Systematic review reveals multiple sexually antagonistic polymorphisms affecting human disease and complex traits

Jon Alexander Harper\textsuperscript{1*}, Tim Janicke\textsuperscript{2,3} and Edward H. Morrow\textsuperscript{4}

* Corresponding author

\textsuperscript{1}Evolution, Behaviour and Environment Group, School of Life Sciences, John Maynard Smith Building, University of Sussex, Brighton, BN1 9QG, UK.

\textsuperscript{2}Centre d’Écologie Fonctionnelle et Évolutive, UMR 5175, CNRS, Université de Montpellier, École Pratique des Hautes Études, Montpellier, France.

\textsuperscript{3}Applied Zoology, Technical University Dresden, Zellescher Weg 20b, 01062 Dresden, Germany.

\textsuperscript{4}Department of Environmental and Life Sciences, Karlstad University, 651 88, Karlstad, Sweden.

Jon Alexander Harper ORCID: 0000-0002-1765-9990

Tim Janicke ORCID: 0000-0002-1453-6813

Edward H Morrow ORCID: 0000-0002-1853-7469

Short title: Systematic review of sexually antagonistic genes in human disease

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

An evolutionary model for sex differences in disease risk posits that alleles conferring higher risk in one sex may be protective in the other. These sexually antagonistic (SA) alleles are predicted to be maintained at frequencies higher than expected under purifying selection against unconditionally deleterious alleles, but there are apparently no examples in humans. Discipline-specific terminology, rather than a genuine lack of such alleles, could explain this disparity. We undertook a two-stage review of evidence for SA polymorphisms in humans using search terms from (i) evolutionary biology and (ii) biomedicine. While the first stage returned no eligible studies, the second revealed 51 genes with sex-opposite effects, 22 increased disease risk or severity in one sex but protected the other. Those with net positive effects occurred at higher frequencies. None were referred to as SA. Our review reveals significant communication barriers to fields as a result of discipline-specific terminology.
Introduction

In species with separate sexes an evolutionary conflict at the level of individual genetic loci can occur, where alleles that are favoured by selection in one sex are selected against in the other (i.e intralocus sexual conflict, Parker, 1979). This sexually antagonistic (SA) form of selection is thought to be driven by differences in how the two sexes maximize their fitness and concomitant unequal variances in reproductive success. Research into SA selection has expanded recently with the recognition that it feeds into several important evolutionary processes. It is primarily thought to drive the evolution of sexual dimorphism and trait diversification (Lande, 1980; Pennell et al., 2016; Rice, 1984), including sex-biased gene expression (Ellegren and Parsch, 2007). As a potential driver of balancing selection, it has also been implicated in the maintenance of genetic variation (Connallon and Clark, 2013; Grieshop and Arnvist, 2018) that would otherwise be eroded by directional selection. Most recently, it has been suggested that genetic variation at SA loci could contribute to the occurrence of a number of common human diseases (Morrow, 2015; Morrow and Connallon, 2013), which also show considerable variation between the sexes in terms of their prevalence, severity and age of onset (Ober et al., 2008; Rigby and Kulathinal, 2015).

Despite the interest in sexual antagonism as an evolutionary force, determining the identity of SA loci remains a major challenge (Ruzicka et al., 2020). An early empirical milestone in the field of sexual conflict was achieved by quantitative genetic studies demonstrating that genomes in a laboratory adapted population of *Drosophila melanogaster* harbour significant amounts of SA standing genetic variation (Chippindale et al., 2001; Rice, 1992). A number of other systems have shown similar results, albeit using different methods, including invertebrates and vertebrates from both lab and wild populations (Bonduriansky and Chenoweth, 2009; Foerster et al., 2007; Mills et al., 2011). With the advent of advanced genomic tools, two model systems report specific examples of SA genetic loci: in Atlantic salmon (*Salmo salar*) the VGLL3 locus (Barson et al., 2015), and in *D. melanogaster* the DDT-R locus (Rostant et al., 2015), the Ala-278-Thr polymorphism in the mitochondrial genome (Camus et al., 2015), and multiple candidate loci from a recent genome-wide association study (Ruzicka et al., 2019).

Although humans have also been shown to experience SA selection for some quantitative traits (Garver-Apgar et al., 2011; Stearns et al., 2012; Stulp et al., 2012), and there
are numerous reports of genetic loci with sex-specific effects (i.e. effects that differ in magnitude between the sexes) on multiple complex traits or disease phenotypes (Gilks et al., 2014; Winkler et al., 2015), there are apparently no clear examples of SA loci in humans, which is inconsistent with theoretical expectations (Connallon and Clark, 2014). One potential explanation is that sexual antagonism is a weak selective force in humans, with sexual dimorphism being rather limited (Dixson, 2009; Short, 1979). This may indicate that concordant selection pressures between the sexes have dominated our evolutionary history. But the recent quantitative genetic studies challenge this view (Stearns et al., 2012; Stulp et al., 2012), and indeed theory predicts there is an inevitability to SA loci occurring in organisms with separate sexes (Connallon and Clark, 2014; Parker et al., 1972). Thus, there is a clear disparity between, on the one hand, theoretical expectations and quantitative genetic evidence that SA selection does occur in humans, and on the other hand, a complete lack of specific examples of SA loci. All this in the context of decades of research into how individual genetic variants influence disease profiles.

An alternative explanation for the absence of documented SA loci in humans is that since biomedical science does not use the same terminology for SA effects that evolutionary biology uses, there may in fact be examples that have been misclassified and therefore remain hidden in the literature. For example, genetic loci with sex-specific effects are frequently referred to as sex-dependent effects in the biomedical literature. Other terms are also used for related concepts, including sex-different effects. Furthermore, sex-opposite effects (i.e. a positive effect in one sex and a negative in the other) could be SA effects if the phenotype in question is deleterious, as disease phenotypes often are, since they can influence an individual’s survival or reproductive fitness. The biomedical literature also frequently uses the terms sex and gender interchangeably, with males and females more often than not referred to as men and women, or boys and girls (Khramtsova et al., 2019). There is therefore ample scope for discipline-specific terms to be used for describing the same concept, which may hinder communication between disciplines and obscure examples of SA loci. Moreover, although there has been some progress in introducing concepts from evolutionary biology into the field of medicine (Nesse and Williams, 1994), those inroads have been relatively recent. It is therefore likely that the concepts of (intralocus) sexual conflict and sexual antagonism are not generally well known within biomedical science, and variants that may have SA effects on disease risk may not be referred to as such. A secondary related hypothesis
is that SA effects, when discovered, are discounted as errors as they do not match expectations that sex-differences in phenotype or genetic architecture are not important (Clayton and Collins, 2014). As a result, that may also lead them to be misclassified or simply go unreported, leading to a general publication bias.

For these reasons, we propose that a systematic review of the biomedical literature to identify SA loci in humans requires a specific and targeted set of search terms that would not normally be used within the field of evolutionary biology. We test this assumption by dividing our systematic review into two stages. In the first stage, we search for papers identifying specific genetic loci using terms directly relating to the concept of sexual antagonism and others utilized in evolutionary biology. We supplement this search with a second stage, where we develop a set of terms that we hypothesize may capture the same concept as sexual antagonism within the biomedical literature, although not explicitly stated as such, as well as additional terminology relating to possible alternatives for describing the two sexes. From our searches we also extracted data on sex-specific effect sizes and allele frequencies. We were then able to explore how effect sizes varied for the same alleles between the sexes, and test the prediction that SA alleles experiencing net positive selection between the sexes may achieve higher equilibrium frequencies than alleles that experience more symmetric or net negative patterns of selection (Morrow and Connallon, 2013). Our systematic review aims to advance our understanding of SA genes in humans, which show considerable and largely unexplained diversity in sexual dimorphism for disease phenotypes (Ober et al., 2006; Rigby and Kulathinal, 2015), but we also highlight how and why interdisciplinary research (in this case between biology and medicine) can sometimes fail, impeding scientific advances over significant periods of time.

Methods

For this systematic review we followed PRISMA guidance where possible (Moher et al., 2009). PubMed (https://pubmed.ncbi.nlm.nih.gov/) was searched for articles on 2nd December 2020 with no time limit. The searches were carried out in two Stages, with the organism filter set to human in both stages. In Stage 1, eligible studies were required to report specific genetic variants or haplotypes that were referred to as sexually antagonistic or were an example of intralocus sexual conflict. To achieve this, we conducted a Boolean search for articles that used the terms “sexual antagonism” OR “sexually antagonistic” OR “intralocus sexual conflict”
AND “locus” OR “loci”, “gene” OR “snp” OR “polymorphism” OR “variant” OR “allele” in their abstract or title. The Stage 1 search returned 34 articles in total (full search term in the supplementary material; search output is accessible at [https://pubmed.ncbi.nlm.nih.gov/collections/60255050/?sort=pubdate](https://pubmed.ncbi.nlm.nih.gov/collections/60255050/?sort=pubdate)).

In Stage 2, studies were required to report specific genetic variants or haplotypes in humans with opposite effects in the two sexes on either complex traits, the outcome of a medical intervention, or disease risk/severity. We define complex traits as likely with a polygenic genetic architecture but are not directly related to a disease phenotype. In this second stage search terms were specifically designed to include papers from the biomedical literature that may have been missed in the first stage because they do not report their findings with terms normally found within the evolutionary biology literature. Again, we conducted a Boolean search for articles that used terms in their title or abstract to describe an opposite or different effect in the two sexes (“sex dependent”, “sex different”, “gender-dependent”, “sex AND opposite”, or “gender AND opposite”), or that capture this concept with alternative words for sex (“male AND female AND opposite” OR “men AND women AND opposite” OR “boys AND girls AND opposite”) AND (“locus” OR “loci” OR “gene” OR “snp” OR “polymorphism” OR “variant” OR “allele”). Full details of the search terms used and the numbers of articles returned are provided in the supplementary material. The Stage 2 search returned 881 papers (Figure 1) (full search term in the supplementary material; search output is accessible at [https://pubmed.ncbi.nlm.nih.gov/collections/60254985/?sort=pubdate](https://pubmed.ncbi.nlm.nih.gov/collections/60254985/?sort=pubdate)).

The abstracts of the papers from Stage 1 were then examined by EM and from Stage 2 by JH, and any papers that had the possibility of reporting an opposite effect of a specific genetic locus on a complex trait, medical intervention or on disease risk/severity were considered for further screening. From Stage 1, no articles passed the screening. From Stage 2, this screening produced a shortlist of 70 candidate papers ([https://pubmed.ncbi.nlm.nih.gov/collections/57906298/?sort=pubdate](https://pubmed.ncbi.nlm.nih.gov/collections/57906298/?sort=pubdate)) Both JH and EM then reviewed all the full texts of these candidate papers in detail. Papers were included in the final list if they described a sex-opposite or SA effect linked to a specific genetic locus or loci, and reported the effect to be statistically significant (at a P-value cut-off of <0.05). One additional paper was considered from an outside source.
**Figure 1.** PRISMA flow diagram for systematic review of sexually antagonistic loci in humans.
We converted all reported sex-specific effects into a standard effect size (Cohen’s $d$) quantifying the magnitude of how a given variant affects the studied trait expressed in the given sex. Specifically, Cohen’s $d$ was computed based on the reported descriptive statistics ($N$, mean, standard error) or by conversion from other effect sizes (Odds ratio) and test statistics ($F$-values, $t$-values) using formulas reported elsewhere (Borenstein, 2009; Gurevitch et al., 2013, pp. 195–206; Lajeunesse, 2013). We sought information directly from the authors where these metrics were not possible to extract from the papers themselves (12 authors contacted, 5 responded, 3 responded with data, all later excluded as they did not fulfil the criteria for eligibility). We also recorded the gene name, locus (with accession/rs number where possible), trait affected and the frequency of the focal allele having the effect, hereafter referred to as effect allele frequency. Identification for the variants were taken from the studies where possible, but others necessitated searching the National Centre for Biotechnology Information (NCBI) to find the Reference SNP cluster ID for the variants described. Where effect allele frequencies were not reported, we used genotype frequencies to calculate effect allele frequency. Not all studies reported allele or genotype frequencies and so we attempted to supplement these data with allele frequencies from the 1000 genomes database (Auton et al., 2015), since these show a strong correlation with effect allele frequencies reported in the studies reviewed (Pearson correlation: $N = 25 \ r = 0.94, \ P > 0.001$; Figure S1). However, using this approach we were only able to supplement our allele frequency data for one additional locus (rs7341475 in the RELN gene; Table 1).

If $b_f$ and $b_m$ represent the estimated effect of the focal allele in females and males respectively, then positive values of $b_f$ or $b_m$ imply that the allele is positively associated with disease expression, whereas negative values imply that the allele is negatively associated with disease expression. SA alleles are therefore defined as those with opposite effects between the sexes (i.e. $b_f < 0 < b_m$ or $b_m < 0 < b_f$). An evolutionary model of evolution at SA loci predicts that as the magnitude of the positive effect that an SA allele has outweighs the negative effect, the higher the frequency that allele can achieve (Morrow and Connallon, 2013). We therefore expect a negative relationship between the ratio of positive to negative effect sizes (hereafter referred to as the effect size ratio) and observed allele frequency. To investigate this, we calculated the effect size ratio by dividing the positive standardized effect size by the negative standardized effect size; when $b_f < 0 < b_m$ effect size ratio was calculated as $b_m/ b_f$, and when $b_m < 0 < b_f$ effect size ratio was calculated as $b_f/ b_m$. 
<table>
<thead>
<tr>
<th>Trait</th>
<th>Gene</th>
<th>Locus</th>
<th>EAF</th>
<th>M Effect</th>
<th>M Effect SE</th>
<th>F Effect</th>
<th>F Effect SE</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex traits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concentration performance score</td>
<td>FADS1/2</td>
<td>rs1535</td>
<td>0.672</td>
<td>0.2275</td>
<td>0.1065</td>
<td>-0.2394</td>
<td>0.1116</td>
<td>(Lauritzen et al., 2017)</td>
</tr>
<tr>
<td>Reading speed, number of sentences</td>
<td>FADS1/2</td>
<td>rs174448</td>
<td>0.642</td>
<td>0.5511</td>
<td>0.1069</td>
<td>-0.6074</td>
<td>0.1129</td>
<td>(Lauritzen et al., 2017)</td>
</tr>
<tr>
<td>Reading speed, number of sentences</td>
<td>ELOVL5</td>
<td>rs2397142</td>
<td>0.695</td>
<td>-0.1963</td>
<td>0.1042</td>
<td>0.2364</td>
<td>0.1095</td>
<td>(Lauritzen et al., 2017)</td>
</tr>
<tr>
<td>Reading speed, correct sentences</td>
<td>FADS1/2</td>
<td>rs174448</td>
<td>0.642</td>
<td>0.6599</td>
<td>0.1077</td>
<td>-0.3451</td>
<td>0.1129</td>
<td>(Lauritzen et al., 2017)</td>
</tr>
<tr>
<td>Speed-dating success</td>
<td>OPRM1</td>
<td>rs1799971</td>
<td>0.922</td>
<td>0.4697</td>
<td>0.2275</td>
<td>-0.53</td>
<td>0.2856</td>
<td>(Wu et al., 2016)</td>
</tr>
<tr>
<td>Speed-dating success</td>
<td>HTR2A</td>
<td>rs6311</td>
<td>0.418</td>
<td>-0.4826</td>
<td>0.1982</td>
<td>0.4702</td>
<td>0.1964</td>
<td>(Wu et al., 2016)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>SIM1</td>
<td>rs17185536</td>
<td>0.76</td>
<td>-0.0166</td>
<td>0.0081</td>
<td>0.0264</td>
<td>0.0067</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>NMU</td>
<td>rs3805389</td>
<td>0.28</td>
<td>-0.0139</td>
<td>0.0067</td>
<td>0.0261</td>
<td>0.0066</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>GPNAT1</td>
<td>rs4898764</td>
<td>0.53</td>
<td>0.0208</td>
<td>0.0066</td>
<td>-0.0165</td>
<td>0.0059</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>SGCZ</td>
<td>rs17470444</td>
<td>0.71</td>
<td>0.0345</td>
<td>0.0081</td>
<td>-0.0149</td>
<td>0.0068</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>IQGAP2</td>
<td>rs2069664</td>
<td>0.53</td>
<td>0.0244</td>
<td>0.0078</td>
<td>-0.0183</td>
<td>0.0067</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>SLC2A3</td>
<td>rs741361</td>
<td>0.6</td>
<td>0.0272</td>
<td>0.0079</td>
<td>-0.0196</td>
<td>0.0067</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>IRS1</td>
<td>rs2673140</td>
<td>0.38</td>
<td>0.022</td>
<td>0.0066</td>
<td>-0.0209</td>
<td>0.0059</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>TTN</td>
<td>rs2042995</td>
<td>0.23</td>
<td>0.0192</td>
<td>0.0066</td>
<td>-0.022</td>
<td>0.0078</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>RXRA</td>
<td>rs10881574</td>
<td>0.07</td>
<td>0.0268</td>
<td>0.008</td>
<td>-0.0191</td>
<td>0.0067</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>PTPRD</td>
<td>rs7042428</td>
<td>0.98</td>
<td>0.0244</td>
<td>0.0071</td>
<td>-0.0166</td>
<td>0.0063</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Waist-hip-ratio (BMI adjusted)</td>
<td>CECR2</td>
<td>rs17809093</td>
<td>0.04</td>
<td>0.0294</td>
<td>0.0091</td>
<td>-0.024</td>
<td>0.0077</td>
<td>(Winkler et al., 2015)</td>
</tr>
<tr>
<td>Reaction time (RT) in the Stroop task, congruent condition</td>
<td>CHRNAS</td>
<td>rs3841324</td>
<td>0.406</td>
<td>-1.5421</td>
<td>0.3857</td>
<td>0.9945</td>
<td>0.2517</td>
<td>(Schneider et al., 2015)</td>
</tr>
<tr>
<td>Reaction time (RT) in the Stroop task, incongruent condition</td>
<td>CHRNAS</td>
<td>rs3841324</td>
<td>0.406</td>
<td>-1.6215</td>
<td>0.389</td>
<td>1.0215</td>
<td>0.2521</td>
<td>(Schneider et al., 2015)</td>
</tr>
<tr>
<td>Reaction time (RT) in the Negative priming task</td>
<td>CHRNAS</td>
<td>rs3841324</td>
<td>0.406</td>
<td>-1.675</td>
<td>0.3912</td>
<td>0.6299</td>
<td>0.247</td>
<td>(Schneider et al., 2015)</td>
</tr>
<tr>
<td>Condition</td>
<td>Gene</td>
<td>SNP</td>
<td>Beta</td>
<td>SE</td>
<td>p-value</td>
<td>Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------</td>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-----------</td>
<td>------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borderline personality traits after maltreatment</td>
<td>OXTR</td>
<td>rs53576</td>
<td>0.738</td>
<td>-0.3429</td>
<td>0.0881</td>
<td>0.2066</td>
<td>0.0857</td>
<td>(Cicchetti et al., 2014)</td>
</tr>
<tr>
<td>Borderline personality traits after maltreatment</td>
<td>FKBP5</td>
<td>CATT haplotype</td>
<td>-</td>
<td>0.2041</td>
<td>0.0877</td>
<td>-0.195</td>
<td>0.0857</td>
<td>(Cicchetti et al., 2014)</td>
</tr>
<tr>
<td>Fluid intelligence</td>
<td>C16orf62 (OXTR)</td>
<td>rs226849</td>
<td>0.46</td>
<td>-0.3949</td>
<td>0.1869</td>
<td>0.494</td>
<td>0.1814</td>
<td>(Kiy et al., 2013)</td>
</tr>
<tr>
<td>Face perception</td>
<td>C16orf62 (OXTR)</td>
<td>rs226849</td>
<td>0.46</td>
<td>-0.5819</td>
<td>0.1909</td>
<td>0.3498</td>
<td>0.1788</td>
<td>(Kiy et al., 2013)</td>
</tr>
<tr>
<td>Face memory</td>
<td>C16orf62 (OXTR)</td>
<td>rs226849</td>
<td>0.46</td>
<td>-0.3679</td>
<td>0.1864</td>
<td>0.3562</td>
<td>0.1789</td>
<td>(Kiy et al., 2013)</td>
</tr>
<tr>
<td>Hippocampal volume</td>
<td>SLC6A4</td>
<td>rs25531</td>
<td>0.51</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(Herman et al., 2011)</td>
</tr>
<tr>
<td>Sociopathy (in alcoholics)</td>
<td>SLC6A4</td>
<td>rs25531</td>
<td>0.51</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(Herman et al., 2011)</td>
</tr>
<tr>
<td>Systolic blood pressure S_NIGHT</td>
<td>PON1</td>
<td>rs3917542</td>
<td>-</td>
<td>-0.1047</td>
<td>0.0637</td>
<td>0.1129</td>
<td>0.0645</td>
<td>(Padmanabhan et al., 2010)</td>
</tr>
<tr>
<td>Systolic blood pressure S_NIGHT</td>
<td>PON3</td>
<td>rs17884563</td>
<td>-</td>
<td>-0.1482</td>
<td>0.0637</td>
<td>0.1061</td>
<td>0.0645</td>
<td>(Padmanabhan et al., 2010)</td>
</tr>
<tr>
<td>Systolic blood pressure S_CLIN</td>
<td>EDNRB</td>
<td>rs1335803</td>
<td>-</td>
<td>-0.1308</td>
<td>0.0637</td>
<td>0.1129</td>
<td>0.0645</td>
<td>(Padmanabhan et al., 2010)</td>
</tr>
<tr>
<td>Diastolic blood pressure D_NIGHT</td>
<td>DRD1</td>
<td>rs265974</td>
<td>-</td>
<td>-0.1482</td>
<td>0.0637</td>
<td>0.1502</td>
<td>0.0646</td>
<td>(Padmanabhan et al., 2010)</td>
</tr>
<tr>
<td>Diastolic blood pressure D_HOME</td>
<td>SLC4A2</td>
<td>rs2303933</td>
<td>-</td>
<td>-0.1482</td>
<td>0.0637</td>
<td>0.1325</td>
<td>0.0646</td>
<td>(Padmanabhan et al., 2010)</td>
</tr>
<tr>
<td>Diastolic blood pressure D_CLIN</td>
<td>LPL</td>
<td>rs17410577</td>
<td>-</td>
<td>-0.1047</td>
<td>0.0637</td>
<td>0.1775</td>
<td>0.0647</td>
<td>(Padmanabhan et al., 2010)</td>
</tr>
<tr>
<td>Psychological Stress</td>
<td>BDNF</td>
<td>rs6265</td>
<td>0.149</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(Shalev et al., 2009)</td>
</tr>
<tr>
<td>BMI standard deviation score (children)</td>
<td>FAS</td>
<td>rs2228305</td>
<td>0.0312</td>
<td>-0.552</td>
<td>0.2259</td>
<td>0.4427</td>
<td>0.2075</td>
<td>(Korner et al., 2007)</td>
</tr>
<tr>
<td>Pain sensitivity</td>
<td>OPRM1</td>
<td>rs1799971</td>
<td>0.8817</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>(Fillingim et al., 2005)</td>
</tr>
<tr>
<td>Central DBP</td>
<td>REN</td>
<td>rs6693954</td>
<td>0.22</td>
<td>-0.4444</td>
<td>0.2132</td>
<td>1.1052</td>
<td>0.3923</td>
<td>(Ljungberg et al., 2014)</td>
</tr>
</tbody>
</table>

### Disease risk/severity

<table>
<thead>
<tr>
<th>Condition</th>
<th>Gene</th>
<th>SNP</th>
<th>Beta</th>
<th>SE</th>
<th>p-value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graves ophthalmopathy in GD</td>
<td>PRR3</td>
<td>rs2074503</td>
<td>0.3586</td>
<td>0.4397</td>
<td>0.1682</td>
<td>-0.2185</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>LTA</td>
<td>rs1041981</td>
<td>0.32</td>
<td>-0.0722</td>
<td>0.0368</td>
<td>0.1027</td>
</tr>
<tr>
<td>Arterial blood pressure</td>
<td>PPAR</td>
<td>rs1801282</td>
<td>0.1213</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Schizophrenia risk</td>
<td>RELN</td>
<td>rs7341475</td>
<td>0.851*</td>
<td>1.2396</td>
<td>0.2819</td>
<td>-0.762</td>
</tr>
<tr>
<td>Colon cancer survival</td>
<td>ESR2</td>
<td>(CA)n repeats</td>
<td>-</td>
<td>0.2272</td>
<td>0.0962</td>
<td>-0.2046</td>
</tr>
<tr>
<td>Schizophrenia risk</td>
<td>DAO</td>
<td>rs2070586</td>
<td>0.2729</td>
<td>0.1982</td>
<td>0.0888</td>
<td>-0.2537</td>
</tr>
<tr>
<td>Condition</td>
<td>Gene</td>
<td>SNP</td>
<td>p-value (case-control)</td>
<td>p-value (association)</td>
<td>Minor Allele Effect</td>
<td>Reference</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>--------</td>
<td>---------</td>
<td>------------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>T2DM, fasting glucose</td>
<td>SCARB1</td>
<td>rs9919713</td>
<td>0.04</td>
<td>0.1865</td>
<td>0.0853 -0.2515</td>
<td>0.0697 (McCarthy et al., 2009)</td>
</tr>
<tr>
<td>Depression</td>
<td>SLC6A4</td>
<td>rs25531</td>
<td>0.4586</td>
<td>0.5404</td>
<td>0.2277 -0.3996</td>
<td>0.1859 (Sjoberg et al., 2006)</td>
</tr>
<tr>
<td>Infertility in cystic fibrosis</td>
<td>CFTR</td>
<td>rs213950</td>
<td>0.417931</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bone mineral density</td>
<td>APOE</td>
<td>rs429358</td>
<td>0.0798</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Narcolepsy severity</td>
<td>COMT</td>
<td>Long allele</td>
<td>0.4588</td>
<td>1.3596</td>
<td>0.3878 -2.2152</td>
<td>0.4871 (Dauvilliers et al., 2001)</td>
</tr>
<tr>
<td>Nonsyndromic cleft lip</td>
<td>IMPK</td>
<td>rs11142081</td>
<td>0.863</td>
<td>1.2156</td>
<td>0.0937 -1.1161</td>
<td>0.0783 (Carlson et al., 2018)</td>
</tr>
<tr>
<td>Nonsyndromic cleft lip</td>
<td>IMPK</td>
<td>rs72804706</td>
<td>0.898</td>
<td>-0.816</td>
<td>0.0906 2.0665</td>
<td>0.0908 (Carlson et al., 2018)</td>
</tr>
<tr>
<td>Allergy</td>
<td>ACP1</td>
<td>A* allele (CG haplotype)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Hepatitis C susceptibility</td>
<td>IL-12B</td>
<td>rs3212227</td>
<td>0.3463</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RSV severity</td>
<td>IL-9</td>
<td>rs2069885</td>
<td>0.15</td>
<td>0.1888</td>
<td>0.0933 -0.2794</td>
<td>0.0957 (Schuurhof et al., 2010)</td>
</tr>
<tr>
<td>CV risk marker (plasma fatty acid composition)</td>
<td>FABP2</td>
<td>rs1799883</td>
<td>0.287</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Bone mineral density</td>
<td>NR3C1</td>
<td>GCAG haplotype</td>
<td>-</td>
<td>0.6854</td>
<td>0.3676 -0.9029</td>
<td>0.4311 (Peng et al., 2008)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>NGF</td>
<td>rs6330</td>
<td>0.451</td>
<td>0.4562</td>
<td>0.18 -0.4674</td>
<td>0.1817 (Lang et al., 2008)</td>
</tr>
<tr>
<td>Melanoma risk</td>
<td>RNASEL, miR-146a</td>
<td>AG/AA and C haplotype</td>
<td>-</td>
<td>0.4525</td>
<td>0.1564 -0.3553</td>
<td>0.174 (Sangalli et al., 2017)</td>
</tr>
</tbody>
</table>
Thus, SA alleles with a greater beneficial effect will have a smaller, more negative effect size ratio (<-1), while SA alleles that have a greater deleterious effect will have a larger effect size ratio (> -1), up to a maximum value of 0 (positive values occur when the effect is positive or negative in both sexes, but the allele would then no longer be defined as a SA allele).

We modelled how effect allele frequency changes with effect size ratio using a generalized linear model (GLM), weighted by the inverse of the variance of the effect sizes such that data points with smaller variance have a higher weight, since smaller sample sizes were associated with larger and more variable effect sizes (Figure S2). We initially also included trait class and its interaction with effect size ratio as a fixed factor with two levels (complex trait and disease trait/severity), since alleles that influence complex traits in opposite directions are not necessarily under SA selection and so may not behave in the way predicted, whereas alleles influencing disease traits, unless very late acting, are more likely to show a closer relationship with marginal effects on fitness. As a response variable, effect allele frequency is limited between 0 and 1, so we looked at allele counts to allow allele frequency to vary freely. GLMs with binomial error distribution showed substantial overdispersion, so we used a quasibinomial function to address this issue (Payne et al., 2018).

We used a Chi-squared test to infer significance of the two predictor variables and their interaction. The full model indicated that the interaction term and trait class have no significant effect, so that we report the fit of the reduced model which only includes effect size ratio. We also subsequently modelled the data for the two trait classes (complex traits and disease risk/severity) separately, again using a quasibinomial distribution function, to see if the result was replicated in these smaller subsets of the data. The raw data and R script is available for replicating the analyses and figures we present (Supplementary Material).

**Results**

The Stage 1 search found no articles that described genetic loci in humans with effects that were sexually antagonistic. In contrast, the Stage 2 search identified 32 articles that described variants with statistically significant sex-opposite or SA effects (Figure 1; https://pubmed.ncbi.nlm.nih.gov/collections/60278165/?sort=pubdate). From the studies examined, 51 SA variants were identified (Table 1), affecting 21 different complex traits (30 loci) and 19 disease risk/severity traits (21 loci), with a large range of effect sizes. There were
no variants affecting medical intervention traits that passed the screening process. The majority of alleles had effect sizes of similar absolute values in the two sexes, with larger effects tending to show greater differences on absolute effect size (Figure S3).

We found 22 studies provided sufficient data to allow the relationship between effect size ratio and effect allele frequency to be explored by statistical modelling. Trait class and its interaction with effect size ratio were initially included in the model but neither were found to have a significant effect (GLM: trait class dfs = 1,35, deviance = $15.5 \times 10^3$, $P = 0.965$; interaction dfs = 1,34, deviance = 1476, $P = 0.989$) and were therefore excluded from the subsequent model in which effect size ratio was found to relate negatively to effect allele frequency, as predicted (GLM: estimate $\pm$ SE = -1.10 $\pm$ 0.50, dfs = 1,36, deviance = $38.7 \times 10^6$, $P = 0.024$; Figure 2). For three loci there were data on effect sizes for more than one trait (Table 1), which introduced a degree of non-independence between these values in the dataset. However, resampling the data with effect sizes for only one trait for these three loci in turn and modelling these datasets in the same way (18 individual GLMs, dfs = 1,31) did not change the results qualitatively (minimum, maximum: model estimate $\pm$ SE = -1.0960, -1.0961; deviance = $38.72 \times 10^6, 38.73 \times 10^6$; $P = 0.03618, 0.03620$).

Although the interaction term between trait type and effect size ratio was not significant we wanted to investigate whether the negative relationship between effect size ratio and effect allele frequency across all loci was repeated when the data was divided according to trait class, since while opposite effects on disease risk/severity may show a direct relationship with SA fitness effects, complex traits may or may not be experiencing SA selection, even if effects are in opposite directions in the two sexes. However, although the trends were both again negative as predicted, the generalized linear models of these smaller datasets were only marginally significant (complex traits model estimate $\pm$ SE = -1.10 $\pm$ 0.61, dfs = 1,24, deviance = $38.7 \times 10^6$, $P = 0.065$; disease risk/severity model estimate $\pm$ SE = -1.34 $\pm$ 0.77, dfs = 1,10, deviance = $5.14 \times 10^4$, $P = 0.053$; Figure S4). As for the full dataset, there was some lack of independence for the complex traits subset which contained loci with effect sizes for more than one trait. Again, resampling and modelling these datasets in the same way (18 individual glms, dfs 1, 22) did not change the results quantitatively (minimum, maximum: model estimate = -1.0960, -1.0961; deviance = $38.669 \times 10^6, 38.675 \times 10^6$; $P = 0.07774, 0.07776$).
Figure 2. The relationship between effect allele frequency and effect size ratio. Point size varies according to the variance of effect size ratio - larger points have smaller variance and therefore a larger weighting in the model. The vertical dotted line represents the switch point between a net-negative effect of a particular locus (effect size ratio >-1) and a net positive effect (effect size ratio <-1). The blue line represents the predicted values derived from the generalized linear model fitted to the data (see Results).
Discussion

The Stage 1 literature search of our systematic review focussed on finding reports of specific genetic loci in humans with evidence for SA effects using terminology normally associated with the evolutionary concepts of sexual conflict or sexual antagonism. The Stage 2 search sought equivalent evidence, but used search terms that we anticipated would be used by scientists outside the field of evolutionary biology, who may not use the same terminology. Although the Stage 1 search did not find any examples of SA loci occurring in humans, the Stage 2 search identified 51 genetic loci across 39 studies that had ostensibly SA effects, but were not described as such. Clearly then the Stage 1 review failed to identify multiple relevant reports going back some 20 years because the terminology used in those reports did not match the conceptual framework of the search. The Stage 2 search may also be a lower limit given searches of other larger databases may also harbour further examples (see below). Although not yet validated, these reports nonetheless represent a substantial body of evidence that humans, like other organisms with separate sexes (Bonduriansky and Chenoweth, 2009), inevitably experience SA selection for a wide range of complex traits (Connallon and Clark, 2014; Garver-Apgar et al., 2011; Stearns et al., 2012; Stulp et al., 2012), as well as for a range of diseases.

A key prediction from a population genetic model of SA genetic variation is that SA alleles are expected to achieve higher equilibrium frequencies, without necessarily going to fixation, as the relative magnitude of the positive effect in one sex outweighs the negative effect in the opposite sex i.e. an increasingly negative effect size ratio (Morrow and Connallon, 2013). We found clear support for this prediction, with a negative relationship between effect size ratio and allele frequency when examining all complex traits and diseases together (Fig 2), although this relationship was only marginally significant when complex or disease trait classes were modelled separately (Fig S3). These results lend support to the view that the loci identified are genuinely experiencing SA selection, and although it is difficult to discern how selection acts on trait size in males and females for complex traits we would not expect the frequencies of disease-causing alleles to be so high under a mutation-selection balance. We were not able to look at the evolutionary history or age of these alleles due to limited information (n = 8) (Albers and McVean, 2018), it would nonetheless be valuable to examine the population dynamics of these alleles over a broad timescale.
The loci showed a very broad range of effect sizes, from small to very large, with a generally symmetrical inverse relationship between effect size in one sex and the other (Figure S2), which is generally expected, since our screening process necessarily excluded studies reporting effects in the same direction across the two sexes. Nonetheless, it is striking just how large some of the effects were, with very large negative effects in one sex simultaneously occurring with similarly large positive effects. Larger and more variable effect sizes are expected when sample sizes are small, a pattern we also found, which motivates further investigation of the traits and loci included in this review to elucidate more accurate estimates of real effect sizes.

The loci themselves fall within genes that influence a broad range of phenotypes, including morphological, physiological and behavioural complex traits, and a similarly diverse range of disease types, including various cancers, neurological disorders and immune system processes (see Table 1). For the most part, the diseases appear to be early-acting rather than late-onset, with the exception of perhaps bone mineral density. As such, it seems reasonable that an allele that increases disease risk or severity in these cases will indeed have a concomitant reduction in marginal fitness. Although a single large study identified several loci related to BMI adjusted waist-to-hip ratio, there is generally no overall bias towards one particular trait or disease class. The total number of loci is relatively small however, which may limit power to identify such biases if they exist. We also did not find any examples of SA loci on the X-chromosome, which may or may not be a hotspot of sexual antagonism (Fry, 2010; Rice, 1984; Ruzicka and Connallon, 2020). This was expected given that it is commonly not included in genome-wide analyses (Wise et al., 2013), although most of the papers we reviewed were candidate gene studies. Nonetheless, we anticipate further examples may be identified should data from the X-chromosome be included systematically in association studies.

We also found evidence that some traits are influenced by more than one genetic variant in sex-opposite or SA ways (n=8), or that some specific variants have pleiotropic effects on more than one trait or disease in a sex-opposite or SA way (n=6; see Table 1). This complexity in the genetic architecture of traits or diseases may make conflict resolution particularly difficult, as a change in the allele at a single genetic locus experiencing SA selection may simultaneously influence multiple genes or phenotypes in both sexes in divergent ways (Fitzpatrick, 2004; Pennell and Morrow, 2013). Consequently, this may explain
the persistence of intralocus sexual conflict at these loci. It may be that some of the remaining loci also have as yet unidentified pleiotropic effects with other traits or there are as yet unidentified SA loci influencing those same traits or diseases. Of the 51 variants identified, 4 were confirmed to be haplotypes. Such variants, consisting of two or more SNPs in linkage disequilibrium having a joint effect, could also present a problem for conflict resolution, since the ability of selection to act on any individual locus independent of the others in the linkage block is reduced for linked loci. This issue may also extend to other single variants reported here if they also occur in linkage blocks.

Although we have identified multiple genetic loci with either sex-opposite or SA effects these are likely outnumbered by those with either sex-specific (same direction but different magnitude) or sex-limited effects. For instance, Winkler et al. (2015) report 44 loci with sex-specific or sex-limited effects but only 11 with sex-opposite effects. Several other genome-wide association studies report sex-specific effects in several human diseases (Khramtsova et al., 2019), and a recent (non-systematic) review identified 37 SNPs with sex-dependent effects (Gilks et al., 2014). Detection of loci with SA versus sex-specific or sex-limited effects may differ systematically since we expect SA loci are more likely to persist at intermediate frequencies than loci with sex-specific effects (i.e. that differ between the sexes in magnitude but not sign). A larger systematic review that targeted variants with sex-specific or sex-dependent effects would therefore be a valuable contribution and enable us to more clearly understand how important SA alleles are relative to the broader context of genes with non-identical effects in the two sexes.

A key gap in our knowledge is whether the putative examples of sex-opposite or SA alleles presented here can be validated using independently derived datasets in the same or different sub-populations. For this, focussing on those variants associated with sex-differential disease risk would naturally be the most fruitful for advancing our knowledge of sexual antagonism in human disease. We therefore encourage specialists for the particular disease groups presented in Table 1 to include in the future the potential for SA genetic effects to occur when designing studies and analysing data. It is also important for all studies to report results in sufficient detail so that effect sizes can be calculated (if not given) as well as allele frequency data. Unfortunately, 17 studies failed to do this, so could not be included in our review and subsequent test of the evolutionary model. More generally, our study reinforces the view that sex is an important factor in shaping genetic associations with human
traits and disease, even in divergent and contradictory ways, and so should always be considered when investigators examine genetic associations with phenotypes (Lee, 2018; Ober et al., 2008).

There is a possibility that some bias exists in the PubMed database against papers from the field of evolutionary biology, and if we had used another database then both stages would have returned valid papers. We subsequently (14th Dec 2020) sought to investigate this possibility with additional searches of Scopus and Web of Science. Repeating the Stage 1 search initially returned 85 papers in Web of Science (Basic search of Topic, refined by “Human” within results) and 71 papers in Scopus (include keywords “Human”, exclude “Non-Human”), but again no papers made it through the screening process. Repeating the Stage 2 search returned a much larger number of studies in both Web of Science (4,573; Basic search of Topic, refined by “Human” within results) and Scopus (6,616; Advanced search of Title-Abstract, filter: include human and humans, exclude nonhumans, Articles only), with 8,604 unique records. This large number of records is too many to currently screen and may harbour further examples of SA loci in humans, but it is noteworthy that the majority of the initial list of studies found during Stage 2 using PubMed were also found in the other two databases (712 out of 881) as well as 28 out of the screened list of 32 studies identified in the current systematic review.

The disparity between the two stages of the review process provides a remarkable example of how discipline-specific terminology and concepts can hinder scientific communication between fields for substantial periods of time. For many of the studies identified in this review, the focus of the research was not on quantifying or even identifying sex-specific genetic effects in traits or disease, although they were tested for. The responses to such findings varied. Some framed these results as a major finding, others merely made note of them, with one suggesting that since the genetic associations were in opposite directions in the two sexes then it should be regarded as a false positive (Wong et al., 2005).

It is possible, or even likely then that there exists a publication bias in the biomedical sciences against studies with apparently incongruous sex-opposite or SA effects, perhaps in part because the evolutionary framework that may provide a context for understanding those results is not generally appreciated or known. Of course, for evolutionary biologists interested in the role of SA selection on human traits and disease, this is clearly a problem. For many years no specific examples of SA loci were known in humans, and yet the first putative
examples in fact appeared in the literature two decades ago. Our systematic review therefore provides a concrete illustration of how advances to both fields have likely been hampered by a limited understanding of the theoretical framework and terminology used by the neighbouring field.

How can communication between the fields of medicine and evolutionary biology be facilitated? We do not have a comprehensive answer but clearly training in, or an increased awareness of, evolutionary medicine for scientists in both camps would probably be beneficial. The topic of evolutionary biology now forms a key part of the education for some biomedical scientists and clinicians. Similarly, evolutionary medicine is now frequently included in fundamental textbooks on evolution, and is being taught to students of evolutionary biology as part of their degree courses. These changes have been relatively recent, are far from being universally employed, but they should be encouraged. Funders of primary research are also now actively encouraging or obliging researchers to include sex as a factor when designing their studies (Clayton and Collins, 2014), although this may simply result in studies that aim to control for rather than investigate these effects explicitly. Consequently even this remedy may still preclude a deeper understanding of the origins of some human diseases.

Acknowledgements

We thank Jessica Abbott, Tim Connallon and Filip Ruzicka for comments on a previous version of the MS. Funding was provided by the Swedish Research Council (Grant number: 2019-03567) and by the Royal Society to EHM as a University Research Fellowship and Enhancement Award. TJ was funded by the Centre national de la recherche scientifique (CNRS) and the German Research Foundation (DFG grant number: JA 2653/2-1).

Competing Interest Statement

The authors have declared no competing interest.
References


Auton, A., Abecasis, G.R., Altshuler, D.M., Durbin, R.M., Abecasis, G.R., Bentley, D.R.,

Chakravarti, A., Clark, A.G., Donnelly, P., Eichler, E.E., Flicek, P., Gabriel, S.B., Gibbs,

R.A., Green, E.D., Hurles, M.E., Knoppers, B.M., Korbel, J.O., Lander, E.S., Lee, C.,


Sherry, S.T., Wang, J., Wilson, R.K., Gibbs, R.A., Boerwinkle, E., Dondapati, H., Han,


Feng, Q., Fang, X., Guo, X., Jia, M., Jian, H., Jin, X., Lan, T., Li, G., Li, J., Li, Y., Liu, S.,


Zhang, D., Zhang, W., Zhao, J., Zheng, X., Lander, E.S., Altshuler, D.M.,

Gabriel, S.B., Gupta, N., Gharani, N., Toji, L.H., Gerry, N.P., Resch, A.M., Flicek, P.,

Barker, J., Clarke, L., Gil, L., Hunt, S.E., Kelman, G., Kulesha, E., Leinonen, R., Mardis,

W.M., Radhakrishnan, R., Roa, A., Smirnov, D., Smith, R.E., Streeter, I., Thornham, A.,

Toneva, I., Vaughan, B., Zheng-Bradley, X., Bentley, D.R., Grocock, R., Humphray, S.,

James, T., Kingsbury, Z., Lehrach, H., Sudbrak, R., Albrecht, M.W., Amstislavskiy, V.S.,

Borodina, T.A., Lienhart, M., Mertes, F., Sultan, M., Timmermann, B., Yoo, M.-L.,


Beloslyudtsev, D., Bour, N., Chen, C., Church, D., Cohen, R., Cook, C., Garner, J.,

Hefferon, T., Kimelman, M., Liu, C., Lopez, J., Merc, P., O’Sullivan, C., Ostapchuk, Y.,

Phan, L., Ponomarev, S., Schneider, V., Shekhtman, E., Sirotkin, K., Slotta, D., Zhang,

H., McVean, G.A., Durbin, R.M., Balasubramaniam, S., Burton, J., Danecek, P., Keane,

T.M., Kolb-Kokocinski, A., McCarthy, S., Stalker, J., Quail, M., Schmidt, J.P., Davies,

C.J., Gollub, J., Webster, T., Wong, B., Zhan, Y., Auton, A., Campbell, C.L., Kong, Y.,

Marcketta, A., Gibbs, R.A., Yu, F., Antunes, L., Bainbridge, M., Muzny, D., Sabo, A.,

Huang, Z., Wang, J., Coin, L.J.M., Fang, L., Guo, X., Jin, X., Li, G., Li, Q., Li, Y., Li, Z., Lin,


C., Dal, E., Kahveci, F., Marth, G.T., Garrison, E.P., Kural, D., Lee, W.-P., Fung Leong,

W., Stromberg, M., Ward, A.N., Wu, J., Zhang, M., Daly, M.J., DePristo, M.A.,

Handsaker, R.E., Altshuler, D.M., Banks, E., Bhatia, G., Bish, A., Daniel, G., Gabriel, S.B.,

Genovese, G., Gupta, N., Li, H., Kashin, S., Lander, E.S., McCarron, S.A., Nemes, J.C.,

Poplin, R.E., Yoon, S.C., Lihm, J., Makarov, V., Clark, A.G., Gottipati, S., Keinan, A.,

Rodriguez-Flores, J.L., Korbel, J.O., Rausch, T., Fritz, M.H., Stütz, A.M., Flicek, P., Beal,

K., Clarke, L., Datta, A., Herrero, J., McLaren, W.M., Ritchie, G.R.S., Smith, R.E.,

Zerbin, D., Zheng-Bradley, X., Sabati, P.C., Shylakhter, I., Schaffner, S.F., Vitti, J.,

Cooper, D.N., Ball, E.V., Stenson, P.D., Bentley, D.R., Barnes, B., Bauer, M., Keira

Cheetham, R., Cox, A., Eberle, M., Humphray, S., Kahn, S., Murray, L., Peden, J.,


M., Sudbrak, R., Amstislavskiy, V.S., Herwig, R., Mardis, E.R., Ding, L., Koboldt, D.C.,

Larson, D., Ye, K., Gravel, S., The 1000 Genomes Project Consortium, Corresponding authors, Steering committee, Production group, Baylor College of Medicine, BGI-Shenzhen, Broad Institute of MIT and Harvard, Coriell Institute for Medical Research, European Molecular Biology Laboratory, EBI, Illumina, Max Planck Institute for Molecular Genetics, McDonnell Genome Institute at Washington University, US

National Institutes of Health, University of Oxford, Wellcome Trust Sanger Institute,
Analysis group, Affymetrix, Albert Einstein College of Medicine, Bilkent University, Boston College, Cold Spring Harbor Laboratory, Cornell University, European Molecular Biology Laboratory, Harvard University, Human Gene Mutation Database, Icahn School of Medicine at Mount Sinai, Louisiana State University, Massachusetts General Hospital, McGill University, National Eye Institute, N., 2015. A global reference for human genetic variation. Nature 526, 68–74. https://doi.org/10.1038/nature15393


Both stages of the PubMed literature search employed a Species filter set to Humans only.

**Stage 1 Search term:**


**Stage 2 Search term:**

841 different[Title/Abstract] AND loci[Title/Abstract]) OR (sex different[Title/Abstract] AND
842 gene[Title/Abstract]) OR (sex different[Title/Abstract] AND snp[Title/Abstract]) OR (sex
843 different[Title/Abstract] AND polymorphism[Title/Abstract]) OR (sex
different[Title/Abstract] AND variant[Title/Abstract]) OR (gender-dependent[Title/Abstract]
844 AND locus[Title/Abstract]) OR (gender-dependent[Title/Abstract] AND loci[Title/Abstract])
845 OR (gender-dependent[Title/Abstract] AND gene[Title/Abstract]) OR (gender-
dependent[Title/Abstract] AND snp[Title/Abstract]) OR (gender-dependent[Title/Abstract]
846 AND polymorphism[Title/Abstract]) OR (gender-dependent[Title/Abstract] AND
847 variant[Title/Abstract]) OR (male[Title/Abstract] AND female[Title/Abstract] AND
848 opposite[Title/Abstract] AND locus[Title/Abstract]) OR (male[Title/Abstract] AND
849 female[Title/Abstract] AND opposite[Title/Abstract] AND loci[Title/Abstract]) OR
850 (male[Title/Abstract] AND female[Title/Abstract] AND opposite[Title/Abstract] AND
851 gene[Title/Abstract]) OR (male[Title/Abstract] AND female[Title/Abstract] AND
852 opposite[Title/Abstract] AND snp[Title/Abstract]) OR (male[Title/Abstract] AND
853 female[Title/Abstract] AND opposite[Title/Abstract] AND polymorphism[Title/Abstract])
854 OR (male[Title/Abstract] AND female[Title/Abstract] AND
855 opposite[Title/Abstract] AND variant[Title/Abstract]) OR (men[Title/Abstract] AND
856 women[Title/Abstract] AND opposite[Title/Abstract] AND locus[Title/Abstract]) OR
857 (men[Title/Abstract] AND women[Title/Abstract] AND opposite[Title/Abstract] AND
858 gene[Title/Abstract]) OR (men[Title/Abstract] AND women[Title/Abstract] AND
859 opposite[Title/Abstract] AND snp[Title/Abstract]) OR (men[Title/Abstract] AND
860 women[Title/Abstract] AND opposite[Title/Abstract] AND polymorphism[Title/Abstract])
861 OR (men[Title/Abstract] AND women[Title/Abstract] AND
862 opposite[Title/Abstract] AND variant[Title/Abstract]) OR (boys[Title/Abstract] AND
863 girls[Title/Abstract] AND opposite[Title/Abstract] AND locus[Title/Abstract]) OR
864 (boys[Title/Abstract] AND girls[Title/Abstract] AND opposite[Title/Abstract] AND
865 gene[Title/Abstract]) OR (boys[Title/Abstract] AND girls[Title/Abstract] AND
866 opposite[Title/Abstract] AND snp[Title/Abstract]) OR (boys[Title/Abstract] AND
867 girls[Title/Abstract] AND opposite[Title/Abstract] AND polymorphism[Title/Abstract])
868 OR (boys[Title/Abstract] AND girls[Title/Abstract] AND
869 opposite[Title/Abstract] AND variant[Title/Abstract]) OR (sex dependent[Title/Abstract] AND
870 allele[Title/Abstract]) OR (sex different[Title/Abstract] AND allele[Title/Abstract]) OR (gender-
dependent[Title/Abstract] AND allele[Title/Abstract]) OR (male[Title/Abstract] AND
871 female[Title/Abstract] AND opposite[Title/Abstract] AND allele[Title/Abstract]) OR
872 (men[Title/Abstract] AND women[Title/Abstract] AND opposite[Title/Abstract] AND
873 allele[Title/Abstract]) OR (boys[Title/Abstract] AND girls[Title/Abstract] AND
874 opposite[Title/Abstract] AND allele[Title/Abstract])
Figure S1. Study reported effect allele frequencies and database allele frequencies are strongly correlated. The correlation between reported effect allele frequency and frequencies for the same alleles obtained from the 1000 genomes database (0.94). Filled circles represent disease risk/severity related variants while open circles represent complex trait variants.
Figure S2. Relationship between study sample size and average extracted effect size across all studies.
Figure S3. Female effect size and male effect size are negatively correlated in SA alleles.

Effect size (Cohen’s D) of variants in females against effect size in males. Triangular points represent sex-opposite complex trait variants, circular points represent disease risk/severity variants.
Figure S4. The relationship between effect allele frequency and effect size ratio, grouped by trait class. Points to the right of the vertical dotted line have a greater negative effect than positive. **A.** Disease risk/severity variants. **B.** Complex traits. Point size is based on the variance of the effect size ratio, with smaller variance having larger point sizes and greater weighting in the statistical model.