnL) | Null(LnL) | LRT |
| Group 1 | -4870.312772 | -4870.322533 | 0.019522 |
| Group 2 | -4869.72755 | -4870.365022 | 1.274944 |
| Group 3 | -4870.734803 | -4871.920207 | 2.370808 |
| Group 4 | -4871.812956 | -4871.968784 | 0.311656 |

LRT did not show significant support for the hypotheses proposing greater rates of evolution in TSD🡪GSD transition branches relative to other branches (Table 5).

Table 5. Likelihood ratio tests (LTR) for hypothesis that TSD 🡪 GSD transition branches evolve at a faster rate than non-transition branches for the genes SF1 and WT1. LnL are log likelihood values produced from CODEML as a part of the PAML program.

|  |  |  |  |
| --- | --- | --- | --- |
| TSD🡪GSD | Alternate(LnL) | Null(LnL) | LRT |
| SF1 | -4871.655479 | -4872.032349 | 0.75374 |
| WT1 | -3703.582263 | -3704.323549 | 1.482572 |

LRT did not show significant support for greater evolution rates in ASP relative to CPI for either SF1 or WT1 (Table 6).

Table 6. Likelihood ratio tests (LTR) for hypothesis that CPI and ASP branches would exhibit different rates of evolution. LnL are log likelihood values produced from CODEML as a part of the PAML program.

|  |  |  |  |
| --- | --- | --- | --- |
| CPI v. ASP | Alternate(LnL) | Null(LnL) | LRT |
| SF1 | -4871.675995 | -4871.685883 | 0.019776 |
| WT1 | -3702.576398 | -3702.577609 | 0.002422 |

**Discussion**

Results indicate that the rate of evolution of the PEX branch was faster than the EMY branch for WT1. Additionally, there was a trend for the CMY branch rate to be faster than the rate for the STR branch for WT1. Both test cases refutes the hypothesis that GSD species would have faster evolutionary rates than TSD species given that TSD is ancestral. There was no clear trend in any other groups. Additionally, these differences were observed in the gene WT1 which is an example of a gene that has exhibited conserved thermosensitivity across two greatly diverged turtle lineages, counter to what was predicted. Rather, SF1 (which has lost thermosensitivity) showed no differences in evolutionary rates between any of the lineages tested.

The differences present in this group could be due to greater divergence present between these two species relative to other groupings as test groups are not comparable because they are not controlled for divergence times. Using test groups (as was done here) is perhaps not the most desirable way to test for differences based on sex determination mechanisms as it divides up the data, limited the power of the analysis.

Indeed, when the CPI vs ASP was tested (the pairing that led to the proposed hypotheses) the LRT did not show support for faster rates in ASP relative to CPI, even though SF1 has exhibited lost thermosensitivity.

Results for WT1 were inconsistent with the hypothesis that TSD and GSD species would exhibit similar rates of evolution since it represents a gene with conserved thermosensitivity. Likewise results for SF1 were inconsistent with the hypothesis that TSD and GSD species would exhibit different rates of evolution as a result of its lost thermosensitivity. It is possible there is a large constraint on the evolution of these genes as they are important for sex determination (ref). It is also possible that coding sequences are not the best region in which to test this question. Coding sequences are more likely to be under strong negative selection because of the potential importance of a gene product’s function (ref?). Perhaps promoter sequence would better explain the patterns of thermosensitivity observed in these turtles. Due to turtles being non-model organisms, obtaining promoter sequences is much more difficult than coding sequences.

Testing for differences in rates of evolution on branches believed to show TSD🡪GSD transitions did not show any difference between those branches and the rest of the tree for either gene suggesting that transitions in sex determination mechanisms may not result in increased evolutionary results for these two genes. These results also refute the hypotheses as they show evidence against transitions of GSD being associated with faster evolutionary rates. Another similar test that could be done, would to be test the rates on branches that exhibited reversions from GSD back to TSD. PEX is hypothesized to have exhibited a second transition (Valenzuela and Adams 2011), perhaps owing to its greater evolution rate. (I should test second transition branches as well as well as all transition branches)

{Bob’s findings on the species level when using more genes}

References

Literman et al 2017

Radhakrishnan et al 201#

Valenzuela and Adams 2011