

Amm_transcriptome_results_update5

After Skype meeting, I re-calculated tau based on a new formula: which separate sex in all 5 developmental larva stages and three adult tissues, so in total 16 tissues (instead of 11 in previous analysis).

A. Tissue specificity index tau

Approach:

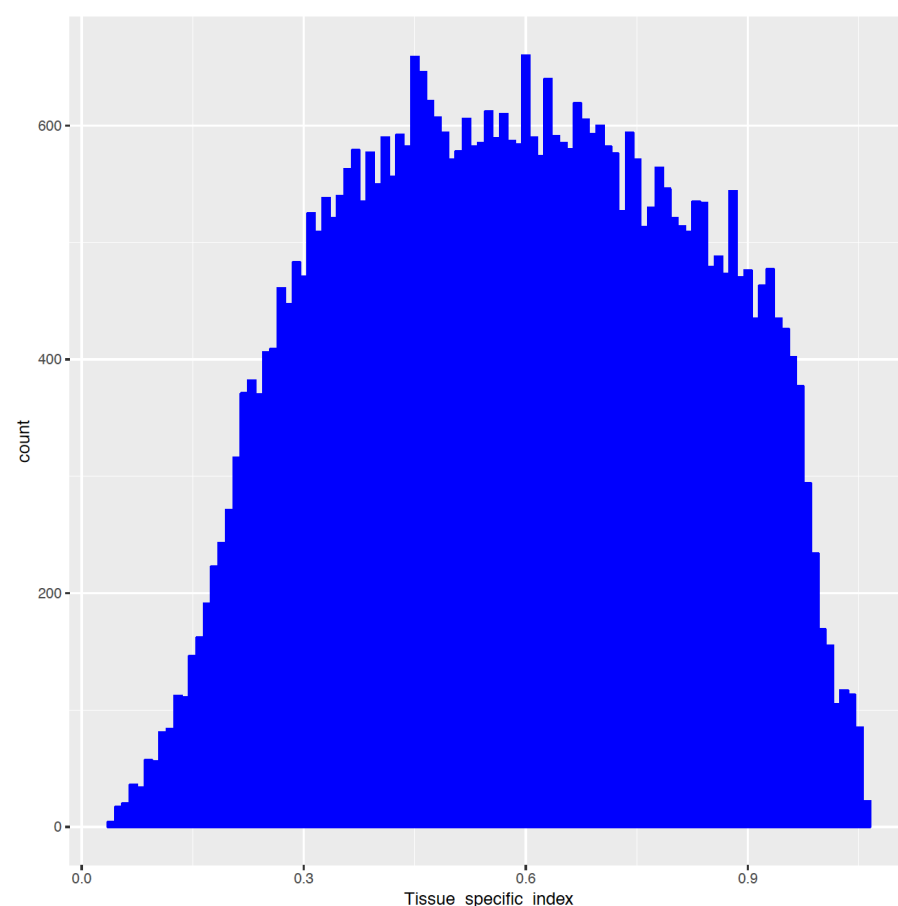
1. Define in total 16 tissues, see above explanation.
2. Use Kallisto to quantify transcript expression, and then calculate Tau from the generated output of TPM matrix.
3. Use tissue specificity index Tau formula below (Mank et al. 2008; Brown & Bachtrog 2014):

$$\tau = \frac{\sum_{i=1}^N 1 - \frac{\log E_i}{\log E_{max}}}{N - 1},$$

i is one tissue, Ei is the expression (non-normalized TPM value) of certain transcript.

4. Keep transcripts which are expressed in at least one of the 16 tissues with TPM >=1.

Figure 1. Histogram distribution of Tau for all transcripts in Ammarnas transcriptome.



For G43 stage
#####

Figure 2.1 Boxplot of sex-biased genes and the tissue specificity index Tau at stage G43.

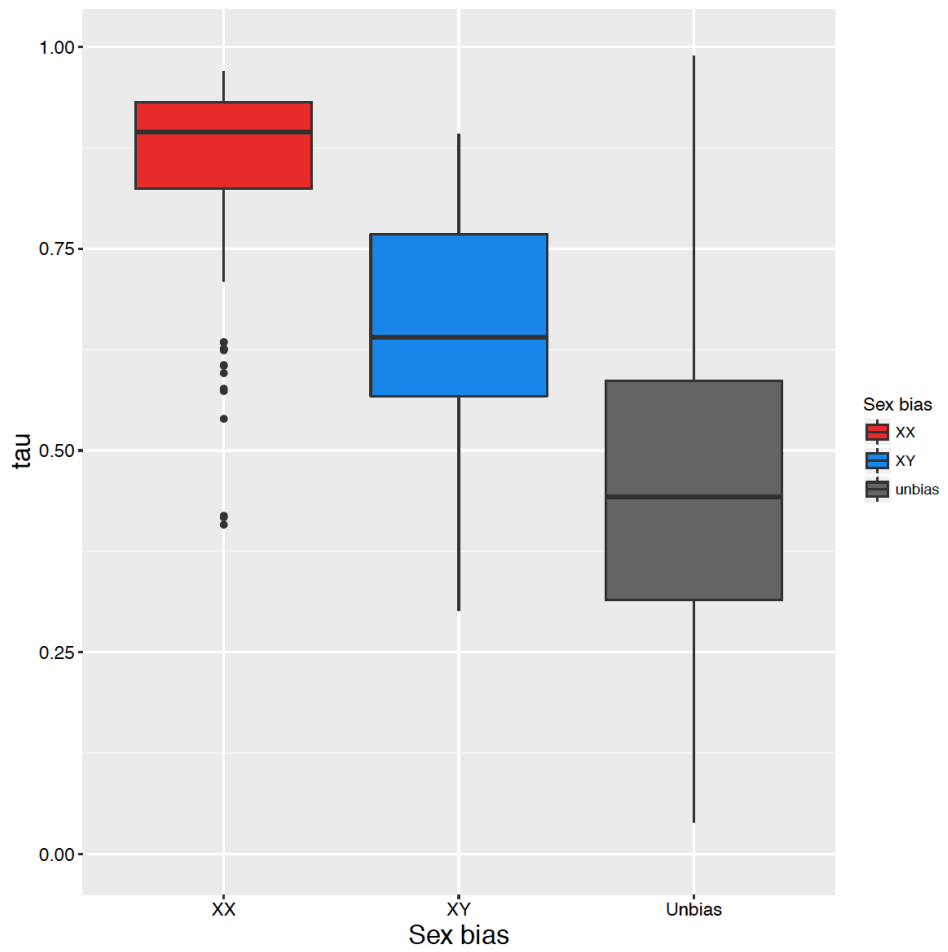


Figure 2.2 Boxplot of dN/dS among female-, male-biased and unbiased genes at G43 stage.

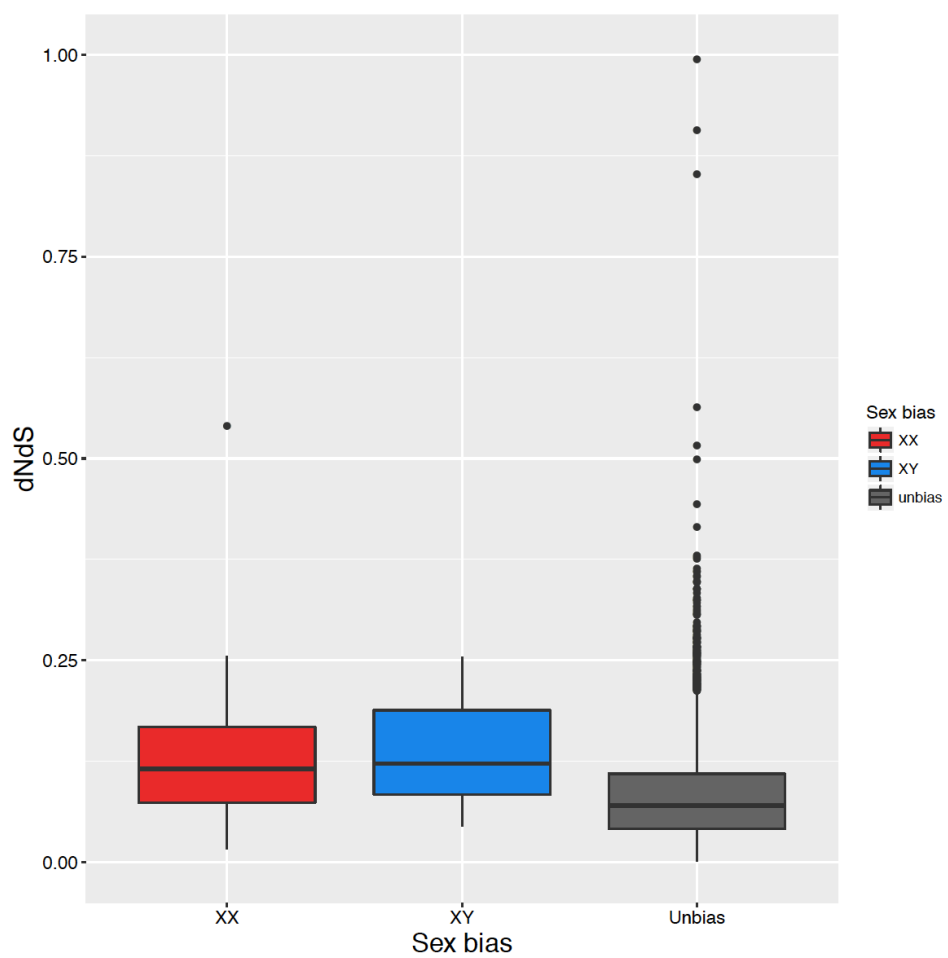
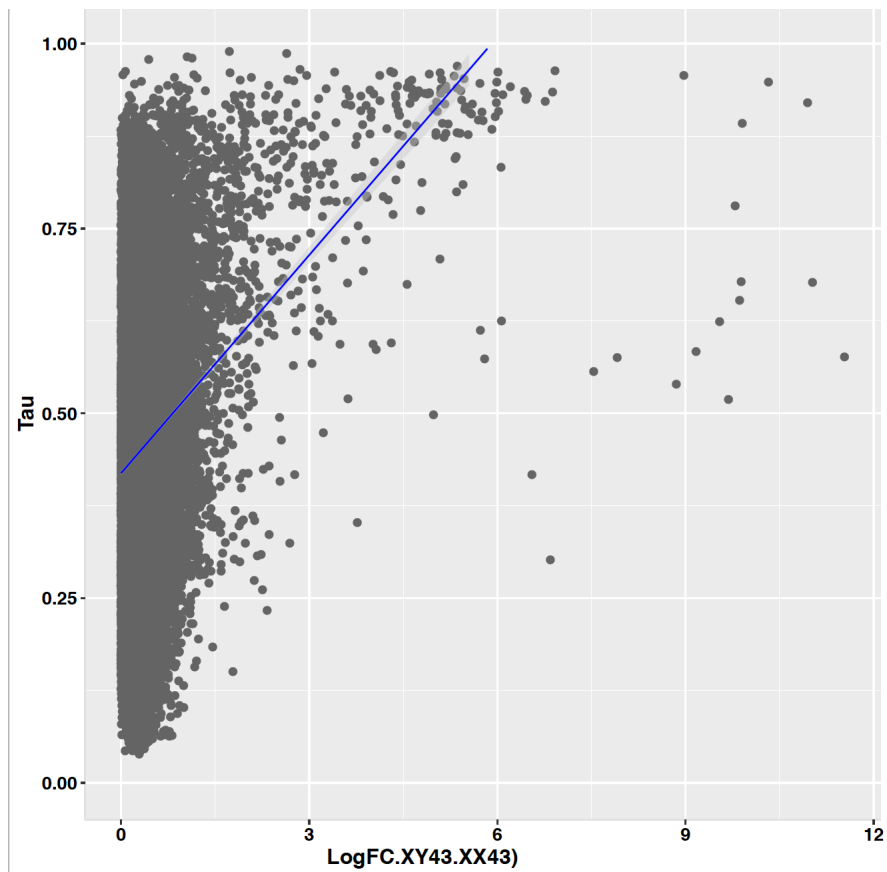


Figure 3. Correlation between absolute values of gene expression ratio $\text{Log}_2(\text{XY43}/\text{XX43})$ (include sex-biased and unbiased genes) and tissue specificity index tau.



##investigate whether dN/dS can be better explained by tissue specificity or sex biased gene expression.

```
y <- lm(sqrt(tau)~sqrt(abs(logFC.XY43.XX43))*bias, data=g43_tau)
anova(y)
```

```
#####
Response: sqrt(tau)
Df Sum Sq Mean Sq F value Pr(>F)
sqrt(abs(logFC.XY43.XX43))      1 35.32 35.321 1924.0158 < 2.2e-16 ***
bias                            2  0.85  0.424  23.0809 9.663e-11 ***
sqrt(abs(logFC.XY43.XX43)):bias  2  0.31  0.154   8.3743 0.0002314 ***
Residuals                    25568 469.38  0.018
```

```
#####
For G46 stage
#####
```

Figure 4.1. Boxplot of sex-biased genes and the tissue specificity index Tau at stage G46 (gene expression from sex-reversed XX male is not included here).

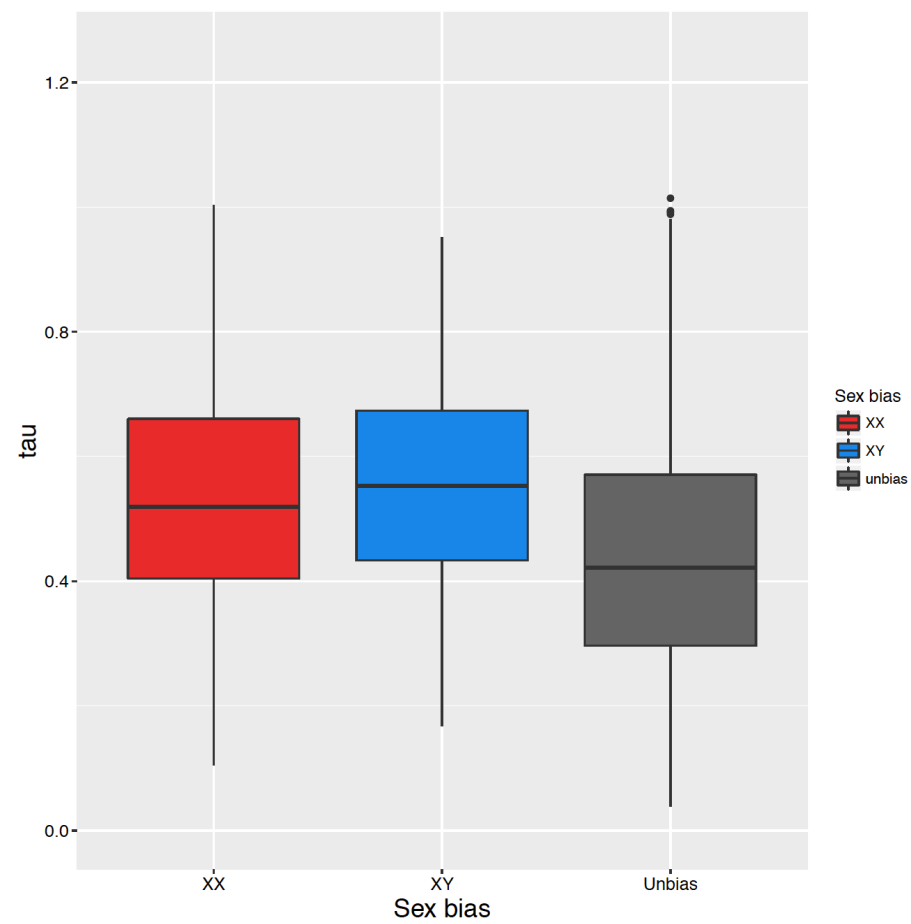


Figure 4.2 Boxplot of dN/dS among female-, male-biased and unbiased genes at G46 stage.

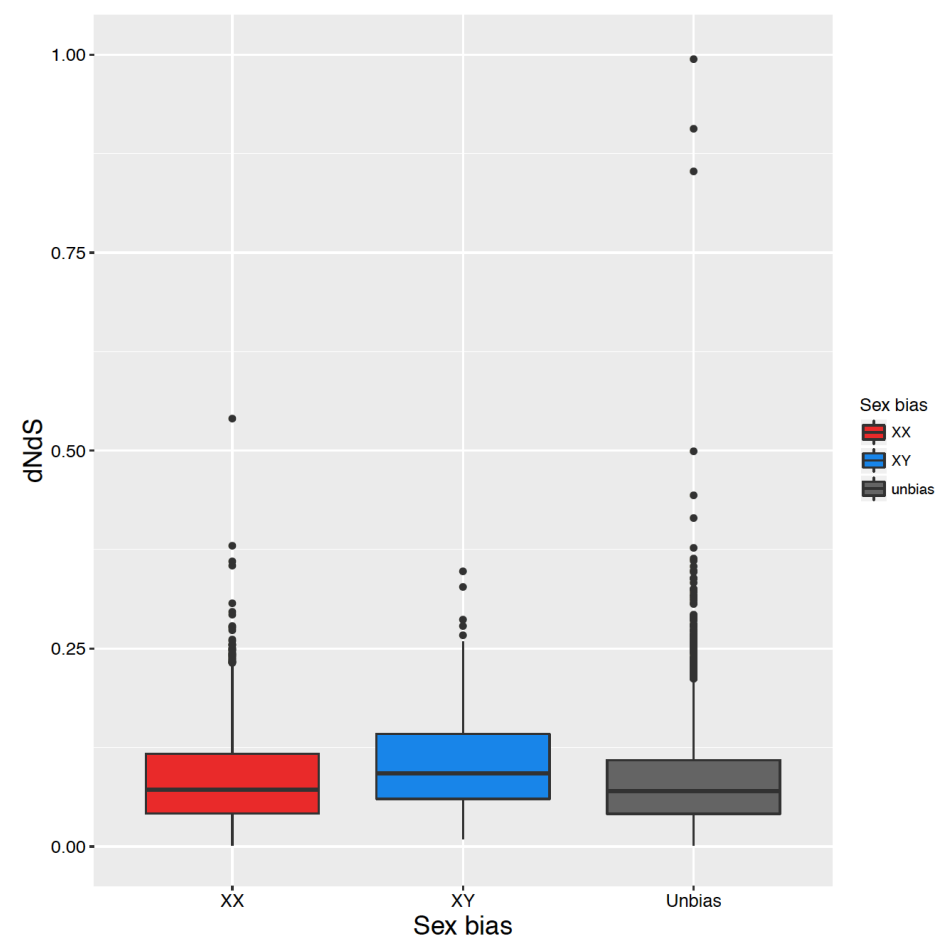
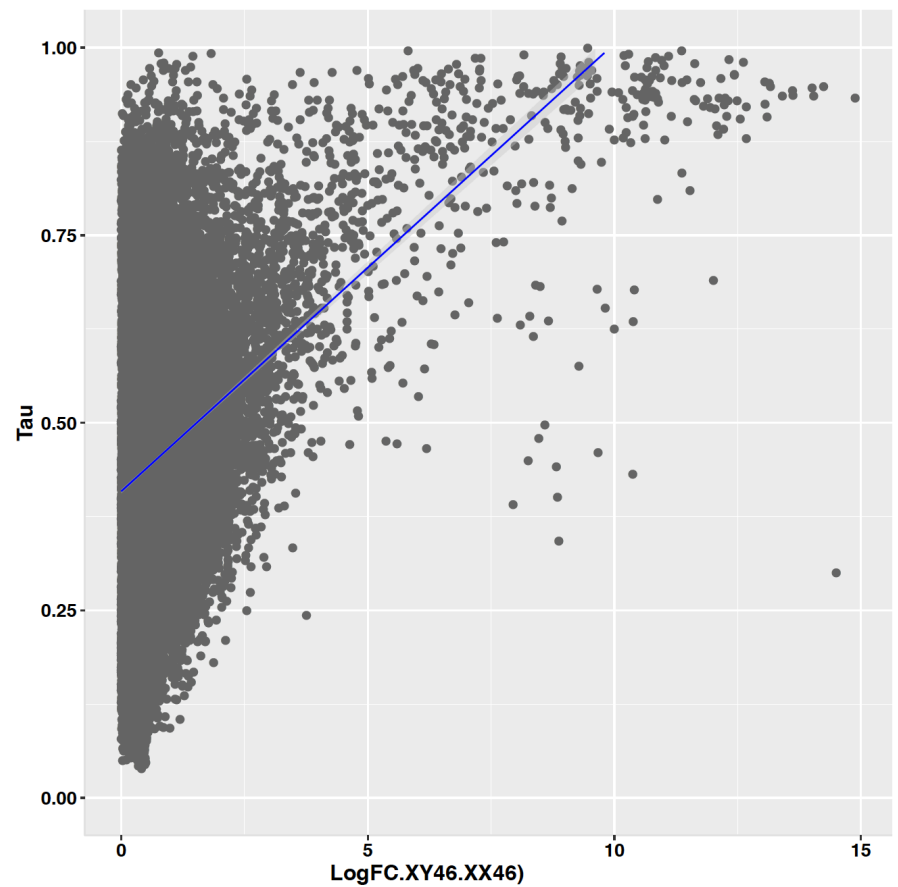


Figure 5. Correlation between absolute values of gene expression ratio $\text{Log}_2(\text{XY}_{46}/\text{XX}_{46})$ (include sex-biased and unbiased genes) and tissue specificity index τ .



Furthermore, we could also ask whether the evolutionary rate of coding sequence can be explained by tissue specificity, or sex bias, or the interaction of the two factors.

```
y1 <- lm(sqrt(dNdS)~sqrt(tau)*bias, g46_tau_dnds)
anova(y1)

#####
Df Sum Sq Mean Sq F value    Pr(>F)
sqrt(tau)      1  1.941  1.94074 212.689 < 2.2e-16 ***
bias           2  0.200  0.10001  10.960 1.770e-05 ***
sqrt(tau):bias  2  0.375  0.18746  20.544 1.275e-09 ***
Residuals    6589 60.123  0.00912
#####
```


For gonad tissues
#####

Figure 6.1. Boxplot of sex-biased genes and the tissue specificity index Tau in gonad tissues.

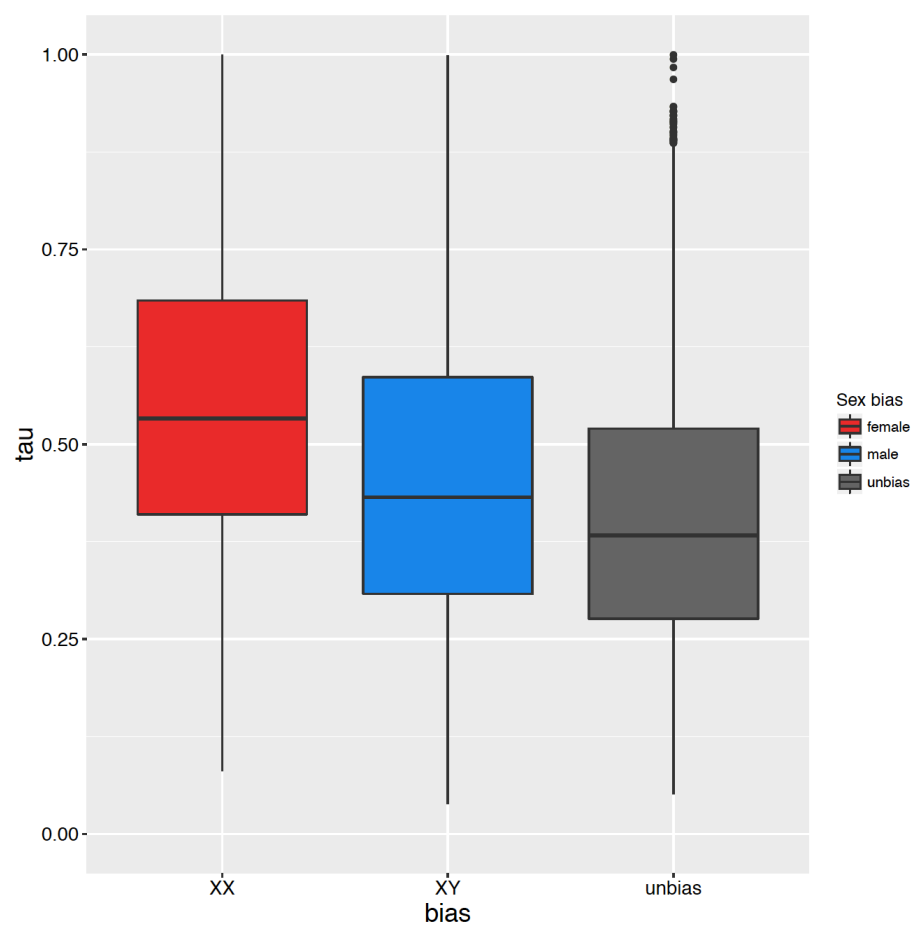


Figure 6.2 Boxplot of dN/dS among female-, male-biased and unbiased genes for gonad tissues.

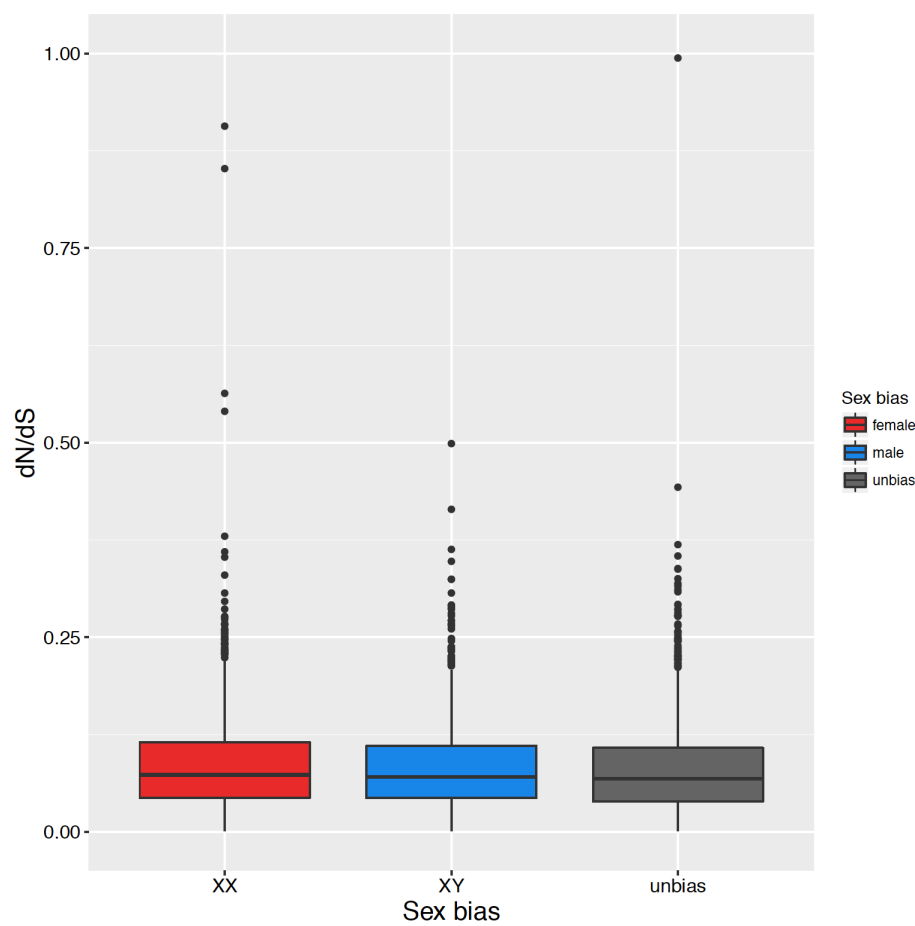
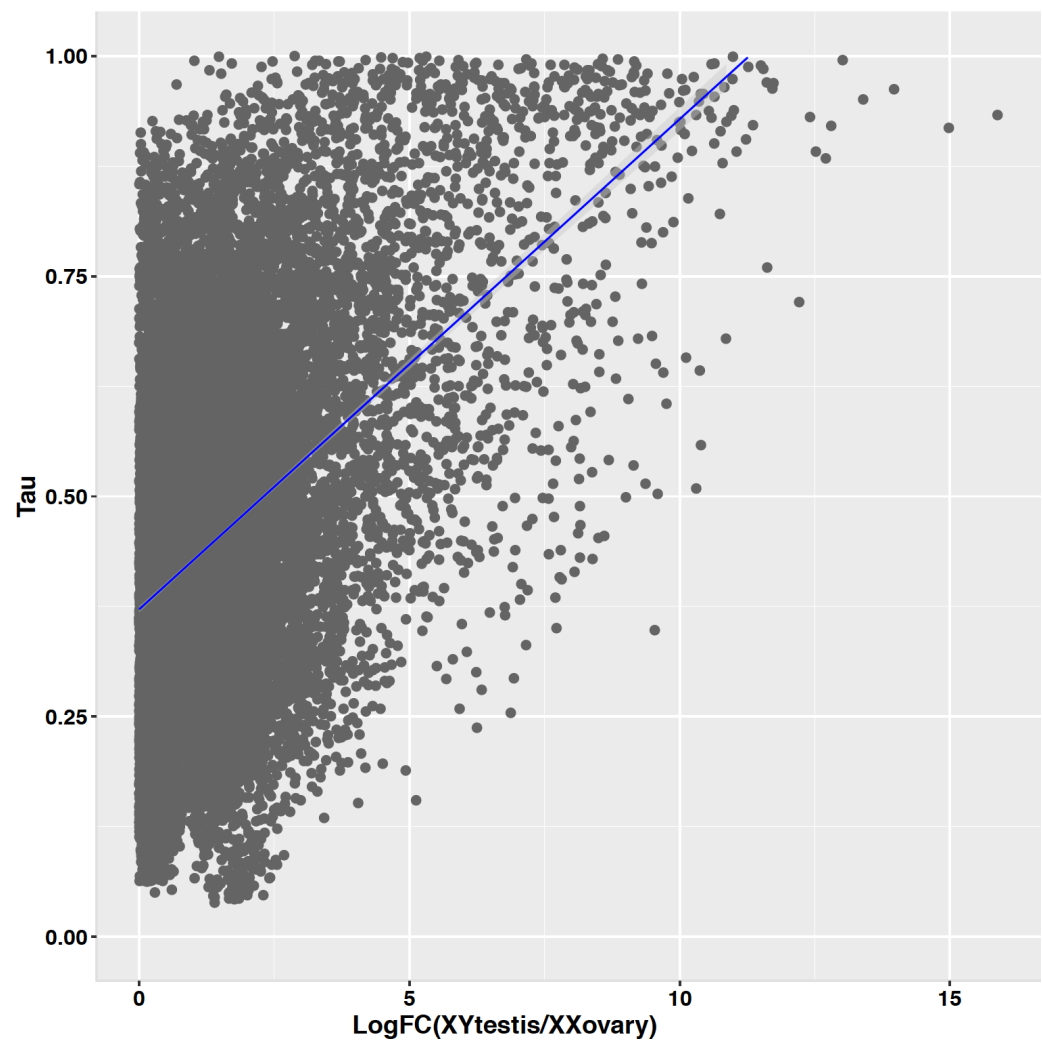


Figure 7. Correlation between absolute values of gene expression ratio $\text{Log}_2(\text{male/female})$ (only for sex-biased and unbiased genes) and tissue specificity index tau.



Similarly, for gonad tissues, we could also ask whether the evolutionary rate of coding sequence can be explained by tissue specificity, or sex bias, or the interaction of the two factors.

```
y <- lm(sqrt(tau)~sqrt(abs(logFC.XY43.XX43))*bias, data=g43_tau)
anova(y)

#####
Response: sqrt(tau)
Df Sum Sq Mean Sq F value Pr(>F)
sqrt(abs(logFC.XY43.XX43))      1 35.32 35.321 1924.0158 < 2.2e-16 ***
bias                          2  0.85  0.424  23.0809 9.663e-11 ***
sqrt(abs(logFC.XY43.XX43)):bias  2  0.31  0.154   8.3743 0.0002314 ***
Residuals                    25568 469.38  0.018
---

#####
For brain tissues
#####
```

Figure 8.1. Boxplot of sex-biased genes and the tissue specificity index Tau in brain tissues.

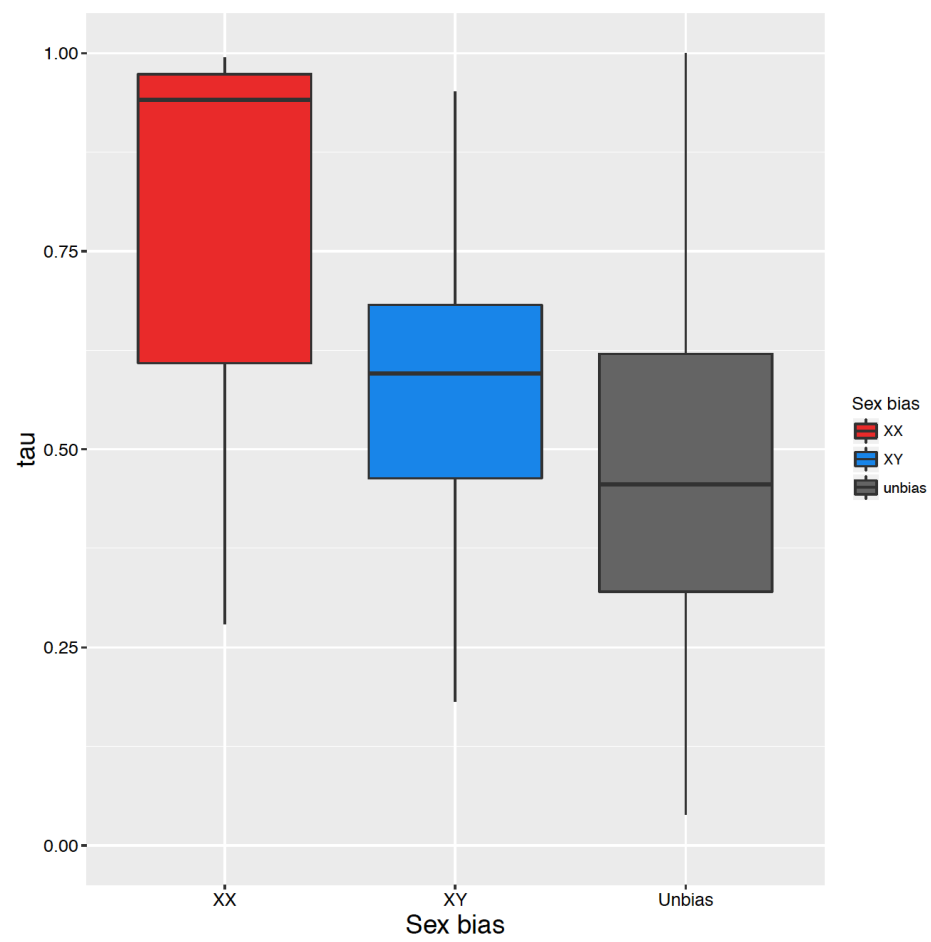


Figure 8.2. Boxplot of dN/dS among female-, male-biased and unbiased genes for brain tissues.

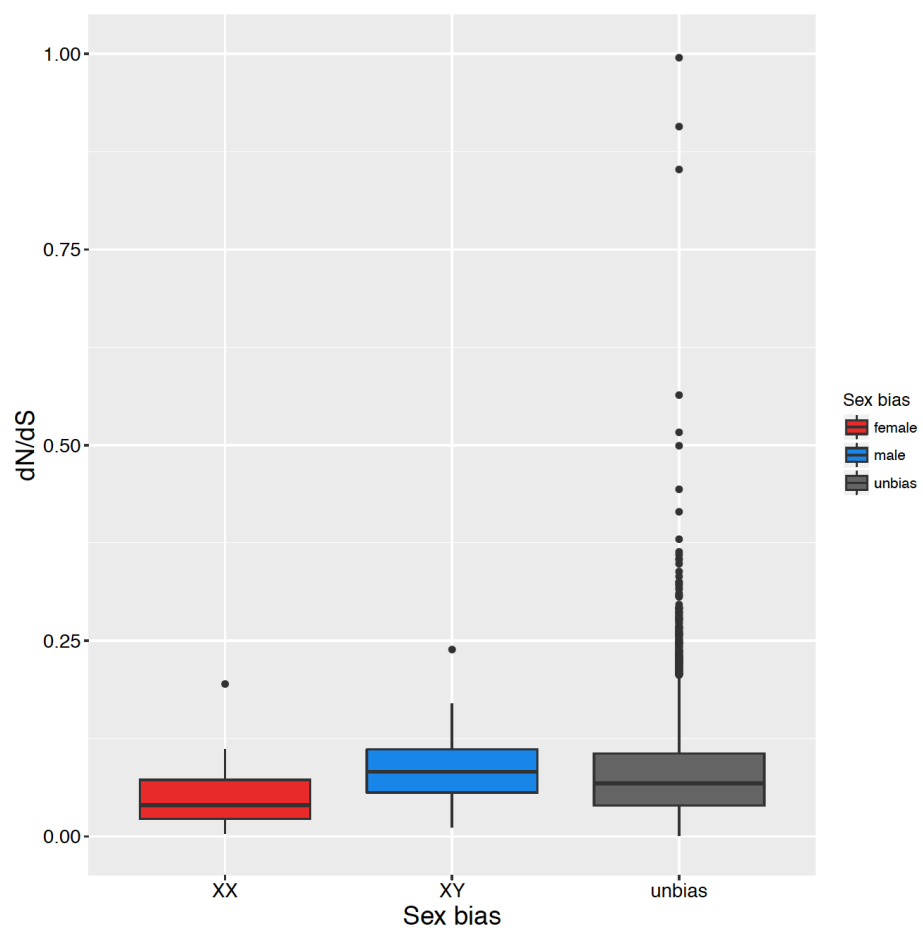
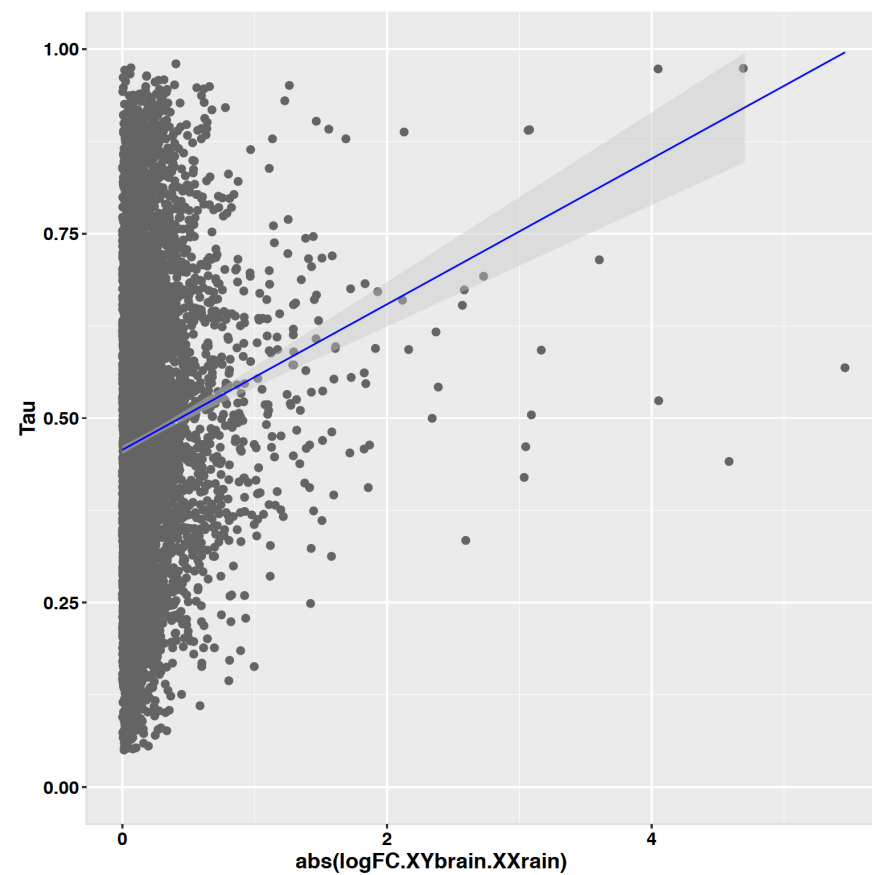


Figure 9. Correlation between absolute values of gene expression ratio $\text{Log}_2(\text{male/female})$ (only for sex-biased and unbiased genes) and tissue specificity index τ .



```
y2 <- lm(sqrt(dNdS) ~ sqrt(tau) * bias, brain_tau_dnds)
anova(y2)
```

```
#####
Df Sum Sq Mean Sq F value Pr(>F)
sqrt(tau)      1  0.645 0.64482 70.8861 <2e-16 ***
bias           2  0.034 0.01724  1.8947 0.1504
sqrt(tau):bias  2  0.034 0.01711  1.8808 0.1525
Residuals    6830 62.129 0.00910
```

```
#####
For liver tissues
#####
```

Figure 10.1. Boxplot of sex-biased genes and the tissue specificity index Tau in liver tissues.

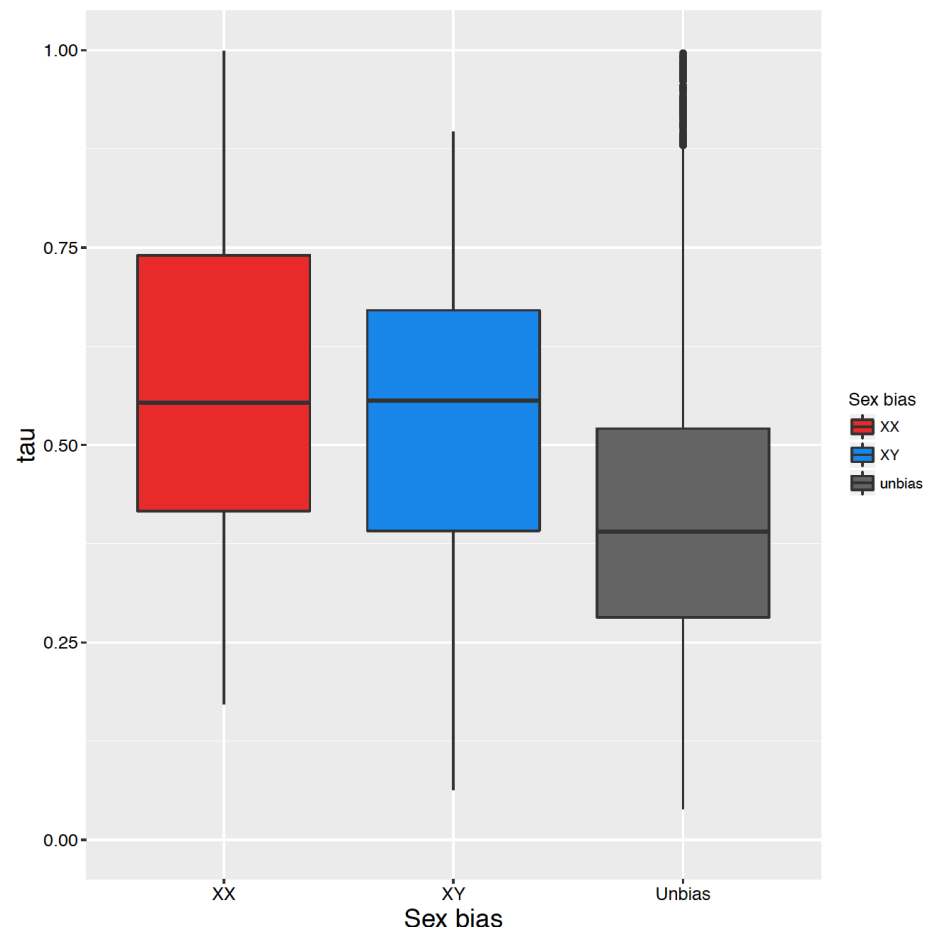


Figure 10.2. Boxplot of dN/dS among female-, male-biased and unbiased genes for liver tissues.

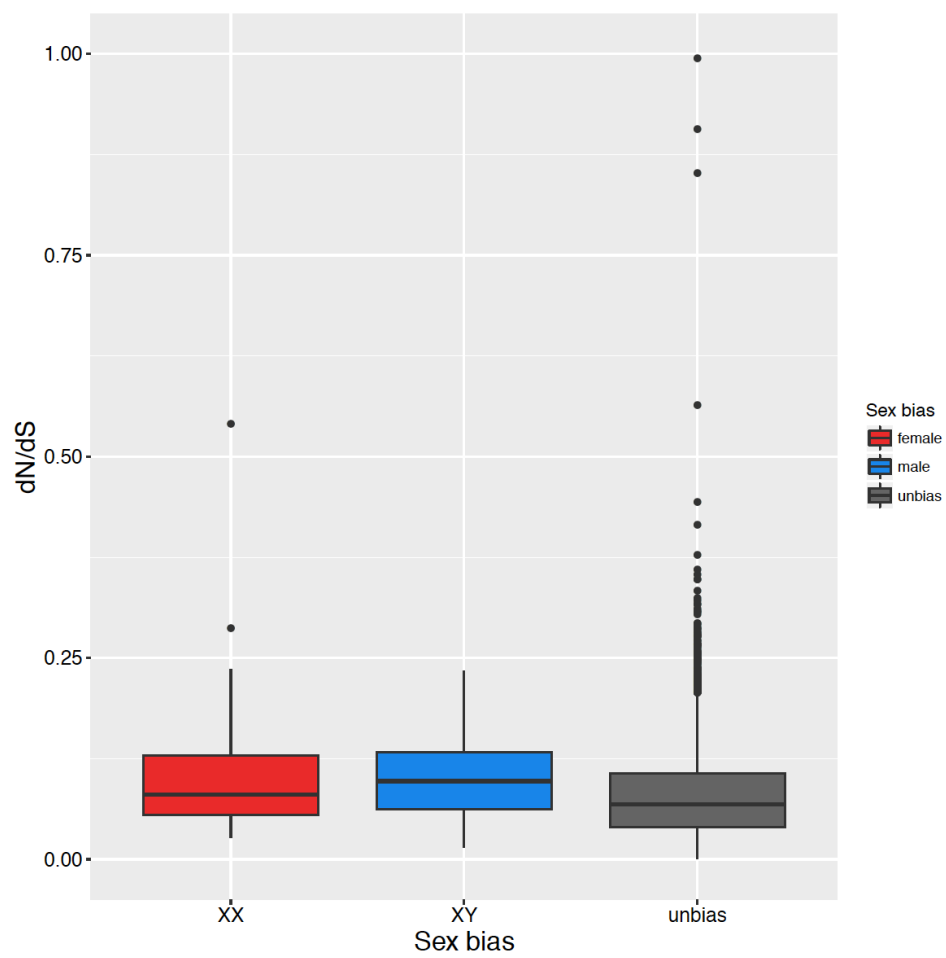
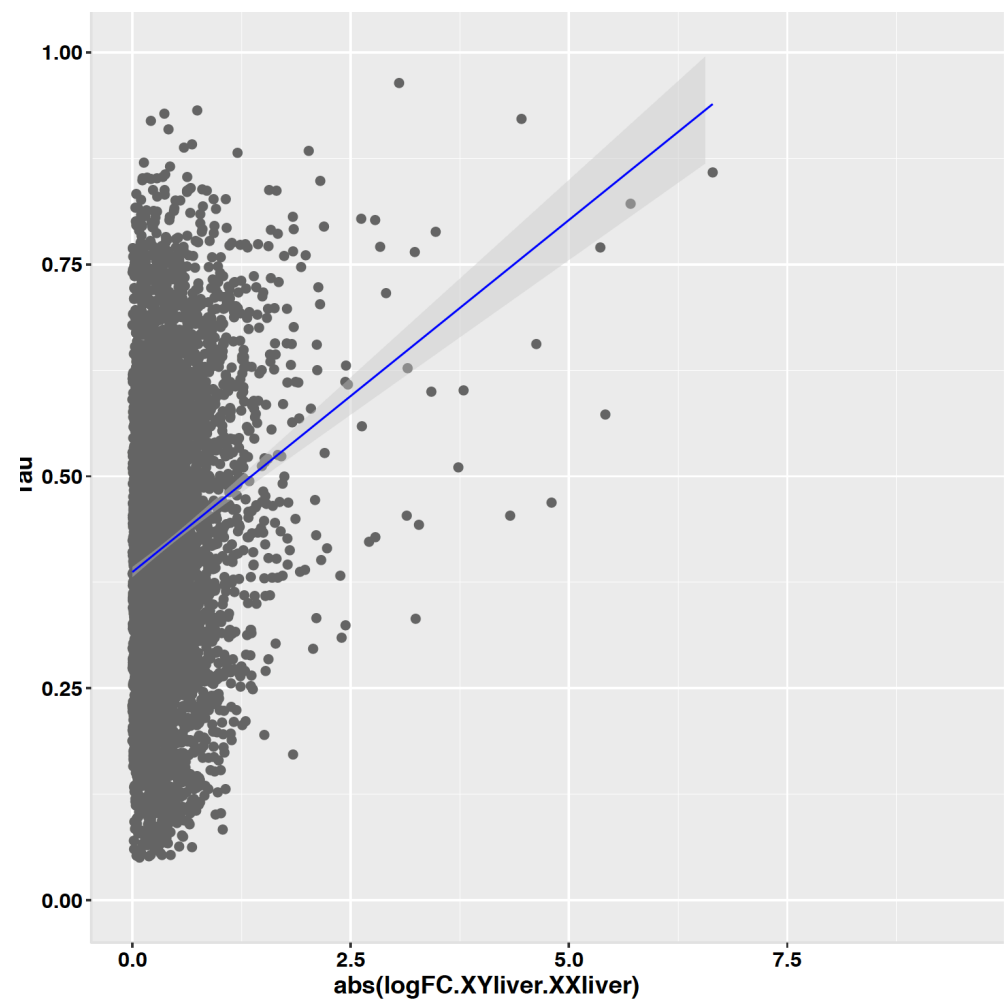


Figure 11. Correlation between absolute values of gene expression ratio $\text{Log}_2(\text{male/female})$ (only for sex-biased and unbiased genes) and tissue specificity index τ .



```
y3 <- lm(sqrt(dNdS) ~ sqrt(tau) * bias, liver_tau_dnds)
anova(y3)
```

```
##
Df Sum Sq Mean Sq F value Pr(>F)
sqrt(tau)      1  1.910  1.91017 212.7090 < 2e-16 ***
bias           2  0.082  0.04080   4.5432 0.01068 *
sqrt(tau):bias  2  0.051  0.02525   2.8121 0.06016 .
Residuals    5453 48.969  0.00898
```