1   SEXUAL CONFLICT OVER SURVIVAL IN TRINIDADIAN GUPPIES

3   Yuying Lin[1*], Iulia Darolti[1], Benjamin L. S. Furman[1], Pedro Almeida[2], Benjamin A. Sandkam[1], Felix

4   Breden[3], Alison E. Wright[4], Judith E. Mank[1,2*]

5   1 Department of Zoology and Biodiversity Research Centre, University of British Columbia,
6   Canada
7
8   2 Department of Genetics, Evolution and Environment, University College London, United
9   Kingdom
10
11   3 Department of Biological Sciences, Simon Fraser University, Canada
12
13   4 Department of Animal and Plant Sciences, University of Sheffield, United Kingdom
14
15   *Corresponding author E-mail: linyuying@zoology.ubc.ca

ABSTRACT

Sexual conflict over survival produces two distinct population genetic signatures. Fluctuating selection in males and females produces balancing selection. Additionally, at conception, allele frequencies are the same in males and females. However, loci with alleles that benefit the survival of one sex at some survival cost to the other should diverge over the course of a generation. We therefore expect that sexual conflict over survival would produce both signatures of allelic differentiation between the sexes and balancing selection. However, given the substantial mortality costs required to produce allelic differences between males and females, it remains unclear how many loci within the genome, if any at all, experience significant sexual conflict over survival. We assessed the genomes of 120 wild-caught guppies, which are expected to experience substantial predation- and pathogen-induced mortality. We identified a core list of 15 high confidence genes that show allelic differences between male and female adults. However, eight of these show evidence of having duplicated copies on the Y chromosome, suggesting that the male-specific region of the guppy Y chromosome may act as a hotspot for the resolution of conflict. We recovered just seven genes with significant male-female allelic differentiation without evidence of Y duplication, and these show elevated Tajima's D, consistent with balancing selection from sexual conflict. Only one of these seven genes, Puf60b, shows substantial intersexual F_{ST}. Puf60b has roles in cognition and the immune system, suggesting substantial ongoing, unresolved sexual conflict related to predator and pathogen avoidance strategies.

Intra-locus sexual conflict, where an allele increases the fitness of one sex at a cost to the other,

leaves distinct genomic signatures (Mank 2017). All forms of intra-locus sexual conflict are

expected to result in balancing selection, as alleles are selected for or against depending on

whether they are present in males or females (Foerster et al. 2007; Johnston et al. 2013;

Connallon and Clark 2014; Barson et al. 2015; Hawkes et al. 2016; Lonn et al. 2017; Ruzicka et

al. 2019). Balancing selection resulting from sexual conflict may be a major factor in the

maintenance of genetic polymorphisms within populations (Connallon and Clark 2014),

therefore, the proportion of loci subject to unresolved sexual conflict within the genome, as

well as the types of loci affected, may have important implications for a range of evolutionary

factors, such as the speed and genetic basis of adaptation.

However, many factors can produce balancing selection in addition to sexual conflict, and a

signature of balancing selection alone is not sufficient to identify conflict loci within the

genome. Numerous recent studies have used allelic differences between males and females

within a population (F_{ST}, D_{xy}, etc) as a way to infer sex-differences in viability or survival, and

therefore sexual conflict over mortality (Cheng and Kirkpatrick 2016; Lucotte et al. 2016;

Flanagan and Jones 2017; Dutoit et al. 2018; Wright et al. 2018; Wright et al. 2019).This

approach assumes that allele frequencies are the same in males and females at conception, but

diverge over the course of a generation for loci with alleles that benefit the survival of one sex

at some survival cost to the other. We would expect that sexual conflict over survival would

produce both a signature of allelic differentiation between the sexes and balancing selection,

and these two measures together offer a powerful approach to identify potentially sexually

antagonistic loci within the genome.

Intra-locus conflict over survival implies a significant mortality cost each generation. Modelling

and simulation methods (Bissegger et al. 2019; Kasimatis et al. 2019; Ruzicka et al. 2020)

suggest that the sex-specific mortality rates necessary to generate significant allelic differences

between the sexes within each generation are quite high for any one locus. This precludes the

presence of large numbers of sites subject to sexual conflict due to survival, as the associated

mortality load would simply be too great. Moreover, recent work has highlighted the potential

that many perceived allelic sex differences actually are the result of sequencing homology

between autosomal and sex-linked loci (Bissegger et al. 2019; Kasimatis et al. 2019; Ruzicka et

al. 2020). The Y chromosome may preferentially retain duplicates that play an important role in

male development or fitness (Bachtrog 2013; Carvalho et al. 2015), and the process of Y

duplication ironically offers a route to sexual conflict resolution even though it may lead to the

mistaken perception of unresolved conflict acting on autosomal loci.

It therefore remains unclear how common and pervasive sexual conflict over survival is in

natural populations, or what types of loci, if any, it is expected to target. Here we assess the

potential for intersexual F_{ST} based on a sample of 120 wild-caught adult guppies, which would

be expected to experience substantial predation- and pathogen-induced mortality. We

identified a core list of 15 high confidence genes that show evidence of allelic differences

between adult males and females. Of these, eight show evidence of Y duplicates, suggesting

that the Y chromosome in guppies may indeed act as a hotspot for the resolution of conflict.

We recovered just seven genes with evidence of male-female allelic differentiation without

evidence of Y duplication, suggesting the number of loci subject to sex-specific mortality

selection in wild guppies is relatively small. Importantly, these genes show elevated Tajima's D,

consistent with expectations for balancing selection due to sexual conflict. Six of the seven

genes exhibit modest levels of allelic differentiation, and only one, Puf60b, shows F_{ST} levels

consistent with high selection coefficients. Importantly, all high F_{ST} SNPs in this gene are

localized within one exon. Puf60b is broadly expressed, and has both neurological and immune

system functions (Low et al. 2017; Kew et al. 2020). This, in combination with the differences

between female and male guppies in predation risk and predator avoidance (Magurran and

Nowak 1991; Croft et al. 2006) as well as pathogen infection and avoidance (Stephenson et al.

2015), suggests that sexual conflict over survival may primarily target loci that play a role in

behavior, cognition and the immune system.

Results

We individually sequenced 60 male and 60 female wild-caught adults from three rivers in

Trinidad to ~30X coverage in males and ~20X coverage in females after filtering, trimming and

quality control (see Materials and Methods for further details). The greater read depth in males

was designed to aid detection of the Y chromosome, which is present in only one copy in each

genome.

Because the size of the non-recombining region of the sex chromosome varies across guppy

populations (Wright et al. 2017; Almeida et al. 2020), and sex chromosome divergence results

in allelic differences on the X and Y between males and females, we confined all our analyses to

autosomal sequence, although we present the X chromosome (Chromosome 12) here in figures

for comparison. From our sequencing data, we obtained 8,054,872 biallelic high quality filtered

autosomal SNPs. We further filtered these to 268,416 SNPs present in annotated autosomal

coding sequence. We focused on coding sequence for three reasons. First, non-coding

sequence is enriched for repetitive elements, and the accumulation of repetitive elements on

the ancestral guppy Y chromosome (Almeida et al. 2020) would lead to male-specific SNPs in

repetitive elements that could bias our estimates of Y duplications. Second, previous work has

suggested that unresolved sexual conflict occurs primarily in coding sequence (Ruzicka et al.

2019), and this fits with the substantial evidence that sexual conflict over gene regulation can

be relatively quickly resolved through sex-specific gene regulation (Kopp et al. 2000; Mank

2017; Wright et al. 2019). Finally, by focusing on functional coding sequence, we are able to

determine the potential functional role of genes that are subject to sexual conflict.

Intersexual F_{ST}

Due to the limited number of loci expected, and the low level of allelic divergence that is

tolerable due to the required mortality loads (Kasimatis et al. 2019), it is critical to minimize

false positives and to avoid potential genomic biases in allele frequency estimation. In order to

reduce the possibility of false positives, we used three independent methods to estimate

intersexual F_{ST}. We identified SNPs that were 1) in the top 1% of the autosomal F_{ST} distribution

2) were significant after permutation testing of samples (1000 replicates, P < 0.001) and 3)

showed significant differences in male and female allele frequency based on Fisher's exact test

(P < 0.001) (Supplementary Fig. 1). We identified 504 autosomal coding sequence SNPs that

were significant by all three of these measures (Supplementary Figs. 1 and 2), designated as

high intersexual F_{ST} SNPs.

Table 1. Summary statistics of sexually differentiated genes. Those values that are statistically
significant are in bold.

Gene Name	Chrom	Mean F_{ST}[1]	M:F read depth[2]	M:F SNP density[2]	Gene functions[3]
Puf60b	LG20	**0.1386**	0.9759	1.00	Poly-U binding factor, mutations result in neurological conditions
Spata6	LG4	**0.0645**	1.0186	1.00	Spermiogenesis
Syngap1a	LG11	**0.0363**	1.0558	1.00	development of cognition & proper synapse function
MARK4	LG13	**0.0259**	1.0911	1.00	microtubule organization in neuronal cells
ENSPREG13206	LG11	**0.0113**	0.9338	1.00	N/A
ENSPREG21185	LG2	**0.0108**	0.9848	1.00	N/A
ENSPREG12132	LG2	**0.0096**	1.0652	1.00	N/A
Olr1496	LG7	**0.1802**	**1.3854**	1.23	immune response & germline differentiation
Atat1	LG11	**0.1203**	**1.1219**	1.00	sperm flagellar function, microtubules, neuronal migration and maturation, embryogenesis
si:rp71-17i16.5	LG7	**0.0889**	**1.1817**	**3.35**	Phosphatidylinositol 3-kinase signalling, involved in immune, inflammatory and allergic responses
ENSPREG16391	LG14	**0.0462**	**2.0197**	0.96	N/A
Fezf1	LG6	**0.0407**	**1.5379**	1.01	Nervous system development, migration of gonadotropin-releasing hormone neurons
ENSPREG05754	LG15	**0.0244**	**1.6949**	0.99	N/A
ENSPREG18780	LG9	**0.0236**	**1.5771**	0.96	N/A
ENSPREG12441	LG7	**0.0176**	**1.1185**	1.04	N/A

144 1 Significance based on Wilcoxon's rank-sum test (P < 0.01), compared to autosomal average.
145 2 Significance based on Fisher's Exact Test (P < 0.05), relative to genome-wide M:F SNP density.
146 3 From Gene Ontology (Ashburner et al. 2000; Anon 2021) and Gene Cards (Stelzer et al. 2016).
147

To further reduce false positives, we identified autosomal coding sequence with ≥ 5 high

intersexual F_{ST} SNPs (Tobler et al. 2017; Bissegger et al. 2019), resulting in 15 sexually

differentiated genes. For each of the 15 sexually differentiated genes, average intersexual F_{ST}

significantly greater than the autosomal average (Table 1, Supplementary Fig. 3).

Estimates of intersexual F_{ST} and Tajima's D can be biased due to relatedness of individuals

within groups. Although our sampling design was balanced, with 10 females and 10 males

collected at two separate sites for each of three rivers across the island of Trinidad (see

Almeida et al. 2020 for details), we could not account for relatedness in wild-caught samples.

We therefore assessed pairwise kinship coefficients among all our samples using both KING

(Manichaikul et al. 2010) and NgsRelate (Korneliussen and Moltke 2015), as implemented in

ANGSD (Korneliussen et al. 2014). Neither method identified greater relatedness among male

or female samples (Supplementary Fig. 4).

Y duplications

When mapping whole-genome data to a female reference genome, sequence similarity

between the male-specific Y chromosome and the autosomes can lead to the perception of

allelic differences between males and females. Recent work has shown that that many genes

with allelic sex differences are in fact autosomal loci that are either have recent Y duplicates

(Bissegger et al. 2019; Mank et al. 2020) or otherwise display sequence homology to the Y

(Kasimatis et al. 2020) . We can use differences in M:F read depth to identify these genes (Hall

et al. 2013). Duplication of the complete coding sequence for an autosomal gene on the Y

chromosome would produce an average M:F read depth of 1.5 (three copies in males, two in

females), and subsequent tandem duplications on the Y would result in M:F read depth > 1.5.

Partial Y duplications, or a full duplication followed by significant differentiation, would result in

average M:F read depth > 1 and < 1.5.

Eight of our 15 sexually differentiated genes showed average M:F read depth significantly > 1

(Table 1, Supplemental Fig. 5) consistent with at least a partial duplication of the coding

sequence on the Y chromosome. Four of our eight Y-duplicated genes have annotated function,

three of which (Olr1496, Atat1 and Fezf1) have a known role in sexual differentiation, and two

(Olr1496 and si:rp71-17i16.5) with immune functions (Table 1).

Given sufficient evolutionary time, Y duplications will also accumulate male-specific SNPs

(Tobler et al. 2017), leading to a signal of elevated M:F SNP density. Of the eight sexually

differentiated genes with greater read depth in males, one showed significantly higher M:F SNP

density (Table 1), suggesting it is older than the remaining seven Y duplicated genes. This high

number of male-specific SNPs in this gene, si:rp71-17i16.5, accounts for the very high spike in

significant F_{ST} at the proximal end of Chromosome 7 (Supplemental Fig. 1), illustrating the

potential for duplications on the Y chromosome to affect M:F F_{ST} estimates of autosomal

paralogs.

Across our 15 sexually differentiated genes, we did not observe evidence of duplications to the

X chromosome, which would result in three copies in males and four in females and therefore

an average M:F read depth of 0.75

Sexual conflict over survival for autosomal genes

We identified seven sexually differentiated genes without evidence of Y duplication (Table 1),

six of which displayed relatively modest levels of intersexual F_{ST},(0.009-0.0645), consistent with

the expectation that large numbers of high intersexual F_{ST} loci would require unsustainable

levels of sex-specific mortality (Bissegger et al. 2019; Kasimatis et al. 2019; Ruzicka et al. 2020).

Genes subject to ongoing sexual conflict are expected to experience balancing selection, which

is often measured with Tajima's D (Cheng and Kirkpatrick 2016; Mank 2017; Wright et al. 2018;

Bissegger et al. 2019). We compared Tajima's D for these seven genes to the autosomal

average, as well as average Tajima's D for genes with known immune function, which are

known to be subject to high levels of balancing selection (Ferrer-Admetlla et al. 2008; Andrés et

al. 2009; Van Oosterhout 2009; Weedall and Conway 2010). Tajima's D for the seven sexually

differentiated genes without evidence of Y duplicates is significantly higher than the autosomal

average (Fig. 1), and is also significantly higher than that for immune genes. The combined

signatures of both significantly elevated intersexual F_{ST} and Tajima's D are consistent with these

genes experiencing sexual conflict over survival, with alleles benefiting the survival of one sex at

a survival cost to the other (Mank 2017; Wright et al. 2018).

Fig. 1. Distribution of Tajima's D among autosomal genes, immunity genes, and seven sexually
differentiated genes without significantly elevated M:F read depth. * indicates pair-wise
significantly elevated Tajima's D (Wilcoxon rank-sum test). *P < 0.01, **P < 0.001, ***P <
0.0001

One of these genes, Puf60b, exhibits substantially higher M:F F_{ST} (0.1386). In order to

understand the nature of sexual conflict at this locus, and to validate our results, we mapped

our intersexual F_{ST} SNPs to this coding sequence, identifying 8 high F_{ST} SNPs, all localized in exon

6 of this gene (Fig. 2A). We also verified the M:F read depth for these SNPs (Fig. 2B), which do

not show elevated read depth in males and is therefore inconsistent with Y duplication. Of the

eight high intersexual F_{ST} SNPs, seven are synonymous, and just one is non-synonymous (Fig.

2A). The single non-synonymous SNP results in an amino acid change from aspartic acid to

glutamic acid, two amino acids with broadly similar biochemical properties.

We also examined sex-differences in haplotype structure for this exon (Fig. 2C), recovering two

major haplotypes. Haplotype 1 includes the non-synonymous SNP for aspartic acid (Asp), and

Haplotype 2 contains the glutamic acid SNP (Glu). All rare haplotypes include Glu, and were

included as Haplotype 2. There are significantly more Haplotype 1 homozygotes among females

than expected, and Haplotype 1/Haplotype 2 heterozygotes are significantly under-represented

(P < 0.008 in both cases, based on χ^2, 1.d.f, Fig. 2D, Supplementary Table 1). The relative

survival for female heterozygotes is 0.315 (Fig. 2D, Supplementary Table 2). In contrast, there are significantly less Haplotype 1 homozygotes among males and significantly more Haplotype 1/Haplotype 2 heterozygotes (P < 0.0005 in both cases, χ^2, 1.d.f.). The relative inferred survival for Haplotype 1 homozygotes in males is just 0.158. In both sexes, we observed no Haplotype 2 homozygotes, although this is marginally non-significant when assessing the sexes separately (P = 0.061, χ^2, 1.d.f.), if we assume concordant selection against Haplotype 2 homozygotes, this becomes significant (P = 0.008, χ^2, 1.d.f.) due to the increased sample size.

Fig. 2. (A) All high F_{ST} SNPs localize in Exon 6 of Puf60b. Gene structure shown, with exons (grey boxes) and introns (blue lines). Within Exon 6, vertical black lines represent synonymous mutations, vertical red line represents non-synonymous mutation (Asp→ Glu). (B) F_{ST} and M:F read depth for all SNPs in Puf60b. Of 11 SNPs in Puf60b coding sequence, 8 high intersexual SNPs do not show evidence of elevated M:F read depth. Synonymous SNPs in black, non-synonymous SNP in red. (C) Haplotype network recovers two main haplotypes of Puf60b, with Haplotype 1 includes the aspartic acid, and Haplotype 2 contains glutamic acid SNP variant. Each circle represents one haplotype, and rare haplotypes all contain glutamic acid (Haplotype 2). The size of circle is proportion to the number of individuals belonging to that haplotype. Red and blue of each circle represent the number of female and male individuals of each haplotype, respectively. Bars show mutation steps between haplotypes. (D) Genotype frequencies and relative survival for males (blue) and females (red), compared to expected (grey). See Supplementary Tables 1 and 2 for further details.

In order to understand the mortality load these allele frequencies might impose, we calculated the proportion of individuals of each sex that would be expected to be lost due to selection

over survival based on Pub60b genotypes (Fig. 2D, Supplementary Table 2). Selection at Pub60b

would remove 39% of females and 54% of males at each generation. The frequencies of

Haplotype 1 and Haplotype 2 did not differ between High and Low Predation populations

(Supplementary Tables 3 and 4, Fisher's Exact Test, P = 0.36).

We next assessed whether we observe the evidence of sexual conflict over survival in Puf60b in

our lab population, which experiences no predation, as well as greatly reduced parasite load

and pathogens due to our biosafety protocols. To do this, we identified the Puf60b coding

sequence from previously collected RNA-Seq data for three families (Darolti et al. 2020).

Haplotype 2 was not present in any of our lab samples. We also used this data to quantify

expression in males and females, and found that Puf60b is expressed at a significantly higher

level in males (Supplementary Fig. 6).

Functional annotations

Of the seven genes without evidence of Y duplications, four of the seven genes have functional

annotations in the *Danio rerio* Gene Ontology (Ashburner et al. 2000; Anon 2021) or GeneCards

(Stelzer et al. 2016) databases. Three, Puf60b, Syngap1a and MARK4 have neurological

functions (Table 1). Puf60b is broadly expressed, and has additional immunological functions

(Kew et al. 2020).

Discussion

Using a conservative approach, we identified 504 coding sequence SNPs in the guppy

autosomal genome which showed significant differences in allele frequency between males and

females. We used these to identify 15 autosomal genes with significant average intersexual F_{ST}.

Our approach, based on the intersection of three statistical methods, reduces the likelihood of

false positives and results in a high confidence gene list of intersexual F_{ST} (Table 1,

Supplementary Figs. 1, 2, and 3). This list of genes can be used to assess the role of Y

duplication in resolving conflict and driving intersexual F_{ST}, and the role of sex differences in

mortality in sexual conflict of autosomal genes.

The Y chromosome in guppies as a locus for conflict resolution.

Over half (8 out of 15) of our high-confidence sexually differentiated genes showed evidence of

Y duplications based on elevated male-to-female read depth ratios (Table 1, Supplementary Fig.

4). Four of these genes have sex-specific functions based on Gene Ontology and Gene Cards

annotations, consistent with theory and empirical studies showing that Y chromosomes can

accumulate gene duplicates with male-specific functions (Carvalho et al. 2015; Mahajan and

Bachtrog 2017). Specifically, Olr1496 (Olfactory factor 1496-like gene) plays a role in multiple

aspects of germline development in *C. elegans* (Cho et al. 2007), Atat1 (Alpha-tubulin N-

acetyltransferase 1) plays a role in sperm flagellar functions (Yanai et al. 2020), and Fezf1 (FEZ

Family Zink Finger 1) plays an important role in the migration of gonadotropin-releasing

neurons (Damla Kotan et al. 2014).

Y chromosomes in general represent a unique genomic environment. Although Y chromosome

gene content degrades quickly following recombination suppression (Bachtrog 2013), several

recent studies indicate that older, large Y chromosomes, such as that in *Drosophila* (Koerich et

al. 2008; Carvalho et al. 2015; Tobler et al. 2017), sticklebacks (Bissegger et al. 2019) and

humans (Kasimatis et al. 2020), contain substantial numbers of genes duplicated from

autosomes. Duplications may represent an important mechanism of sex chromosome

divergence, and the Y chromosome may preferentially retain duplicates that play an important

role in male development or fitness (Bachtrog 2013; Carvalho et al. 2015), offering a hotspot for

sexual conflict resolution within the genome.

However, it remains unclear how common Y duplications of autosomal genes occur in younger

systems with a far smaller male-specific region. Recent work has suggested that the ancestral

non-recombining region of the Y chromosome in guppies is relatively small, spanning at most 5

Mb (Darolti et al. 2019; Almeida et al. 2020). Others have failed to recover evidence for even

this limited region of Y degeneration (Charlesworth et al. 2020; Fraser et al. 2020). Our findings

add further support for a region of recombination suppression across populations on Trinidad,

as only a male-specific region of the Y can explain the M:F read depth differences we observe.

Moreover, our work suggests that the guppy Y chromosome is dynamic with regard to gene

content, and acts as a hotspot for gene duplications with male-specific functions despite its

recent origin and small size. It is also possible that Y genes have duplicated to the autosomes,

which would also produce a pattern of increased male read depth, although this is arguably less

likely. Taken together, our work suggests that even homomorphic sex chromosomes may act as

a hotspot of sexual conflict resolution. Moreover, our results further emphasize the importance

of accounting for Y gene duplications in scans for M:F F_{ST}, as over half of our sexually

differentiated genes show evidence of Y duplication, rather than sex-specific mortality, as the

cause of allelic differences between the sexes.

These eight Y-duplicated genes are not present in our previous list of Y-linked genes based on

male-specific sequence (Almeida et al. 2020), or in other similar analyses (Fraser et al. 2020).

This is not surprising, as duplications, particularly if recent, will still retain substantial homology

to the autosomal copy and will not be detected when bioinformatically identifying sequence

that is unique to male genomes. Consistent with recent duplications and limited divergence,

seven of our eight Y-duplicated genes do not exhibit elevated average M:F SNP density (Table

1).

We have previously (Wright et al. 2018) noted evidence of intersexual F_{ST} in guppies (*Poecilia*

reticulata). This finding was curious given that the data from this earlier work derived from a

lab population, free of most of the pathogens and all the predators that would exacerbate sex-

differences in mortality and predation (Wright et al. 2018). It was not clear how much of this

signal, if any, was due to Y duplications. However, it is worth noting that elevated intersexual

F_{ST} was highest for genes with male-biased expression, as would be expected for genes with Y

duplicates. We also did not observe a concomitant pattern of elevated Tajima's D for these

genes, which is inconsistent with sexual conflict over survival.

Sexual conflict over survival targets neurological and immune functions

Intra-locus sexual selection over survival or viability leads to allele frequency differences

between the sexes over the course of a generation, as an allele increases the survival of one sex

at a mortality cost to the other (Mank 2017; Wright et al. 2018; Kasimatis et al. 2019). The

significant mortality costs required each generation to generate allele frequency differences

between the sexes preclude large numbers of genes with significant M:F F_{ST} (Bissegger et al.

2019; Kasimatis et al. 2019; Ruzicka et al. 2020). At most, we might expect a limited number of

loci with significant allelic differentiation between males and females, and this would be most

evidence in wild species where males and females experience differences in predation, parasite

or pathogen loads. Consistent with this, we observe just seven loci in the genome with high

confidence evidence of sexual conflict over survival based on M:F F_{ST}. Additionally, these seven

genes exhibit elevated levels of balancing selection (Fig. 1) as expected under sexual conflict

over survival. Six of these seven genes have quite modest F_{ST} levels (ranging from 0.0096 to

0.0645).

Puf60b F_{ST} is substantially higher (0.1386), and requires a mortality load of 39% in females and

54% in males. Although high, it is worth noting that survival is significantly greater in females,

which would allow for significant population growth despite the strong levels of selection.

These results were similar for both High and Low Predation populations, and suggest similar

selection coefficients across all populations in Trinidad. However, we did not detect Haplotype

2 in our lab colony, originally collected from the Quare River. This colony has been maintained

in order to maximize genetic diversity, and it may be that the loss of predation, and the

reduction in pathogen and parasite loads associated with our biosecurity protocols have

eliminated sexual conflict and therefore the selection for heterozygotes in males. However, we

also cannot rule out the possibility that Haplotype 2 has been lost due to genetic drift.

Recent work in salmon (Barson et al. 2015) has suggested that some sexually antagonistic loci

experience sex-specific dominance effects. This reduces the mortality load associated with

sexual conflict over survival. Our genotype data (Supplemental Tables 1 & 2) suggests that this

is not the case for Puf60b, as male heterozygotes have the highest survival relative to both

homozygous genotypes. This is inconsistent with sex-specific dominance.

Four of the seven fully autosomal sexually differentiated genes have cognitive or neurological

functions. Puf60b (Poly(U) binding splicing factor 60) is related to RNA-binding, mRNA

processing and RNA-splicing and mutations can result in neurological conditions (El Chehadeh

et al. 2016). Syngap1 (Synaptic Ras GTPase-activating protein 1), encodes a Ras GTPase

activating protein, regulating synaptic plasticity and neuronal homeostasis, and mutations

affect forebrain and cognitive development (Ozkan et al. 2014). Mark4 (microtubule affinity

regulating kinase 4) has been shown to be involved in microtubule organization in neuronal

cells (Trinczek et al. 2004). Puf60b has also been recently implicated as well in immune system

functions (Kew et al. 2020).

At first it may seem counterintuitive that sexual conflict over survival would preferentially

target neurological genes. However, in light of the differences between females and males in

predation risk and predator avoidance (Magurran and Nowak 1991; Croft et al. 2006) as well as

pathogen infection and avoidance (Stephenson et al. 2015). Our results suggest that sexual

conflict over survival may primarily target genes that play a role in behavior, cognition and the

immune system.

Concluding remarks

We found profound sexual conflict over survival at one locus in the genome, Puf60b, and

limited evidence at six additional loci in wild-caught guppy samples from across the island of

Trinidad. Our results highlight the limited but important role of sexual conflict over survival in

shaping patterns of genomic diversity. Additionally, we identified eight loci that have at least

partial Y homology, indicating the potential of the guppy Y chromosome to act as a hotspot of

sexual conflict resolution.

Materials and Methods

Data Collection and Genotyping

Samples were collected from three rivers, Aripo, Yarra, Quare, in Trinidad in December 2016, in

accordance with national collecting guidelines. In total, 10 males and 10 females were collected

from one high predation and one low predation population in each river, resulting in 120

samples, which were sequenced individually with Illumina HISEQX. Further sequencing details

are available in Almeida et al., (2020).

We used FastQC v0.11 (www.bioinformatics. babraham.ac.uk/projects/fastqc) and

Trimmomatic 0.36 (Bolger et al. 2014) to remove adapter sequences and low-quality reads.

After quality control, we recovered ~30X average sequencing depth for males and ~20X

sequencing depth for females. High quality reads were aligned against the *Poecilia reticulata*

female reference genome (Ensembl GCA_000633615.2) (Künstner et al. 2016), using BWA-MEM

(Li and Durbin 2009) with default parameters. We filled in mate coordinates and mate related

flags, sorted alignment by coordinates, and marked PCR duplications with SAMtools-1.9 (Li et

al. 2009).

We called genotypes across all the samples using the `mpileup` function from SAMtools-1.9

with the following parameters: --min-MQ 20 --min-BQ 20 --skip-indels -a FORMAT/AD,

FORMAT/DP. After genotyping, we used VCFtools v0.1.16 (Danecek et al. 2011) to exclude SNPs

that had either : (1) genotype quality < 20; (2) sequencing depth <0.5x or >10x of average

depth; (3) missing data in > 10% of individuals or (4) minor allele frequency < 0.05. In total, the

autosomal filtered SNP dataset consisted of 8,054,872 biallelic SNPs. We extracted 268,416

SNPs in annotated coding sequences (Ensembl build Guppy Female 1.0) for downstream

analysis. Finally, we confined our analysis to autosomal genes, but included the X chromosome

(Chromosome 12) in figures as a means of comparison.

Intersexual F_{ST}

In order to estimate intersexual allele frequency differences, we implemented Weir &

Cockerham's estimator of F_{ST} (Weir and Cockerham 1984) between males and females using

VCFtools v0.1.16 for each SNP in genome-wide coding sequence regions. We employed three

methods jointly to identify SNPs exhibiting high F_{ST}. First, we used a cut-off method, retaining

SNPs in only the top 1% of autosomal F_{ST} values. Second, we performed permutation tests by

randomly assigning individuals to one of two groups to generate a null distribution of F_{ST} across

the genome. We determined significance for each SNP from 1000 replicates, using a P < 0.001

threshold. Finally, we performed Fisher's exact test on SNPs to determine significance of allele

frequency differences between males and females (P < 0.001). We denoted SNPs that were

significant in all three of these measures as high F_{ST} SNPs.

Using approaches to further limit false positives (Tobler et al. 2017; Bissegger et al. 2019), we

identified 15 genes with ≥ 5 high intersexual F_{ST} SNPs, which we designated as sexually

differentiated genes. We calculated average intersexual F_{ST} for all genes using VCFtools v0.1.16,

respectively. We used Wilcoxon rank-sum test to indicate statistical difference in intersexual F_{ST}

between autosomal genes and other categorical genes (sexually differentiated genes and genes

on the sex chromosome).

Haplotype network

To further investigate allelic differences of Puf60b, we used SHAPEIT4 v4.2.0 (Delaneau et al.

2019) with default parameters to phase all the samples. We then reconstructed the haplotype

sequence of Puf60b for all samples. We used MEGAX (Kumar et al. 2018) for multiple sequence

alignment, and the inferred haplotype network using PopART v1.7 (Leigh JW and Bryant D

2015).

Relatedness Inference

In order to avoid biases in calculating intersexual allele frequency differences due to

relatedness, we used KING (Manichaikul et al. 2010) to infer the pairwise degree of relatedness

between individuals from estimated kinship coefficients. We first converted genotype data

from the raw, unfiltered SNPs dataset to plink binary format using PLINK (Purcell et al. 2007). In

order to avoid potential biases from KING software (Ramstetter et al. 2017) and validations, we

also used NgsRelate (Korneliussen and Moltke 2015), implemented in ANGSD (Korneliussen et

al. 2014) to infer genetic relatedness coefficients for each pair of individuals.

Assessing Y duplications of sexually differentiated genes

When using a female reference genome, reads from genes which have duplicates to the male-

specific region of the Y chromosome will map back to original autosomal or X chromosome

regions, resulting in elevated M:F coverage ratio (Bissegger et al. 2019; Mank et al. 2020). For

example, if an autosomal gene has one Y duplication, we would expect three copies in males

(two autosomal and one Y-linked) and two copies in females, and therefore an average M:F

read depth of 1.5. We first calculated M:F read depth for each coding sequence SNP from

genotype data, as male coverage/female coverage, correcting for differences in average

genomic read depth between males and females.

Genes on the Y chromosome will accumulate male-specific mutations over time, leading to an

increased number of male-specific mutations as well as elevated M:F SNP density (Bissegger et

al. 2019; Mank et al. 2020). M:F SNP density for each gene was calculated as number of male

SNPs /number of female SNPs.

In order to validate the M:F read depth ratio of sexually differentiated genes, we first extracted

coding sequence of our 15 sexually differentiated genes. We then calculated male and female

normalized read depth for each gene, based on male and female pooled reads. Additionally, we

calculated M:F read depth from our pools for each high FST SNP identified in reads mapping to

Puf60b.

Tajima's D

Based on the filtered genotype data, we calculated Tajima's D for the coding sequence of all

autosomal genes using VCFtools v1.19. We compared mean Tajima's D for autosomal genes

(minus those with immune function), genes with immune function (defined following Wright et

al. 2017, Wright et al. 2018), and our seven sexually differentiated genes.

Expression and diversity in lab colony

We used the *Poecilia reticulata* RNA-Seq data from Darolti et al., (2020) to first determine

differences in male and female expression for Puf60b. We first BLASTed the reference sequence

for Puf60b against our *de novo* transcriptome (see Darolti et al. 2019 for details), and

determined FPKM for male and female samples by mapping reads back to this contig. We also

mapped reads back to the Ensembl reference genome to determine isoform variation, and

examined this RNA-Seq data to determine whether we observe Haplotype 2 in our lab colony.

Functional Annotations

The small number of sexually differentiated genes precludes Gene Ontology enrichment

analysis. We therefore cataloged functional annotations from the Danio rerio Gene Ontology

(Ashburner et al. 2000; Anon 2021) and Gene Cards (Stelzer et al. 2016).

Acknowledgements

This was funded by a Chinese Scholarship Council Doctoral Scholarship to YL (No. 201906040216), a grant from the European Research Council (grant agreement 680951) and a Canada 150 Research Chair to JEM. We thank The Trinidad Ministry of Agriculture, Land and Fisheries and Indar Ramnarine for field permits. We thank Dolph Schluter for helpful discussions, and XXX for helpful comments and suggestions on the manuscript.

References Cited:

Almeida P, Sandkam BA, Morris J, Darolti I, Breden F, Mank JE. 2020. Divergence and Remarkable Diversity of the Y Chromosome in Guppies. *Mol. Biol. Evol.*:1–15.

Andrés AM, Hubisz MJ, Indap A, Torgerson DG, Degenhardt JD, Boyko AR, Gutenkunst RN, White TJ, Green ED, Bustamante CD, et al. 2009. Targets of balancing selection in the human genome. *Mol. Biol. Evol.*

Anon. 2021. The Gene Ontology resource: enriching a GOld mine. *Nucleic Acids Res.*

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. 2000. Gene ontology: Tool for the unification of biology. *Nat. Genet.*

Bachtrog D. 2013. Y-chromosome evolution: Emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* 14:113–124.

Barson NJ, Aykanat T, Hindar K, Baranski M, Bolstad GH, Fiske P, Jacq C, Jensen AJ, Johnston SE, Karlsson S, et al. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature* 528:405–408.

Bissegger M, Laurentino TG, Roesti M, Berner D. 2019. Widespread intersex differentiation across the stickleback genome – The signature of sexually antagonistic selection? *Mol. Ecol.*:1–10.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.

Carvalho AB, Vicoso B, Russo CAM, Swenor B, Clark AG. 2015. Birth of a new gene on the Y chromosome of Drosophila melanogaster. *Proc. Natl. Acad. Sci. U. S. A.*

Charlesworth D, Bergero R, Graham C, Gardner J, Yong L. 2020. Locating the sex determining region of linkage group 12 of guppy (Poecilia reticulata). *G3 Genes, Genomes, Genet.*

El Chehadeh S, Kerstjens-Frederikse WS, Thevenon J, Kuentz P, Bruel AL, Thauvin-Robinet C, Bensignor C, Dollfus H, Laugel V, Rivière JB, et al. 2016. Dominant variants in the splicing factor PUF60 cause a recognizable syndrome with intellectual disability, heart defects and short stature. In: European Journal of Human Genetics.

Cheng C, Kirkpatrick M. 2016. Sex-Specific Selection and Sex-Biased Gene Expression in Humans and Flies. *PLoS Genet.*

Cho S, Rogers KW, Fay DSS. 2007. The C. elegans Glycopeptide Hormone Receptor Ortholog, FSHR-1, Regulates Germline Differentiation and Survival. *Curr. Biol.*

Connallon T, Clark AG. 2014. Balancing selection in species with separate sexes: Insights from

fisher's geometric model. *Genetics*.

Croft DP, Morrell LJ, Wade AS, Piyapong C, Ioannou CC, Dyer JRG, Chapman BB, Wong Y, Krause J. 2006. Predation risk as a driving force for sexual segregation: A cross-population comparison. *Am. Nat.*

Damla Kotan L, Ian Hutchins B, Ozkan Y, Demirel F, Stoner H, Cheng PJ, Esen I, Gurbuz F, Kenan Bicakci Y, Mengen E, et al. 2014. Mutations in FEZF1 cause kallmann syndrome. *Am. J. Hum. Genet.*

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. *Bioinformatics*.

Darolti I, Wright AE, Mank JE. 2020. Guppy Y Chromosome Integrity Maintained by Incomplete Recombination Suppression. *Genome Biol. Evol.* 12:965–977.

Darolti I, Wright AE, Sandkam BA, Morris J, Bloch NI, Farré M, Fuller RC, Bourne GR, Larkin DM, Breden F, et al. 2019. Extreme heterogeneity in sex chromosome differentiation and dosage compensation in livebearers. *Proc. Natl. Acad. Sci. U. S. A.* 116:19031–19036.

Delaneau O, Zagury JF, Robinson MR, Marchini JL, Dermitzakis ET. 2019. Accurate, scalable and integrative haplotype estimation. *Nat. Commun.*

Dutoit L, Mugal CF, Bolívar P, Wang M, Nadachowska-Brzyska K, Smeds L, Yazdi HP, Gustafsson L, Ellegren H. 2018. Sex-biased gene expression, sexual antagonism and levels of genetic diversity in the collared flycatcher (Ficedula albicollis) genome. *Mol. Ecol.*

Ferrer-Admetlla A, Bosch E, Sikora M, Marquès-Bonet T, Ramírez-Soriano A, Muntasell A, Navarro A, Lazarus R, Calafell F, Bertranpetit J, et al. 2008. Balancing Selection Is the Main Force Shaping the Evolution of Innate Immunity Genes. *J. Immunol.*

Flanagan SP, Jones AG. 2017. Genome-wide selection components analysis in a fish with male pregnancy. *Evolution (N. Y).*

Foerster K, Coulson T, Sheldon BC, Pemberton JM, Clutton-Brock TH, Kruuk LEB. 2007. Sexually antagonistic genetic variation for fitness in red deer. *Nature*.

Fraser BA, Whiting JR, Paris JR, Weadick CJ, Parsons PJ, Charlesworth D, Bergero R, Bemm F, Hoffmann M, Kottler VA, et al. 2020. Improved Reference Genome Uncovers Novel Sex-Linked Regions in the Guppy (Poecilia reticulata). *Genome Biol. Evol.* 12:1789–1805.

Hall AB, Qi Y, Timoshevskiy V, Sharakhova M V., Sharakhov I V., Tu Z. 2013. Six novel Y chromosome genes in Anopheles mosquitoes discovered by independently sequencing males and females. *BMC Genomics*.

Hawkes MF, Gamble CE, Turner ECR, Carey MR, Wedell N, Hosken DJ. 2016. Intralocus sexual conflict and insecticide resistance. *Proc. R. Soc. B Biol. Sci.*

Johnston SE, Gratten J, Berenos C, Pilkington JG, Clutton-Brock TH, Pemberton JM, Slate J. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature*.

Kasimatis KR, Abraham A, Ralph PL, Kern AD, Capra JA, Phillips PC. 2020. Evaluating Human Autosomal Loci for Sexually Antagonistic Viability Selection in Two Large Biobanks. *Genetics*.

Kasimatis KR, Ralph PL, Phillips PC. 2019. Limits to genomic divergence under sexually antagonistic selection. *G3 Genes, Genomes, Genet.* 9:3813–3824.

Kew C, Huang W, Fischer J, Ganesan R, Robinson N, Antebi A. 2020. Evolutionarily conserved regulation of immunity by the splicing factor RNP-6/PUF60. *Elife* 9:1–23.

Koerich LB, Wang X, Clark AG, Carvalho AB. 2008. Low conservation of gene content in the Drosophila Y chromosome. *Nature*.

Kopp A, Duncan I, Carroll SB. 2000. Genetic control and evolution of sexually dimorphic
characters in Drosophilia. *Nature*.
Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation
Sequencing Data. *BMC Bioinformatics* 15:1–13.
Korneliussen TS, Moltke I. 2015. NgsRelate: A software tool for estimating pairwise relatedness
from next-generation sequencing data. *Bioinformatics*.
Kumar S, Stecher G, Li M, Knyaz C, Tamura K. 2018. MEGA X: Molecular evolutionary
genetics analysis across computing platforms. *Mol. Biol. Evol.*
Künstner A, Hoffmann M, Fraser BA, Kottler VA, Sharma E, Weigel D, Dreyer C. 2016. The
genome of the trinidadian guppy, Poecilia reticulata, and variation in the Guanapo
population. *PLoS One* 11:1–25.
Leigh JW, Bryant D. 2015. POPART: full-featrure software for haplotype network construction.
Methods Ecol. Evol.
Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform.
Bioinformatics.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.
2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*.
Lonn E, Koskela E, Mappes T, Mokkonen M, Sims AM, Watts PC. 2017. Balancing selection
maintains polymorphisms at neurogenetic loci in field experiments. *Proc. Natl. Acad. Sci.*
U. S. A.
Low KJ, Ansari M, Abou Jamra R, Clarke A, El Chehadeh S, Fitzpatrick DR, Greenslade M,
Henderson A, Hurst J, Keller K, et al. 2017. PUF60 variants cause a syndrome of ID, short
stature, microcephaly, coloboma, craniofacial, cardiac, renal and spinal features. *Eur. J.*
Hum. Genet.
Lucotte EA, Laurent R, Heyer E, Ségurel L, Toupance B. 2016. Detection of allelic frequency
differences between the sexes in humans: A signature of sexually antagonistic selection.
Genome Biol. Evol. 8:1489–1500.
Magurran AE, Nowak MA. 1991. Another battle of the sexes: The consequences of sexual
asymmetry in mating costs and predation risk in the guppy, Poecilia reticulata. *Proc. R. Soc.*
B Biol. Sci.
Mahajan S, Bachtrog D. 2017. Convergent evolution of y chromosome gene content in flies. *Nat.*
*Commun.*8
Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. 2010. Robust
relationship inference in genome-wide association studies. *Bioinformatics*.
Mank JE. 2017. Population genetics of sexual conflict in the genomic era. *Nat. Rev. Genet.*
18:721–730.
Mank JE, Shu JJ, Wright AE. 2020. Signature of sexual conflict is actually conflict resolved.
Mol. Ecol.
Van Oosterhout C. 2009. A new theory of MHC evolution: Beyond selection on the immune
genes. *Proc. R. Soc. B Biol. Sci.*
Ozkan ED, Creson TK, Kramár EA, Rojas C, Seese RR, Babyan AH, Shi Y, Lucero R, Xu X,
Noebels JL, et al. 2014. Reduced cognition in Syngap1 mutants is caused by isolated
damage within developing forebrain excitatory neurons. *Neuron*.
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, De
Bakker PIW, Daly MJ, et al. 2007. PLINK: A tool set for whole-genome association and
population-based linkage analyses. *Am. J. Hum. Genet.*

Ramstetter MD, Dyer TD, Lehman DM, Curran JE, Duggirala R, Blangero J, Mezey JG,
Williams AL. 2017. Benchmarking relatedness inference methods with genome-wide data
from thousands of relatives. *Genetics*.
Ruzicka F, Dutoit L, Czuppon P, Jordan CY, Li X, Olito C, Runemark A, Svensson EI, Yazdi
HP, Connallon T. 2020. The search for sexually antagonistic genes: Practical insights from
studies of local adaptation and statistical genomics. *Evol. Lett.* 4:398–415.
Ruzicka F, Hill MS, Pennell TM, Flis I, Ingleby FC, Mott R, Fowler K, Morrow EH, Reuter M.
2019. Genome-wide sexually antagonistic variants reveal long-standing constraints on
sexual dimorphism in fruit flies.
Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Iny Stein T, Nudel R,
Lieder I, Mazor Y, et al. 2016. The GeneCards suite: From gene data mining to disease
genome sequence analyses. *Curr. Protoc. Bioinforma.*
Stephenson JF, Van Oosterhout C, Mohammed RS, Cable J. 2015. Parasites of Trinidadian
guppies: Evidence for sex- and age-specific trait-mediated indirect effects of predators.
Ecology.
Tobler R, Nolte V, Schlötterer C. 2017. High rate of translocation-based gene birth on the
Drosophila Y chromosome. *Proc. Natl. Acad. Sci. U. S. A.* 114:11721–11726.
Trinczek B, Brajenovic M, Ebneth A, Drewes G. 2004. MARK4 Is a Novel Microtubule-
associated Proteins/Microtubule Affinity-regulating Kinase That Binds to the Cellular
Microtubule Network and to Centrosomes. *J. Biol. Chem.*
Weedall GD, Conway DJ. 2010. Detecting signatures of balancing selection to identify targets of
anti-parasite immunity. *Trends Parasitol.*
Weir BS, Cockerham CC. 1984. Estimating F-statistics for the analysis of population structure.
Evolution (N. Y).
Wright AE, Darolti I, Bloch NI, Oostra V, Sandkam B, Buechel SD, Kolm N, Breden F, Vicoso
B, Mank JE. 2017. Convergent recombination suppression suggests role of sexual selection
in guppy sex chromosome formation. *Nat. Commun.* 8.
Wright AE, Fumagalli M, Cooney CR, Bloch NI, Vieira FG, Buechel SD, Kolm N, Mank JE.
2018. Male-biased gene expression resolves sexual conflict through the evolution of sex-
specific genetic architecture. *Evol. Lett.* 2:52–61.
Wright AE, Rogers TF, Fumagalli M, Cooney CR, Mank JE. 2019. Phenotypic sexual
dimorphism is associated with genomic signatures of resolved sexual conflict. *Mol. Ecol.*
Yanai R, Yamashita Y, Umezu K, Hiradate Y, Hara K, Tanemura K. 2020. Expression and
localization of alpha-tubulin N-acetyltransferase 1 in the reproductive system of male mice.
J. Reprod. Dev
638