


Disrupted Gene Networks in Subfertile Hybrid House Mice

Katy Morgan ¹, Bettina Harr,² Michael A. White,³ Bret A. Payseur,⁴ and Leslie M. Turner*¹

¹Milner Centre for Evolution, Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom

²Department of Evolutionary Genetics, Max Planck Institute for Evolutionary Biology, Plön, Germany

³Department of Genetics, University of Georgia, Athens, GE

⁴Laboratory of Genetics, University of Wisconsin, Madison, WI

*Corresponding author: E-mail: l.m.turner@bath.ac.uk.

Associate editor: Yoko Satta

Abstract

The Dobzhansky–Muller (DM) model provides a widely accepted mechanism for the evolution of reproductive isolation: incompatible substitutions disrupt interactions between genes. To date, few candidate incompatibility genes have been identified, leaving the genes driving speciation mostly uncharacterized. The importance of interactions in the DM model suggests that gene coexpression networks provide a powerful framework to understand disrupted pathways associated with postzygotic isolation. Here, we perform weighted gene coexpression network analysis to infer gene interactions in hybrids of two recently diverged European house mouse subspecies, *Mus mus domesticus* and *M. m. musculus*, which commonly show hybrid male sterility or subfertility. We use genome-wide testis expression data from 467 hybrid mice from two mapping populations: F₂s from a laboratory cross between wild-derived pure subspecies strains and offspring of natural hybrids captured in the Central Europe hybrid zone. This large data set enabled us to build a robust consensus network using hybrid males with fertile phenotypes. We identify several expression modules, or groups of coexpressed genes, that are disrupted in subfertile hybrids, including modules functionally enriched for spermatogenesis, cilium and sperm flagellum organization, chromosome organization, and DNA repair, and including genes expressed in spermatogonia, spermatocytes, and spermatids. Our network-based approach enabled us to hone in on specific hub genes likely to be influencing module-wide gene expression and hence potentially driving large-effect DM incompatibilities. A disproportionate number of hub genes lie within sterility loci identified previously in the hybrid zone mapping population and represent promising candidate barrier genes and targets for future functional analysis.

Key words: speciation, Dobzhansky–Muller incompatibilities, gene networks, *Mus musculus*, hybrid sterility, reproductive isolation.

Introduction

According to the classic Dobzhansky–Muller (DM) model of speciation, mutations that accumulate independently and in different genomic regions may be incompatible when brought together in a hybrid background, resulting in disrupted epistasis and the development of postzygotic reproductive barriers (Muller 1942; Dobzhansky 1937). These barriers, which include reductions in hybrid fertility and/or hybrid viability, in turn restrict gene flow, enabling the further divergence of incipient species. The applicability of the DM model to allopatric speciation scenarios has been demonstrated through both theoretical and empirical studies (reviewed in Presgraves 2010). However, although many genes and loci influencing hybrid fertility have been described (reviewed in Maheshwari and Barbash 2011), characterizing the specific interactions between loci underpinning Dobzhansky–Muller incompatibilities (DMIs) remains a challenge.

Although DMIs were originally assumed to act independently of one another, such that individual loci are involved in a single DMI (Orr 1995), accumulating evidence raises questions about this assumption. A computational modeling study

based on RNA folding demonstrated that as two populations evolve in allopatry, DMIs become increasingly complex over time (Kalirad and Azevedo 2017). In this context, the majority of DMIs could involve more than two loci and individual loci are expected to participate in multiple DMIs (Kalirad and Azevedo 2017). Evidence from empirical studies in house mice (Turner and Harr 2014; Turner et al. 2014), *Drosophila* (Phadnis 2011), as well as plant and fungal taxa, supports the prevalence of complex DMIs involving multiple partners (reviewed in Fraïsse et al. 2014). Theoretical studies have also implicated the role of divergence in complex regulatory pathways in driving postzygotic reproductive isolation via DMIs, with reproductive isolation being more likely to develop when regulatory pathways contain larger numbers of loci (Johnson and Porter 2000) and are under the influence of directional selection (Porter and Johnson 2002; Johnson and Porter 2006). Empirical evidence from house mice suggests that divergence in regulatory elements between subspecies disrupts epistatic interactions between sets of genes, resulting in significant reductions or elevations of gene expression in hybrid relative to pure subspecies individuals (Mack et al. 2016).

© The Author(s) 2020. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

Misexpression in hybrid individuals is commonly observed and has been associated with reduced fertility in house mice (Good et al. 2010; Turner and Harr 2014; Turner et al. 2014; Mack et al. 2016; Larson et al. 2017), *Drosophila* (Michalak and Noor 2003; Moehring et al. 2007; Gomes and Civetta 2015), and cats (Davis et al. 2015). Understanding the role of potentially complex DMIs in driving patterns of misexpression may be facilitated by exploring gene interactions in a network context, in which expression patterns of sets of genes are allowed to depend upon one another. Network-based approaches cluster genes into coexpression modules, which are likely to be associated with common biological pathways and functions (e.g., Ayroles et al. 2009; Miller et al. 2010; reviewed in Mackay 2014), and have been employed to identify sets of genes associated with fitness-related phenotypes (Jumbo-Lucioni et al. 2010; Mack et al. 2018) and genes with disrupted interactions associated with disease (e.g., Saris et al. 2009; Miller et al. 2010; reviewed in De la Fuente 2010). Lee et al. (2013) demonstrated that SNPs associated with complex, polygenic traits were significantly more likely to lie within highly connected network “hub genes,” suggesting information regarding network topology may be useful in identifying biologically important genes. Network approaches have also been used to identify mechanisms involved in adaptive ecological divergence (Kelley et al. 2016; Gould et al. 2018; Zhao et al. 2019), and to explore the mechanisms underlying ecological speciation of sympatric lake whitefish ecotypes (Filteau et al. 2013). However, the power of network-based analyses for exploring the disrupted gene interactions associated with reduced hybrid fertility and postzygotic reproductive isolation is yet to be realized.

As mutations and incompatibilities continue to accumulate between independently evolving lineages over time, identifying the incompatibilities that initially caused reproductive isolation requires studying recently diverged lineages, with incomplete reproductive barriers. The European house mouse (*Mus musculus*) provides one such system. Three subspecies, *M. m. musculus*, *M. m. domesticus*, and *M. m. castaneus*, diverged ~500,000 years ago (Gerald et al. 2011). Following this initial divergence, *M. m. musculus* and *M. m. domesticus* (henceforth referred to as *musculus* and *domesticus*, respectively) used different routes to colonize Europe providing the opportunity for the accumulation of DMIs in allopatry. Laboratory crosses between *musculus* and *domesticus* have demonstrated reduced fertility in hybrid males (Forejt and Ivanyi 1974; Oka et al. 2004; Good, Dean, et al. 2008), although the degree of sterility varies depending on the nature of the cross (Britton-Davidian et al. 2005; Vyskočilová et al. 2005; Good, Handel, et al. 2008; Larson et al. 2018). Genetic studies of *musculus*–*domesticus* hybrids have also led to the first characterization of a mammalian hybrid incompatibility gene, *Prdm9* (Mihola et al. 2009), an autosomal histone methyltransferase which binds DNA at recombination hotspots and plays an important role in the initiation of meiotic recombination (Baudat et al. 2010; Davies et al. 2016). Negative interactions between some *domesticus* *Prdm9* alleles and loci on the *musculus* X-chromosome disrupt expression of the X-chromosome and chromosome

synapsis resulting in meiotic arrest (Bhattacharyya et al. 2013, 2014; Campbell et al. 2013; Larson et al. 2017).

Following colonization of Europe, *musculus* and *domesticus* expanded their ranges to form a zone of secondary contact traversing Central Europe (Macholán et al. 2003; Janoušek et al. 2012). Hybridization along this secondary contact zone is common resulting in a cline of genomic admixture (Payseur et al. 2004; Macholán et al. 2007; Teeter et al. 2007; Janoušek et al. 2012). Reduced male fertility is frequent in wild-caught hybrids, but complete sterility is rare (Britton-Davidian et al. 2005; Albrechtova et al. 2012; Turner et al. 2012), suggesting that this postzygotic barrier reduces gene flow and contributes to maintenance of the subspecies boundary, but is incomplete and variable in strength across populations.

In the present study, we are building on genetic mapping of fertility phenotypes and gene expression traits in F₂ *musculus*–*domesticus* hybrids generated through a laboratory cross (White et al. 2011; Turner et al. 2014), and in offspring of wild-caught *musculus*–*domesticus* hybrids (Turner and Harr 2014), which have identified many autosomal and X-linked sterility loci. DMIs were mapped in both studies, using different methods, yet many loci and interactions are shared between mapping populations (Turner and Harr 2014). Most loci have multiple interaction partners supporting the presence of complex DMIs (Turner and Harr 2014; Turner et al. 2014). Identifying the underlying mechanisms and causative genes remains a challenge, because many loci encompass a large number of genes.

Here, we characterize disruptions in gene networks associated with hybrid male sterility in mice by analyzing patterns of gene coexpression using microarray data from a total of 467 mice from two *musculus*–*domesticus* hybrid mapping populations (Turner and Harr 2014; Turner et al. 2014). We use weighted gene coexpression analysis (WGCNA; Langfelder and Horvath 2008) to: 1) characterize a consensus “fertile” network, 2) identify modules of coexpressed genes associated with biological pathways and processes, 3) identify modules disrupted in subfertile hybrids, and 4) identify specific candidate genes likely to be driving network disruptions in subfertile hybrids.

Results

Concordant Genome-Wide Testis Expression Patterns in F₂ and Hybrid Zone Hybrids

We analyzed genome-wide expression patterns in testes from 295 F₂ hybrids between two inbred strains of *musculus* and *domesticus* (White et al. 2011; Turner et al. 2014) and 172 lab-bred male offspring of mice wild-caught from a hybrid zone (HZ) (Turner and Harr 2014). Principal components analysis (PCA) of the F₂ and HZ data sets showed similar overall patterns (supplementary fig. 1, Supplementary Material online). For both populations, PC1, which explains 20.1% of variation in F₂ hybrids and 27.8% of variation in HZ hybrids, is clearly associated with fertility classified on the basis of relative testis weight (testis weight/body weight) and sperm count (White et al. 2011; Turner et al. 2012; supplementary fig. 1A

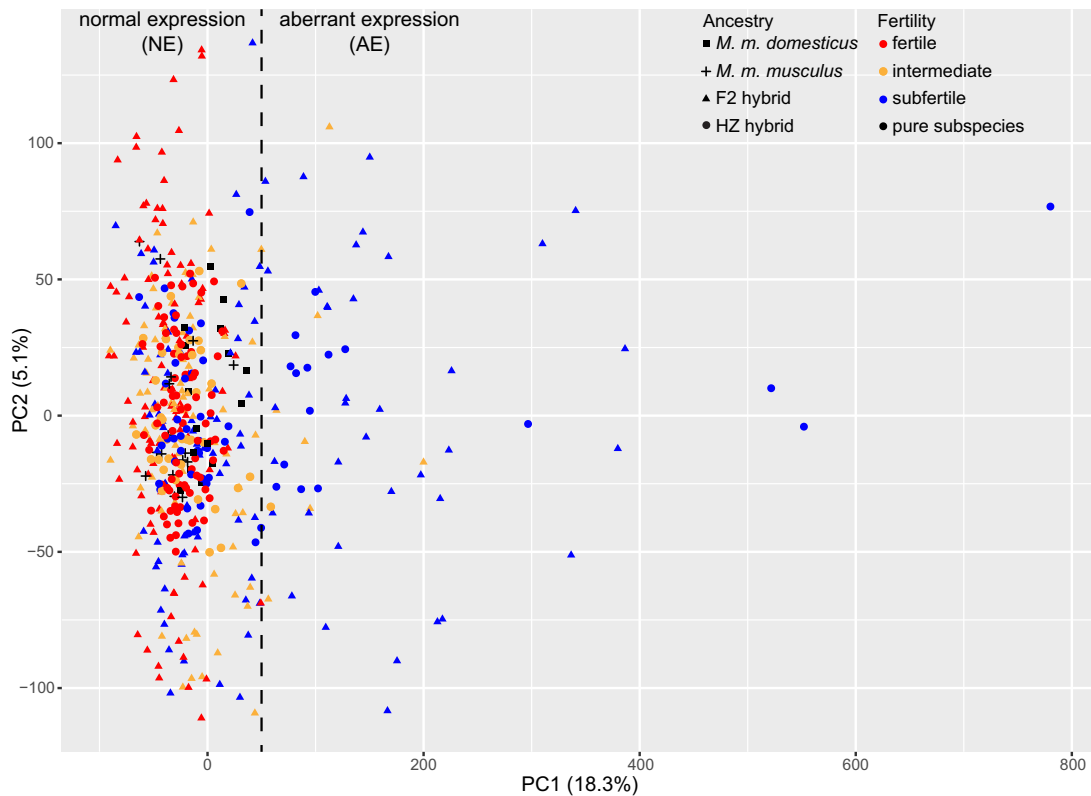


Fig. 1. Principal components analysis (PCA) of genome-wide expression in testis of pure *Mus musculus domesticus*, *M. m. musculus* and hybrids, PC1 versus PC2. Point shape indicates subspecies or hybrid mapping population for each individual and point color indicates fertility class (see Materials and Methods). Subfertile hybrids with PC1 scores outside the range observed in pure subspecies males and fertile hybrids are classified as “Subfertile Aberrant Expression” (SFAE), whereas subfertile hybrids within the pure subspecies and fertile hybrid range were classified as “Subfertile Normal Expression” (SFNE). The dashed line indicates the cutoff between SFNE and SFAE hybrid groups.

and B, [Supplementary Material](#) online). The PC1 loadings for probes in the two data sets are strikingly highly correlated ($r = 0.91$, $P < 2.2e-16$) suggesting a common set of genes contribute to the “sterile” expression pattern ([supplementary fig. 1C](#), [Supplementary Material](#) online). The remarkable similarities in overall testis expression pattern between mapping populations, and previous evidence for shared incompatibilities ([Turner and Harr 2014](#)), motivated us to combine expression data from F₂ and HZ hybrids to characterize gene network disruptions associated with sterility.

After combining and batch-correcting data for the 36,896 probes (representing 17,775 unique Entrez genes) with above-background levels of expression from F₂ and HZ hybrid and pure subspecies males, PC1 explains 18.26% of variation in gene expression ([fig. 1](#)) and is significantly negatively correlated with both relative testis weight and sperm count ($r = -0.527$ and -0.161 , $P = 2.2e-16$ and $3.3e-4$, respectively). The PC1 scores of fertile hybrid and pure subspecies males show similar levels of variation (variance: 609.36 and 616.10, respectively), whereas variation in PC1 scores is considerably higher in hybrids with subfertile phenotypes (variance: 15,149.70). Hybrids for which both fertility phenotypes fall within the pure subspecies range, yet at least one of the phenotypes is >1 SD from the pure subspecies mean, were classified as “intermediate phenotype” (see Materials and Methods) and show intermediate levels of variation along PC1 (variance: 1,816.78).

A sizeable proportion (33.5%) of the subfertile hybrids have PC1 scores that exceed the fertile hybrid and pure subspecies range ([fig. 1](#)); we classified these individuals as “Subfertile Aberrant Expression” (SFAE), and classified subfertile hybrids that cluster with fertile hybrids and pure subspecies males along PC1 as “Subfertile Normal Expression” (SFNE). A total of 69 F₂ and 38 HZ hybrids were categorized as SFNE, whereas 37 F₂ and 17 HZ hybrids were categorized as SFAE ([fig. 1](#)). PC3, which explains 4.28% of variation, is associated with subspecies ancestry: *musculus* individuals have high scores, *domesticus* individuals have low scores, and hybrids have intermediate scores ([supplementary fig. 2](#), [Supplementary Material](#) online).

Fertile Hybrid Consensus Network

To identify potentially interacting sets of genes that are coexpressed in fertile hybrids from different genomic backgrounds, we constructed a consensus fertile network using expression data from the fertile F₂ ($n = 102$) and HZ ($n = 79$) hybrids and 18,411 probes representing 10,531 genes (see Materials and Methods for details). A total of 14,346 probes, representing 7,989 unique genes, were assigned to 1 of the 15 coexpression modules ([fig. 2A](#) and [supplementary data 1](#), [Supplementary Material](#) online); 4,065 probes could not be assigned to a module and are shown in the gray, “bin” module. Thirteen modules are significantly enriched for specific functions on the basis of Gene Ontology (GO) analysis, of

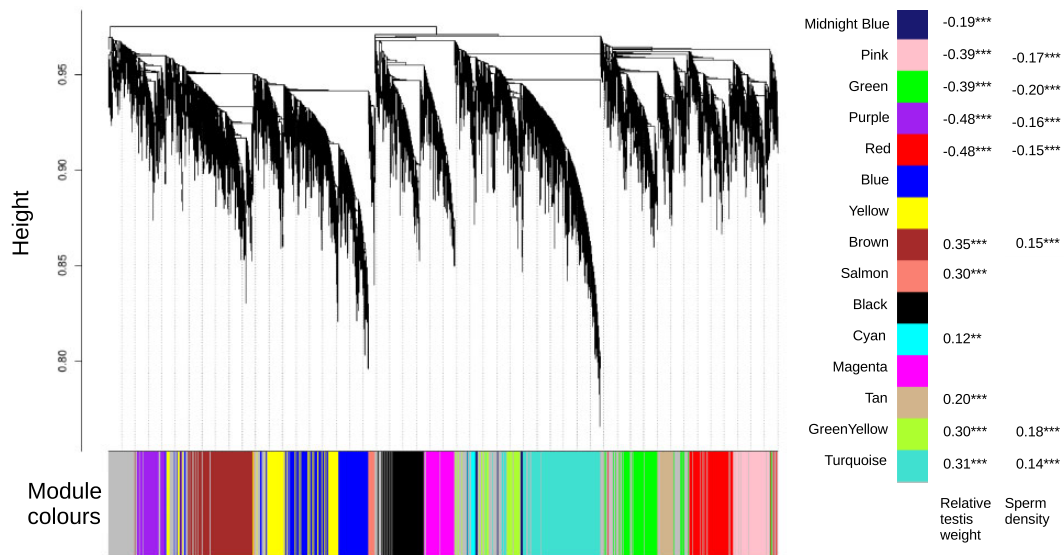


Fig. 2. Gene coexpression modules. (A) Consensus fertile network generated using weighted gene coexpression network analysis (WGCNA; Langfelder and Horvath 2008) of testis expression from 102 fertile F_2 and 79 fertile HZ hybrid males. The dendrogram shows the clustering of probes based on the topological overlap distance within fertile hybrids. Color bar beneath the dendrogram indicates coexpression modules. (B) The correlation between the module eigengene (ME), representing overall module expression, and sterility phenotypes. Significant positive correlations are indicated in red and significant negative correlations are indicated in blue; ** $P < 0.01$, *** $P < 0.001$.

which seven were significantly enriched for spermatogenesis or potentially related functions (table 1 and supplementary data 2, Supplementary Material online). The module eigengene (ME), which describes the overall expression level of each module across the full data set of fertile and subfertile hybrids, was significantly positively correlated with both sperm count and relative testis weight for three modules, and significantly positively correlated with relative testis weight alone for an additional three modules (fig. 2B). Four modules have an ME which is significantly negatively correlated with both sterility phenotypes, whereas the ME of one additional module is significantly negatively correlated with relative testis weight alone. Of the seven modules that are significantly enriched for spermatogenesis or potentially related functions, two have an ME which is significantly positively correlated with at least one of the fertility phenotypes (Brown and Tan), whereas two have an ME which is significantly negatively correlated with at least one of the fertility phenotypes (Green and Pink) (fig. 2B). These correlations suggest that the inferred modules of coexpression are informative about fertility.

In addition to the two fertility traits used to classify hybrid males, data describing several sperm motility traits were also available for a subset of the HZ males ($n = 121$) (Turner et al. 2012). Interestingly, the ME of the Brown, Greenyellow, and Pink modules were all significantly positively correlated with beat-cross frequency (BCF) ($r = 0.19, 0.36, \text{ and } 0.23, P = 0.034, 5.9e-5, \text{ and } 0.012$, respectively), which is defined as the frequency with which the sperm head crosses the middle plane of the straight trajectory (Turner et al. 2012). In human sperm, BCF is roughly proportional to the frequency with which the sperm flagellum beats (Serres et al. 1984). This is of note since the Brown module is enriched for genes expressed in the

ciliary part ($P = 0.049$), likely to be important for the functioning of the sperm flagellum, and the Pink module is enriched for genes involved in the mitochondrial respiratory chain ($P = 4.3e-11$), which is highly active in motile sperm (Amaral et al. 2013).

Network Preservation in Subfertile Hybrids

To determine whether coexpression networks are disrupted in subfertile and intermediate phenotype hybrids, we estimated module preservation using two approaches based on a set of metrics developed by Langfelder et al. (2011; see Materials and Methods). Module preservation was assessed independently for F_2 and HZ hybrids, because the presence and prevalence of specific DMIs and associated network disruptions may vary between mapping populations. Although significant preservation was detected for all modules in intermediate phenotype hybrid groups, several modules showed a lack of significant preservation in the subfertile phenotype hybrids (see supplementary table 1, Supplementary Material online). Figure 3A shows the estimated preservation of each module from the consensus fertile network in F_2 and HZ hybrids with subfertile phenotypes and either aberrant or normal overall expression patterns. Note that median rank shows relative preservation within a group, whereas Z statistics allow for comparison between groups. Modules showing significant evidence for preservation in subfertile hybrids, according to the preservation statistics presented in supplementary table 1, Supplementary Material online, are represented by circles, whereas modules showing a lack of significant preservation are represented by squares (fig. 3A). Figure 3B and C illustrate coexpression patterns within a well preserved (Red) versus a poorly preserved (Brown) module in the F_2 SFAE hybrids. Pairwise coexpression in the Red module

Table 1. Module Enrichment within the Fertile F₂ and HZ Networks, Using the Testis-Expressed Genes As a Background.

Network Module	Number of Genes	Most Significant GO Enrichment Term			Testis Cell Type(s) in Which a Significant Proportion of Genes are Expressed ^b
		GO Term	Number Genes in GO Term	P Value ^a	
Black	719	Ubiquitin-like protein transferase activity	44	3.4e-7	SG, eP1, Sertoli
Blue	1,421	GTPase binding	88	1.2e-11	SG, eP1, mP, MII, S1, S11, Sertoli
Brown	1,274	Spermatogenesis	72	1.1e-4	SG, D, MI, MII, S11, Sertoli
		Ubiquitin-dependent protein catabolic process	74	1.4e-4	
Cyan	102	Ciliary part	50	4.6e-2	—
		—	—	—	
Green	894	Protein ubiquitination	57	2.2e-4	SG, eP1, S1, Sertoli
Greenyellow	455	Chromosome segregation	31	1.7e-2	S8
		—	—	—	
Magenta	602	Nuclear chromosome	55	1.3e-10	SG, eP1, eP2, MII, Sertoli
		Regulation of chromosome organization	38	1.8e-7	
Midnightblue	90	Phagocytosis, engulfment	5	6.3e-3	SG, Leydig
Pink	571	Mitochondrial protein complex	50	1.7e-22	
		Mitochondrial respiratory chain	17	4.3e-11	
Purple	435	Monocarboxylic acid metabolic process	38	9.3e-6	Leydig
Red	717	Monocarboxylic acid metabolic process	54	5.8e-7	Leydig, Sertoli
Salmon	203	Microtubule	16	4.5e-3	SG, S11
Tan	320	Ribosome	11	1.8e-2	SG
		Spermatogenesis	22	4.3e-2	
Turquoise	1,436	Olfactory receptor activity	139	9.9e-39	SG, eP1, eP2, MII, S1, Sertoli
		mRNA processing	50	6.4e-5	
Yellow	933	Spermatogenesis	49	4.6e-3	

NOTE.—GOs potentially related to spermatogenesis are listed in bold.

^aP values were corrected using the Benjamini–Hochberg correction (Benjamini and Hochberg 1995).

^bSignificant overlaps between genes expressed in specific testis cell types (Ernst et al. 2019) and module gene content were identified using Fisher’s exact tests with Benjamini–Hochberg correction for multiple tests, using all testis-expressed genes as the background. Testis cell types include spermatogonia (SG), early pachytene spermatocytes (eP1 and eP2, respectively), diplotene spermatocytes (D), metaphase I and II spermatocytes (MI and MII, respectively), stage 1–11 spermatids (S1–11), and Sertoli and Leydig cells.

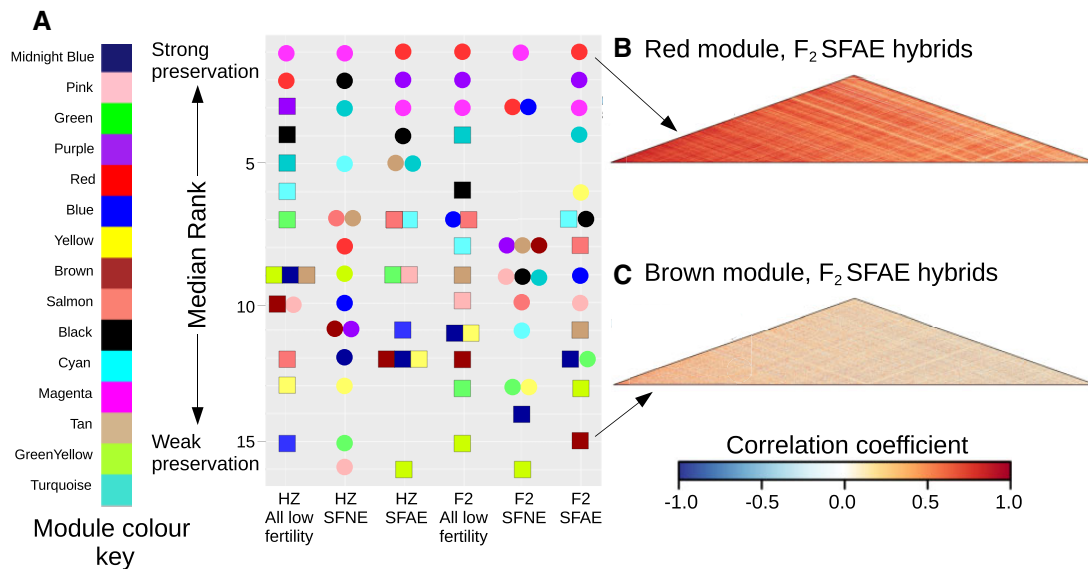


FIG. 3. Preservation of the fertile consensus network within subfertile hybrids classified according to mapping population derived and expression profile (SFNE or SFAE, see fig. 1). (A) Module preservation estimated using median rank statistics. Color indicates module identity. Circles represent significantly preserved modules, and squares represent modules not significantly preserved. (B and C) Coexpression heatmaps for examples of well-preserved (Red) and poorly preserved (Brown) modules. Heatmaps show pairwise correlations between expression values for all genes within modules for SFAE F₂ hybrids.

is characterized by strong positive correlations (fig. 3B). In contrast, many pairwise expression correlations in the Brown module are weakened or even reversed in direction (fig. 3C), suggesting substantial disruption in the coexpression of these gene pairs.

The level of preservation of modules from the fertile network was similar in subfertile HZ and F₂ hybrids. Three modules (Magenta, Red, Pink) showed strong evidence for preservation in subfertile HZ hybrids (fig. 3A and supplementary table 1, Supplementary Material online); the remaining 12 modules had either a $Z_{\text{summary}} < 10$ and/or nonsignificant NetRep statistics, indicating weak or a lack of preservation (Langfelder et al. 2011; Ritchie et al. 2016, see Materials and Methods). In F₂ hybrids, four modules (Magenta, Red, Purple, and Blue) showed strong evidence for preservation in subfertile mice.

As expected, module preservation was much higher in subfertile hybrids with normal expression based on PC1 (SFNE); all modules showed strong evidence of preservation in the HZ SFNE hybrids, and 13 out of 15 modules were strongly preserved in F₂ SFNE hybrids (Greenyellow and Midnightblue modules had nonsignificant NetRep scores; supplementary table 1, Supplementary Material online). By contrast, preservation was lower in subfertile hybrids with aberrant expression; 6 out of 15 modules were preserved in HZ SFAE hybrids, and 9 modules in F₂ SFAE hybrids.

Many modules show consistent levels of preservation across subfertile hybrid classes (fig. 3A). The Red, Purple, Magenta, and Turquoise modules consistently rank among the most strongly preserved modules; these modules are significantly enriched for metabolic processes, histone binding, and olfactory receptor activity, respectively (table 1). By contrast, the Brown, Green, Yellow, Midnightblue, and Greenyellow modules consistently rank among the modules with the weakest preservation; these modules are significantly enriched for histone binding and the ciliary part, DNA repair and chromosome segregation, mRNA processing and spermatogenesis, and phagocytosis (the Greenyellow module has no significant enrichments). Notably, although the Magenta and Green modules are significantly enriched for similar processes (chromosome organization and segregation, respectively), the Magenta module appears to be one of the best-preserved modules, whereas the Green module is among the least preserved across hybrid groups (fig. 3A).

Although consistencies across subfertile groups are apparent, notable differences were also detected between the HZ and F₂ hybrids. For example, the Blue module, enriched for spermatogenesis, is significantly preserved in the F₂ SFAE hybrids, yet shows relatively poor preservation in the HZ SFAE hybrids (fig. 3A and see supplementary table 1, Supplementary Material online). The Yellow module, which is also significantly enriched for spermatogenesis, is among the least preserved modules in all subfertile hybrid groups except for the F₂ SFAE, within which it is relatively well preserved.

In summary, broad similarity of module preservation statistics across subfertile classes provides further evidence that specific functions and pathways are commonly disrupted in

sterile hybrids. By contrast, modules showing differences in conservation in subfertile HZ versus F₂ hybrids suggest some network disruptions are unique to specific mapping populations.

Differentially Correlated Genes

Our next aim was to identify specific genes responsible for the signal interpreted as network disruption in subfertile hybrids. We identified genes with significant changes in coexpression pattern in subfertile versus fertile hybrids using differential correlation analysis, which was performed independently for each module using the DGCA R package (McKenzie et al. 2016). A total of 2,800 genes showed a significant loss or reversal of coexpression pattern within at least one of the subfertile groups (supplementary data 1, Supplementary Material online). Although the power to detect changes in coexpression patterns is likely to differ between subfertile groups due to varying sample size, differentially correlated genes were detected within all four subfertile hybrid groups (SFNE and SFAE F₂ and HZ hybrids). As expected, a higher proportion of genes were differentially correlated in weakly preserved (28.89–46.59%) versus strongly preserved (1.12–3.65%) modules (supplementary table 2, Supplementary Material online).

Modules Associated with Specific Stages of Spermatogenesis or Testis Cell Types

To determine whether coexpression modules are associated with specific stages of spermatogenesis, we tested for significant enrichment of genes expressed in different testis cell types, which was recently determined at high resolution using a combination of single-cell RNAseq and bulk RNAseq at different time points during the first stage of spermatogenesis (gene lists from Supplementary Data 2 in Ernst et al. 2019). The gene content of nine modules (Black, Blue, Brown, Green, Magenta, Pink, Salmon, Tan, and Yellow) overlaps significantly with genes expressed in spermatogonia (table 1). Six modules (Black, Blue, Brown, Green, Magenta, and Yellow) overlap with genes expressed in spermatocytes during stages of meiosis (early pachytene, diplotene, and metaphase). Six modules (Blue, Brown, Green, Greenyellow, Salmon, and Yellow) overlap with genes expressed postmeiotically in at least 1 of the 11 stages of developing spermatids. For somatic cell types, seven modules (Black, Blue, Brown, Green, Magenta, Red, Yellow) overlap with genes expressed in Sertoli cells and three modules (Pink, Purple, and Red) are enriched for genes expressed in Leydig cells. Thus, of the 15 modules, 3 overlap significantly with genes expressed during spermatogenesis only (Greenyellow, Salmon, and Tan), 7 modules overlap significantly with both genes expressed in spermatogenesis and with genes expressed in somatic cells (Black, Blue, Brown, Green, Magenta, Pink, and Yellow), 2 modules overlap only with somatic cells (Purple and Red), and 3 modules do not overlap with genes expressed in any of specific testis cell types (Cyan, Midnightblue, and Turquoise). The complete lists of genes with cell type-specific expression (Ernst et al. 2019) are in supplementary data 3, Supplementary Material online. Our findings indicate that

Table 2. Significant Overlap between Genes within Coexpression Modules and Quantitative Trait Transcripts (QTT) Associated with *trans*-eQTL Hotspots (Turner et al. 2014).

Coexpression Module	<i>trans</i> -eQTL Hotspot, Direction ^a	Sterile Expression Pattern ^b	Number of Genes in Module	Number of QTT	Overlap between Gene Sets	Fisher's Exact P Value	Corrected P Value ^c
Black	18–38 cM Chr 15, Mhigh	Low	719	511	62	1.1e-5	5.1e-5
Brown	38–44 cM Chr 3, Dhigh	Low	1,271	102	30	3.2e-6	4.8e-5
Green	18–38 cM Chr 15, Dhigh	High	894	867	102	7.2e-4	3.6e-3
Greenyellow	18–38 cM Chr 15, Mhigh	Low	455	511	47	1.2e-6	8.4e-6
Red	26–38 cM Chr 2, Dhigh	High	717	347	85	7.1e-26	1.1e-24
	54–62 cM Chr 11, Dhigh	High	717	763	166	5.1e-44	7.1e-43
Purple	18–38 cM Chr 15, Dhigh	High	717	246	59	1.1e-17	8.5e-17
	46–50 cM Chr 15, Dhigh	High	717	867	293	3.5e-140	5.2e-139
	0–16 cM Chr 17, Mhigh	High	717	100	31	4.7e-13	4.7e-13
	0–42 cM X Chr, Mhigh	High	717	1,117	171	3.3e-25	2.5e-24
	26–38 cM Chr 2, Dhigh	High	435	347	36	5.1e-7	3.8e-6
	4–24 cM Chr 10, Dhigh	High	435	763	88	7.7e-19	5.4e-18
	54–62 cM Chr 11, Dhigh	High	435	246	58	5.2e-28	8.8e-27
	18–38 cM Chr 15, Dhigh	High	435	867	130	1.0e-40	7.5e-40
	46–50 cM Chr 15, Dhigh	High	435	100	14	7.4e-5	5.6e-4
	0–16 cM Chr 17, Mhigh	High	435	1,117	155	3.3e-45	5.0e-44
Turquoise	0–42 cM X Chr, Mhigh	High	435	3,329	238	6.2e-23	4.7e-22
	18–38 cM Chr 15, Mhigh	Low	1,434	511	115	4.8e-8	6.7e-7

^aChromosome and cM position for *trans*-eQTL hotspots, as reported in Turner et al. (2014), and subset of QTTs associated with either the *domesticus* (Dhigh) or *musculus* (Mhigh) eQTL allele.

^bQTTs showed higher or lower expression associated with the “sterile” eQTL allele (classified on the basis of multiple criteria, see Table 2 in Turner et al. 2014).

^cBenjamini–Hochberg correction for multiple tests.

several coexpression modules can be linked with one or more stages of the spermatogenesis process.

Modules Associated with *trans*-eQTL Hotspots

Turner et al. (2014) identified 11 *trans*-expression quantitative trait loci (eQTL) hotspots, that is, regions of the genome with significant effects on the expression of hundreds to thousands of quantitative trait transcripts (QTTs) in F₂ hybrids, and provided several lines of evidence linking these hotspots to sterility phenotypes. We investigated whether *trans*-eQTL hotspots affected specific parts of the gene coexpression network by testing for overlap between QTT associated with each hotspot and genes in modules from the fertile network. Genes in seven modules overlap with QTT from one to seven hotspots (table 2). Modules negatively correlated with relative testis weight and/or sperm count (Green, Red, Purple) overlap significantly with QTT showing high expression associated with the sterile allele, as established from expression patterns in sterile F₁ hybrids (Turner et al. 2014), whereas modules positively correlated with fertility phenotypes (Brown, Greenyellow, Turquoise) overlap QTT with low expression associated with the sterile allele. That is, in all the cases, “sterile” versus “fertile” patterns were consistent between coexpression modules and QTT associated with sterile versus fertile alleles.

Modules with higher preservation in subfertile hybrids (Red, Purple) overlap with QTT associated with multiple *trans*-eQTL hotspots (table 2), highlighting potential interactions between eQTL hotspots. In contrast, modules with intermediate (Black, Turquoise) or weak (Brown, Green, and Greenyellow) preservation in subfertile hybrids each overlap significantly with QTT associated with a single *trans*-eQTL hotspot. In summary, five modules with intermediate or weak preservation overlap significantly with QTT associated

with a specific *trans*-eQTL hotspot supporting an influence of the underlying sterility alleles on specific parts of the fertile gene network.

Module Hub Genes

We next identified hub genes in the fertile network. Hub genes are highly connected within modules, and thus are likely to have a disproportionate effect on module-wide expression patterns. We identified 281 hub genes across the 15 modules on the basis of degree (number of connections) and module membership (correlation between the expression of a gene and the ME) (Horvath and Dong 2008). Of the 281 hub genes, 95 genes within 10 modules show a significant loss or reversal of coexpression pattern in at least one of the subfertile hybrid groups relative to the fertile hybrids (table 3). Of those 95 differentially correlated hub genes, 29 and 57 show a significant loss of coexpression pattern in the SFAE and SFNE F₂, respectively, whereas 11 and 10 show a significant loss of coexpression pattern the SFAE and SFNE HZ, respectively (supplementary data 4, Supplementary Material online). Hence, although the WGCNA and NetRep analyses show evidence for stronger module preservation in SFNE relative to SFAE hybrids, differential correlation analysis shows evidence for the disruption of module hub gene interactions in both the SFAE and SFNE hybrid groups.

Table 3 lists the module hub genes indicating the following as potential candidates for DMIs: genes with different coexpression patterns in subfertile relative to fertile hybrids ($n = 95$); genes that have been found to be expressed within spermatogonia, spermatocytes, or spermatids (Ernst et al. 2019) ($n = 152$); genes with GOs including those related to the regulation of gene expression and/or male fertility ($n = 52$); genes that fall within previously identified *trans*-eQTL hotspots (Turner et al. 2014) ($n = 40$); and genes within

Table 3. Candidate Module Hub Genes.

Module	Hub Genes
Black	1700081L11Rik , <i>Asxl2^a</i> , <i>Atxn2^b</i> , <i>Fto^{b,c}</i> , <i>Gle1^{b,c}</i> , <i>Hectd2</i> , <i>Lpp^b</i> , <i>Nr2c2^{a,b,d}</i> , <i>Rab14^{b,c}</i> , <i>Sbno1^{a,b}</i> , <i>Sltm^{a,b}</i> , <i>Srpk2^{a,b}</i> , <i>Stim1</i> , <i>Pibf1^b</i> , <i>Ppp6r3</i> , <i>Ube2e2</i> , <i>Znrf3^b</i>
Blue	1700029G01Rik^b , <i>Agpat6^b</i> , <i>Akap1^{b,d}</i> , <i>Ankrd13a^b</i> , <i>Ccdc147^b</i> , <i>Dnhd1^b</i> , <i>Gm101^b</i> , <i>Gm14857^b</i> , <i>Hira^{a,b}</i> , <i>Hk1^{b,d}</i> , <i>Inpp5e^{b,c}</i> , <i>Nupl2^{a,b}</i> , <i>Phf12^{a,b}</i> , <i>Rangap1^{b,d}</i> , <i>Rnf19b^b</i> , <i>Sbf2^b</i> , <i>Slain2^b</i> , <i>Sort^{a,b}</i> , <i>Triml1^b</i> , <i>Zfp445^a</i> , <i>1110037F02Rik</i> , <i>1700029F09Rik^b</i> , <i>4930473A06Rik^b</i> , <i>4932425I24Rik^b</i> , <i>Adam1a^{a,b}</i> , <i>Ankrd40^d</i> , <i>Brap^b</i> , <i>Cnot1^{a,b,c}</i> , <i>Ctnn</i> , <i>Cpsf2^b</i> , <i>D1Pas1^a</i> , <i>Fam48a^b</i> , <i>Gigyf1^b</i> , <i>Ggcx</i> , <i>Hsp90ab1^{b,c}</i> , <i>Ift172^b</i> , <i>Insig2^b</i> , <i>Ints10^b</i> , <i>Ipo11</i> , <i>Kifs5b</i> , <i>Mtmr9^b</i> , <i>Mum1^b</i> , <i>Nasp^b</i> , <i>N4bp1^b</i> , <i>Nelf^{b,c}</i> , <i>Net1^b</i> , <i>Nxf1^a</i> , <i>Pfas^{b,c}</i> , <i>Ranbp10</i> , <i>Rnd3</i> , <i>Sarm1^b</i> , <i>Shprh</i> , <i>Smg5^{b,d}</i> , <i>Snx27^d</i> , <i>Spag5</i> , <i>Tapt1</i> , <i>Tex2^b</i> , <i>Trpc4ap^b</i> , <i>Usp33</i> , <i>Zfp541^a</i>
Brown	Dbnl^b , <i>Evisl^b</i> , <i>Fkbp4^b</i> , <i>Kif17^b</i> , <i>Pacs2^b</i> , <i>Prkcd^{b,c}</i> , <i>Ptdss2^b</i> , <i>Smarcd1</i> , <i>Traf1^b</i> , <i>Xpo6</i>
Cyan	Foxl2^a , <i>Arhgap20</i> , <i>N4bp2l2^b</i> , <i>Olfr1350</i> , <i>Plekha5^b</i> , <i>Wac^b</i>
Green	Clint1^b , <i>Dpy19L4</i> , <i>Maml1^a</i> , <i>Pik3c2a</i> , <i>Zfp711^{a,b,d}</i> , <i>Zfp770^{a,b,d}</i> , <i>Dld^{a,b,c}</i> , <i>Ganc^b</i>
Greenyellow	D10Bwg1379e , <i>E130304I02Rik</i> , <i>Fbxo2^b</i> , <i>Mapkap1^{b,c}</i> , <i>Olfr279</i> , <i>Rapgef1^{b,c}</i> , <i>Shc2</i> , <i>Sprr2b^d</i> , <i>Tbc1d2^b</i>
Magenta	4921528I01Rik , <i>B230208H17Rik^c</i> , <i>Rab6l^c</i> , <i>Akap13^b</i> , <i>Casc5^{b,d}</i> , <i>Cdk5rap2^{a,b}</i> , <i>Cep290^{a,b}</i> , <i>Chd4^{a,b}</i> , <i>Ckaps^{b,d}</i> , <i>Clasp1</i> , <i>Culs5^b</i> , <i>Ddx21^{a,b,d}</i> , <i>Ddx46^b</i> , <i>Dek^b</i> , <i>Dicer1^{a,b}</i> , <i>Eea1^b</i> , <i>Eif5b^b</i> , <i>Eif3c^b</i> , <i>Eprs^b</i> , <i>Golgb1</i> , <i>Gpatch8</i> , <i>Hcfc1^d</i> , <i>Heatr6^{b,d}</i> , <i>Hectd1^{b,c}</i> , <i>Kdm5b^{a,b}</i> , <i>Kif20b^b</i> , <i>Man1a2^d</i> , <i>Mlh3^{a,b}</i> , <i>Nol8^b</i> , <i>Nvl^{b,c}</i> , <i>Parp10^{b,d}</i> , <i>Pcm1^b</i> , <i>Ppp2r5e^b</i> , <i>Prpf40a^b</i> , <i>R3hdm1</i> , <i>Rapgef6^b</i> , <i>Rif1^{a,b}</i> , <i>Rock1^b</i> , <i>Rrbp1</i> , <i>Setd5</i> , <i>Sfrs18</i> , <i>Smc2^{a,b}</i> , <i>Spnb2</i> , <i>Stlm^{a,b}</i> , <i>Thoc2^{a,d}</i> , <i>Thrap3^{a,b}</i> , <i>Tnrc18^b</i> , <i>Tpr^{a,b}</i> , <i>Trip11^b</i> , <i>Ubxn4^b</i> , <i>Wasf2^b</i> , <i>Wapal^{a,b}</i> , <i>Zc3h13^b</i>
Midnightblue	Adam28^c , <i>Hoxb6^{a,d}</i> , <i>Krt10^b</i> , <i>Lcn9^c</i> , <i>Cst11</i> , <i>Cyp4a12a</i> , <i>Lcn10^c</i> , <i>Ly6g5c^d</i> , <i>Krt10</i> , <i>Krt14</i> , <i>Rnf186</i> , <i>Serpinf2</i>
Pink	1810027O10Rik , <i>Bola2^b</i> , <i>Cox17</i> , <i>Med31^{a,b}</i> , <i>Pop5^b</i> , <i>Pop7^b</i> , <i>Ucqr</i>
Purple	Cry2^{a,d} , <i>Frk^a</i> , <i>Gart</i> , <i>Glb1l2</i> , <i>Gnas^a</i> , <i>Mboat1</i> , <i>Kcnk1</i> , <i>Zfp185^{b,c,d}</i>
Red	Acaa2 , <i>Chst8^c</i> , <i>Ifnar2^a</i> , <i>Itm2b^c</i> , <i>Gpx1^a</i> , <i>Cyp11a1^c</i> , <i>Rnf128^{b,c,d}</i> , <i>Tmem30b</i>
Salmon	Arhgef2^{b,d} , <i>Fam188b^b</i> , <i>Fgl1^b</i> , <i>Gm7416^d</i> , <i>Itgb3^b</i> , <i>Krtap3-1</i> , <i>Pomt2^b</i> , <i>Slc11a2</i>
Tan	Cops2^{a,b} , <i>Oxr1^b</i> , <i>Ralgapa1^{a,b}</i> , <i>Tmem30a</i>
Turquoise	2410017117Rik^d , <i>A730081D07Rik</i> , <i>Adamts14^d</i> , <i>Atp8b5^b</i> , <i>Atpif1</i> , <i>Calcoco1^a</i> , <i>Chst11</i> , <i>Csf2rb^d</i> , <i>Fam155a</i> , <i>Gm10229</i> , <i>Gm13033</i> , <i>Gpr152</i> , <i>Gsx1^{a,d}</i> , <i>Helq^b</i> , <i>Ifna13</i> , <i>Lce1l^d</i> , <i>Lrrc8e^b</i> , <i>Mc3r</i> , <i>Mis18bp1^b</i> , <i>Mlxip^{a,b}</i> , <i>Msx1^a</i> , <i>Ndor1^{b,d}</i> , <i>Olfr1131^d</i> , <i>Olfr166</i> , <i>Olfr414^c</i> , <i>Olfr656</i> , <i>Osr1^{a,c}</i> , <i>Phxr4^b</i> , <i>Prune2</i> , <i>Rasgrp4^b</i> , <i>Rsl1d1^{a,b,c}</i> , <i>Slc22a8</i> , <i>Slco2a1^b</i> , <i>Speg</i> , <i>Sprr2e^d</i> , <i>Sprr2j-ps^d</i> , <i>Tmc8</i> , <i>Vmn1r192</i> , <i>Ctrl</i> , <i>Cyp2d34^d</i> , <i>Gja5^d</i> , <i>Gm10319</i> , <i>Gm11019</i> , <i>Helq^b</i> , <i>Mdga2^b</i> , <i>Mfap2^b</i> , <i>Mlxip^b</i> , <i>Ndor^{b,c}</i> , <i>Nlrp3</i> , <i>Olfr374</i> , <i>Olfr434</i> , <i>Olfr549</i> , <i>Olfr577</i> , <i>Olfr60</i> , <i>Sall4^{a,b}</i> , <i>Serhl^{b,d}</i> , <i>Tbkbp1^{b,d}</i> , <i>Tmem8c^c</i>
Yellow	Ar18b^b , <i>Atr^b</i> , <i>Gm9805^d</i> , <i>Gmps^b</i> , <i>Lnp^b</i> , <i>Tcerg1^a</i> , <i>Aars^{a,b,c}</i> , <i>Nom1^b</i> , <i>Pex13^b</i> , <i>Psm2^b</i> , <i>Rab5a^{b,c}</i> , <i>Scaf11^d</i>

NOTE.—Differentially correlated hub genes, for which patterns of coexpression detected in fertile hybrids are significantly lost or reversed in at least one subfertile hybrid group, are highlighted in bold.

^aGenes with GO terms related to regulation of gene expression and/or male fertility.

^bGenes that have been found to be expressed in one or more class of testis germ cell (Ernst et al. 2019).

^cGenes within sterility loci identified by GWAS in HZ hybrids (Turner and Harr 2014).

^dGenes found within *trans*-eQTL hotspots (Turner et al. 2014).

sterility loci identified by GWAS in HZ mice (Turner and Harr 2014) ($n = 31$). Permutation tests revealed that the module hubs include a significantly larger number of genes that fall within GWAS sterility regions than do random draws from all testis-expressed genes ($P = 0.04$) and from all module genes (genes assigned to 1 of the 15 modules in the consensus fertile network, $P = 0.05$), supporting an enrichment for genes within GWAS sterility regions in the module hubs. Although the number of genes within eQTL hotspots did not differ significantly between the module hub genes and random draws from the testis-expressed genes ($P = 0.50$) and from all module genes ($P = 0.65$), this result is not unexpected as the eQTL hotspot regions span large regions of the genome, and likely contain many genes that play no role in fertility.

All hub genes in the Brown module are differentially correlated in at least one subfertile hybrid group relative to fertile hybrids. Figure 4 illustrates the widespread loss of connectivity within the SFAE relative to the fertile F_2 network. The loss of interactions for four of the module hub genes, *Prkcd*, *Dbnl*, *Ptdss2*, and *Fkbp4*, are shown in figure 4A, and the overall pattern of reduced connectivity in the SFAE network is shown in figure 4B. The Brown module is significantly enriched for genes involved in cilium organization, and several of the genes with cell component GOs including cilium and/or sperm flagellum were found to interact with the module hub genes (fig. 4A).

We found some hub genes in significantly preserved modules that nevertheless show disrupted coexpression patterns in subfertile hybrids. For example, the Blue module is significantly preserved in the SFAE F_2 hybrid group (fig. 3), however, several hub genes show significant loss of interactions in SFAE F_2 hybrids, including *Hk1*, *Akap1*, *Agpat6*, and *Slain2* (table 3 and supplementary fig. 3A and supplementary data 4, Supplementary Material online). The Blue module coexpression heatmaps for fertile and SFAE F_2 hybrids (supplementary fig. 3B and C, Supplementary Material online) show weakening of positive correlations in these subfertile hybrids. Meanwhile, the Midnightblue module was preserved in SFNE HZ hybrids, despite showing a lack of preservation in all other subfertile groups (fig. 3). However, two hub genes in this module have changes in interactions in both SFNE and SFAE HZ hybrids (fig. 5A). Moreover, an intermediate level of reduced connectivity overall in SFNE HZ hybrids in the Midnightblue module is apparent from visual comparison of the coexpression heatmap of the module compared with fertile hybrids and the more severely disrupted SFAE hybrids (fig. 5B–D).

Discussion

Although the DM model provides a well-accepted mechanism for the development of reproductive isolation between diverging lineages, the specific epistatic interactions

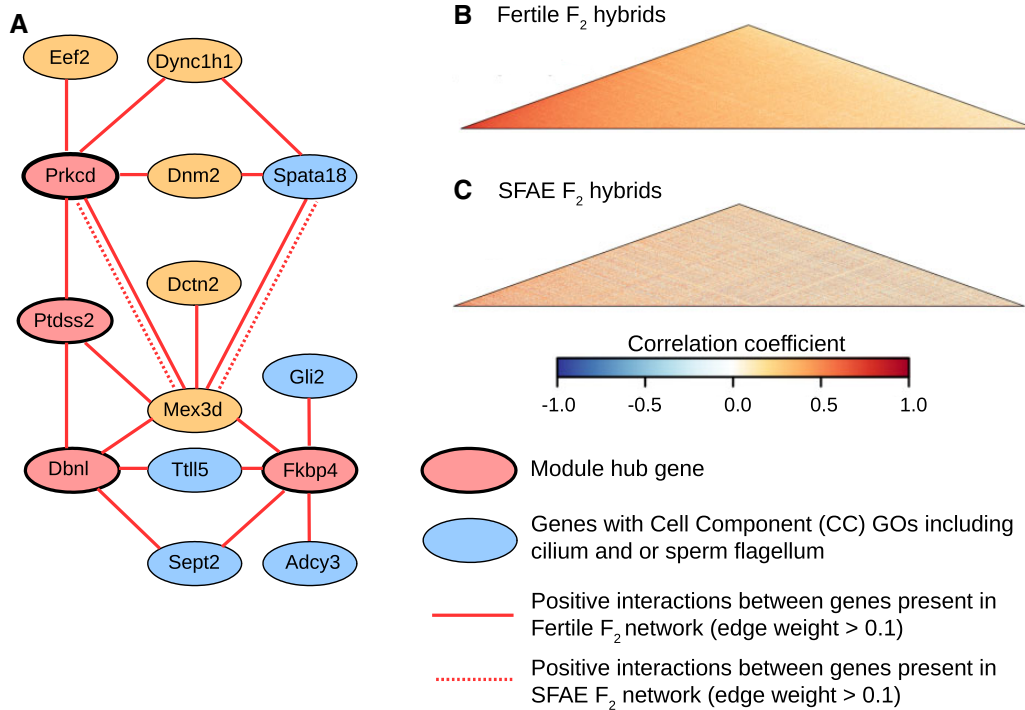


Fig. 4. Disrupted interactions of Brown module genes in subfertile F_2 hybrids. (A) Interactions between Brown module hub genes (red nodes) and genes with GOs including cilium and/or sperm flagellum (blue nodes) in fertile and SFAE F_2 hybrids. Orange nodes indicate intermediate genes with functions potentially related to male fertility. Gene interactions with an edge-weight exceeding 0.1, as estimated using topological overlap matrices, are indicated using continuous and dashed lines for the fertile and SFAE hybrids, respectively. (B and C) Coexpression heatmaps showing pairwise correlations between expression values of Brown module genes in fertile and SFAE F_2 hybrids, respectively.

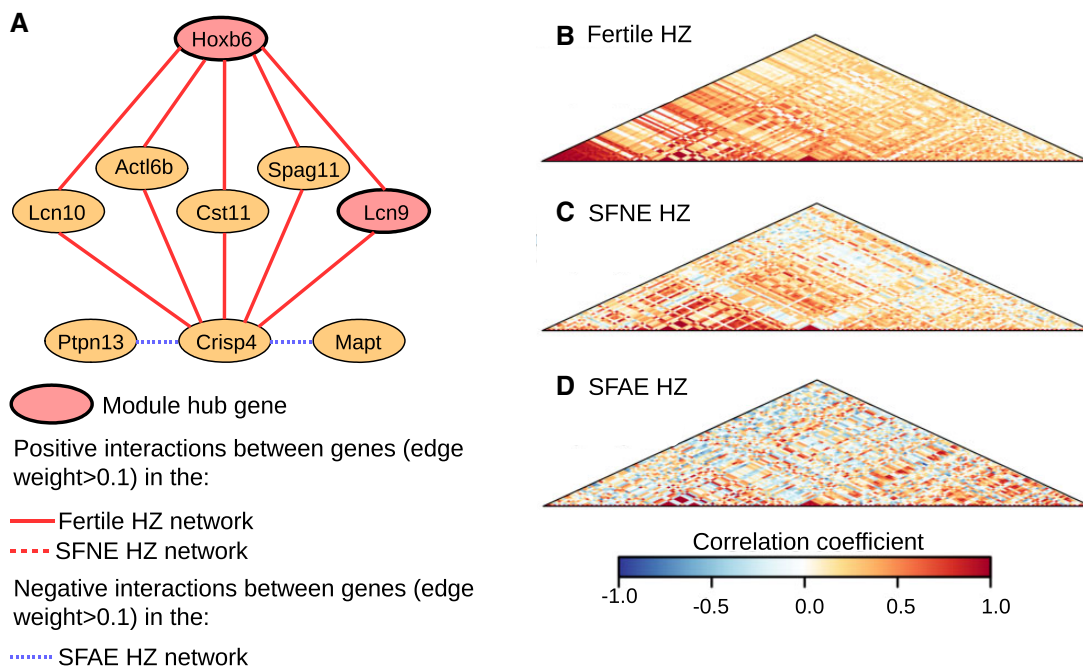


Fig. 5. Disrupted interactions of Midnightblue module genes in subfertile HZ hybrids. (A) Interactions between Midnightblue module hub genes (red nodes) and genes with functions potentially related to sperm maturation (orange nodes) in HZ hybrids. Gene interactions with an edge-weight exceeding 0.1, as estimated using topological overlap matrices, are indicated using continuous, dashed, and fine-dashed lines for the fertile, SFNE, and SFAE hybrids, respectively. Positive interactions are shown in red and negative interactions are shown in blue. (B–D) Coexpression heatmaps showing pairwise correlations between expression values of Midnightblue module genes in fertile, SFNE, and SFAE HZ hybrids, respectively. Heatmaps reveal moderate weakening of gene interactions in the SFNE hybrids and weakening or reversal of interactions in the SFAE HZ hybrids, indicating more severe network disruption in the SFAE hybrid group.

underlying DMIs remain mostly uncharacterized. A network approach is ideal for identifying complex DMIs, as pairwise interactions between genes are likely to be nonindependent. Here, using hybrids of two house mouse subspecies between which reproductive barriers are incomplete (Britton-Davidian et al. 2005; Albrechtova et al. 2012; Turner et al. 2012), we take a network-based approach to characterize gene interactions in testis of fertile versus subfertile males from two hybrid mapping populations. Combining testis expression data from an F_2 cross between wild-derived inbred strains (Turner et al. 2014) and offspring of mice wild-caught from a HZ (Turner and Harr 2014) enabled us to generate a consensus fertile network and identify modules of genes that are commonly coexpressed. We identified disruptions in this network in subfertile hybrids at the module and gene level, some of which are associated with specific functions, stages of spermatogenesis, or testis cell types. Integrating our results with previous mapping of subfertility phenotypes and gene expression traits in the same mice (White et al. 2011; Turner and Harr 2014; Turner et al. 2014) reveals candidate pathways and genes for subfertility.

Network Disruptions in Subfertile Hybrids

Several modules show consistent patterns of poor preservation in subfertile hybrids within both the F_2 and HZ mapping populations, including modules enriched for cilium organization (Brown), spermatogenesis and the regulation of cell cycle (Blue and Yellow), and DNA repair and chromosome segregation (Green). Of particular note is the Brown module, for which expression is significantly correlated with sperm BCF, a measure of sperm motility which has been associated with fertility in humans (Larsen et al. 2000). The disrupted modules are enriched for genes expressed in spermatogonia (Blue, Green, and Yellow), spermatocytes (Blue, Brown, Green, and Yellow), and spermatids (Blue, Brown, Greenyellow, and Yellow). Hence, our findings suggest that multiple stages of spermatogenesis are impacted by DMIs in both natural and laboratory-bred hybrid mice. These findings are somewhat supported by previous histological analyses of testis defects in the F_2 hybrid mapping population, which revealed a range of phenotypic defects linked to reduced fertility (Schwahn et al. 2018). Although the majority of these defects could be explained by a failure to complete meiosis I, so potentially implicating genes expressed in primary spermatocytes, possible downstream errors in meiosis II and postmeiotic errors in spermiogenesis were also implicated (Schwahn et al. 2018).

Although most modules enriched for functions potentially related to spermatogenesis show a lack of preservation in at least one of the subfertile hybrid groups, the Magenta module is an exception. This module is relatively well preserved in all subfertile hybrid groups, despite being enriched for similar processes to the poorly preserved Green module (chromosome organization) and despite being enriched with genes expressed in spermatogonia and spermatocytes. This pattern suggests that while several different stages of spermatogenesis are potentially impacted by DMIs, some pathways and processes remain intact.

Patterns of network disruption are broadly similar in the F_2 and HZ mapping populations, but differences are also evident. The Blue and Yellow modules, for example, showed higher levels of preservation in the F_2 relative to HZ subfertile hybrids, and several module hub genes showed significant changes in coexpression pattern in only one of the hybrid populations. As noted earlier, power to detect disruptions varies among subfertile classes, due to sample size, but we expect there are also true biological differences, because incompatibility loci are segregating within *musculus* and *domesticus* (Good, Handel, et al. 2008; Larson et al. 2018). Moreover, the HZ mapping population is more genetically and phenotypically diverse (Turner et al. 2012), hence the specific genes driving network disruptions may vary across hybrid populations. It is also possible that severe DMIs might have been purged by selection in the HZ, and are consequently detectable in F_2 but not HZ hybrids. This phenomenon may explain the lack of evidence for a prominent role for *Prdm9* in the HZ; neither candidate DMI loci based on genomic clines or GWAS for sterility phenotypes overlap *Prdm9* (Janoušek et al. 2012; Turner and Harr 2014). However, ongoing gene flow from pure subspecies populations is likely to counter or subdue the effects of selection, reintroducing even DMI alleles causing severe deleterious effects into the HZ. Cline analysis of HZ populations revealed high variation between markers across the genome in levels of introgression, and linkage disequilibrium between nonadjacent markers in the center of the HZ (Teeter et al. 2007), supporting the influence of selection on gene flow and the presence and prevalence of DMIs.

Variation in Patterns of Network Disruption

We split hybrid mice with subfertile phenotypes into two broad categories: mice with similar overall patterns of gene expression to those seen in fertile hybrids (SFNE) and those with relatively aberrant patterns of gene expression (SFAE). Unsurprisingly, network disruption appears more severe in the SFAE hybrid groups from both the F_2 and HZ populations, with fewer modules showing evidence of significant preservation. However, there is evidence for more subtle network disruption in the SFNE hybrids. Two modules are not preserved in the SFNE F_2 hybrid group (Midnightblue and Greenyellow, fig. 3). In SFNE HZ hybrids, all modules are preserved but intermediate levels of disruption are apparent upon detailed examination of module-specific coexpression patterns within the Midnightblue module (fig. 5).

We identified more fine-scale disruptions in networks by performing differential correlation analysis, which detects genes showing a significant change in coexpression pattern between groups. Genes showing a significant loss or reversal of coexpression patterns were detected in both SFNE and SFAE hybrid groups relative to fertile hybrids, in both the F_2 and HZ mapping populations, supporting an influence of subtle network disruptions in all subfertile hybrid groups. The 2,800 genes with different interactions in at least one subfertile group include 95 module hub genes, some of which are in relatively well-preserved modules, suggesting hybrid incompatibilities can cause minor or major perturbations in gene interactions. Hence, our findings support varying levels

of network disruption within and among subfertile hybrid groups, as expected because sterility loci are segregating within the mapping populations and putative DMIs previously identified show a range of complexity and effect size (White et al. 2011; Turner and Harr 2014; Turner et al. 2014).

Overlap with Previously Identified Sterility Regions

Several modules, including the Black, Brown, Green, Greenyellow, Red, Purple, and Turquoise modules, have gene content that overlaps significantly with QTT that are associated with a specific *trans*-eQTL hotspot. The majority of these modules show weak (Brown, Green, Greenyellow) or intermediate (Black, Turquoise) preservation in subfertile hybrids, potentially supporting a role for network disruptions involving specific *trans*-eQTL. The Purple and Red modules, however, are well preserved across subfertile hybrid groups and overlap significantly with QTT associated with multiple *trans*-eQTL hotspots. Unlike the more weakly preserved modules, very few of the Purple and Red module genes are expressed during spermatogenesis (Ernst et al. 2019), rather both modules are enriched for genes expressed in Leydig cells. The overall expression of both modules is negatively correlated with both fertility phenotypes, suggesting expression tends to be higher in subfertile relative to fertile hybrids. These observations are consistent with previous reports that genes expressed in somatic cells in testis (i.e., Leydig, Sertoli) have relatively high expression in subfertile hybrids (Turner et al. 2014), likely reflecting reduction/absence of germ cells.

Candidate DMI Genes

We highlight the following as candidate DMI genes: genes that show a significant loss or reversal in coexpression patterns in subfertile relative to fertile hybrid groups; genes that fall within sterility regions previously identified in the HZ and F₂ mapping populations; and genes that are highly connected hubs within disrupted modules. The latter are good candidates for large-effect DMIs, as disrupted epistatic interactions involving these genes may have knock-on effects on module-wide gene expression. Moreover, candidate DMIs previously identified in both the F₂ and HZ mapping populations suggest most DMI loci have more than one incompatible interacting partner locus (Turner and Harr 2014; Turner et al. 2014), as would be expected for hub genes. Module hub genes were found to be significantly more likely than a random sample of testis-expressed genes to fall within GWAS sterility regions, supporting their likely involvement in DMIs. Hub genes previously prioritized as candidate DMI loci include *Nr2c2*, *Zfp711*, *Zfp770*, *Hoxb6*, *Cry2*, and *Gsx1*, all of which lie within *trans*-eQTL hotspots, and *Cyp11a1*, which is located within a GWAS sterility locus on chromosome 9. All of these genes have GO categories that include the regulation of transcription and/or spermatogenesis. The Black module hub gene *Nr2c2* is of particular interest. *Nr2c2* is expressed in mid- and late-stage pachytene spermatocytes and round spermatids (Mu et al. 2004; Ernst et al. 2019) and a lack of expression has been associated with disruptions to late meiotic prophase and consequent delays to spermiogenesis (Mu

et al. 2004). The transcription factors *Zfp711* (Green module), *Zfp770* (Green module), *Hoxb6* (Midnightblue module), and *Gsx1* (Turquoise module) are also good candidate loci for large-effect DMIs, as they show a loss of coexpression patterns in subfertile relative to fertile hybrids and are hub genes within modules that show weak (Green and Midnightblue) or intermediate (Turquoise) preservation in subfertile hybrids.

Trans-eQTL hotspots and many GWAS sterility loci each contain numerous genes (Turner and Harr 2014; Turner et al. 2014), and honing in on the specific causative genes driving DMIs has been challenging. Genes that lie within candidate regions but lack GOs relating to the regulation of transcription and/or spermatogenesis are likely to have been overlooked. Our study identified several hub genes that lie within previously identified sterility regions yet have not previously been highlighted as likely candidate DMI loci. One such gene is *Prkcd*, a hub gene within the poorly conserved Brown module that lies within a GWAS sterility region on chromosome 14 (Turner and Harr 2014). Although this gene lacks GOs related to spermatogenesis, a knockout study has reported a role for *Prkcd* in male fertility. Specifically, the sperm of male mice lacking *Prkcd* expression have a reduced ability to penetrate the zona pellucida, potentially impairing fertilization (Ma et al. 2015). *Hk1*, a hub gene within the Blue module, lies within a *trans*-eQTL hotspot on chromosome 10 yet also lacks GOs related to spermatogenesis. This gene encodes the enzyme that initiates the glycolysis pathway, which is important for sperm motility (Mori et al. 1998), hence suggesting an important role in male fertility. Other genes that lie within *trans*-eQTL hotspots and/or GWAS sterility regions yet have been overlooked as candidate DMI loci include *Rapgef1* and *Mapkap1*, which are involved in the regulation of signal transduction and establishment of actin cytoskeleton polarity, respectively, *Akap1*, which is known to be involved in meiosis in female mice (Newhall et al. 2006), and *Atr*, which plays a role in preventing DNA damage and is associated with reduced testis weight, abnormal DNA replication and cell cycle in knockout male mice (Murga et al. 2009).

Although several hub genes within the poorly preserved Midnightblue module lie within previously identified sterility regions, only *Hoxb6* has been highlighted as a candidate DMI locus. *Adam28*, *Lcn10*, and *Lcn9* all lie within GWAS sterility regions, and both *Adam28* and *Lcn9* show a loss of positive coexpression patterns in subfertile relative to fertile hybrids. Although all three genes lack GOs relating to spermatogenesis, *Adam28*, *Lcn9*, and *Lcn10* are all known to be expressed in the epididymis, and are likely to be involved in male fertility (Suzuki et al. 2004; Oh et al. 2005). Although our expression data are from testis, we suspect most of these genes are expressed in both tissues as testes are relatively easy to separate from epididymis during dissection. In fertile hybrids, the Midnightblue module hub genes are coexpressed with several genes thought to be involved in sperm maturation, including *Cst11* and *Spag11*, both of which have antimicrobial activity and are thought to be important for sperm maturation in other mammal species (Hamil et al. 2002; Avellar et al. 2007), and *Crisp4*, which is implicated in the acrosome reaction

required for the binding of sperm to the zona pellucida (Turunen et al. 2012). Hence disrupted interactions in Midnightblue module hub genes may have an impact on sperm motility and functioning.

Our gene network analysis is largely independent of phenotype data, hence potentially enabling us to identify entirely novel candidate DMI loci. Candidate DMI genes that lie outside of previously identified sterility regions include *Fkbp4* and *Ptdss2*, which are hub genes in the poorly preserved Brown module and show loss of coexpression patterns in subfertile relative to fertile hybrids. Mice lacking *Ptdss2* expression have reduced testis weight and can be infertile (Bergo et al. 2002), whereas the lack of *Fkbp4* expression is associated with abnormal sperm morphology (Hong et al. 2007). *Helq*, a differentially correlated hub gene in the Turquoise module, also lies outside of previously identified sterility regions and is associated with subfertile phenotypes in male mice (Adelman et al. 2013). Finally, *D1Pas1* and *Adam1a*, both hub genes in the Blue module have GOs relating to spermatogenesis and the binding of sperm to the zona pellucida, respectively, yet have not been previously identified as candidates for DMI speciation genes.

Conclusion

We demonstrate widespread disruptions to gene-interaction networks in association with reduced fertility in hybrid *musculus–domesticus* house mice. Disruptions are variable in magnitude among hybrid mapping populations and appear to affect multiple stages of spermatogenesis, including chromosome segregation and cell cycle, assembly of the sperm flagellum, and sperm maturation. We identify specific candidate DMI genes, several of which fall within previously identified sterility loci and have been previously associated with reduced fertility phenotypes in male mice.

Materials and Methods

Microarray and Phenotype Data

We used testis gene expression data from two previous studies, the first included F₂ hybrid males from a cross between wild-derived inbred strains of *M. m. domesticus* (WSB/Eij) and *M. m. musculus* (PWD/Phj) (White et al. 2011; Turner et al. 2014) and the second included first-generation offspring of mice wild-caught in the HZ (Turner and Harr 2014). Gene expression in the testis of hybrid males sacrificed at a similar developmental stage (70 ± 5 days and 9–12 weeks for F₂ and HZ mice, respectively) was measured using Whole Mouse Genome Microarrays (Agilent), as described in Turner et al. (2014) and Turner and Harr (2014).

To investigate changes in gene expression networks associated with sterility, we first classified individuals as “fertile” versus “subfertile.” As fertility (i.e., ability to father offspring) was not directly measured in F₂ or HZ individuals, we used two phenotypes to categorize males as fertile or subfertile: relative testis weight (testis weight/body weight) and sperm count (White et al. 2011; Turner et al. 2012). These phenotypes were measured comparably in mice from both mapping populations and have been associated with reduced fertility

in multiple studies of *musculus–domesticus* hybrids (Britton-Davidian et al. 2005; reviewed in Good, Dean, et al. 2008; Good et al. 2010); we will henceforth refer to these traits as “fertility phenotypes.” A total of 102 F₂ and 79 HZ hybrid males have fertility phenotypes that fall within 1 SD of the pure subspecies mean and were categorized as “fertile.” “Subfertile” hybrids, for which one or both of the sterility phenotypes fall outside of the pure subspecies range include 107 F₂ and 55 HZ males. For a total of 92 F₂ and 41 HZ hybrids, both sterility phenotypes fall within the pure subspecies range yet at least one of the phenotypes is >1 SD from the pure subspecies mean. These hybrids were categorized as “intermediate phenotype.” Finally, data were available for 32 pure subspecies males: 16 *domesticus* and 16 *musculus*. Eight individuals each of pure *domesticus* and pure *musculus* were offspring of mice wild-caught at the edges of the HZ (Turner and Harr 2014), and the remaining pure subspecies males were from wild-derived inbred strains WSB/Eij (*domesticus*) and PWD/Phj (*musculus*), whose expression was reported in Turner et al. (2014). In total, microarray and sterility phenotype data were available for 467 hybrid and 32 pure subspecies males.

Microarray Data Processing

The Whole Mouse Genome Microarray (Agilent) contains 43,379 probes including 22,210 transcripts from 21,326 genes. We started from raw array data from each study rather than processed expression values, to ensure data sets were comparable in network analyses. Preprocessing of raw expression data was performed in the R package *limma* (Smyth 2005). Background correction was performed by specifying the “half” setting, which resets intensities that fall <0.5 following background subtraction to 0.5, and by adding an offset of 50. To identify probes with consistently low expression, the 98th percentile of the expression of negative control probes was calculated and only probes that were at least 10% brighter than this background expression level in at least 10% of sampled individuals were retained, reducing the data set to a total of 36,896 probes. The quantile method was used to normalize expression between arrays.

As the expression data set includes data generated within different laboratories and over different time periods, non-biological systematic bias or “batch effects” must be considered. We adjusted for known batch effects using the empirical Bayes framework implemented via the ComBat function (Johnson et al. 2007). We also tried detecting and adjusting for hidden batch effects using the SVA R package (Leek and Storey 2007) and obtained similar results, with several of the detected surrogate variables clustering the data by the known batches. Adjusting for the batch effect may result in losing potential heterogeneity in gene expression between the F₂ and HZ mapping populations. Nevertheless, this effect should be equal for both fertile and subfertile phenotypes, and should therefore have minimal impacts on observed patterns of network disruption in subfertile hybrids. Variation in gene expression across individuals was summarized using PCA, as implemented using the *prcomp* function in R (R Core Team 2018), which uses a singular value decomposition of the

centered and scaled data matrix, and extreme outliers were identified visually and removed prior to downstream analyses. Batch-corrected and normalized expression data are available on the Gene Expression Omnibus (GEO) (GSE136886).

Weighted Gene Coexpression Network Analysis

We used The WGCNA R package (Langfelder and Horvath 2008) to identify groups of genes, “modules,” showing similar patterns of expression within and across the F₂ and HZ fertile hybrid groups. As network analysis is computationally intensive, we further filtered the data set prior to network construction. Specifically, the connectivity of each probe was estimated using the softConnectivity function, which is available within the WGCNA R package (Langfelder and Horvath 2008) and calculates the sum of the adjacency of each probe to all other probes within the data set. The median connectivity was calculated and probes with above-average connectivity within fertile hybrids were retained, resulting in a final data set of 18,411 probes representing 10,531 genes. We used the blockwiseConsensusModules function to perform signed network construction and identify consensus modules across the fertile F₂ and HZ data sets, assigning each probe to a single module. Briefly, this process involved calculating a pairwise coexpression matrix for each of the F₂ and HZ fertile groups, in which coexpression is estimated using Pearson correlation values. Raising the coexpression matrix to a defined soft-threshold power introduces scale-free topology, in which a small proportion of nodes (hub genes) have a large number of connections within the network (Zhang and Horvath 2005). Such scale-free topology is thought to be a fundamental property of most biological networks (Barabasi and Albert 1999). Network topology analysis was performed for a range of soft-threshold values, and an optimal soft-threshold value of five was chosen as the lowest value at which median connectivity reached a low plateau. The coexpression matrix was raised to this soft-threshold power to create an adjacency matrix, which was then converted to a topological overlap matrix (TOM). The TOM describes the network interconnectivity or coexpression between each pair of genes in relation to all others in the network. A consensus TOM was then used to cluster genes using average linkage hierarchical clustering. A dynamic tree cutting algorithm was used to cut the clustering tree, so defining consensus modules of similarly expressed genes. The deepSplit and minimum module size parameters were set to 0 and 50, respectively, and the ME distance threshold was set to 0.2 to merge similar modules.

Module eigengenes are defined as the first principal component describing the expression of a given module. To determine whether specific modules are more or less expressed in fertile versus subfertile individuals, Pearson correlations were computed between the eigengene of each module and each of the two fertility phenotypic trait values (relative testis weight and sperm count). To determine whether the expression of specific modules was associated with sperm motility, we took phenotype data available for a subset ($n = 121$) of HZ hybrid males from Turner et al. (2012) and used Pearson correlation tests to estimate the significance of relationships between the expression of each module and the

following sperm motility traits: track velocity (VCL), smoothed path velocity (VAP), straight line velocity (VSL), amplitude of latitudinal head displacement (ALH), and beat-cross frequency (BCF).

GO Enrichment Analysis

We tested for functional enrichment within each module on the basis of GO overrepresentation analysis with Benjamini–Hochberg P -value adjustment (Boyle et al. 2004), performed using the enrichGO function from the clusterProfiler R package (Yu et al. 2012). We performed this analysis using two “gene universes” as the background: firstly, a gene universe consisting of the 17,514 genes with above-background levels of expression in the testis and secondly, a universe consisting of all 21,200 genes associated with the G4122F microarray (Agilent Whole Mouse Genome Microarray), as established using the Gene Expression Omnibus entry for that platform (Edgar et al. 2002; Barrett et al. 2012). GO terms associated with more than 10 and fewer than 500 genes in the gene universe were available for assignment. As there were few clear differences in the identity and significance of GO annotations using the full data set of genes versus the testis-expressed subset, results are presented for the analysis using the testis-expressed gene universe only.

Network Preservation between Fertile and Low-Fertility Hybrid Groups

To identify gene interactions which are present in fertile hybrids and disrupted in hybrids with low fertility, we tested for preservation of modules from the fertile network in intermediate and subfertile phenotype hybrids. Levels of genetic variability are likely to vary between the HZ and F₂ mapping populations, as the F₂ hybrids were created through crosses of inbred *domesticus* and *musculus* strains, whereas the HZ population was created through crosses of mice caught wild in the HZ. As incompatibility loci are likely to be segregating in natural *domesticus* and *musculus* populations, the presence or absence of specific sterility loci is likely to differ between the mapping populations. We therefore tested for module preservation in subfertile hybrids independently within the F₂ and HZ populations. As the PCA revealed a strong association between the low-fertility phenotype and variation along PC1 (see Results), we further split the subfertile individuals into those clustering together with the fertile individuals along PC1 and those with a PC1 score that falls outside of the fertile range (−91.25 to 49.33; see fig. 1). We refer to these groupings as SFNE and SFAE. To explore whether the lack of preservation for several modules in the subfertile phenotypes was exclusive to the SFAE group, we tested for module preservation between the fertile and each of the SFNE and SFAE groups within the F₂ and HZ populations.

We used the statistical frameworks implemented in the WGCNA and NetRep R packages to estimate module preservation (Langfelder and Horvath 2008; Ritchie et al. 2016). Both of these permutation-based approaches use the seven preservation metrics developed by Langfelder et al. (2011). The modulePreservation function (WGCNA package, 500 permutations) was used to generate Z_{summary} scores, which

combine several preservation statistics that compare the density and pattern of connections within modules and between data sets. Z_{summary} scores of ≥ 10 , $2-10$, and < 2 indicate strong, weak, and a lack of module preservation between data sets, respectively (Langfelder et al. 2011). Modules were ranked according to their relative preservation using the median rank statistic, which is based on the Z_{summary} score and module size (Langfelder et al. 2011).

In addition, we used the NetRep R package (Ritchie et al. 2016) to test the significance of all seven of Langfelder's statistics summarizing the preservation of modules between test and discovery data sets. If one or more of the NetRep statistics was found to be nonsignificant, then this was considered evidence for a lack of significant module preservation in subfertile hybrids. The Z_{summary} scores and least significant NetRep statistics for each module are presented in supplementary table 1, Supplementary Material online.

Differential Correlation Analysis

Genes showing significantly different patterns of pairwise coexpression in the subfertile relative to the fertile hybrids were identified using the R package DGCA (McKenzie et al. 2016). Once again, differential correlation analyses were performed independently for each of the F_2 and HZ populations, and for each of the SFNE and SFAE hybrid groups. The median log-fold change in pairwise coexpression was estimated for each gene in each of the 15 modules, and the significance of median log-fold change values was estimated using 100 permutations.

Identification of Module Hub Genes

Two methods were used to identify hub genes. First, genes with a Module Membership (kME) ≥ 0.85 were identified as hub genes, where kME represents the Pearson correlation between the expression of an individual gene and the ME (Horvath and Dong 2008). Second, connectivity statistics including the average number of neighbors, which describes the average connectivity of nodes in a module, and the network density, which summarizes the overall module connectivity, were calculated independently for fertile F_2 and HZ hybrids, using Cytoscape v3.7.1 (Shannon et al. 2003). The top five most connected genes within each of the F_2 and HZ networks were also classified as hub genes for each module. We then used permutation tests to determine whether the module hub genes were significantly more likely than a random subset of genes to fall within either GWAS sterility regions or *trans*-eQTL hotspots. We performed 10,000 random draws of 281 genes from the full set of testis-expressed genes, as well as from the subset of genes that were assigned to 1 of the 15 modules in the consensus fertile network. We then calculated the 95% quantiles for the number of genes within GWAS sterility region genes and the number of genes within eQTL hotspots for these randomly drawn gene sets, and assumed significant enrichment of GWAS sterility region or eQTL hotspot genes if the values observed in the module gene data set equaled or exceeded these 95% quantiles.

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We are grateful to colleagues in the Milner Centre for Evolution, University of Bath, for valuable feedback and discussions. We would also like to thank the anonymous reviewers for their helpful comments and suggestions. Expression data processed in this study have been submitted to the Gene Expression Omnibus (GSE136886). An R script for the analyses performed in this study is included in Supplementary Material online. This work was supported by funding from the Deutsche Forschungsgemeinschaft (TU500/2-1 awarded to L.M.T.).

References

- Adelman CA, Lolo RL, Birkbak NJ, Murina O, Matsuzaki K, Horejsi Z, Parmar K, Borel V, Skehel JM, Stamp G, et al. 2013. *HELQ* promotes *RAD51* paralogue-dependent repair to avert germ cell loss and tumorigenesis. *Nature* 502(7471):381–384.
- Albrechtova J, Albrecht T, Baird SJ, Macholan M, Rudolfson G, Munclinger P, Tucker PK, Pialek J. 2012. Sperm-related phenotypes implicated in both maintenance and breakdown of a natural species barrier in the house mouse. *Proc R Soc B*. 279(1748):4803–4810.
- Amaral A, Lourenço B, Marques M, Ramalho-Santos J. 2013. Mitochondria functionality and sperm quality. *Reproduction* 146(5):R163–R174.
- Avellar MC, Honda L, Hamil KG, Radhakrishnan Y, Yenugu S, Grossman G, Petrusz P, French FS, Hall SH. 2009. Systems genetics of complex traits in *Drosophila melanogaster*. *Nat Genet*. 41:299.
- Avellar MCW, Honda L, Hamil KG, Radhakrishnan Y, Yenugu S, Grossman G, Petrusz P, French FS, Hall SH. 2007. Novel aspects of the *sperm-associated antigen 11 (SPAG11)* gene organization and expression in cattle (*Bos taurus*). *Biol Reprod*. 76(6):1103–1116.
- Ayroles JF, Carbone MA, Stone EA, Jordan KW, Lyman RF, Magwire MM, Rollmann SM, Duncan LH, Lawrence F, Anholt RR, et al. 2009. Systems genetics of complex traits in *Drosophila melanogaster*. *Nat Genet*. 41(3):299.
- Barabasi AL, Albert R. 1999. Emergence of scaling in random networks. *Science* 286:509–512.
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, et al. 2012. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res*. 41(D1):D991–D995.
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. 2010. *PRDM9* is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327(5967):836–840.
- Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Methodol*. 57(1):289–300.
- Bergo MO, Gavino BJ, Steenbergen R, Sturbois B, Parlow AF, Sanan DA, Skarnes WC, Vance JE, Young SG. 2002. Defining the importance of *phosphatidylserine synthase 2* in mice. *J Biol Chem*. 277(49):47701–47708.
- Bhattacharyya T, Gregorova S, Mihola O, Anger M, Sebestova J, Denny P, Simecek P, Forejt J. 2013. Mechanistic basis of infertility of mouse intersubspecific hybrids. *Proc Natl Acad Sci U S A*. 110(6):E468–E477.
- Bhattacharyya T, Reifova R, Gregorova S, Simecek P, Gergelits V, Mistrik M, Martincova I, Pialek J, Forejt J. 2014. X chromosome control of meiotic chromosome synapsis in mouse inter-subspecific hybrids. *PLoS Genet*. 10(2):e1004088.
- Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, Sherlock G. 2004. GO: termFinder—open source software for accessing Gene

- Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. *Bioinformatics* 20(18):3710–3715.
- Britton-Davidian J, Fel-Clair F, Lopez J, Alibert P, Boursot P. 2005. Postzygotic isolation between the two European subspecies of the house mouse: estimates from fertility patterns in wild and laboratory-bred hybrids. *Biol J Linn Soc.* 84(3):379–393.
- Campbell P, Good JM, Nachman MW. 2013. Meiotic sex chromosome inactivation is disrupted in sterile hybrid male house mice. *Genetics* 193(3):819–828.
- Davies B, Hlatton E, Altemose N, Hussin JG, Pratto F, Zhang G, Hinch AG, Moralli D, Biggs D, Diaz R, et al. 2016. Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature* 530(7589):171–176.
- Davis BW, Seabury CM, Brashear WA, Li G, Roelke-Parker M, Murphy WJ. 2015. Mechanisms underlying mammalian hybrid sterility in two feline interspecies models. *Mol Biol Evol.* 32(10):2534–2546.
- De la Fuente A. 2010. From 'differential expression' to 'differential networking'—identification of dysfunctional regulatory networks in diseases. *Trends Genet.* 26(7):326–333.
- Dobzhansky TH. 1937. *Genetics and the Origin of Species*. New York: Columbia University Press.
- Edgar R, Domrachev M, Lash AE. 2002. Gene Expression Omnibus: nCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* 30(1):207–210.
- Ernst C, Eling N, Martinez-Jimenez CP, Marioni JC, Odom DT. 2019. Staged developmental mapping and X chromosome transcriptional dynamics during mouse spermatogenesis. *Nat Commun.* 10(1):1251.
- Filteau M, Pavay SA, St-Cyr J, Bernatchez L. 2013. Gene coexpression networks reveal key drivers of phenotypic divergence in lake whitefish. *Mol Biol Evol.* 30(6):1384–1396.
- Forejt J, Ivanyi P. 1974. Genetic studies on male sterility of hybrids between laboratory and wild mice (*Mus musculus* L.). *Genet Res.* 24(2):189–206.
- Fraïsse C, Elderfield JAD, Welch JJ. 2014. The genetics of speciation: are complex incompatibilities easier to evolve? *J Evol Biol.* 27(4):688–699.
- Geraldes A, Basset P, Smith KL, Nachman MW. 2011. Higher differentiation among subspecies of the house mouse (*Mus musculus*) in genomic regions with low recombination. *Mol Ecol.* 20(22):4722–4736.
- Gomes S, Civetta A. 2015. Hybrid male sterility and genome-wide misexpression of male reproductive proteases. *Sci Rep.* 5(1):11976.
- Good JM, Dean MD, Nachman MW. 2008. A complex genetic basis to X-linked hybrid male sterility between two species of house mice. *Genetics* 179(4):2213–2228.
- Good JM, Giger T, Dean MD, Nachman MW. 2010. Widespread over-expression of the X chromosome in sterile F1 hybrid mice. *PLoS Genet.* 6(9):e1001148.
- Good JM, Handel MA, Nachman MW. 2008. Asymmetry and polymorphism of hybrid male sterility during the early stages of speciation in house mice. *Evolution* 62:50–65.
- Gould BA, Chen Y, Lowry DB. 2018. Gene regulatory divergence between locally adapted ecotypes in their native habitats. *Mol Ecol.* 27(21):4174–4188.
- Hamil KG, Liu Q, Sivashanmugam P, Yenugu S, Soundararajan R, Grossman G, Richardson RT, Zhang YL, O'Rand MG, Petrusz P, et al. 2002. *Cystatin 11*: a new member of the cystatin type 2 family. *Endocrinology* 143(7):2787–2796.
- Hong J, Kim ST, Tranguch S, Smith DF, Dey SK. 2007. Deficiency of co-chaperone immunophilin FKBP52 compromises sperm fertilizing capacity. *Reproduction* 133(2):395–403.
- Horvath S, Dong J. 2008. Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol.* 4(8):e1000117.
- Janoušek V, Wang L, Luzynski K, Dufková P, Vyskočilová MM, Nachman MW, Munclinger P, Macholán M, Piálek J, Tucker PK. 2012. Genome-wide architecture of reproductive isolation in a naturally occurring hybrid zone between *Mus musculus musculus* and *M. m. domesticus*. *Mol Ecol.* 21(12):3032–3047.
- Johnson NA, Porter AH. 2000. Rapid speciation via parallel, directional selection on regulatory genetic pathways. *J Theor Biol.* 205(4):527–542.
- Johnson NA, Porter AH. 2006. Evolution of branched regulatory genetic pathways: directional selection on pleiotropic loci accelerates developmental system drift. *Genetica* 129(1):57–70.
- Johnson WE, Li C, Rabinovic A. 2007. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8(1):118–127.
- Jumbo-Lucioni P, Ayroles JF, Chambers MM, Jordan KW, Leips J, Mackay TF, De Luca M. 2010. Systems genetics analysis of body weight and energy metabolism traits in *Drosophila melanogaster*. *BMC Genomics* 11(1):297.
- Kalirad A, Azevedo RB. 2017. Spiraling complexity: a test of the snowball effect in a computational model of RNA folding. *Genetics* 206(1):377–388.
- Kelley JL, Arias-Rodriguez L, Patacsil Martin D, Yee MC, Bustamante CD, Tobler M. 2016. Mechanisms underlying adaptation to life in hydrogen sulfide-rich environments. *Mol Biol Evol.* 33(6):1419–1434.
- Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9(1):559.
- Langfelder P, Luo R, Oldham MC, Horvath S. 2011. Is my network module preserved and reproducible? *PLoS Comput Biol.* 7(1):e1001057.
- Larsen L, Scheike T, Jensen TK, Bonde JP, Ernst E, Hjollund NH, Zhou Y, Skakkebaek NE, Giwercman A; Danish First Pregnancy Planner Study Team. 2000. Computer-assisted semen analysis parameters as predictors for fertility of men from the general population. *Hum Reprod.* 15(7):1562–1567.
- Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM. 2017. The composite regulatory basis of the large X-effect in mouse speciation. *Mol Biol Evol.* 34(2):282–295.
- Larson EL, Vanderpool D, Sarver BA, Callahan C, Keeble S, Provencio LL, Kessler MD, Stewart V, Nordquist E, Dean MD, et al. 2018. The evolution of polymorphic hybrid incompatibilities in house mice. *Genetics* 209(3):845–859.
- Lee Y, Li H, Li J, Rebman E, Achour I, Regan KE, Gamazon ER, Chen JL, Yang XH, Cox NJ, et al. 2013. Network models of genome-wide association studies uncover the topological centrality of protein interactions in complex diseases. *J Am Med Inform Assoc.* 20(4):619–629.
- Leek JT, Storey JD. 2007. Capturing heterogeneity in gene expression studies by surrogate variable analysis. *PLoS Genet.* 3(9):e161.
- Ma W, Baumann C, Viveiros MM. 2015. Lack of protein kinase C-delta (PKC δ) disrupts fertilization and embryonic development. *Mol Reprod Dev.* 82(10):797–808.
- Macholán M, Kryštufek B, Vohralík V. 2003. The location of the *Mus musculus/M. domesticus* hybrid zone in the Balkans: clues from morphology. *Acta Theriol.* 48(2):177–188.
- Macholán M, Munclinger P, Šugerková M, Dufková P, Bímová B, Božíková E, Zima J, Piálek J. 2007. Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution* 61(4):746–771.
- Mack KL, Ballinger MA, Phifer-Rixey M, Nachman MW. 2018. Gene regulation underlies environmental adaptation in house mice. *Genome Res.* 28(11):1636–1645.
- Mack KL, Campbell P, Nachman MW. 2016. Gene regulation and speciation in house mice. *Genome Res.* 26(4):451–461.
- Mackay TF. 2014. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat Rev Genet.* 15(1):22–33.
- Maheshwari S, Barbash DA. 2011. The genetics of hybrid incompatibilities. *Annu Rev Genet.* 45(1):331–355.
- McKenzie AT, Katsyov I, Song WM, Wang M, Zhang B. 2016. DGCA: a comprehensive R package for differential gene correlation analysis. *BMC Syst Biol.* 10(1):106.
- Michalak P, Noor MA. 2003. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. *Mol Biol Evol.* 20(7):1070–1076.
- Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* 323(5912):373–375.

- Miller JA, Horvath S, Geschwind DH. 2010. Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc Natl Acad Sci U S A*. 107(28):12698–12703.
- Moehring AJ, Teeter KC, Noor MA. 2007. Genome-wide patterns of expression in *Drosophila* pure species and hybrid males. II. Examination of multiple-species hybridizations, platforms, and life cycle stages. *Mol Biol Evol*. 24(1):137–145.
- Mori C, Nakamura N, Welch JE, Gotoh H, Goulding EH, Fujioka M, Eddy EM. 1998. Mouse spermatogenic cell-specific type 1 hexokinase (*mHk1-s*) transcripts are expressed by alternative splicing from the *mHk1* gene and the HK1-S protein is localized mainly in the sperm tail. *Mol Reprod Dev*. 49(4):374–385.
- Mu X, Lee YF, Liu NC, Chen YT, Kim E, Shyr CR, Chang C. 2004. Targeted inactivation of testicular nuclear orphan receptor 4 delays and disrupts late meiotic prophase and subsequent meiotic divisions of spermatogenesis. *Mol Cell Biol*. 24(13):5887–5899.
- Muller HJ. 1942. Isolating mechanisms, evolution and temperature. *Biol Symp*. 6:71–125.
- Murga M, Bunting S, Montaña MF, Soria R, Mulero F, Cañamero M, Lee Y, McKinnon PJ, Nussenzweig A, Fernandez-Capetillo O. 2009. A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. *Nat Genet*. 41(8):891–898.
- Newhall KJ, Criniti AR, Cheah CS, Smith KC, Kafer KE, Burkart AD, McKnight GS. 2006. Dynamic anchoring of PKA is essential during oocyte maturation. *Curr Biol*. 16(3):321–327.
- Oh J, Woo J-M, Choi E, Kim T, Cho B-N, Park ZY, Kim YC, Kim DH, Cho C. 2005. Molecular, biochemical, and cellular characterization of epididymal ADAMs, ADAM7 and ADAM28. *Biochem Biophys Res Commun*. 331(4):1374–1383.
- Oka A, Mita A, Sakurai-Yamatani N, Yamamoto H, Takagi N, Takano-Shimizu T, Toshimori K, Moriwaki K, Shiroishi T. 2004. Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics* 166(2):913–924.
- Orr HA. 1995. The population genetics of speciation: the evolution of hybrid incompatibilities. *Genetics* 139(4):1805–1813.
- Payseur BA, Krenz JG, Nachman MW. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* 58(9):2064–2078.
- Phadnis N. 2011. Genetic architecture of male sterility and segregation distortion in *Drosophila pseudoobscura* Bogota–USA hybrids. *Genetics* 189(3):1001–1009.
- Porter AH, Johnson NA. 2002. Speciation despite gene flow when developmental pathways evolve. *Evolution* 56(11):2103–2111.
- Presgraves DC. 2010. Darwin and the origin of interspecific genetic incompatibilities. *Am Nat*. 176(S1):S45–S60.
- R Core Team. 2018. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.
- Ritchie SC, Watts S, Fearnley LG, Holt KE, Abraham G, Inouye M. 2016. A scalable permutation approach reveals replication and preservation patterns of network modules in large datasets. *Cell Syst*. 3(1):71–82.
- Saris CG, Horvath S, van Vught PW, van Es MA, Blauw HM, Fuller TF, Langfelder P, DeYoung J, Wokke JHJ, Veldink JH, et al. 2009. Weighted gene co-expression network analysis of the peripheral blood from Amyotrophic Lateral Sclerosis patients. *BMC Genomics* 10(1):405.
- Schwahn DJ, Wang RJ, White MA, Payseur BA. 2018. Genetic dissection of hybrid male sterility across stages of spermatogenesis. *Genetics* 210(4):1453–1465.
- Serres C, Feneux D, Jouannet P, David G. 1984. Influence of the flagellar wave development and propagation on the human sperm movement in seminal plasma. *Gamete Res*. 9(2):183–195.
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*. 13(11):2498–2504.
- Smyth GK. 2005. Limma: linear models for microarray data. In: Bioinformatics and computational biology solutions using R and bioconductor. In: Gentleman R, Carey V, Dudoit S, R Irizarry WH, editors. New York: Springer. p. 397–420.
- Suzuki K, Lareyre JJ, Sánchez D, Gutierrez G, Araki Y, Matusik RJ, Orgebin-Crist MC. 2004. Molecular evolution of epididymal lipocalin genes localized on mouse chromosome 2. *Gene* 339:49–59.
- Teeter KC, Payseur BA, Harris LW, Bakewell MA, Thibodeau LM, O'Brien JE, Krenz JG, Sans-Fuentes MA, Nachman MW, Tucker PK. 2007. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res*. 18(1):67–76.
- Turner LM, Harr B. 2014. Genome-wide mapping in a house mouse hybrid zone reveals hybrid sterility loci and Dobzhansky-Muller interactions. *Elife* 3:e02504.
- Turner LM, Schwahn DJ, Harr B. 2012. Reduced male fertility is common but highly variable in form and severity in a natural house mouse hybrid zone. *Evolution* 66(2):443–458.
- Turner LM, White MA, Tautz D, Payseur BA. 2014. Genomic networks of hybrid sterility. *PLoS Genet*. 10(2):e1004162.
- Turunen HT, Sipilä P, Krutskikh A, Toivanen J, Mankonen H, Hämäläinen V, Björkgren I, Huhtaniemi I, Poutanen M. 2012. Loss of *cysteine-rich secretory protein 4* (*Crisp4*) leads to deficiency in sperm-zona pellucida interaction in mice. *Biol Reprod*. 86:12–11.
- Vyskočilová M, Trachtulec Z, Forejt J, Piálek J. 2005. Does geography matter in hybrid sterility in house mice? *Biol J Linn Soc*. 84(3):663–674.
- White MA, Steffy B, Wiltshire T, Payseur BA. 2011. Genetic dissection of a key reproductive barrier between nascent species of house mice. *Genetics* 189(1):289–304.
- Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16(5):284–287.
- Zhang B, Horvath S. 2005. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. 4:Article17.
- Zhao Y, Xu F, Liu J, Guan F, Quan H, Meng F. 2019. The adaptation strategies of *Herpetospermum pedunculatum* (Ser.) Baill at altitude gradient of the Tibetan plateau by physiological and metabolomic methods. *BMC Genomics* 20(1):451.