Sex differences in disease genetics: Evidence, evolution and detection

William P. Gilks†, Jessica K. Abbott*, Edward H. Morrow†

† Evolution, Behaviour and Environment group, School of Life Sciences, University of Sussex, Falmer BN1 9QG, United Kingdom
* Experimental Evolution, Ecology & Behaviour group, Department of Biology, Section for Evolutionary Ecology, Lund University, Sölvegatan 37, 22362 Lund, Sweden

Corresponding author: Gilks, W.P. (w.gilks@sussex.ac.uk)

Keywords
Intra-locus sexual conflict; GWAS; sexual selection; GxS; complex trait; disease.
Abstract

Understanding the genetic architecture of disease is an enormous challenge, and should be guided by evolutionary principles. Recent studies in evolutionary genetics show that sexual selection can have a profound influence on the genetic architecture of complex traits. Here, we summarise data from heritability studies and genome-wide association studies showing that common genetic variation influences many diseases and medically relevant traits in a sex-dependent manner. In addition, we discuss how the discovery of sex-dependent effects in population samples is improved by joint interaction analysis (rather than separate-sex), as well as by recently developed software. Finally, we argue that although genetic variation that has sex-dependent effects on disease risk could be maintained by mutation-selection balance and genetic drift, recent evidence indicates that intra-locus sexual conflict could be a powerful influence on complex trait architecture, and maintain sex-dependent disease risk alleles in a population because they are beneficial to the opposite sex.
Can sex differences explain the missing heritability?

Heritable diseases are loosely classified as being rare or common (prevalence >0.1%). Rare diseases have a monogenic aetiology, whereas common diseases are caused by multiple genetic variants, each with high population frequency but small individual contribution to disease risk [1,2]. For the latter, genome-wide association studies (GWAS) (Glossary) have been successful at identifying contributing loci, but the heritability accounted for by main effects, and by polygenic risk score, remains conspicuously low [3,4]. This deficit (generally referred to as ‘missing heritability’) is stimulating integration of other evidence-based factors such as the environment, epigenetics, and epistasis into analyses [5]. Here, we consider the role of sex (gender), in the genetic architecture of common, heritable medical disorders.

The difference in gamete size between males and females is a fundamental property of almost all sexual species. Sexual dimorphism also exists throughout the body in cellular and anatomical specialisation, secondary sexual traits such as ornamentation and behaviour, and in gene co-expression networks [6-8]. It is therefore unsurprising that in the field of medicine, males and females frequently differ in core phenotypic features of disease [9]. Appreciating the magnitude and extent of these sex differences is important for the effective design of therapies, but at a fundamental level, it would also add to our understanding of how these differences evolve.

The simplest way in which a sex-dependent disease risk allele can be maintained in frequency is through mutation-selection balance and genetic drift. Selection alone is not a necessary condition, because a new allele can easily have a
sex-dependent effect regardless of the selection on the trait that it might affect. An alternative mechanism for the maintenance of sex-dependent risk alleles is sexual antagonism, whereby an allele that is deleterious to one sex is maintained because it is beneficial to the other sex (Box 1) [10,11]. We refer here to intra-locus sexual conflict because it occurs across a single locus, in contrast with inter-locus sexual conflict, which concerns conflict between different sets of genes in males and females, e.g. competition between seminal fluid and the female immune system in Drosophila melanogaster) [12]. An example of intra-locus sexual conflict in humans is relative body height, which is positively selected in men, yet negatively selected in women despite being controlled by the same molecular genetic variation [13].

Insights from evolutionary biology are of great value here because theory about the ultimate origin and evolution of sex differences is well developed, both on the phenotypic and on the genetic level. Asymmetrical selection pressures operating between the sexes on genetic variants offer a long-term, evolutionary explanation for the existence of sexually dimorphic phenotypes, including those identified in human diseases. Sex differences in the genetic architecture of common diseases have been known for some time [14], and recent analysis of large GWAS datasets has resulted in an unprecedented rise in the identification of sex-specific loci for human diseases and quantitative traits (Table 1). Whilst this fact alone should encourage further investigation, evolutionary theory also predicts the existence of sex-specific genetic architecture for complex traits via sex-specific or sexually antagonistic selection.

In this review we summarise recent evidence for the sex-specific genetic architecture of common diseases and offer guidelines for the identification of sex-specific genetic effects in population-based samples. We also discuss the relationship between sexual antagonism and sexual dimorphism, and propose new mechanisms
through which the genetic architecture of disease might be determined by the existence of two sexes and the different selection pressures that they experience.

Evidence for sex-specific genetic architecture

Broad-sense heritability is the proportion of phenotypic variance in a population sample that can be attributed to genetic variation [15]. Precisely how the genetic variation of complex traits maps to the phenotype is the focus of a large research effort but remains largely unknown. It is clear that the effect of the genotype is often context dependent, whereby factors such as age, environment or sex can have important influences. One clue as to whether a complex trait is influenced by loci with sex dependent effects is the difference in the heritability estimates between males and females (although identical heritabilities in males and females may nonetheless mask underlying differences in sex-specific genetic architecture). For example, in a study of twenty quantitative traits in humans, eleven showed significant sex differences in heritability [16]. Following a PubMed literature search, we identified eighteen independent studies in humans that provided separate heritability estimates for males and females (thirty-one traits), and also stated whether the difference was significant. Of the thirty-one traits, fifteen showed no sex difference in heritability, thirteen had a higher heritability in females, and three had a higher heritability in males (Figure 1). The apparent excess of female-biased heritability estimates, compared to those that are male-biased, requires proper statistical analysis in order to be confirmed. Nevertheless, this observation may be due to the more risky behaviour or more
dangerous working environments that men partake in, which over-ride the genetic risk factors [17].

Non-genetic factors such as behaviour, environmental exposure, anatomical differences, and sex hormones create systemic differences between males and females for trait expression, which in turn affect disease risk and heritability. One example is the protective effect of high oestrogen levels in women on heart disease [18]. Experiments using hormone treatment and gonadectomy show that sex differences in measurements of immune response, behaviour, and toxin resistance are determined by sex chromosome dosage and not by sex hormone levels [19-21]. One possible cause of this may be sex-specific epigenetic modification i.e. regulation of gene expression in one sex only, independent of sex hormone levels. The attenuation of deleterious alleles via sex-specific epigenetic modification is beneficial only if the silencing of that gene can be sufficiently tolerated in that sex. One interesting example of sex-dependent epigenetic modification is a 9% reduction in methylation of the ZPBP2 gene promoter in young males compared to females. The resulting increase in ZPBP2 expression in young males likely explains why common genetic variation in the region increases risk of asthma in this this patient subgroup [22]. In male mice, knock-out of Zpbp2 causes sperm abnormalities and infertility in males, yet has no effect in females. This fact hints that the hypo-methylation of ZPBP2 that increases asthma risk is maintained in the male population because of the demand for proper sperm production [23]. As an extension to sex-dependent regulation by hormones and epigenetic modification, gene co-expression networks also exhibit distinctive sexual dimorphism (although these networks themselves may be a result of sex-dependent hormones and methylation) [7,24,25]. These mechanisms provide a proximal
explanation as to how a genetic variant could have a sex-dependent effect on phenotype.

Initial reports of sex-dependent genetic effects came from linkage mapping and candidate gene studies but have since been surpassed by high-powered, high-coverage GWAS, most of which have been published in the past five years. Testing males and females separately in a GWAS revealed that 15% of SNPs that regulate gene expression in cell lines do so in a sex-dependent manner, even in the absence of sex hormones [26]. For complex traits, GWAS have identified many SNPs with sex-dependent effects on diseases and quantitative traits. These results are summarised in Table 1, which shows thirty-three loci with sex-dependent effects in the twenty-two traits studied. The majority of the SNPs effects were in one sex only (twenty-eight loci) although in five instances, the direction of effect was the same between sexes but differed significantly in magnitude. There are also two well-powered, sex-sensitive GWAS that were negative (for rheumatoid arthritis and for bone mineral density) [27,28]. There is theoretical evidence that existing sexually antagonistic variation promotes the evolution of more sexually antagonistic variation, and is likely to occur in distinct clusters across the genome [29]. Similarly, sex-dependent regulatory variation has been observed in clusters encompassing up to fifty genes [30]. Thus, we have organised the list of SNPs with sex-dependent effects on disease phenotypes in Table 1 by chromosomal position. Although no clustering is visible, the identification of sex-dependent genetic effects in additional phenotypes should provide enough data with which to test the predicted clustering.

Sex-dependent effects of common, genetic variation on quantitative traits have also been documented in non-human organisms [31-36]. Gene manipulation studies in model organisms have identified sexually pleiotropic and sex-reversed effects. For
example, murine vitamin D receptor disruption causes weight loss in males but decreased bone density in females [37], and p53 over-expression in *D. melanogaster* increases male life-span but reduces that of females [38]. There is also good evidence for sex-specific *trans*-eQTLs [30,32], sex-specific residual genetic variance [39], sex-specific epistasis [40], and sex-specific genetic modifiers of age-at-onset [41]. The proximal, biochemical cause of each sex-dependent effect will likely involve sex hormones, sex-specific methylation, interaction with sex chromosomes, or small dimorphisms in the sex determination pathway. It remains to be determined whether the identified sex-dependent genetic effects are the result of on-going or past intralocus sexual conflict, or other evolutionary processes (See Outstanding questions).

**Methods for identifying sex-specific genetic architecture in case-control samples**

A common approach is to test for association in each sex separately (i.e., sex-stratified). If a SNP is significant in one sex but not in the other, authors often conclude that there is a sex-dependent effect. However, a formal test of male versus female association statistics should be made before concluding that the effect is truly sex-dependent. This approach is limited in comparison to joint tests, because of loss in power caused by partitioning of the sample [42]. A joint analysis incorporates a genotype-by-sex interaction term that tests the difference in allele frequencies between male and female cases, given their allele frequencies in controls. It is more suited to identifying genetic differences in trait architecture between males and females rather than for main effects [27]. The regression model with which to test for
genotype-by-sex interactions in an unrelated population sample is: $Y = \beta_0 + \beta_G G + \beta_S S + \beta_{G\times S}(G \times S) + \epsilon$, where $Y$ is the phenotype value, $G$ is the genotype, $S$ is the sex, $\beta$ is the standardised regression coefficient of each variable, and $\epsilon$ is the error [43]. Other covariates, such as environmental variables or those used to correct for population stratification, can also be incorporated into this model. The tests can be performed using open-source software (e.g., PLINK [44] and GenAbel [45]). For family trio data, an interaction analysis is also possible, exemplified by use of a case/pseudo-control test that detected two loci for autism risk [46].

Statistical power should always be calculated for any association study [47]. The behaviour of GxS tests is comparable to a genotype-by-environment test, but specifically one in which the interaction term is binary and equally distributed in the population. Thus, software designed for power calculations in GxE tests is likely to be accurate for GxS tests. Known examples include Gene-Environment iNteraction Simulator (GENS) [48], GxEscan [49] and GWASGxE [50]. For case-control GxE tests, several alternatives have been presented which are potentially applicable to GxS in order to improve power. These include case-only GxE, two-stage, and ‘cocktail’ methods [50,51]. Depending on the method used, 4,000-8,000 cases and the same number of controls confer 80% power to detect a small interaction effect of 1.5, although this is strongly dependent on balanced sex ratios in cases and controls [49,50]. Analytical hazards when using an interaction term include population substructure [52] and incorrect control of covariates [53], such as age, ethnicity, or socio-economic background. Meta-analysis of GWAS data is a routine approach for large heterogeneous sample collections, and a powerful algorithm has been developed in which both sex-specific and main effects can be tested for in a meta-analysis [54,55].
As more sex-specific analyses of GWAS datasets are performed, it would be informative for authors to present sex-specific values for (i) trait heritability, (ii) the phenotypic variance accounted for by significant SNPs, and (iii) genomic prediction/Risk profile score. Finally, given the extent of sexually dimorphic interaction networks [7,24,25], pathway enrichment and epistasis testing should be informative.

Evolutionary processes leading to sex-dependent genetic architecture

Alleles that increase risk of disease, and often reduce fitness in an individual can occur in the human population at high frequency. The reasons for this are not well-understood but may include ancestral neutrality, balancing selection and polygenic mutation-selection balance [56]. A classic example of balancing selection in human disease is sickle-cell anaemia and malaria. The mild form of disease conferred by the heterozygous genotype also protects against malaria, thus maintaining anaemia risk alleles in malaria-endemic regions. So how might the processes maintaining sex-dependent disease risk alleles in a population differ from those which maintain sexually concordant disease variation?

By definition, sex-dependent disease risk alleles are only required to differ in their effect between the sexes – there is no obligation for them to be under differential selection between the sexes. Mutation-selection balance and drift may therefore be sufficient to maintain sex-dependent risk alleles. This could occur in several different ways. Firstly, new mutations might be more deleterious in one sex than in the other.
Indeed, laboratory experiments using *Drosophila melanogaster* indicate that males are more likely to suffer a loss of fitness than females in the presence of novel sex-linked [57] or autosomal mutations [58]. This could either be due to overall reduced genetic robustness in males compared to females, or due to stronger sexual selection on males, effectively making the same phenotype more deleterious [57,59].

Secondly, sex-limited or sex-biased genes might be more likely to accumulate deleterious alleles than sexually monomorphic genes. This is because the efficiency of selection will be reduced if only half of the population expresses the phenotype under selection, causing reduced purging of deleterious alleles. There is recent evidence that this is the case for genes that are expressed exclusively in men [60].

Thirdly, age of onset is likely to be an important factor in determining the influence of drift on sex-dependent risk alleles. Many diseases have an onset age well after reproduction and so should not affect fitness in terms of number of viable offspring produced, making drift a potentially potent force. In addition, men’s potential reproductive lifespan is considerably longer than women’s, which is limited by menopause. This means that late-onset female-specific risk alleles might be expected to experience weaker selection than male-specific risk alleles in humans.

Some models indicate that only diseases that have a low impact on fitness will be caused by alleles that are common in the population, whereas diseases which do affect fitness (i.e. early-onset) are more likely to be caused by rare or unique alleles [61,62]. However, given that there has been some success recently in the identification of high-frequency risk alleles for early-onset diseases such as type II diabetes [63] and schizophrenia [64], these models may not be sufficient to explain all segregating disease variation. Although most of the genome-wide association studies that have tested for sex-dependent effects have targeted quantitative traits, the identification of
sex-dependent risk alleles for young-onset diseases such as Crohn’s Disease and Type 2 diabetes (Table 1) also demonstrates that this type of genetic variation can exist at high frequency in the population despite a likely impact on fitness.

One implication of sexual antagonism is the maintenance of deleterious genetic variation at higher population frequency than would be expected from mutation-selection balance [65,66]. This leads us to consider its role in susceptibility to common, genetically complex disorders: an allele that increases disease risk and fitness in one sex only can be maintained at a frequency and duration greater than that expected by mutation-selection balance or genetic drift, if it is under positive selection in to the other sex. Consistent with this reasoning, mathematical simulation predicts that alleles that are under sex-differential selection (including sexually antagonistic ones) will make up a disproportionately large subset of alleles underlying disease phenotypes [67] (i.e. that among disease-causing alleles, alleles that are subject to sex-specific or sexually antagonistic selection will be overrepresented compared to alleles which experience concordant selection). Below we discuss in greater depth how sexual antagonism and sex-specific selection might contribute to the genetic architecture of complex traits in humans.

Unequal endophenotype outcome

An accepted model of causation for common disease risk alleles is that they do not cause disease directly, rather they affect a quantitative trait that confers increased risk to the disease as its value becomes more extreme [68]. This concept is exemplified by the endophenotype hypothesis of psychiatric disorders [69], but other examples include the relationship between adiponectin level and Type 2 Diabetes [70] as well as between triglyceride level and coronary artery disease [71].
The risk of disease due to an extreme trait value can differ between the sexes even if the genetic architecture of the trait is identical between the sexes. An example of this is for cholesterol levels. High levels of non-LDL cholesterol (>4.9mmol/L) increase risk of myocardial infarction in men more so than in women (hazard ratio 3.09 versus 2.07) [72]. This illustrates the point that although the genetic architecture behind a quantitative trait might be the same (or similar), its impact on morbidity and mortality differs between the sexes, and thus so does natural selection.

**Equal disease risk but with unequal fitness effects**

The effects of disease on fecundity, a major component of fitness, are not always equal between the sexes. One example of this is schizophrenia, which leads to a consistently greater reduction in reproductive success for men than women [73-75]. A second example is for congenital hypothyroidism, associated with loss of fecundity in women but not in men [76]. These examples show that although the genetic architecture of disease may be the same, the fitness effect, and therefore the strength and direction of selection on each sex, differs as a result of the disease.

**Sex-specific migration**

It has been proposed that the genetic variation for a sexually antagonistic trait may vary between populations [66], and thus immigration results in the introduction of novel, sexually antagonistic alleles into the host population. Sex-specific immigration will cause preferential transmission of alleles that are beneficial to that sex (and thus under net positive selection) into a host population, only for the opposite sex to inherit novel deleterious alleles, in addition to those that it already has for that trait.
Although obtaining empirical evidence for these processes may be challenging, there is good evidence for large-scale, sex-specific migrations amongst historical human populations from Central Asia [77,78], the Iberian Peninsula [79], the British Isles [80,81], Central Africa [82], Indonesia [83], and globally [84,85]. Indeed, a recent study of polycystic ovary syndrome suggests that a combination of migration and sexual antagonism might explain observed geographic patterns in risk allele frequencies [86]. Furthermore, these mechanisms could provide a novel explanation for the outbreeding depression observed in some wild animal populations [87].

Sexually antagonistic pleiotropy

We define sexually antagonistic pleiotropy as the deleterious effect of an allele on a fitness-related trait in one sex, combined with a gain in fitness in the other sex through a different trait (Box 1, Figure I, stages B-C). One example of this is comes from a study of evolutionary selection on biometric traits in the Framingham Heart Study [88]. Body height is already known to exhibit sexual antagonism in humans (with short females and tall males being favoured by selection) [13] but the example study additionally identified a negative correlation between selection for body height and cholesterol levels. The authors interpret this as an example of negative pleiotropy in which selection for shorter females maintains the population frequency of high-cholesterol alleles, and thus causes a response in a different male phenotype [88]. Indirect empirical evidence indicates that pleiotropic genes are indeed less able to escape sexual antagonism [89,90], and thus the involvement of pleiotropic genes in disease risk seems likely to be amplified by sex-specific selection.

Resolution of sexual antagonism creates targets for sex-dependent genetic effects
Sexual antagonism can also contribute to sex differences in genetic architecture indirectly because it is resolved through the evolution of sexual dimorphism in the previously shared trait. For example, if a gene is de-activated in one sex, functional genetic variation in that gene can only contribute toward the genetic architecture of a trait in the other sex. See Box 2 for more details about the resolution of sexual antagonism via the evolution of sexual dimorphism.

Concluding remarks

Despite sharing genetic variation, there are profound biological differences between males and females. This can result in different optima for shared traits, sexual antagonism, and sexual dimorphism. Sex-specific selection on an allele can have important effects on its maintenance within a population, allowing deleterious alleles to persist and restrict the fitness of a population [65-67]. The solution is to allow genes to function and evolve independently in each sex, i.e. sexual dimorphism. We expect common, heritable disorders to have sex-dependent genetic architecture because sexual antagonism and sexual dimorphism exist in human populations [13,88] and because disease, in many instances, causes loss of evolutionary fitness to the individual.

One remaining question is how much of the heritability of complex traits is accounted for by sex-dependent genetic effects (See Outstanding questions)? A recent study of sex effects in heritability of 122 complex traits did not find any significant effects [91], although sample sizes varied from 300 to 30,000 in this study, and there were many medically-relevant traits with known sexual dimorphism not tested. A
recent QTL study of 55 complex traits in mice found that only 0.14-4.3% of the phenotypic variation in a quarter of the traits was explained by GxS effects [100]. The authors nevertheless concluded that due to the skewed distribution of effect sizes, some traits have a strong sex effect arising from a few key loci [92]. Given the strong empirical evidence for sex-dependent genetic effects in anthropomorphic traits, serum metabolites, recombination rate (Table 1), sex-dependent genetic modifiers [30,32,39-41], and the many phenotypes yet to be fully investigated for sex effects, researchers should not be discouraged. Analytical approaches have varied, and we hope that researchers will use the most powerful and accurate approaches available [42,49,50,54,55]. High-resolution genotyping, and appropriate analysis of common genetic variation on mitochondrial, X, and Y chromosomes would be hugely beneficial to understanding complex trait genetic architecture, given their widely-known contribution to monogenic disorders, and as likely locations for sexually antagonistic variation [93]. The incentive for investigating sex-dependent effects in a trait is often stated as visible sexual dimorphism but, as outlined above, monomorphic traits may experience the strongest sexually antagonistic selection pressures, and thus also have sex-dependent genetic effects. Although existing evidence indicates that immune genes can be sexually antagonistic [24], it remains to be empirically demonstrated which human, disease-related phenotypes are sexually antagonistic (see Outstanding questions).

Much of what is known about sexual antagonism has been obtained through studies on wild and laboratory animal populations, as well as mathematical modelling. Identification of the molecular genetic basis of fitness and of sexual antagonism in model organisms would not only confirm the empirical observations but also provide grounding for studies of sex-specific genetic architecture in humans. Equally so,
ecological studies in humans could also provide interesting perspectives, for example
how ecological factors influence selection on specific traits to produce varying
degrees of sexually concordant or sex-specific selection across populations [94].

We anticipate that analysis of GWAS data with respect to sex, encouraged by
both evolutionary genetics and recent results presented in this review, will generate
many more significant findings, and reinforce the role that sex-specific and sexually
antagonistic selection may have in contributing to the genetic architecture of complex
traits. Finally, we hope that the identification of sex-specific genetic aetiologies in
what otherwise appears to be the same disease will result in the development of more
effective, sex-specific therapies.

Acknowledgements
The authors thank two anonymous reviewers and Dr Fiona Ingleby for their helpful
comments. This work was supported by the European Research Council (WPG and
EHM; Starting Grant #280632), a Royal Society University Research Fellowship
(EHM), the Swedish Research Council (JKA), and the Volkswagen Foundation
(JKA). The funders had no role in study design, data collection and analysis, decision
to publish, or preparation of the manuscript.

Glossary
Fitness: An evolutionary concept, applicable to individuals, comprised of (i) the
ability to survive, and (ii) the number of offspring produced (fecundity). Ideally
measured as lifetime reproductive success.
**Genetic architecture:** The number, allele frequency in the population and effect size of genetic variants that contribute toward phenotypic variance of a particular trait.

**Genome-wide association study (GWAS):** Method for identifying molecular genetic variation controlling heritable traits in a population sample. Involves assessing the correlation between allele frequencies and phenotype value, at millions of markers of common genetic variation across the genome.

**Intra-locus sexual conflict:** Opposing direction of selection between males and females for a particular locus or single trait, for instance where a sequence variant improves the fitness of one sex but reduces fitness in the other.

**Sexual antagonism:** Opposing direction of selection between males and females for a particular heritable trait which has a positive genetic correlation between the sexes. In contrast to intra-locus sexual conflict, sexual antagonism can involve different traits in each sex, and is therefore a more inclusive term.

**Sexual dimorphism:** A statistical difference between males and females in a population for the value of a particular trait. May include anything from anatomical measurements to expression level of a gene.

**Sex-specific selection:** Difference in magnitude but not direction of selection between the sexes, for example if a trait experiences stronger selection in one sex, or if a trait is sex-limited and therefore only subject to selection in one sex. Compare with sexually antagonistic selection.

**Sexually antagonistic selection:** Difference in direction (and possibly magnitude) of selection between the sexes, for example if a trait experiences positive selection in one sex and negative selection in the other.

**Single nucleotide polymorphism (SNP):** DNA sequence variation occurring in multiple unrelated individuals in a population; stably inherited and caused by
replacement of a nucleotide base with one of the remaining three. Depending on exact location within functional DNA sequence, SNPs can alter biological metrics, and contribute to complex traits and disease susceptibility.

**Box 1: Sexual antagonism and its role in the maintenance of genetic variation**

Sexual antagonism results from sexually discordant (antagonistic) selection acting on a shared genome. Sexual antagonism has now been demonstrated in a wide variety of taxa, including plants, birds, mammals, and insects [11,96]. Anisogamy (difference in gamete size) is considered to be the ultimate source of sex-specific selection [97,98], although ecological factors can also play a role in shaping patterns of sex-specific selection [99]. Sex-specific selection is thought to result in the evolution of sexual dimorphism [100]. However, these divergent phenotypes must be developed from a shared gene pool, making it difficult to simultaneously achieve optimum trait values in both sexes. Thus, for certain traits a conflict will be maintained and the sexes will be displaced from their optimum phenotypes. For example, in fruit flies *Drosophila melanogaster*, when selection on females was removed, they became more masculinized, demonstrating that males had previously been displaced from their phenotypic optimum by counter-selection in females [101]. Pedigree analysis of wild animal populations has also demonstrated a negative intersexual genetic correlation for fitness, i.e., genotypes producing successful males produce unsuccessful females and vice versa [102,103].

More formally, sexual antagonism occurs when genetically correlated traits have opposite effects on male and female fitness. In the simplest case, increasing values of a single trait would increase fitness in one sex and decrease it symmetrically...
in the other sex (Figure I, A). In this case, it is assumed that the trait is positively correlated between the sexes. However more complicated patterns are also possible, such as opposite fitness effects of different correlated traits (Figure I, B-C) or asymmetric patterns of selection (Figure I, D). Consistent with this, a recent study demonstrated that human height was likely to be subject to sexual antagonism: within sibling pairs, men of average height had higher fitness while shorter women had higher fitness [13]. This means that the fitness effect of a given height-determining allele will be context-dependent in terms of sex, and that the population as a whole will be unlikely to evolve towards a shorter phenotype, despite directional selection in females, because of counter-selection in males. Sexual antagonism has also been observed for tolerance to infection in the fruit fly Drosophila melanogaster [104]. One of the major evolutionary implications of sexual antagonism is the maintenance of genetic variation that is deleterious to one sex. Although this has not been fully demonstrated at the molecular level, the population dynamics of a synthetic sexually antagonistic allele in a laboratory D. melanogaster study accurately follows predictions [65,66].

Figure I: The different forms of sexual antagonism. Female fitness functions are shown with red lines, male with blue lines, and the intersexual genetic correlation with black lines. A. The simplest case (also known as intralocus sexual conflict) is where the same trait has opposite and approximately symmetric fitness effects on males and females. The intersexual genetic correlation for the traits is high and positive. B. Sexual antagonism can also occur when different traits have a high positive intersexual genetic correlation, but are selected in opposite directions in males relative to females. In the unselected sex (broken lines), selection for the trait
in question might be weakly positive, neutral, or even absent if the trait is sex-limited.

C. Although no empirical examples of this type have yet been demonstrated, it is also possible that traits with a strong negative intersexual genetic correlation could be subject to sexual antagonism, assuming both traits are selected concordantly across the sexes. A negative intersexual genetic correlation could occur when the same gene product is incorporated in competing alternative pathways. D. It should also be pointed out that selection pressures need not be completely symmetric. Non-linear relationships are also possible.

**Box 2. Sexual dimorphism and resolution of sexual antagonism**

Most research on sexual antagonism to date has focused on sexually dimorphic traits, under the assumption that this dimorphism is an indicator of sex-specific phenotypic optima (Box 1). However, the stage of the most severe sexual antagonism [95,105] is in fact before the trait in question becomes sexually dimorphic (Figure I, stage B), and gene expression data from *D. melanogaster* suggested that most sex-biased genes had already reached their phenotypic optima and were no longer sexually antagonistic [90]. In addition, if sexual antagonism results from correlated expression of different traits across the sexes, monomorphism in a given trait may not be informative about its likelihood of being subject to sexual antagonism [106]. This speaks in favour of casting a broad net when searching for sexually antagonistic loci, and not only investigating traits that are already sexually dimorphic.

Proposed mechanisms for the resolution of sexual antagonism include the evolution of sex-linked modifiers, alternative splicing, or gene duplication [100,107].
Gene duplication is a popular theory as to how genes can escape sexual antagonism, by allowing each copy to evolve independently for each sex [108]. Specifically, this would include genes that are activated by sex hormones or have sex-specific methylation, and are thus expressed at different levels in each sex. Determining which mechanisms of conflict resolution apply or are common is still very much an open question. There is also debate about the time-scale of the resolution of sexual antagonism [107,109-113], but regardless of whether the process is fast or slow in evolutionary time, the outcome is always sex-specific genetic architecture. In this sense, sex-specific genetic architecture in disease is likely to be an indirect result of past sex-specific or sexually antagonistic selection.

**Figure I**: Predicted stages in the resolution of sexual antagonism. A. Initially, the trait is monomorphic and under weak stabilizing selection. B. A change in the physical or social environment causes the previously concordant trait to become subject to opposite patterns of sex-specific directional selection. C. Sexual dimorphism then evolves, causing the sexes to come closer to their respective phenotypic optima, but some antagonism remains. D. The sexes reach their independent optima and the antagonism is completely resolved. Redrawn after information presented in [11].

**Outstanding questions box**

1. How much heritability – broad sense, narrow sense and residual - do sex-dependent loci really account for?
2. Are the identified sex-dependent genetic effects on disease risk sexually antagonistic, sexually dimorphic, or both? How can we show this experimentally?
3. Are some traits or genes more prone to sex dependent genetic effects? Because reproduction and fecundity is key component of fitness, disease with sexually antagonistic genetic risk alleles should have an onset prior to or during reproductive age. Evidence suggests that sexually antagonistic genes are more likely to be pleiotropic, and at least some are likely to be involved in the immune response to infection.

4. Are sexually antagonistic disease alleles distributed non-randomly across the genome? The sex chromosomes have been suggested to be hotspots for sexual antagonism. However recent models also predict that sexual antagonism should increase linkage disequilibrium, which could cause physical clustering of disease alleles [29,114].
Table 1: SNPs with sex-dependent effects on human phenotypes, identified through genome-wide association studies.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>Individuals tested</th>
<th>Gene</th>
<th>Chromosome band</th>
<th>SNP</th>
<th>MAF</th>
<th>Male effect(\dagger)</th>
<th>Female effect(\dagger)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitochondrial DNA levels</td>
<td>384</td>
<td>MRPL37</td>
<td>1p32.3</td>
<td>rs10888838</td>
<td>0.11</td>
<td>0.81</td>
<td>ns</td>
<td>[115]</td>
</tr>
<tr>
<td>Heart beat rate (QT interval)</td>
<td>3761</td>
<td>NOS1AP</td>
<td>1q23.3</td>
<td>rs10494366</td>
<td>0.29</td>
<td>3.08</td>
<td>2.09</td>
<td>[116]</td>
</tr>
<tr>
<td>Waist-height ratio 175585 LYPAL1/SLC30A10</td>
<td>1q41</td>
<td>rs4846567</td>
<td>1q41</td>
<td>rs2820443</td>
<td>0.29</td>
<td>ns</td>
<td>0.06</td>
<td>[117]</td>
</tr>
<tr>
<td>Waist-height ratio 190803 LYPAL1/SLC30A10</td>
<td>2q4.3</td>
<td>rs10195252</td>
<td>2q4.3</td>
<td>rs6717858</td>
<td>0.44</td>
<td>ns</td>
<td>0.05</td>
<td>[118]</td>
</tr>
<tr>
<td>Visceral adiposity</td>
<td>117857</td>
<td>THNSL2</td>
<td>2p11.2</td>
<td>rs1659258</td>
<td>0.05</td>
<td>ns</td>
<td>Z-score 1.5</td>
<td>[119]</td>
</tr>
<tr>
<td>Mitochondrial DNA levels</td>
<td>384</td>
<td>RNFI44</td>
<td>2p25.1</td>
<td>rs2140855</td>
<td>0.39</td>
<td>ns</td>
<td>0.32</td>
<td>[115]</td>
</tr>
<tr>
<td>Waist-height ratio 175585 GRB14/COBBL1</td>
<td>2q2.3</td>
<td>rs10195252</td>
<td>2q2.3</td>
<td>rs1047891</td>
<td>0.30</td>
<td>ns</td>
<td>0.04</td>
<td>[120]</td>
</tr>
<tr>
<td>Waist-height ratio 190803 GRB14/COBBL1</td>
<td>2q2.3</td>
<td>rs6717858</td>
<td>2q2.3</td>
<td>rs6717858</td>
<td>0.30</td>
<td>ns</td>
<td>0.04</td>
<td>[120]</td>
</tr>
<tr>
<td>Plasma homocysteine</td>
<td>1679</td>
<td>CPS1</td>
<td>2q34</td>
<td>rs1047891</td>
<td>0.30</td>
<td>ns</td>
<td>0.04</td>
<td>[120]</td>
</tr>
<tr>
<td>Glycine levels</td>
<td>3343</td>
<td>CPS1</td>
<td>2q34</td>
<td>rs715</td>
<td>0.24</td>
<td>ns</td>
<td>0.23</td>
<td>[6]</td>
</tr>
<tr>
<td>Crohn’s Disease</td>
<td>8463</td>
<td>ATG16L1</td>
<td>2q37.1</td>
<td>rs3792106</td>
<td>0.40</td>
<td>ns</td>
<td>OR 1.48</td>
<td>[121]</td>
</tr>
<tr>
<td>Waist-height ratio 175585 PARG</td>
<td>3p2.5</td>
<td>rs4684854</td>
<td>3p2.5</td>
<td>rs6795735</td>
<td>0.19</td>
<td>ns</td>
<td>0.05</td>
<td>[117]</td>
</tr>
<tr>
<td>Waist-height ratio 175585 ADAMTS9</td>
<td>3p4.1</td>
<td>rs1047891</td>
<td>3p4.1</td>
<td>rs6795735</td>
<td>0.19</td>
<td>ns</td>
<td>0.05</td>
<td>[117]</td>
</tr>
<tr>
<td>Recombination rate</td>
<td>35927</td>
<td>RNF212</td>
<td>4p16.3</td>
<td>rs4054581</td>
<td>0.33</td>
<td>+64cM</td>
<td>ns</td>
<td>[122]</td>
</tr>
<tr>
<td>Recombination rate</td>
<td>35927</td>
<td>RNF212</td>
<td>4p16.3</td>
<td>rs658846</td>
<td>0.02</td>
<td>ns</td>
<td>+95cM</td>
<td>[122]</td>
</tr>
<tr>
<td>Uric acid concentration</td>
<td>28141</td>
<td>SL2C2A9</td>
<td>4p16.1</td>
<td>rs734553</td>
<td>0.26</td>
<td>-0.22</td>
<td>-0.40</td>
<td>[123]</td>
</tr>
<tr>
<td>Sex-hormone binding globulin</td>
<td>2179</td>
<td>UGT2B15</td>
<td>4q13.2</td>
<td>rs293428</td>
<td>0.30</td>
<td>ns</td>
<td>-0.03</td>
<td>[124]</td>
</tr>
<tr>
<td>Uric acid concentration</td>
<td>28141</td>
<td>ABCG2</td>
<td>4q22.1</td>
<td>rs2231142</td>
<td>0.12</td>
<td>0.22</td>
<td>0.13</td>
<td>[123]</td>
</tr>
<tr>
<td>Waist circumference</td>
<td>199499</td>
<td>MAP3K1</td>
<td>5q11.2</td>
<td>rs1743303</td>
<td>0.19</td>
<td>ns</td>
<td>0.03</td>
<td>[118]</td>
</tr>
<tr>
<td>Low-density lipoprotein (LDL)</td>
<td>20512</td>
<td>HMGRCR</td>
<td>5q13.3</td>
<td>rs1265426</td>
<td>0.38</td>
<td>-0.03</td>
<td>ns</td>
<td>[125]</td>
</tr>
<tr>
<td>Thyroid stimulating hormone</td>
<td>26420</td>
<td>PDE8B</td>
<td>5q13.3</td>
<td>rs6885099</td>
<td>0.29</td>
<td>-0.17</td>
<td>-0.12</td>
<td>[126]</td>
</tr>
<tr>
<td>Waist-height ratio 175585 VEGFA</td>
<td>6p21.1</td>
<td>rs6905288</td>
<td>6p21.1</td>
<td>rs753760</td>
<td>0.45</td>
<td>ns</td>
<td>0.05</td>
<td>[117]</td>
</tr>
<tr>
<td>Thyroid stimulating hormone</td>
<td>26420</td>
<td>PDE10A</td>
<td>6q27</td>
<td>rs297325</td>
<td>0.20</td>
<td>1.48</td>
<td>ns</td>
<td>[127]</td>
</tr>
<tr>
<td>Pro-insulin levels</td>
<td>27079</td>
<td>DDX31</td>
<td>9q34.13</td>
<td>rs306549</td>
<td>0.24</td>
<td>0.04</td>
<td>ns</td>
<td>[55]</td>
</tr>
<tr>
<td>Body-mass index, Bone density</td>
<td>4355</td>
<td>SOX6</td>
<td>11p15.1</td>
<td>rs297325</td>
<td>0.20</td>
<td>1.48</td>
<td>ns</td>
<td>[127]</td>
</tr>
<tr>
<td>Triglyceride levels</td>
<td>24273</td>
<td>APOA5/BUD1</td>
<td>11q23.3</td>
<td>rs28927680</td>
<td>0.07</td>
<td>0.13</td>
<td>ns</td>
<td>[125]</td>
</tr>
<tr>
<td>Type II Diabetes</td>
<td>149000</td>
<td>CCND2</td>
<td>12p13.32</td>
<td>rs11063069</td>
<td>0.21</td>
<td>OR 1.08-1.16</td>
<td>ns</td>
<td>[128]</td>
</tr>
<tr>
<td>Recombination rate</td>
<td>35927</td>
<td>CCNB1IP1</td>
<td>14q1.1.2</td>
<td>rs1132644</td>
<td>0.48</td>
<td>+16cM</td>
<td>+57cM</td>
<td>[122]</td>
</tr>
<tr>
<td>Recombination rate</td>
<td>35927</td>
<td>C14orf39</td>
<td>14q2.3</td>
<td>rs1254319</td>
<td>0.30</td>
<td>ns</td>
<td>+72cM</td>
<td>[122]</td>
</tr>
<tr>
<td>Recombination rate</td>
<td>35927</td>
<td>SMEK</td>
<td>14q42.12</td>
<td>rs1013595</td>
<td>0.40</td>
<td>ns</td>
<td>+73cM</td>
<td>[122]</td>
</tr>
<tr>
<td>Type I Diabetes</td>
<td>27530</td>
<td>CTSH</td>
<td>15q25.1</td>
<td>rs3825932</td>
<td>0.30</td>
<td>OR 1.13-1.27</td>
<td>ns</td>
<td>[129]</td>
</tr>
<tr>
<td>Thyroid levels (FT4)</td>
<td>17498</td>
<td>LPCAT2/CAENPS2</td>
<td>16q12.2</td>
<td>rs6499766</td>
<td>0.48</td>
<td>0.02</td>
<td>ns</td>
<td>[126]</td>
</tr>
<tr>
<td>Thyroid stimulating hormone</td>
<td>26420</td>
<td>MAF</td>
<td>16q23.2</td>
<td>rs3813582</td>
<td>0.38</td>
<td>0.12</td>
<td>0.06</td>
<td>[126]</td>
</tr>
<tr>
<td>Recombination rate</td>
<td>35927</td>
<td>17q21.31 region</td>
<td>17q21.31</td>
<td>rs56162163</td>
<td>0.18</td>
<td>ns</td>
<td>+60cM</td>
<td>[122]</td>
</tr>
<tr>
<td>Thyroid levels (FT4)</td>
<td>17146</td>
<td>NETO1/FBXO15</td>
<td>18q22.3</td>
<td>rs724077</td>
<td>0.47</td>
<td>ns</td>
<td>-0.08</td>
<td>[126]</td>
</tr>
<tr>
<td>Type II Diabetes</td>
<td>149000</td>
<td>GIPR</td>
<td>19q13.32</td>
<td>rs8108269</td>
<td>0.31</td>
<td>ns</td>
<td>OR 1.06-1.14</td>
<td>[128]</td>
</tr>
<tr>
<td>High-density lipoprotein (HDL)</td>
<td>11528</td>
<td>PLTP</td>
<td>20q13.12</td>
<td>rs7679</td>
<td>0.18</td>
<td>ns</td>
<td>1.68</td>
<td>[125]</td>
</tr>
</tbody>
</table>

Bold font indicates loci that were confirmed as having sex-dependent effects via an explicit test of male and female association statistics, as opposed to just testing male and female groups separately.  
MAF Minor allele frequency. Value for similar HapMap population sample stated when study sample MAF not available.  
\(\dagger\) Effect value is for the correlation coefficient \(\beta\) unless otherwise stated. OR Odds ratio, 95% confidence intervals. Ns, not significant.  
\(\dagger\) Y Result of separate-sex analyses of SNPs previously identified in a standard, main-effects analysis.  
* GWAMA ‘Genome-wide analysis, meta-analysis’  
SNP rs1047891 previously known as rs7422339.
Reference List


18 Bhupathy,P. et al. (2010) Influence of sex hormones and phytoestrogens on heart disease in men and women. Women's Health (Lond Engl) 6, 77-95


22 Naumova,A.K. et al. (2013) Sex- and age-dependent DNA methylation at the 17q12-q21 locus associated with childhood asthma. Human Genetics 132, 811-822


29 Patten,M.M. et al. (2010) Fitness variation due to sexual antagonism and linkage disequilibrium. Evolution 64, 3638-3642

30 Derome,N. et al. (2008) Pervasive sex-linked effects on transcription regulation as revealed by expression quantitative trait loci mapping in lake whitefish species pairs (Coregonus sp., Salmonidae). Genetics 179, 1903-1917


42 Behrens, G. et al. (2011) To stratify or not to stratify: power considerations for population-based genome-wide association studies of quantitative traits. *Genetic Epidemiology* 35, 867-879


44 Purcell, S. et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559-575


49 Gauderman, W.J. et al. (2013) Finding novel genes by testing G × E interactions in a genome-wide association study. *Genetic Epidemiology* 37, 603-613

50 Murcray, C.E. et al. (2011) Sample size requirements to detect gene-environment interactions in genome-wide association studies. *Genetic Epidemiology* 35, 201-210


53 Keller, M.C. Gene × Environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol. Psychiatry* (in press)


55 Strawbridge, R.J. et al. (2011) Genome-wide association identifies nine common variants associated with fasting proinsulin levels and provides new insights into the pathophysiology of type 2 diabetes. *Diabetes* 60, 2624-2634


72 Madssen, E. et al. (2013) Risk of acute myocardial infarction: dyslipidemia more detrimental for men than women. *Epidemiology* 24, 637-642


74 Power, R.A. et al. (2013) Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. *JAMA Psychiatry* 70, 22-30


118 Randall, J.C. et al. (2013) Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. *PLoS Genetics* 9, e1003500


120 Lange, L.A. et al. (2010) Genome-wide association study of homocysteine levels in Filipinos provides evidence for CPS1 in women and a stronger MTHFR effect in young adults. *Human Molecular Genetics* 19, 2050-2058


130 Keski-Rahkonen, A. et al. (2005) Body dissatisfaction and drive for thinness in young adult twins. *Int. J. Eat. Disord* 37, 188-199


**Figure legends**

**Figure 1:** Comparison of male and female narrow-sense heritability estimates from human studies. Red and blue-coloured data points indicate that a significant difference was identified in heritability between the sexes in that study. Data points are numbered by study, and a letter is added if more than one phenotype was tested in each study. 1a: Drive for thinness. 1b Body Dissatisfaction [130]. 2a: Waist diameter. 2b: Waist-height ratio. 2c: Body-mass index. 2d: Peripheral body fat. 2e: Hip diameter. 2f: Body weight. 2g: Body height [131]. 3a: Triglyceride serum level. 3b: LDL cholesterol serum level [132]. 4a: Lung FEV1 (forced exit volume). 4b: Lung DLCO (diffusing capacity). 4c: Lung VC (vital capacity) [133]. 5: Geriatric depression [134]. 6a: Smoking initiation. 6b: Regular tobacco use [135]. 7: Sleep reactivity (insomnia) [136]. 8: Alcohol dependence [137]. 9: Subjective well-being [138]. 10: Reading disability [139]. 11: Reading difficulties [140]. 12: Self-esteem [141]. 13a: