

## Evolution of Sex-Biased Genes

### 1. Background

Sexual dimorphism – males and females differ in physical traits and behavior. This occurs in most higher eukaryotes.

In most cases, it is the male that shows extravagant or exaggerated secondary sexual characteristics (e.g. the peacock's tail). Often these traits appear to be deleterious to the individual's survival.

Charles Darwin explained the occurrence of such traits by the theory of sexual selection, which has two components:

- a) Male-male competition – males use these traits to compete with each other for food, territory, access to females, *etc.*
- b) Female choice – females prefer to mate with males with “attractive” phenotypes.

In general, traits involved in male-reproduction (either directly or indirectly) tend to evolve fast. That is, they show extensive changes between species.

Does this also occur at the molecular level? (DNA or protein sequences)

### 2. Sex-biased genes

With the exception of a very small number of genes (such as those on the Y chromosome), males and females share the same genome. Almost all of the differences between the sexes are caused by the differential expression of genes that are present in both sexes.

Sex-biased genes are those that are expressed at a higher level in one sex than in the other. The expression difference is usually determined from male vs. female microarray or RNA-seq experiments.

How does one determine which genes are sex biased? Often a fold-change cutoff or a statistical test is used (or sometimes a combination of the two).

For example, using a 2-fold cutoff:

$M/F$  = Male/female expression ratio

$M/F > 2$  = male-biased gene (MBG)

$M/F < 0.5$  = female-biased gene (FBG)

$0.5 < M/F < 2$  = unbiased gene (UBG)

Often,  $M/F$  is presented on a  $\log_2$  scale. In this case:

$\log_2(M/F) > 1$  = MBG

$\log_2(M/F) < -1$  = FBG

$-1 < \log_2(M/F) < 1$  = UBG

Using the 2-fold cutoff for *Drosophila melanogaster* (whole flies), there are roughly: 2000 MBG, 2000 FBG, 9000 UBG expressed in adult flies (some genes are unclassified due to lack of data).

The above numbers are not exact, as they depend on the experimental conditions and the amount of replication.

Around 30% of *Drosophila* genes are sex-biased by the 2-fold criterion. An even larger fraction can be identified as sex-biased using statistical tests. For example, using a meta-analysis and a false discovery rate (FDR) of 5%, there are around 2800 MBG and 4000 FBG. With very high replication, it is possible to classify about 85% of the genes as sex-biased, although for some genes the difference in expression between sexes may only be a few percent.

In *Drosophila*, MBG tend to have a stronger expression bias than FBG:

The average M/F for MBG = 6, while the average F/M for FBG = 2.5 (using 2-fold cutoff).

In studies of whole *Drosophila*, MBG are under-represented on the X chromosome, while FBG are enriched on the X. These patterns are sometimes referred to as "demasculinization" or "feminization" of the X, respectively.

The X contains about 16% of the genes in the genome:

11% of MBG are on the X

21% of FBG are on the X

16% of UBG are on the X

Why are MBG under-represented on the X? Some ideas:

a) Dosage: males have only 1 copy of the X, while females have 2 copies. Thus, one might expect X-linked genes to have twice as much expression in females as in males. However, most species have mechanisms of dosage compensation to equalize expression between the sexes. In *Drosophila*, dosage compensation occurs in somatic tissues by doubling the transcription of the male X. Dosage compensation does not appear to occur in the *Drosophila* male germline, which may explain why an under-representation of male-biased genes on the X is seen in studies using whole flies or samples that include reproductive organs. In somatic tissues, variation in the extent of dosage compensation along the X chromosome may also contribute to sex-biased expression. In many ZW taxa, there does not appear to be Z-chromosome dosage compensation in females, leading to an excess of MBG on the X.

b) X-inactivation (or X suppression): in *Drosophila* and many other species, including mammals, it appears that the X chromosome becomes transcriptionally inactive in the male germline. Thus, the expression of MBG is suppressed during spermatogenesis if they are X-linked. This could lead to an under-representation of MBG on the X if many MBG are expressed in testes/sperm.

c) Sexual conflict: expression of a MBG may have a deleterious effect on females. Since the X spends twice as much time in females as it does in males, the female interests will outweigh the male interests. This may lead to reduced expression of X-linked MBG.

In general, the X appears to be a bad environment for MBG and there may be selective pressure for MBG to "escape" the X and move to an autosome by retrotransposition. It has been observed that there is an excess of *Drosophila* retrotransposed genes that have moved from the X to the autosomes. Most of these are expressed in testes.

### 3. Rates of evolution

By combining expression data from microarrays and comparative genomic data, one can compare rates of evolution of sex-biased genes.

The rate of evolution is usually measured by  $K_a$  or by  $K_a/K_s$  (also known as  $dN/dS$ ). The latter controls for possible differences in mutation rates.

In *Drosophila*, MBG show greater divergence between species than FBG and UBG. The biggest difference is at nonsynonymous sites, indicating a faster rate of protein evolution in MBG. FBG appear to evolve slightly faster than UBG. The strongest difference is seen for genes expressed in reproductive organs.

In a comparison of the human and chimpanzee genomes, it was also found that genes expressed in testes evolve faster than those expressed in other tissues.

### 4. Why do male-biased genes evolve fast?

There are two possibilities:

a) MBG genes are under less selective constraint. They can accumulate many neutral amino acid changes that have no effect on fitness.

b) MBG are subject to more positive (or sexual) selection. They accumulate more adaptive amino acid replacements.

How can we test the two hypotheses?

#### 1) Correlation between evolutionary rate and recombination rate

Selection is more efficient in regions of high recombination, because each mutation can be selected independently. If changes are due to positive selection, then we should see a positive correlation between divergence and recombination rate.

This is what is seen for MBG in *Drosophila*.

FBG and UBG show a slightly negative correlation, suggesting that most amino acid changes in these genes are deleterious and are removed by purifying selection.

#### 2) Polymorphism within species

Hypothesis (a) from above predicts that MBG should have more amino acid polymorphism within species than other genes and that it should be proportional to the amount of divergence between species.

This can be tested with the McDonald-Kreitman (MK) test, which compares the ratio of divergence at nonsynonymous and synonymous sites ( $D_n/D_s$ ) to polymorphism at nonsynonymous and synonymous sites ( $P_n/P_s$ ).

If the ratios are equal (do not differ significantly) -> neutral evolution

If there is an excess of nonsynonymous divergence -> positive selection

If there is an excess of nonsynonymous polymorphism -> balancing selection or weak purifying selection

For *Drosophila* MBG,  $(Dn/Ds)/(Pn/Ps) \approx 2$ .

For FBG it is  $\approx 1.2$

For UBG it is  $\approx 0.9$

These results support hypothesis (b). MBG tend to show an excess of nonsynonymous divergence. FBG also show some evidence for positive selection when compared to UBG.

X-linked MBG show a stronger signal of positive selection than autosomal MBG.

For X-linked MBG,  $(Dn/Ds)/(Pn/Ps) \approx 3.4$ .

X-linked MBG also show the highest divergence (measured by  $Ka/Ks$ ).

These observations can be explained if beneficial mutations in MBG are often recessive. When X-linked, these mutations can be instantly “seen” by selection in hemizygous males. Recessive autosomal mutations can only be selected after they are in high enough frequency to be homozygous.

#### Additional reading:

Ellegren, H., and J. Parsch (2007) The evolution of sex-biased genes and sex-biased gene expression. *Nature Reviews Genetics* 8: 689-698.

Parsch, J., and H. Ellegren (2013) The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews Genetics* 14: 83-87.

Grath, S., and J. Parsch (2016) Sex-biased gene expression. *Annual Review of Genetics* 50: 29-44.