Supplementary Material

Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution

Richard P. Meisel

Data Sets

In the following figures and tables, all data were taken from the following sources:

- Sex-biased expression measurements in *D. melanogaster* are from Parisi et al. (2003), Ranz et al. (2003), Gibson et al. (2004), Stolc et al. (2004), McIntyre et al. (2006), and Ayroles et al. (2009). Sex-biased expression data in mouse are from Yang et al. (2006) and analyzed in Mank et al. (2008).

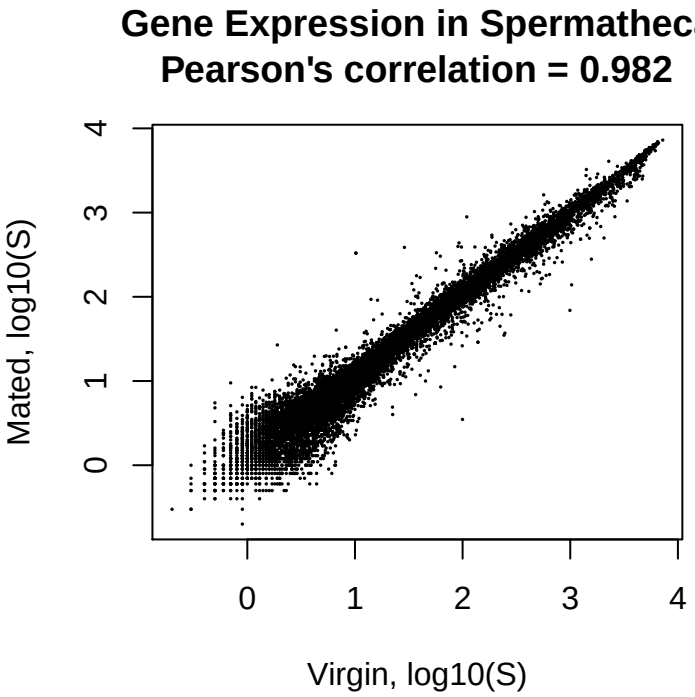
- P-values of sex-biased expression and M/F expression ratios in *D. melanogaster* for individual experiments and the meta-analysis were obtained from SEBIDA (`http://www.sebida.de`; Gnad and Parsch, 2006).

- Expression measurements from individual body parts and whole flies (combined sexes) were obtained from FlyAtlas (`http://www.flyatlas.org`; Chintapalli, Wang and Dow, 2007). Expression measurements from individual mouse tissues were obtained from UniGene (`http://www.ncbi.nlm.nih.gov/unigene/`) build #186.

- Rates of protein coding sequence evolution (d_N/d_S) along the *D. melanogaster* branch and within the *melanogaster* subgroup and group were obtained from Larracuente et al. (2008). Pair-wise d_N/d_S between *D. melanogaster* and *D. simulans* or *D. yakuba* were obtained from `http://www.sebida.de` (Gnad and Parsch, 2006). Pair-wise d_N/d_S between mouse and rat orthologs were obtained from Ensembl (`http://www.ensembl.org`).

Evaluating spermatheca expression data

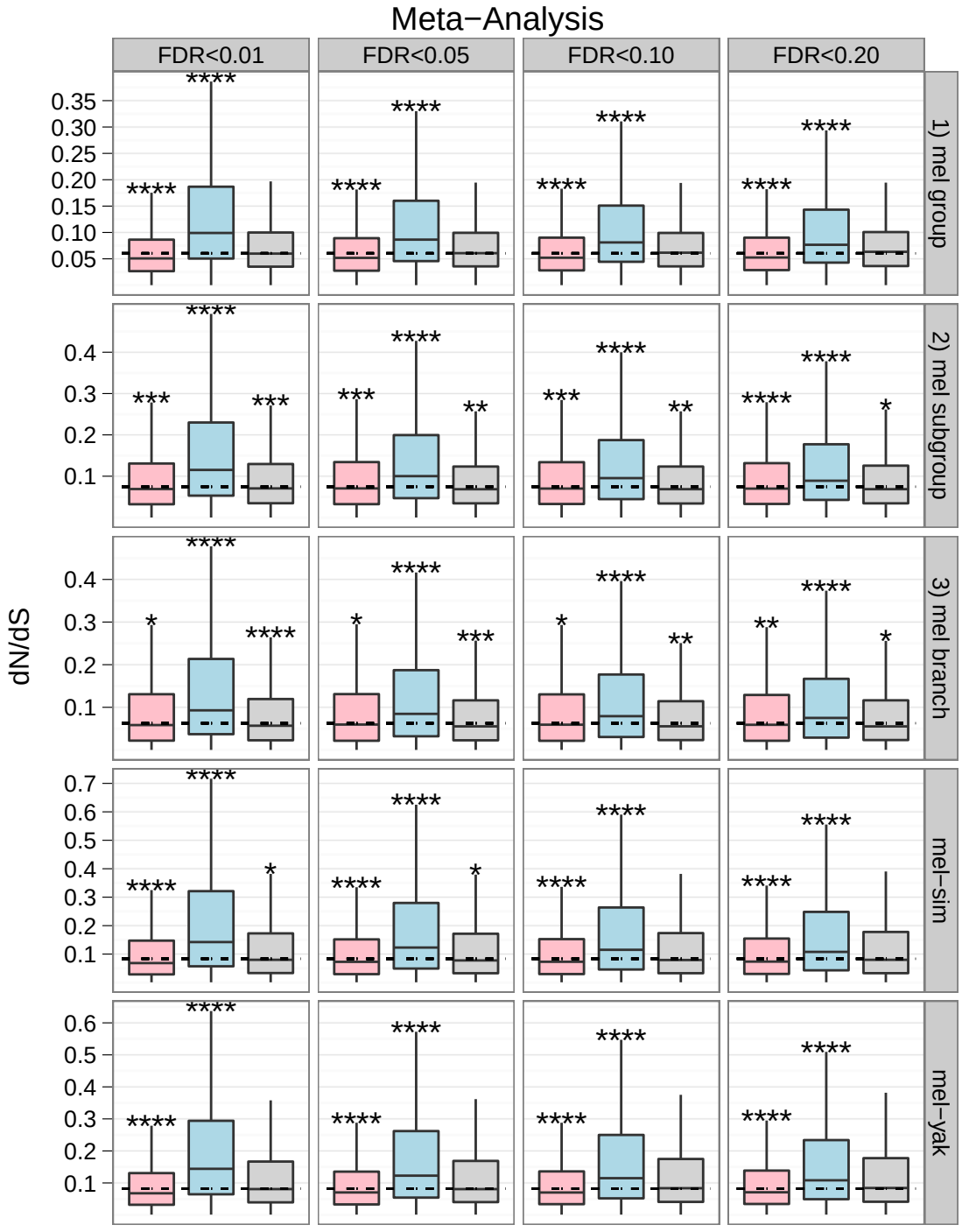
Expression signal intensity (S) is strongly correlated when measured in virgin and mated spermatheca.

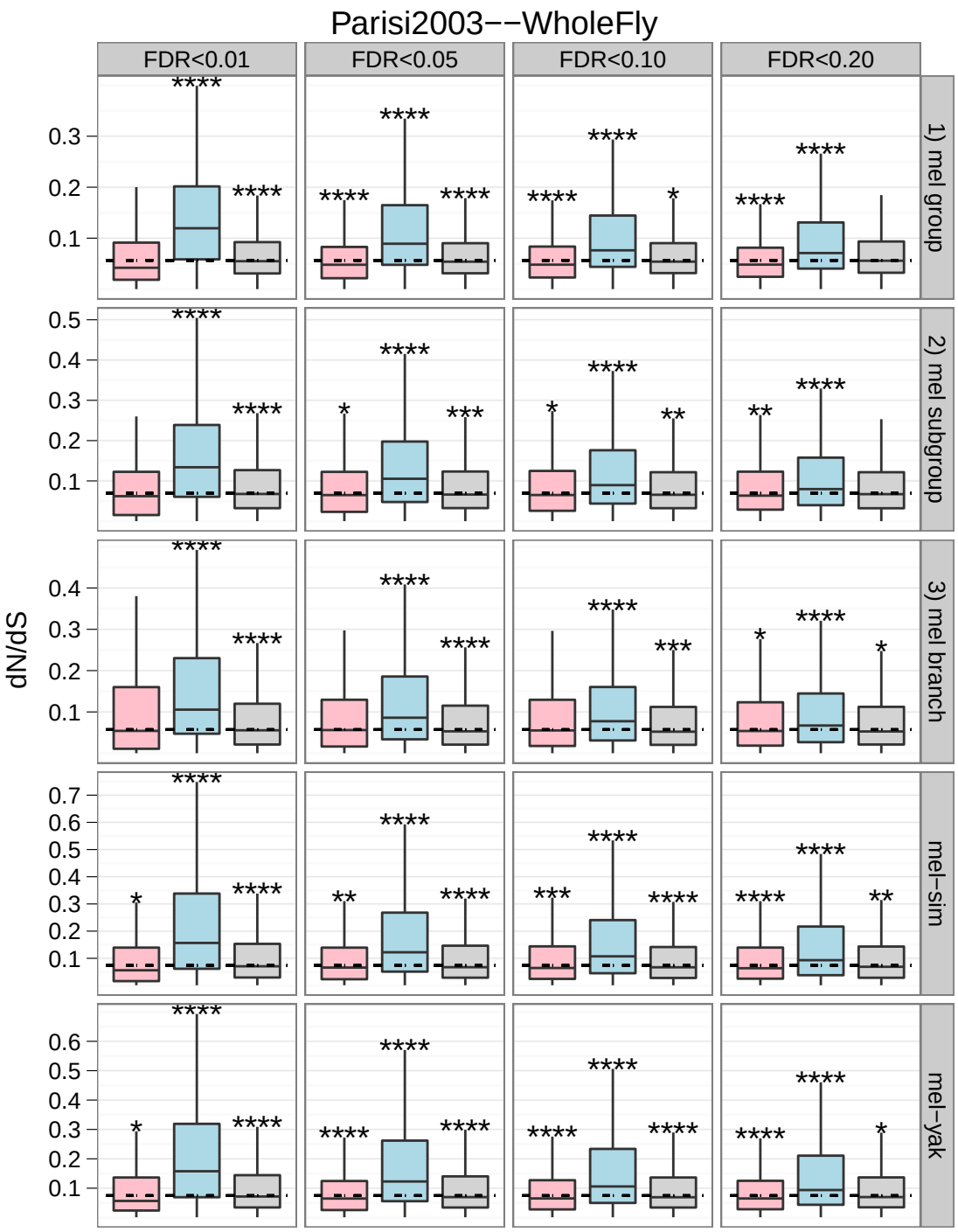


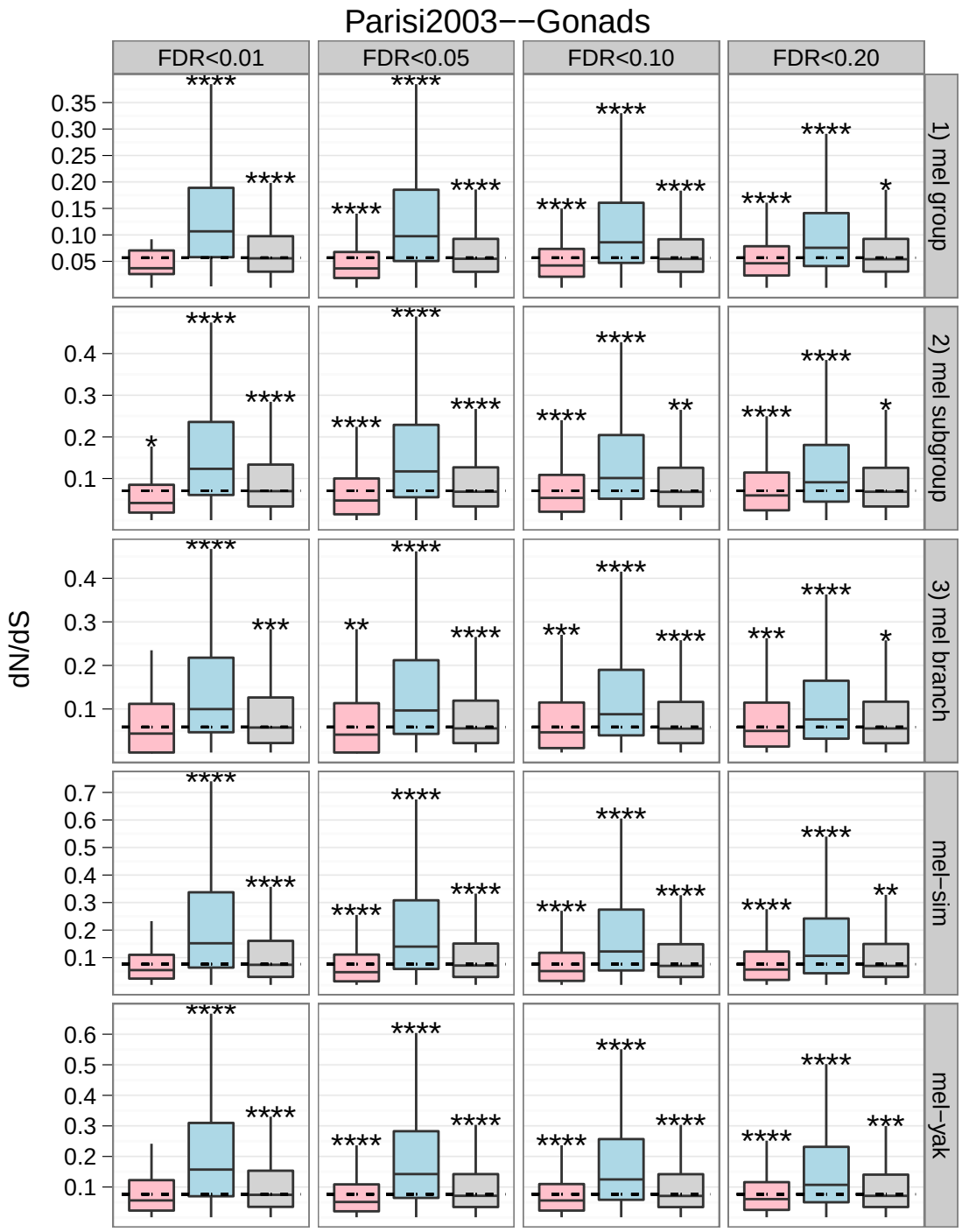
Sex-biased expression and rates of evolution

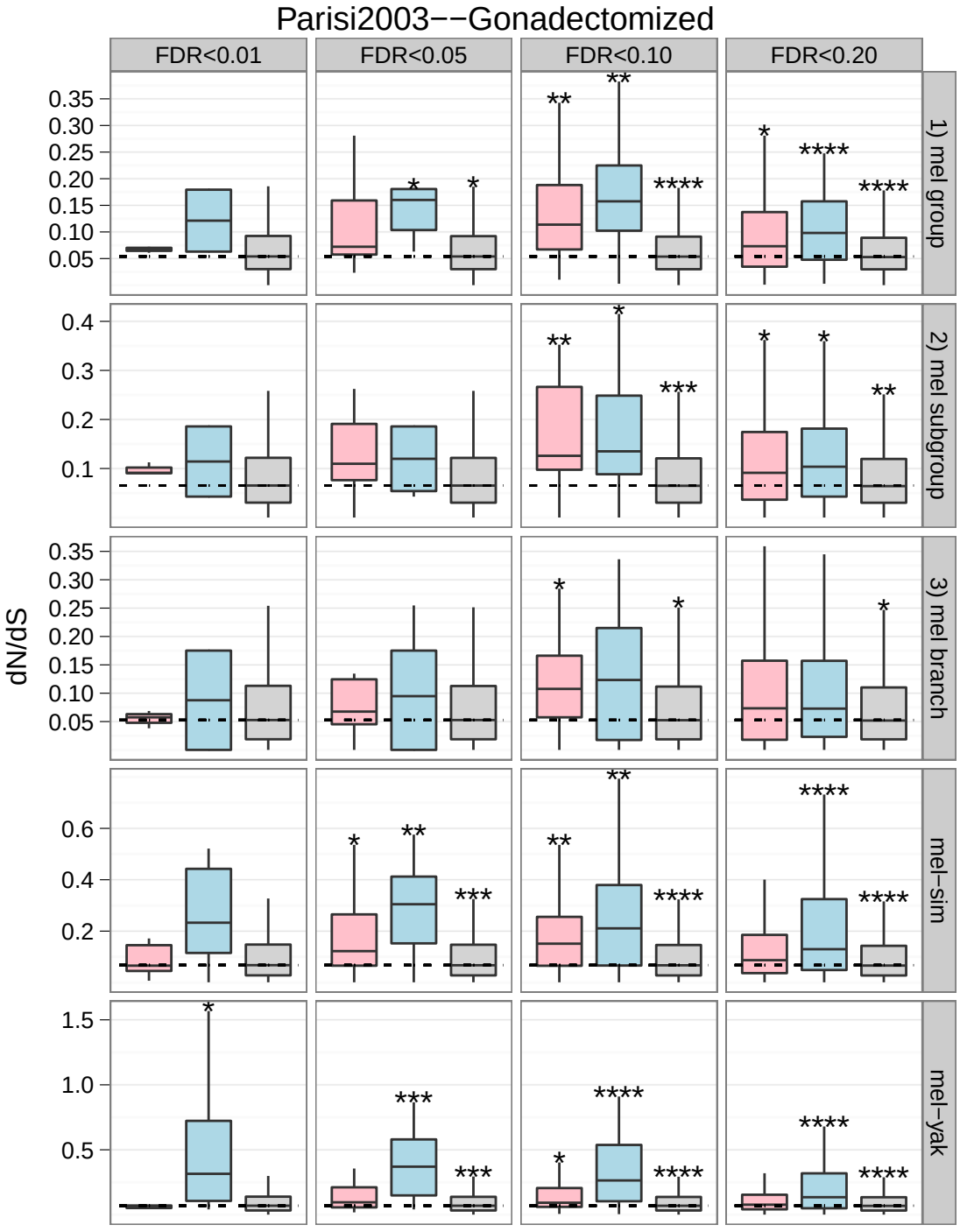
Evolutionary rates of female-, male-, and un-biased genes in *D. melanogaster*

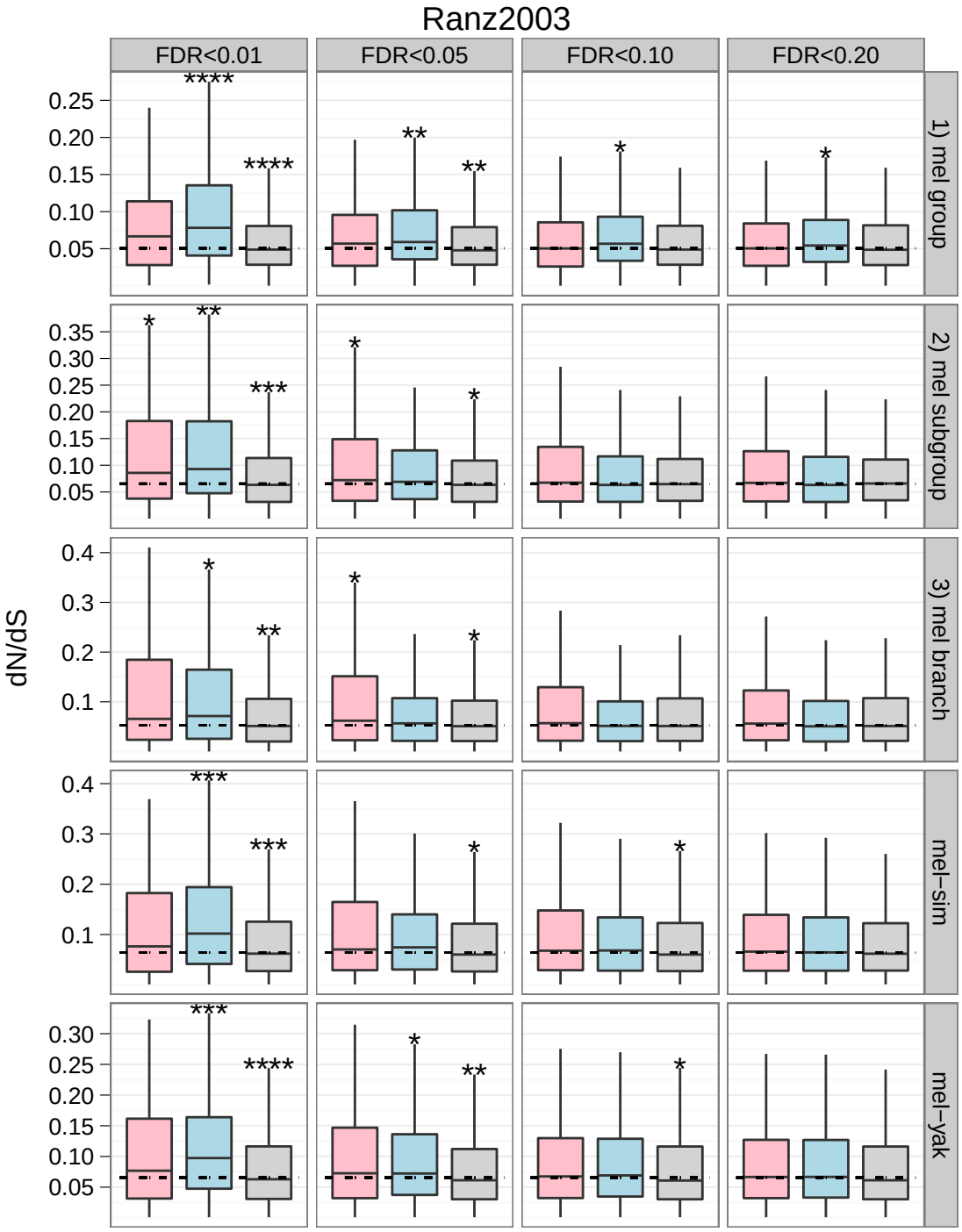
Boxplots show the distribution of d_N/d_S for female-biased (red), male-biased (blue), and unbiased (gray) genes, with outliers omitted. Sex-biased genes were identified for each data set with a large enough sample size to compare d_N/d_S between sex-bias classes (the data sets are indicated at the top of each page), using four different FDR cut-offs (0.01, 0.05, 0.10, 0.20). d_N/d_S was calculated five ways: along all the branches in the *melanogaster* species group (mel group); along on the branches in the *melanogaster* subgroup (mel subgroup); along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (mel branch); pair-wise between *D. melanogaster* and *D. simulans* (mel-sim); and pair-wise between *D. melanogaster* and *D. yakuba* (mel-yak). The median d_N/d_S value across all genes within each panel is indicated by a dashed line. Estimates of d_N/d_S for each sex-bias class were compared to the other classes using a Mann-Whitney test, and asterisks indicate significant differences ($*P < 10^{-2}$, $**P < 10^{-4}$, $***P < 10^{-6}$, $****P < 10^{-8}$).

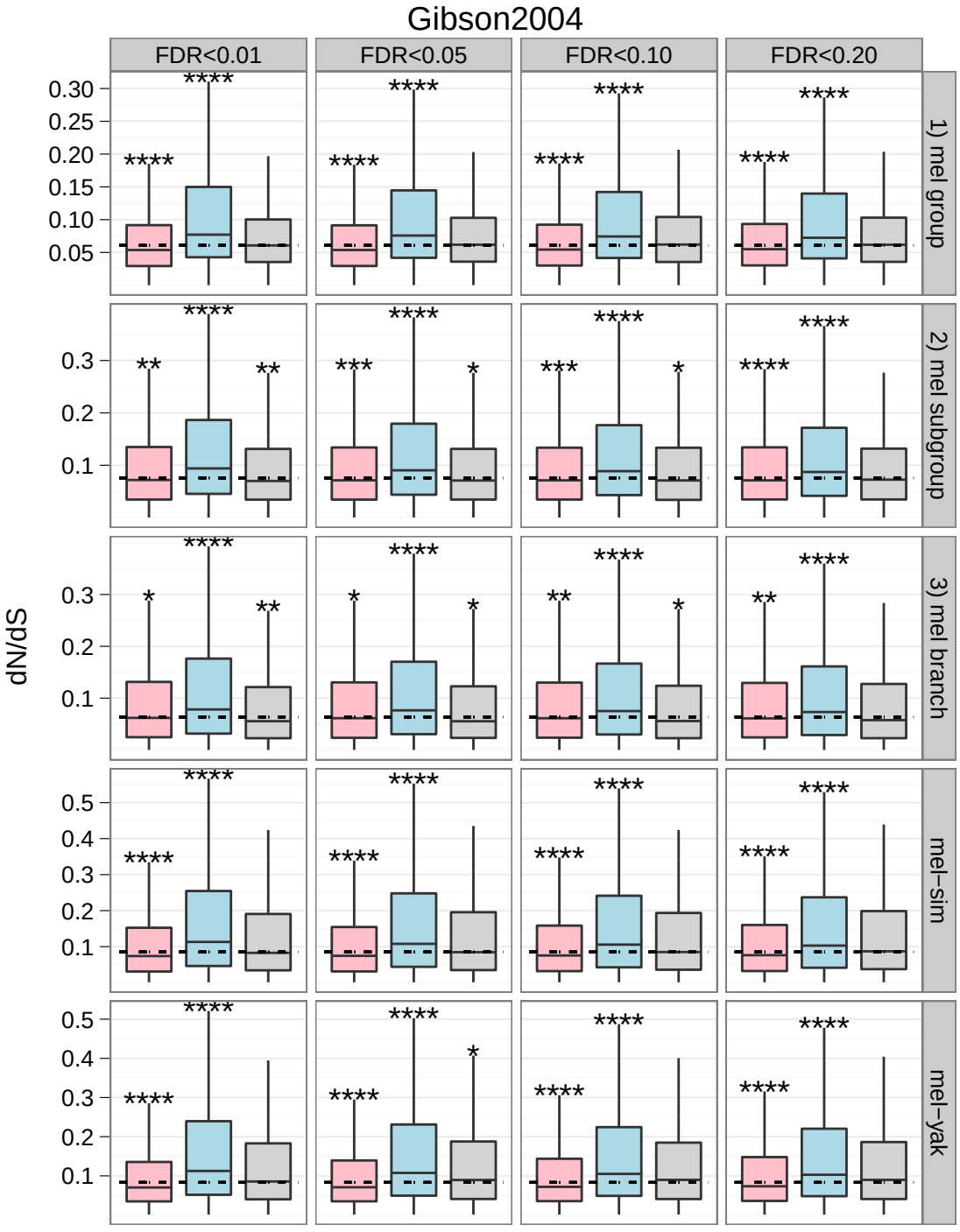

Meta−Analysis
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS

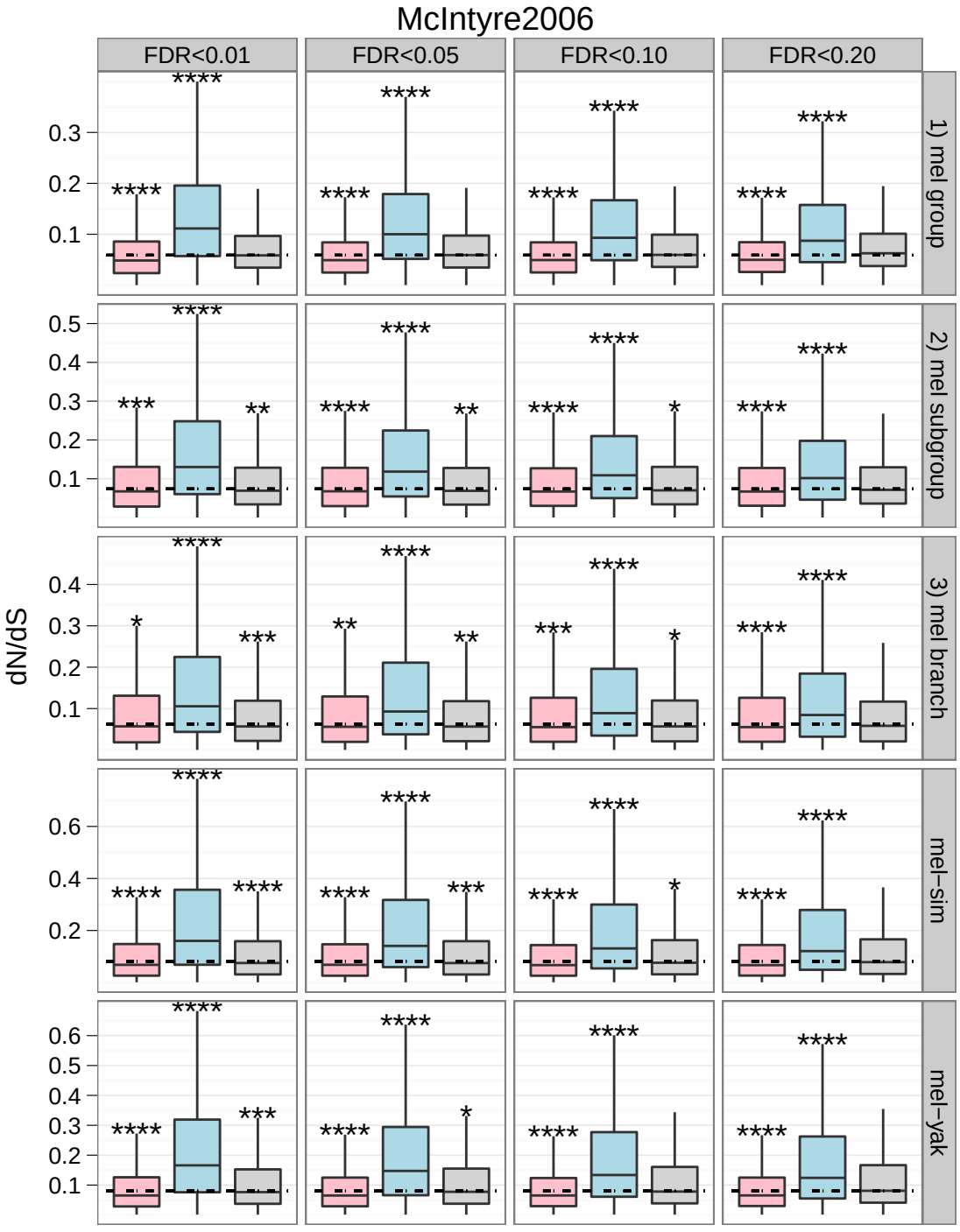

Parisi2003−−WholeFly
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS

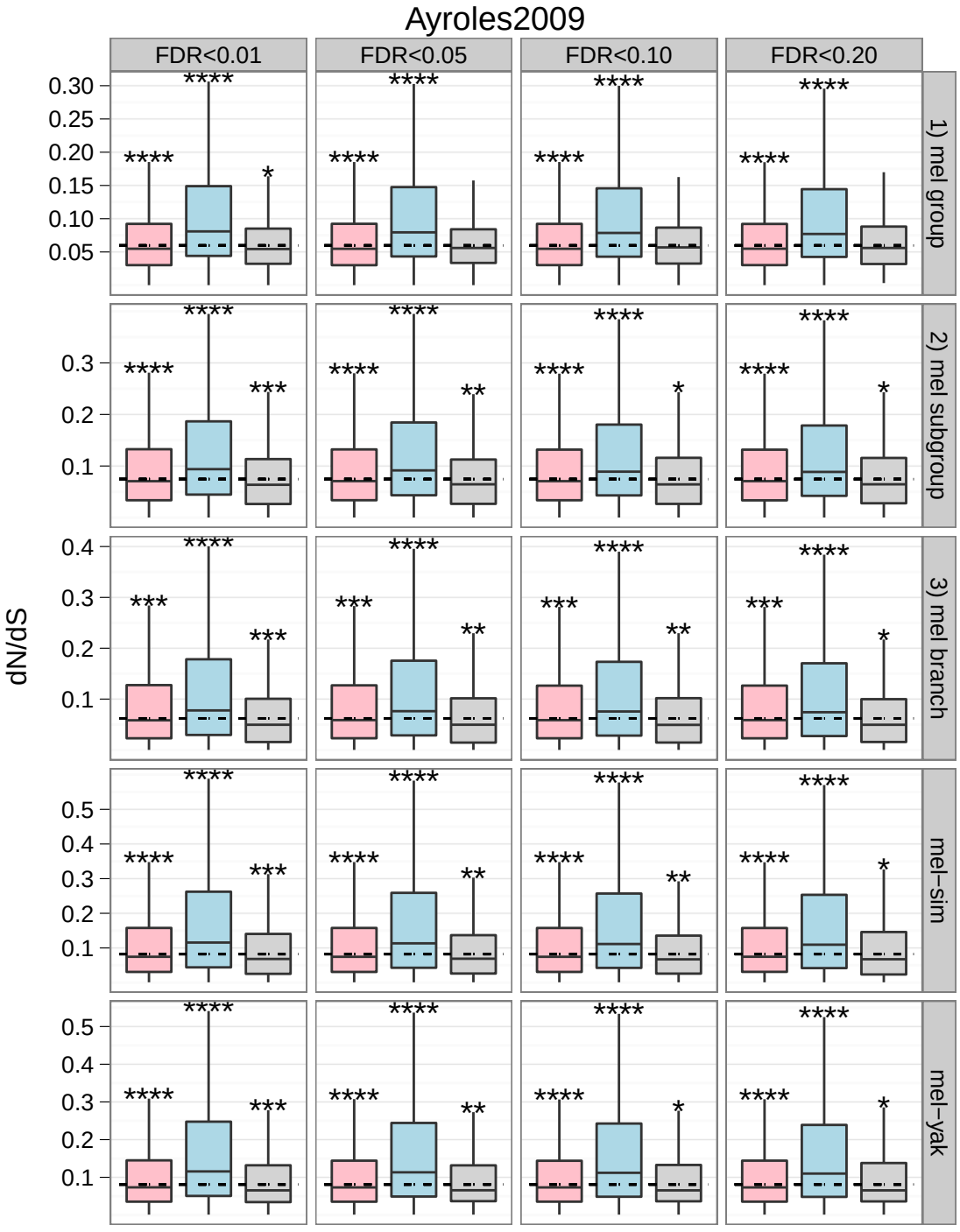

Parisi2003−−Gonads
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS


Parisi2003−−Gonadectomized
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS


Ranz2003
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS

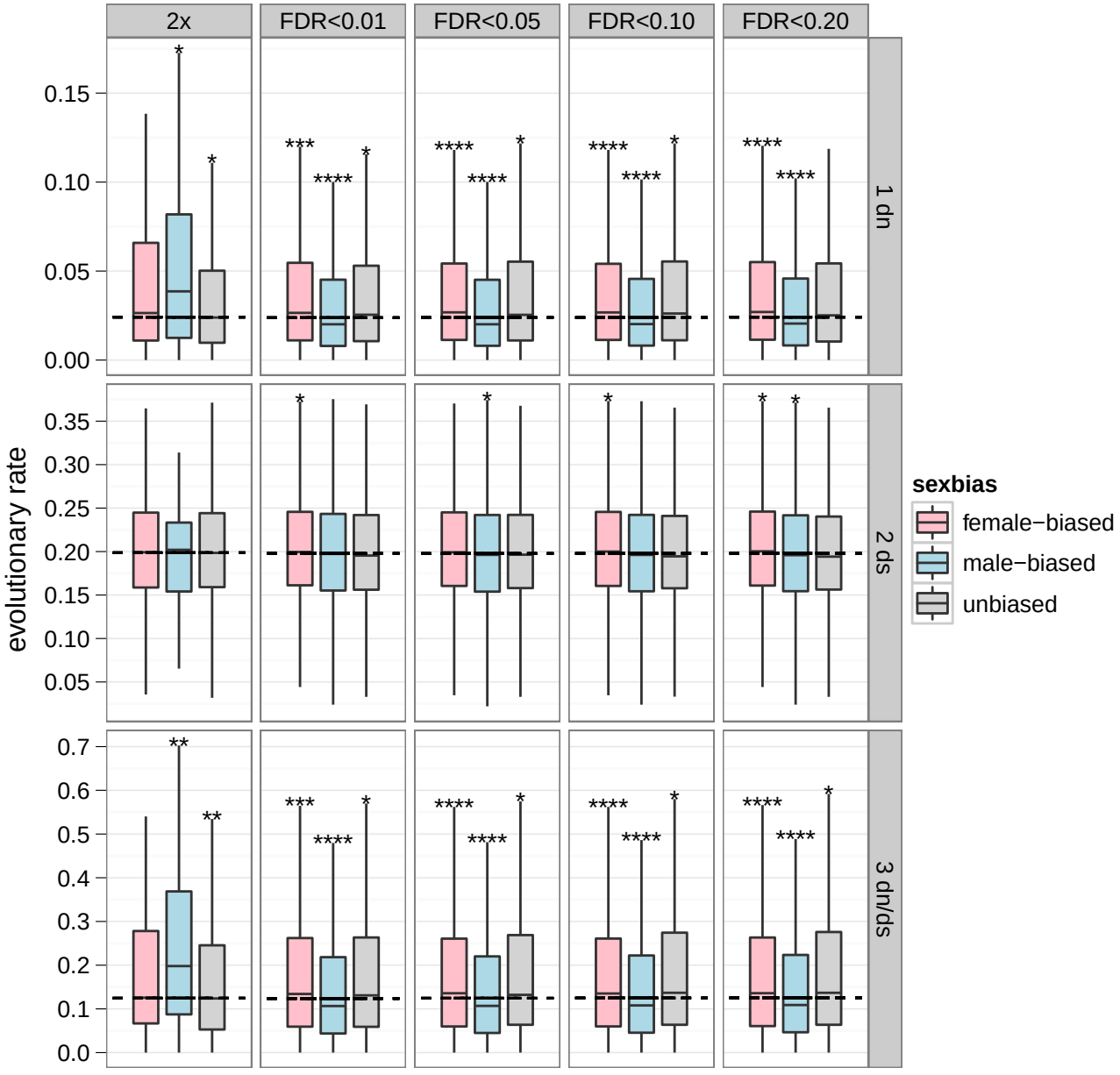

Gibson2004
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS


McIntyre2006
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS


Ayroles2009
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS

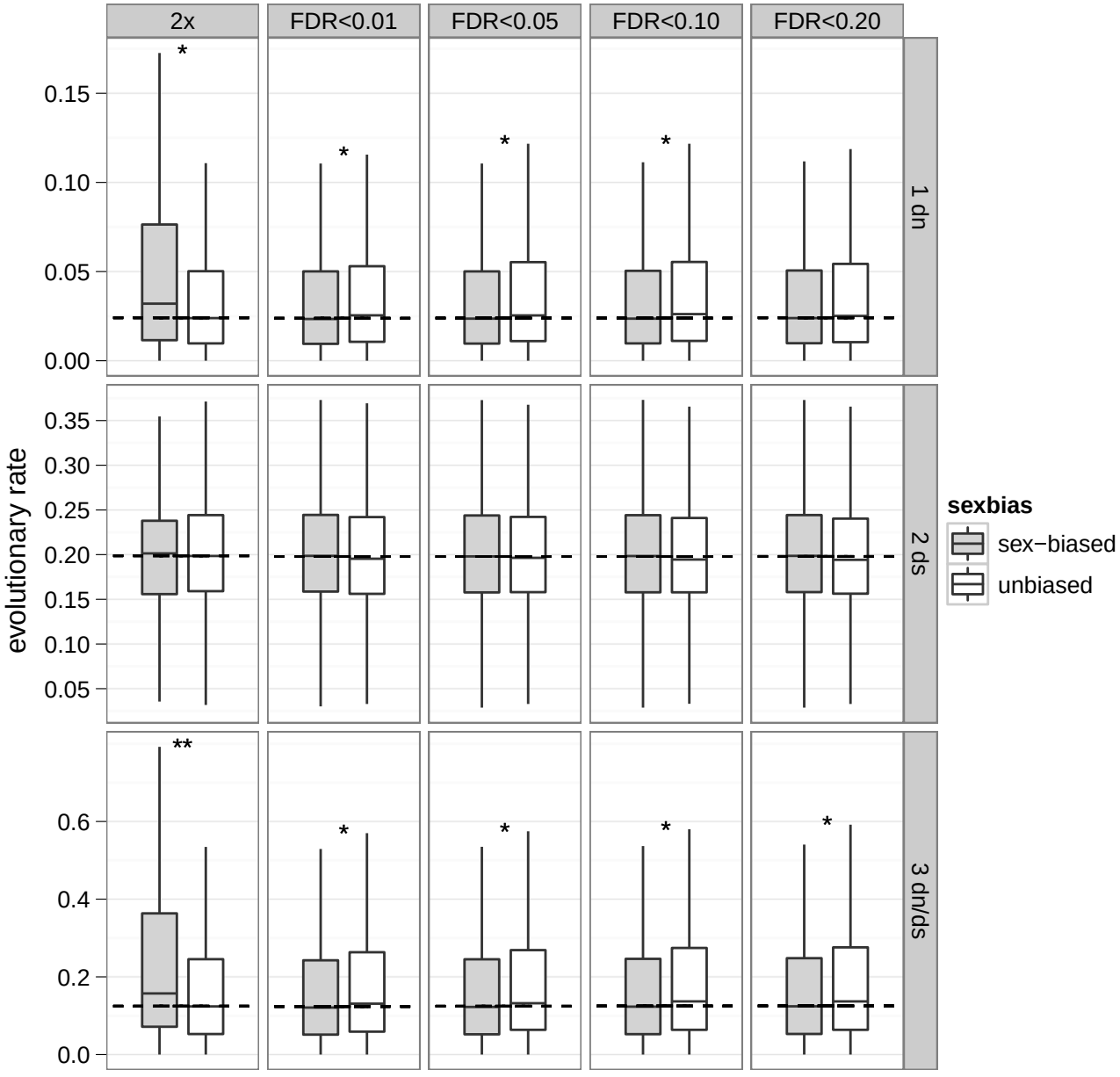
Evolutionary rates of female-, male-, and un-biased genes in mouse

Boxplots show the distribution of d_N/d_S for female-biased (red), male-biased (blue), and unbiased (gray) genes, with outliers omitted. Sex-biased genes were identified in each of 4 tissues (liver, adipose, brain, and muscle) using either a 2-fold cutoff (2x) or one of four FDR cutoffs (0.01, 0.05, 0.10, or 0.20). A sex-biased gene is sex-biased, in the same direction, in at least one tissue and unbiased in all other tissues. Evolutionary rates were estimated as d_N, d_S, or d_N/d_S. The median evolutionary rate within each panel is indicated by a dashed line. Estimates of evolutionary rate for each sex-bias class were compared to the other classes within each panel using a Mann-Whitney test, and asterisks indicate significant differences ($*P < 10^{-2}$, $**P < 10^{-4}$, $***P < 10^{-6}$, $****P < 10^{-8}$).


2x
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1 dn
2 ds
3 dn/ds
evolutionary rate
sexbias
female−biased
male−biased
unbiased

Evolutionary rates of sex-biased and unbiased genes in mouse

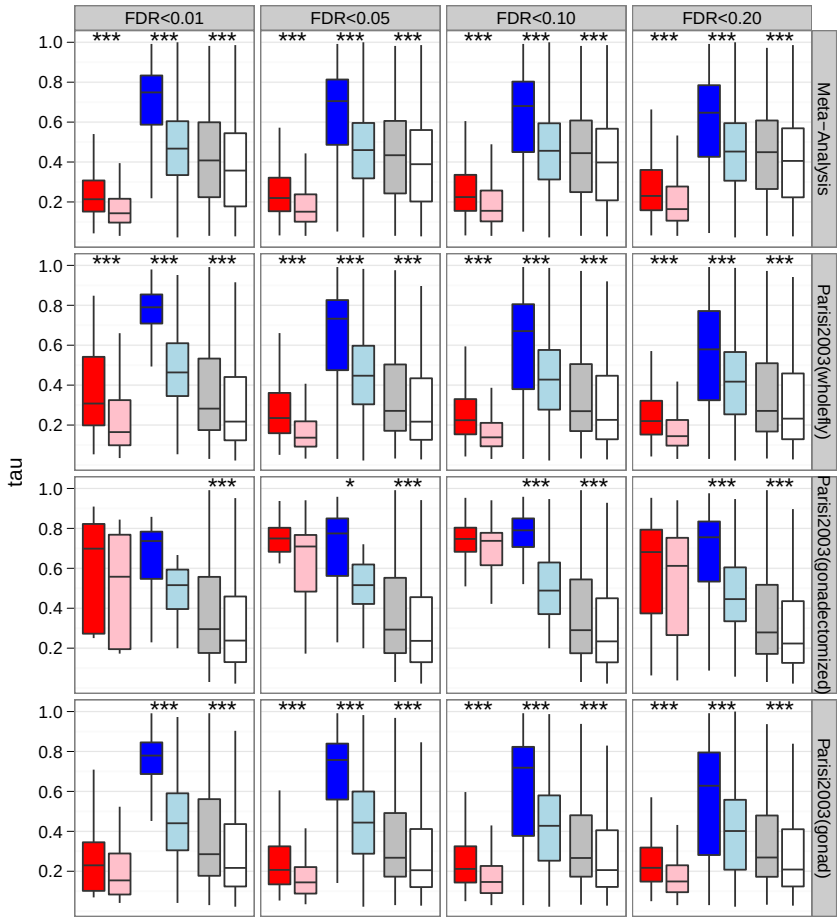
Boxplots show the distribution of d_N/d_S for sex-biased (gray) and unbiased (white) genes, with outliers omitted. Sex-biased genes were identified in each of 4 tissues (liver, adipose, brain, and muscle) using either a 2-fold cutoff (2x) or one of four FDR cutoffs (0.01, 0.05, 0.10, or 0.20). A sex-biased gene is sex-biased, in the same direction, in at least one tissue and unbiased in all other tissues. Evolutionary rates were estimated as d_N, d_S, or d_N/d_S. The median evolutionary rate within each panel is indicated by a dashed line. Estimates of evolutionary rate were compared between sex-biased and unbiased genes within each panel using a two sample Wilcoxon test, and asterisks indicate significant differences ($^*P < 10^{-2}$, $^{**}P < 10^{-4}$, $^{***}P < 10^{-6}$, $^{****}P < 10^{-8}$).

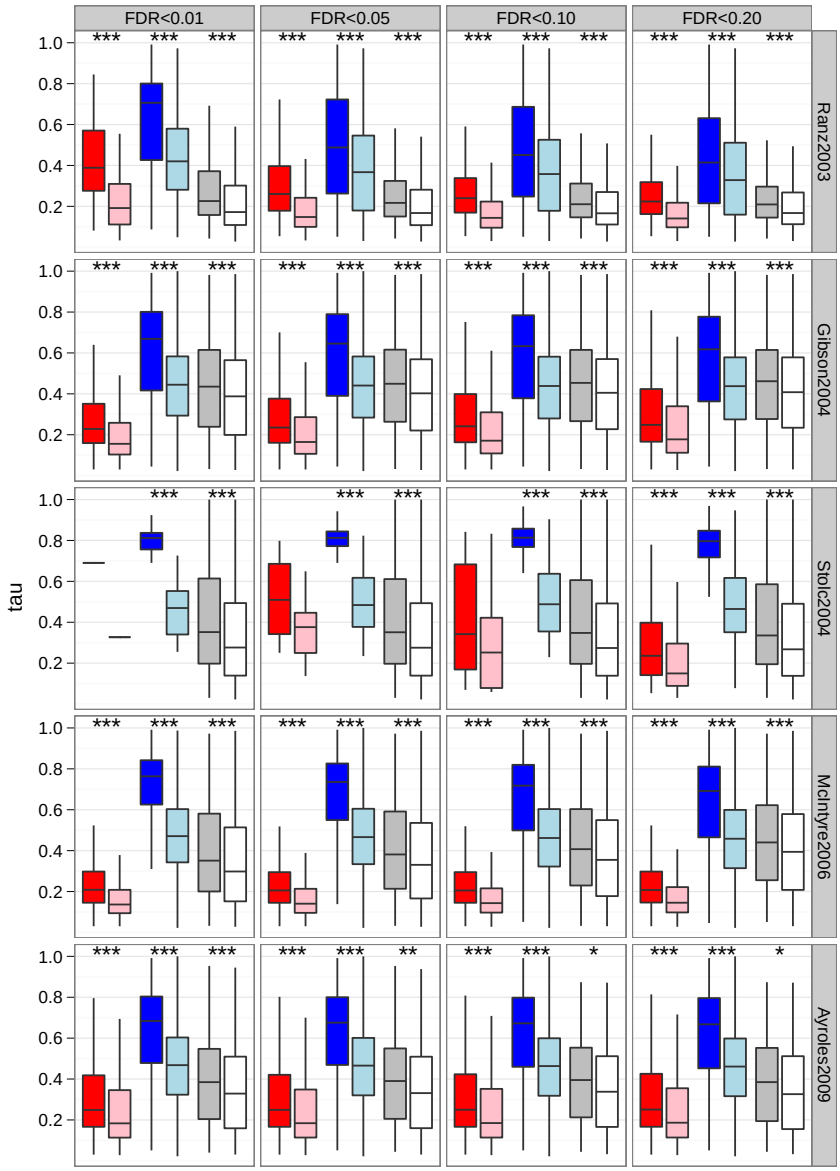

2x
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1 dn
2 ds
3 dn/ds
evolutionary rate
0.15
0.10
0.05
0.00
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.6
0.4
0.2
0.0
*
**
sexbias
sex−biased
unbiased

Relationship between sex-biased expression and tissue-specificity in *D. melanogaster*

Sex-biased genes and tissue-specificity

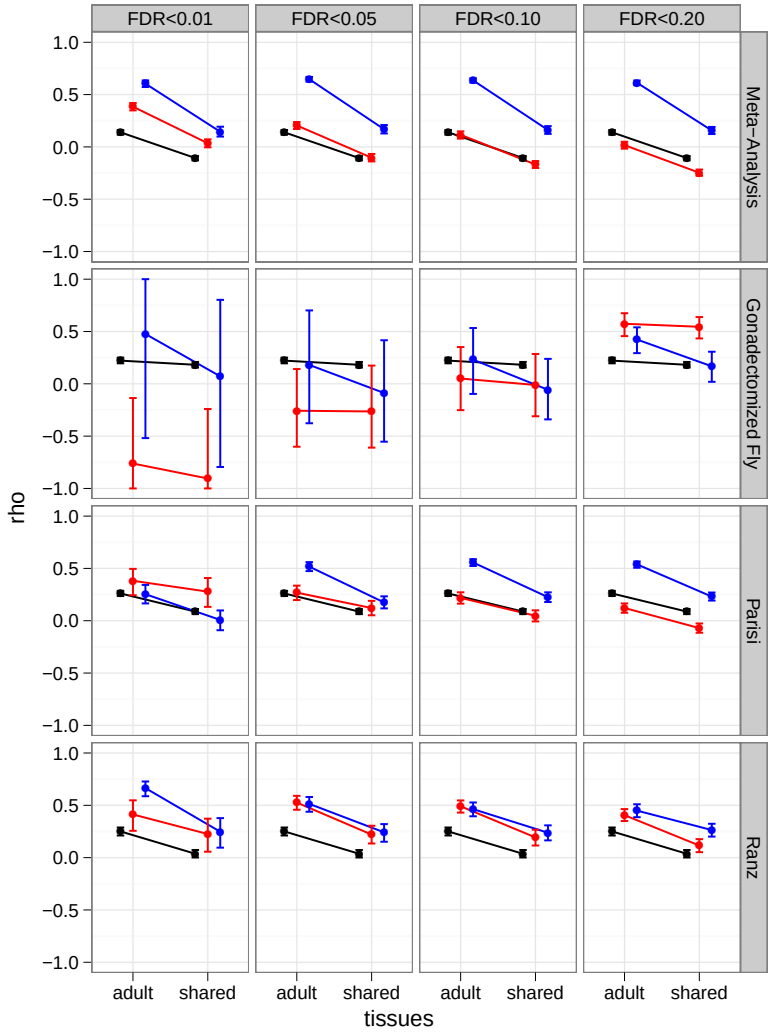
Boxplots show the distribution of τ for female-biased (red), male-biased (blue), and unbiased (gray/white) genes as determined using the meta-analysis and individual data sets. For each data set, sex-biased genes were called with four FDR cut-offs (0.01, 0.05, 0.10, 0.20). Measurements of τ were carried out on all adult tissues (dark) and on non-sex-limited adult tissues (light). Asterisks indicate significant differences in τ when measured in all adult tissues compared to when measured in non-sex-specific tissues using a Mann-Whitney test $*P < 10^{-2}$, $**P < 10^{-4}$, $***P < 10^{-6}$).

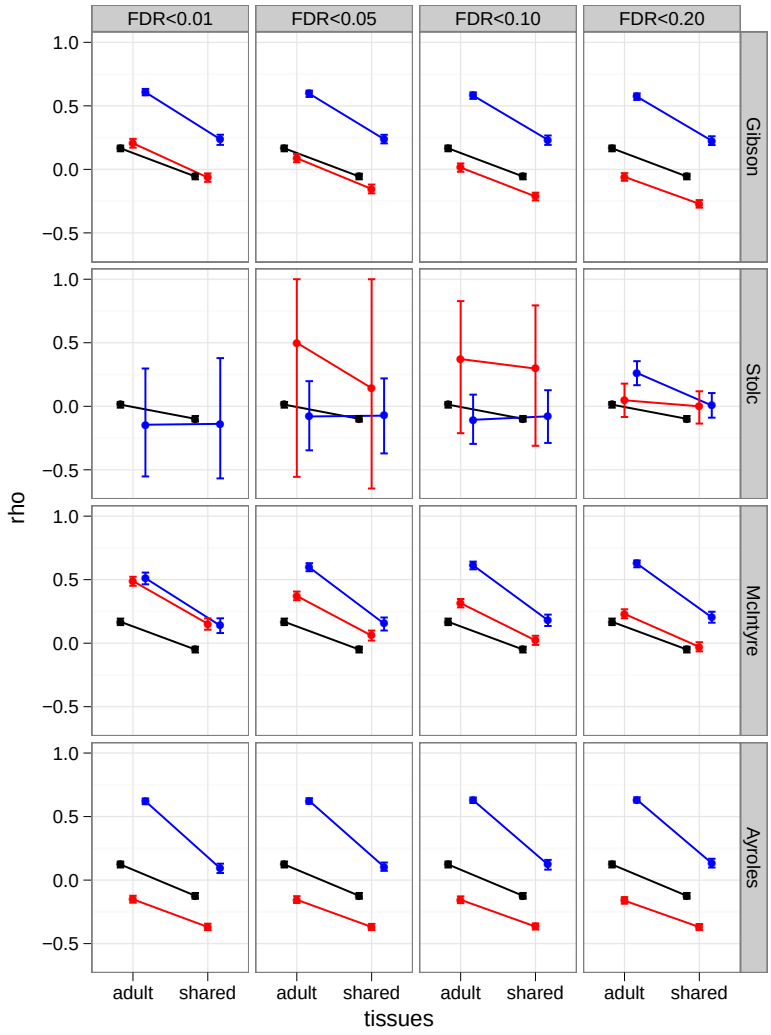

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Parisi2003(wholefly)
Parisi2003(gonadectomized)
Parisi2003(gonad)
tau
1.0
0.8
0.6
0.4
0.2
***
*


FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Ranz2003
Gibson2004
Stolc2004
McIntyre2006
Ayroles2009
tau
1.0
0.8
0.6
0.4
0.2
***
**
*

Correlation between sex-biased expression and tissue-specificity

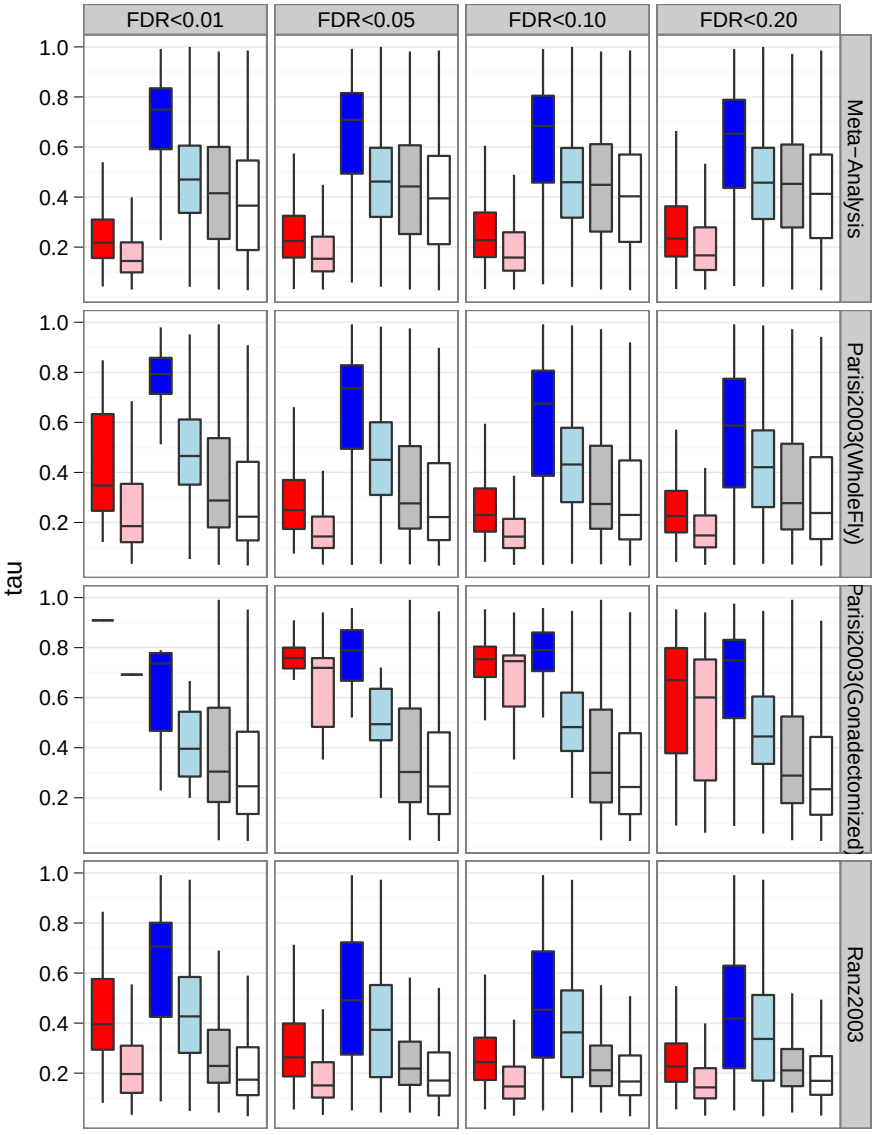
Plots show the point estimate of the correlation coefficient (ρ) between $|\log(M/F)|$ and tissue-specificity (τ) and the 95% CI of the estimate. The data set for which M/F was estimated and the FDR cut-off used to assign genes into sex-bias classes are given. Estimates of τ were calculated using all adult tissues (adult) or shared somatic tissues (shared). Correlations were calculated using all genes (black), female-biased genes (red), and male-biased genes (blue).

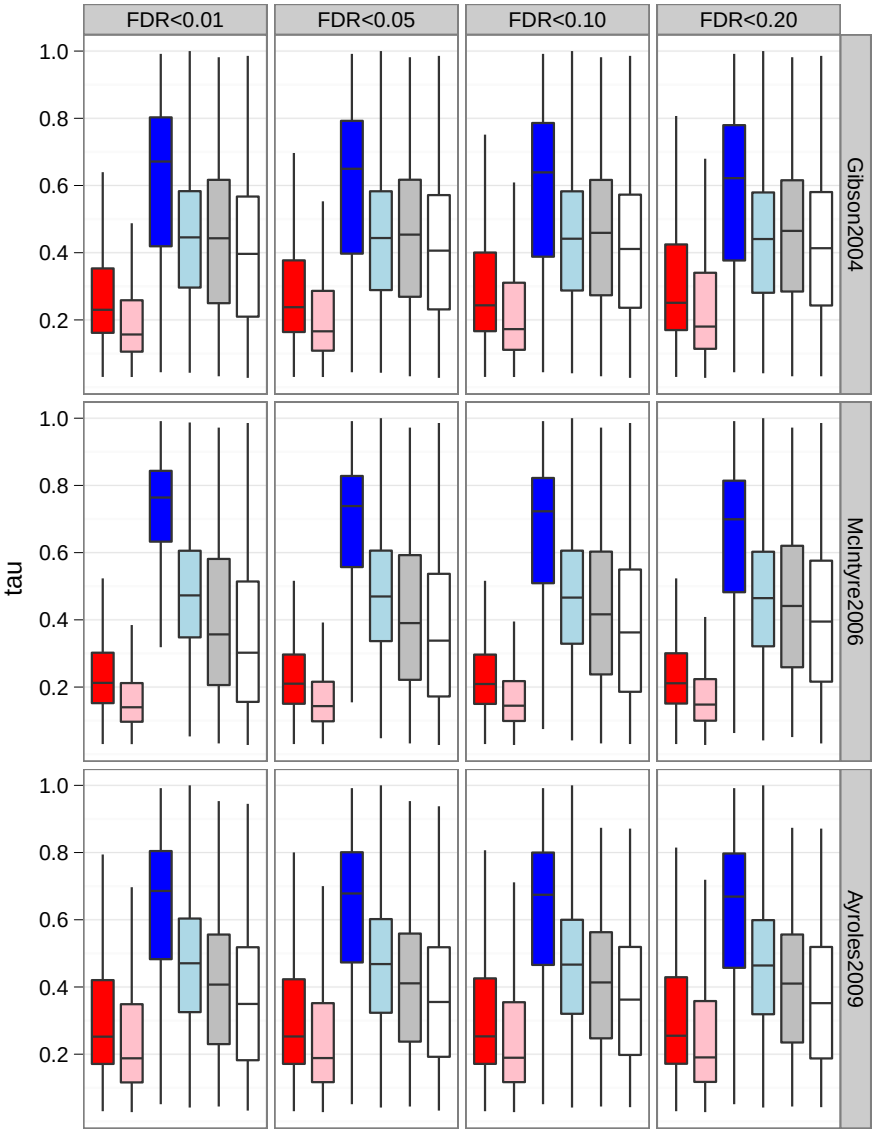


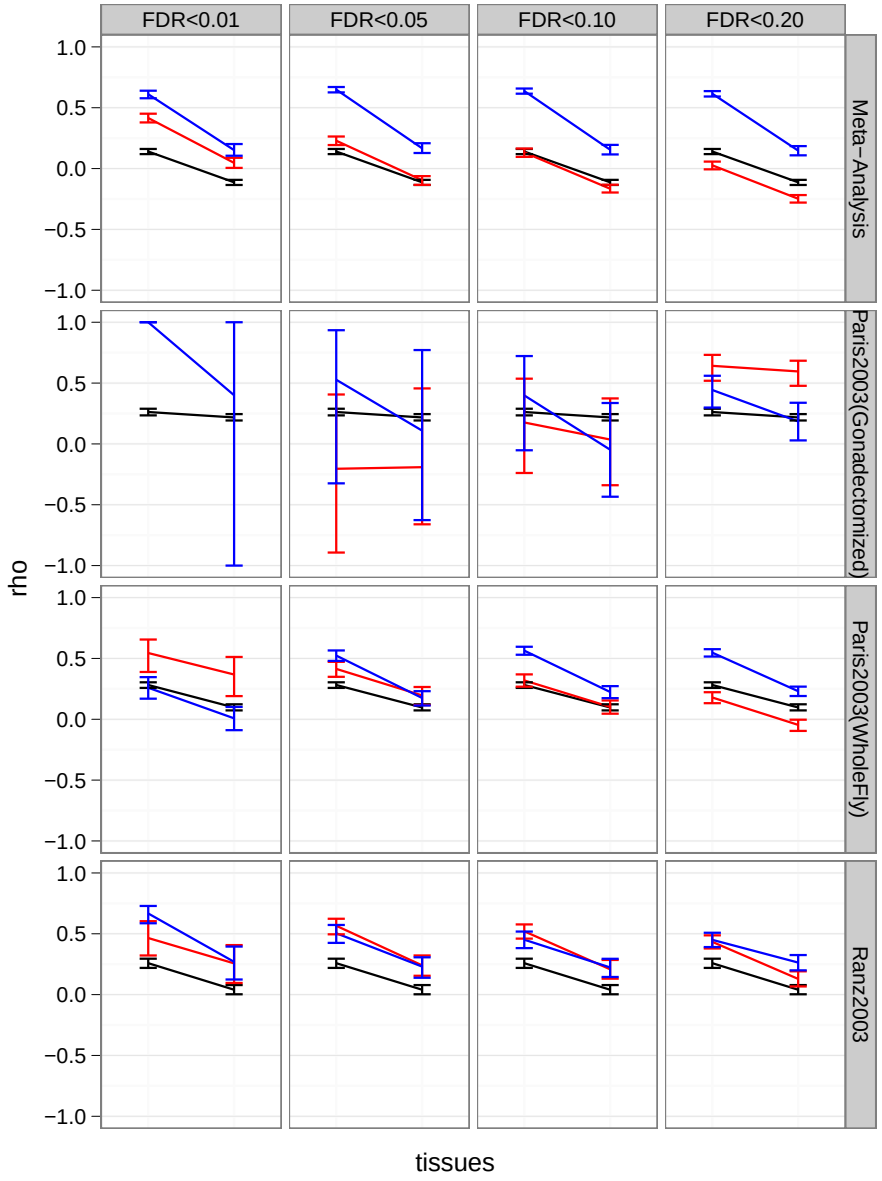

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Gibson
Stolc
McIntyre
Ayroles
rho
1.0
0.5
0.0
−0.5
adult
shared
tissues

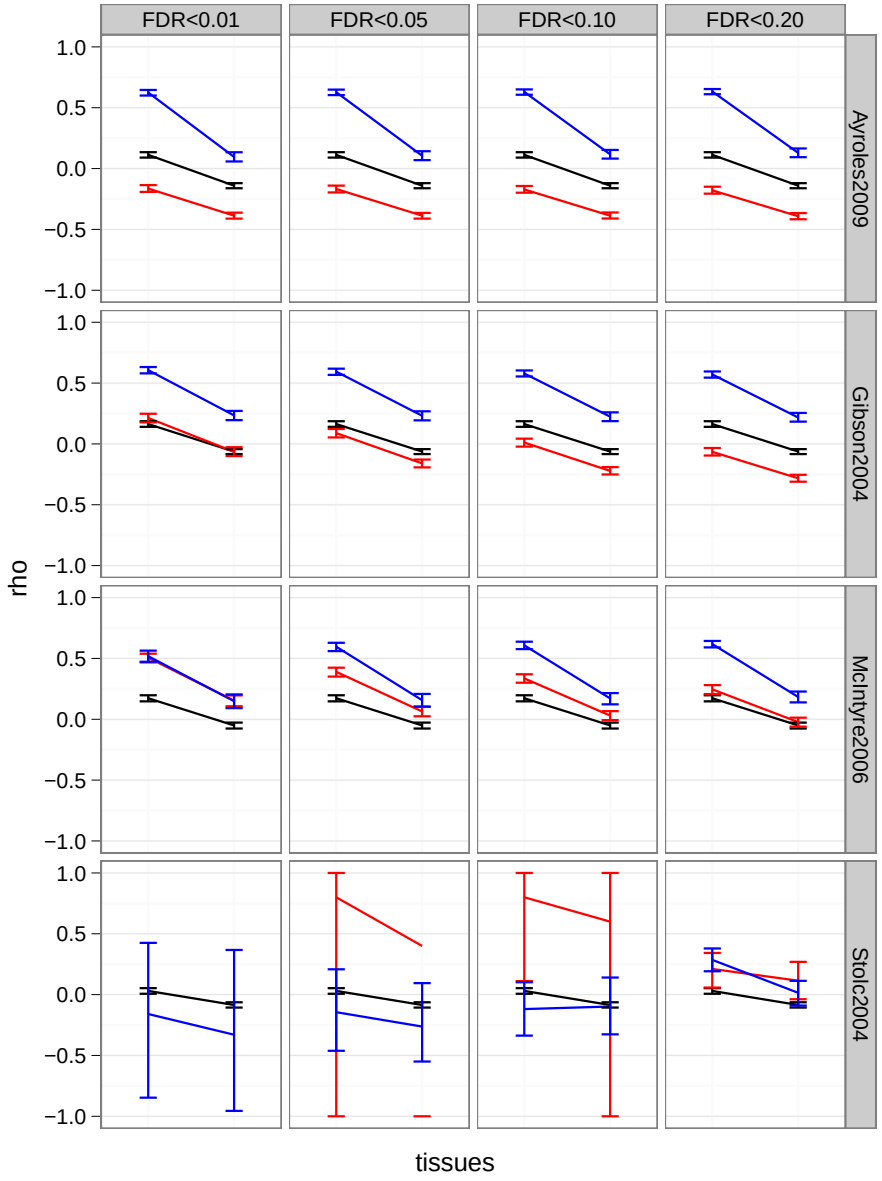
Excluding highly expressed genes

Broadly expressed genes are known to be more highly expressed (Larracuente et al., 2008), and the signal intensity of highly expressed genes may reach the point of saturation in microarray data (Hsiao et al., 2002). Therefore, there may be less power to detect sex-biased expression in broadly expressed genes because the signal may be saturated in the more highly expressed sex. To test whether signal saturation can explain the tissue-specificity of male-biased genes, I repeated the analyses above after excluding genes expressed above a signal threshold (signal intensity > 1000). I observe the same patterns when these genes are excluded, indicating that the tissue-specificity of male-biased genes is not a byproduct of signal saturation.


FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Parisi2003(WholeFly)
Parisi2003(Gonadectomized)
Ranz2003
tau
1.0
0.8
0.6
0.4
0.2

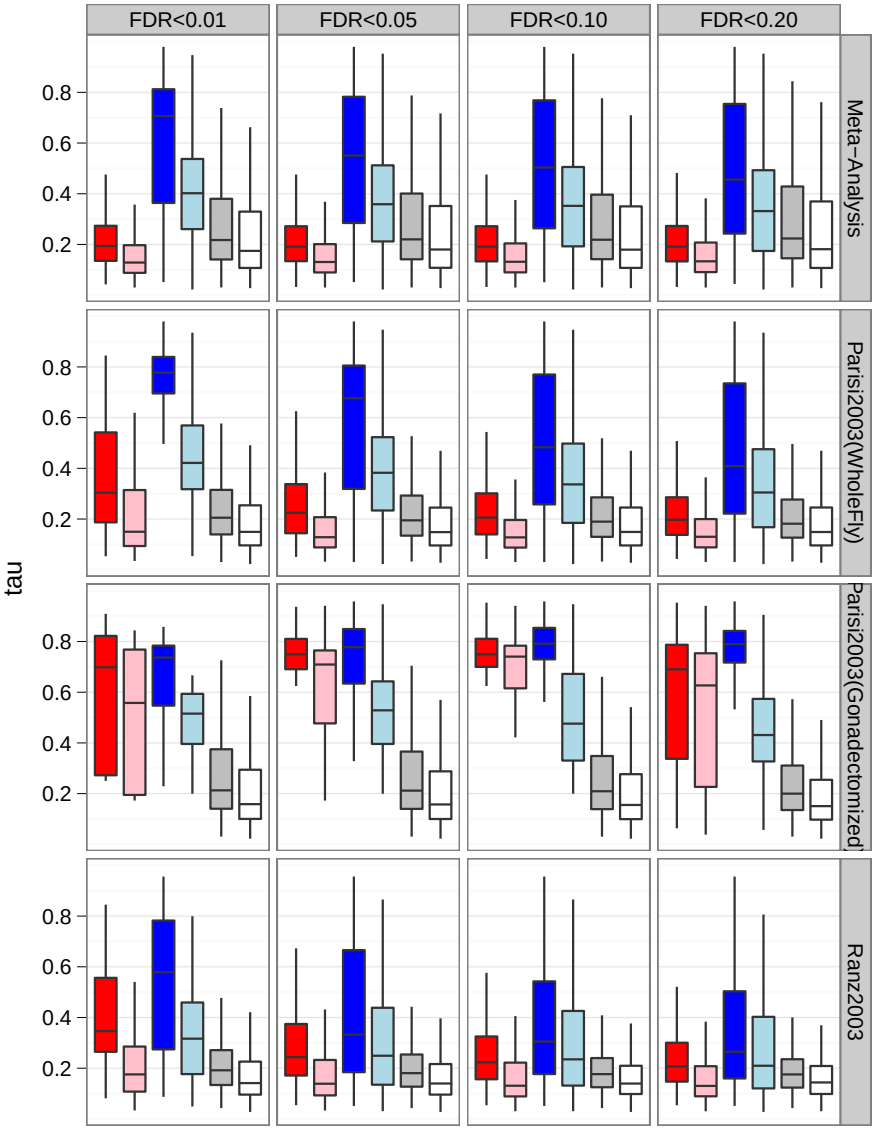

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Gibson2004
McIntyre2006
Ayroles2009
tau
1.0
0.8
0.6
0.4
0.2

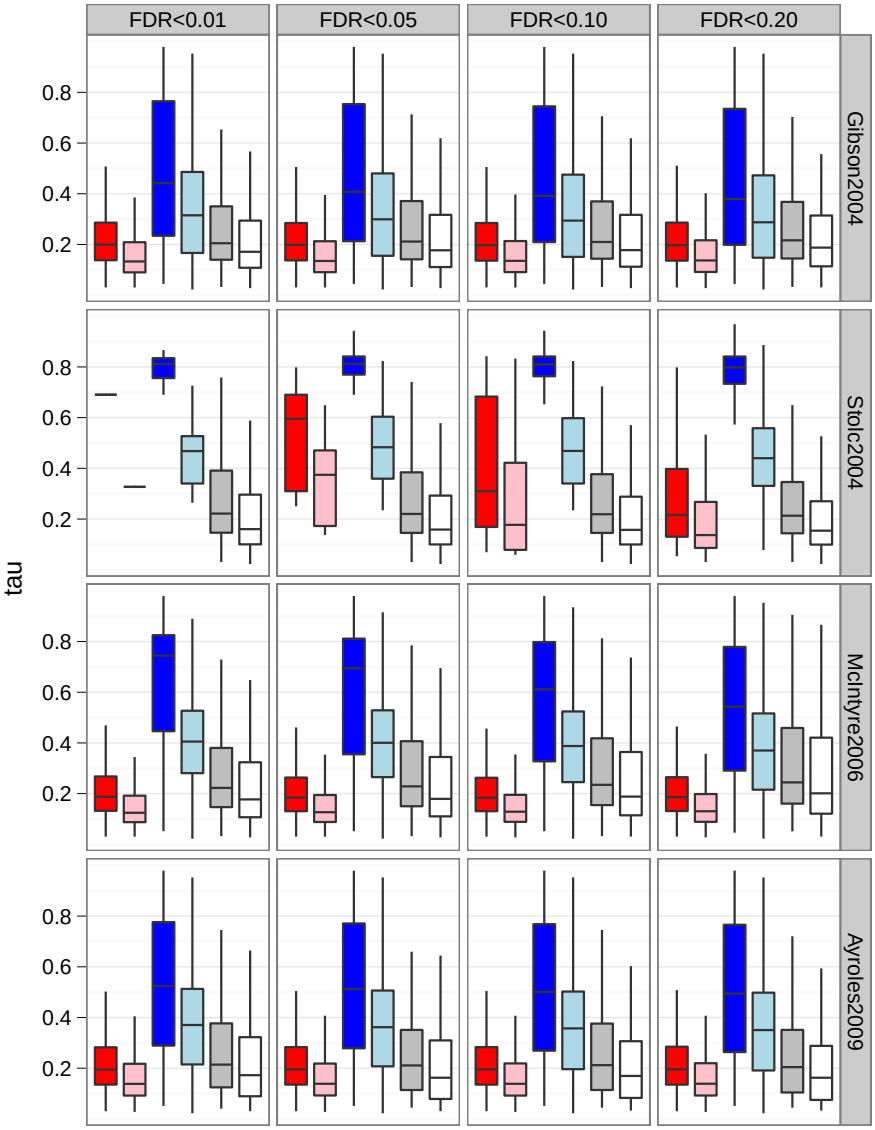

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Paris2003(Gonadectomized)
Paris2003(WholeFly)
Ranz2003
rho
1.0
0.5
0.0
−0.5
−1.0
tissues

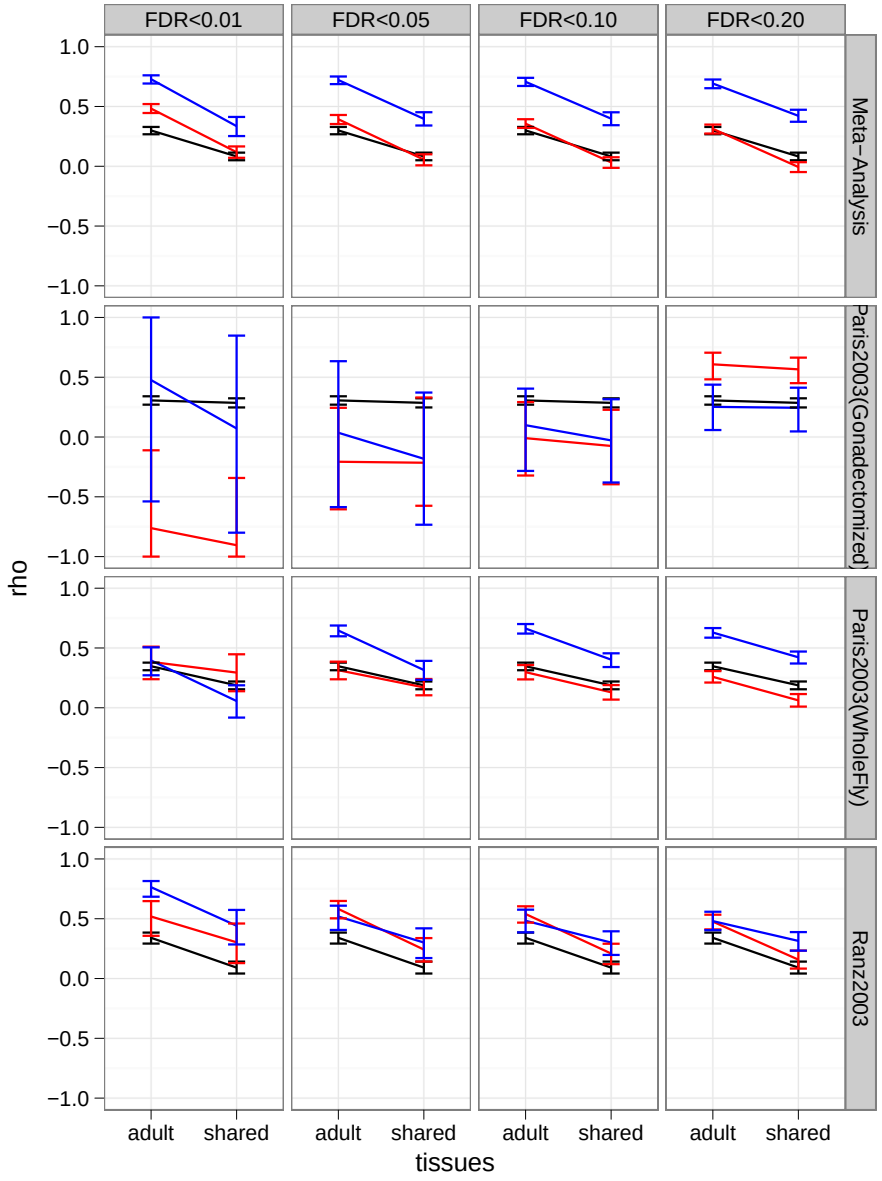

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Ayroles2009
Gibson2004
McIntyre2006
Stolc2004
rho
1.0
0.5
0.0
−0.5
−1.0
tissues

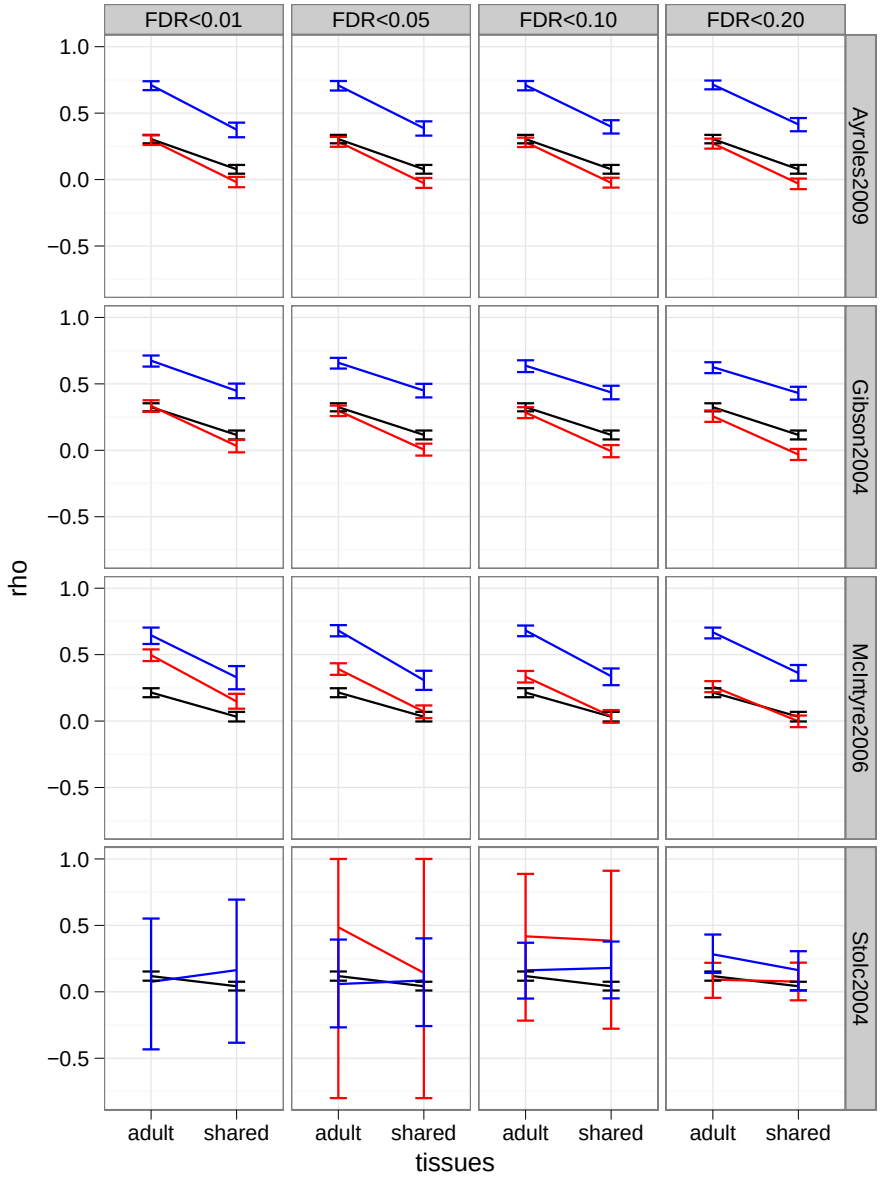
Excluding lowly expressed genes

This analysis may be affected by lowly expressed, tissue-specific genes. Many genes in the *D. melanogaster* genome are lowly expressed in whole fly because they are tissue-specific (Chintapalli, Wang and Dow, 2007). Slight changes in expression in one sex or experimental errors may cause these lowly expressed genes to be called as sex-biased when whole flies are compared. To determine whether these lowly expressed genes drive the correlation between $|\log(M/F)|$ and τ, I estimated ρ using only genes with moderate expression when measured in whole fly (signal intensity > 100). The results are similar to those obtained when all genes are included. I thus conclude that any correlations between $|\log(M/F)|$ and τ are not driven by the sex-biased expression of lowly expressed genes.


FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Parisi2003(WholeFly)
Parisi2003(Gonadectomized)
Ranz2003
tau
0.2
0.4
0.6
0.8

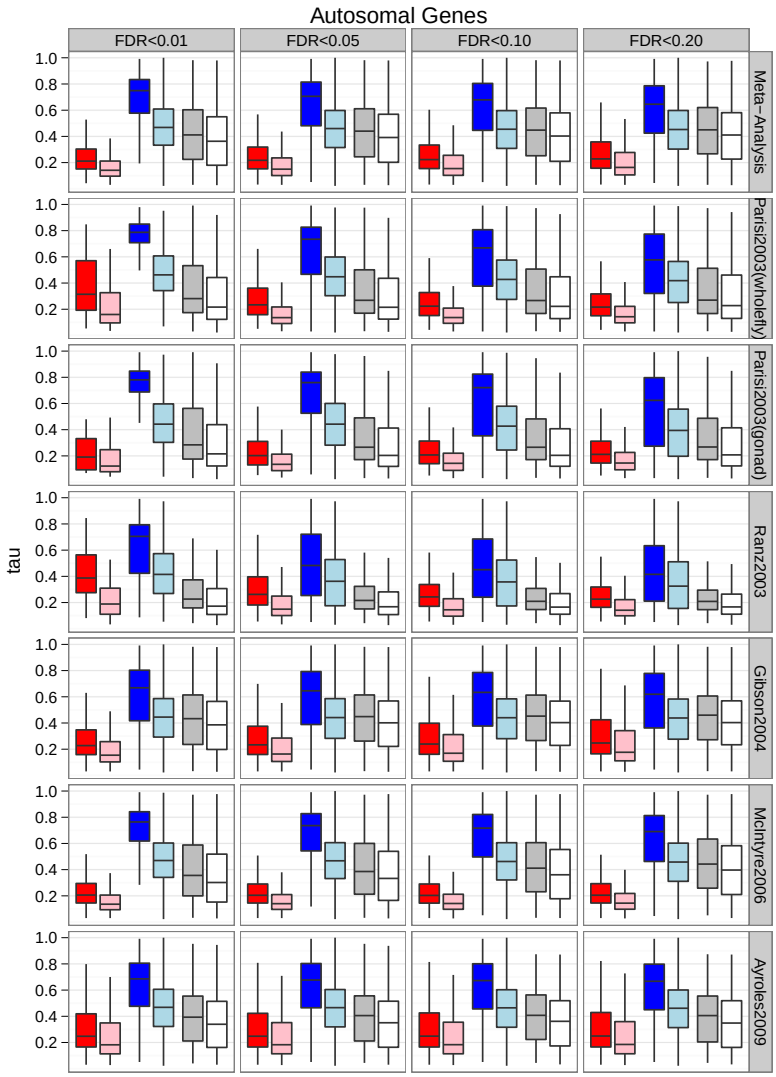

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Gibson2004
Stolc2004
McIntyre2006
Ayroles2009
tau
0.2
0.4
0.6
0.8

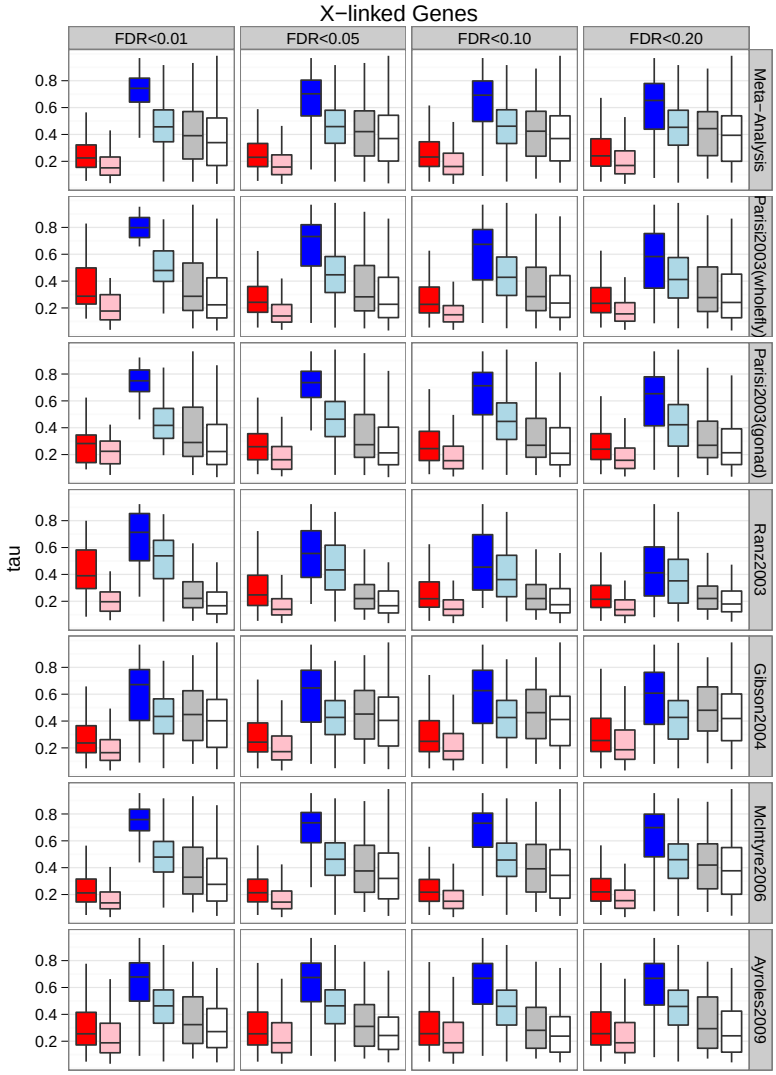

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Paris2003(Gonadectomized)
Paris2003(WholeFly)
Ranz2003
rho
1.0
0.5
0.0
−0.5
−1.0
adult
shared
tissues


FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Ayroles2009
Gibson2004
McIntyre2006
Stolc2004
rho
1.0
0.5
0.0
−0.5
adult
shared
tissues

There is no effect of the X chromosome on the relationship between sex-bias and tissue-specificity

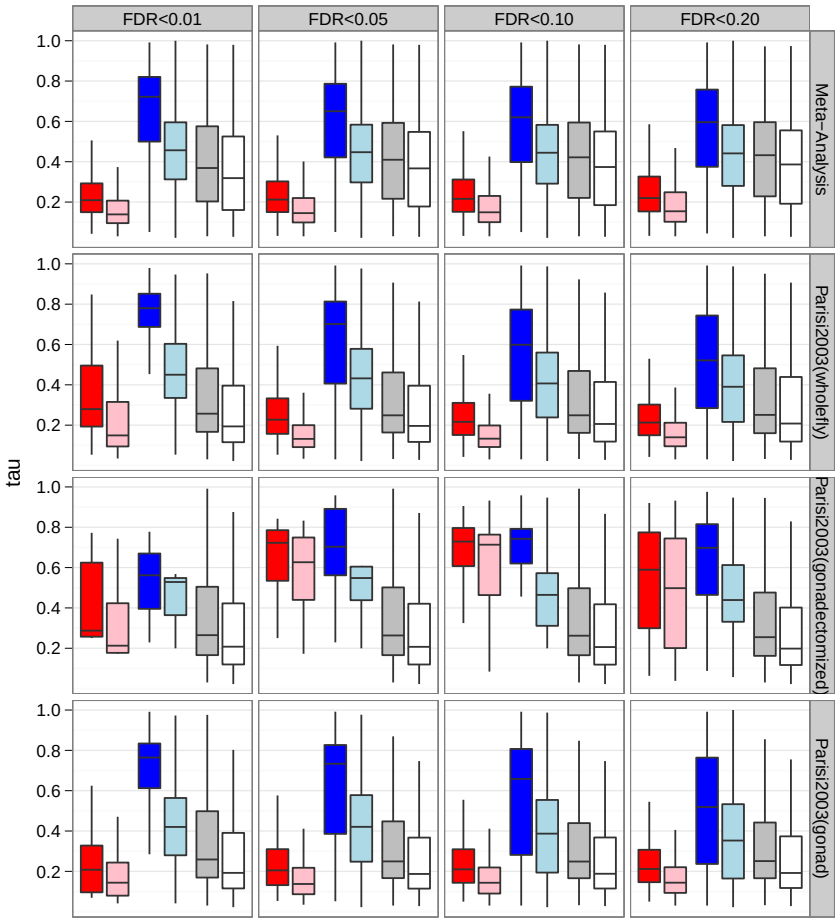
The *D. melanogaster* X chromosome contains a paucity of female-biased genes and an excess of male-biased genes (Parisi et al., 2003; Ranz et al., 2003), and the relationship between sex-bias and tissue-specificity may differ between the X chromosome and the autosomes as a result of differences in sex-biased gene content. To test this possibility, I analyzed X-linked and autosomal genes separately. One observes the same patterns as above for X-linked and autosomal genes, indicating that the tissue-specificity of male-biased genes is true throughout the genome.

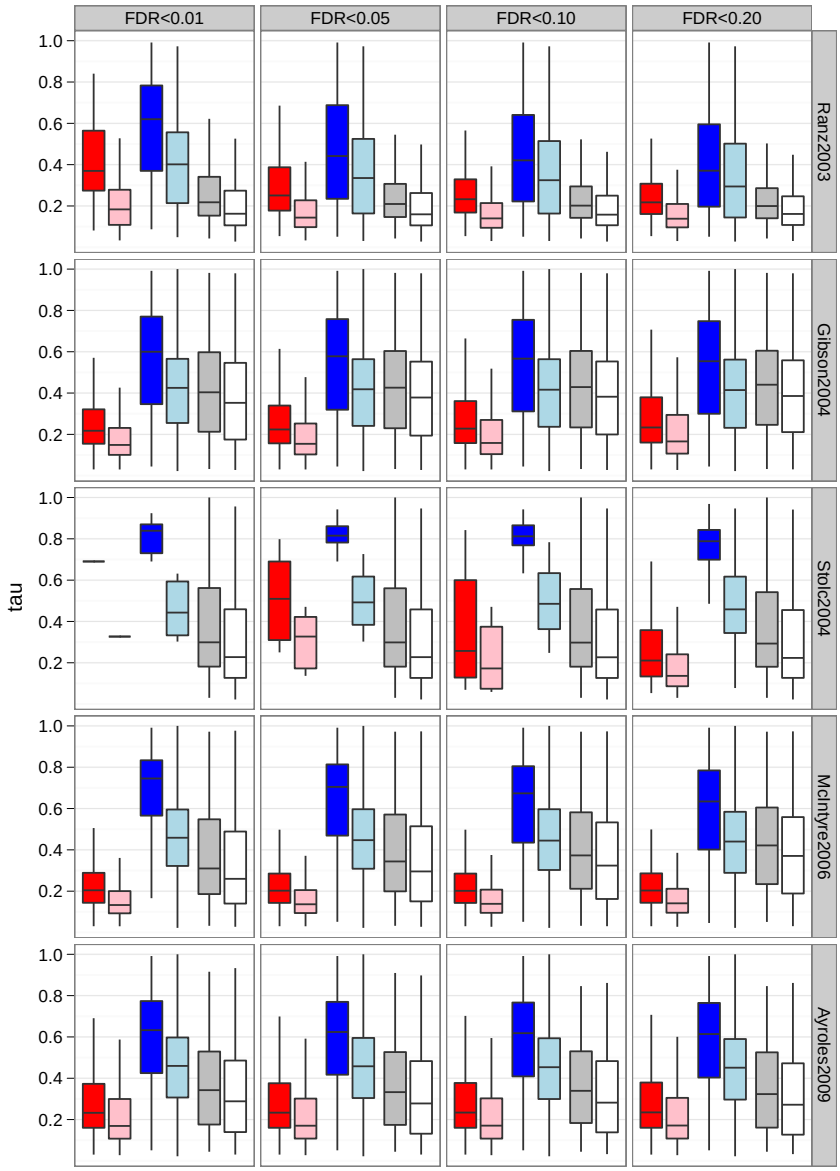

Autosomal Genes
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Paris2003(wholefly)
Paris2003(gonad)
Ranz2003
Gibson2004
McIntyre2006
Ayroles2009
tau
1.0
0.8
0.6
0.4
0.2


X−linked Genes
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta−Analysis
Parisi2003(wholefly)
Parisi2003(gonad)
Ranz2003
Gibson2004
McIntyre2006
Ayroles2009
tau
0.2
0.4
0.6
0.8

There is no effect of duplicated genes on the relationship between sex-bias and tissue-specificity

The derived copies of duplicated genes tend to have tissue-specific and male-biased expression (Meisel, Han and Hahn, 2009), and the inclusion of these genes may be partially responsible for the tissue-specific expression of male-biased genes. I repeated my analysis by excluding genes with a closely related paralog in the *D. melanogaster* genome (using fuzzy reciprocal blast gene families (Drosophila 12 Genomes Consortium, 2007)), and I obtained the same results as when all genes were included. Therefore, the tissue-specificity of male-biased genes is not a byproduct of tissue-specific male-biased duplicated genes.

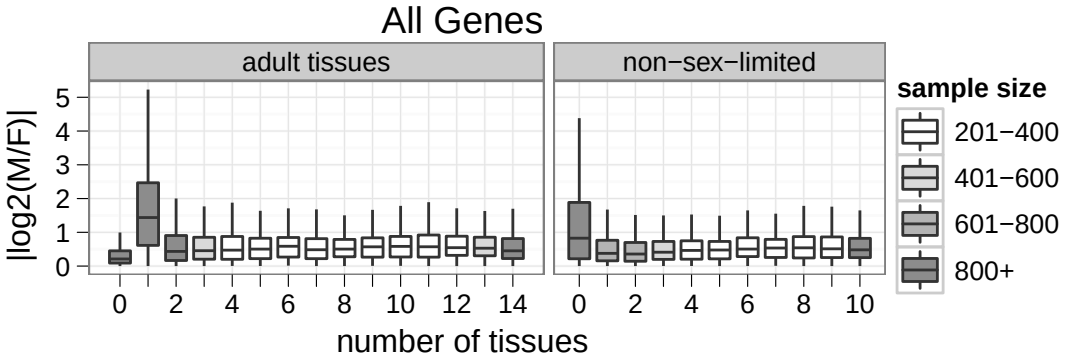

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Meta-Analysis
Parisi2003(wholefly)
Parisi2003(gonadectomized)
Parisi2003(gonad)
tau
1.0
0.8
0.6
0.4
0.2

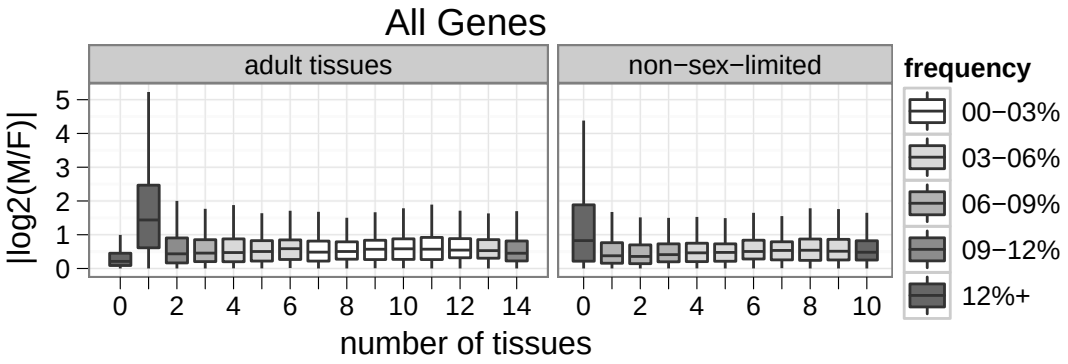

FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
Ranz2003
Gibson2004
Stolc2004
McIntyre2006
Ayroles2009
tau
1.0
0.8
0.6
0.4
0.2

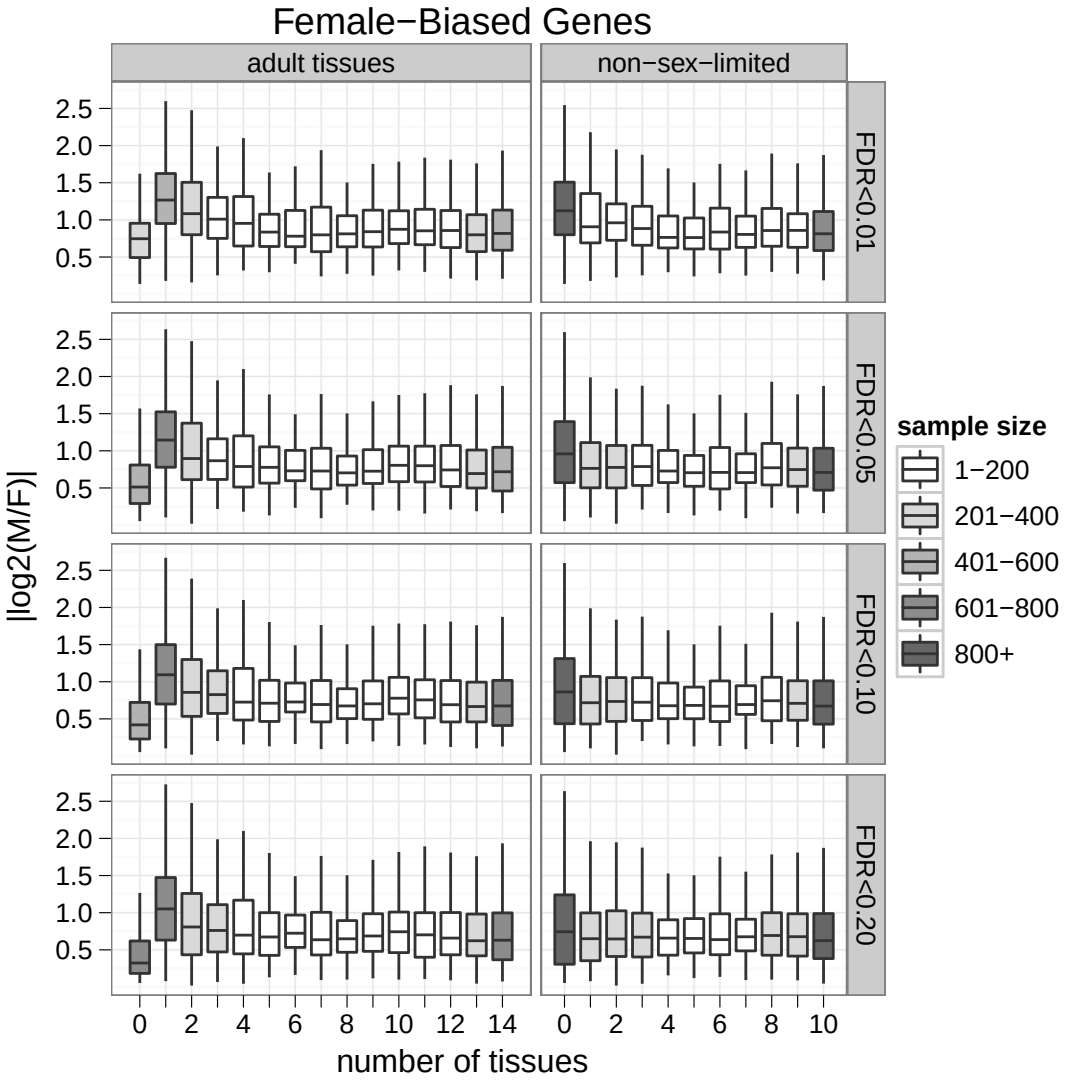
Measuring tissue-specificity as the number of tissues in which a gene is expressed

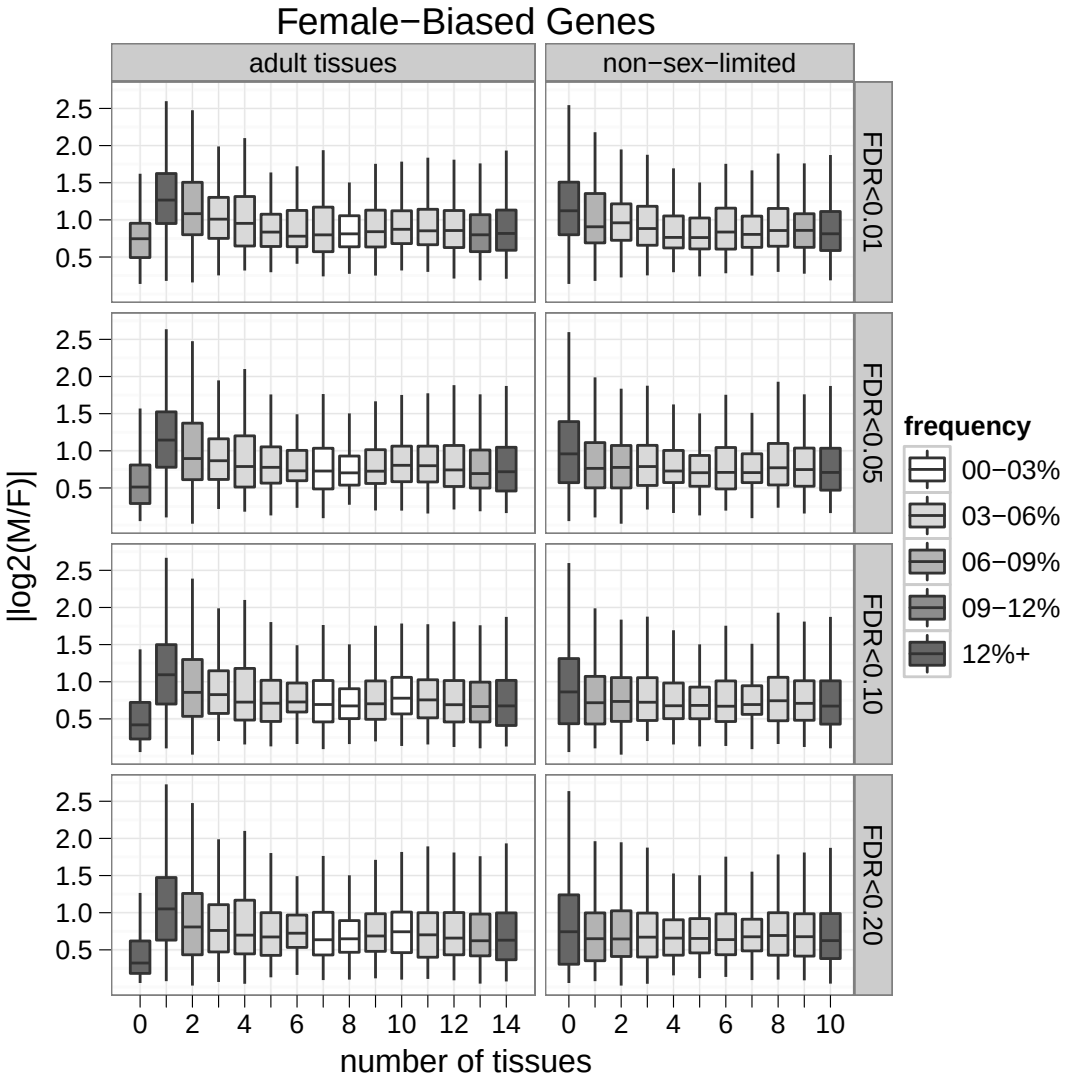
As an alternative to τ, I measured tissue-specificity by counting the number of tissues in which a gene is expressed above a certain threshold (signal intensity > 100). When examining the expression of all genes in all adult tissues, a large fraction of genes are expressed in a single tissue, and these single-tissue genes tend to be more sex-biased than other genes. If I remove sex-limited tissues from the estimation of expression breadth, a large fraction of genes are extremely sex-biased and expressed in no shared somatic tissues. Therefore, there is a large group of genes with sex-biased expression that is driven by expression in sex-limited tissues. Male-biased genes have the same pattern of a large group of highly sex-biased genes expressed in a single male-specific tissue. However, there is not as large of a group of as highly female-biased genes expressed in a single tissue. This provides further evidence that male-biased genes have tissue-specific expression that is driven by sex-limited tissues, while female-biased genes are more broadly expressed.

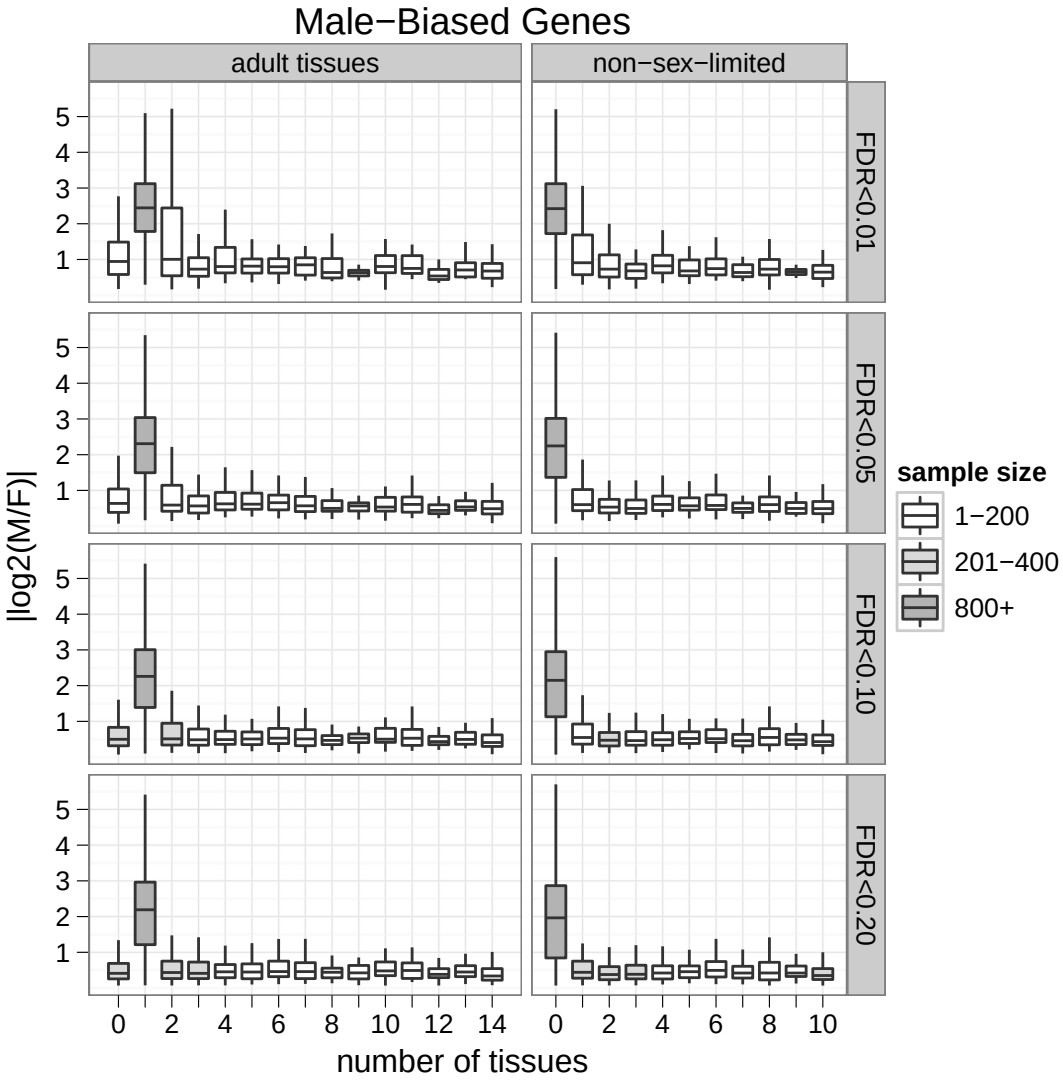
Boxplots show the distribution of $|\log_2(M/F)|$ for genes expressed in a given number of tissues. Counts of tissues are either for all adult tissues (left) or non-sex-limited tissues (right). Data are either taken from all genes (first two rows) or sex-biased genes (subsequent graphs). In the case of sex-biased genes, the FDR cut-off used is shown for each row of plots. The color of each box represents either the number of genes expressed in a given number of tissues (shown first) or the percent of genes expressed in a given number of tissues (shown second). Darker colors indicate larger sample sizes.

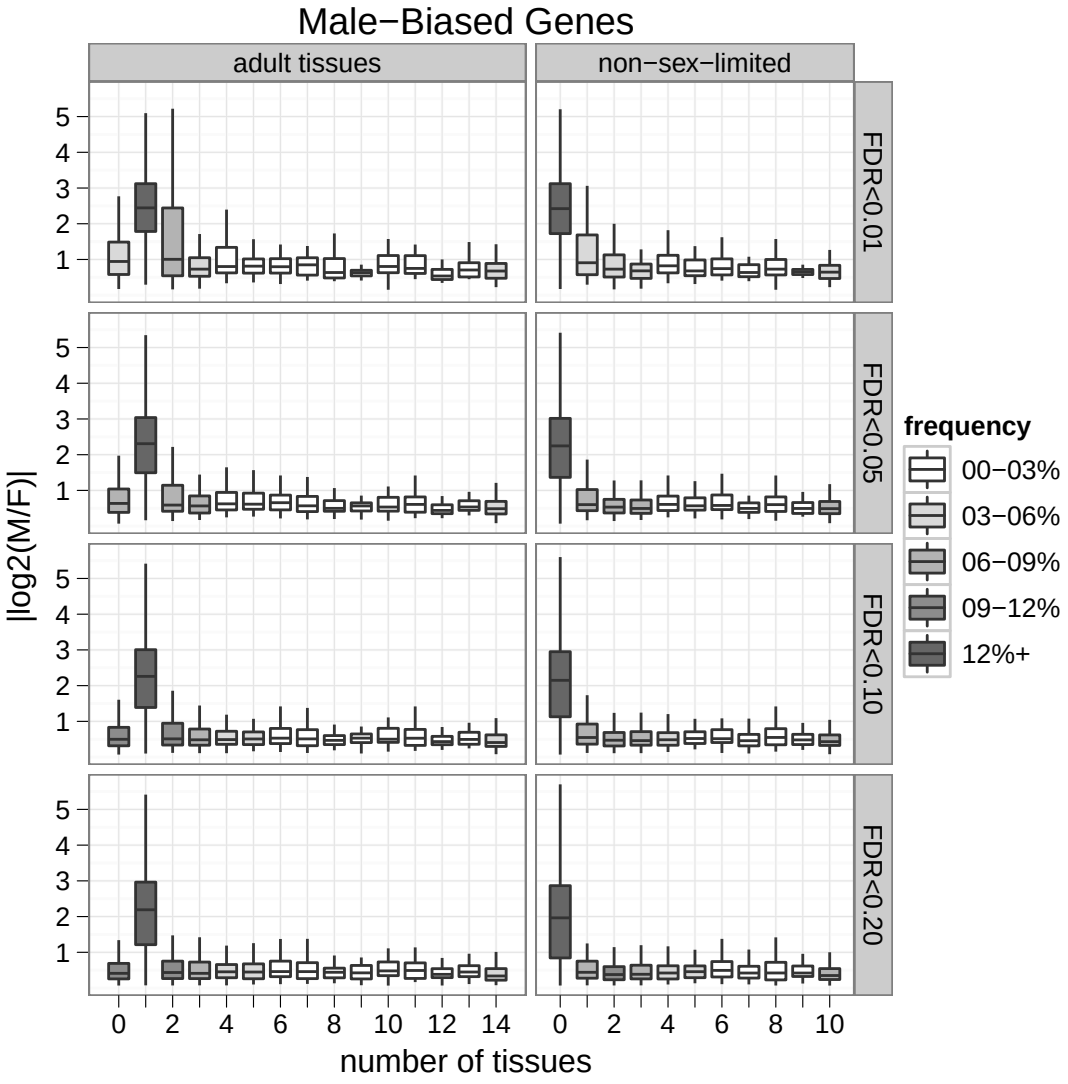

All Genes
adult tissues
non−sex−limited
|log2(M/F)|
number of tissues
sample size
201−400
401−600
601−800
800+


All Genes
adult tissues
non−sex−limited
|log2(M/F)|
number of tissues
frequency
00−03%
03−06%
06−09%
09−12%
12%+


Female−Biased Genes
adult tissues
non−sex−limited
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
|log2(M/F)|
2.5
2.0
1.5
1.0
0.5
0
2
4
6
8
10
12
14
number of tissues
sample size
1−200
201−400
401−600
601−800
800+


Female−Biased Genes
adult tissues
non−sex−limited
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
|log2(M/F)|
2.5
2.0
1.5
1.0
0.5
0
2
4
6
8
10
12
14
number of tissues
frequency
00−03%
03−06%
06−09%
09−12%
12%+

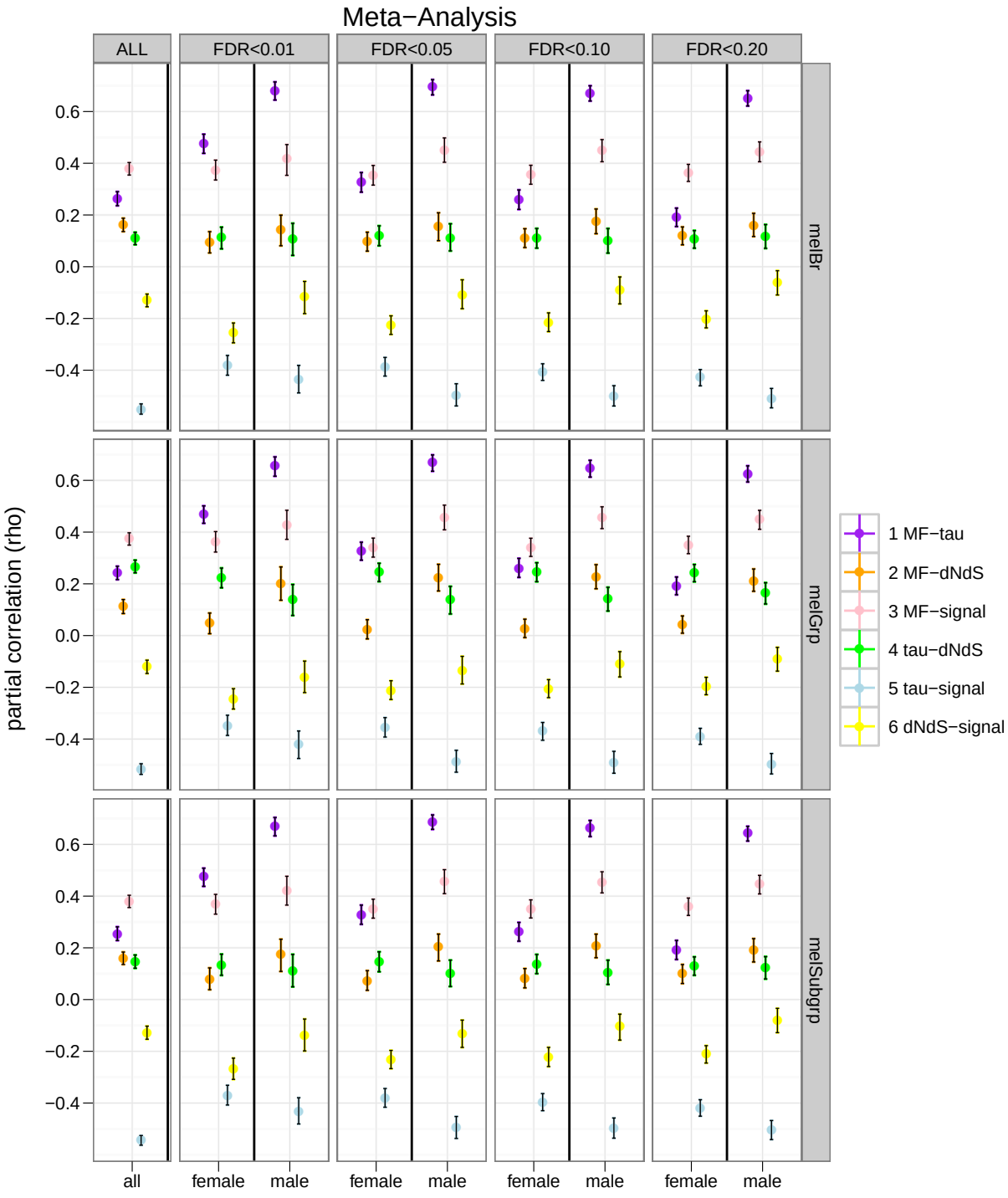

Male−Biased Genes
adult tissues
non−sex−limited
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
|log2(M/F)|
number of tissues
sample size
1−200
201−400
800+

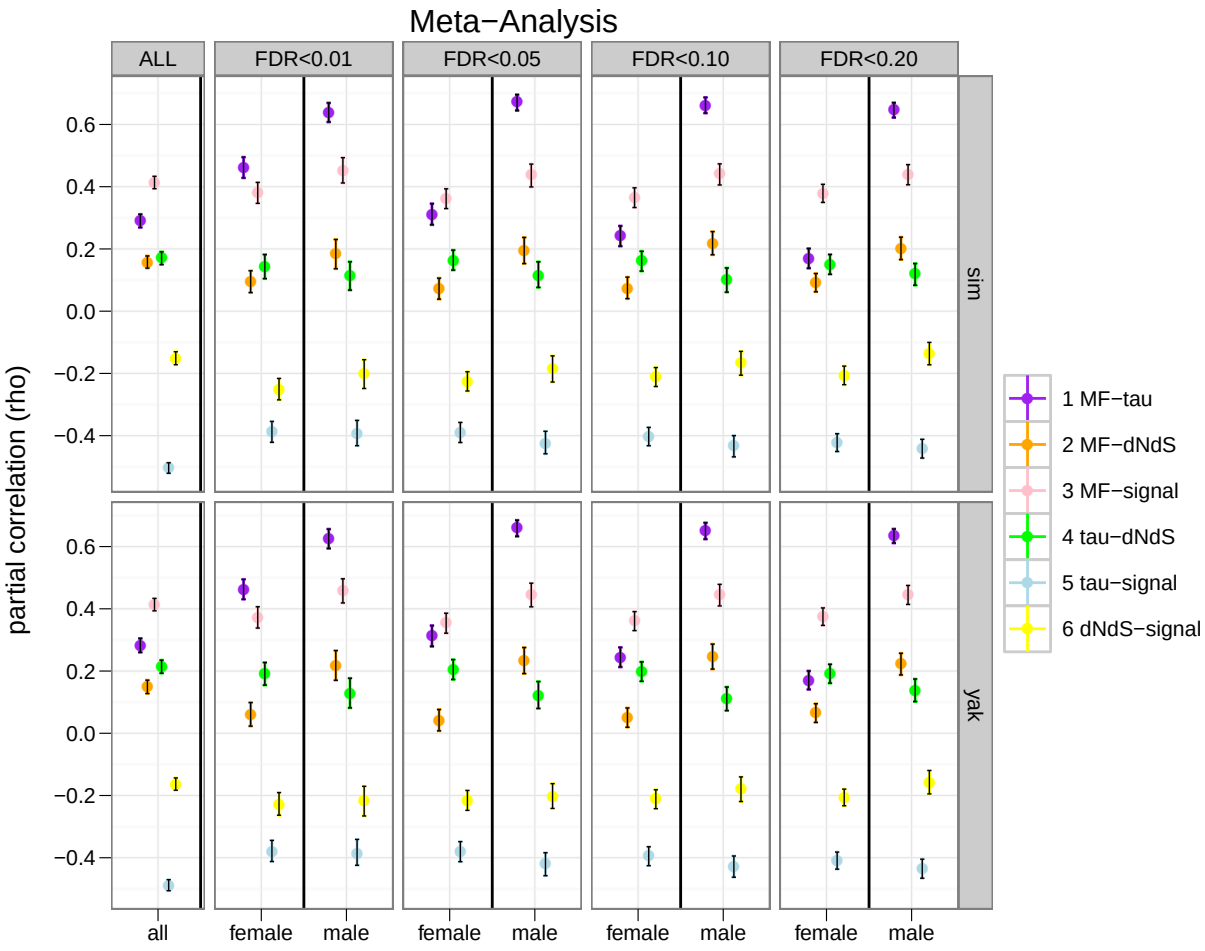

Male−Biased Genes
adult tissues
non−sex−limited
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
|log2(M/F)|
number of tissues
frequency
00−03%
03−06%
06−09%
09−12%
12%+

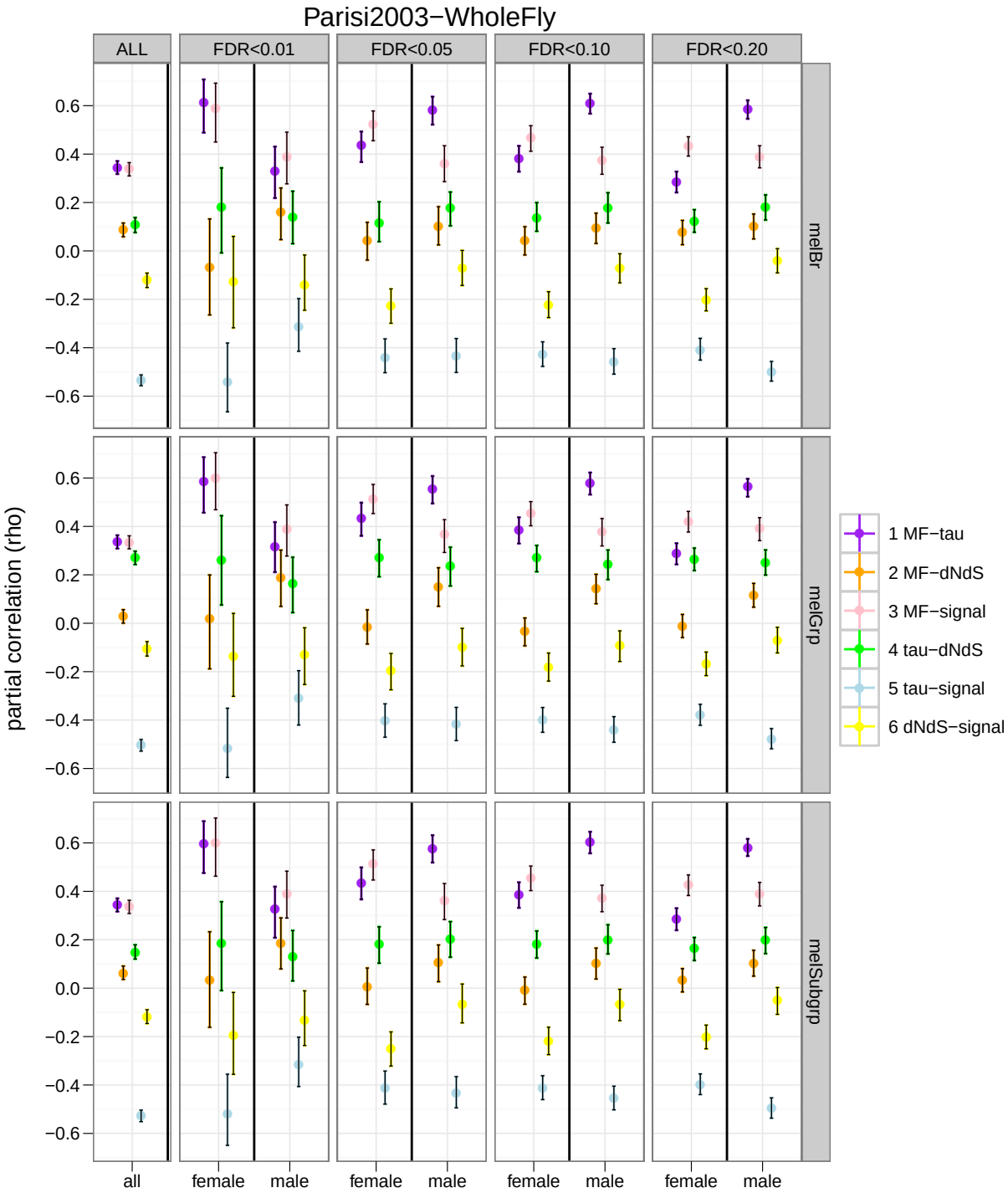
Effect of tissue-specificity on the relationship between sex-biased expression and rates of evolution

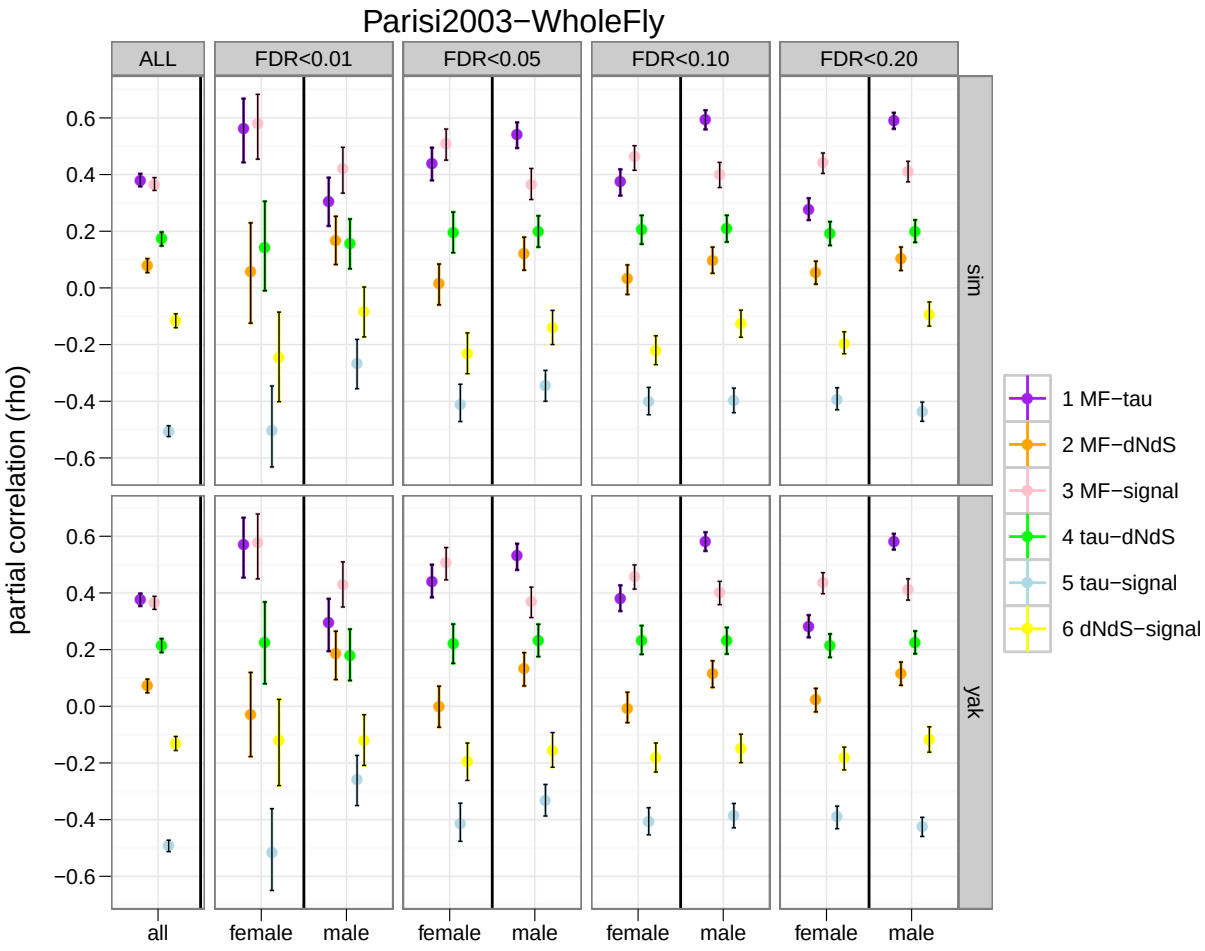
Partial correlations between sex-biased expression, tissue-specificity, rate of evolution, and expression level in *D. melanogaster*

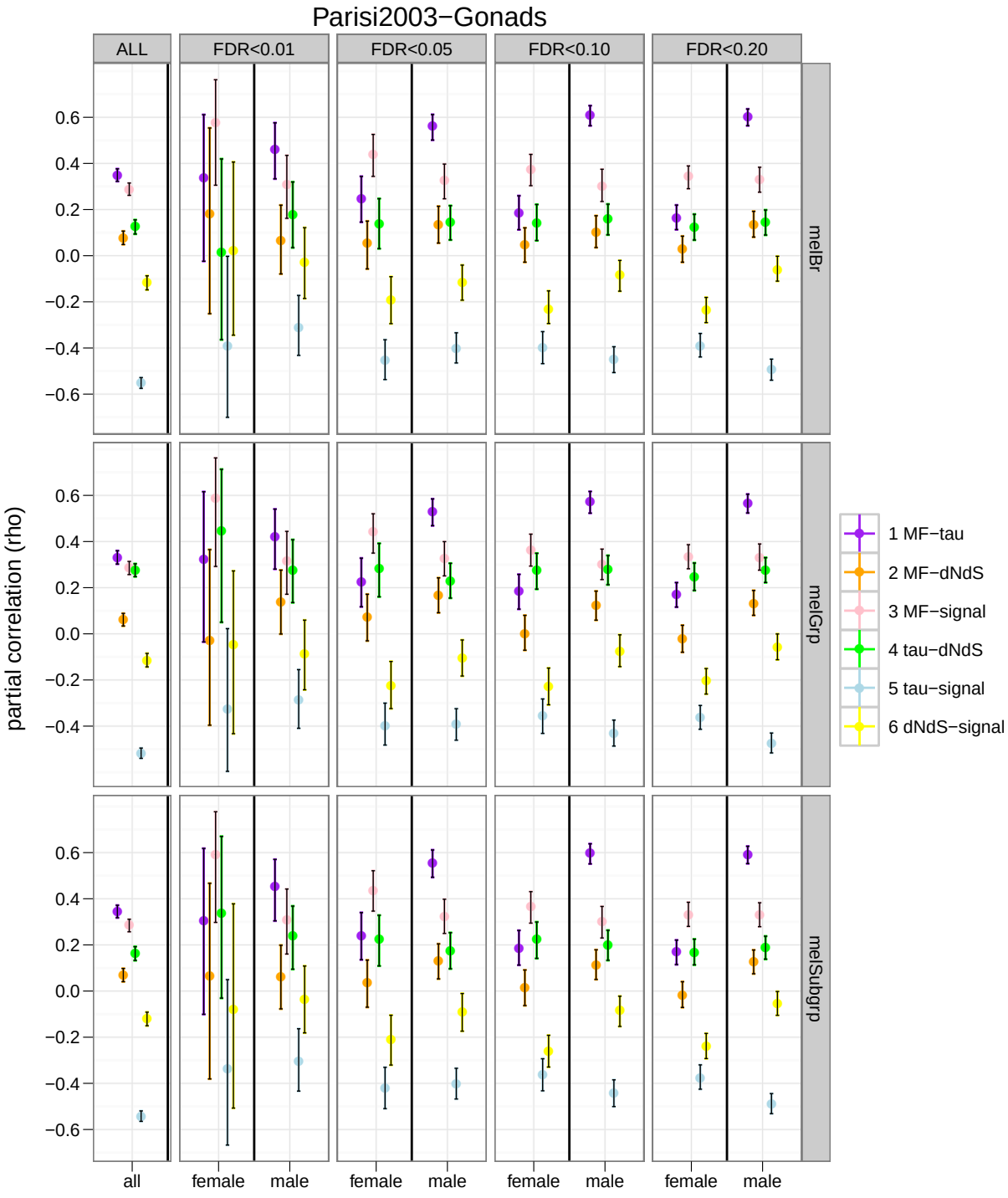
Point estimates and 95% CI of the partial correlation (ρ) between $|\log(M/F)|$ and τ (purple), $|\log(M/F)|$ and d_N/d_S (orange), $|\log(M/F)|$ and expression level (pink), τ and d_N/d_S (green), τ and expression level (light blue), and d_N/d_S and expression level (yellow) are shown for all genes, female-biased genes, and male-biased genes in the *D. melanogaster* genome. Sex-biased genes were called using different data sets and different FDR cut-offs within each data set (shown at the top of each page). Tissue-specificity (τ) was calculated using all adult tissues. Rates of coding sequence evolution (d_N/d_S) were calculated five different ways: along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (melBr); along all branches within the *melanogaster* species group (melGrp); along all the branches within the *melanogaster* species subgroup (melSubgrp); pair-wise between *D. melanogaster* and *D. simulans* (sim); pair-wise between *D. melanogaster* and *D. yakuba* (yak).

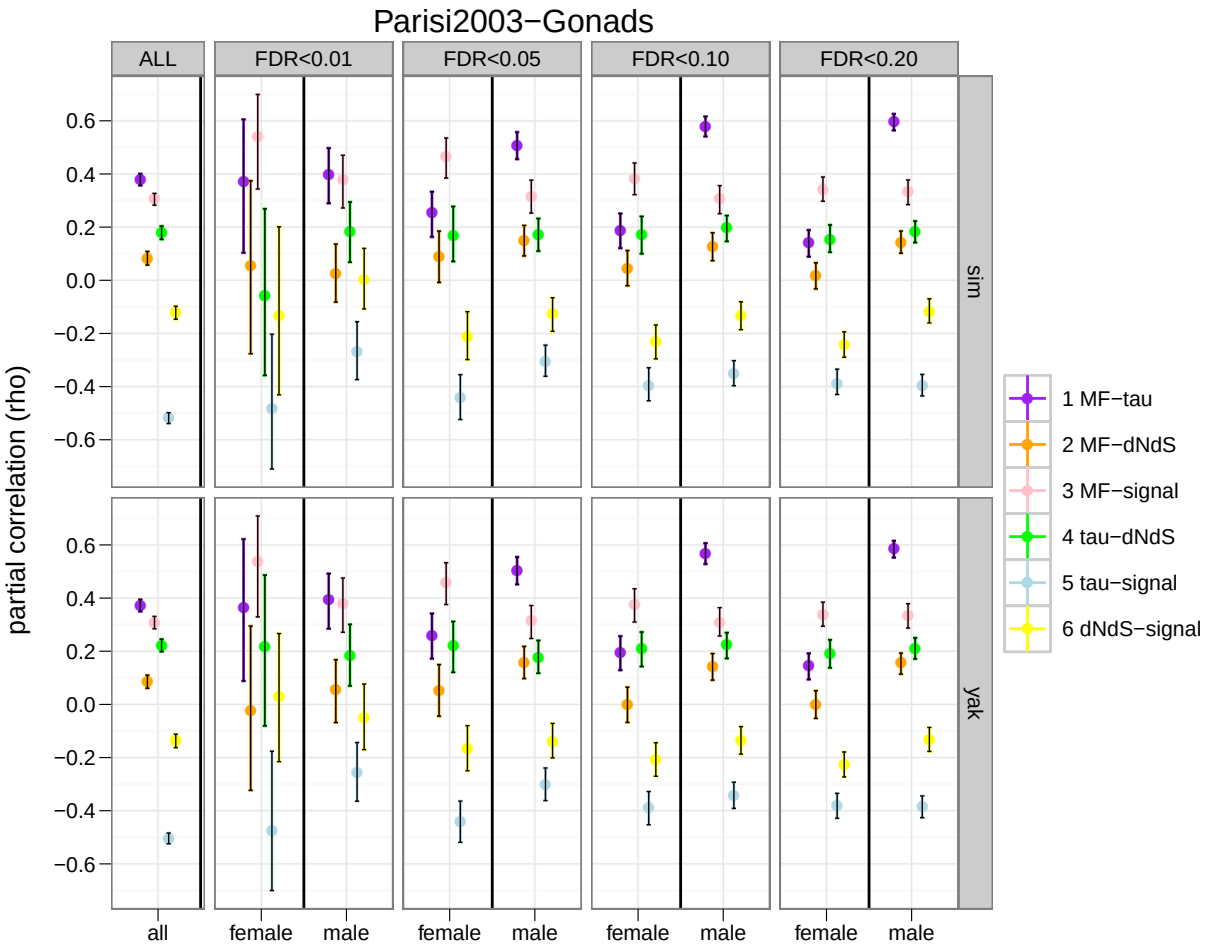

Meta−Analysis
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

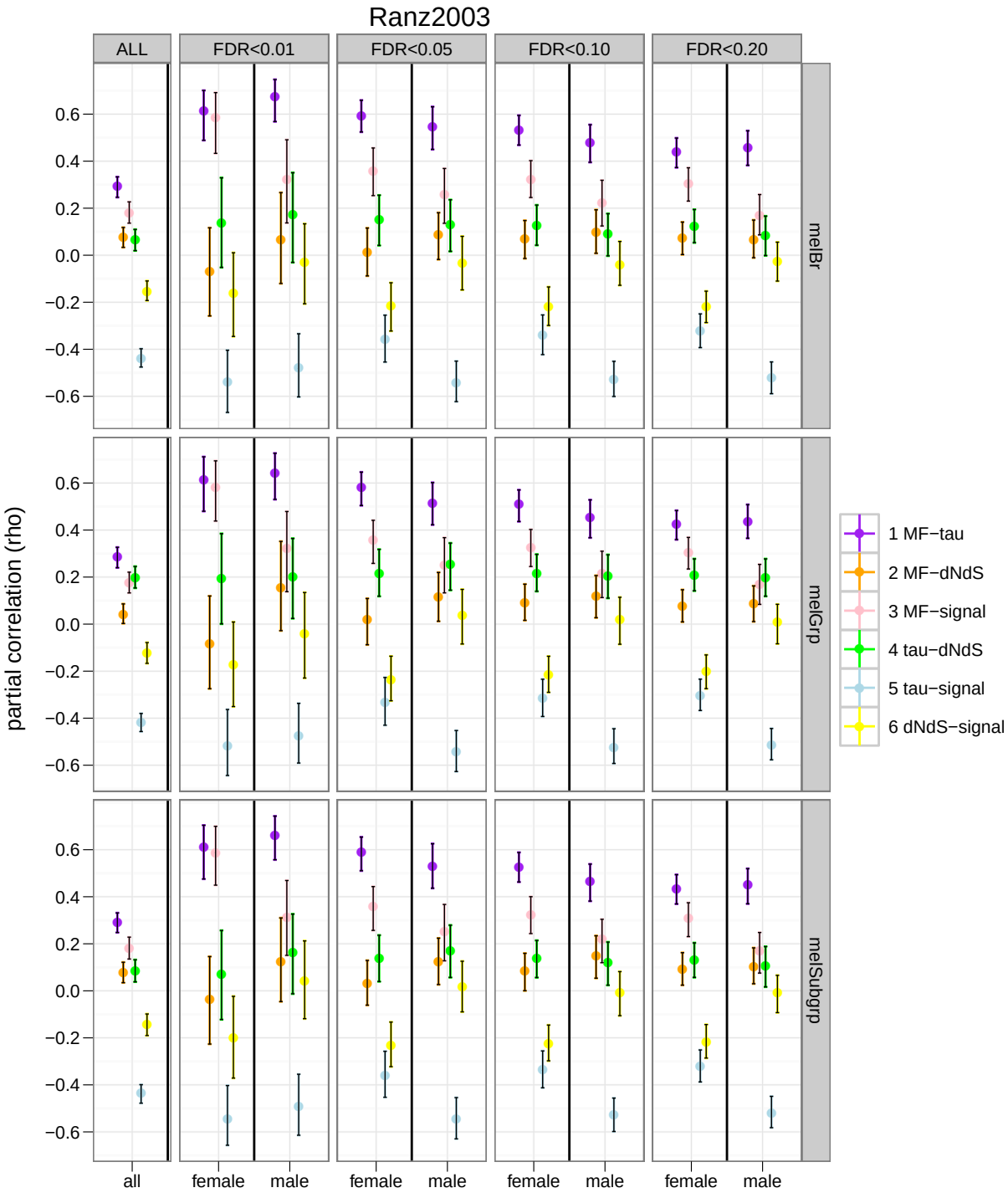

Meta−Analysis
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

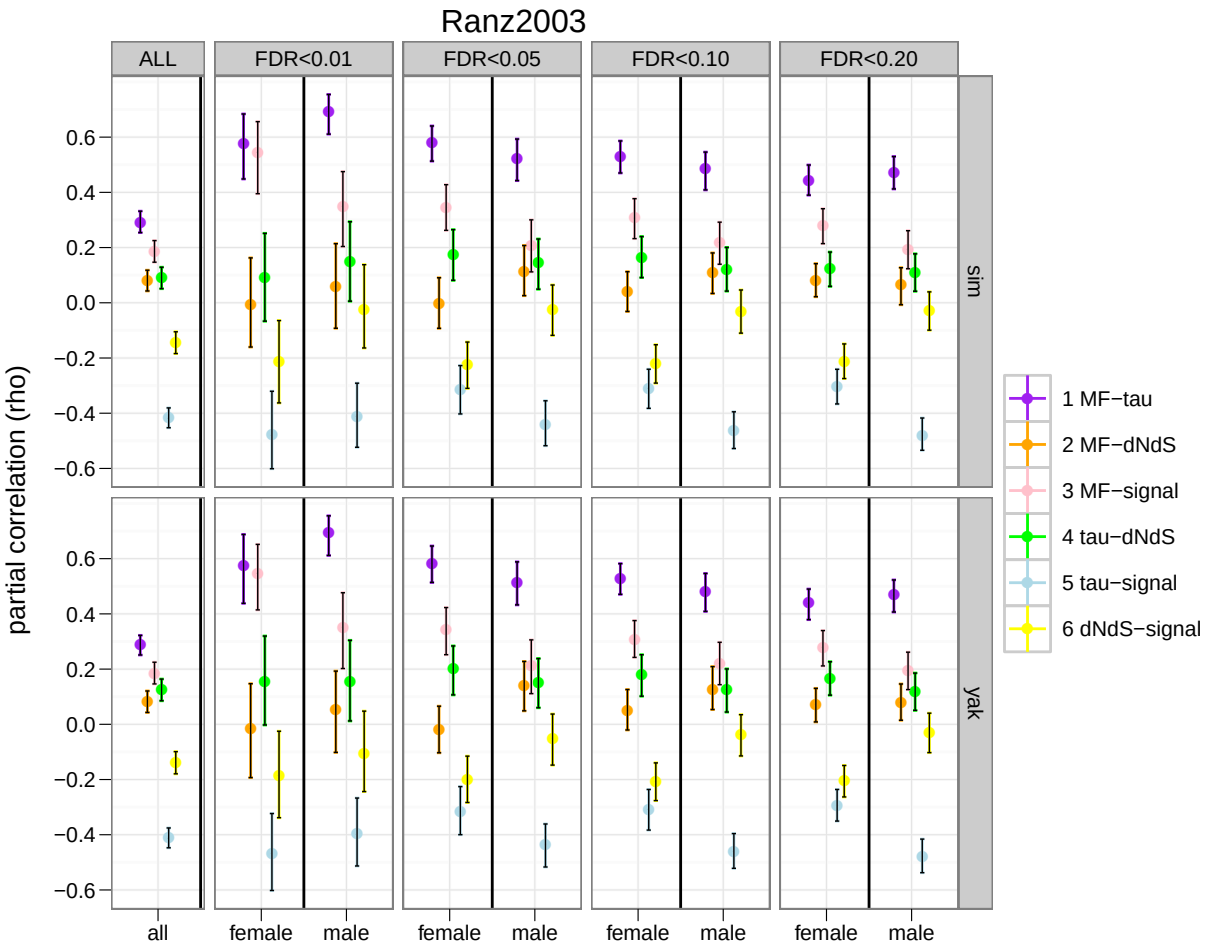

Parisi2003−WholeFly
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

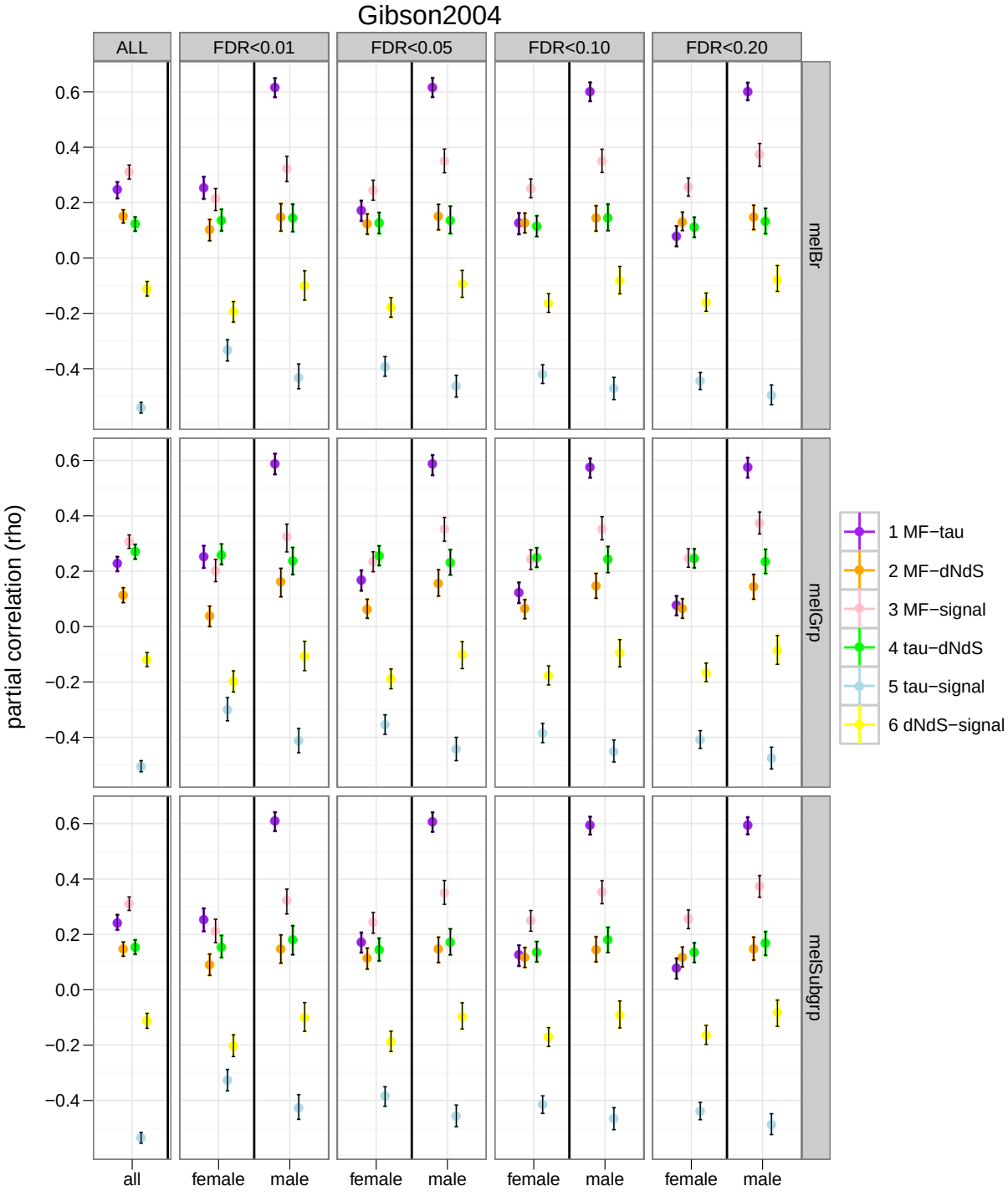

Parisi2003−WholeFly
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

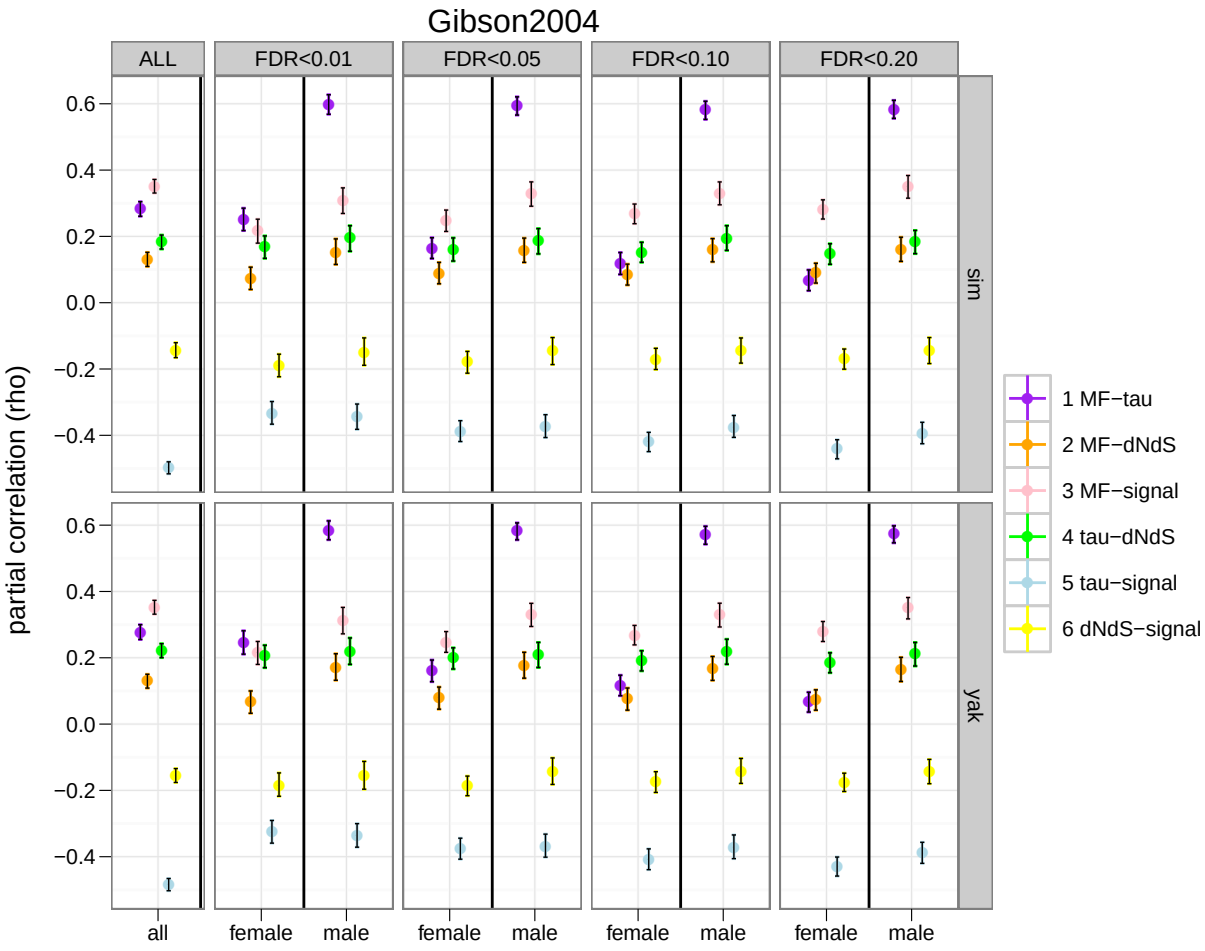

Parisi2003−Gonads
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

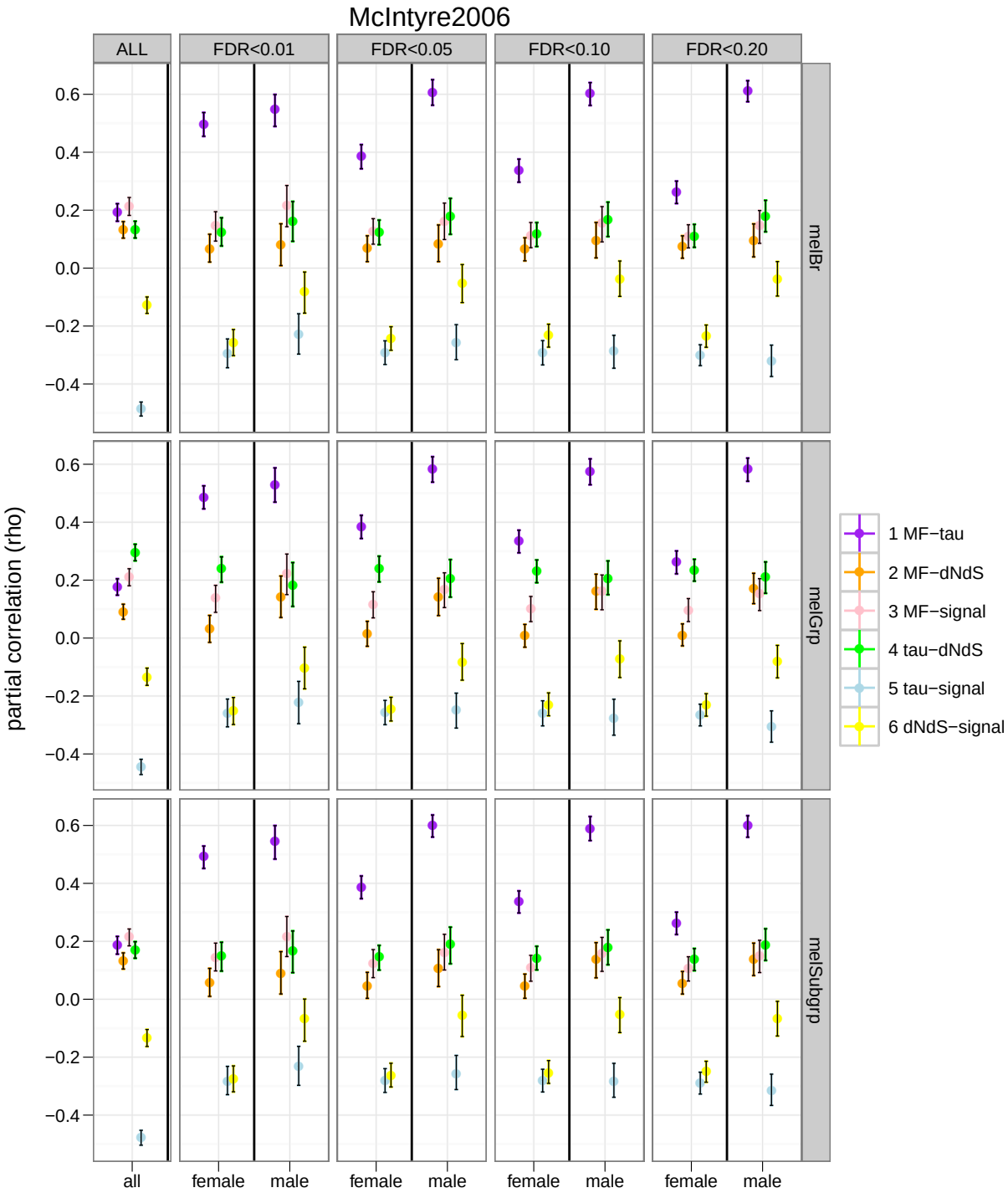

Parisi2003−Gonads
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

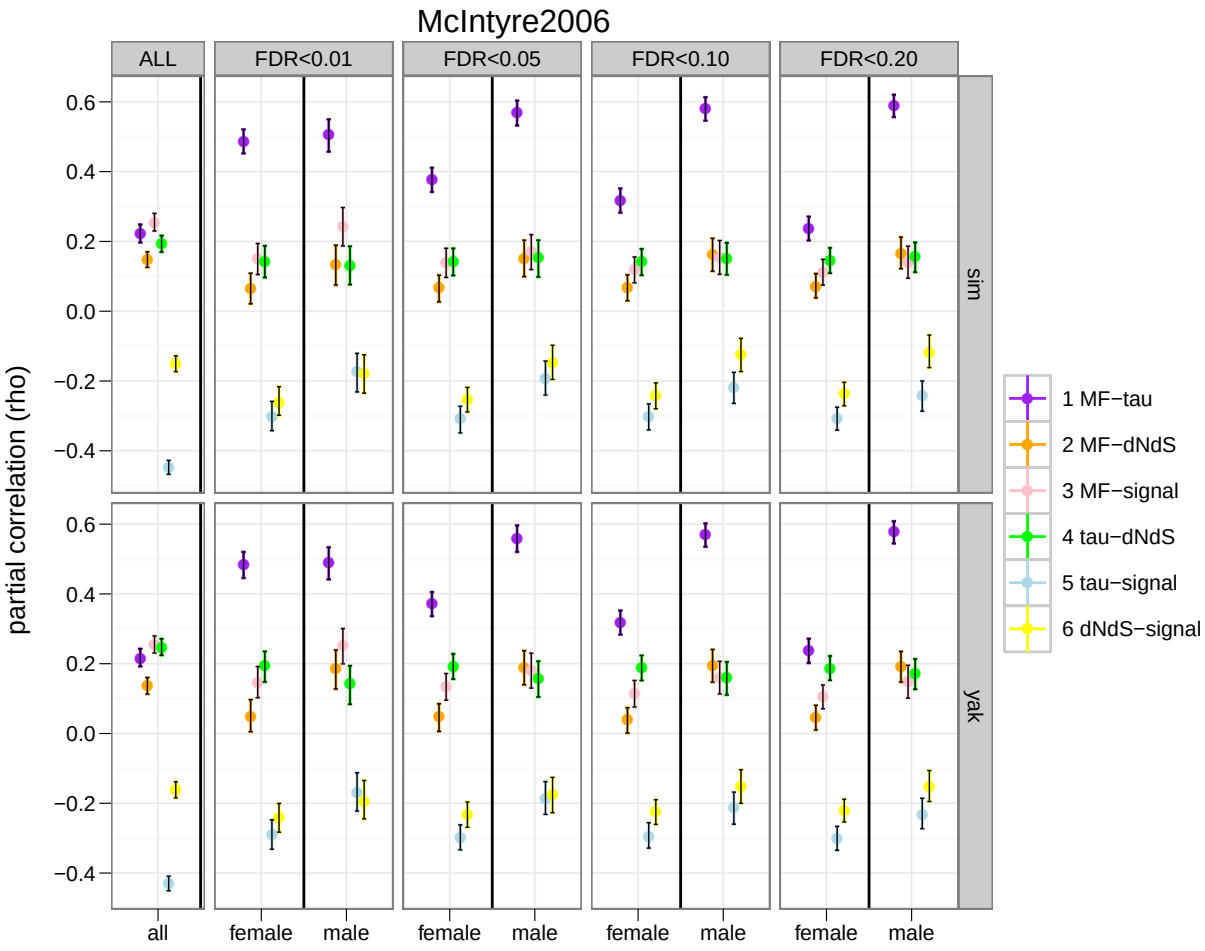

Ranz2003
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

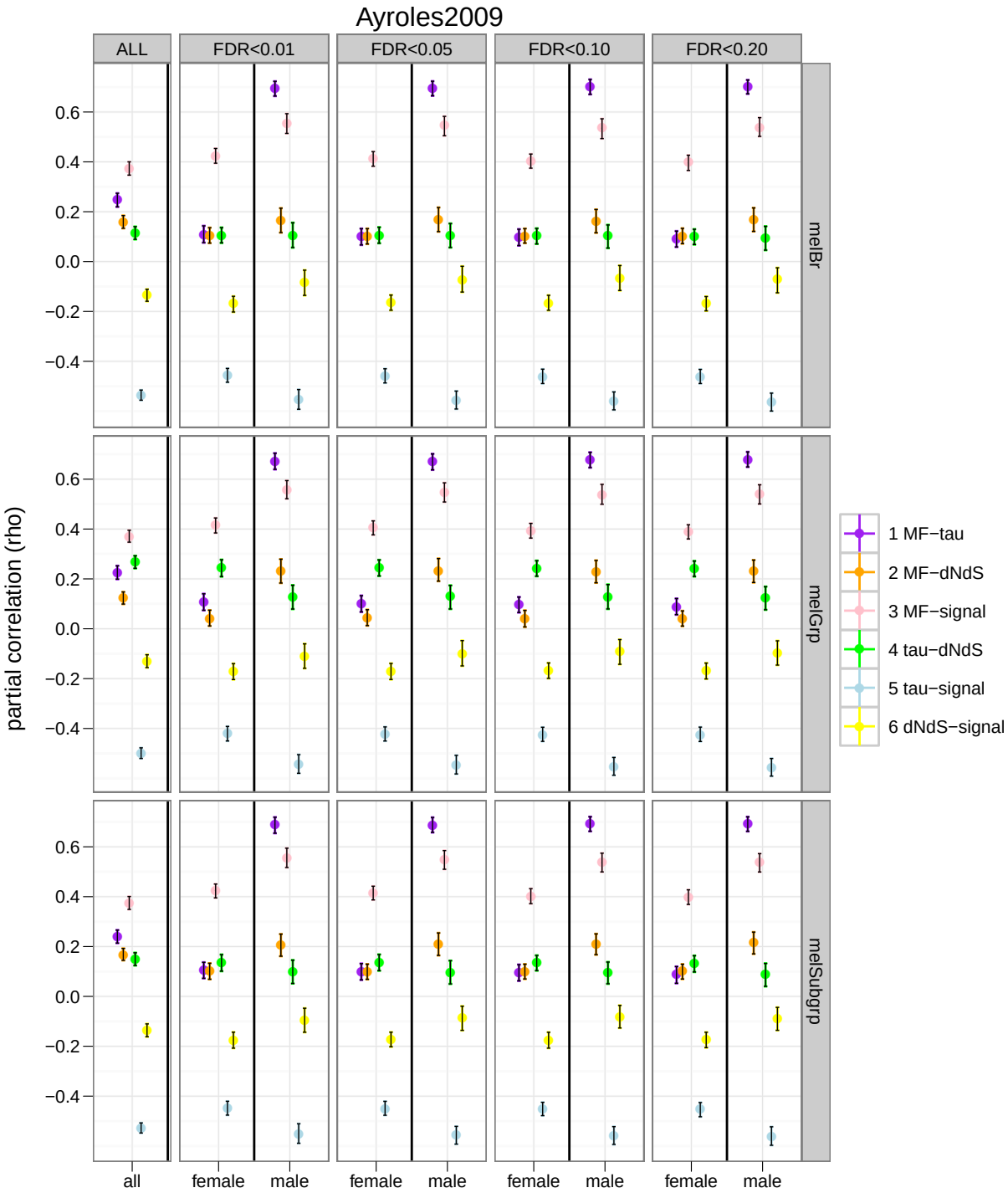

Ranz2003
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

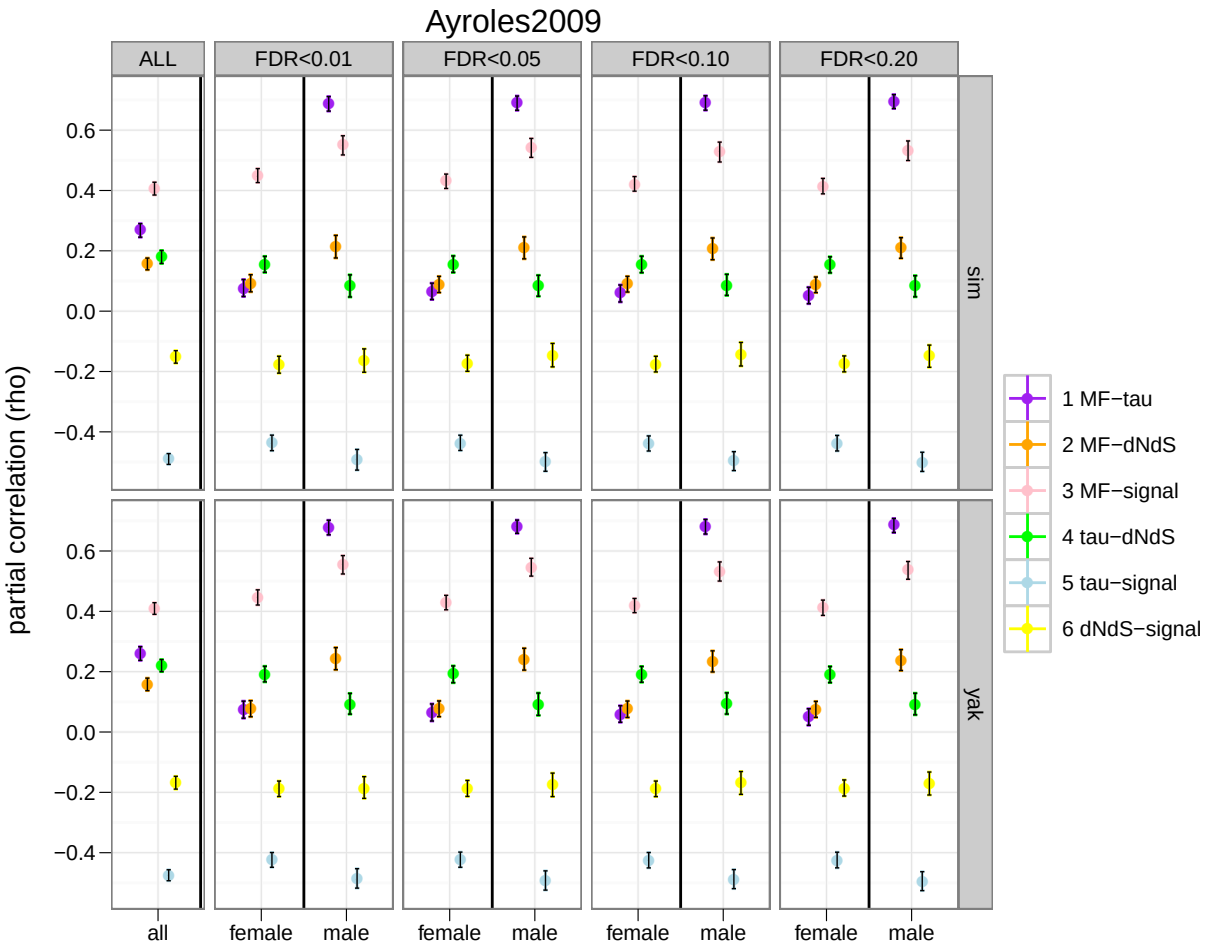

Gibson2004
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal


Gibson2004
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal


McIntyre2006
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal


McIntyre2006
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

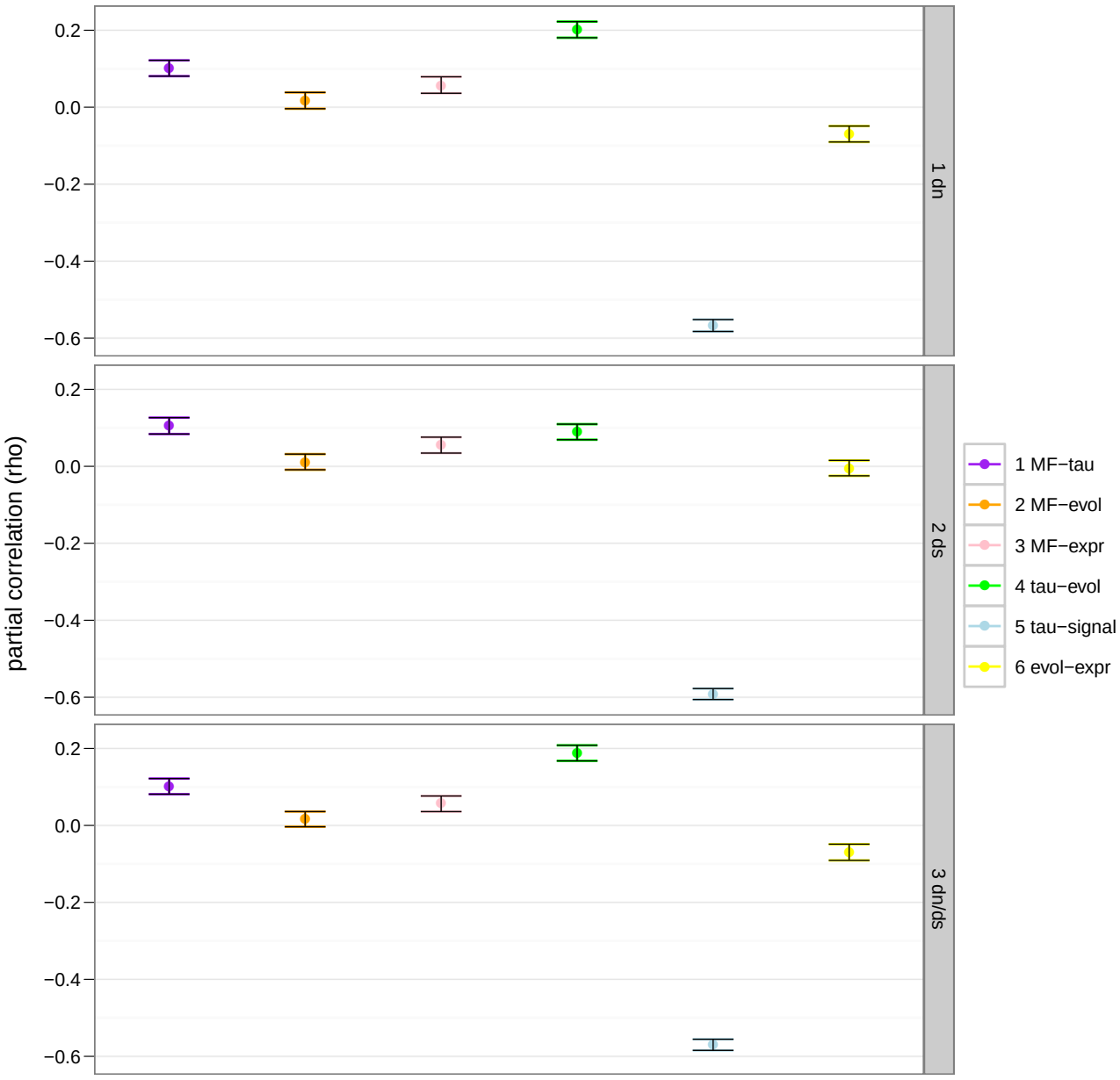

Ayroles2009
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
melBr
melGrp
melSubgrp
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

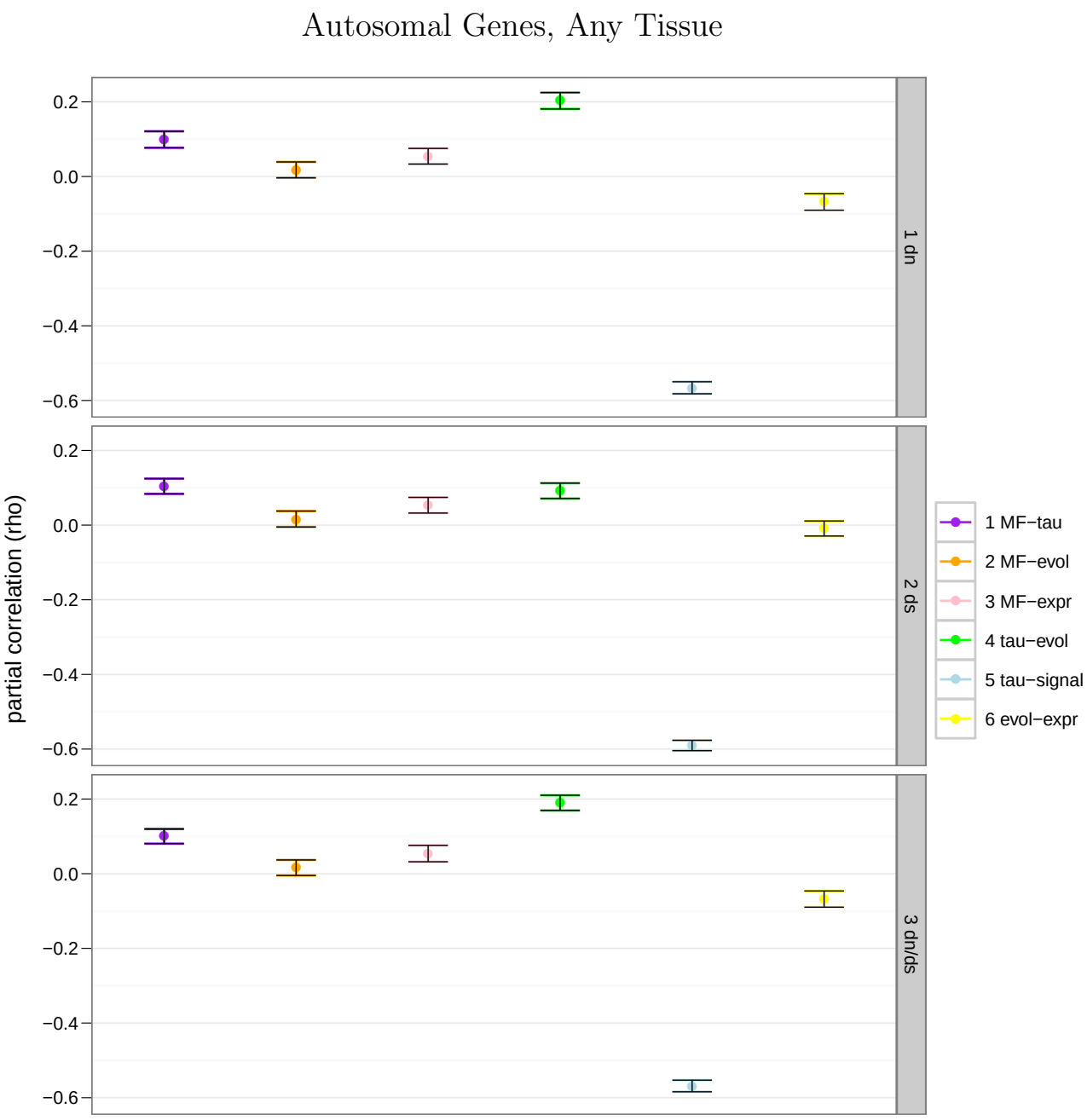

Ayroles2009
ALL
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
sim
yak
partial correlation (rho)
0.6
0.4
0.2
0.0
−0.2
−0.4
all
female
male
1 MF−tau
2 MF−dNdS
3 MF−signal
4 tau−dNdS
5 tau−signal
6 dNdS−signal

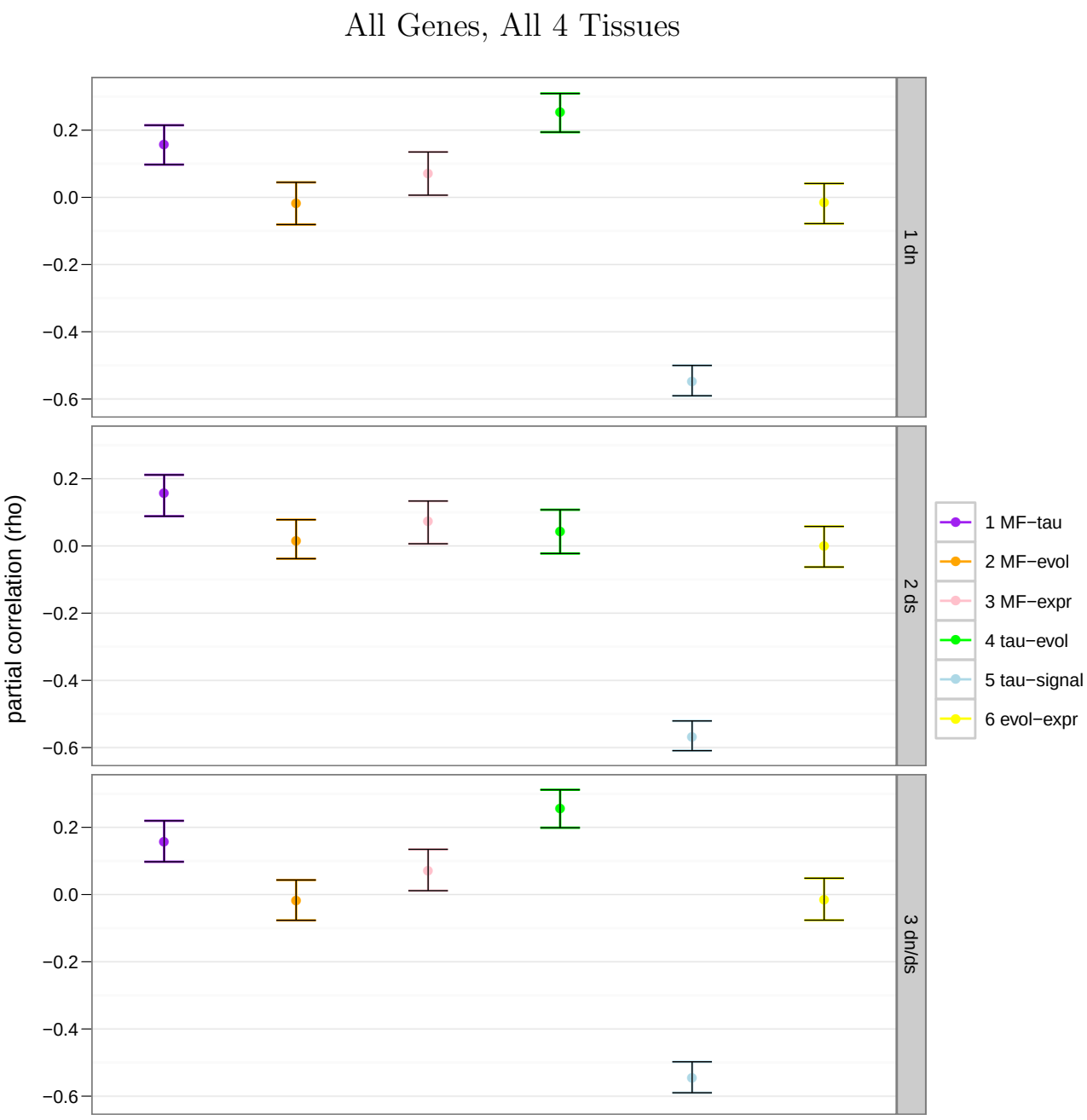
Partial correlations between sex-biased expression, tissue-specificity, rate of evolution, and expression level in mouse

Point estimates and 95% CI of the partial correlation (ρ) between $|\log(M/F)|$ and τ (purple), $|\log(M/F)|$ and evolutionary rate (orange), $|\log(M/F)|$ and expression level (pink), τ and evolutionary rate (green), τ and expression level (light blue), and evolutionary rate and expression level (yellow) are shown for *M. musculus*genes. Because high d_N/d_S can be driven by both high d_N and low d_S, I performed the partial correlations using d_N, d_S, or d_N/d_S. Note how the correlation between τ and d_S is significantly less than the correlation between τ and d_N or d_N/d_S, indicating that the relationship between tissue specificity and evolutionary rate only applies to changes in the protein coding sequence and is likely not an artifact of higher mutation rates in tissue-specific genes. Additionally, these patterns are not driven by special features of the X chromosome. I analyzed that data using all genes in the genome (All Genes) or only autosomal genes (Autosomal Genes), and I found similar patters. Finally, the estimation of M/F was performed on genes expressed in at least one tissue (Any Tissue) or on genes that were expressed in all 4 tissues (All 4 Tissues). The most notable difference between these results is that d_N/d_S is negatively correlated with expression level (i.e., rapidly evolving genes are lowly expressed) in the Any Tissue dataset but not in the All 4 Tissues dataset.

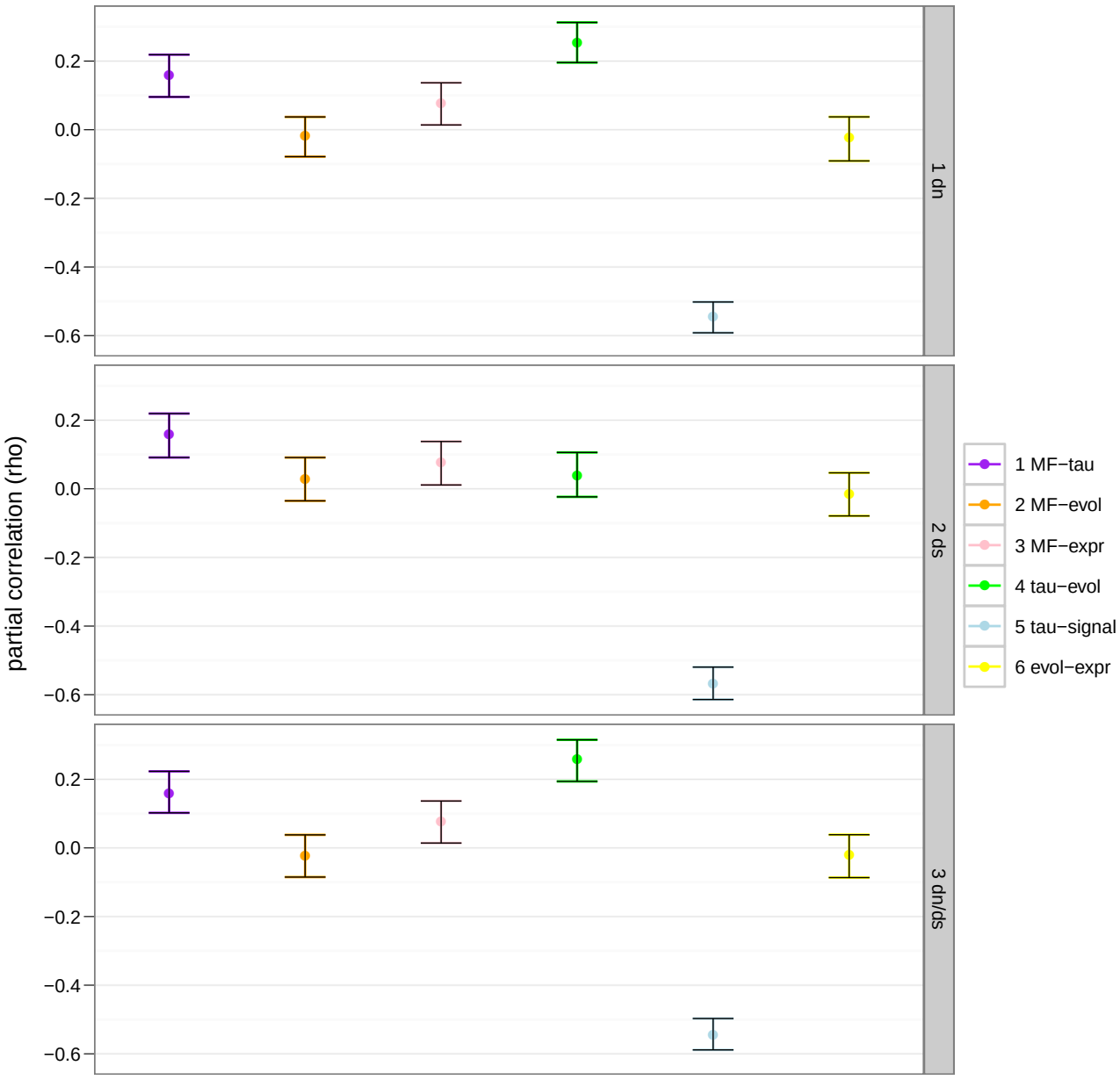
All Genes, Any Tissue




Autosomal Genes, Any Tissue
partial correlation (rho)
0.2
0.0
−0.2
−0.4
−0.6
1 dn
2 ds
3 dn/ds
1 MF−tau
2 MF−evol
3 MF−expr
4 tau−evol
5 tau−signal
6 evol−expr


All Genes, All 4 Tissues
partial correlation (rho)
0.2
0.0
−0.2
−0.4
−0.6
1 dn
2 ds
3 dn/ds
1 MF−tau
2 MF−evol
3 MF−expr
4 tau−evol
5 tau−signal
6 evol−expr

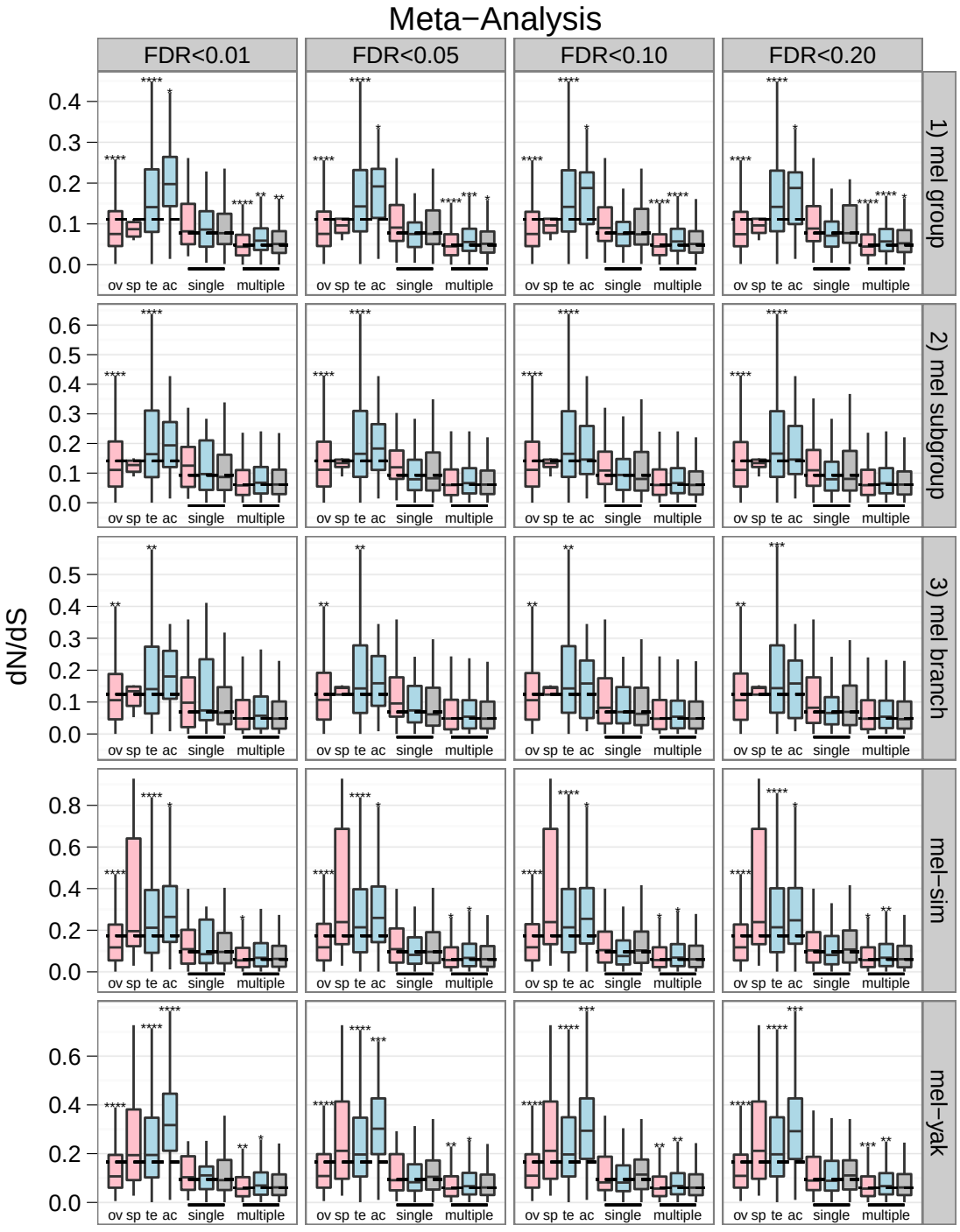
Autosomal Genes, All 4 Tissues

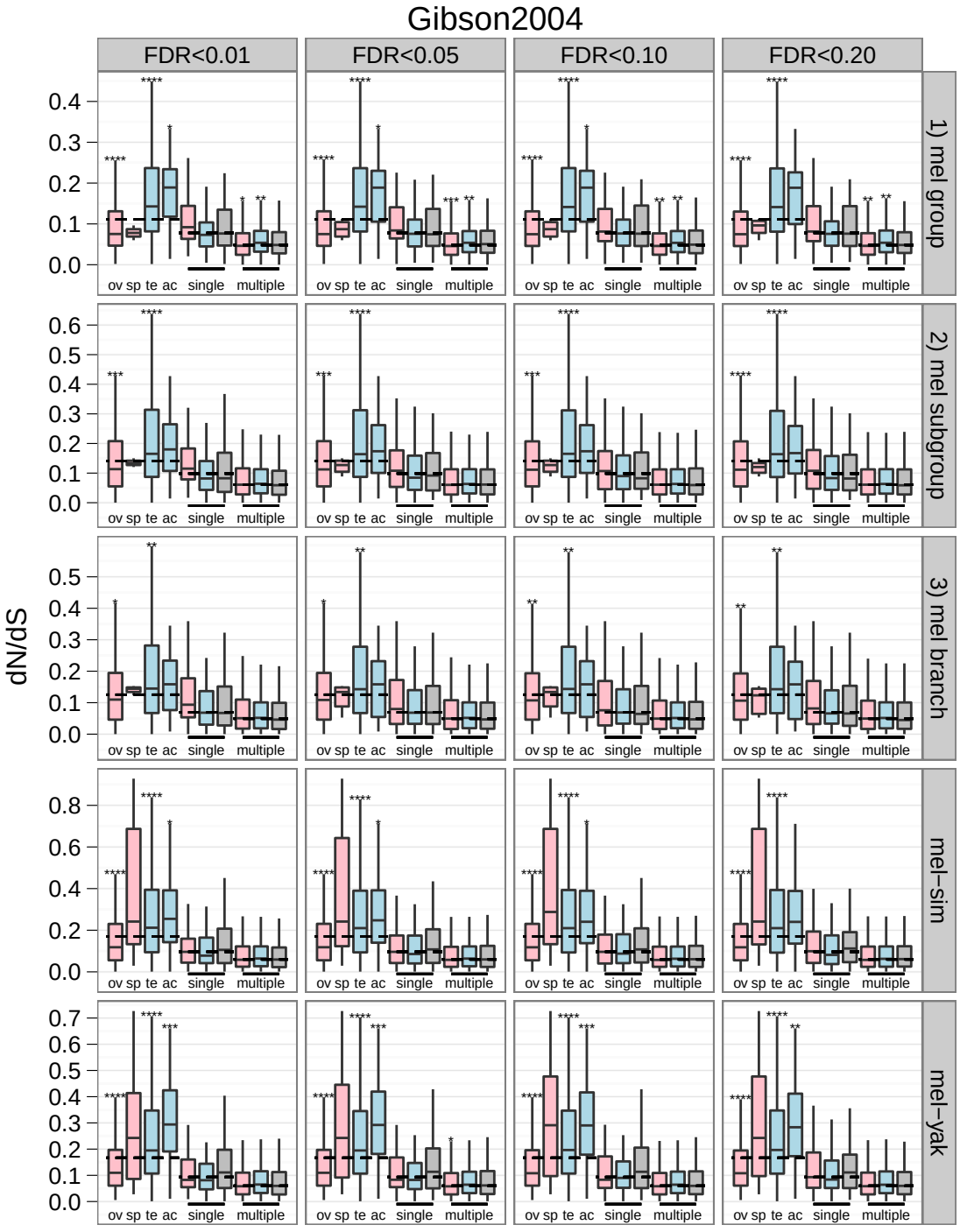


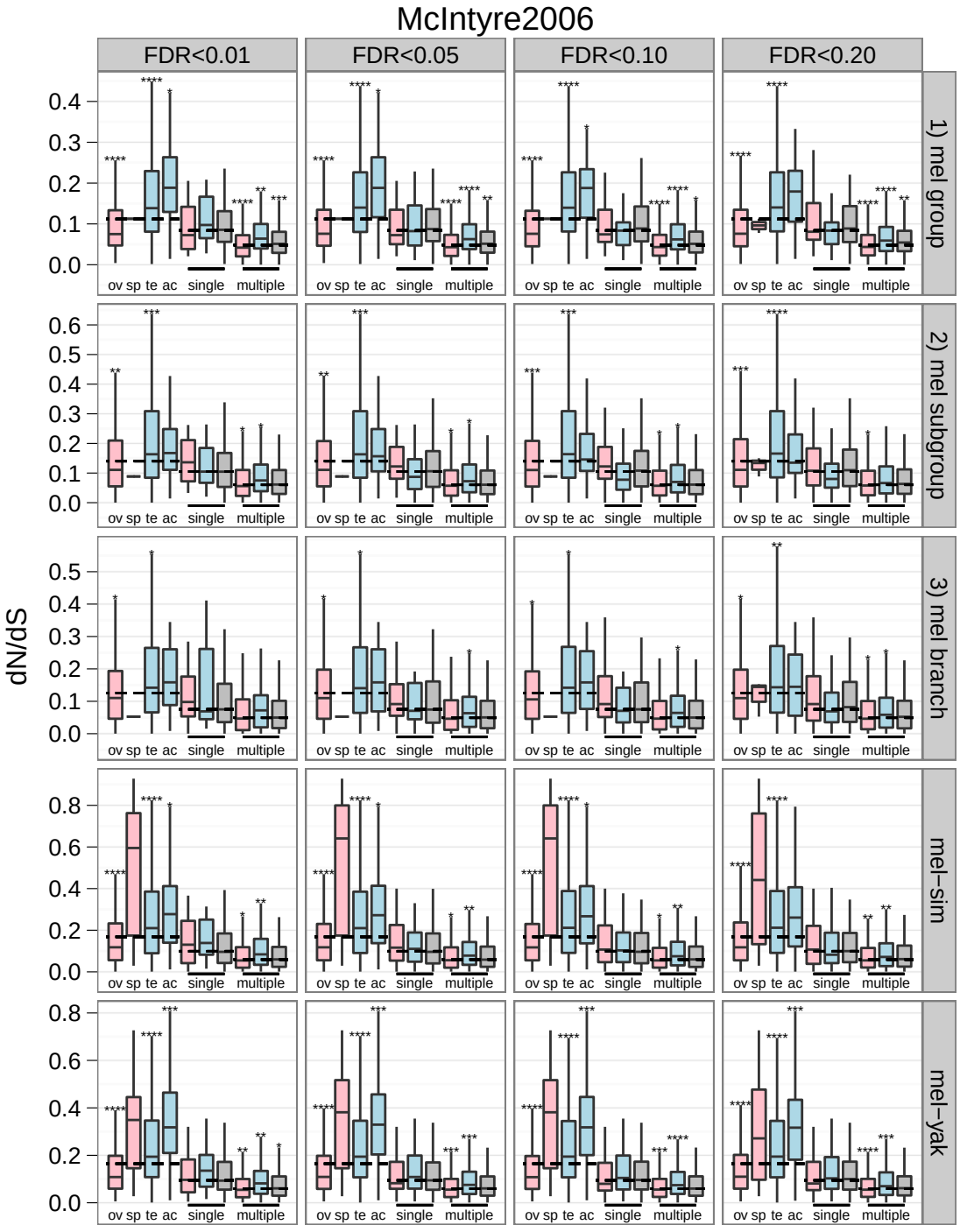
Effect of sex-limited tissues on the relationship between sex-biased expression and rates of evolution

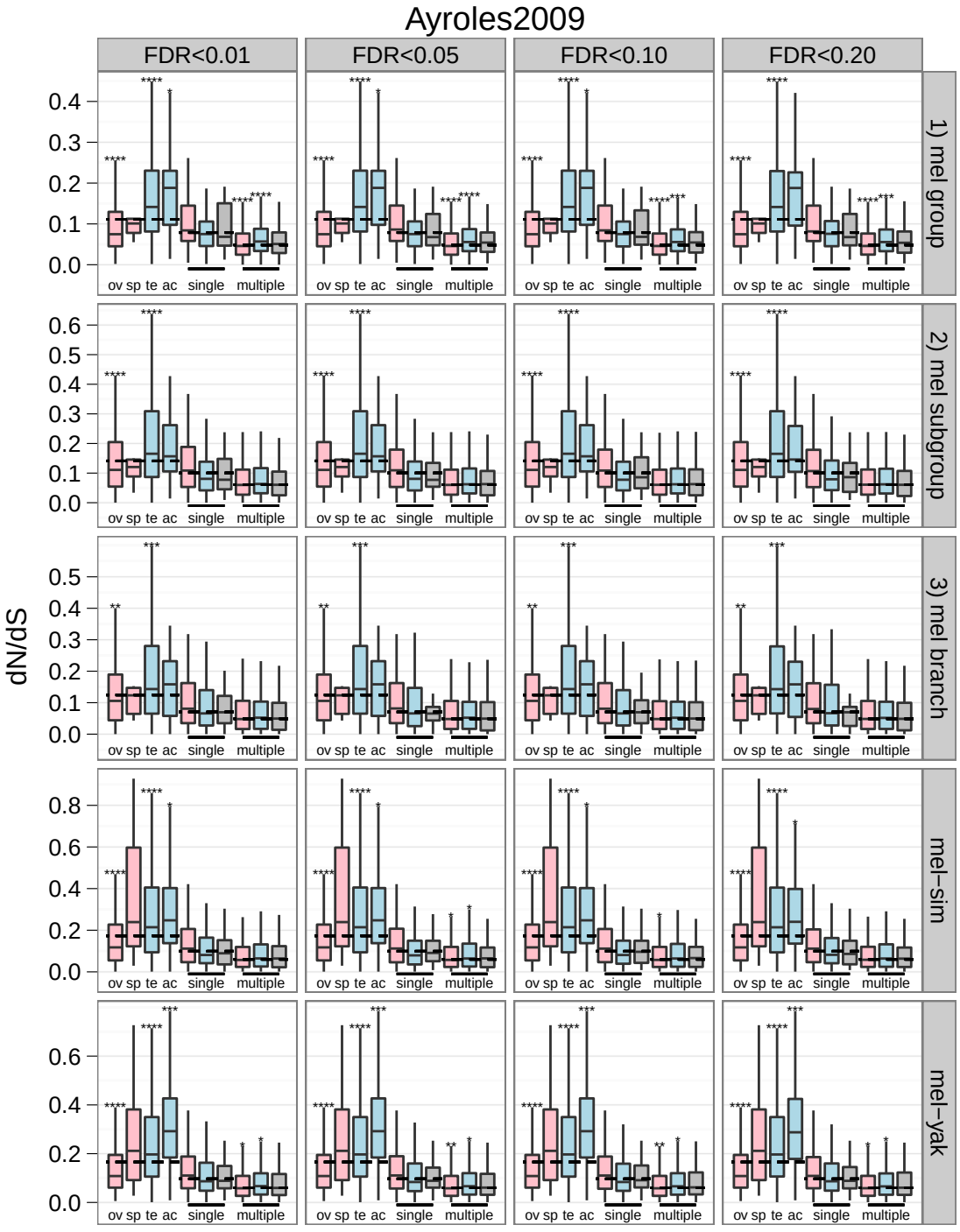
Evolutionary rates, sex-bias, and breadth of expression

Boxplots show the distribution of d_N/d_S for female-biased (red), male-biased (blue), and unbiased (gray) genes, with outliers omitted. Genes were divided into those that are expressed in a single sex-limited tissue (ov=ovary, sp=spermatheca, te=testis, ac=accessory gland), those that are expressed in a single shared somatic tissue (single), and those that are expressed in multiple tissues (multiple). Sex-biased genes were identified for each data set with a large enough sample size to compare d_N/d_S between sex-bias classes (the data sets are indicated at the top of each page), using four different FDR cut-offs (0.01, 0.05, 0.10, 0.20). d_N/d_S was calculated five ways: along all of the branches in the *melanogaster* species group (mel group); along all of the branches in the *melanogaster* subgroup (mel subgroup); along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (mel branch); pair-wise between *D. melanogaster* and *D. simulans* (mel-sim); and pair-wise between *D. melanogaster* and *D. yakuba* (mel-yak). The median d_N/d_S value across all genes within each panel is indicated by a dashed line. Estimates of d_N/d_S for each sex-bias class were compared to the other classes within each expression breadth category using a Mann-Whitney test, and asterisks indicate significant differences ($*P < 10^{-2}$, $**P < 10^{-4}$, $***P < 10^{-6}$, $****P < 10^{-8}$).


Meta−Analysis
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS
ov sp te ac single multiple

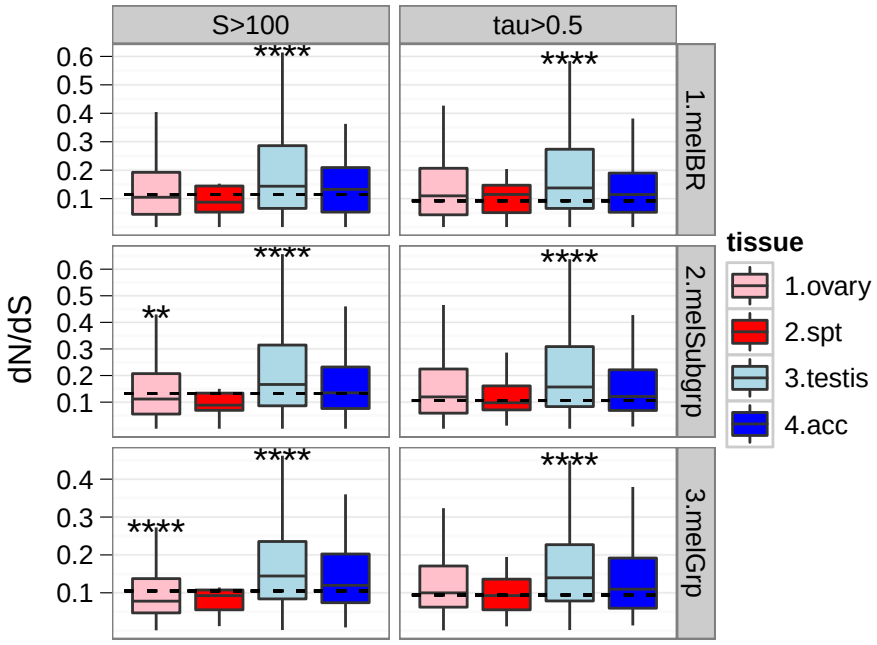

Gibson2004
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS
ov sp te ac single multiple


McIntyre2006
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS
ov sp te ac single multiple


Ayroles2009
FDR<0.01
FDR<0.05
FDR<0.10
FDR<0.20
1) mel group
2) mel subgroup
3) mel branch
mel−sim
mel−yak
dN/dS
ov sp te ac single multiple

Evolutionary rates and expression in sex-limited tissues

The previous analyses of genes expressed in sex-limited tissues required that the genes also be sex-biased when expression is measured in whole flies. I also looked for differences in the evolutionary rate of genes primarily expressed in sex-limited tissues regardless of whether they are sex-biased in measurements from whole flies. Boxplots show the distribution of d_N/d_S for genes that are primarily expressed in sex-limited tissues. Tissue-biased expression was determined two ways: 1) Genes with a signal intensity, S, > 100 in one tissue and $S < 100$ in all other tissues are said to be tissue-biased (left column); 2) Genes with $\tau > 0.5$ are said to be tissue-biased in the tissue in which they are most highly expressed (right column). The rate of protein coding evolution (d_N/d_S) was calculated three ways: along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (melBR); along all of the branches in the *melanogaster* subgroup (melSubgrp); along all of the branches in the *melanogaster* species group (melGrp). Dashed lines represent the median d_N/d_S for all genes that have tissue-biased expression in both sex-limited and shared somatic tissues. Asterisks indicate sex-limited tissues with d_N/d_S values significantly greater or less than the average across all tissue-biased genes ($^*P < 10^{-2}$, $^{**}P < 10^{-4}$, $^{***}P < 10^{-6}$, $^{****}P < 10^{-8}$ using a Mann-Whitney test).


S>100
tau>0.5
1.melBR
2.melSubgrp
3.melGrp
dN/dS
0.6
0.5
0.4
0.3
0.2
0.1
****
**
tissue
1.ovary
2.spt
3.testis
4.acc

Faster-X Evolution of Male-Biased Genes

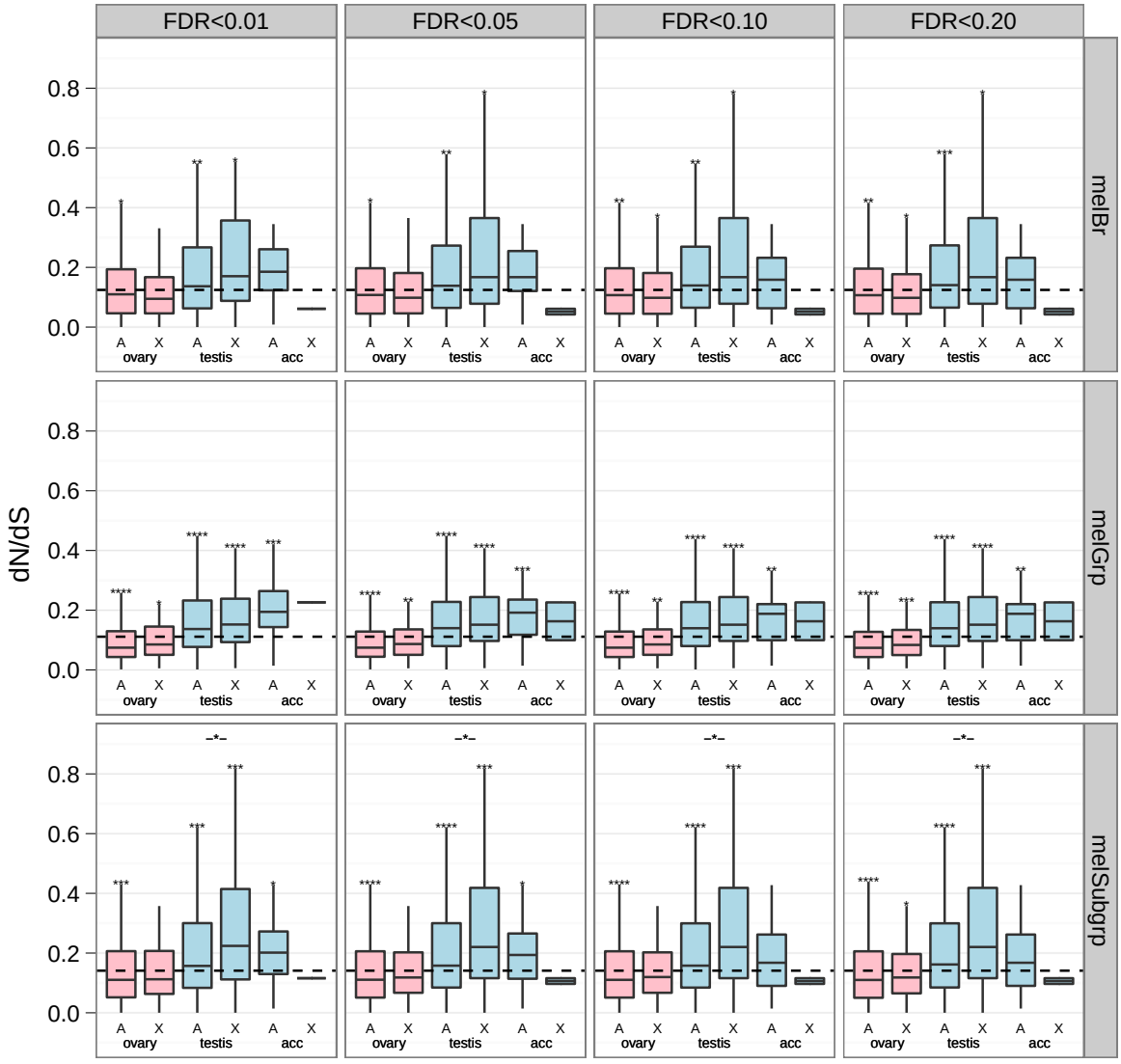
The X chromosome is hemizygous in males. This means that *de novo* recessive mutations will be exposed to selection in males, while *de novo* recessive mutations are not exposed to selection in females or on the autosomes. Therefore, if natural selection acts on *de novo* recessive beneficial mutations, male-expressed X-linked genes should experience more adaptive evolution than X-linked female-biased genes and autosomal genes (Charlesworth, Coyne and Barton, 1987). This has been termed the "faster-X" hypothesis, and numerous studies have attempted to identify faster-X evolution using both polymorphism and divergence estimates. Of note is the finding that X-linked male-biased genes in *D. melanogaster* experience more adaptive evolution than autosomal male-biased genes (Baines et al., 2008).

My results indicate that the rapid evolution of *D. melanogaster* male-biased genes is limited to those genes expressed in male-limited reproductive tissues. It is thus possible that the faster-X evolution of male-biased genes may also be limited to only those genes expressed in male-limited reproductive tissues. I tested this hypothesis by repeating my analysis of rates of evolution of sex-biased genes expressed in different tissues on X-linked and autosomal genes separately. The results are presented below fow the analysis of two types of data: 1) sex-biased genes were divided into those expressed in a single sex-limited tissue, a single shared somatic tissue, and multiple tissues; 2) genes expressed in a single sex-limited tissue were considered regardless of whether they are sex-biased in whole fly measurements. I find some evidence for faster-X evolution of testis-biased genes, but this pattern depends on how d_N/d_S and sex-biased expression are measured. I find no evidence for faster-X evolution of male-biased genes expressed in shared somatic tissues.

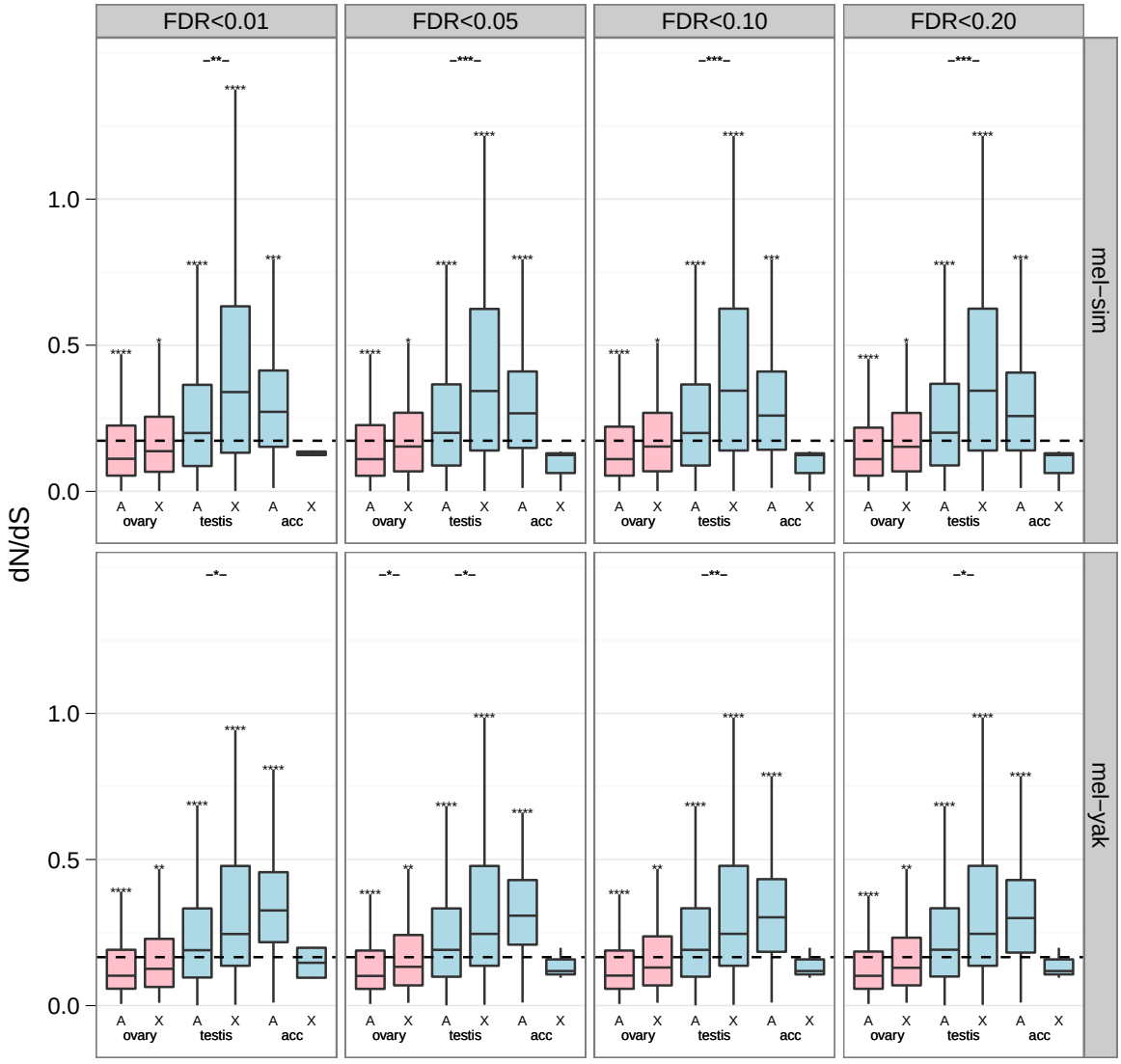
Faster-X evolution, sex-bias, and breadth of expression

Boxplots show the distribution of d_N/d_S for female-biased (red), male-biased (blue), and unbiased (gray) genes, with outliers omitted. Autosomal (A) and X-linked (X) genes are plotted separately. Genes were divided into those that are expressed in a single sex-limited tissue (ovary, testis, accessory gland [acc]), those that are expressed in a single shared somatic tissue, and those that are expressed in multiple tissues (the type of genes under consideration is given atop each page). Spermatheca expressed female-biased genes were excluded because of small sample sizes. Sex-biased genes were identified for each data set with a large enough sample size to compare d_N/d_S between sex-bias classes (the data sets are indicated at the top of each page), using four different FDR cut-offs (0.01, 0.05, 0.10, 0.20). d_N/d_S was calculated five ways: along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (melBr); along all of the branches in the *melanogaster* species group (melGrp); along all of the branches in the *melanogaster* subgroup (melSubgrp); pair-wise between *D. melanogaster* and *D. simulans* (mel-sim); and pair-wise between *D. melanogaster* and *D. yakuba* (mel-yak). The median d_N/d_S value across all genes within each panel is indicated by a dashed line. Estimates of d_N/d_S for each set of genes was compared to all other genes within each panel using a Mann-Whitney test, and asterisks above a box and whisker indicate a significant difference from the other genes in the panel ($*P < 10^{-2}$, $**P < 10^{-4}$, $***P < 10^{-6}$, $****P < 10^{-8}$). Additionally, autosomal and X-linked genes were compared, and asterisks flanked by dashes indicate a significant difference (-*-$P < 10^{-2}$, -**-$P < 10^{-4}$, -*** - $P < 10^{-6}$, -****-$P < 10^{-8}$).

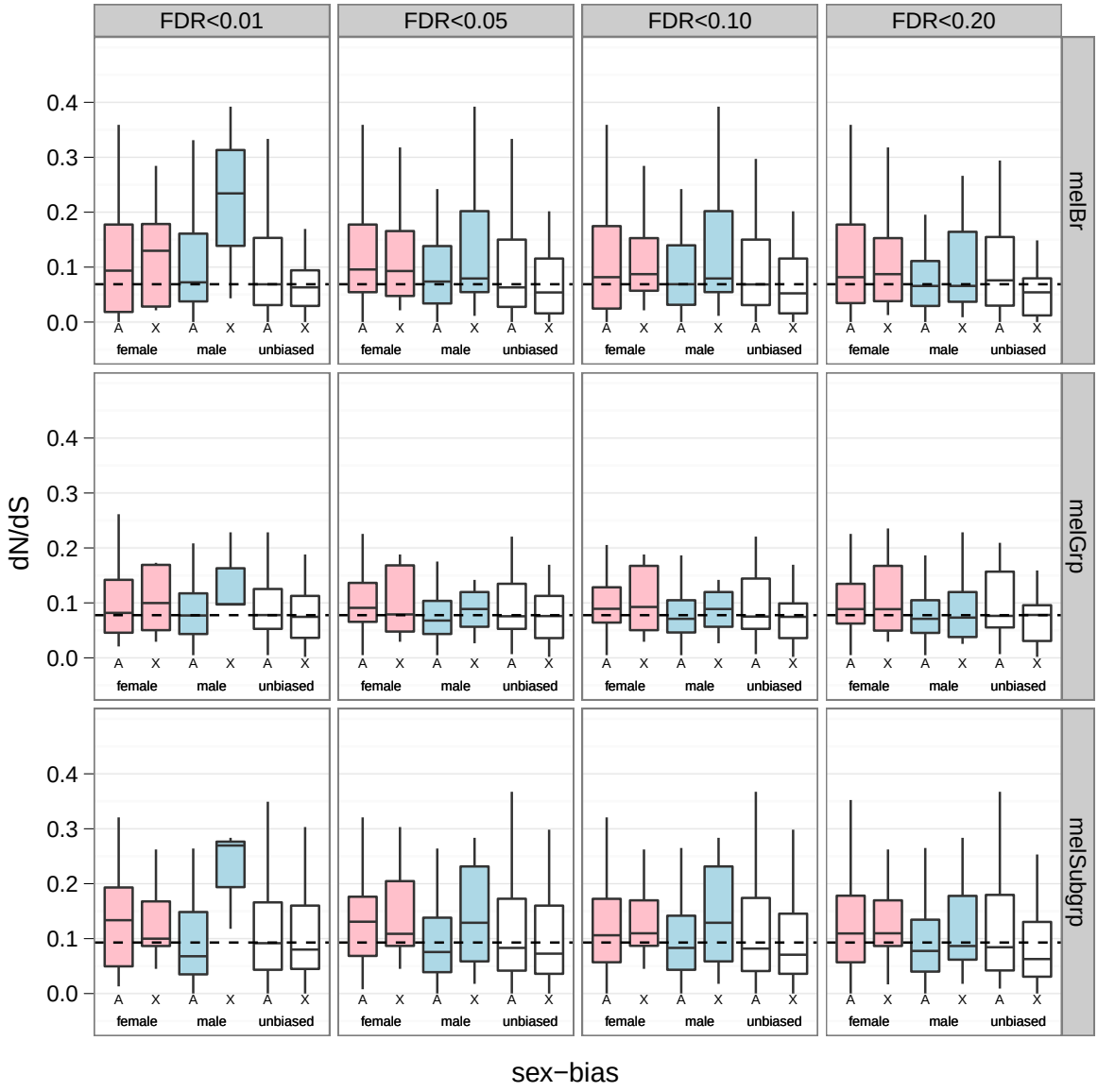
Meta-Analysis, Sex-Limited Tissue



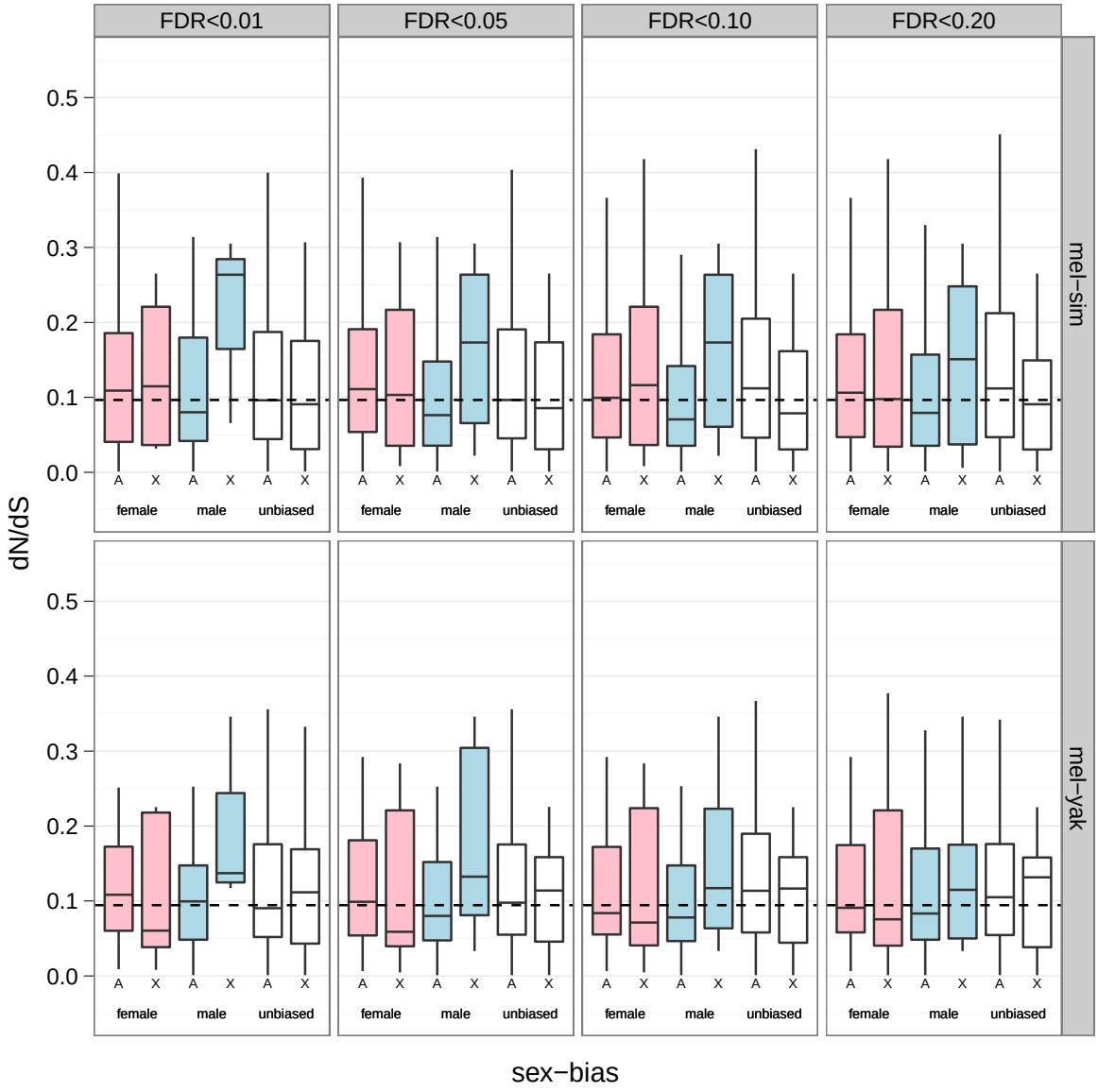
Meta-Analysis, Sex-Limited Tissue



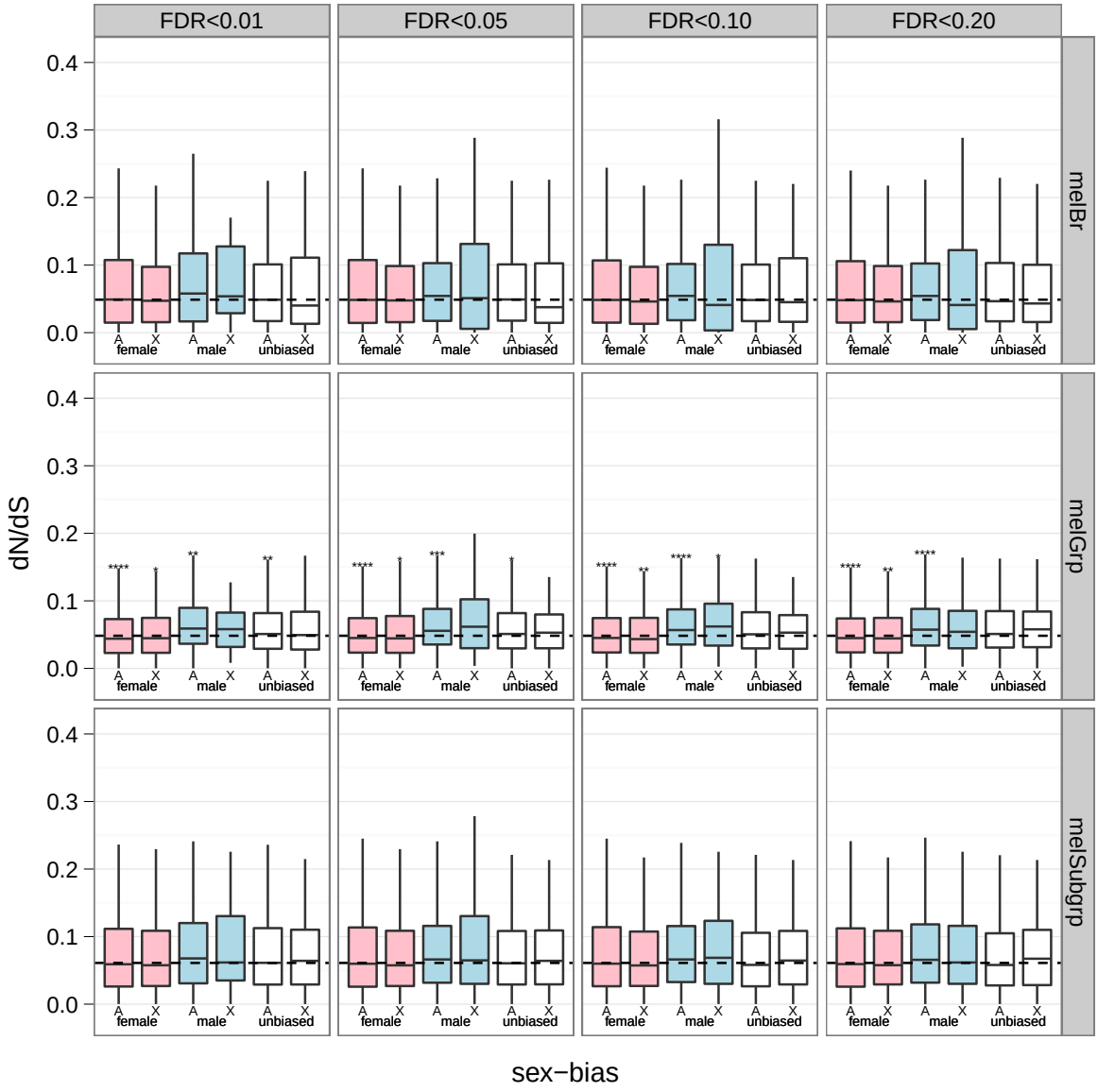
Meta-Analysis, Single Shared Somatic Tissue



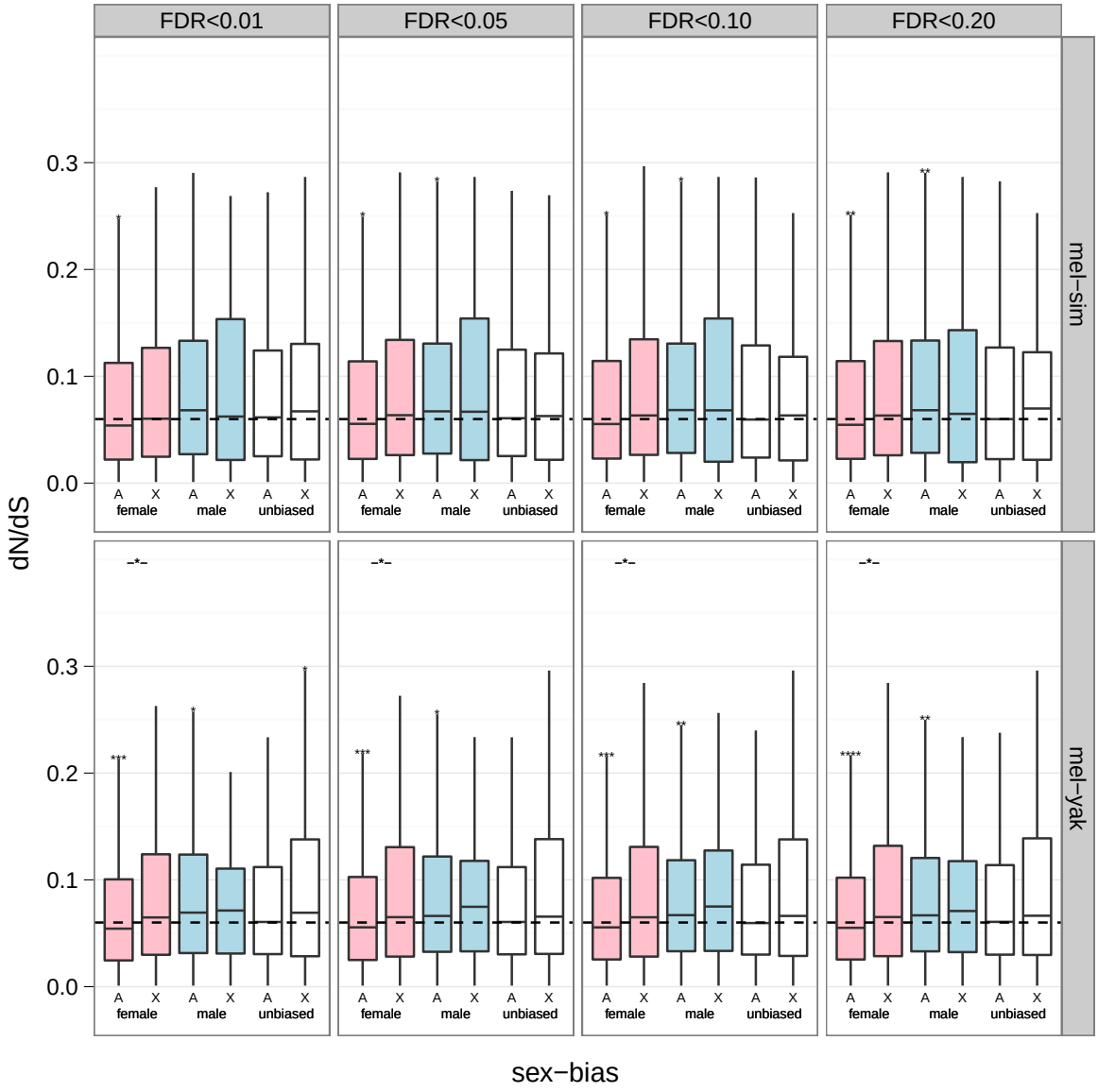
Meta-Analysis, Single Shared Somatic Tissue



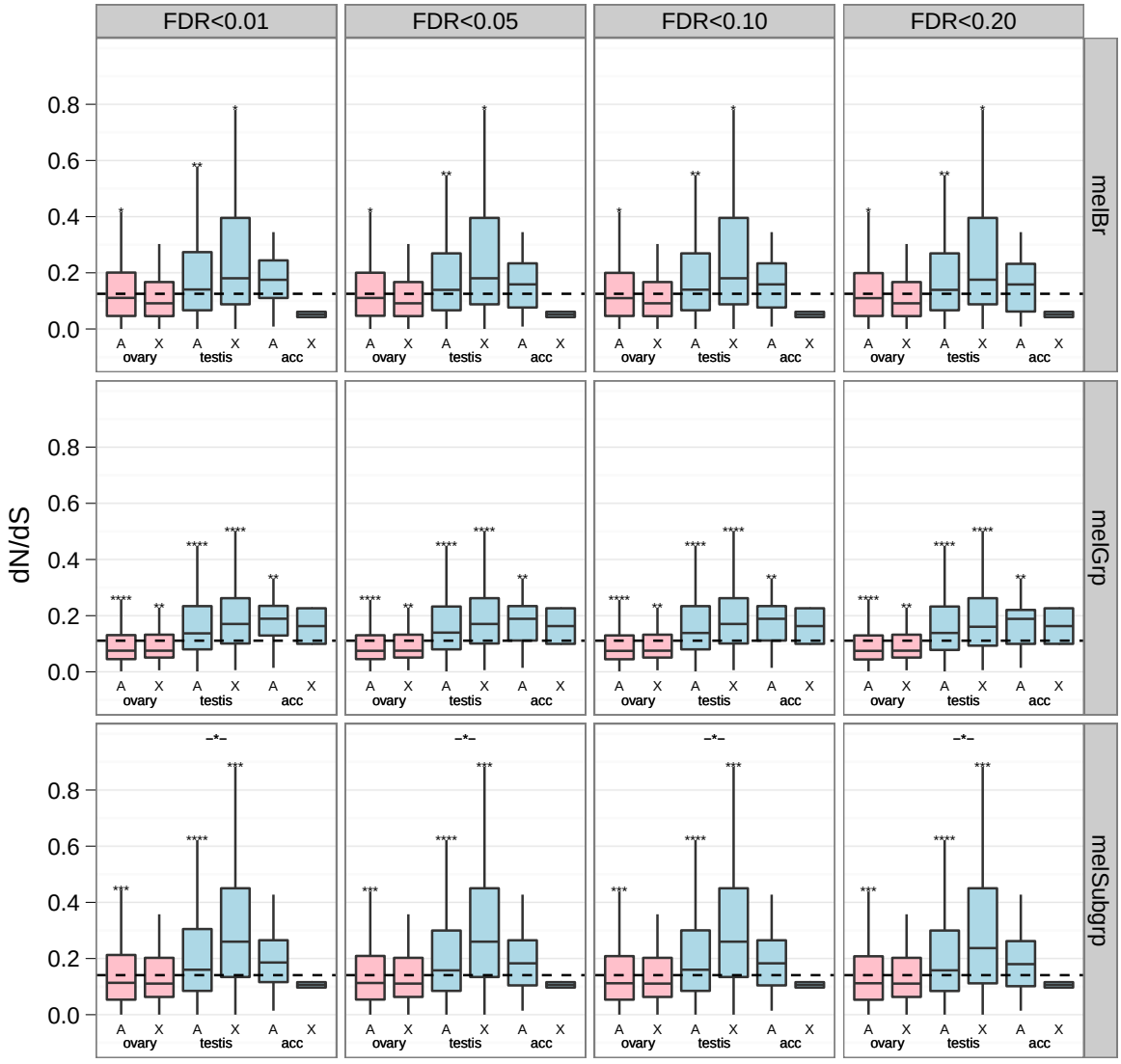
Meta-Analysis, Multiple Tissues



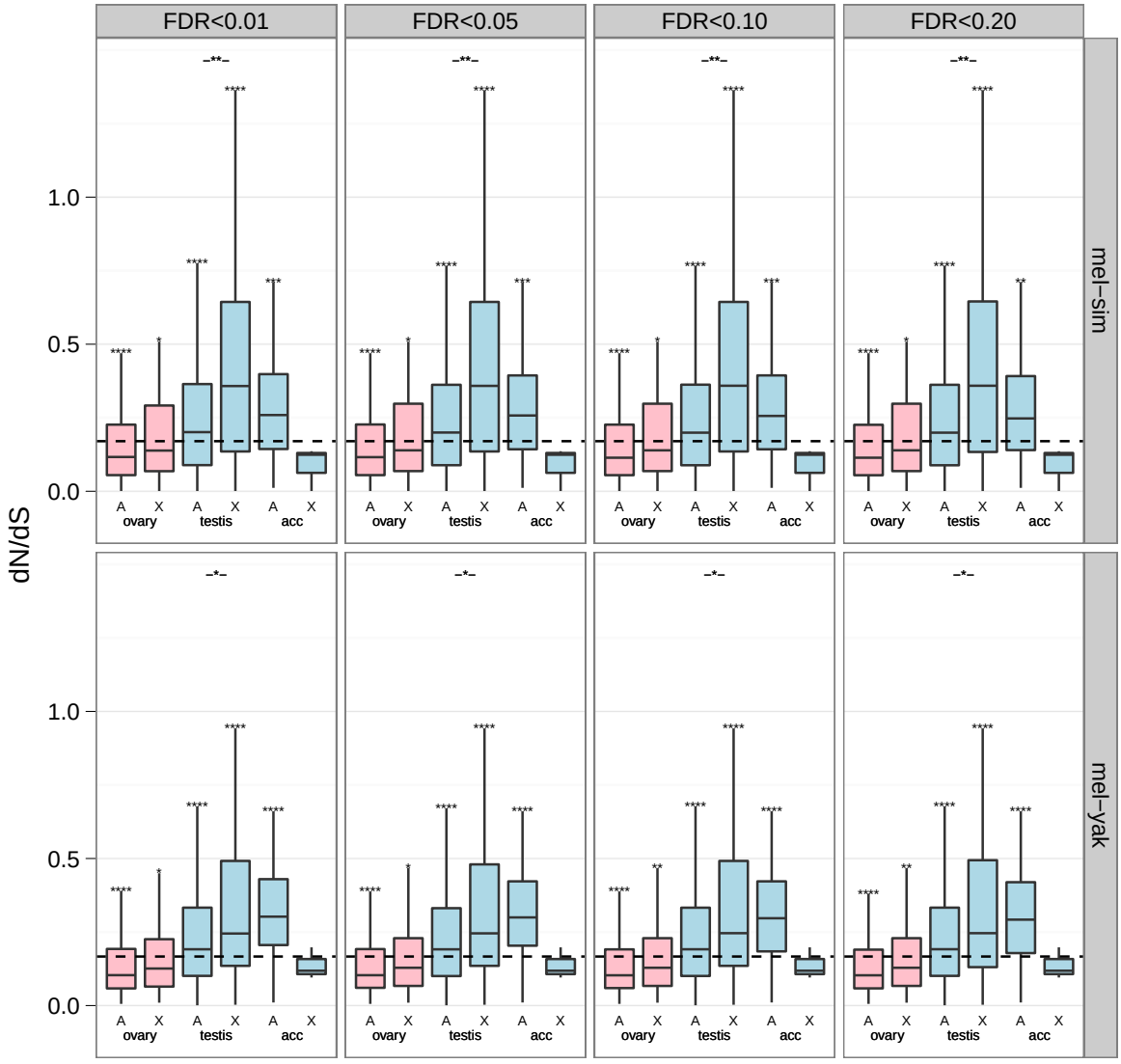
Meta-Analysis, Multiple Tissues



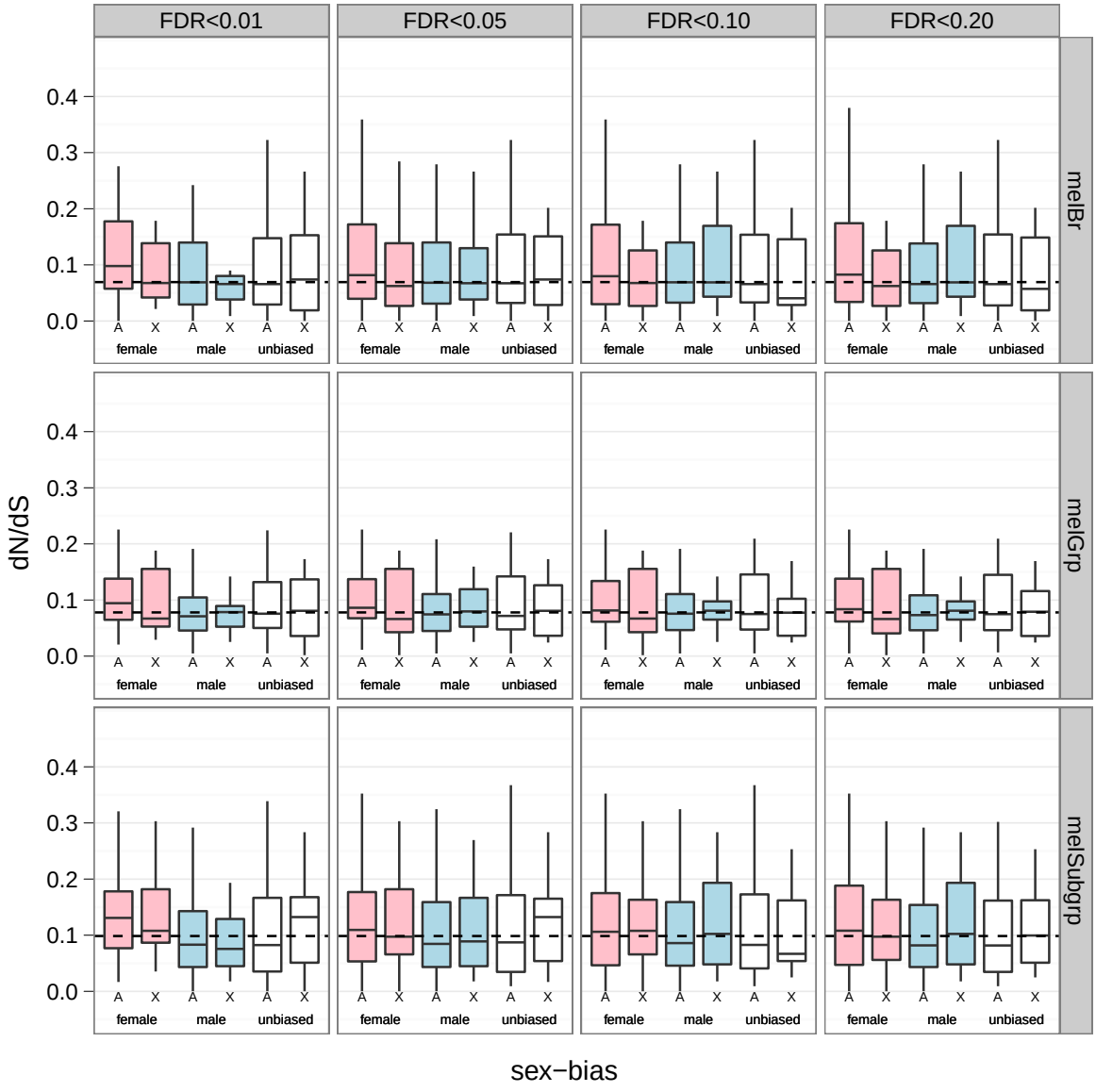
Gibson (2004), Sex-Limited Tissue



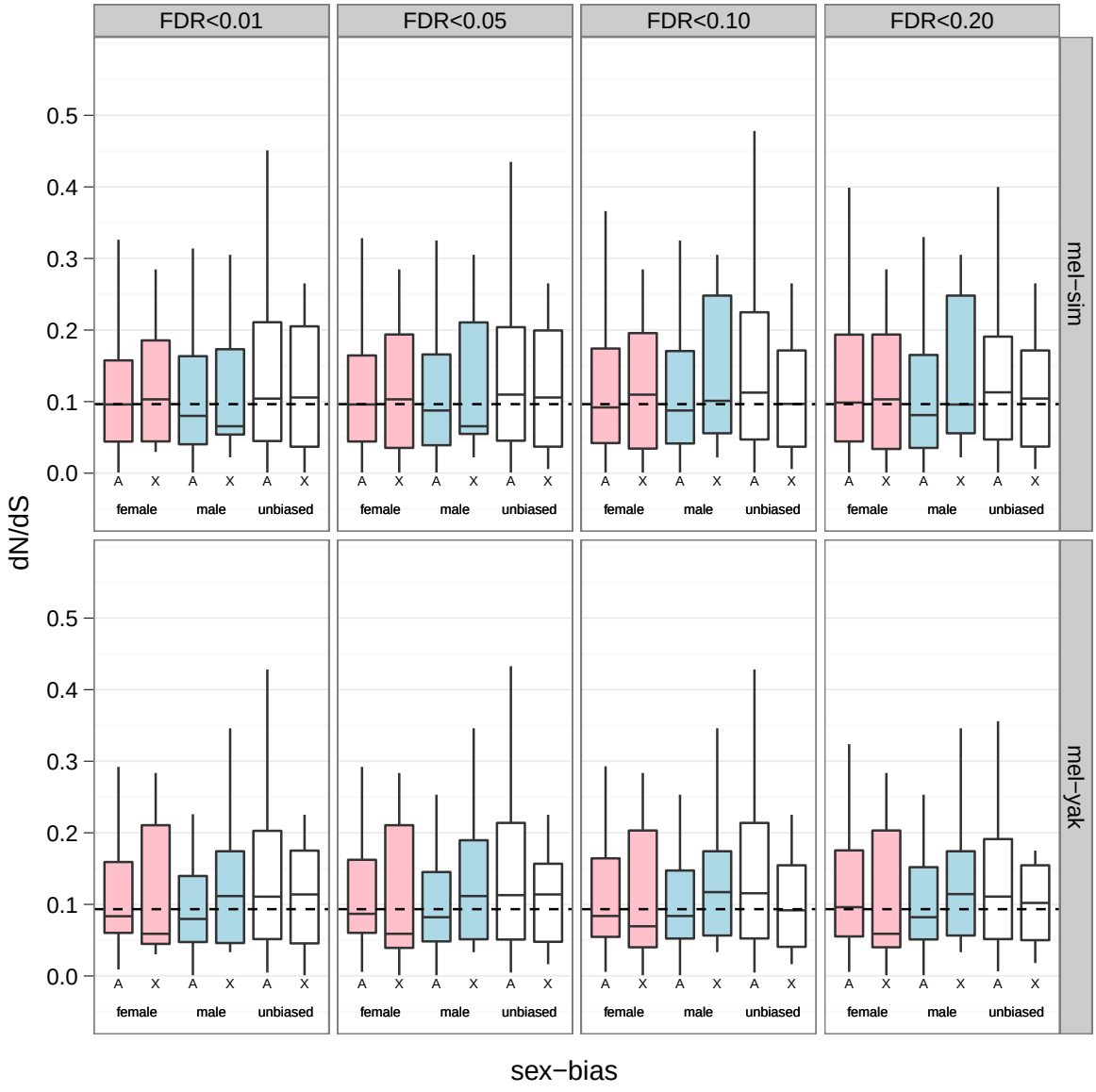
Gibson (2004), Sex-Limited Tissue



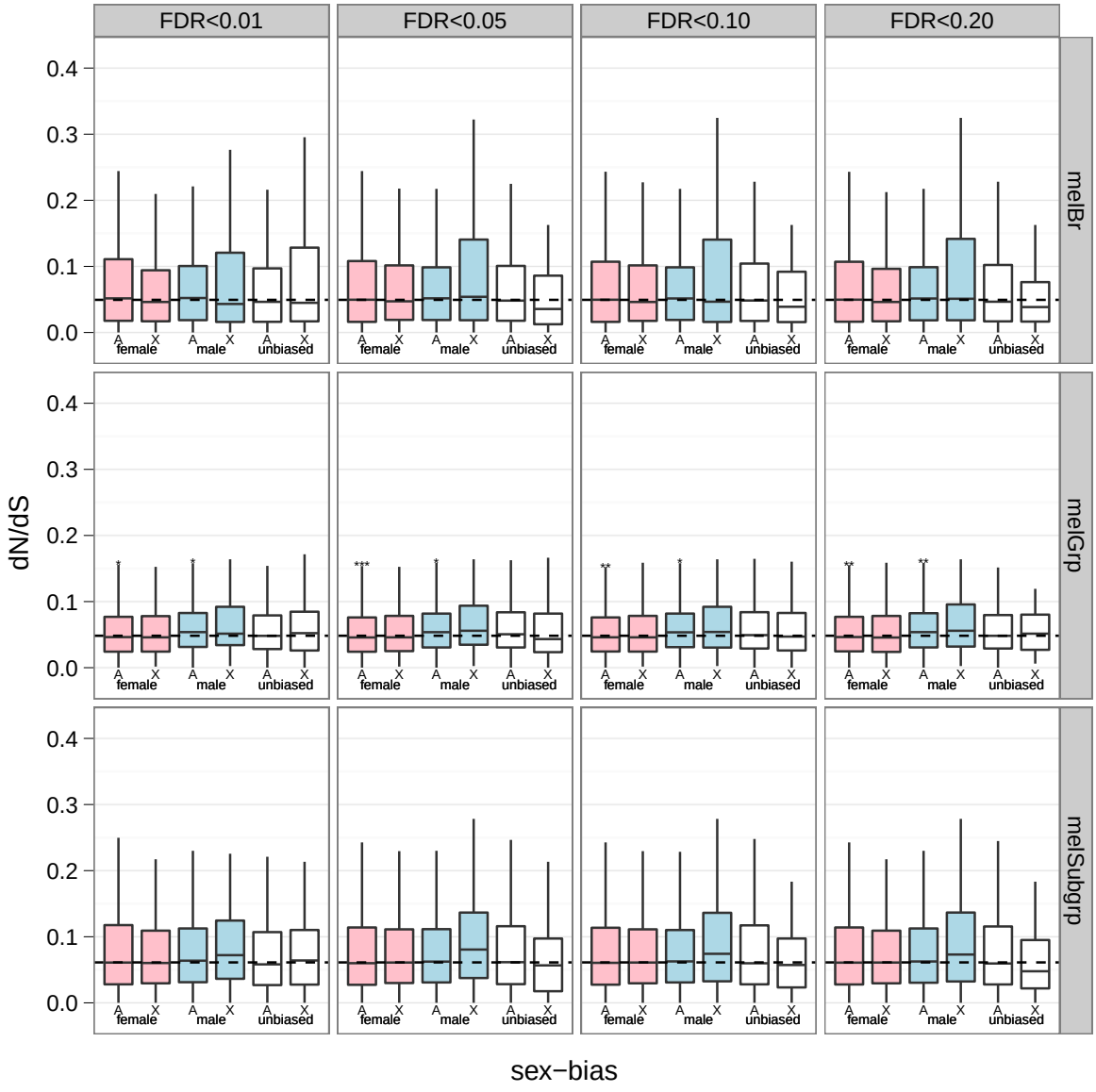
Gibson (2004), Single Shared Somatic Tissue



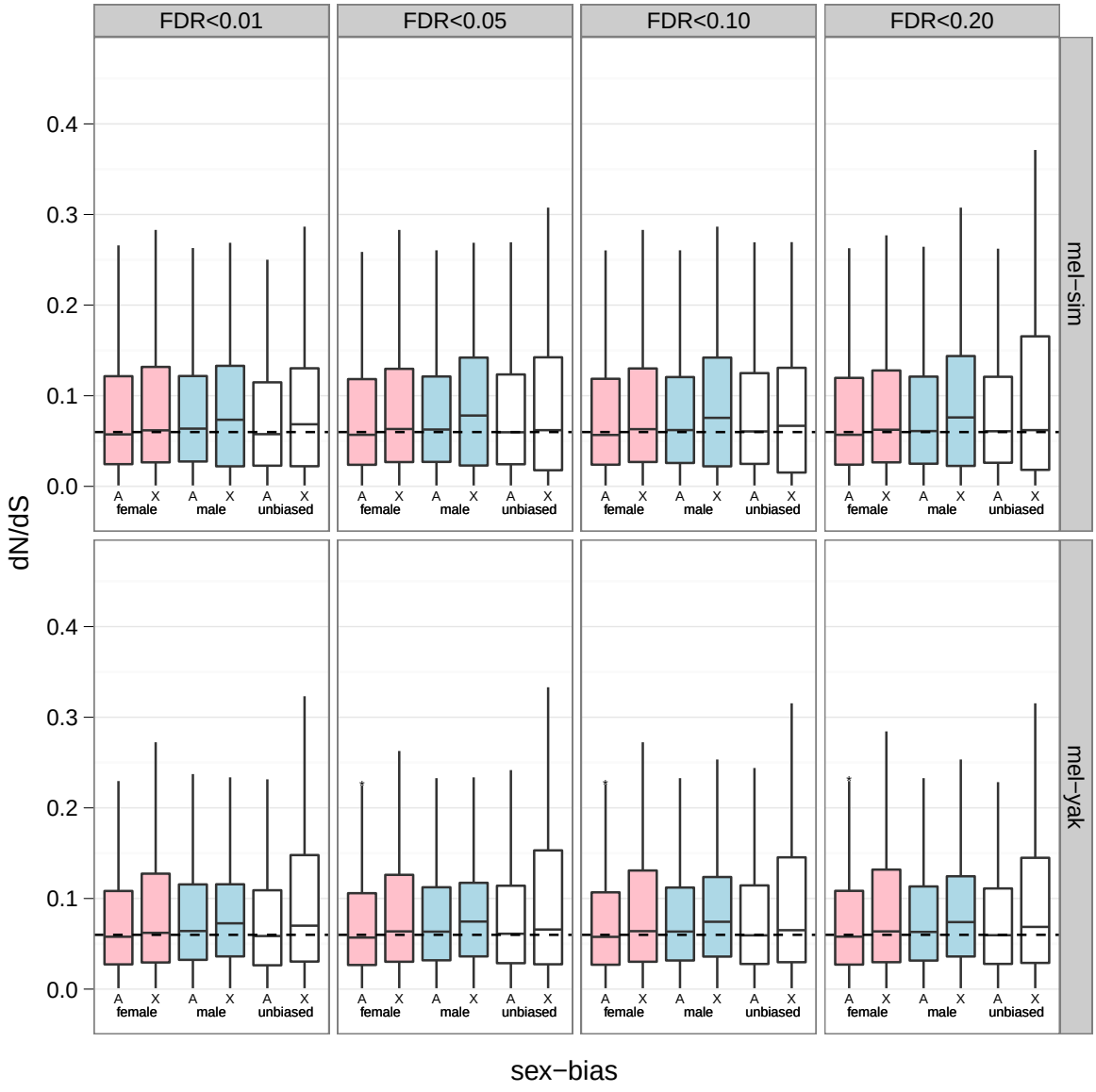
Gibson (2004), Single Shared Somatic Tissue



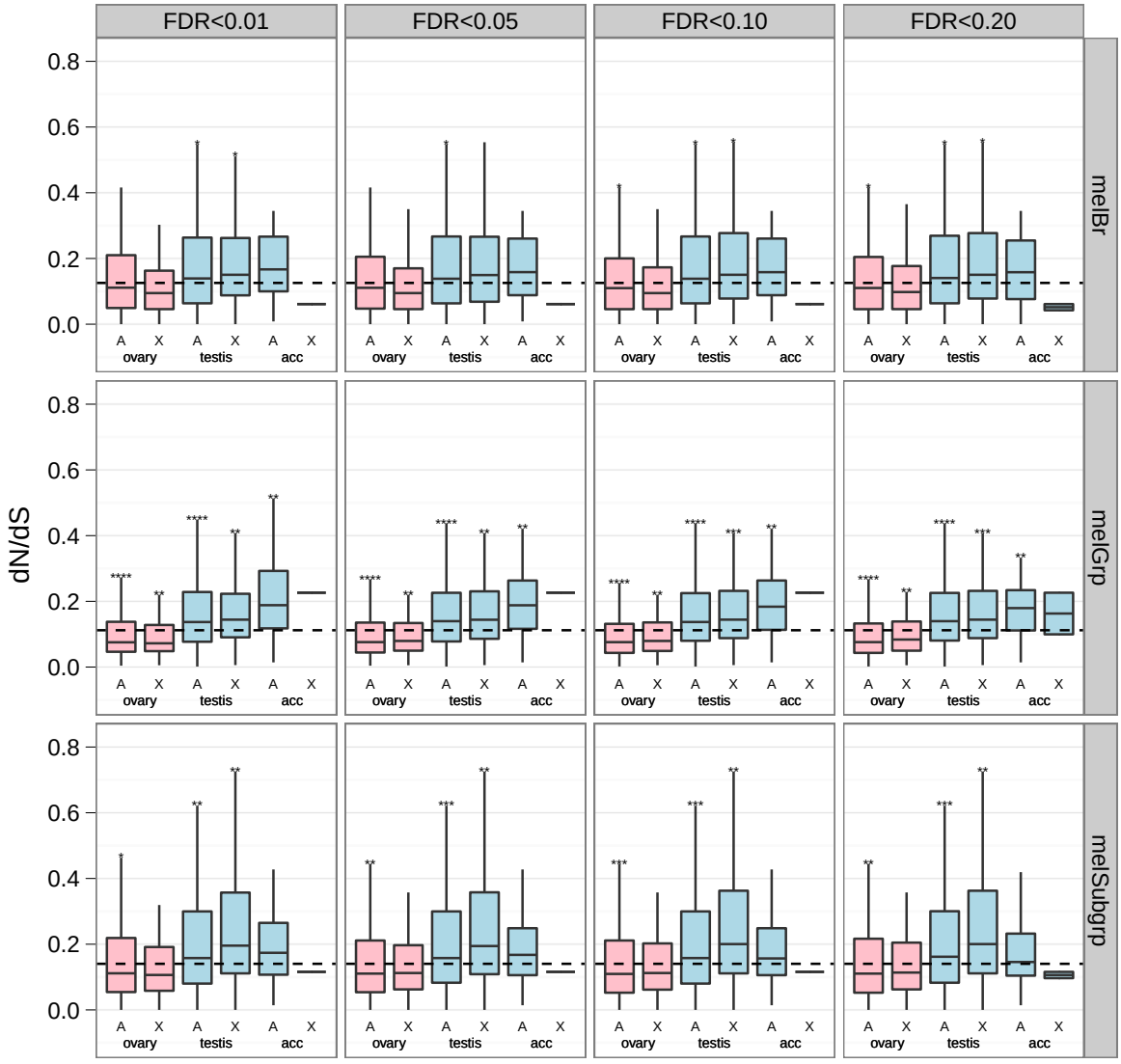
Gibson (2004), Multiple Tissues



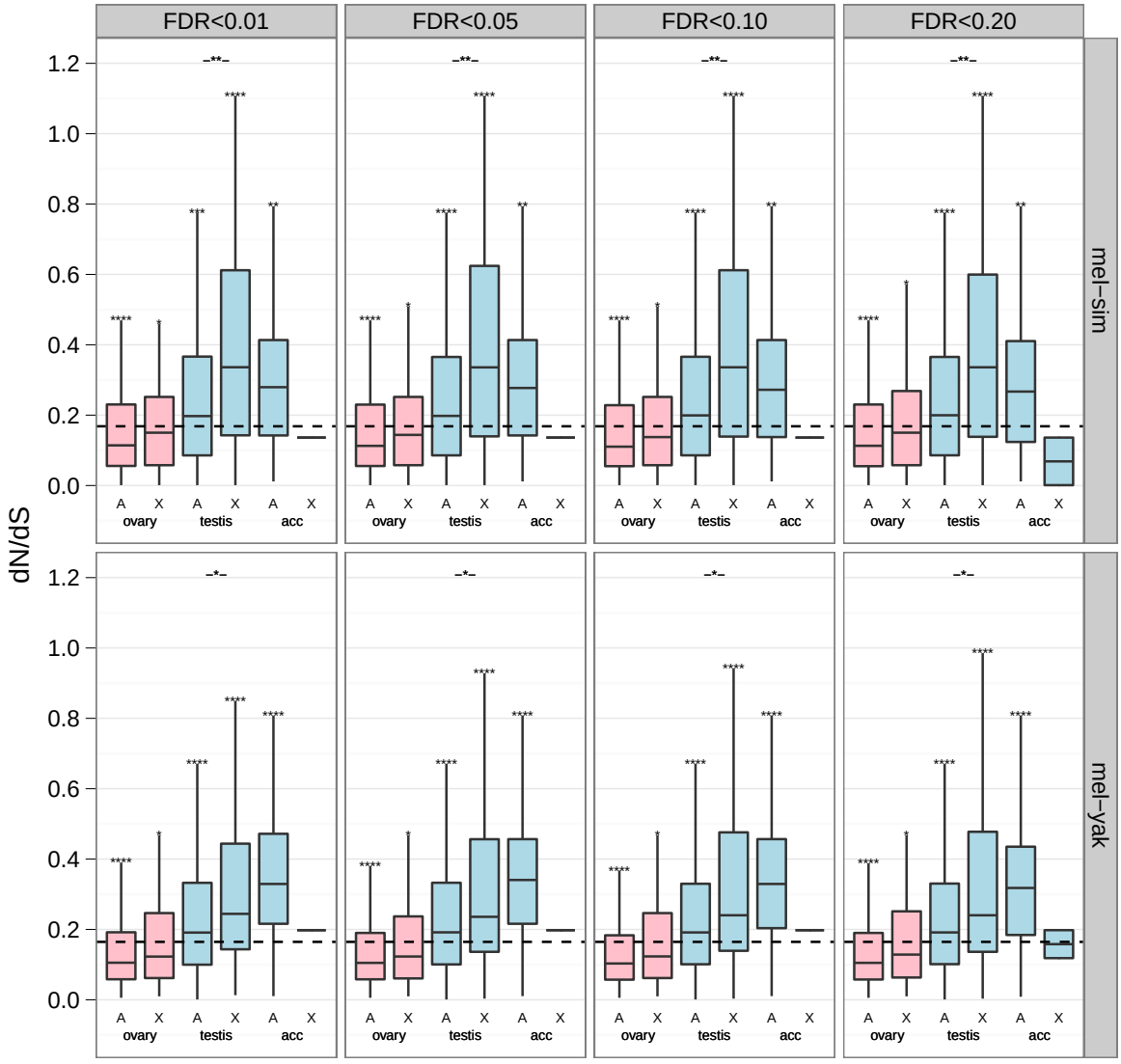
Gibson (2004), Multiple Tissues



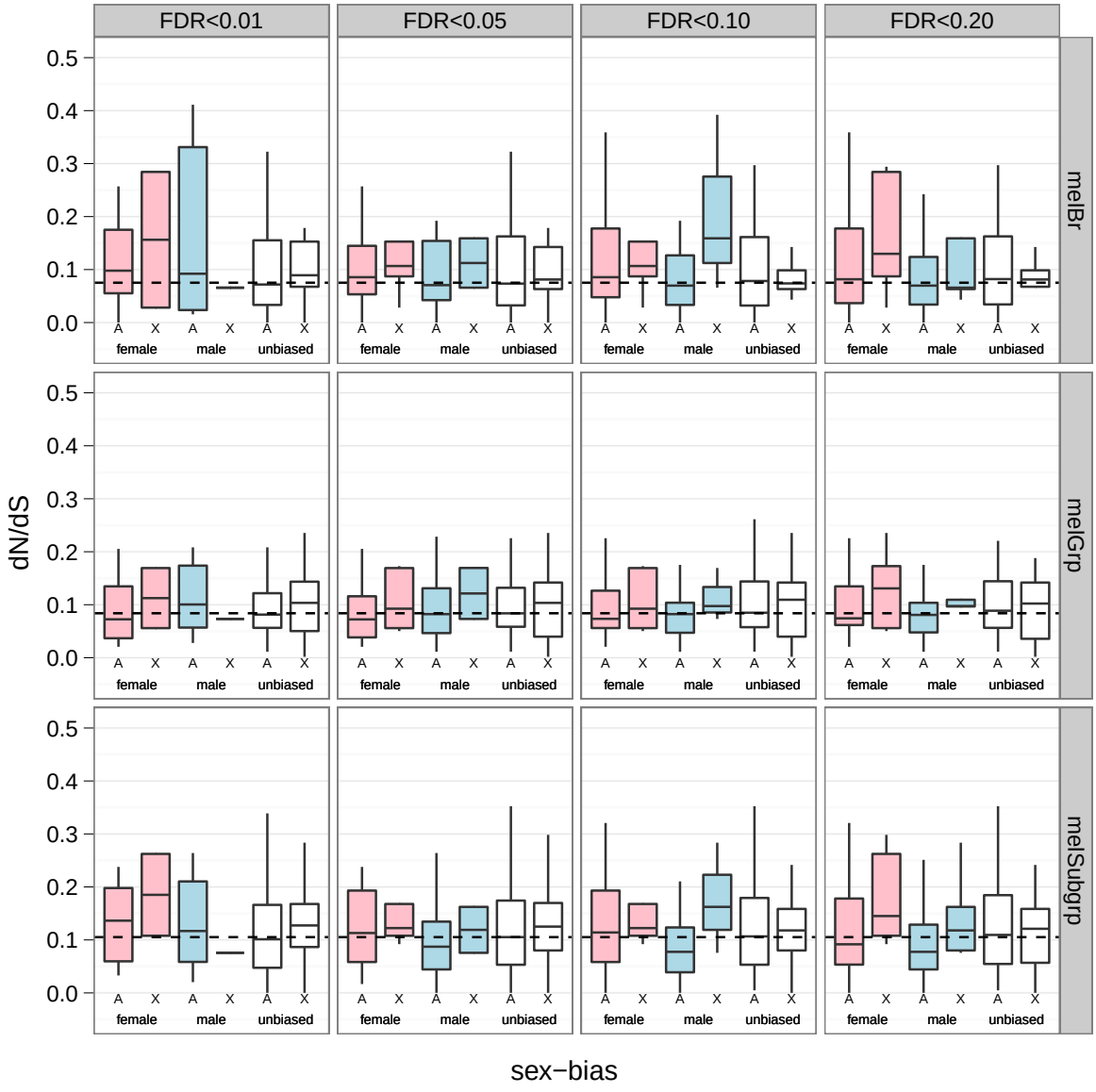
McIntyre (2006), Sex-Limited Tissue



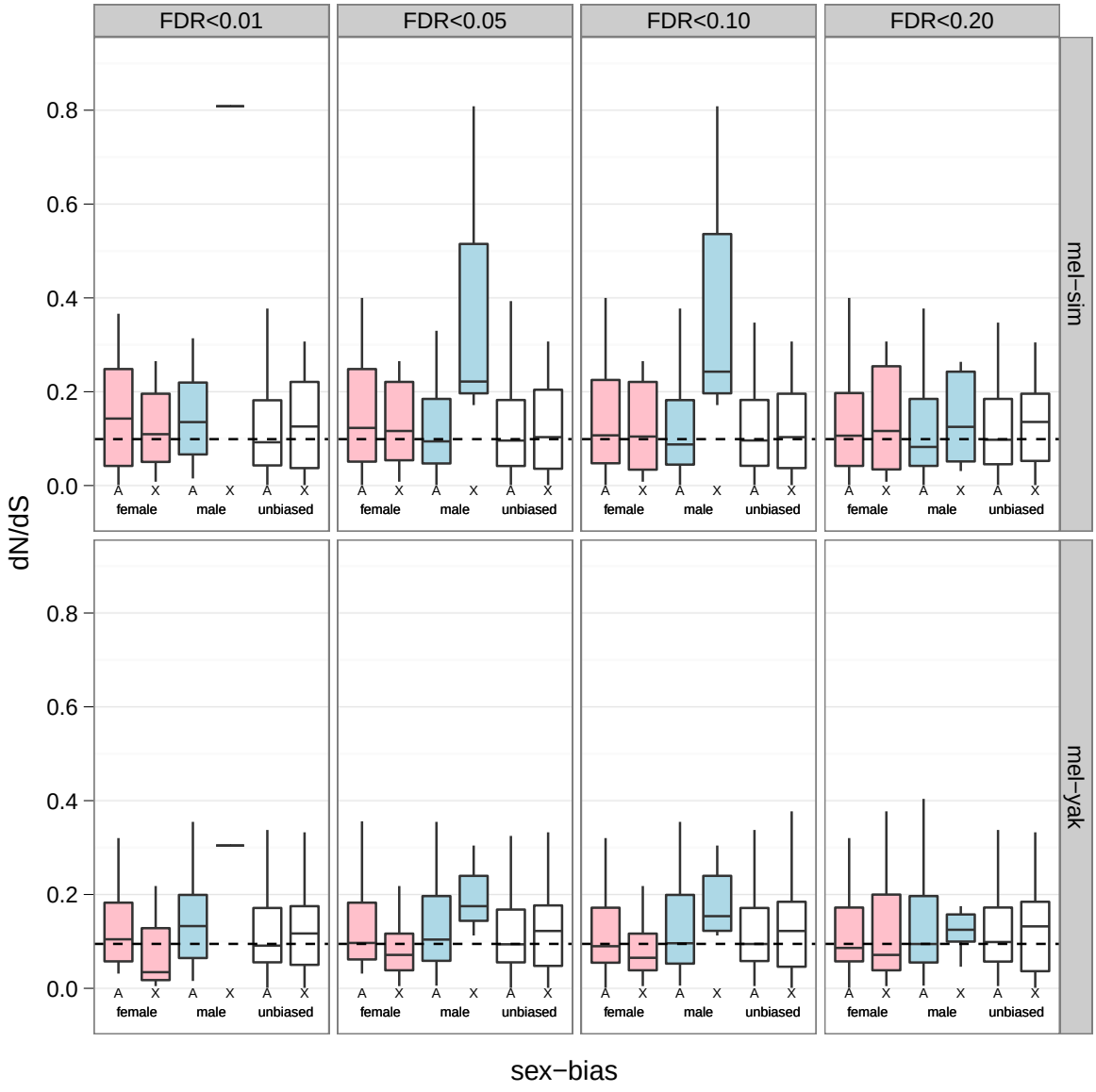
McIntyre (2006), Sex-Limited Tissue



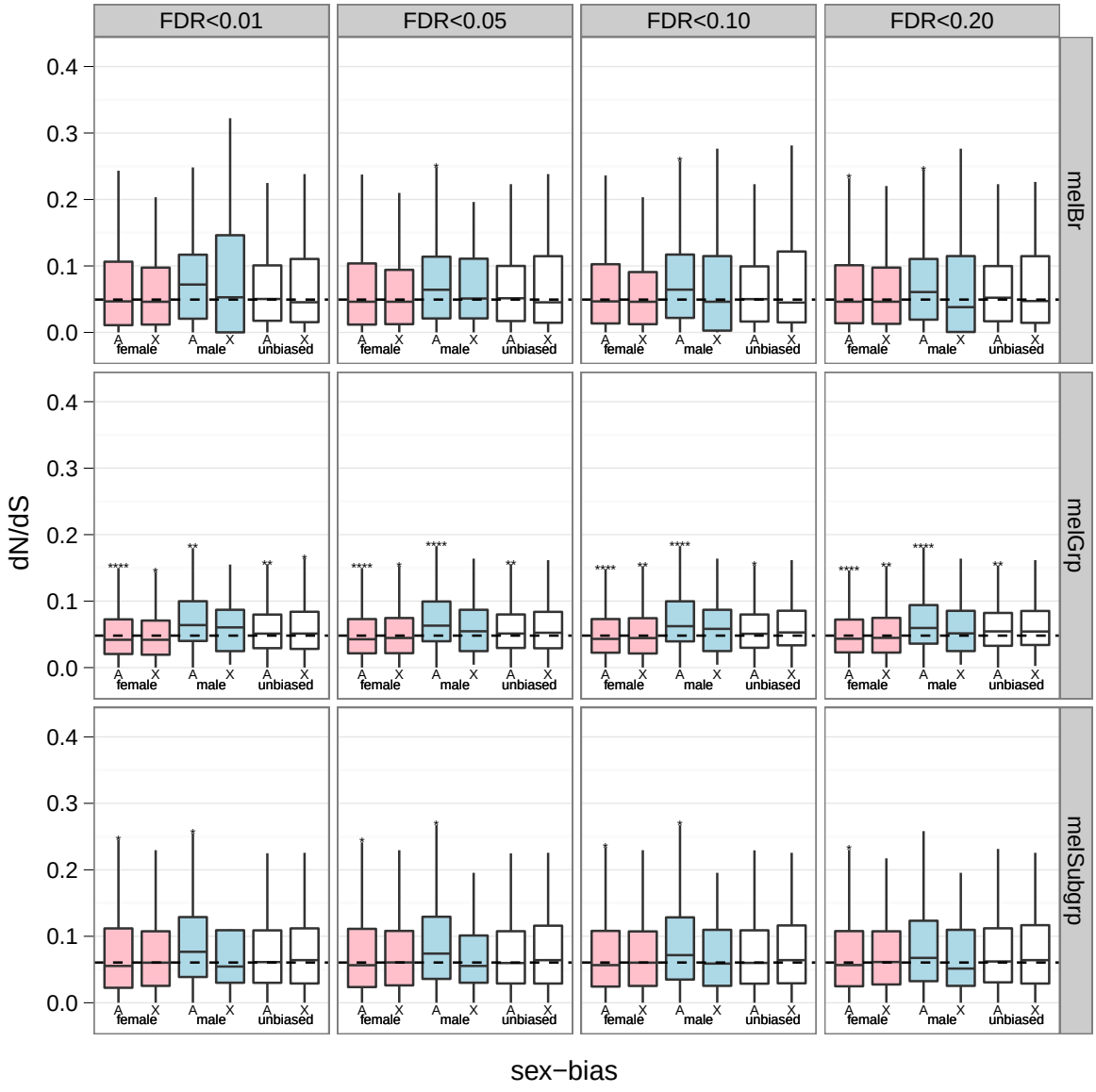
McIntyre (2006), Single Shared Somatic Tissue



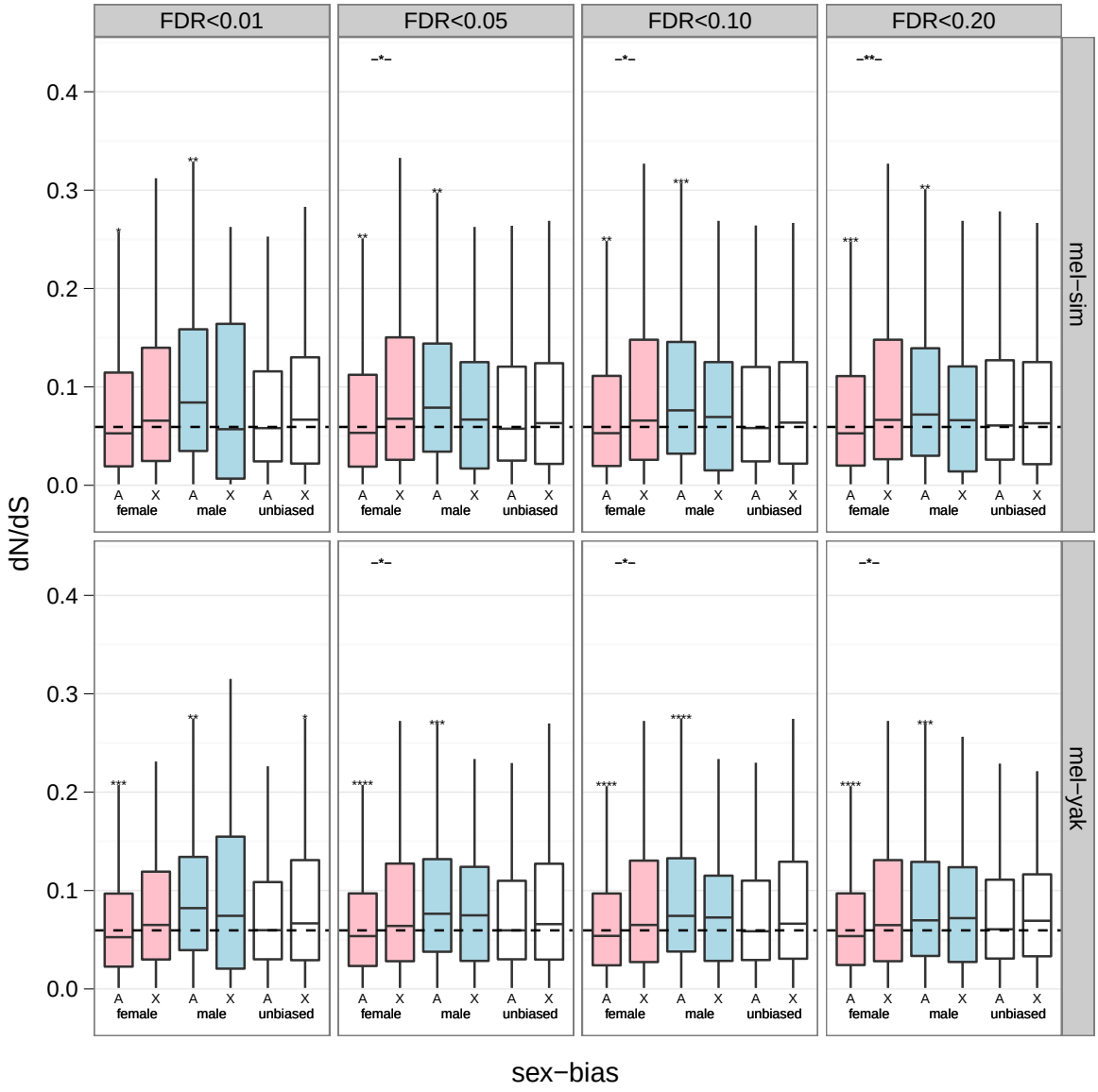
McIntyre (2006), Single Shared Somatic Tissue



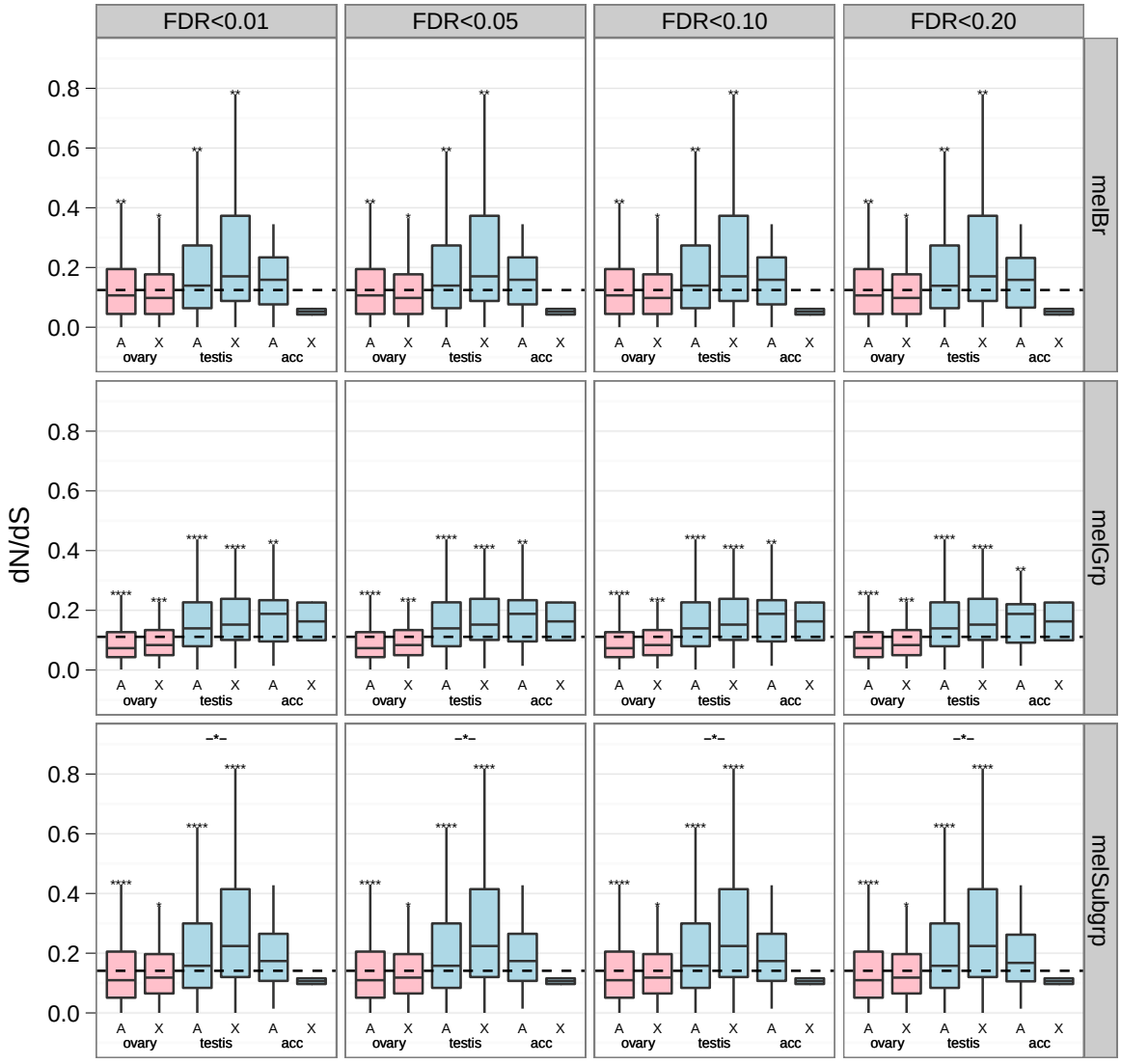
McIntyre (2006), Multiple Tissues



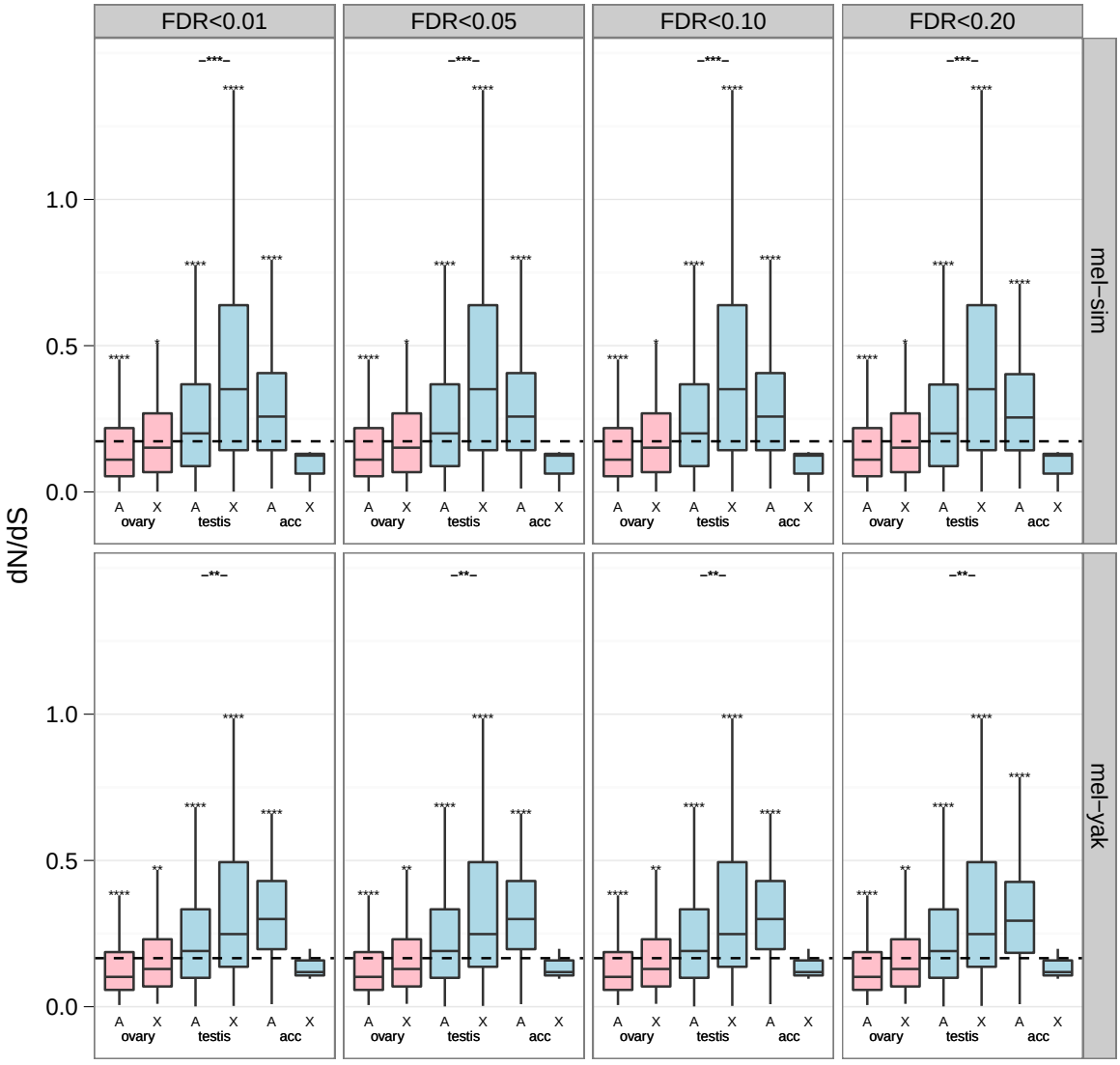
McIntyre (2006), Multiple Tissues



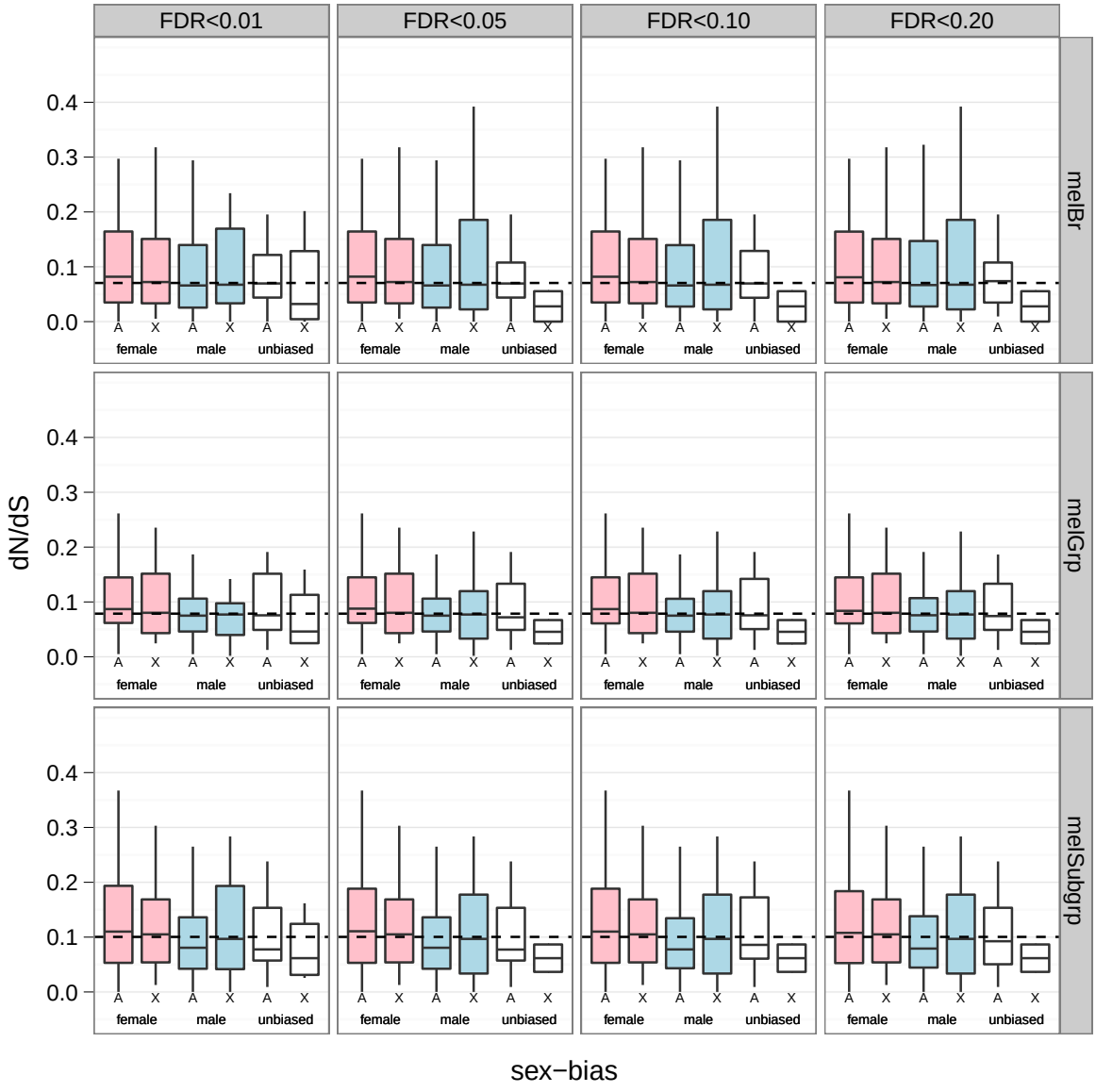
Ayroles (2006), Sex-Limited Tissue



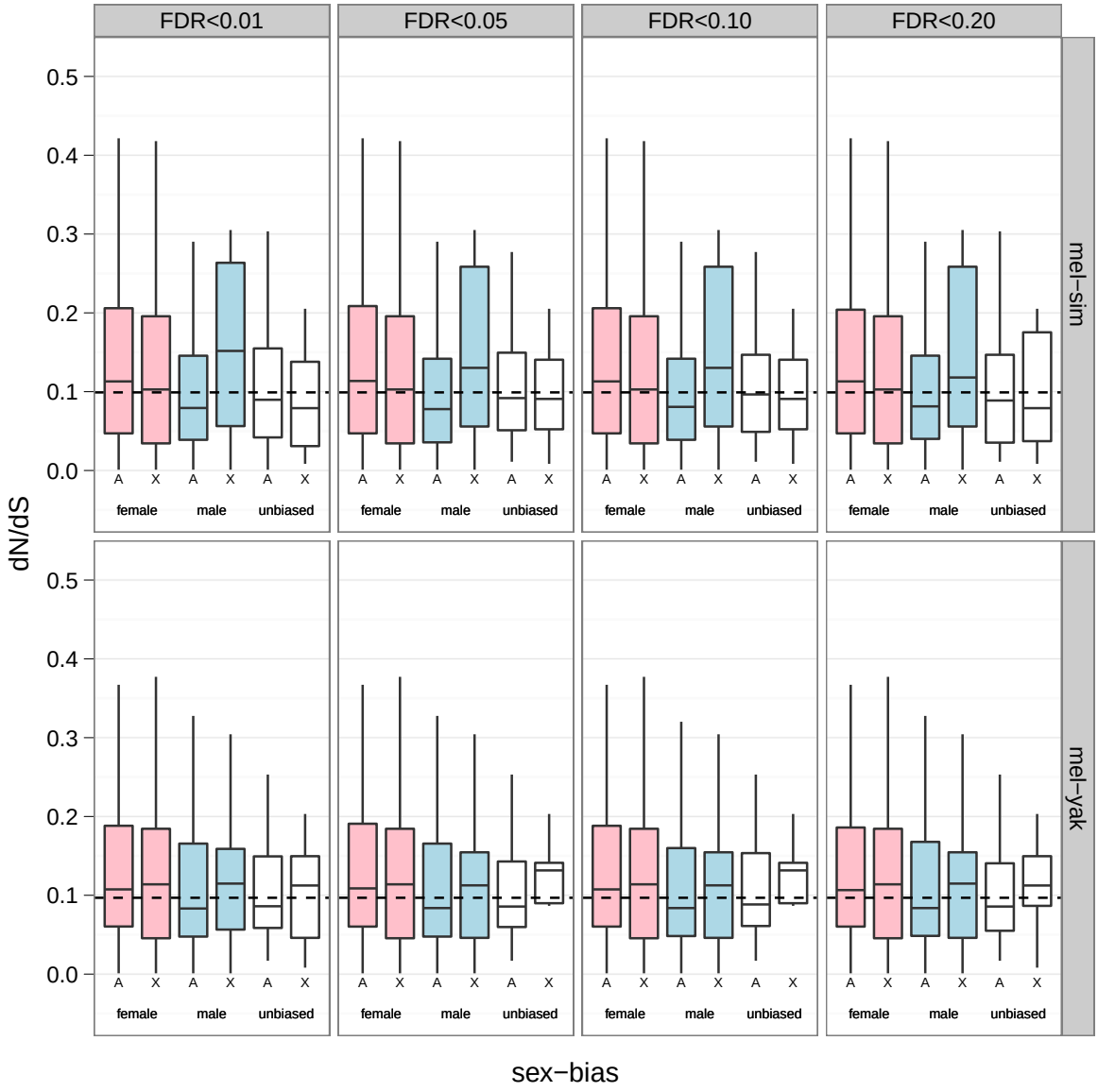
Ayroles (2009), Sex-Limited Tissue



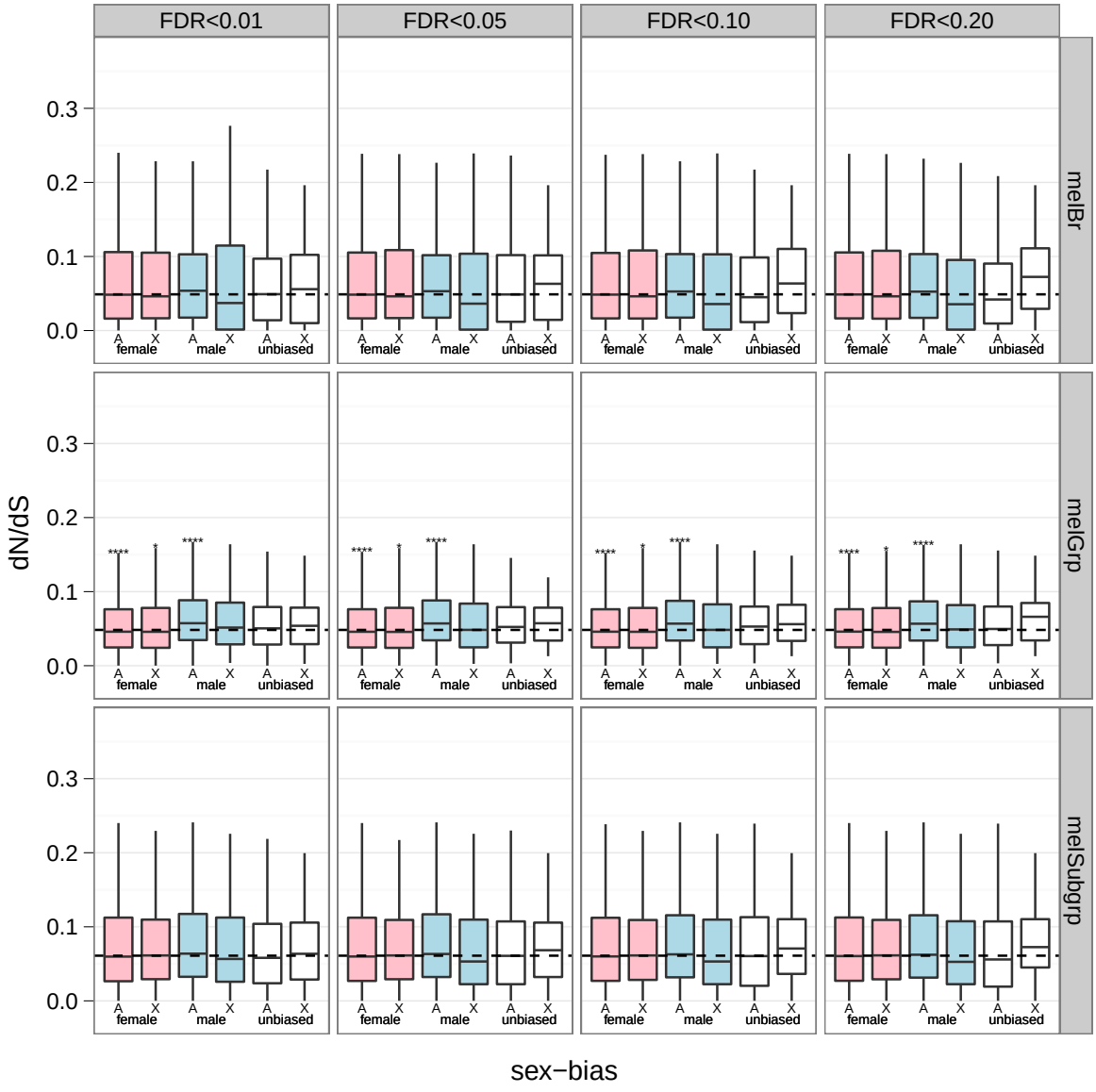
Ayroles (2009), Single Shared Somatic Tissue



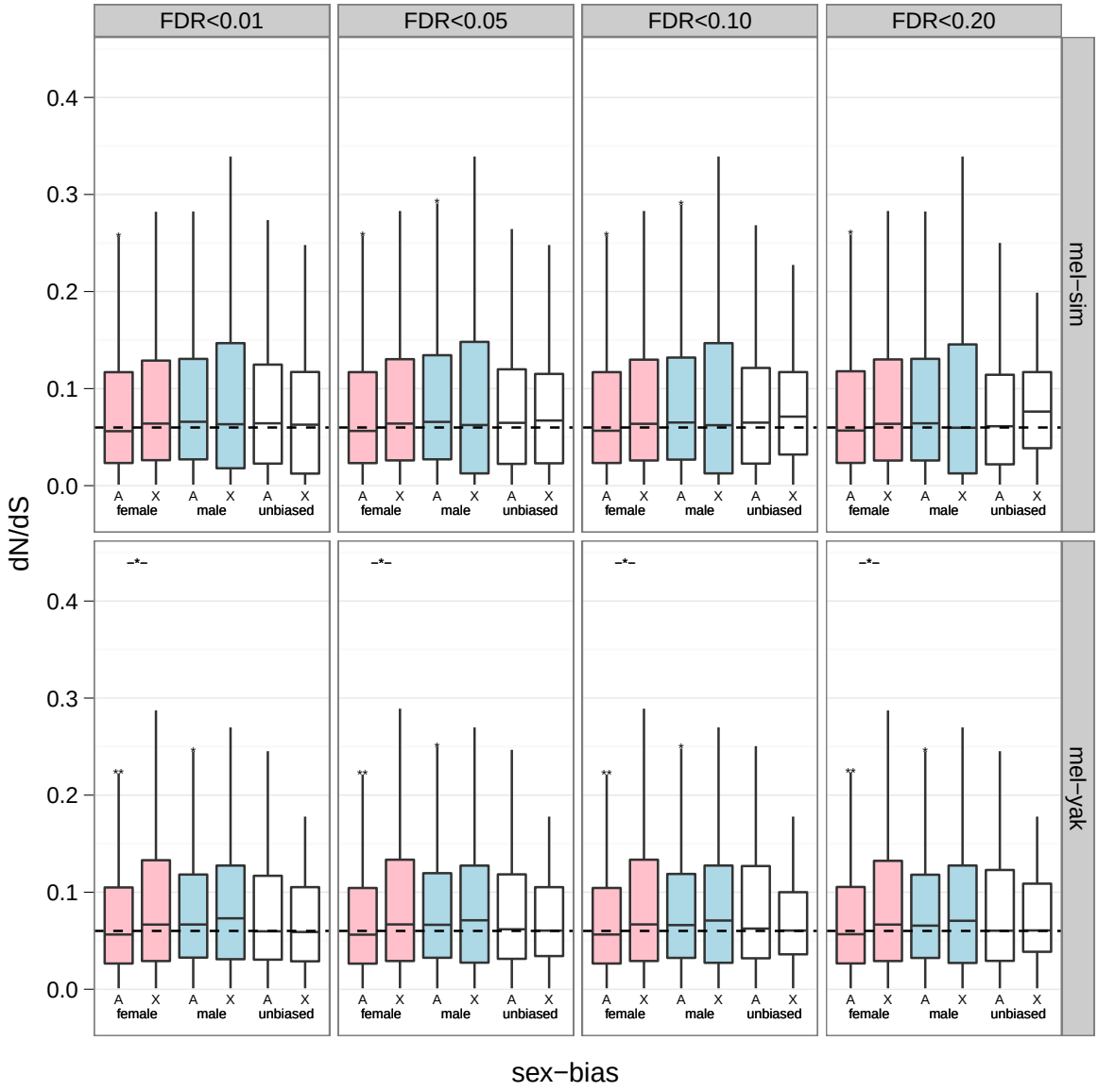
Ayroles (2009), Single Shared Somatic Tissue



Ayroles (2009), Multiple Tissues

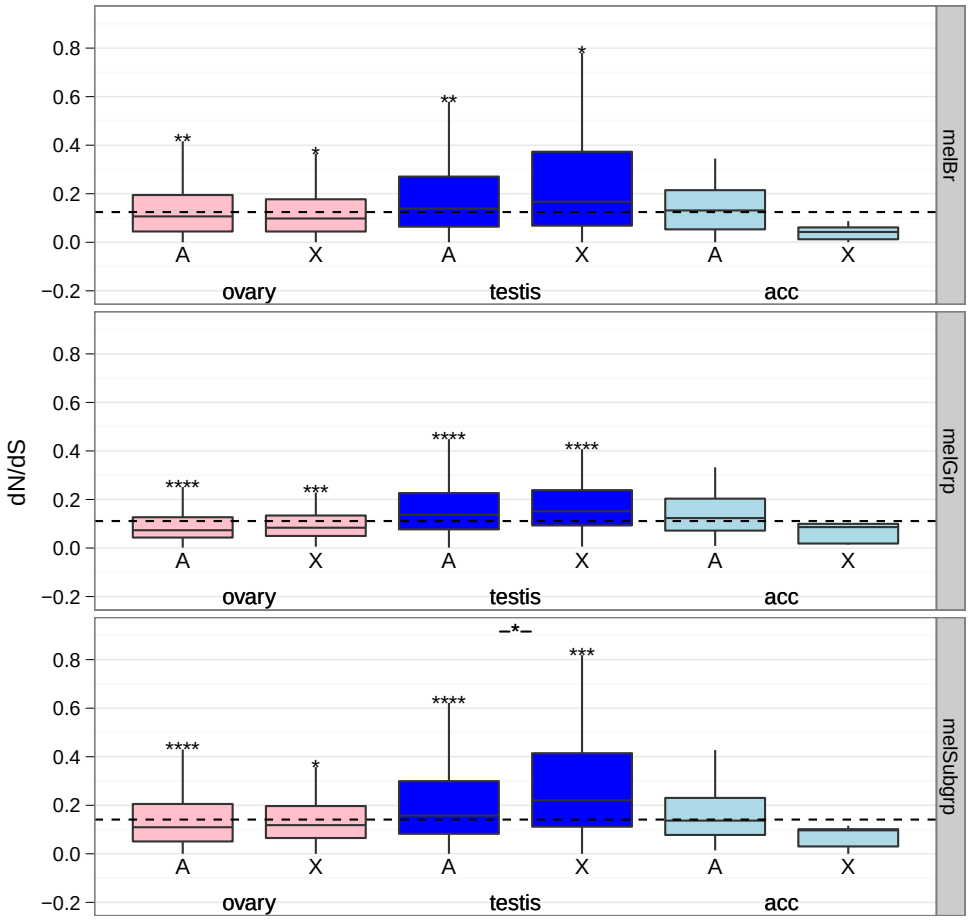


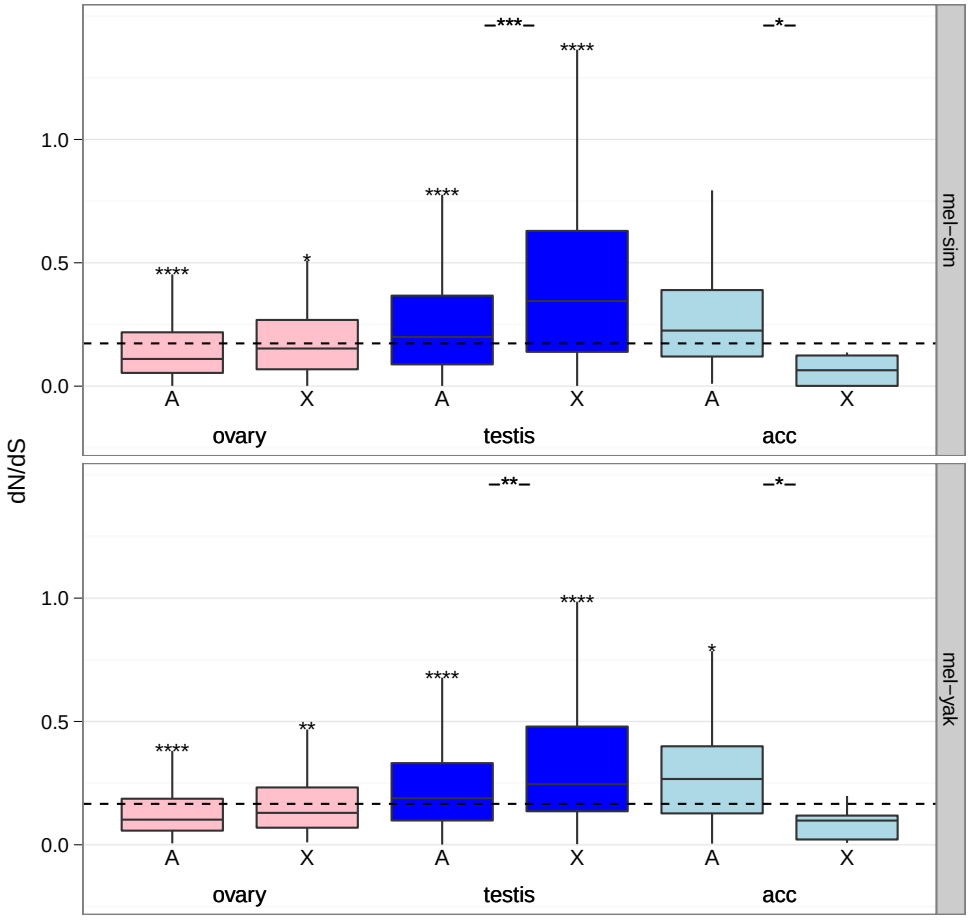
Ayroles (2009), Multiple Tissues



Faster-X evolution and expression in sex-limited tissues

Boxplots show the distribution of d_N/d_S for genes that are primarily expressed in sex-limited tissues: ovary (pink), testis (blue), and accessory gland (light blue). Autosomal (A) and X-linked (X) genes are plotted separately.Spermatheca-biased genes were excluded because of small sample sizes. Tissue-biased expression was determined using genes with signal intensity, S, > 100 in one tissue and $S < 100$ in all other tissues. The rate of protein coding evolution (d_N/d_S) was calculated five ways: along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (melBr); along all of the branches in the *melanogaster* species group (melGrp); along all of the branches in the *melanogaster* subgroup (melSubgrp); pair-wise between *D. melanogaster* and *D. simulans* (mel-sim); and pair-wise between *D. melanogaster* and *D. yakuba* (mel-yak). Dashed lines represent the median d_N/d_S for all genes that have tissue-biased expression in both sex-limited and shared somatic tissues. Estimates of d_N/d_S for each set of genes was compared to all other genes within each panel using a Mann-Whitney test, and asterisks above a box and whisker indicate a significant difference from the other genes in the panel ($*P < 10^{-2}$, $**P < 10^{-4}$, $***P < 10^{-6}$, $****P < 10^{-8}$). Additionally, autosomal and X-linked genes were compared, and asterisks flanked by dashes indicate a significant difference (-*-$P < 10^{-2}$, -**-$P < 10^{-4}$, -*** - $P < 10^{-6}$, -****-$P < 10^{-8}$).


dN/dS
0.8
0.6
0.4
0.2
0.0
−0.2
A
X
ovary
testis
acc
melBr
melGrp
melSubgrp


dN/dS
1.0
0.5
0.0
mel−sim
mel−yak
****
*
**
–***–
–**–
–*–
A
X
ovary
testis
acc

Gene ontology analysis

To determine if any particular classes of genes are responsible for the rapid evolution of testis-expressed genes, I used GOrilla (Eden et al., 2009) in an attempt to identify gene ontology (GO) categories that are over-represented amongst rapidly evolving testis-expressed genes (by ranking testis-expressed genes by d_N/d_S). This analysis was not feasible because the sample size of testis-expressed genes was too small for identifying over-represented GO terms. However, I was able to compare genes with testis-biased expression to other narrowly expressed genes using GOrilla. Testis-biased genes were identified three ways: 1) genes with male-biased expression when expression was measured in whole fly, and $S > 100$ in testis and $S < 100$ in all other adult tissues; 2) genes with $S > 100$ in testis and $S < 100$ in all other tissues, but not necessarily male-biased expression in whole fly; 3) genes with $\tau > 0.5$ and maximal S in testis, but not necessarily male-biased expression in whole fly. Each of these groups of testis-biased genes was compared to all genes, all genes expressed in a single adult tissue, and all genes expressed in a single sex-limited tissues to identify GO terms that are over-represented in the testis-biased genes. In these analyses, I continuously observed that genes involved in energy metabolism (i.e., ATP biosynthesis, glycolysis) are over-represented in the testis-biased genes.

The following table shows d_N/d_S measured along the branch leading to *D. melanogaster* after the split with the *D. simulans* lineage (MelBr), within the *melanogaster* subgroup (MelSubgrp), and within the *melanogaster* group (MelGrp) for genes with testis-biased expression (determined using $\tau > 0.5$) that are involved in ATP biosynthesis. Additionally, the fraction of other genes with lower d_N/d_S values is shown (percentile). Note that multiple genes have elevated d_N/d_S when compared to the rest of the genome.

d_N/d_S of testis-biased genes involved in ATP biosynthesis.

	d_N/d_S			percentile		
Gene	MelBr	MelSubgrp	MelGrp	MelBr	MelSubgrp	MelGrp
CG5421	0.1850	0.2094	0.2002	0.8261	0.8476	0.9171
CG7026	0.1320	0.1242	0.0808	0.7290	0.6845	0.6181
CG13167	0.0754	0.0756	0.0996	0.5473	0.4879	0.7066
CG7813	0.1406	0.1451	0.1778	0.7480	0.7402	0.8897
CG4624	0.0070	0.0498	0.0667	0.1407	0.3297	0.5204
CG32090	0.0001	0.0001	0.0565	0.0001	0.0001	0.4450
CG9013	0.2133	0.0673	0.1344	0.8552	0.4400	0.8142
CG30329	0.0258	0.0514	0.1124	0.2533	0.3406	0.7552
CG5250	0.0987	0.1925	0.2588	0.6375	0.8242	0.9571
CG6737	0.0673	0.0891	0.1118	0.5084	0.5554	0.7534
CG1076	0.0914	0.1464	0.1512	0.6138	0.7440	0.8495
CG7211	0.0703	0.1244	0.1177	0.5249	0.6851	0.7703
CG12027	0.2252	0.2908	0.1998	0.8671	0.9154	0.9166
CG32686	0.4883	0.4059	0.3153	0.9679	0.9603	0.9744
CG32089	0.1241	0.0787	0.1037	0.7096	0.5054	0.7230
CG14909	0.0001	0.0346	0.1460	0.0001	0.2288	0.8394

References

Ayroles JF, Carbone MA, Stone EA, et al. (11 co-authors). 2009. Systems genetics of complex traits in *Drosophila melanogaster*. Nat. Genet. 41:299–307.

Baines JF, Sawyer SA, Hartl DL, Parsch J. 2008. Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. Mol. Biol. Evol. 25:1639–1650.

Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. Am. Nat. 130:113–146.

Chintapalli VR, Wang J, Dow JAT. 2007. Using FlyAtlas to identify better Drosophila melanogaster models of human disease. Nat. Genet. 39:715–720.

Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the Drosophila phylogeny. Nature. 450:203–218.

Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics. 10:48.

Gibson G, Riley-Berger R, Harshman L, Kopp A, Vacha S, Nuzhdin S, Wayne M. 2004. Extensive sex-specific nonadditivity of gene expression in *Drosophila melanogaster*. Genetics. 167:1791–1799.

Gnad F, Parsch J. 2006. Sebida: a database for the functional and evolutionary analysis of genes with sex-biased expression. Bioinformatics. 22:2577–2579.

Hsiao LL, Jensen RV, Yoshida T, Clark KE, Blumenstock JE, Gullans SR. 2002. Correcting for signal saturation errors in the analysis of microarray data. Biotechniques. 32:330–332.

Larracuente AM, Sackton TB, Greenberg AJ, Wong A, Singh ND, Sturgill D, Zhang Y, Oliver B, Clark AG. 2008. Evolution of protein-coding genes in Drosophila. Trends Genet. 24:114–123.

Mank JE, Hultin-Rosenberg L, Zwahlen M, Ellegren H. 2008. Pleiotropic constraint hampers the resolution of sexual antagonism in vertebrate gene expression. Am. Nat. 171:35–43.

McIntyre LM, Bono LM, Genissel A, Westerman R, Junk D, Telonis-Scott M, Harshman L, Wayne ML, Kopp A, Nuzhdin SV. 2006. Sex-specific expression of alternative transcripts in *Drosophila*. Genome Biol. 7:R79.

Meisel RP, Han MV, Hahn MW. 2009. A complex suite of forces drives gene traffic from Drosophila X chromosomes. Genome Biol. Evol. 1:176–188.

Parisi M, Nuttall R, Naiman D, Bouffard G, Malley J, Andrews J, Eastman S, Oliver B. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. Science. 299:697–700.

Ranz JM, Castillo-Davis CI, Meiklejohn CD, Hartl DL. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. Science. 300:1742–1745.

Stolc V, Gauhar Z, Mason C, et al. (12 co-authors). 2004. A gene expression map for the euchromatic genome of Drosophila melanogaster. Science. 306:655–660.

Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res. 16:995–1004.