Chromosome Y genetic variants: impact in animal models and on human disease

J. W. Prokop¹ and C. F. Deschepper²

¹HudsonAlpha Institute for Biotechnology, Huntsville, Alabama; and ²Institut de recherches cliniques de Montréal (IRCM) and Université de Montréal, Montreal, Quebec, Canada

Prokop JW, Deschepper CF. Chromosome Y genetic variants: impact in animal models and on human disease. Physiol Genomics 47: 525–537, 2015. First published August 18, 2015; doi:10.1152/physiolgenomics.00074.2015.—Chromosome Y (chrY) variation has been associated with many complex diseases ranging from cancer to cardiovascular disorders. Functional roles of chrY genes outside of testes are suggested by the fact that they are broadly expressed in many other tissues and correspond to regulators of basic cellular functions (such as transcription, translation, and protein stability). However, the unique genetic properties of chrY (including the lack of meiotic crossover and the presence of numerous highly repetitive sequences) have made the identification of causal variants very difficult. Despite the prior lack of reliable sequences and/or data on genetic polymorphisms, earlier studies with animal chrY consomic strains have made it possible to narrow down the phenotypic contributions of chrY. Some of the evidence so far indicates that chrY gene variants associate with regulatory changes in the expression of other autosomal genes, in part via epigenetic effects. In humans, a limited number of studies have shown associations between chrY haplotypes and disease traits. However, recent sequencing efforts have made it possible to greatly increase the identification of genetic variants on chrY, which promises that future association of chrY with disease traits will be further refined. Continuing studies (both in humans and in animal models) will be critical to help explain the many sex-biased disease states in human that are contributed to not only by the classical sex steroid hormones, but also by chrY genetics.

chromosome Y; consomic strains; phenotypic traits; rat; mouse; human disease; sex differences

IN MAMMALS, SEX IS DETERMINED by the complementation of sex chromosomes, such that males are XY and females are XX. The mammalian chromosomes X and Y (chrX and chrY) diverged from an ancestral autosomal chromosome pair around 300 million years ago through four sequential events (60). The first one is believed to have been the introduction of the sex-determining region of chrY (Sry) gene as the locus driving differentiation of gonads into testes, followed by chrY losing the ability for most of the chromosome to recombine with chrX. Ultimately, this led to the separation of chrY into two distinct domains: 1) the pseudoautosomal region (PAR; 5% in human) and 2) the nonpseudoautosomal region (95% in human), also known as the male-specific region on chrY (MSY). ChrYPAR allows for mitotic pairing and constitutes the only part of chrY that is able to recombine with the chrX (115). Sry is found on the MSY, along with several additional protein coding genes that have homologous counterparts on chrX (115).

The MSY has features that set it apart from all other chromosomal regions, particularly because of the lack of recombination. First, there is an elevated rate of mutation in male gametes (45), but since the resulting de novo variants [either single nucleotide variants (SNVs) or insertion-deletions (inserts)] cannot be removed, they become fixed in the transmission from father to son. Likewise, variants that increase fitness cannot be separated from variants that decrease fitness, such that positive selection of beneficial traits can result in the maintenance of associated negative traits. Second, most ancestral genes from the chromosome have been lost on MSY in the course of vertebrate evolution, with the only few genes that have been retained showing specialization in either sex determination, male fertility, and X/Y dosage compensation (10, 27, 58). Finally, repetitive sequences (such as transposable elements) have accumulated in many species, resulting in large heterochromatic regions.

Given the high rate of degeneration of MSY across evolution, some had made the prediction that the human MSY might become extinct in around 10 million years (1). In addition to the fact that this perspective might have contributed to a relative lack of interest, many experimental hurdles have made it difficult to understand whether MSY has broader implications to biologic functions and/or disease state. Indeed, without recombination, variants within large population association studies cannot be segregated by classical linkage disequilibrium (LD) genetic methods. Therefore, causal variants cannot be narrowed down to smaller regions either by defining LD blocks of variants or by performing congenic mapping in
animals. Second, knockouts and other targeted mutations are difficult to generate, because of the presence of large stretches of repetitive sequence and the fact that many MSY genes are required for reproductive success. Third, since large fractions of MSY comprise large repeat units (amplicons), often organized as inverted repeats (palindromes), MSY sequencing is particularly challenging, especially with short-read next-generation technologies (8). In fact, when certain species’ genomes are sequenced, MSY is either ignored altogether, or the corresponding sequence exists as hundreds of independent scaffolds with little to no confidence on chromosomal organization. Unlike the majority of genomic regions that can be reliably sequenced with short reads (60–100 base pairs) of next-generation sequencing, the sequencing of the MSY is still dependent on the generation of bacterial artificial chromosomes and/or the use of long read sequencing technology (such as PacBio), thus increasing the difficulty and the overall cost of sequencing. Finally, in addition to the fact that there are considerable interspecies differences in MSY sequences (due to the elevated rate of variation, the process of gene conversion, and the absence of segregation of traits), genes on the MSY vary considerably between species in terms of which particular genes have been retained from the ancestral chromosome and in the expression profiles of retained or gained genes in several tissues. Thus, experimental evidence obtained in one vertebrate species may not always be readily applicable to another one.

Nonetheless, data obtained in recent years have shown that MSY genetic variants may contribute to phenotypic diversity in traits as varied as cardiovascular functions, immune cell properties, or cancer susceptibility, as discussed below. Likewise, investigation in new animal models have, in combination with recent sequencing data of MSY in several mammalian species, provided clues concerning possible mechanisms whereby MSY could contribute to such diverse phenotypes. In this review, we discuss how the growing body of research is contributing to changing our view of MSY contributions to complex phenotypic traits.

The Contributions of Genetic MSY Variants to Phenotypic Traits May Segregate from Those of Sex Hormones

One of the prevailing views of mammalian sex determination is that male and female somatic cells are initially not different, with sexual dimorphisms being imposed by the type of gonads initially developed (137). In this view, the Sry gene in mammals initiates testicular development and suppresses ovarian differentiation (87, 127). Once the gonads in respective sexes are formed, hormones secreted by the gonads (estrogens by the ovary or testosterone by the testes) result in further sex differentiation (112, 132). However, there is evidence that sex chromosome complementation (i.e., XX vs. XY) plays roles that are independent from those of gonadal steroids. For instance, several reports show that in placental mammals, early XY embryos are more advanced than XX counterparts at gestational times that occur before the time of gonadal differentiation (16). Likewise, almost one-third of transcripts in bovine blastocyst (including MSY-encoded genes other than Sry in male embryos) show sexual dimorphism in their expression levels (11). However, expression of MSY genes may not be the only cause responsible for these differences, because inactivation of chrX-linked genes is incomplete in blastocysts, so that expression of most corresponding transcripts is upregulated in female vs. male blastocysts (11).

Outside of initial gonad development, most phenotypic traits with differences between males and females are commonly thought to result from differences in circulating estrogen and testosterone between the sexes. For example, the primary contributor to the increased risk of cardiovascular disease in males is thought to be the lack of estrogen’s protective effectiveness in males, thus explaining why risk increases in females after menopause (69). Many reviews have focused on the hormonal contributions to sex differences in phenotypic traits for human and animal models (67, 80, 95, 96, 129). However, many MSY genes are expressed outside of gonads. In humans, the contributions of MSY genes and contributions of protein products are illustrated by the data provided by the Human Protein Atlas (126) (Fig. 1). Both in higher primates and rats (but not mice), genes such as Sry show a similar pattern of expression across tissues (124), including a high level of expression in kidneys (3). Expression of MSY genes in tissues other than male gonads thus represent potential for genetic contributions of MSY to sex-dependent differences and highlight the need to elucidate the functions of these genes outside of sex determination and the physiology of male gonads.

Is there evidence that the contributions of sex hormones to phenotypic traits segregate from those of sex chromosomes? In humans, the majority of sex reversal disorders (XX males and XY females) involve the translocation or mutations of the SRY gene. In such, an XX male commonly has SRY translocated to autosomes, whereas an XY male common has SRY mutations that perturb the function of the protein it encodes for. In XY female athletes, MSY has been suggested to associate with stature in an androgen-independent manner (42). While future studies of these patients may elucidate the genetics of sex differences in disease susceptibility (82), progress of this field of work is greatly limited by the rarity of such subjects.

By dissociating Sry from MSY in animal models, it has been possible to mimic the sex reversal genetics of humans and thus create XY “gonadal” females and XX “gonadal” males, which provide models to study the genetic contributions of sex chromosome in greater detail. In particular, a panel of four mouse strains constitutes what is referred to as the model of “four core genotypes” (FCG) (6). The panel comprises two types of “gonadal males” [i.e., the wild-type XY males (XYM) and the transgenic XX males (XXM)] and two types of gonadal females [i.e., the wild-type XX females (XXF) and the Sry-deleted XY females (XYF)]. Comparisons between strains have made it possible to differentiate between the effects of “gonadal” vs. that of “chromosomal” sex. For instance, differences between XYM and XXM should be due to sex chromosome complementation rather than to androgens. Likewise, differences between XXF and XYF should be due to sex complementation rather than to estrogens.

Although the FCG model has been used predominantly to show the impact of sex chromosome complementation on either behavioral traits or neural functions, the model has also made it possible to show that sex complementation affects systems other than the central nervous system (5, 6). For instance, in mouse models of autoimmune diseases, XX mice are more affected than XY mice, with the difference being
irrespective of their type of gonad (116). Likewise, treatment of FCG mice with angiotensin II causes different changes in blood pressure in XX vs. XY mice, with the effect of chromosomal sex going even in the opposite direction from that of sex steroids (50). Sex chromosome complement has also been shown to play a role in the susceptibility to some viral infections (99) as well as to several metabolic phenotypes, including adiposity, feeding behavior, fatty liver, and glucose homeostasis (65).

Despite the utility of the FCG model, it cannot fully differentiate between the respective influence of chrX and chrY. In that model, chrX may be implicated via differences in either: 1) chrX gene dosages (in case of incomplete inactivation of chrX genes), 2) chrX gene imprinting (as ChrX genes receive different paternal imprints in XX and XY mice); or 3) in chrX allelic mosaicism (present in XX, but not in XY animals) (6). However, additional lines of evidence indicate that MSY itself may have an impact on a variety of biologic functions, either in humans or in animal models.

**Genetic MSY Haplotypes and Human Disease**

In human males, comparison between groups carrying genetic variants of MSY has provided evidence implicating MSY more directly in the regulation of biologic processes and disease susceptibility. In initial studies, a *HindIII* restriction fragment polymorphism in the centromeric MSY region was shown to associate with either lower blood pressure (35), higher blood pressure (22), or higher plasma cholesterol (21) in Caucasian populations. However, these associations were not replicated in more recent and well-powered subsequent studies (56, 100, 107). One possible explanation may be the existence of multiple strata within a *HindIII* haplotype, such that a single marker does not segregate populations with enough resolution for association studies (100).

More recently, a system using a panel of binary single nucleotide polymorphisms (SNP) markers has made it possible to divide human MSY into a phylogenetic tree containing well-defined haplogroups (51). This system also provides investigators with tools to perform more robust association...
analyses. Using such markers in Caucasian British populations, investigators found the MSY haplogroup I to be associated with increased risk of coronary artery disease (20). The increased risk was independent of traditional cardiovascular risk factors, but associated with differential expression of inflammation and immunity-related genes in macrophages (20). In human immunodeficiency virus-infected humans, the MSY haplogroup I also associated with increased morbidity and decreased responses to antiviral therapies (111). In populations of African origin, there are SNPs within MSY-encoded TBL1Y and USP9Y genes, with the rs768983 SNP in the TBL1Y gene defining the MSY E1b1a haplogroup (106, 122). In these populations, the most frequent allelic combination of that SNP with another one in USP9Y associates with lower levels of triglycerides and higher HDL-cholesterol and thus with a more favorable lipoprotein pattern (106). In Japanese men, the O2b* and O2b1 haplogroups (which are confined to East Asian populations) have much lower risk of prostate cancer than other haplogroups (41). Expression analyses of chrY genes have also associated MSY genes with neurological phenotypes such as anxiety, attention deficit hyperactivity disorder (ADHD), depression, and autistic behaviors (103).

Animal Models for the Study of MSY-associated Phenotypic Traits

Several lines of evidence implicating MSY genes in physiology and disease have been obtained in animal models (particularly mice and rats). The first suggestions about possible roles of MSY genes were derived from comparisons of the progenies of reciprocal F1 crosses between two inbred strains (110, 114). As parents in such crosses represent either of the two possible strain combinations, differences in their respective progenies suggest a parent-of-origin effect for sex chromosomes. Nonetheless, reciprocal F1 hybrids also differ in terms of maternal environment, of ova cytoplasm, and in the origin of chrX. Less ambiguous models are those provided by chromosome substitution strains (also known as “consomic” strains), where one particular chromosome within a host strain A is substituted by its counterpart from a donor strain B (77). MSY consomic strains are created by first intercrossing a male from strain B to a female from strain A, and then using the resulting F1 males for multiple generations of backcrosses to the host strain A (Fig. 2). This creates the consomic strain that shares with its host strain the same 1) genetic background (including chrYPAR, autosomes, chrX, mitochondrial DNA) and 2) pre- and postnatal maternal environment (including inoculated microbiota). In the resulting model, the only differences between the consomic and the host strains relate to the composition of MSY, thus providing a way to bypass the experimental problem resulting from the lack of recombination of MSY. In mouse and rats, nearly all studies rely on identifying first sex-dependent differences between two inbred strains and then showing that the differences cosegregate with the MSY in consomic strains, a procedure that minimizes the possibility that the MSY-dependent phenotypic differences derive from genetic alteration of the MSY in the course of generation of the consomic strains. Phenotyping of consomic strains has been performed for a variety of selected phenotypes in either mice or rats (Table 1). More recently, larger-scale standardized phenotyping efforts have been performed in rats (http://pga.mcw.edu).

In both rats and mice, MSY variants were found to affect a large variety of phenotypes, and several possible mechanisms have been tested to account for these differences. In mice, the genetic contributions of MSY represent a particular situation: although classical mouse laboratory inbred strains are overwhelmingly derived from Mus musculus domesticus (of European origin), MSY is derived from M. musculus musculus (of Asian origin) for most strains (12, 78). It was therefore hypothesized that some effects of MSY may relate to the subspecies origin of MSY in mouse, and this possibility was tested in models where MSY from a large number of donor strains were introgressed into one particular host strain (18, 120). In one study, MSY from M. musculus musculus associated with significantly lower plasma HDL-cholesterol levels than in strains carrying MSY from M. musculus domesticus (120). In contrast, associations cosegregating with the MSY subspecies origin were not found as a possible explanation for some immune-related phenotypes (18).
Table 1. MSY consomic rodent models

<table>
<thead>
<tr>
<th>Species</th>
<th>Host strain(s)</th>
<th>MSY donor strain(s)</th>
<th>Phenotypes</th>
<th>Possible MSY genetic variation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mouse</td>
<td>PHL C57BL/FaSt</td>
<td>PHH PHH/PHL</td>
<td>adult plasma testosterone</td>
<td>subspecies origin</td>
<td>Jutley et al. (52)</td>
</tr>
<tr>
<td></td>
<td>NZB/BinJ</td>
<td>CBA/HGnc</td>
<td>adult plasma testosterone</td>
<td>subspecies origin</td>
<td>Botbol et al. (14)</td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>RF/F</td>
<td>adult plasma testosterone</td>
<td>n.d.</td>
<td>Case et al. (17)</td>
<td></td>
</tr>
<tr>
<td>C57BL/FaSt</td>
<td>CBA/FaSt</td>
<td>testosterone-dependent weight of seminal vesicle</td>
<td>subspecies origin</td>
<td>Jutley et al. (52)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6</td>
<td>AKR</td>
<td>chemosensory identity</td>
<td>n.d.</td>
<td>Yamazaki et al. (135)</td>
<td></td>
</tr>
<tr>
<td>DBA1</td>
<td>C57BL10</td>
<td>testosterone-dependent urinary chemosignals</td>
<td>n.d.</td>
<td>Monahan and Maxson (74)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>A/J</td>
<td>testosterone-dependent adult cardiomyocyte size</td>
<td>n.d.</td>
<td>Llamas et al. (66)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>A/J</td>
<td>perinatal testosterone exposure</td>
<td>n.d.</td>
<td>Praktiknjo et al. (89)</td>
<td></td>
</tr>
<tr>
<td>NZB/BinJ</td>
<td>CBA/HGnc</td>
<td>brain neurotransmitters</td>
<td>subspecies origin</td>
<td>Botbol et al. (14)</td>
<td></td>
</tr>
<tr>
<td>DBA1</td>
<td>C57BL10</td>
<td>aggressive behavior</td>
<td>n.d.</td>
<td>Monahan and Maxson. (74)</td>
<td></td>
</tr>
<tr>
<td>SAL/LAL</td>
<td>LAL/SAL</td>
<td>hippocampal morphology</td>
<td>n.d.</td>
<td>Hensbroek et al. (48)</td>
<td></td>
</tr>
<tr>
<td>DH/Sgn</td>
<td>KK/Ta</td>
<td>body weight/length</td>
<td>n.d.</td>
<td>Suto (119)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>SJL</td>
<td>T cell transcriptome/chromatin remodeling genes</td>
<td>copy numbers of Sry and Rbmy</td>
<td>Case et al. (18)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>A/J</td>
<td>cardiac transcriptome</td>
<td>n.d.</td>
<td>Llamas et al. (66)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>A/J</td>
<td>autosomal chromatin architecture/chromatin remodeling genes</td>
<td>n.d.</td>
<td>Praktiknjo et al. (89)</td>
<td></td>
</tr>
<tr>
<td>MF1</td>
<td>RIII</td>
<td>cell number in blastocysts</td>
<td>deletion</td>
<td>Burgyne et al. (16)</td>
<td></td>
</tr>
<tr>
<td>DH/Sgn</td>
<td>16 inbred strains</td>
<td>plasma HDL-cholesterol</td>
<td>subspecies origin</td>
<td>Suto and Satou (120)</td>
<td></td>
</tr>
<tr>
<td>C57BL/6J</td>
<td>12 inbred strains</td>
<td>EAE/CVB3 myocarditis</td>
<td>copy number of Sry and Rbmy genes</td>
<td>Case et al. (17)</td>
<td></td>
</tr>
<tr>
<td>Rat</td>
<td>WKY/Hsd</td>
<td>SHR/Hsd</td>
<td>blood pressure/sympathetic activity/behavior response</td>
<td>Sry</td>
<td>Case et al. (18)</td>
</tr>
<tr>
<td></td>
<td>SHR/Hsd</td>
<td>WKY/Hsd</td>
<td></td>
<td></td>
<td>Ely and Turner (39)</td>
</tr>
<tr>
<td>WKY</td>
<td>SHRSP</td>
<td>WKY/Hsd</td>
<td>blood pressure/metabolic syndrome</td>
<td>n.d.</td>
<td>Turner et al. (124)</td>
</tr>
<tr>
<td>BN</td>
<td>SHR/Hsd</td>
<td>BN</td>
<td>blood pressure/blood lipids</td>
<td>n.d.</td>
<td>Davidson et al. (30)</td>
</tr>
<tr>
<td>BN</td>
<td>SS</td>
<td>urinary albumin</td>
<td>n.d.</td>
<td>Strahorn et al. (118)</td>
<td></td>
</tr>
</tbody>
</table>

MSY, male-specific region on chromosome Y; n.d., not determined.

Interestingly, MSY-dependent phenotypes in mouse often appear to relate to several aspects of androgen biology, including adult plasma testosterone levels (14, 17, 52, 121) or the effects of either postpubertal (52, 66, 74) or perinatal testosterone (89) on several organs. In the course of evolution, where genes are usually selected on the basis of their advantage for the survival of the species, it makes evolutionary sense that the father-related transmission of MSY variants could relate with androgen-related biology. Indeed, evolutionary advantages of variants transmitted through the paternal cell line are bound to concern preferentially male-related biologic features, which includes the production of androgens and/or their biologic effects (45). Of note, the effects of sex steroid hormones fall in two categories: 1) the organizational effects (which are permanent and occur early in development) and 2) the activational effects (which occur after puberty and are transient) (4, 24). In essence, the early organizational phase programs the developing organism for how it will respond to sex hormones later during adult life. The organizational effects of sex steroids are limited to a restricted period of sensitivity called the “critical period of sexual differentiation”. The timing of that critical period is species specific: in rodents, it includes the last 2 days of gestation and the first days of postnatal life (28, 132). In all mammals studied to date, there is an abrupt discharge of testosterone in the first few hours after birth (26). Postnatal testosterone appears to have actions that are distinct from those of testosterone during gestation (108). Altogether, perinatal testosterone has important effects on several aspects of sexual differentiation, as it shapes several behavioral patterns, the development of external genitalia, the programing of the adult pattern of secretion of gonadotrophins and growth hormone, and several brain neuroanatomical features (129, 132). However, it appears that perinatal testosterone may also have long-term consequences on other phenotypes, including the rate of body weight increase (9), several metabolic functions (31, 76, 79, 101), and the general organization of circadian rhythms (2, 139). In at least one instance, the effects of MSY on adult phenotypic traits were found to be a consequence of differential effects of testosterone occurring previously at perinatal times. Male C57BL/6J mice differ from their C57BL/6J.MSYA/J counterparts as: 1) postpubertal testosterone has hypertrophic effects on adult C57BL/6J cardiomyocytes, but not in those from C57BL/6J.MSYA/J counterparts (89). When the perinatal actions of androgens in developing C57BL/6J mice are prevented by administration of flutamide (an androgen antagonist), testosterone no longer exerts hypertrophic effects on cardiomyocytes later during adult life; conversely, when C57BL/6J.MSYA/J are exposed to higher levels of testosterone during the perinatal period, their cardiomyocytes become sensitive to the hypertrophic effects of testosterone later during adult life.

Despite the several studies showing possible roles of androgens in the mediation of the phenotypic effects of MSY variants, there are also other studies reporting that these effects may be independent from testosterone. One example is that of differences in the concentration of neurotransmitters in certain...
As mentioned previously, the expression of several MSY genes outside of the testis (Fig. 1) indicates that they may play roles unrelated to male reproductive functions. Among those genes, the one that may have been characterized in greatest detail for its roles outside of testis is Sry. In the brain, Sry has been tied to a variety of functions, as reviewed recently (124). In the human digestive system (particularly in the colon), SRY is expressed at readily detectable levels, and it has been suggested as a potential regulator of genes involved in Hirschprung disease, potentially explaining the 5:1 male to female ratio of this disease (64). Likewise, the readily detectable levels of SRY expression in human skin may be significant in light of studies showing correlations between SRY expression levels and male pattern baldness (23). Interestingly, SRY is also expressed in male pluripotent stem cells being induced by reprogramming (102), which is of interest in light of the well-known role of the SRY homologous protein SOX2 in pluripotency (70). All of the SOX genes (including SRY) contain a structurally homologous HMG domain sequence that binds and regulates DNA, with Sry being most similar to the SoxB subgroup that includes Sox1, Sox2, and Sox3 (92). The homology of Sry with SoxB genes may also be consistent with the documented functions of Sry in the brain (124), as it may overlap with some of the critical functions played by Sox3 (the Sry orthologous gene on chrX) in brain development (25).

Even in mouse, where expression of Sry is almost nondetectable in tissues other than testes, low levels of expression have been found in some brain nuclei, where it may promote sexually dimorphic expression of particular genes and yield a genetically “masculinized” brain (32, 33, 132).

Studies comparing the expression profiles of both Sry and Sox3 in rat tissues showed expression of Sry (but not of Sox3) within kidneys, a pattern that was further shown to hold true in multiple primates (3). Delivery of a Sry expression vector into the kidneys of normotensive rats can elevate blood pressure through changes in the sympathetic nervous system (40) and regulatory changes of components of the renin-angiotensin system (RAS) (36). The tissue expression profile for Sry suggests that it gained a kidney-specific function over Sox3 (3). An additionally duplicated Sry3 gene has been found in the SHR strain (125). Consomic rats suggest that this duplicated Sry3 contains a nonsynonymous variant, P131T, that alters regulation of promoters of RAS genes (95). The overlap between Sry and Sox3 in their ability to regulate promoters of RAS genes may be pertinent to their roles within the brain as well, because RAS genes share with Sox3 the ability to contribute to development of the hypothalamic-pituitary-adrenal axis (3). Both the rat Sry and the human SRY proteins have similar effects in their abilities to regulate multiple RAS genes (73, 95). Direct binding of SRY protein to promoter sequence of RAS genes such as MASI has also been shown for human (3).

Sry is not the only gene on the MSY that may be linked to functions outside of testes. In XYY males, expression of the NLGN4Y gene correlates with neurological phenotypes such as anxiety, ADHD, depression, and autistic behaviors (103). In cancer, a recent review has detailed the potential roles of several MSY genes, including TSPY, RBMY, TGIF2LY, VCY, BPY2, DAZ, CDY1, and SRY (55). On the basis of network-based expression analysis, up to 22 genes on the MSY have been identified as being possibly linked to cancers in several tissues, including the prostate (54).

In C57BL/6J and C57BL/6J.MSY/−/− mice, strain-specific differences were found in the cardiac expression levels of Uty and Kelm5d, two MSY-encoded genes that both belong to the family of H3 lysine demethylases (109, 134) and are also ubiquitously expressed in humans (Fig. 1). Although differences in the expression of these genes were limited in mouse hearts to just the first postnatal day, that period corresponds to the time known as the critical period when androgens are believed to exert programming effects. Interestingly, UTY in humans was also found to be expressed at lower levels in
macrophages of men with MSY haplogroup I (which associates with increased risk of coronary artery disease) (13).

In mammals, chrY and chrX evolved from an ancestral pair of autosomes. In humans, only 3% of the ancestral genes have been retained on chrY, compared with 98% on chrX (10). However, recent sequencing efforts of the euchromatic part of MSY across eight mammals has revealed that the genes that have resisted decay belong to a group that cannot result from a random selection process. Since these genes correspond to ancestral X-Y gene pairs that escape inactivation in females, the surviving gene pairs (X-X in females, X-Y in males) appear to be dosage-sensitive regulators. Moreover, these genes are expressed in a wide spectrum of tissues and can be categorized into five overlapping functional families, i.e., regulators of chromatin modification, of ubiquitination, of transcription, of splicing, and of translation (10). These genes therefore appear to be potentially important regulators of a wide range of vital cellular functions in several tissues, which provides possible mechanisms whereby MSY gene variants could contribute to many different aspects of health and disease.

**MSY Variants May Regulate Autosomal Genes via Epigenetic Chromatin Remodeling**

Several lines of evidence indicate that some of the effects of MSY genes are mediated via changes in chromatin organization. Hearts from both castrated and intact C57BL/6J and C57BL/6J.MSYA/J mice show differences in their transcriptomic profile, along with strain-specific differences in the identity of testosterone-sensitive genes (66). These differences in gene expression are accompanied by strain-specific differences in the distribution of ARs in cardiac chromatin, with some of the differentially occupied regions matching the loci of differentially responding cardiac genes (89). Differential distribution of ARs is also observed in cardiac chromatin from 1-day-old pups, along with differences in the distribution of accessible chromatin sites and of H3K4me3 marks (which characterize active promoters). Within the genes closest to sites showing differential H3K4 trimethylation, functional enrichment was observed for genes related to nucleosome organization and/or regulation of nucleic acids. Other studies using C57BL/6J and C57BL/6J.MSYA/J mice have also documented strain-specific differences in the transcriptomes of CD4+ T cells (18). When differentially expressed genes were tested for enrichment for functional Gene Ontology (GO) terms, many terms were associated with chromatin modification events whose dysregulation may affect gene transcription. MSY was also associated with alternative and differential splicing of genes in both CD4+ T cells and macrophages in C57BL/6J and C57BL/6J.MSYA/J male mice (18).

The above results are compatible with recent results describing the functions of genes retained on the MSY of several vertebrates. The numbers of such genes vary across species, ranging from 17 in humans and only nine in mice (10). However, five out of these nine genes (Sry, Rbmy, Ube1l, Usp9y, and Zfy) have expressed testis-specific expression in mice. The remaining broadly expressed genes are Uty and Kdm5d (involved in the processes of chromatin modification and/or ubiquitination) and Ddx3y and Eif2s3y (involved in the processes of RNA splicing and/or translation). Uty is the counterpart on MSY of the Utx gene on chrX; in mice, it encodes a protein that forms functional complexes with other factors and has functions that are partially redundant with those of Utx (61, 113). Despite being part of the H3 lysine demethylases family, Uty (in contrast to Utx) is devoid of such enzymatic activity but acts primarily as a recruiter of the H3K4 methyl-transferase complex (113). In addition, UTX and UTY have the ability to regulate the activity of several promoters by association with BRG1 (the ATP-dependent subunit of the chromatin remodeling Swi/Snf complex) (61, 113). BRG1 has also been shown to induce chromatin modifications that allow ARs to bind to their cognate response elements and is required for effective transcriptional regulation of androgen-regulated genes (29, 63). The data suggesting a role of Uty in MSY-dependent phenotypes in mice might be interesting in light of the fact that in humans, MSY haplogroups show differences in the level of UTY expression in macrophages (13).

Besides histone demethylases, other data have suggested alternative mechanisms whereby MSY variants may exert their effects via chromatin regulatory factors in mice. By comparing C57BL/6J mice to consomic strains carrying MSY from 11 other inbred strains, Case et al. (19) found that the numbers of copies of the multicopy Sly and Rbmy genes correlated with the severity scores of either EAE or viral myocarditis, as well as with upregulation of genes in immune cells. However, Rbmy and Sly are unlikely to exert their regulatory actions in immune cells via their protein products, because their expression is testis and germ-line specific. Alternatively, it has been proposed that higher frequencies of multicopy MSY genes may create domains that bind to chromatin remodeling proteins, leading to fewer euchromatic DNA regions and less transcriptional activity by restricting the overall availability of the chromatin remodeling proteins (18, 19). Interestingly, similar mechanisms have recently been proposed for Drosophila melanogaster, where comparison of MSY consomic strains showed that polymorphic MSY associate with quantitative effects on the expression of several hundred autosomal genes (62, 138). These effects do not appear to be linked to the products of protein-coding genes but, rather, to effects of MSY heterochromatin that may sequester chromatin regulatory factors.

The SRY protein is also known to play roles in the remodeling of chromatin architecture. It recruits heterochromatic protein 1 (HP1) to DNA through the most abundant repressor domain in the human genome, KRAB (81). KRAB subsequently recruits the epigenetic regulation complex of KAP1, which includes HP1, DNA methyltransferases, and the NuRD Histone deacetylase complex (49, 85). The physical interaction between SRY and the KRAB domain has been mapped for both proteins (86, 93). Knockdown of KRAB containing proteins decreases some of the pathways needed for testis determination, namely Sox9 (88). Even without the recruitment of the KRAB domain to regions of the genome, following an initial recruitment of KAP1 to genomic regions, epigenetic modifications are maintained for >50 population doublings (7). This suggests a difficult to test hypothesis, where expression of MSY genes such as Sry during early development could epigenetically “masculinize” cells, and where the masculinized state is maintained following the removal of expression of the MSY gene.
To expand on previous identification efforts of SNPs that allow the classification of MSY variants into different haplogroups (51), a high-coverage sequencing study was recently performed in a wide range of samples that covers the majority of clades of the phylogeny. These efforts yielded 13,261 high-confidence SNPs, 66% of which were previously unreported (46). When these results were combined to that of five other recent high-coverage sequencing studies, a total of 33,479 potential SNPs were identified (albeit with lower confidence thresholds), and it is anticipated than even higher numbers of SNPs will be identified in the near future (46). These newly discovered SNPs, when applied to populations originating from as many as 128 countries, have allowed for an expansion and redefinition of chrY haplotypes and subhaplotype groups (83, 96, 131). Mapping SNVs in larger cohorts by next-generation sequencing is becoming more common, allowing for mapping the characteristics of human migration and the discovery of de novo variants (47, 53). Interestingly, some have recently developed a new genotyping technique that combines validated multiplex PCR with Ion Torrent-based sequencing of amplified fragments, along with software tools to identify MSY SNPs (96). This technique promises to make it possible to genotype simultaneously many more SNPs than what is possible with techniques based on single-based primer extension chemistry and yields tools to perform efficient high-throughput genotyping for future phenotypic trait association studies.

The recent sequencing efforts have also be useful to further understand the idiosyncrasies of chrY genetics. For instance, out of the 13,261 high-definition MSY variants recently reported (46), 259 result in missense, nonsense, and frame shift mutations within MSY genes (Fig. 1). The frequency distribution of damaging variations in MSY single-copy genes also suggests that purifying selection is ongoing (46), which is in keeping with the findings that the genes conserved from the ancestral chromosome represent in fact a group of highly conserved genes (10). As mentioned previously, the lack of interchromosomal recombination impairs the removal of inserted retrotransposed sequences. This has resulted in the occurrence on the MSY of retrotransposed genes such as Med14y in rat (94) and CDY in humans (59). Also, studies have revealed that sequences on the MSY are less static than previously anticipated. For instance, gene conversion is the process of nonreciprocal transfer of sequence without crossover (either within or between chromosomes). Gene conversion events were found to be abundant between arms of palindromes on the MSY (105), as well as between the MSY and chrX (104, 123). In the bovine MSY, conversion has driven massive gene duplications (10), with copy number variants of TSPY in bull (Bos taurus and Bos indicus) being associated with the quality of bull semen (75). Sequencing of the rat MSY (10) revealed duplication of the Sry gene with break points of the gene duplication occurring at LINE elements, which suggests convergence (94). Similarly, other wild-caught rat species have undergone Sry duplication independently (15, 68, 94), domestic cats have multiple Sry genes (84), and humans with Turner syndrome or exposure to radiation can undergo SRY gene duplication (90, 91).

Progress in sequencing of the MSY has also revealed the extent of differences between several species. In particular, the mouse MSY appears to be spectacularly different from humans or other higher primates (117): 1) only 2.2% of the MSY derives from the ancestral autosomes that gave rise to the mammalian sex chromosomes; 2) the remaining of the MSY is dominated by acquired amplicons of three protein-coding and rodent-specific gene families, comprising 126 Sly, 197 Srsy, and 306 Ssty copies; and 3) in contrast to other species, 99.9% of the mouse MSY is euchro-
matic. As mentioned above, copy number variation of Sly in the mouse MSY shows linkage to a paternal parent-of-origin effect on autoimmune disease in female offspring (20).

Of note, sequencing approaches still struggle with identification of indels on chrY, which may be interesting in light of data showing that loss of chrY (LOY) may have an impact on cell functions. In prostate cancer, chrY is one of the most frequently lost chromosomes (128). In the PC-3 human prostate cancer cell line (which is devoid of chrY), back-transfer of chrY suppresses the tumorigenicity of these cells, thus suggesting a functional importance of chrY (128). LOY is also an event that occurs frequently in normal hematopoietic cells from elderly men (44). In cancer-free men, it was determined that the median survival times of subjects with LOY were 5.5 times shorter than that of their counterparts without LOY, even after adjustment for common confounders, including age, hypertension, exercise, smoking, diabetes, body mass index, blood lipid profiles, and ancestry (44). Likewise, LOY in hematopoietic cells associates with increased risk for nonhematological cancer mortality (44). Interestingly, even smoking has been associated in humans with transient and dose-dependent mutagenic effects of LOY, including the removal of stretches of MSY sequences (34).

The Future of MSY Research

ChrY has long constituted one of the most challenging regions of the mammalian genome. The past several years have seen an expansion of data on chrY, allowing for a better understanding of its evolutionary characteristics, of its sequence variations, and the impact of the latter on disease and complex traits. With the continuing decrease in the cost of genome sequencing, future chrY variants will be identified within the several thousands of genomes that will be sequenced in the coming years. Further resolution of variants will be useful to refine our understanding of how chrY may associate with disease traits. Additional future human research will likely focus on somatic mutations, LOY, and the impact of gene convergence.

The new emerging picture of MSY genes, as evidenced by the fact that they are broadly expressed in many tissues outside of testes and correspond to regulators of processes as fundamental as transcription, translation, and protein stability, is that they may have a significant impact on a variety of biologic functions. One of the mechanisms whereby MSY variants may associate with disease traits is that it appears that they may ultimately lead to changes in expression of other autosomal genes. Animal models therefore remain critical for narrowing down such mechanisms. Having reliable sequence for chrY in rodent models will make it possible to genetically modify MSY genes to study disease associations. Future large-scale analyses of phenotypic traits in chrY consomic mouse and rat strains may also increase our ability to correlate chrY gene variants to phenotype and ultimately understand how orthologous genes may drive disease state in humans. One area that has not received much attention yet is the potential involvement of noncoding RNA transcripts originating from chrY, as these genes are likely to play a role in chrY biology.

Altogether, there is increasing evidence for connections between chrY genetic variation and disease states, as summarized in Fig. 3. Continuing association studies will be critical to help explain the many sex-biased disease states in human that are contributed to not only by the classical sex steroid hormones, but also by chrY genetics.

GRANTS

J. W. Prokop was supported by National Institutes of Health Office of the Director Grant K01ES-025435. C. F. Deschepper was supported by Canadian Institutes of Health Research Grant MOP-93583.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: J.W.P. and C.F.D. conception and design of research; J.W.P. analyzed data; J.W.P. and C.F.D. prepared figures; J.W.P. and C.F.D. drafted manuscript; J.W.P. and C.F.D. edited and revised manuscript; J.W.P. and C.F.D. approved final version of manuscript.

REFERENCES

6. Arnold AP, Chen X. What does the “four core genotypes” mouse model tell us about sex differences in the brain and other tissues? Front Neuroendocrinol 30: 1–9, 2009.


120. Turner ME, Fishkind J, Dunmore J, Fishkind J, Milsted A. By which locus is the Y chromosome haplogroup E1b1b1 (E-P2) revealed through the use of newly characterized binary polymorphisms. PLoS One 6: e16073, 2011.
121. Turner ME, Fishkind J, Dunmore J, Fishkind J, Milsted A. By which locus is the Y chromosome haplogroup E1b1b1 (E-P2) revealed through the use of newly characterized binary polymorphisms. PLoS One 6: e16073, 2011.
123. Turner ME, Fishkind J, Dunmore J, Ely D, Milsted A. By which locus is the Y chromosome haplogroup E1b1b1 (E-P2) revealed through the use of newly characterized binary polymorphisms. PLoS One 6: e16073, 2011.
