

UNIVERSIDADE DE SÃO PAULO
Instituto de Biociências
Departamento de Genética e Biologia Evolutiva

Carolina de Athayde Mendonça

**Funções de genes novos do cromossomo Y de *Drosophila*
por meio da análise de expressão na espermatogênese**

**A study on the functions of new Y-linked genes through
expression analysis of *Drosophila*'s spermatogenesis**

Profa. Dra. Maria Dulcetti Vibranovski

São Paulo

2020

VERSÃO
CORRIGIDA

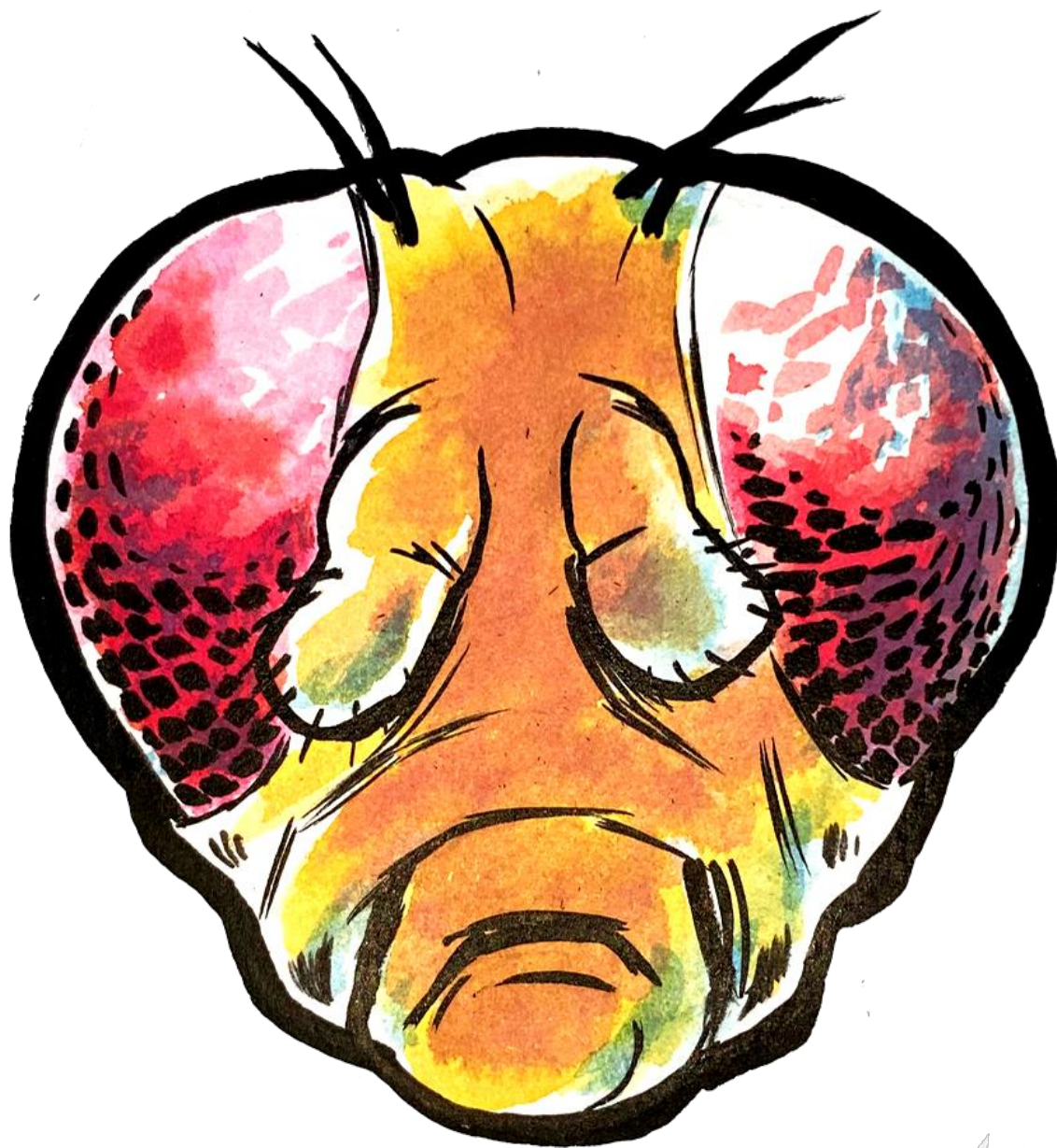
Resumo

Os estudos do cromossomo Y de *Drosophila melanogaster* sempre caminharam de maneira paulatina mesmo com o advento da era genômica devido principalmente à sua natureza heterocromática. Métodos engenhosos permitiram o mapeamento de seis fatores de fertilidade e a identificação de 12 genes de cópia-única no Y. Tais genes foram originados a partir de duplicações autossômicas. A comparação do perfil de expressão de um gene que é ligado ao Y em uma espécie com seu ortólogo autossômico em uma segunda espécie pode revelar aspectos importantes acerca da evolução dos perfis de expressão dos genes do cromossomo Y. Foram realizados sequenciamentos de RNA estágio-específico (mitose, meiose e pós-meiose) da espermatogênese para três espécies: *D. melanogaster*, *D. willistoni* e *D. mojavensis*. A análise de expressão diferencial mostrou que o aumento de expressão da mitose para a meiose é comumente encontrado nos genes do cromossomo Y e em seus ortólogos autossômicos, sugerindo que esse é um caráter ancestral que foi selecionado entre os genes duplicados para o Y. O aumento de expressão na meiose é coincidente com a formação de *loops* nos espermatócitos intermediários. Nós propomos que genes duplicados para o cromossomo Y, cujo perfil de expressão ancestral era de aumento de expressão da mitose para a meiose, têm maior probabilidade de fixação no Y.

Abstract

The study of *Drosophila melanogaster* Y chromosome has always been slow even in the genomic era mainly because of its heterochromatic nature. Ingenious methods lead to the mapping of six fertility factors and the identification of 12 single-copy genes. These genes were originated from duplications of autosomes. The comparison of the expression profile of a gene that is Y-linked in one species and autosomic in a second species may shed light on the evolution of the expression profiles of Y-linked genes. Stage-specific (mitosis, meiosis and post-meiosis) RNAseq from spermatogenesis were conducted for *D. melanogaster*, *D. willistoni* and *D. mojavensis*. Differential expression analysis revealed that the increase in expression from mitosis to meiosis is commonly found in Y-linked genes and in their autosomic orthologs. This suggests that the increase in meiosis is an ancestral character that was selected among genes duplicated to the Y. The increased expression in meiosis coincides with the formations of Y-loops in intermediary spermatocytes. We propose that genes duplicated to the Y chromosome, which ancestrally had an increase in expression from mitosis to meiosis, have a higher probability of being fixated on the Y.

I - Introdução



1. O cromossomo Y de *D. melanogaster*

Os estudos do cromossomo Y de *D. melanogaster* são quase tão antigos quanto a Genética como disciplina. Em seu trabalho histórico sobre a base cromossômica da herança, Bridges (1916) trabalhou com não-disjunção em gametas de *Drosophila*. Duas importantes observações devem ser destacadas desse trabalho. A primeira delas é que um indivíduo X0 é um macho viável, porém estéril. A segunda é que um indivíduo XXY é uma fêmea viável e fértil. Com base nessas observações, foi possível inferir que o cromossomo Y em *Drosophila* não determina o sexo masculino, mas é essencial para a fertilidade masculina. Atualmente, é conhecido que a determinação do sexo está sob controle do gene *Sex lethal* (CLINE, 1989), que está no topo de uma cascata gênica que leva, em última instância, a estruturas e comportamentos sexualmente dimórficos (revisado em: SALZ; ERICKSON, 2010).

Em 1929, Stern (*apud* BROSSEAU, 1960) demonstrou que ambos os braços do cromossomo Y contêm fatores essenciais para a fertilidade masculina, mas foi apenas três décadas depois que eles foram identificados. Em 1960, Brosseau fez experimentos com linhagens irradiadas de *Drosophila* e, por meio de testes de complementação, identificou sete regiões do cromossomo Y essenciais para a fertilidade masculina: cinco delas no braço longo (kl-1, kl-2, kl-3, kl-4 e kl-5) e duas no braço curto (ks-1 e ks-2). A essas regiões deu-se o nome de *fatores de fertilidade*. Em 1981, Kennison fez uso de translocações entre os cromossomos X-Y e confirmou seis dos fatores previamente identificados (kl-4 não foi confirmado), sendo que cada um deles corresponderia presumivelmente a um gene (Figura 1).

Nesse mesmo ano, foi demonstrado que deleções das regiões do kl-3 ou kl-5 resultavam na perda do braço externo da dineína do axonema flagelar (HARDY; TOKUYASU; LINDSLEY, 1981). O gene responsável pela deficiência (nomeado *Dhc-Yh3*, mas referido daqui em diante como *kl-5*) foi localizado no kl-5 e foi verificado que sua expressão se dá nas fases mais tardias da espermatogênese (GEPNER; HAYS, 1993) (Figura 2).

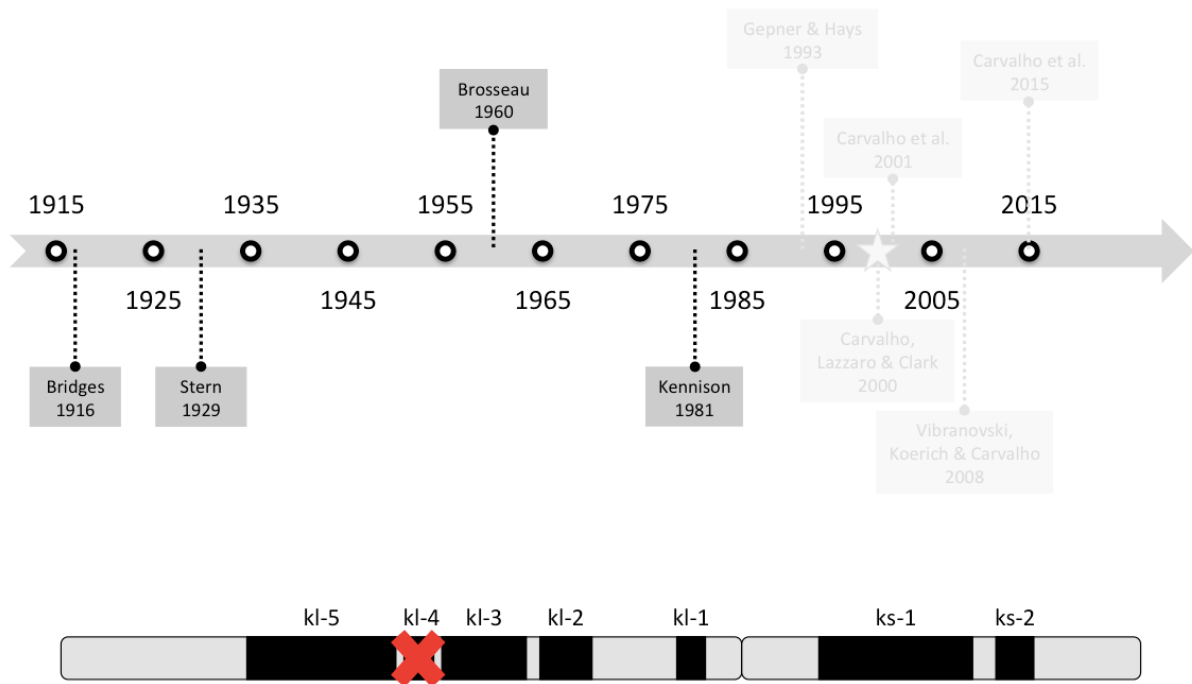


Figura 1: Os primeiros trabalhos a respeito do cromossomo Y de *Drosophila* mostraram que ele é essencial para a fertilidade masculina e que possui seis fatores de fertilidade.

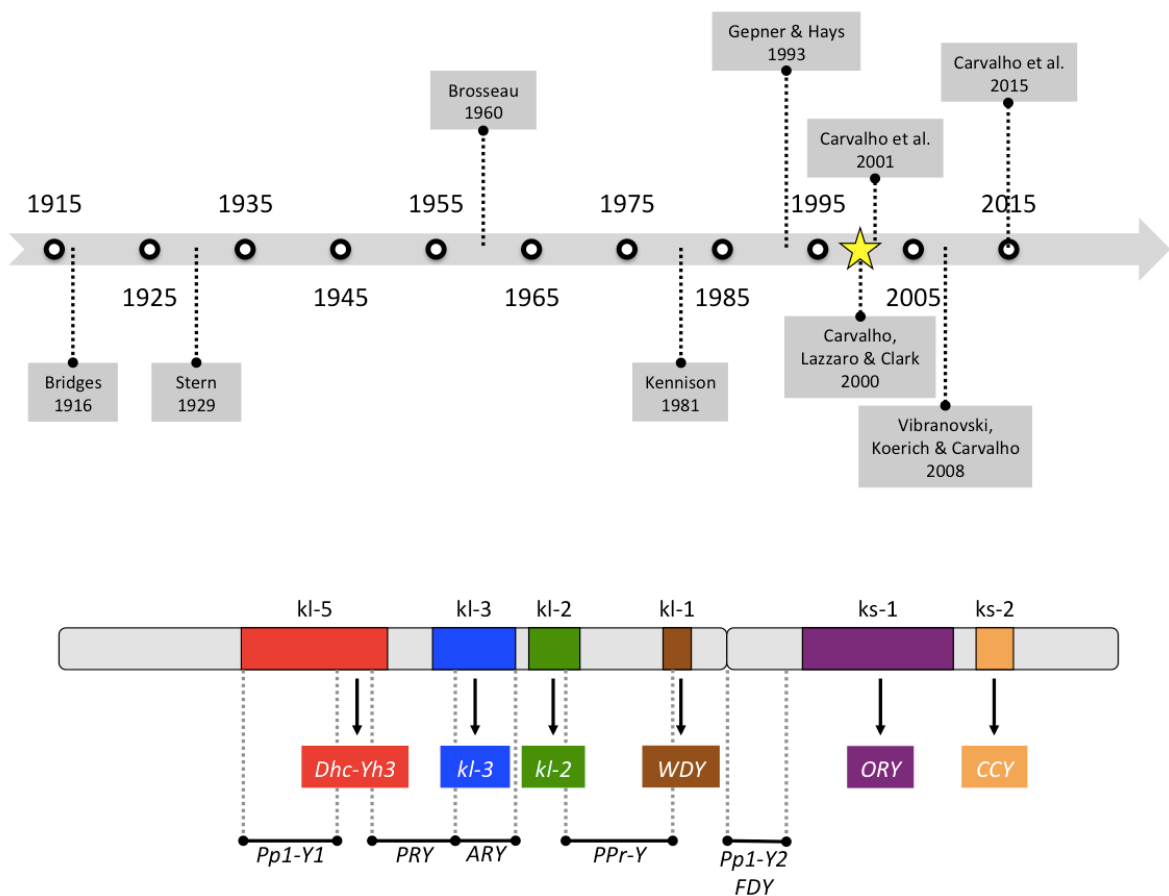


Figura 2: Estado-da-arte dos genes do cromossomo Y de *D. melanogaster*. Foram identificados 12 genes de cópia única, sendo que existe pelo menos um candidato a gene essencial em cada um dos fatores de fertilidade. A estrela amarela marca o ano do sequenciamento do genoma de *D. melanogaster*. Para distinguir os fatores de fertilidade dos genes, os últimos estão sempre destacados em *itálico*.

O avanço paulatino na identificação dos genes do cromossomo Y talvez fosse acelerado com o sequenciamento do genoma de *D. melanogaster* (ADAMS *et al.*, 2000). No entanto, as sequências repetitivas da heterocromatina do Y não são estáveis dentro dos vetores usados em projetos de sequenciamento. Além disso, elas dificultam a montagem dos *reads* individuais em *scaffolds* e destes em cromossomos completos (CARVALHO; LAZZARO; CLARK, 2000). Consequentemente, poucas sequências foram atribuídas ao cromossomo Y. Provavelmente, o restante delas estaria no chamado *unmapped arm* (*armU*), que são as sequências que não foram mapeadas em nenhum dos braços cromossômicos. Partindo dessa hipótese, Carvalho, Lazzaro e Clark (2000), desenvolveram um engenhoso método para identificar genes ligados ao Y.

Em suma, sequências do *armU* filtradas pelo *RepeatMasker* (SMIT; HUBLEY; GREEN, 1996-2010) são usadas para identificação de proteínas candidatas (BLASTX). Em seguida, essas proteínas são traduzidas em seus seis possíveis quadros de leitura (TBLASTN) e usadas como *queries* em buscas de similaridade contra o próprio *armU*. Segue-se então a confirmação experimental por reações em cadeia da polimerase (PCR) em machos e fêmeas para verificar se tal gene é de fato ligado ao Y. Nos casos positivos, linhagens com translocações recíprocas X-Y são usadas para localizar o gene no cromossomo Y. Por último, com uso de transcrição reversa-PCR (RT-PCR), as lacunas entre as sequências são preenchidas. Nesse trabalho, os dois genes dos fatores de fertilidade *kl-2* e *kl-3* foram identificados, sendo que eles também codificam cadeias pesadas da dineína. Adicionalmente, foi identificado o gene *PRY* que não é um gene essencial, pois localiza-se na região entre os fatores *kl-3* e *kl-5*.

Em 2001, cinco novos genes foram identificados (CARVALHO *et al.*, 2001): *Ppl-Y1*, *Ppl-Y2*, *PPr-Y*, *ORY* e *CCY*. Os três primeiros codificam fosfatases de proteínas e *ORY* é um gene relacionado à ocludina. *Ppl-Y2* e *PPr-Y* localizam-se em regiões que não correspondem a fatores de fertilidade e, portanto, não são genes essenciais. Apesar de *Ppl-Y1* fazer parte da região do *kl-5*, provavelmente ele não é um gene essencial pois o gene essencial desse

fator de fertilidade já foi identificado (*kl-5*). *ORY*, por sua vez, é o único candidato a gene essencial da região do *ks-1*. *CCY* merece um pouco mais de detalhamento pois ele foi identificado usando um método ligeiramente diferente. Sequências expressas em testículos foram comparadas ao *armU* (BLASTN), partindo-se da premissa de que genes do Y são expressos em testículos. A única pista a respeito da função de *CCY* é sua similaridade com proteínas *coiled-coils* e ele permanece como o principal candidato a gene essencial do *ks-2*.

Entre 2001 e 2008, o número de sequências expressas em testículos aumentou de 3.000 para 31.000, o que possibilitou uma nova busca por genes ligados ao Y. Em Vibranovski, Koerich e Carvalho (2008), *ARY* e *WDY* foram identificados. O primeiro é um potencial membro da família de genes de aldol-redutases e, por localizar-se na região do *kl-3*, não deve se tratar de um gene essencial. A função de *WDY* permanece um mistério mas, por ser o único gene identificado na região do *kl-1*, ele é o melhor candidato a gene essencial desse fator de fertilidade.

Mais recentemente, Carvalho *et al.* (2015) identificaram *FDY*. Ele foi originado a partir de uma duplicação da região autossômica do gene *vig2* para o cromossomo Y. Com exceção do *FDY*, que resguarda uma identidade de 98% a nível de nucleotídeos com seu parental, os outros genes duplicados para o Y foram pseudogenizados. Apesar da semelhança, *vig2* é expresso em machos e fêmeas, enquanto *FDY* é expresso exclusivamente em testículos. *FDY* foi o último gene de cópia única identificado no cromossomo Y de *D. melanogaster*, resultando em um total de 12 (Tabela I), sendo que existe pelo menos um candidato a gene essencial para cada um dos fatores de fertilidade.

2. A origem do cromossomo Y de *Drosophila*

Como mencionado anteriormente, o gene *FDY* foi originado a partir da duplicação de uma região autossômica de 3R para o cromossomo Y (CARVALHO *et al.*, 2015). Como o próprio nome do gene sugere (*flagrante delicto Y*), esse é um evento “pego em flagrante” pois trata-se da duplicação

mais recente (e exclusiva) de *D. melanogaster*. A história contada pelo *FDY* está longe de ser um caso anedótico. Na verdade, a origem de genes do cromossomo Y a partir de duplicações autossômicas parece ser a regra em *Drosophila*.

Tabela I: Genes de cópia única do cromossomo Y de *D. melanogaster*, suas localizações, funções e as referências de suas primeiras descrições.

Gene	Região do cromossomo Y	Função	Referência
<i>kl-2</i>	kl-2	cadeia pesada 1 β da dineína	CARVALHO; LAZZARO; CLARK, 2000.
<i>kl-3</i>	kl-3	cadeia pesada γ da dineína	CARVALHO; LAZZARO; CLARK, 2000.
<i>Dhc-Yh3</i>	kl-5	cadeia pesada da dineína	GEPNER; HAYS, 1993.
<i>ORY</i>	ks-1	relacionado à ocludina	CARVALHO <i>et al.</i> , 2001.
<i>PRY</i>	kl-3 - kl-5	relacionado à policisteína fosfatase de proteína	CARVALHO; LAZZARO; CLARK, 2000.
<i>PPr-Y</i>	kl-1 - kl-2	fosfatase de proteína	CARVALHO <i>et al.</i> , 2001.
<i>CCY</i>	ks-2	?	CARVALHO <i>et al.</i> , 2001.
<i>ARY</i>	kl-3	aldo-ceto redutase	VIBRANOVSKI; KOERICH; CARVALHO, 2008.
<i>WDY</i>	kl-1	?	VIBRANOVSKI; KOERICH; CARVALHO, 2008.
<i>Pp1-Y1</i>	kl-5	fosfatase de proteína	CARVALHO <i>et al.</i> , 2001.
<i>Pp1-Y2</i>	próximo ao centrômero, entre as bandas h15 e h20	fosfatase de proteína	CARVALHO <i>et al.</i> , 2015.
<i>FDY</i>	próximo ao centrômero, entre as bandas h15 e h20	organização da heterocromatina	CARVALHO <i>et al.</i> , 2015.

Cromossomos sexuais estão presentes em diferentes táxons, sendo que em alguns deles o par heteromórfico leva ao desenvolvimento de machos (XX/XY) e em outros de fêmeas (ZZ/ZW). Tradicionalmente (teoria canônica), postula-se que o par heteromórfico teria se originado a partir de um par de autossomos. O processo de diferenciação seria iniciado com a aquisição de um gene de determinação sexual pelo Y (ou W). Uma combinação de processos evolutivos levaria à diminuição da recombinação entre os cromossomos e consequente degeneração do Y. Em última instância, o Y poderia ser completamente

degenerado, levando ao sistema de determinação sexual XX/X0 (revisado em: CHARLESWORTH; CHARLESWORTH, 2000) (Figura 3).

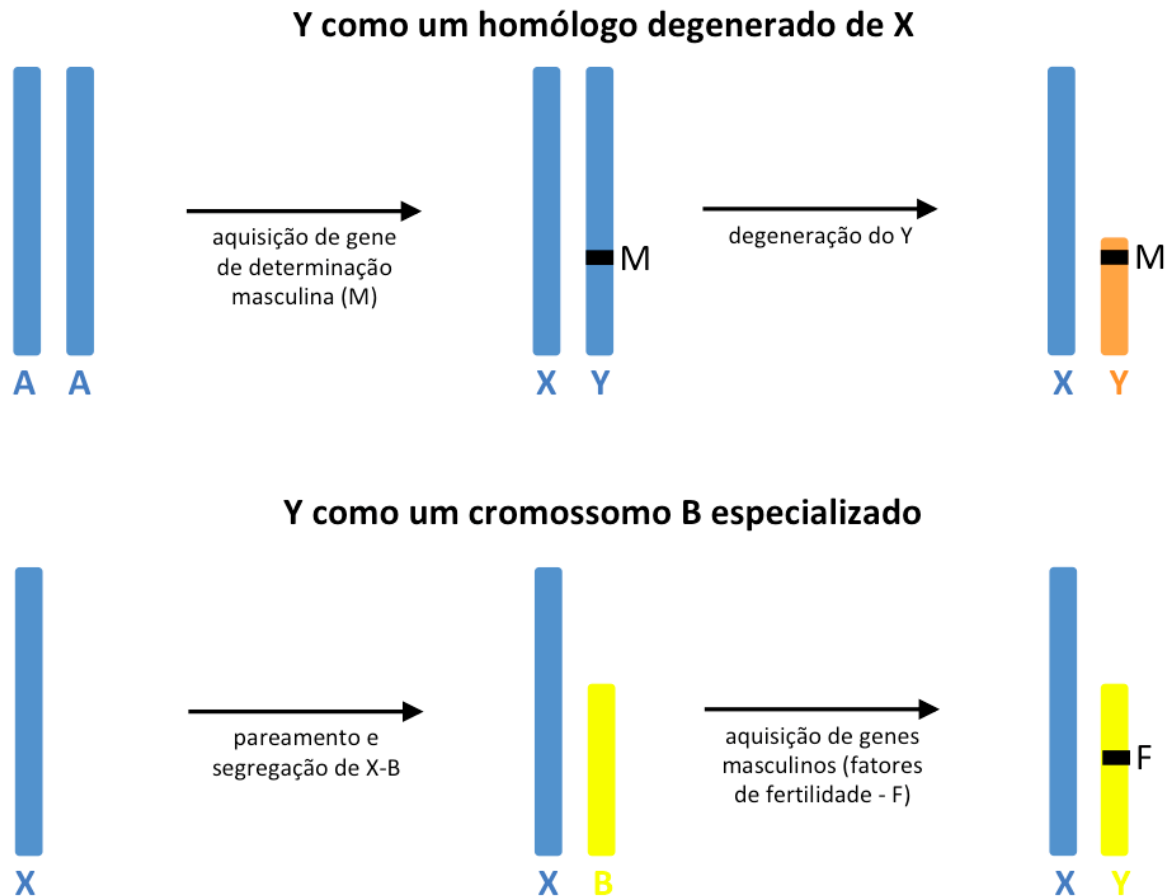


Figura 3: Duas hipóteses que explicam a origem do cromossomo Y. Segundo a hipótese canônica, um par ordinário de autossomos (A) teria originado XY a partir da aquisição de um gene de determinação sexual masculina (M) por Y e sua consequente degeneração. Alternativamente, um cromossomo B poderia parear e segregar regularmente com X em uma espécie ancestralmente X0. Com a aquisição de genes masculinos (F), esse cromossomo passaria a ser indispensável e a segregação seria reforçada pela seleção natural. - Reproduzida e traduzida de Carvalho, 2002.

Em mamíferos, existem evidências de que o cromossomo Y é de fato um homólogo degenerado do X. A primeira delas é a aquisição pelo Y de um gene de determinação sexual masculina, o *SRY*. A segunda é a presença no Y de genes (ou genes degenerados) homólogos aos genes do X (revisado em: HUGHES; PAGE, 2015). Em *Drosophila*, por outro lado, não há um gene de determinação sexual masculina no cromossomo Y. Ademais, genes ribossômicos (rDNA) são os únicos encontrados de forma homóloga nos cromossomos X e Y,

mas não constituem evidência suficiente de homologia pois cromossomos B comumente adquirem rDNAs (HACKSTEIN *et al.*, 1996). Na verdade, os parálogos mais próximos de genes do cromossomo Y localizam-se em autossomos. Com base nessas observações, foi proposta uma teoria alternativa para a origem do cromossomo Y de *Drosophila*, segundo a qual ele teria se originado a partir de um cromossomo B que adquiriu genes masculinos (CARVALHO, 2002).

Para além dos parálogos em *D. melanogaster*, os ortólogos em outras espécies de *Drosophila* também corroboram a origem não-canônica do Y. Vários dos genes ligados ao Y de *D. melanogaster* são autossômicos em outras espécies. *kl-2*, *kl-3*, *ORY*, *PRY* e *Ppr-Y* são os genes ligados ao Y mais antigos, mas na ausência de um grupo externo, não é possível inferir a direção de seus movimentos (perdas ou ganhos). *CCY* e *ARY* são ancestralmente autossômicos, mas passam a ser ligados ao Y em *Sophophora*. *WDY*, *kl-5*, *Pp1-Y1* e *Pp1-Y2* são ligados ao Y apenas no grupo *melanogaster*. A [Figura 4](#) mostra os ganhos e perdas do Y, incluindo não apenas os 12 genes que são ligados ao Y em *D. melanogaster* (KOERICH *et al.*, 2008), mas também outros quatro genes (*JYa*, *GJ19835*, *CG11719* e *CG2964*) que são autossômicos em *D. melanogaster* e ligados ao Y em outras espécies (CARVALHO; CLARK, 2013).

A taxa de ganho de genes no cromossomo Y é aproximadamente 11 vezes maior do que a taxa de perda, o que mostra que seu conteúdo é mais dinâmico do que o de outros cromossomos. A constante aquisição de genes pelo Y a partir de duplicações autossômicas explicaria como um cromossomo B poderia adquirir genes masculinos (CARVALHO; KOERICH; CLARK, 2009). Uma pergunta que permanece aberta é quais genes são duplicados e mantidos no cromossomo Y. Uma estratégia para responde-la é comparar genes que são ligados ao Y em uma espécie com seus ortólogos autossômicos em uma outra espécie. A comparação entre os ortólogos pode revelar aspectos importantes da evolução desses genes e, portanto, do cromossomo Y.

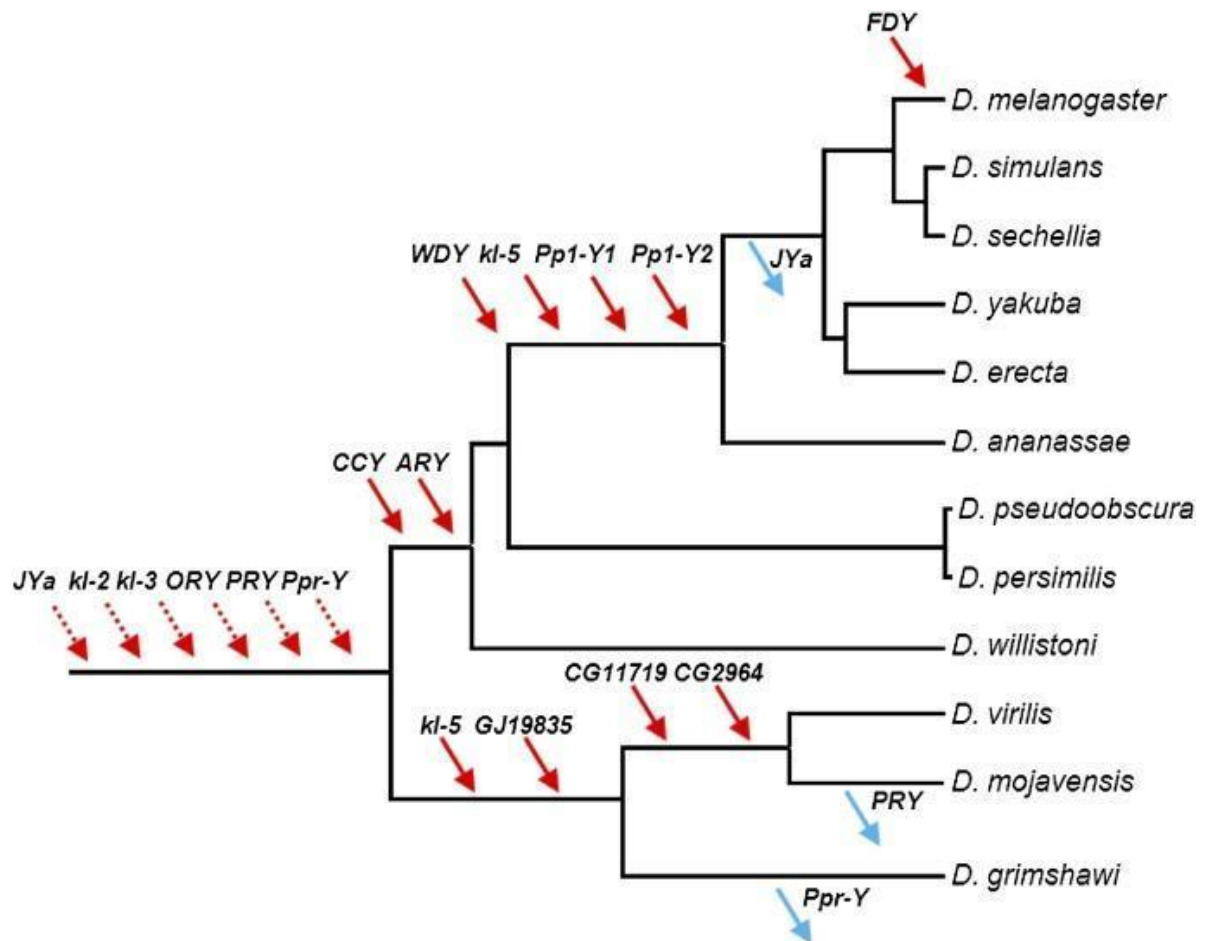


Figura 4: Ganhos e perdas de genes do cromossomo Y de *Drosophila*. As setas vermelhas simbolizam ganhos e as setas azuis perdas. Para os seis genes ancestrais (setas pontilhadas), não é possível inferir a direção do movimento sem um grupo externo. *JYa*, *CG11719* e *CG2964* são os nomes dos ortólogos em *D. melanogaster*. *GJ19835* é o nome do ortólogo em *D. virilis*. - Extraída de Carvalho e Clark, 2013.

3. A espermatogênese de *Drosophila*

Avanços significativos foram feitos no sentido de identificação de genes do cromossomo Y, não apenas em *D. melanogaster*, mas também em outras espécies de *Drosophila*. Apesar dos números serem modestos (apenas 12 genes de cópia única em *D. melanogaster*), essa lista provavelmente está próxima à saturação. Mesmo com o uso de diferentes abordagens, os mesmos genes vêm sendo repetidamente recuperados e o número de genes novos identificados é cada vez menor (VIBRANOVSKI; KOERICH; CARVALHO, 2008). Assim, o desafio não é mais a identificação de genes do cromossomo Y.

Para os poucos genes cujas funções são conhecidas (*kl-2*, *kl-3* e *kl-5*), observa-se uma coesão funcional em relação à fertilidade masculina: todos

codificam cadeias pesadas da dineína, essenciais para o batimento flagelar do espermatozóide. No entanto, para os outros genes, o máximo que se sabe são especulações com base na similaridade de domínios de proteínas (Tabela I). Um primeiro passo para se compreender a função de um gene é conhecer o seu perfil de expressão. Posto que tratam-se de genes do cromossomo Y, uma premissa razoável é que eles são expressos em testículos.

A espermatogênese de *Drosophila* inicia-se com a divisão assimétrica de células-tronco germinativas. Essa divisão tem a dupla-função de manter a população de células germinativas e produzir um espermatócito primário. As células-tronco (que vão originar um espermatócito) passam por quatro divisões mitóticas consecutivas com citocinese incompleta (fase de espermatogônias), resultando em um cisto de espermatócitos primários de 16 células interconectadas. Elas passam por uma etapa de crescimento e diferenciação, dando origem aos espermatócitos primários maduros. Eles, por sua vez, passam pela meiose I e II, transformando-se em um cisto de 64 espermátides (o chamado *estágio de cebola*). Por último, as espermátides passam por uma fase de alongamento e individualização (HIRST; CARMICHAEL, 2011).

Apesar dos diferentes tipos celulares não estarem confinados a compartimentos separados, eles são distribuídos em um gradiente ao longo do testículo de *D. melanogaster*. Podem ser identificadas três regiões, cada qual com predomínio de determinados tipos celulares (Figura 5): (1) a extremidade apical do testículo, contendo células em mitose (enriquecida com células-tronco apicais, cistos espermatogoniais e níveis reduzidos de cistos de espermatócitos primários); (2) a porção mediana do órgão, contendo células em meiose (enriquecida com cistos de espermatócitos primários e secundários, níveis reduzidos de espermátides); (3) a porção final, contendo células pós-meióticas (enriquecida com cistos de espermátides em processo de alongamento, grupos de espermátides e baixos níveis de cistos de espermatócitos) (VIBRANOVSKI *et al.*, 2009).

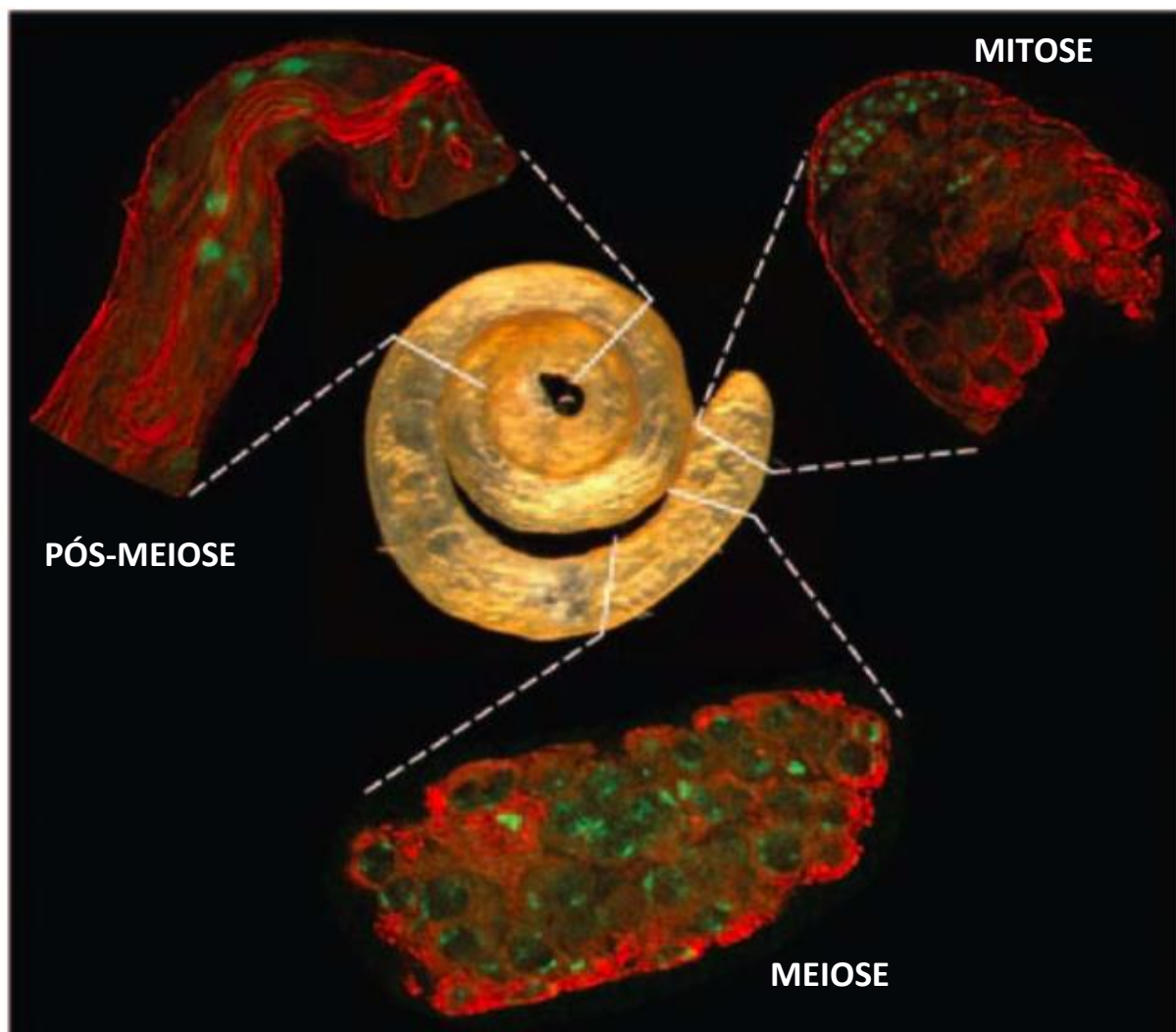


Figura 5: No centro, está mostrado um testículo inteiro de *D. melanogaster* e, nas regiões periféricas, amostras de tecidos isoladas. Os tecidos foram corados por imunofluorescência indireta: α -tubulina em vermelho e DNA em verde. - Extraída de Vibranovski *et al.*, 2009.

Além da descrição do método para separação do testículo, Vibranovski *et al.* (2009) também foram responsáveis pelos primeiros dados de expressão estágio-específico de testículos com dados de microarranjo. Apesar da relevância do trabalho, sondas para os genes do cromossomo Y não estavam disponíveis na época. Uma segunda base de dados relevante de expressão na espermatogênese de *Drosophila* é o FlyTed (ZHAO *et al.*, 2010). Ele reúne 2.762 fotos de hibridizações *in situ* (ISH) de RNA de 817 genes. Apesar de menos abrangente do que os dados de microarranjo, a ISH permite uma descrição mais detalhada do perfil de expressão do gene. No entanto, o FlyTed também não inclui genes ligados Y. Assim, o único gene do Y cujo perfil de expressão é conhecido é o *kl-5* (GEPNER; HAYS, 1993).

4. Evolução da expressão de genes ligados ao Y

O cromossomo Y de *Drosophila* oferece um sistema único para se estudar a evolução dos perfis de expressão de genes do cromossomo Y. Nos casos de duplicações mais recentes, é possível comparar o perfil de expressão de um gene do cromossomo Y com o do seu parental autossômico. Por exemplo, em *D. melanogaster*, ainda é possível reconhecer *vig2* como parental de *FDY*. Em outros casos, é possível identificar parálogos autossômicos de genes do Y, mas pelo fato da duplicação ser antiga, o gene autossômico pode não recapitular mais o perfil de expressão ancestral. Nesses casos, uma segunda estratégia seria comparar genes que são ligados ao Y em uma espécie com seus ortólogos autossômicos em uma outra espécie. Por exemplo, o gene *CCY* é ligado ao Y em *D. melanogaster*, mas autossômico em todas as espécies fora de *Sophophora*.

A comparação de genes ligados ao cromossomo Y com o seus ortólogos autossômicos permite distinguir entre algumas hipóteses evolutivas. Em um cenário nulo, não haveria qualquer similaridade entre os perfis de expressão dos genes ligados ao Y e nem deles com seus respectivos ortólogos autossômicos (Figura 6 - sem influência). Isto é, genes seriam duplicados para o Y aleatoriamente e não haveria qualquer tipo de seleção atuando sobre eles, resultando em perfis de expressão variáveis para os genes do Y. Além disso, a partir do momento da duplicação, esses genes passariam a evoluir de maneira independente de seus parentais autossômicos. Assim, também não seria esperado encontrar uma relação entre os perfis de expressão dos genes do cromossomo Y e de seus ortólogos autossômicos.

No entanto, a extrema especialização funcional do cromossomo Y na fertilidade masculina sugere que os genes que são duplicados para o Y podem estar sujeitos à ação da seleção natural. Um fenômeno conhecido da espermatogênese de *D. melanogaster* é a formação de alças (os chamados *lampbrush-like loops*) nos espermatócitos primários (BONACCORSI *et al.*, 1988). Os *loops* são formados nos genes *kl-3*, *kl-5* e *ORY* que possuem introns gigantes. Recentemente, foi demonstrado que a transcrição de *kl-3* e *kl-5* se dá

de maneira espaço-temporal ao longo das ~90 horas de diferenciação do espermatócito (FINGERHUT; MORAN; YAMASHITA, 2019). Considerando a abertura da cromatina durante a transcrição desses genes, uma segunda hipótese seria que determinados perfis de expressão de genes duplicados para o Y seriam selecionados (Figura 6 - pré-adaptação). Por exemplo, um gene autossômico com pico de expressão na meiose que fosse duplicado para o Y poderia ser selecionado caso sua função estivesse relacionada à fertilidade masculina. Por outro lado, um gene duplicado para o Y que possuísse expressão em uma fase em que a cromatina está condensada provavelmente seria eliminado. Diante deste cenário, seria esperado encontrar perfis similares entre os genes do Y e seus ortólogos autossômicos, já que apenas certos perfis seriam selecionados e mantidos no Y. Conseqüentemente, seria esperado também que os genes ligados ao Y possuam perfis de expressão similares entre si.

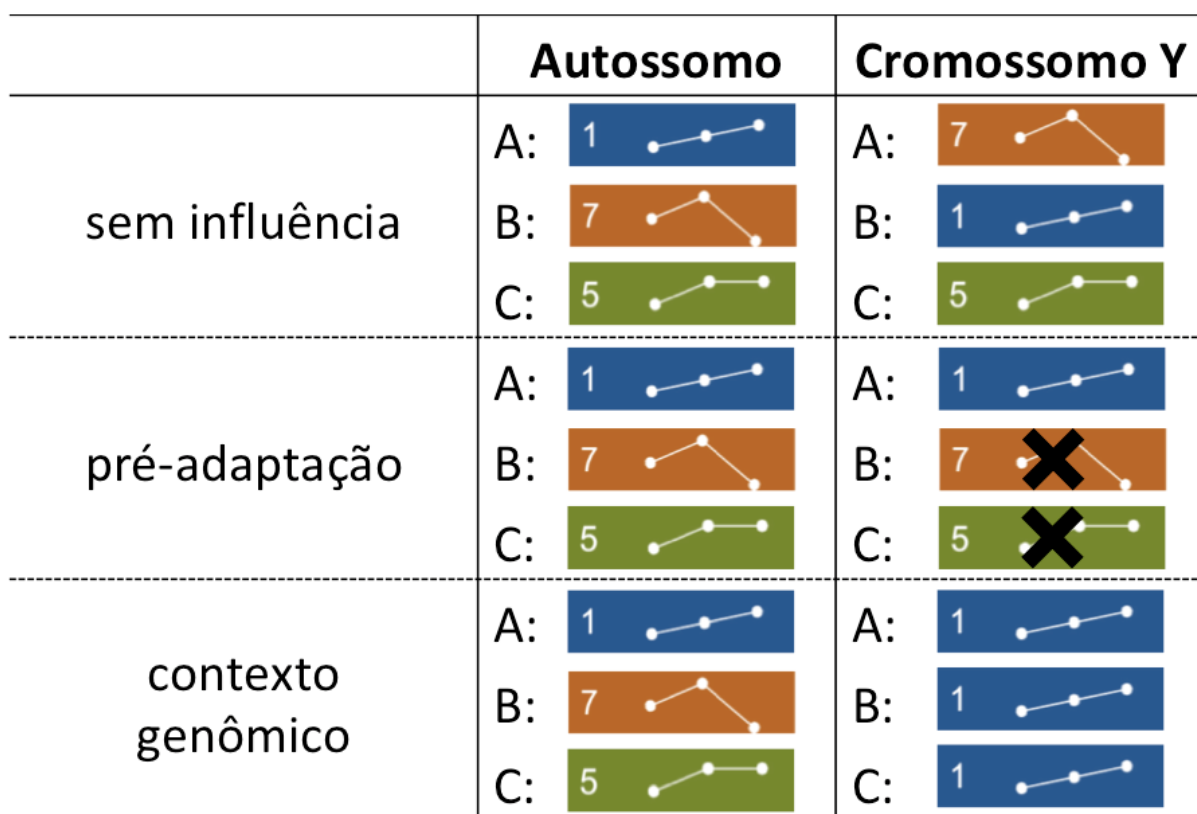


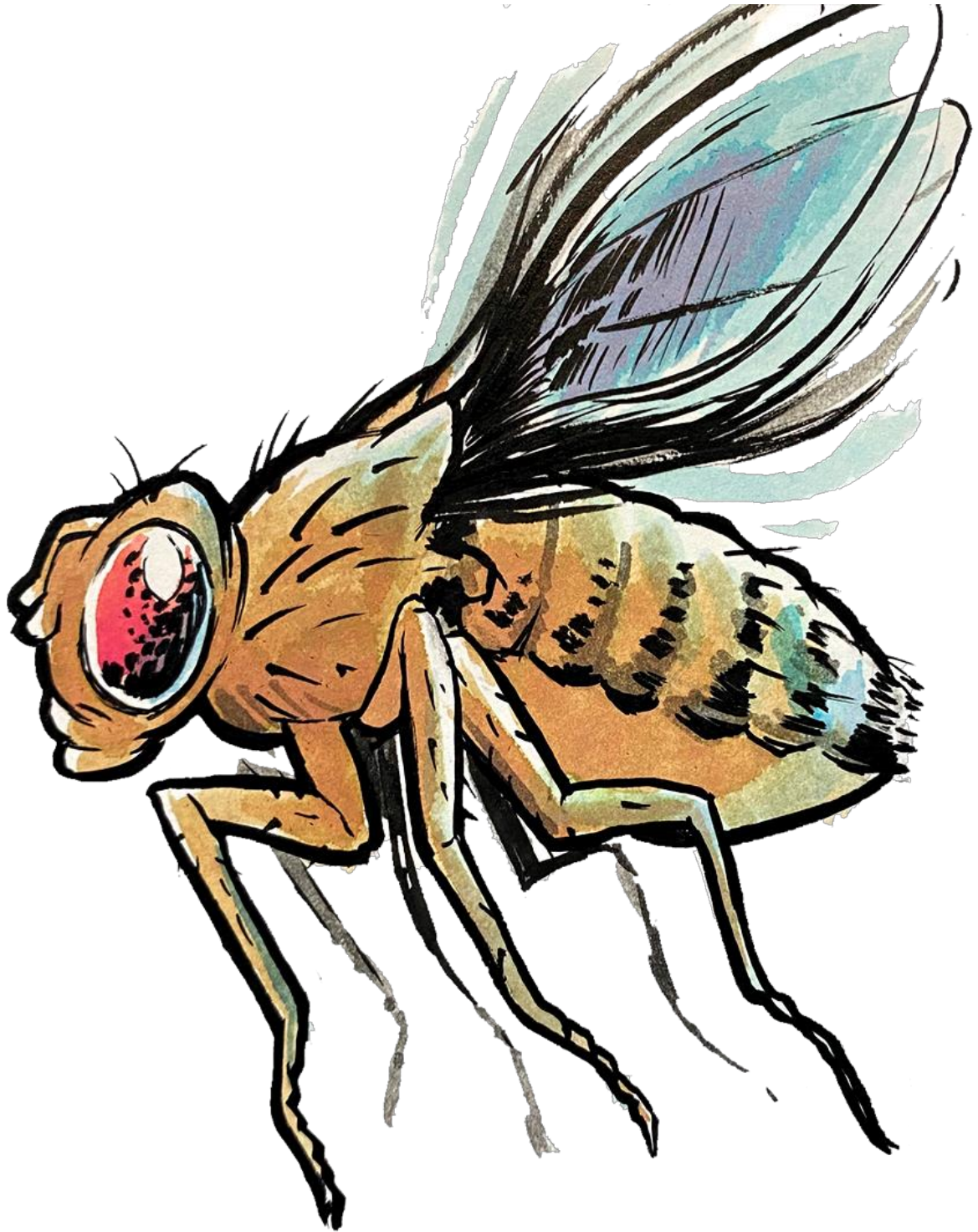
Figura 6: Hipóteses evolutivas acerca dos perfis de expressão de genes ligados ao cromossomo Y. Os quadrados coloridos representam a expressão (eixo Y) do gene ao longo da espermatogênese, sendo que cada um dos pontos (eixo X) corresponde respectivamente a: mitose, meiose e pós-meiose (adaptados de RAICES; OTTO; VIBRANOVSKI, 2019). A, B e C representam três genes diferentes. Ver texto para detalhes sobre cada uma das hipóteses.

No trabalho de Sackton *et al.* (2011), foram construídas linhagens em que os cromossomos Y de *D. simulans* ou *D. sechellia* foram introgridos em uma linhagem com *background* genômico de *D. simulans*. Análises da aptidão reprodutiva mostraram que a linhagem com cromossomo Y homoespecífico (Y-sim) apresentava performance superior à linhagem com cromossomo Y heteroespecífico (Y-sec). Além disso, foram identificados 122 genes diferencialmente expressos entre as linhagens, sendo que os genes com menor expressão em Y-sec em comparação a Y-sim eram enriquecidos com genes masculinos e genes expressos em testículos. Mais especificamente, eles apresentavam um viés de expressão na direção de genes relacionados à diferenciação terminal/espermiogênese. *D. simulans* e *D. sechellia* são próximas evolutivamente (Figura 4) e seus cromossomos Y apresentam conteúdo similar (*kl-5*, *kl-2*, *kl-3*, *PRY*, *ORY*, *PPr-Y*, *CCY*, *Pp1-Y1*, *Pp1-Y2*, *ARY* e *WDY*). Portanto, as diferenças observadas entre Y-sim e Y-sec devem estar relacionadas a diferenças epigenéticas entre os cromossomos Y das duas espécies.

Levando em consideração a relevância das marcas epigenéticas do cromossomo Y, uma terceira hipótese seria que genes que são duplicados para o Y adquiririam um novo perfil de expressão característico desse novo contexto genômico (Figura 6 - contexto genômico). Neste cenário, não seria esperado encontrar uma similaridade no perfil de expressão dos genes do cromossomo Y com seus respectivos ortólogos ancestrais. Isso porquê, a partir do momento que o gene fosse duplicado para o Y, ele adquiriria marcas epigenéticas específicas deste novo contexto genômico, podendo adquirir em última instância um novo perfil de expressão. Uma segunda previsão dessa hipótese seria que genes do cromossomo Y possuiriam perfis de expressão semelhantes por compartilharem o mesmo contexto genômico.

As hipóteses apresentadas não são necessariamente mutuamente exclusivas, mas cada uma delas gera previsões claras que permitem a identificação de quais processos/mecanismos são responsáveis pela evolução dos perfis de expressão dos genes do cromossomo Y de *Drosophila*.

V - Discussão



1. Os perfis de expressão de genes ligados ao Y

O que significa um gene ser expresso em uma fase da espermatogênese e não em outra? De maneira mais ampla: que perguntas podem ser respondidas ao se conhecer o perfil de expressão de um gene? Entender o controle da expressão gênica é crítico para se compreender a relação entre genótipo e fenótipo (MANTIONE *et al.*, 2014). A expressão de um gene em determinado local e em determinado momento sugere que o seu produto, seja ele um RNA funcional ou uma proteína, é necessário para certo processo celular ou algum processo subsequente do desenvolvimento. Assim, conhecer o perfil de expressão de um gene revela aspectos importantes sobre a sua função. Ademais, a comparação de perfis de expressão permite responder perguntas ainda mais interessantes. Quais foram os processos evolutivos envolvidos? Qual é a relevância do contexto genômico? Em última instância: o que determina o perfil de expressão de um gene?

Um primeiro resultado do nosso trabalho foi descrever o perfil de expressão dos genes do cromossomo Y de *D. melanogaster*. Até o momento, apenas o perfil do *kl-5* era conhecido (GEPNER; HAYS, 1993). Nós verificamos uma extrema coesão dos perfis de expressão, pois 8 dos 12 genes ligados ao Y apresentam expressão crescente ao longo da espermatogênese, isto é: a expressão na meiose é maior do que na mitose e a expressão na pós-meiose é maior do que na meiose. Por muito tempo, a transcrição na pós-meiose de *Drosophila* foi tida como inexistente (OLIVIERI, G.; OLIVIERI, 1965) e postulava-se que os transcritos traduzidos durante a pós-meiose eram armazenados de fases anteriores. No entanto, nos trabalhos de Barreau *et al.* (2008) e Vibriantovski *et al.* (2010), foi demonstrado pela primeira vez evidências de transcrição pós-meiótica. Nós não apenas confirmamos esses resultados, mas verificamos também que genes pós-meióticos parecem ser a regra para cromossomo Y de *D. melanogaster*. Cabe ressaltar que, em experimentos de transcriptoma, se a quantidade de mRNA de um gene é significativamente maior do que na fase anterior do desenvolvimento celular

(como no caso da pós-meiose em relação à meiose), a interpretação mais parcimoniosa é que houve transcrição gênica após a fase anterior.

Se a maior atividade transcricional ocorre na pós-meiose, podemos inferir que as proteínas codificadas pelos genes ligados ao Y são usadas nas fases mais tardias da espermatogênese. Para os três genes cujas funções são conhecidas (*kl-2*, *kl-3* e *kl-5*), tal inferência procede pois eles codificam cadeias pesadas da dineína que são essenciais para o batimento flagelar (CARVALHO; LAZZARO; CLARK, 2000; GEPNER; HAYS, 1993). Para os outros genes, o pouco que se conhece baseia-se em similaridades de domínios de proteínas. Assim, nosso trabalho dá um primeiro passo na direção de se compreender suas funções ao mostrar em quais fases da espermatogênese a transcrição é mais relevante. De fato, é na pós-meiose que ocorre o amadurecimento dos espermatozóides e seria esperado que falhas em genes que afetam esse processo levem à infertilidade masculina.

Em acordo com os nossos resultados, dados de *single cell* RNAseq de testículos de larvas mostram uma baixa expressão do Y em células somáticas e espermatogônias e aumento de expressão em espermatócitos primários (MAHADEVARAJU; FEAR *et al.*, 2021). Em um trabalho recente (GÄRTNER *et al.*, 2019), foram construídos proteomas de testículo de *D. melanogaster* para várias fases do desenvolvimento, incluindo larvas, pupas e adultos. Das proteínas codificadas por genes ligados ao Y, apenas a correspondente ao *FDY* apresentou um enriquecimento nos testículos de larvas (discutido a seguir). Nos outros casos analisados (*ARY*, *ORY*, *Pp1-Y2*, *Ppr-Y*, *WDY*, *kl-2*, *kl-3* e *kl-5*), as proteínas se acumularam ao longo do desenvolvimento com maior abundância nas fases de pupa e adultos. Esses resultados corroboram as nossas observações da importância dos genes do cromossomo Y nas fases mais tardias da espermatogênese. Além disso, combinados à nossa observação da expressão crescente, eles mostram que a tradução dos genes do cromossomo Y se dá nas fases tardias da espermatogênese.

A notável exceção ao perfil de expressão crescente discutido até agora é o gene *FDY*. Ele foi o único cuja expressão decresce ao longo da espermatogênese. Trata-se do gene mais recentemente duplicado para o cromossomo Y de *D. melanogaster* (CARVALHO *et al.*, 2015). Diferentemente dos outros genes ligados ao Y, ele resguarda alta semelhança com seu parental autossômico *vig2* e a comparação mostrou que ambos têm o mesmo perfil de expressão. A situação única do *FDY* revela o que ocorre com um gene logo após a sua duplicação para o Y: ele ainda retém o perfil de expressão de seu parental autossômico. As comparações dos outros genes do cromossomo Y de *D. melanogaster* com seus parálogos/resquírios autossômicos foram inconclusivas. Várias das duplicações de autossomos para o Y são antigas e mesmo nos casos em que ainda é possível reconhecer um parálogo autossômico, os genes podem ser muito divergentes. Portanto, eles possivelmente não recapitulam o perfil de expressão ancestral dos genes do cromossomo Y. Neste sentido, fazer a comparação com as outras espécies torna-se necessário.

Os genes ligados ao Y de *D. willistoni* e *D. mojavensis* apresentam perfis de expressão mais variados do que os de *D. melanogaster*. Nos dois casos, um perfil bastante típico foi aquele em que a expressão aumenta na meiose em relação à mitose e permanece com o mesmo nível na pós-meiose. Além dele, em *D. willistoni*, observamos vários casos em que a expressão cresce na meiose e decresce na pós-meiose. Já em *D. mojavensis*, o outro perfil mais comum é o de expressão crescente (como em *D. melanogaster*). Uma característica comum entre as três espécies é o aumento de expressão da meiose em relação à mitose. De um total de 49 genes (12 de *D. melanogaster*, 23 de *D. willistoni* e 14 de *D. mojavensis*), o aumento significativo da mitose para a meiose foi observado em 44 casos.

A transição da meiose para a pós-meiose não apresenta um padrão claro, sendo que as três possibilidades (aumento, manutenção e diminuição da expressão) ocorrem em *D. willistoni* e *D. mojavensis*. Nós avaliamos se as diferenças de comportamento na pós-meiose observadas entre as espécies

poderia ser um viés introduzido pelo ponto de corte da pós-meiose durante as dissecações. Para isso, comparamos as frequências dos perfis de expressão considerando todos os genes (e não apenas os do cromossomo Y) pois elas deveriam ser semelhantes entre as espécies. De fato, os perfis mais comuns são os mesmos em *D. melanogaster* e *D. mojavensis*. No entanto, em *D. willistoni*, os perfis de expressão que diminuem/aumentam de expressão na pós-meiose em relação à meiose parecem estar subrepresentados. Esses resultados indicam que os nossos dados de *D. willistoni* podem ter excluído a porção inicial da pós-meiose mas, caso o corte tivesse sido padronizado em uma posição mais proximal do testículo, ele incluiria células típicas da fase meiótica. Ou seja, talvez não seja possível separar as porções enriquecidas com células meióticas das porções enriquecidas com células pós-meióticas em *D. willistoni* de uma maneira tão precisa como é feito em *D. melanogaster*.

Em conjunto, os resultados dos perfis de expressão dos genes do cromossomo Y das três espécies sugerem que existe uma coerência nos perfis de expressão dos genes ligados ao Y. Para cada uma das três espécies, predominam um ou dois perfis de expressão, sendo que todos eles têm em comum o aumento de expressão da mitose para a meiose. Esses resultados nos permitem descartar a primeira hipótese levantada, segundo a qual não haveria influência da ancestralidade ou do contexto genômico na determinação dos perfis de expressão dos genes do cromossomo Y. Pelo menos um dos dois processos parece estar operando e a comparação de genes que são ligados ao Y em uma espécie com seus ortólogos autossômicos em uma segunda espécie nos permite distinguir entre elas.

2. A evolução dos perfis de expressão

O cromossomo Y apresenta algumas peculiaridades em relação aos outros cromossomos que podem ser relevantes para a determinação do perfil de expressão de seus genes, como a formação dos *loops* nos espermatozoides primários. Esse é um fenômeno associado à presença de introns gigantes nos

genes do *kl-5*, *kl-3* e *ks-1* (BONACCORSI *et al.*, 1988). Nós evidenciamos que em *D. melanogaster*, a formação dos *loops* ocorre nos espermatócitos intermediários/tardios, ao passo que o Y encontra-se condensado nos espermatócitos iniciais. Em *D. pseudoobscura*, os *loops* também são observados, mas eles se originam de diferentes introns ligados ao Y (CHANG; LARRACUENTE, 2017). Como mencionado anteriormente, no grupo *obscura*, o cromossomo Y ancestral da linhagem foi completamente translocado para um autossomo (CARVALHO; CLARK, 2005). Portanto, a formação dos *loops* parece estar associada a características específicas de genes do cromossomo Y e não a genes específicos. A respeito, Fingerhut, Moran e Yamashita (2019) demonstraram por meio de um engenhoso desenho experimental que a transcrição de *kl-3* e *kl-5* em *D. melanogaster* se dá ao longo do desenvolvimento do espermatócito (~90 horas) e sugerem que a presença de introns gigantes adiciona uma camada a mais na regulação gênica. Seguindo essa linha de raciocínio, estar no cromossomo Y favorece o acúmulo de sequências heterocromáticas que, por sua vez, requerem um programa de transcrição específico.

Nós verificamos que um padrão comum a todos os genes do cromossomo Y é o aumento de expressão da mitose para a meiose. O mesmo foi observado para os ortólogos autossômicos/ligados ao X, sugerindo que o aumento de expressão da mitose para a meiose deve ser um caráter ancestral que foi selecionado entre os genes duplicados para o cromossomo Y. Levando em consideração nossos experimentos de citogenética sobre o comportamento do Y na meiose, um gene com pico de expressão na mitose que fosse duplicado para o Y provavelmente não seria selecionado já que o cromossomo encontra-se condensado nessa fase. Por outro lado, um gene com expressão na meiose que fosse duplicado para o Y, poderia ter sua expressão potencializada nesse cromossomo. Esses resultados corroboram a hipótese da pré-adaptação, que previa que genes ligados ao Y possuiriam perfis de expressão semelhantes aos seus ortólogos autossômicos/ligados ao X.

Sabe-se que a taxa de ganho de genes no cromossomo Y é maior do que a taxa de perda (CARVALHO; KOERICH; CLARK, 2009). No entanto, o número de genes do cromossomo Y conhecido está na casa de duas dezenas e provavelmente restam (caso restem) poucos a serem descritos (VIBRANONSKI; KOERICH; CARVALHO, 2008). Neste cenário, estamos diante de um paradoxo: se genes são constantemente duplicados para o Y, qual é o destino deles? Um gene duplicado para o Y cuja expressão decresce na meiose deve ter menor probabilidade de fixação em um cromossomo cuja função primordial está relacionada à espermatogênese. Por outro lado, um gene cuja expressão ancestral era de aumento de expressão na meiose tem maior probabilidade de fixação no Y, especialmente se ele conferir alguma vantagem à fertilidade masculina. O aumento de expressão da mitose para a meiose parece ser um caráter ancestral e ter sido selecionado nos genes que foram duplicados para o Y. Se esse for o caso, o destino evolutivo do *FDY* é de particular interesse.

As comparações da expressão na meiose com a pós-meiose não apresentam um padrão tão claro. Além disso, a subrepresentação das classes de expressão que diminuem/aumentam de expressão na pós-meiose em relação à meiose em *D. willistoni* dificultam conclusões que levem em conta os ortólogos dessa espécie. Alternativamente, nós analisamos os perfis de expressão de genes que são autossômicos/ligados em X em *D. melanogaster* e ligados ao Y em *D. willistoni* e/ou *D. mojavensis*. Esse conjunto de 20 genes é enriquecido com perfis de expressão que aumentam ao longo da espermatogênese, não apenas da mitose para a meiose, mas também da meiose para a pós-meiose. No entanto, o enriquecimento de genes com aumento de expressão na pós-meiose só foi observado para o cromossomo Y de *D. melanogaster*. Assim, ainda é preciso esclarecer se o aumento de expressão da meiose para a pós-meiose é de fato um caráter ancestral compartilhado entre genes do cromossomo Y e seus ortólogos autossômicos/X-ligados ou, se o comportamento na pós-meiose é mais flexível podendo adquirir novos perfis de expressão a partir de sua duplicação para o Y.

Essa maior variação no comportamento da pós-meiose também foi evidenciada nos experimentos com as linhagens com Y introgrado. Os perfis individuais dos genes do cromossomo Y mostram que os genes de Y-sec têm uma tendência para diminuição da expressão na pós-meiose quando comparados aos de Y-sim. As análises de reamostragens sugerem que essas diferenças são estendidas para genes dos outros cromossomos, o que pode ser explicado pelo fato do cromossomo Y ser capaz de modificar a expressão de genes autossômicos e ligados ao X (JIANG; HARTL; LEMOS, 2010). Considerando que a única distinção entre Y-sim e Y-sec são seus cromossomos Y, as diferenças observadas devem ser explicadas por modificações epigenéticas entre as linhagens. Elas podem ser relevantes para explicar as diferenças nos comportamentos de expressão na pós-meiose observadas entre as espécies.

3. Y como modelo para evolução de genes novos

Genes novos possuem viés de expressão nos testículos e são definidos como genes presentes em todos os membros de um grupo monofilético mas ausentes de todos os grupos externos (LONG *et al.*, 2013). Kaessmann (2010) discute a hipótese de que os testículos têm um papel-chave no estabelecimento de novos genes funcionais e serviriam como uma espécie de incubadora para sua evolução. Um trabalho recente do nosso grupo (RAICES; OTTO; VIBRANOVSKI, 2019) propôs a seleção haplóide como processo evolutivo responsável por esse fenômeno. O estudo de genes do cromossomo Y pode trazer contribuições relevantes para a evolução de genes novos já que eles também possuem um viés de expressão em testículos.

Além disso, de acordo com a hipótese não-canônica de evolução do Y, ele teria se originado a partir de um cromossomo B e seus genes seriam oriundos a de duplicações autossômicas (CARVALHO, 2002). Há 50 anos, Ohno (1970) propôs que a duplicação gênica seria o principal mecanismo da evolução. Na ausência de duplicação gênica, a plasticidade do genoma em se adaptar a novos contextos ambientais seria limitada (ZHANG, 2003). Assim, o cromossomo Y

de *Drosophila* oferece um sistema único para se estudar a evolução de genes novos originados a partir de duplicações.

O *FDY*, por exemplo, é um gene novo de *D. melanogaster* que apresenta um perfil de expressão oposto aos outros genes do cromossomo Y. Se o aumento de expressão na meiose for realmente um crivo da seleção natural, o destino provável do *FDY* seria a pseudogenização. Alternativamente, pode existir outra vantagem adaptativa na manutenção desse gene no cromossomo Y. Linhagens mutantes de *vig2* apresentam um nível diminuído de H3K9me2, uma metilação que está fortemente associada à repressão transcricional (GRACHEVA; DU; ELGIN, 2009). Considerando que *vig2* é expresso em machos e fêmeas, pode ser que *FDY* tenha função semelhante a de seu homólogo. Neste cenário, estar localizado no cromossomo Y pode ser uma maneira automática de se regular a expressão de DNAs próximos de maneira testículo-específica, reprimindo na mitose e liberando o aumento da expressão de outros genes nas fases tardias da espermatogênese.

Os resultados de Y-sim são particularmente relevantes para essa hipótese, pois a única diferença entre os cromossomos Y de *D. simulans* e *D. melanogaster* é a presença do *FDY* na última. Apesar disso, os perfis de expressão dos genes do cromossomo Y são diferentes entre as linhagens. Em *D. melanogaster*, o perfil mais comum é o de aumento de expressão ao longo da espermatogênese, enquanto que em Y-sim, o perfil mais comum é o de aumento de expressão na meiose em relação à mitose e manutenção da expressão na pós-meiose. Se a presença do *FDY* no cromossomo Y realmente modular a expressão de genes próximos, as diferenças de comportamento na pós-meiose entre *D. melanogaster* e Y-sim poderiam ser explicadas.

4. Conclusões e perspectivas

Em um sentido amplo, esse trabalho deixa como legado dados de expressão estágio-específico de testículos de *Drosophila* para cinco linhagens diferentes: *D. melanogaster*, *D. willistoni*, *D. mojavensis*, Y-sim e Y-sec.

A comparação dos perfis de expressão entre genes do cromossomo Y e seus ortólogos revelou que o aumento de expressão da mitose para a meiose é um caráter ancestral. Os resultados corroboram a hipótese da pré-adaptação em que genes com determinados perfis de expressão seriam selecionados e mantidos no cromossomo Y. A expressão da meiose para a pós-meiose apresenta perfis mais variados e espécie-específico (em cada espécie, predominam poucos perfis de expressão). Encontramos evidências em *D. melanogaster* de que o aumento de expressão da meiose para a pós-meiose pode também constituir um padrão ancestral, mas comparações com outras espécies podem ser necessárias para se entender se o aumento na pós-meiose é um padrão típico de *D. melanogaster* ou se estende para outras espécies também.

A análise citogenética do comportamento do Y na meiose, nos permitiu concluir que a formação dos *loops* se dá exatamente na transição de espermatócito inicial para espermatócito intermediário. Interessantemente, ela coincide com a fase em que é observada um aumento de expressão dos genes ligados ao Y. Adicionalmente, verificamos que a RNA-Pol II Ser2p não se sobrepõe de maneira alguma às sondas do cromossomo Y. Esses resultados poderiam sugerir que a transcrição nos espermatócitos se dá por uma via alternativa. Essas conclusões, no entanto, são limitadas às porções heterocromáticas do cromossomo Y detectadas pela sonda. Para torna-las mais generalizáveis, o emprego de sondas de *oligopaints* (BELIVEAU *et al.*, 2012) que recubram o cromossomo Y por completo pode ser promissor.

O gene *FDY* é um caso de particular interesse por ser um gene novo de *D. melanogaster* e apresentar perfil de expressão contrário a todos os outros genes do cromossomo Y. Um estudo recente (ZHANG *et al.*, 2020) empregou CRISPR/Cas9 e RNA de interferência para gerar mutantes de genes do cromossomo Y e verificar quais deles são de fato essenciais para a fertilidade masculina. Verificar o perfil de expressão dos genes do cromossomo Y na linhagem nocauteada para o *FDY* pode ajudar a esclarecer se ele é responsável por modular a expressão de outros genes.

Referências bibliográficas

- ÁDÁM, C. *et al.* Conservation of male-specific expression of novel phosphoprotein phosphatases in *Drosophila*. *Development genes and evolution*, v. 220, n. 3-4, p. 123-128, 2010.
- ADAMS, M. D. *et al.* The genome sequence of *Drosophila melanogaster*. *Science*, v. 287, n. 5461, p. 2185-2195, 2000.
- ANDREWS, S. *FastQC: a quality control tool for high throughput sequence data*, 2010. Disponível em: <<http://www.bioinformatics.babraham.ac.uk/projects/fastqc>>. Acesso em: 25 de out. de 2019.
- BALDO, L. *et al.* Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Applied and environmental microbiology*, v. 72, n. 11, p. 7098-7110, 2006.
- BARREAU, C. *et al.* Post-meiotic transcription in *Drosophila* testes. *Development*, v. 135, n. 11, p. 1897-1902, 2008.
- BELIVEAU, B. J. *et al.* Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proceedings of the National Academy of Sciences*, v. 109, n. 52, p. 21301-21306, 2012.
- BENJAMINI, Y.; HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*, v. 57, n. 1, p. 289-300, 1995.
- BONACCORSI, S. *et al.* Y chromosome loops in *Drosophila melanogaster*. *Genetics*, v. 120, n. 4, p. 1015-1034, 1988.
- BRIDGES, C. B. Non-disjunction as Proof of the Chromosome Theory of Inheritance. *Genetics*, v. 1, n. 1, 1916.
- BROSSEAU, G. E. Genetic analysis of the male fertility factors on the Y chromosome of *Drosophila melanogaster*. *Genetics*, v. 45, n. 3, p. 257, 1960.
- CABRAL-DE-MELLO, D. C.; MOURA, R. C.; MARTINS, C. Chromosomal mapping of repetitive DNAs in the beetle *Dichotomius geminatus* provides the first evidence for an association of 5S rRNA and histone H3 genes in insects, and repetitive DNA similarity between the B chromosome and A complement. *Heredity*, v. 104, n. 4, p. 393, 2010.

CARVALHO, A. B.; LAZZARO, B. P.; CLARK, A. G. Y chromosomal fertility factors *kl-2* and *kl-3* of *Drosophila melanogaster* encode dynein heavy chain polypeptides. *PNAS*, v. 97, n. 24, p. 13239-13244, 2000.

CARVALHO, A. B. *et al.* Identification of five new genes on the Y chromosome of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, v. 98, n. 23, p. 13225-13230, 2001.

CARVALHO, A. B. Origin and evolution of the *Drosophila* Y chromosome. *Current opinion in genetics & development*, v. 12, n. 6, p. 664-668, 2002.

CARVALHO, A. B.; CLARK, A. G. Y chromosome of *D. pseudoobscura* is not homologous to the ancestral *Drosophila* Y. *Science*, v. 307, n. 5706, p. 108-110, 2005.

CARVALHO, A. B.; KOERICH, L. B.; CLARK, A. G. Origin and evolution of Y chromosomes: *Drosophila* tales. *Trends in Genetics*, v. 25, n. 6, p. 270-277, 2009.

CARVALHO, A. B.; CLARK, A. G. Efficient identification of Y chromosome sequences in the human and *Drosophila* genomes. *Genome research*, v. 23, n. 11, p. 1894-1907, 2013.

CARVALHO, A. B. *et al.* Birth of a new gene on the Y chromosome of *Drosophila melanogaster*. *Proceedings of the National Academy of Sciences*, v. 112, n. 40, p. 12450-12455, 2015.

CHANG, C. H.; LARRACUENTE, A. M. Genomic changes following the reversal of a Y chromosome to an autosome in *Drosophila pseudoobscura*. *Evolution*, v. 71, n. 5, p. 1285-1296, 2017.

CHARLESWORTH, B.; CHARLESWORTH, D. The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, v. 355, n. 1403, p. 1563-1572, 2000.

CLINE, T. W. Two closely linked mutations in *Drosophila melanogaster* that are lethal to opposite sexes and interact with daughterless. *Genetics*, v. 90, n. 4, p. 683-697, 1978.

CHROSTEK, E. *et al.* *Wolbachia* variants induce differential protection to viruses in *Drosophila melanogaster*: a phenotypic and phylogenomic analysis. *PLoS genetics*, v. 9, n. 12, p. e1003896, 2013.

- CIVETTA, A.; GAUDREAU, C. Hybrid male sterility between *Drosophila willistoni* species is caused by male failure to transfer sperm during copulation. *BMC evolutionary biology*, v. 15, n. 1, p. 75, 2015.
- CLARK, M. E. *et al.* The distribution and proliferation of the intracellular bacteria *Wolbachia* during spermatogenesis in *Drosophila*. *Mechanisms of development*, v. 111, n. 1-2, p. 3-15, 2002.
- CLARK, M. E. *et al.* Induced paternal effects mimic cytoplasmic incompatibility in *Drosophila*. *Genetics*, v. 173, n. 2, p. 727-734, 2006.
- DROSOPHILA 12 GENOMES CONSORTIUM *et al.* Evolution of genes and genomes on the *Drosophila* phylogeny. *nature*, v. 450, n. 7167, p. 203, 2007.
- FINGERHUT, J. M.; MORAN, J. V.; YAMASHITA, Y. M. Satellite DNA-containing gigantic introns in a unique gene expression program during *Drosophila* spermatogenesis. *PLoS genetics*, v. 15, n. 5, p. e1008028, 2019.
- FU, L. *et al.* CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*, v. 28, n. 23, p. 3150-3152, 2012.
- GALLETTA, B. J. *et al.* Asterless is required for centriole length control and sperm development. *Journal of Cell Biology*, v. 213, n. 4, p. 435-450, 2016.
- GÄRTNER, S. M. K. *et al.* Stage-specific testes proteomics of *Drosophila melanogaster* identifies essential proteins for male fertility. *European Journal of Cell Biology*, v. 98, n. 2-4, p. 103-115, 2019.
- GEPNER, J.; HAYS, T. S. A fertility region on the Y chromosome of *Drosophila melanogaster* encodes a dynein microtubule motor. *Proceedings of the National Academy of Sciences*, v. 90, n. 23, p. 11132-11136, 1993.
- GRACHEVA, E.; DUS, M.; ELGIN, S. C. *Drosophila* RISC Component VIG and Its Homolog Vig2 Impact Heterochromatin Formation. *PLoS ONE*, v. 4, n. 7, p. e6182, 2009.
- HACKSTEIN, J. H. P. *et al.* Is the Y chromosome of *Drosophila* an evolved supernumerary chromosome?. *BioEssays*, v. 18, n. 4, p. 317-323, 1996.
- HARDY, R. W.; TOKUYASU, K. T.; LINDSLEY, D. L. Analysis of spermatogenesis in *Drosophila melanogaster* bearing deletions for Y-chromosome fertility genes. *Chromosoma*, v. 83, n. 5, p. 593-617, 1981.

- HIRST, J.; CARMICHAEL, J. A potential role for the clathrin adaptor GGA in *Drosophila* spermatogenesis. *BMC cell biology*, v. 12, n. 1, p. 22, 2011.
- HUGHES, J. F.; PAGE, D. C. The Biology and Evolution of Mammalian Y Chromosomes. *Annu. Rev. Genet.*, v. 49, p. 22.1-22.21, 2015.
- JIANG, P.; HARTL, D. L.; LEMOS, B. Y not a dead end: epistatic interactions between Y-linked regulatory polymorphisms and genetic background affect global gene expression in *Drosophila melanogaster*. *Genetics*, v. 186, n. 1, p. 109-118, 2010.
- KAESSMANN, H. Origins, evolution, and phenotypic impact of new genes. *Genome research*, v. 20, n. 10, p. 1313-1326, 2010.
- KENNISON, J. A. The genetic and cytological organization of the Y chromosome of *Drosophila melanogaster*. *Genetics*, v. 98, n. 3, p. 529-548, 1981.
- KOERICH, L. B. *et al.* Low conservation of gene content in the *Drosophila* Y chromosome. *Nature*, v. 456, n. 7224, p. 949-951, 2008.
- KOERICH, L.B. Baixa conservação do conteúdo gênico do cromossomo Y em 12 espécies de *Drosophila*. Tese (Doutorado em Genética) - Universidade Federal do Rio de Janeiro. Rio de Janeiro, p. 249. 2009.
- KOLDE, R. Package ‘pheatmap’. *R Package*, v. 1, n. 7, p. 790, 2015.
- KUHN, G. C. S. *et al.* The 1.688 repetitive DNA of *Drosophila*: concerted evolution at different genomic scales and association with genes. *Molecular biology and evolution*, v. 29, n. 1, p. 7-11, 2012.
- LEADER, D. P. *et al.* FlyAtlas 2: a new version of the *Drosophila melanogaster* expression atlas with RNA-Seq, miRNA-Seq and sex-specific data. *Nucleic acids research*, v. 46, n. D1, p. D809-D815, 2018.
- LENG, N. *et al.* EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics*, v. 29, n. 8, p. 1035-1043, 2013.
- LI, B.; DEWEY, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. 2011.
- LI, W.; GODZIK, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, v. 22, n. 13, p. 1658-1659, 2006.

- LLOYD, V. K. *et al.* A genetic and molecular characterization of the garnet gene of *Drosophila melanogaster*. *Genome*, v. 42, n. 6, p. 1183-1193, 1999.
- LONG, M. *et al.* New gene evolution: little did we know. *Annual review of genetics*, v. 47, p. 307-333, 2013.
- LOVE, M. I.; HUBER, W.; ANDERS, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, v. 15, n. 12, p. 550, 2014.
- MAGWIRE, M. M. *et al.* Genome-wide association studies reveal a simple genetic basis of resistance to naturally coevolving viruses in *Drosophila melanogaster*. *PLoS genetics*, v. 8, n. 11, p. e1003057, 2012.
- MAHADEVARAJU, S.; FEAR, J. M. *et al.* Dynamic sex chromosomes expression in *Drosophila* male germ cells. *Nature Communications*. 2021. (a ser publicado)
- MANTIONE, K. J. *et al.* Comparing bioinformatic gene expression profiling methods: microarray and RNA-Seq. *Medical science monitor basic research*, v. 20, p. 138, 2014.
- MISTRY, M.; KHETANI, R.; PIPER, M. *Differential Gene Expression workshop: count normalization with DESeq2*, 2017. Disponível em: <https://github.com/hbctraining/DGE_workshop/blob/master/lessons/02_DGE_count_normalization.md>. Acesso em: 15 de ago. de 2020.
- OHNO, S. *Evolution by gene duplication*. Nova Iorque: Springer-Verlag, 2013. 150 p.
- OLIVIERI, G.; OLIVIERI, A. Autoradiographic study of nucleic acid synthesis during spermatogenesis in *Drosophila melanogaster*. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, v. 2, n. 4, p. 366-380, 1965.
- PHATNANI, H. P.; GREENLEAF, A. L. Phosphorylation and functions of the RNA polymerase II CTD. *Genes & development*, v. 20, n. 21, p. 2922-2936, 2006.
- PINKEL, D. *et al.* Cytogenetic analysis by in situ hybridization with fluorescently labeled nucleic acid probes. *In: Cold Spring Harbor symposia on quantitative biology*. Cold Spring Harbor Laboratory Press, 1986. p. 151-157.

- RAICES, J. B.; OTTO, P. A.; VIBRANOVSKI, M. D. Haploid selection drives new gene male germline expression. *Genome Research*, v. 29, n. 7, p. 1115-1122, 2019.
- RICCHIO, J. P. R. O conteúdo gênico do cromossomo Y de *Drosophila willistoni*. Dissertação (Mestrado em Genética) - Universidade Federal do Rio de Janeiro. Rio de Janeiro. p. 122. 2020.
- RICHMOND, M. P.; JOHNSON, S.; MARKOW, T. A. Evolution of reproductive morphology among recently diverged taxa in the *Drosophila mojavensis* species cluster. *Ecology and Evolution*, v. 2, n. 2, p. 397-408, 2012.
- RIPARBELLI, M. G.; CALLAINI, G.; MEGRAW, T. L. Assembly and persistence of primary cilia in dividing *Drosophila* spermatocytes. *Developmental cell*, v. 23, n. 2, p. 425-432, 2012.
- ROBERTS, A.; PACHTER, L. Streaming fragment assignment for real-time analysis of sequencing experiments. *Nature methods*, v. 10, n. 1, p. 71-73, 2013.
- ROBINSON, M. D.; MCCARTHY, D. J.; SMYTH, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, v. 26, n. 1, p. 139-140, 2010.
- SACKTON, T. B. *et al.* Interspecific Y chromosome introgressions disrupt testis-specific gene expression and male reproductive phenotypes in *Drosophila*. *Proceedings of the National Academy of Sciences*, v. 108, n. 41, p. 17046-17051, 2011.
- SALZ, H.; ERICKSON, J. W. Sex determination in *Drosophila*: The view from the top. *Fly*, v. 4, n. 1, p. 60-70, 2010.
- SCHROEDER, A. *et al.* The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC molecular biology*, v. 7, n. 1, p. 3, 2006.
- SONESON, C.; DELORENZI, M. A comparison of methods for differential expression analysis of RNA-seq data. *BMC bioinformatics*, v. 14, n. 1, p. 91, 2013.
- SMIT, A. F. A.; HUBLEY, R.; GREEN, P. *RepeatMasker Open-3.0*. 1996–2004. Disponível em: <<http://www.repeatmasker.org>>. Acesso em: 27 de set. de 2020.

THERMO SCIENTIFIC. *Nucleic Acid - Thermo Scientific Nanodrop Spectrophotometers*, 2010. Disponível em: <<https://assets.thermofisher.com/TFS-Assets/CAD/manuals/ts-nanodrop-nucleicacid-olv-r2.pdf>>. Acesso em: 25 de out. de 2019.

THIBAUD-NISSEN, F. *et al.* Eukaryotic genome annotation pipeline. In: **The NCBI Handbook [Internet]**. 2 ed. National Center for Biotechnology Information, 2013.

THURMOND, J. *et al.* FlyBase 2.0: the next generation. *Nucleic acids research*, v. 47, n. D1, p. D759-D765, 2019.

TSAI, J.; YAN, R.; MCKEE, B. D. Homolog pairing and sister chromatid cohesion in heterochromatin in *Drosophila* male meiosis I. *Chromosoma*, v. 120, n. 4, p. 335-351, 2011.

VANDERLINDE, T. Genes do cromossomo Y e o isolamento em *Drosophila*. Dissertação (Mestrado em Genética) - Universidade Federal do Rio de Janeiro. Rio de Janeiro. p. 68. 2020.

VIBRANOVSKI, M. D.; KOERICH, L. B.; CARVALHO, A. B. Two new Y-linked genes in *Drosophila melanogaster*. *Genetics*, v. 179, n. 4, p. 2325-2327, 2008.

VIBRANOVSKI, M. D. *et al.* Stage-specific expression profiling of *Drosophila* spermatogenesis suggests that meiotic sex chromosome inactivation drives genomic relocation of testis-expressed genes. *PLoS Genet*, v. 5, n. 11, p. e1000731, 2009. Disponível em: <<https://mnlab.uchicago.edu/sppress/>>. Acesso em: 25 de out. de 2019

VIBRANOVSKI, M. D. *et al.* Direct evidence for postmeiotic transcription during *Drosophila melanogaster* spermatogenesis. *Genetics*, v. 186, n. 1, p. 431-433, 2010.

WILSON, A. R.; TOMPKINS, L. Sexual interactions between two distantly related *Drosophila* species, *D. melanogaster* and *D. willistoni* (Diptera: Drosophilidae). *Journal of insect behavior*, v. 10, n. 6, p. 771-781, 1997.

ZHANG, J. Evolution by gene duplication: an update. *Trends in ecology and evolution*, v. 18, n. 6, p. 292-298, 2003.

ZHANG, J. *et al.* The Role of Y Chromosome Genes in Male Fertility in *Drosophila melanogaster*. *Genetics*, v. 215, n. 3, p. 623-633, 2020.

ZHAO, J. *et al.* FlyTED: the *Drosophila* testis gene expression database. *Nucleic acids research*, v. 38, n. suppl_1, p. D710-D715, 2010.

ZHOU, W.; ROUSSET, F.; O'NEILL, S. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proceedings of the Royal Society of London B: Biological Sciences*, v. 265, n. 1395, p. 509-515, 1998.