The use of genomic and gene expression large-scale data for the analyses of sexual evolution

Maria D. Vibranovski

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE BIOCIÊNCIAS
Departamento de Genética e Biologia Evolutiva

15 de abril de 2015 - UFRJ
Acknowledgments

Manyuan Long  
Department of Ecology and Evolution  
The University of Chicago

Timothy L. Karr  
Drosophila Genetic Resource Center  
Kyoto Institute of Technology

Hedibert F. Lopes  
INSPER – São Paulo
Outline

1. Background in Sexual Evolution
2. Biological Problem
3. Large-scale Data
4. Statistical Approaches
5. Ongoing Biological Problems
6. Perspectives
1. Background in Sexual Evolution

Why are there phenotypic differences between sexes?
1. Background in Sexual Evolution

Sex Chromosomes

Male

Female

X
Y

X
X
Gene activity = gene expression = number of mRNA molecules
1. Background in Sexual Evolution

Sex-biased Gene Expression

GENE A

GENE B
2. Biological Problem

What molecular mechanisms and genetics processes are involved?

XY systems

Testis bias
2. Biological Problem

Meiotic Sex Chromosome Inactivation (MSCI)

Xist (X inactive-specific transcripts)

Mouse primary spermatocytes

2. Biological Problem

Time in Spermatogenesis of *Drosophila melanogaster*

- **Spermatogonia (mitotic)**
- **Spermatocytes (meiotic)**

General cell types development

Chromosomes

Gene expression:
- **active**
- **inactive**

Chromosomal Gene movement

Lifschytz and Lindsley, 1972
2. Biological Problem

- Does MSCI exist in *Drosophila*?
- Is MSCI involved in the distribution of sex-biased genes in the genome?

**XY systems**

**Testis bias**
3. Large-scale Data

*Drosophila melanogaster* Spermatogenesis

Clark et al., 2002
Isolation of Spermatogenic Cells

DNA
Cytoplasm
Post-meiotic cells
Mitotic cells
Meiotic cells

Vibranovski et al, 2009, PLoS Genetics
3. Large-scale Data

Tissue Isolation (n=3)

Vibranovski et al, 2009, PLoS Genetics
How to count the number of mRNA molecules for each gene in a cell?

Microarray: A chip containing small pieces of DNA corresponding to all *Drosophila* genes
Microarray

Specific case for gene expression measure: the GeneChip array
3. Large-scale Data

Tissue Isolation (n=3)

RNA extraction

Microarray Hybridization

Gene profile

Vibranovski et al, 2009, PLoS Genetics
4. Statistical Approaches

Meiotic Sex Chromosome Inactivation (MSCI)

Vibranovski et al, 2009, PLoS Genetics
X inactivation occurs when the difference in activity (meiosis-mitosis) in X is lower than the difference in activity in autosome.

\[ \theta = \frac{\text{gene expression}}{\text{activity}} \]

\[ \theta^X_{\text{meiosis}} - \theta^X_{\text{mitosis}} < \theta^A_{\text{meiosis}} - \theta^A_{\text{mitosis}} \]
4. Statistical Approaches

Gene intensity
2982 genes (X chromosomes)
15099 genes (autosomes)

Spermatogenic replicates
3 mitotic
3 meiotic
3 post-meiotic

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mit1</th>
<th>Mit2</th>
<th>Mit3</th>
<th>Mei1</th>
<th>Mei2</th>
<th>Mei3</th>
</tr>
</thead>
<tbody>
<tr>
<td>792</td>
<td>5.273</td>
<td>5.357</td>
<td>5.509</td>
<td>5.548</td>
<td>5.577</td>
<td>5.587</td>
</tr>
<tr>
<td>1966</td>
<td>5.924</td>
<td>5.914</td>
<td>5.945</td>
<td>6.297</td>
<td>6.779</td>
<td>6.643</td>
</tr>
<tr>
<td>2811</td>
<td>5.177</td>
<td>5.197</td>
<td>5.168</td>
<td>5.092</td>
<td>5.166</td>
<td>5.169</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Mit1</th>
<th>Mit2</th>
<th>Mit3</th>
<th>Mei1</th>
<th>Mei2</th>
<th>Mei3</th>
</tr>
</thead>
<tbody>
<tr>
<td>10541</td>
<td>4.932</td>
<td>4.932</td>
<td>4.969</td>
<td>5.073</td>
<td>4.876</td>
<td>4.986</td>
</tr>
</tbody>
</table>
For chromosome $i$ (X or A) and gene $l$, 3 replicates are measured

**Mitosis:** $mit_{i1l}, mit_{i2l}, mit_{i3l}$

$$E(mit_{ikl}) = \theta_{il}^{mit}$$

**Meiosis:** $mei_{i1l}, mei_{i2l}, mei_{i3l}$

$$E(mei_{ikl}) = \theta_{il}^{mei}$$

The objective is to learn whether genes are

**Differently expressed:**

$$H_{1,il} : \theta_{il}^{mei} > \theta_{il}^{mit} \quad \text{(OVER)}$$

$$H_{2,il} : \theta_{il}^{mei} < \theta_{il}^{mit} \quad \text{(UNDER)}$$

or

**Equally expressed:**

$$H_{3,il} : \theta_{il}^{mei} = \theta_{il}^{mit} \quad \text{(EQUAL)}.$$
How do biologists usually approach this problem?

**APPROACH 1:**
Identify differently expressed genes based on 2 fold intensity differences (based on homotypic experiments);

**APPROACH 2:**
Identify differently expressed genes by controlling type I error;

**APPROACH 3:**
Identify differently expressed genes by controlling false discovery rates;

**APPROACH 4:**
Hierarchical Bayes.

How do biologists usually approach this problem?

**APPROACH 1, 2 or 3**
Identify differently expressed genes

Counts

<table>
<thead>
<tr>
<th>Expression</th>
<th>ChrX</th>
<th>ChrA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Under expressed</td>
<td>537(18%)</td>
<td>2416(16%)</td>
</tr>
<tr>
<td>Equally expressed</td>
<td>2147(72%)</td>
<td>10720(71%)</td>
</tr>
<tr>
<td>Over expressed</td>
<td>298(10%)</td>
<td>1963(13%)</td>
</tr>
<tr>
<td>Total</td>
<td>2982</td>
<td>15099</td>
</tr>
</tbody>
</table>
4. Statistical Approaches

![Gene expression diagram](image)

**Proportion of genes (%)**

<table>
<thead>
<tr>
<th>D</th>
<th>Two Fold Change Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVER</td>
<td>Meiosis &gt;&gt; Mitosis</td>
</tr>
<tr>
<td>UNDER</td>
<td></td>
</tr>
</tbody>
</table>

![Bar chart](chart)
4. Statistical Approaches

**Approach 2: False positive rate (FPR)**

<table>
<thead>
<tr>
<th></th>
<th>(E) Evidence against H</th>
<th>(NE) No evidence against H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Over/under expressed</td>
<td>$\tau$</td>
<td>$\eta_1 - \tau$</td>
<td>$\eta_1$</td>
</tr>
<tr>
<td>(H) Equally expressed</td>
<td>$\eta$</td>
<td>$\eta_0 - \eta$</td>
<td>$\eta_0$</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>$x$</td>
<td>$n-x$</td>
<td>$n$</td>
</tr>
</tbody>
</table>

Measures error rate (evidence against H) when H is true: $E(\eta/\eta_0)$

FPR = 0.05 => $E(\eta) < 0.05E(\eta_0)$

**Bonferroni correction**
When evidence is measure by $\{p\text{-value}<0.05/n\}$, then

$Pr(\eta > 0) < 0.05$

Controlling $E(\eta) < 0.05n$ is too liberal!
Controlling $Pr(\eta > 0) < 0.05$ is too conservative!
4. Statistical Approaches

**Approach 3: False discovery rate (FDR)**

<table>
<thead>
<tr>
<th></th>
<th>(E) Evidence against H</th>
<th>(NE) No evidence against H</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Over/under expressed</td>
<td>τ</td>
<td>𝜂_1 - 𝜏</td>
<td>𝜂_1</td>
</tr>
<tr>
<td>(H) Equally expressed</td>
<td>𝜂</td>
<td>𝜂_0 - 𝜂</td>
<td>𝜂_0</td>
</tr>
<tr>
<td>Total</td>
<td>𝑥</td>
<td>𝑛 - 𝑥</td>
<td>𝑛</td>
</tr>
</tbody>
</table>

\[
FDR(\alpha) = E\left( \frac{\eta(\alpha)}{x(\alpha)} \right) \approx \frac{E(\eta(\alpha))}{E(x(\alpha))} = \frac{\eta_0 \alpha}{E(x(\alpha))} \approx \frac{n \hat{\pi}_0 \alpha}{\sum_{i=1}^{n} 1(p_i < \alpha)}
\]

\[
\hat{\pi}_0 = \lim_{i \to 1} \frac{\sum_{i=1}^{n} 1(p_i > \lambda)}{n(1 - \lambda)} \quad \text{(estimate of true nulls)}
\]

\[
FDR(\alpha) \approx \frac{\pi_0 \Pr(E \mid H)}{\Pr(E)} = \Pr(H \mid E) \quad \text{(Looks Bayesian!!)}
\]
4. Statistical Approaches

Approach 4: \textit{mixture} model for differences

Because $\sigma_\bar{x}^2$ and $\sigma_A^2$ (within gene variability) in the hierarchical model are negligible when compared $\tau^2$ (between genes variability), the next model we implemented is a mixture of normals model that accounts for the excess of intensities around zero:

Model:

$$d_l = (\bar{x}^{\text{mei}}_l - \bar{x}^{\text{mit}}_l) \sim \pi N(\theta_1, \tau_1^2) + (1-\pi) N(\theta_2, \tau_2^2) \quad l = 1,\ldots,n$$

Prior:

$$p(\pi, \theta_1, \theta_2, \tau_1^2, \tau_2^2) \propto \tau_1^{-2}\tau_2^{-2}$$
4. Statistical Approaches

Approach 4: **mixture** model for differences

- Bayesian Mixture Model classifies genes as over or under expressed
  - Detect differential gene expression while comparing chromosomes distributions
  - Avoids the arbitrariness of folding
  - Avoids multiple testing distortions

4. Statistical Approaches

Approach 4: mixture model for differences

\[ p(d \mid data) = \int \left( \pi f_N(d; \theta_1, \tau_1^2) + (1 - \pi) f_N(d; \theta_2, \tau_2^2) \right) p(\pi, \theta_1, \theta_2, \tau_1^2, \tau_2^2 \mid data) \, d\pi \, d\theta_1 \, d\theta_2 \, d\tau_1^2 \, d\tau_2^2 \]
Approach 4: *mixture* model for differences

Classification (Prior): \( Z_l \sim Ber(\pi) \)

Classification (Posterior):
- \( (Z_l = 1) \& (d_l > 0) \) – over expressed
- \( (Z_l = 1) \& (d_l < 0) \) – under expressed
- \( (Z_l = 0) \) – equally expressed
4. Statistical Approaches

Approach 4: Bayesian **mixture** model for differences

![Graph showing gene expression during mitosis and meiosis with X inactivation](image)

<table>
<thead>
<tr>
<th>C</th>
<th>Bayesian Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>OVER</td>
<td>UNDER</td>
</tr>
<tr>
<td>Meiosis &gt;&gt; Mitosis</td>
<td>Meiosis &lt;&lt; Mitosis</td>
</tr>
</tbody>
</table>

![Bar chart showing proportion of genes (%)](chart)
4. Statistical Approaches

C | Bayesian Model
---|---
| OVER | UNDER | EQUAL |
| Meiosis >> Mitosis | Meiosis << Mitosis | Meiosis ≈ Mitosis |

D | Two Fold Change Method
---|---
| OVER | UNDER | EQUAL |
| Meiosis >> Mitosis | Meiosis << Mitosis | Meiosis ≈ Mitosis |

![Bar charts showing proportion of genes (%)]
4. Statistical Approaches

Meiotic Sex Chromosome Inactivation (MSCI)

VibranoVski et al, 2009, PLoS Genetics
4. Statistical Approaches

Meiotic Sex Chromosome Inactivation (MCSI)

Vibranski et al, 2009, PLoS Genetics
2. Biological Problem

Expression Intensity

Mitosis  Meiosis  Spermatogenesis

Autosomal genes

X-linked genes

Vibranski et al, 2009, PLoS Genetics
2. Biological Problem

- Does MSCI exist in Drosophila? YES
- Is MSCI involved in the distribution of sex-biased genes in the genome?

XY systems

Testis bias
2. Biological Problem

Out-of-the-X retroposition

Parental gene and retrogene transcription levels during distinct stages of spermatogenesis

- Parental gene mRNA
- Retrogene mRNA

Intensity of sex-chromosome inactivation during and after meiosis

Kaessmann et al, 2009, Nat Rev Genetics
2. Biological Problem

Complementary expression

**Tom40**: Translocase of outer membrane 40 / Chr X
**Tomboy40**: protein transmembrane transporter activity / Chr 2R

Vibranovski et al, 2009, PLoS Genetics
2. Biological Problem

Complementary expression

No Complementary expression

Expression Intensity

Mitosis    Meiosis

Parental gene

Retrogene gene

Mitosis    Meiosis
4. Statistical Approaches

Out of X

[Graph showing data points for gene measurements, with gene numbers from 1 to 30 along the x-axis and measures from 4 to 14 along the y-axis.]
4. Statistical Approaches

Out of A

gene

measures

4 6 8 10 12 14

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31

gene

measures

4 6 8 10 12 14

32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62
4. Statistical Approaches

Approach 1: Counting YESES and NOS

<table>
<thead>
<tr>
<th>Counts</th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out of X</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>Out of A</td>
<td>28</td>
<td>34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Proportions</th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out of X</td>
<td>0.31</td>
<td>0.69</td>
</tr>
<tr>
<td>Out of A</td>
<td>0.45</td>
<td>0.55</td>
</tr>
</tbody>
</table>

**Fisher's Exact Test for Count Data**

Null hypothesis: true odds ratio = 1

Alternative hypothesis: true odds ratio is not equal to 1

\[ p\text{-value} = 0.2546 \]

Estimated odds ratio=0.55 (roughly \( \frac{34}{28}/\frac{20}{9} = 0.546 \))

95% confidence interval: \( (0.1888,1.5086) \)

95% confidence interval for YES (normal approximation)
Out of X: \( (0.518,0.862) \)
Out of A: \( (0.424,0.676) \)
## Approach 2: Controling for FDR

Considering only significant change of expression 2 fold difference \( p<0.05, q < 0.05 \)

<table>
<thead>
<tr>
<th></th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out of X</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Out of A</td>
<td>37</td>
<td>25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>NO</th>
<th>YES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out of X</td>
<td>0.33</td>
<td>0.66</td>
</tr>
<tr>
<td>Out of A</td>
<td>0.60</td>
<td>0.40</td>
</tr>
</tbody>
</table>

### Fisher's Exact Test for Count Data

Null hypothesis: true odds ratio = 1

Alternative hypothesis: true odds ratio is not equal to 1

\[ p\text{-value} = 0.03687 \]

Estimated odds ratio=0.342 (roughly \((25/37)/(18/9)=0.339\))

95% confidence interval: \((0.11,0.95)\)

95% confidence interval for YES (normal approximation)

Out of X: \((0.478,0.842)\)

Out of A: \((0.276,0.524)\)
4. Statistical Approaches

Approach 3: Bayesian hierarchical model

\[
\begin{pmatrix}
\mathbf{mit}_{ijkl} \\
\mathbf{mei}_{ijkl}
\end{pmatrix}
\sim N
\begin{pmatrix}
\begin{pmatrix}
\theta_{ijl}^{mit} \\
\theta_{ijl}^{mei}
\end{pmatrix}, \\
\sigma_i^2 I_2
\end{pmatrix}
\]

\[
\begin{pmatrix}
\theta_{ijl}^{mit} \\
\theta_{ijl}^{mei}
\end{pmatrix}
\sim N
\begin{pmatrix}
\begin{pmatrix}
\theta_{ij}^{mit} \\
\theta_{ij}^{mei}
\end{pmatrix}, \\
\begin{pmatrix}
\tau_{mit}^2 & 0 \\
0 & \tau_{mei}^2
\end{pmatrix}
\end{pmatrix}
\]

Pairs \(\mathbf{mit}_{ijkl}\) and \(\mathbf{mei}_{ijkl}\), for each gene \(l\), each classification group \(i\) (out of X, out of A) and gene type \(j\) (parental, offspring), have individual means \(\theta_{ijl}^{mit}\) and \(\theta_{ijl}^{mei}\), respectively, and common classification group variances \(\sigma_i^2\). Then, the objective is to compute,

\[
\Pr\left[\text{YES}_{\text{Out of } X,l} \right] = \Pr\left[ \left\{ \theta_{\text{Out of } X, \text{par}, l}^{mit} > \theta_{\text{Out of } X, \text{par}, l}^{mei} \right\} \mid \left\{ \theta_{\text{Out of } X, \text{off}, l}^{mit} < \theta_{\text{Out of } X, \text{off}, l}^{mei} \right\} \right]
\]

\[
\Pr\left[\text{YES}_{\text{Out of } A,l} \right] = \Pr\left[ \left\{ \theta_{\text{Out of } A, \text{par}, l}^{mit} > \theta_{\text{Out of } A, \text{par}, l}^{mei} \right\} \mid \left\{ \theta_{\text{Out of } A, \text{off}, l}^{mit} < \theta_{\text{Out of } A, \text{off}, l}^{mei} \right\} \right]
\]
4. Statistical Approaches

Approach 3: Bayesian hierarchical model

Hyperparameters

\[ \Theta = (\theta_{11}^{mit}, \theta_{12}^{mit}, \theta_{21}^{mit}, \theta_{22}^{mit}, \theta_{11}^{mei}, \theta_{12}^{mei}, \theta_{21}^{mei}, \theta_{22}^{mei}) \]

\[ \Lambda = (\sigma_{1}^{2}, \sigma_{2}^{2}, \tau_{mit}^{2}, \tau_{mei}^{2}) \]

Prior distribution

\[ p(\Theta, \Lambda) \propto \sigma_{1}^{-2} \sigma_{2}^{-2} \tau_{mit}^{-2} \tau_{mei}^{-2} \]

This represents vague/noninformative prior views.
4. Statistical Approaches

Approach 3: Bayesian hierarchical model
4. Statistical Approaches

Approach 3: Bayesian hierarchical model
4. Statistical Approaches

Proportion of parental-retrogene pairs w/ Complementary Expression

Retrotransposition Class

Vibranovski et al, 2009, PLoS Genetics
2. Biological Problem

• Does MSCI exist in *Drosophila*? **YES**

• Is MSCI involved in the distribution of sex-biased genes in the genome? **YES**
2. Biological Problem

Stage-Specific Expression Profiling of *Drosophila* Spermatogenesis Suggests that Meiotic Sex Chromosome Inactivation Drives Genomic Relocation of Testis-Expressed Genes

Maria D. Vibranovski\(^1\), Hedibert F. Lopes\(^2\), Timothy L. Karr\(^3\), Manyuan Long\(^1\)

1 Department of Ecology and Evolution, The University of Chicago, Chicago, Illinois, United States of America, 2 The University of Chicago Booth School of Business, Chicago, Illinois, United States of America, 3 The Biodesign Institute, Arizona State University, Tempe, Arizona, United States of America

Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes

MD Vibranovski, HF Lopes, TL Karr, M Long
PLoS genetics 5 (11), e1000731

Citado por 109 2009
2. Biological Problem

Male infertility

In Humans:

• 40% of infertility
• High cost treatment
• Associated with gene expression/function in spermatogenesis
• 30% of infertility are sperm deficiency

Drosophila Model:

• Spermatozoa is one of the few cell types that has homologous function for all sexual organism including humans
• Post-meiotic transcription
2. Biological Problem

Gene expression levels

<table>
<thead>
<tr>
<th>phase</th>
<th>mitosis</th>
<th>meiosis</th>
<th>post-meiosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>value 1</td>
<td>9.32</td>
<td>6.94</td>
<td>4.61</td>
</tr>
<tr>
<td>value 2</td>
<td>9.64</td>
<td>7.37</td>
<td>4.72</td>
</tr>
<tr>
<td>value 3</td>
<td>9.51</td>
<td>7.38</td>
<td>4.56</td>
</tr>
<tr>
<td>average</td>
<td>9.49</td>
<td>7.23</td>
<td>4.63</td>
</tr>
</tbody>
</table>

Values 1-3 correspond to the expression intensity (log transformed) obtained from Microarray hybridization of three different biological replicates of each spermatogenic phase. Averages values correspond to the arithmetic average within replicates for each spermatogenic phase.

Comparison of expression levels

<table>
<thead>
<tr>
<th>phase</th>
<th>mitosis</th>
<th>meiosis</th>
<th>post-meiosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>mitosis</td>
<td>-</td>
<td>Over</td>
<td>Over</td>
</tr>
<tr>
<td>meiosis</td>
<td>-</td>
<td>-</td>
<td>Over</td>
</tr>
<tr>
<td>post-meiosis</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The classifications in this table (Equal, Over or Under) were obtained by our Bayesian Statistical Model (Methods). Each classification is given to a pair-wise comparison between two spermatogenic phases. For example, "Over" mitotic vs. meiotic classification means that, for this gene, the mitotic expression is significantly higher than the meiotic expression.

http://pondside.uchicago.edu/~longlab/spermpress/
5. Ongoing Biological Problems

Vibranovski et al, 2009, PLoS Genetics
5. Ongoing Biological Problems

Post-meiosis Over-expression

~ 20% of the genes

RNA level

Mitosis  Meiosis  Post-meiosis

Spermatogenesis

Vibranovski et al, 2009, PLoS Genetics
RNA synthesis does **NOT** occur in post-meiotic stages

RNA synthesis: incorporation of $[^3\text{H}]$uridine in the *Drosophila melanogaster* testes has been studied by autoradiography.

Olivieri and Olivieri, 1965, Mutation Research
5. Ongoing Biological Problems

- **Transcription**
- **Translation**
- **RNA**

**Spermatogenesis**
- **Mitosis**
- **Meiosis**
- **Post-meiosis**

Diagram showing the processes of spermatogenesis with stages labeled as Transcription, Translation, and RNA, along with mitosis, meiosis, and post-meiosis.
5. Ongoing Biological Problems

Transcription

Translation

RNA

RNA level

Mitosis | Meiosis | Post-meiosis

Spermatogenesis

Yes for post-meiotic transcription

No post-meiotic transcription
5. Ongoing Biological Problems

Bromo-uridine (BrU) incorporation: Direct evidence

(1) Incorporation of BrU
(2) Labeling nascent RNAs with Br-UTP
(3) Binding anti-BrdU antibody
(4) Imaging

Vibranovski et al, 2010, Genetics
Bromo-uridine (BrU) incorporation: 
Post-meiotic transcription - Direct evidence

Spermatid cyst

BrU
DNA
n: nuclei

Vibransovski et al, 2010, Genetics
5. Ongoing Biological Problems

Meiotic cells

Post-meiotic cells

Mitotic cells

Meiotic cells

Vibranovski et al, 2009, PLoS Genetics
5. Ongoing Biological Problems

Zhang, Vibranovski, Krinsky and Long, 2010
5. Ongoing Biological Problems

Mitosis  Meiosis  Post-meiosis

Spermatogenesis

New genes RNA  Old genes RNA
Proportion of Genes by Spermatogenesis Phase and by age

- **Mitosis**
  - Old: 50%
  - New: 30%
  - Fisher Exact Test: P = 0.47

- **Meiosis**
  - Old: 20%
  - New: 40%

- **Pos-Meiosis**
  - Old: 30%
  - New: 50%

Fisher Exact Test:

- * P < 0.05
- ** P < 0.01
- *** P < 0.001

5. Ongoing Biological Problems

Júlia Raices
5. Ongoing Biological Problems

<table>
<thead>
<tr>
<th>Classes</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>Mit Mei Pos</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5. Ongoing Biological Problems

Júlia Raices

Proportion of genes (%)
Perspectives

Development

Molecular Biology

Bioinformatics

Genetics

Cell Biology

Statistics
Acknowledgments

THE VIBRANOVSKI LAB

Gustavo França
IQ /USP

Júlia Raices
IB /USP

Mariana Kanbe
ICB /USP

Joana Carvalho
IB /USP

THE LONG LAB

Manyuan Long
Yuan Huang,
Claus Kemkemer,
Ben Krinsky
Patrick Landback
Nick Vankuren
Jun Wang
Qian Yang
Chengjung Zhang

Post-meiotic transcription

Domitille Chalopin, ENS, Lyon
Hedibert Lopes
Timothy Karr

Funding: