

MIT Open Access Articles

The Biology and Evolution of Mammalian Y Chromosomes

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Hughes, Jennifer F., and Page, David C.. "The Biology and Evolution of Mammalian Y Chromosomes." Annual Review of Genetics 49, no. 1 (November 23, 2015): 507–527. © 2015 by Annual Reviews.

As Published: <http://dx.doi.org/10.1146/annurev-genet-112414-055311>

Publisher: Annual Reviews

Persistent URL: <http://hdl.handle.net/1721.1/108037>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



The Biology and Evolution of Mammalian Y Chromosomes

Jennifer F. Hughes and David C. Page

Whitehead Institute, Howard Hughes Medical Institute, and Department of Biology,

Massachusetts Institute of Technology, Cambridge, Massachusetts 02142;

email: jhughes@wi.mit.edu

Table of Contents

INTRODUCTION	3
The evolutionary history of the Y chromosome	3
Ignorance of Y chromosomes	4
Three biological themes in Y chromosome research	4
BIOLOGICAL THEME #1: TESTIS DETERMINATION	6
Sex chromosomes and sex reversal	6
Searching for the testis-determining gene	7
BIOLOGICAL THEME #2: SPERMATOGENESIS	7
Sequence of the human MSY: Fragile hall of mirrors	8
Sequence of the mouse MSY: Fallout of X vs. Y battle	10
BIOLOGICAL THEME #3: BEYOND THE REPRODUCTIVE TRACT	13
MSY genes and Turner Syndrome	13
Palindromes, isodicentric Y chromosomes, and Turner syndrome	15
Evolutionary analysis reveals that MSY genes are special	16
LOOKING FORWARD: THE Y CHROMOSOME AND SEXUAL DIMORPHISM	18

Keywords

MSY, sex chromosomes, sex determination, spermatogenesis, Turner syndrome, sexual dimorphism

Abstract

Mammals have the oldest sex chromosome system known: the mammalian X and Y chromosomes evolved from ordinary autosomes beginning at least 180 million years ago. Despite their shared ancestry, mammalian Y chromosomes display enormous variation among species in size, gene content, and structural complexity. Several unique features of the Y chromosome – its lack of a homologous partner for crossing-over, its functional specialization for spermatogenesis, and its high degree of sequence amplification – contribute to this extreme variation. However, amidst this evolutionary turmoil many commonalities have been revealed, which have contributed to our understanding of the selective pressures driving the evolution and biology of the Y chromosome. Two biological themes have defined Y chromosome research over the past six decades: testis determination and spermatogenesis. With recent insights into the Y chromosome's roles beyond the reproductive tract, a third biological theme begins to emerge – a theme that promises to broaden the reach of Y chromosome research by shedding light on fundamental sex differences in human health and disease.

INTRODUCTION

The evolutionary history of the Y chromosome

Mammalian sex chromosomes evolved from an ordinary pair of autosomes (85). The X and Y chromosomes began to differentiate at least 180 million years ago, before the divergence of the marsupial and placental mammalian lineages. The autosomes that gave rise to the mammalian sex chromosomes still exist as autosomes in birds today (**Figure 1**). The first step in X-Y differentiation was the acquisition of a testis-determining gene by the proto-Y chromosome. Next, a series of large-scale inversions, most likely on the Y chromosome, suppressed recombination between the X and Y chromosomes in a step-wise fashion, creating at least four evolutionary strata (4, 48, 63, 89). Throughout this process, the X chromosome retained a partner for crossing over in females, but the Y chromosome's opportunities for crossing over became increasingly restricted as stratification progressed. Outside the pseudoautosomal regions (PARs), highly differentiated mammalian sex chromosomes do not ordinarily engage in crossing over with each other. In the absence of crossing over, the male-specific region of the Y chromosome or MSY (which excludes the PARs) was subject to genetic decay, resulting in deletions and gene losses that collectively decimated the MSY (15, 82, 99). In humans, the euchromatin of the present-day MSY is less than one sixth the size of the X chromosome (23 Mb compared to 155 Mb) and has retained only 3% of the genes that were present on the ancestral autosome pair (17 of ~640) (4, 100, 107). The dearth of molecularly defined genes in the human MSY confirmed, in the minds of many biologists, its earlier characterization as a genetic wasteland and even fueled jovial speculation that the demise of the Y chromosome was imminent (1, 38, 39).

Ignorance of Y chromosomes

In diverse species with XX/XY sex chromosomes, biological understanding of Y chromosomes has consistently and dramatically lagged behind that of autosomes and X chromosomes. This is true both in human genetics and in experimental genetic systems like *Drosophila* and mouse.

Consider the human Y chromosome. In the first half of the 20th century, dozens of scientific publications claimed to provide evidence that one trait or another was Y-linked. By the 1950's, however, all such pedigree-based evidence of Y-linked genes had been discredited by the noted *Drosophila* and human geneticist Curt Stern (109). In no case has the existence of an MSY gene been correctly inferred from the transmission of phenotypes in human pedigrees (but see SIDEBAR). This methodological difficulty was misinterpreted as evidence that there are few or no genes in the human MSY. Consequently, the human Y chromosome came to be viewed as having little or no function – and little or no medical relevance.

Three biological themes in Y chromosome research

The advent of molecular genetics and genomics in the closing decades of the 20th century helped bring to light the biological and medical relevance of the Y chromosome. With new molecular tools in hand, such as DNA cloning and nucleic acid hybridization, studies of Y-chromosome gene function came to rely not on classical transmission genetics, but on DNA-based characterization of spontaneously arising sex chromosome anomalies. Three biological themes – one of which is just now beginning to unfold – have come to define Y-chromosome research in the age of molecular genetics and genomics.

The first biological theme, which defined the field of Y chromosome research for more than three decades, involved the pursuit of the Y-linked testis-determining factor, which was the

one and only genetic function attributed to the Y chromosome well into the 1970's, and which offered the only obvious avenue for study. Studies of human individuals carrying part but not all of the Y chromosome, combined with functional studies in mice, culminated in the establishment of *SRY* as the testis-determining gene on the mammalian Y chromosome in 1991 (5, 42, 54, 59).

The second biological theme, which the field has pursued in earnest since the 1990's, concerns the roles of the Y chromosome in sperm production and fertility. Research accelerated with the publication of the DNA sequence of the human MSY in 2003. Over the ensuing decade, and following the sequencing of two additional primate MSYs (49, 50), the first MSY from a model organism – the mouse – was sequenced to completion (108). The mouse MSY sequence dramatically expanded the known range of diversity in mammalian Y-chromosome structure and gene content, enabling investigators to understand more deeply the selective forces driving Y-chromosome evolution that operate at the level of sperm production and fertility.

While our understanding of the Y-chromosome's roles in testis determination and spermatogenesis are firmly established, a third biological theme is just emerging: the quest to comprehend the Y chromosome's influence beyond the reproductive tract. Recent comparative genomic analyses, combined with earlier associations made between the Y chromosome and the multifaceted phenotypes found in females with Turner Syndrome (45,X) (29), provide intriguing clues regarding the roles that many Y-linked genes play throughout the body. The presence or absence of a Y chromosome (i.e., XY or XX sex chromosome constitution) may influence development and physiology on a multitude of levels, contributing to sex differences in health and disease both within and beyond the reproductive tract.

BIOLOGICAL THEME #1: TESTIS DETERMINATION

Sex chromosomes and sex reversal

The first era in molecularly-grounded Y chromosome research focused on one biological role of the Y chromosome: testis determination. In 1959, reports of 45,X females (Turner syndrome, with oocyte-depleted ovaries) and 47,XXY males (Klinefelter syndrome, with germ-cell-depleted testes) established the existence of a testis-determining gene on the human Y chromosome (32, 53), and the ensuing three decades of Y chromosome research focused on the pursuit of this gene. In therian mammals (marsupials and placentals), the presence of a Y chromosome is sufficient to trigger testis formation during fetal development, setting in motion a cascade of events required for anatomic masculinization more broadly, including the production of testosterone (by testicular somatic cells known as Leydig cells). Although the Y chromosome triggers or activates the pathway of testis development, many other genetic factors – most of which are located in autosomes or the X chromosome – are also involved in testicular and subsequent male development (80). For example, 46,XY individuals with complete androgen insensitivity syndrome carry intact Y chromosomes and form testes rather than ovaries, but they develop, externally, as females because their cells cannot respond to testosterone or other androgens (47). In human and mouse, mutations in at least 17 autosomal or X-linked genes cause male-to-female gonadal sex reversal and/or abnormal gonadal development (90).

The first studies investigating the Y-chromosome's role in testis determination focused on individuals with Y-chromosome deletions and translocations that presented with sex-reversal phenotypes (23, 26, 43, 76). It became evident that these Y-chromosome rearrangements frequently arose from aberrant crossing over between the X and Y chromosomes (132). The human X and Y chromosomes normally cross-over only within the small PARs located at their

termini. However, several loci on the short arm of the MSY are sites of occasional ectopic crossing over with the X chromosome (122, 127). Depending on which product of this ectopic crossing-over is transmitted by the father, two very different phenotypes are observed in resultant offspring (**Figure 2**): 1) XX males (with germ-cell-depleted testes) in whom a terminal portion of the short arm of the Y chromosome (Yp) has been translocated to an almost complete paternal X chromosome and 2) XY females (with ovaries lacking normal germ cells) in whom a terminal portion of Yp has been replaced by a terminal portion of Xp. These observations led to the conclusion that the testis-determining gene is located on the terminal portion of Yp.

Searching for the testis-determining gene

Fine mapping of these X-Y crossover products in XX males and XY females eventually narrowed the region on the human Y chromosome carrying the testis-determining gene to a 300-kb segment of Yp (86). The first candidate gene identified within this region was *ZFY*, encoding a zinc-finger protein that has two homologs on the mouse Y chromosome (87). Subsequent studies excluded *ZFY* based on refined mapping data in human and expression analysis in mouse (58, 88). The second gene cloned from the critical region of Yp was the transcription factor *SRY* (*Sex-determining Region Y*), which was found to be conserved on the Y chromosome of multiple mammals and expressed in mouse gonads during the appropriate stage of fetal development (42, 60, 106). After *SRY* was cloned, evidence unequivocally confirming its pivotal role in testis determination soon followed. First, two separate studies identified human XY females (with ovaries lacking normal germ cells) bearing *de novo* point mutations in *SRY* (5, 54). Second, XX mice bearing an autosomal transgene of the mouse *Sry* genomic locus were shown to develop as anatomic males, with testes and male internal and external genitalia (59).

BIOLOGICAL THEME #2: SPERMATOGENESIS

The existence of a testis-determining gene on the MSY, and the molecular identification of that gene as *SRY*, were widely viewed as an isolated, exceptional counterpoint on an otherwise barren and desolate chromosome. The genetic wasteland model of the MSY dominated thinking among biologists into the 1990's, when advances in Y-chromosome genomics ushered in a second era of intensive molecular genetic research. Direct genomic analysis of the human MSY, coupled with molecular characterizations of recurrent MSY deletions and their phenotypic consequences, led to an appreciation of the MSY's central role in spermatogenesis.

Sequence of the human MSY: A fragile hall of mirrors

In 1992, a complete physical map of the human MSY was produced (31), representing one of the first two human chromosomes to be mapped and cloned in its entirety (the other being chromosome 21; (16)). The YAC (yeast artificial chromosome)-based physical map of the human MSY was complemented by a second map, published concurrently and based on naturally occurring deletions as found in children or adults (120). However, certain regions of the human MSY proved to be extremely difficult to map; YACs could not be assembled into large contigs, and markers could not be ordered with confidence. The most problematic region was *AZFc* (*AZoospermia Factor c*), which came to be recognized as the most frequently deleted portion of the Y chromosome in men with spermatogenic failure (37, 57, 83, 94, 105, 119). Several independent efforts to map the *AZFc* region encountered similar difficulties (56, 110, 131). Such difficulties are frequently associated with highly repetitive genomic sequences, of which the Y

chromosome in general and the *AZFc* region in particular would eventually prove to be unmatched exemplars.

Although the technical challenges of these mapping studies hinted at the repetitiveness and structural complexity of the *AZFc* region, the region remained an enigma until it was mapped and sequenced to completion in 2001. Because of the technical challenges that the *AZFc* region posed, its mapping and sequencing required the conception and implementation of a new methodology known as SHIMS (Single Haplotype Iterative Mapping and Sequencing) to disentangle extremely repetitive genomic regions (48, 61)(See SIDEBAR). First, BACs or fosmids are used as sequencing templates because they are less prone to chimerism than the larger YAC clones. Second, clones derived from the Y chromosome of one individual, bearing a single haplotype, are used for sequencing to eliminate polymorphisms, which could otherwise be confused with paralogous duplications. Third, tiling paths are constructed with a high degree of overlap between adjacent clones so that rare, single-nucleotide differences between paralogous repeat copies can be identified. Last, the map is refined through an iterative process, sorting repeat copies into correct tiling paths based on sequence information. The finished *AZFc* sequence, assembled using SHIMS, spanned 4.5 Mb and contained three massive palindromes, three other large inverted repeats, and three direct repeats (**Figure 3**) (61). This region is gene-rich, harboring a total of 15 genes in five distinct gene families, all expressed predominantly if not exclusively in testes.

In 2003, the SHIMS-produced sequence of the entire human MSY was completed (107), and the MSY still stands as the only chromosome in the human reference sequence that was assembled from a single haplotype. Most euchromatic sequence in the human MSY falls into two classes, based on patterns of similarity to other sequences on the X or Y chromosomes (**Figure 3**)

(107). “X-degenerate” regions are strewn with single-copy genes that have homologs on the X chromosome. These X-degenerate genes on the Y chromosome are living fossils that attest to the X and Y chromosomes’ shared evolutionary origins, as an ordinary pair of autosomes.

“Ampliconic” regions of the MSY are composed of sequences that exhibit striking similarity – as much as 99.99% identity over tens or hundreds of kilobases (kb) – to other MSY sequences.

Some ampliconic sequences are derived from autosomal transpositions and others are derived from the X-Y common ancestor. The ampliconic regions of the human MSY are dominated by large palindromes, or mirror-image repeat structures. The X-degenerate and ampliconic sequence classes are also distinct in their functional characteristics, including their gene content (64, 107).

The X-degenerate genes are single-copy, and most are expressed in multiple tissues. The ampliconic genes are multi-copy and expressed predominantly if not exclusively in the testis, suggestive of specialized functions related to spermatogenesis.

The availability of the human MSY sequence enabled studies of the molecular mechanisms underlying recurrent MSY deletions and other types of rearrangements, and this in turn led to a deeper appreciation of the MSY’s central role in spermatogenesis. MSY deletions were known to be the most common genetic cause of spermatogenic failure (reduced or absent sperm production, with few or no sperm in semen) in human populations (91, 93, 94, 116, 119). With the human MSY sequence in hand, it became evident that non-allelic (ectopic) homologous recombination mediated by MSY amplicons was the underlying cause of five major, recurrent, and precisely defined classes of interstitial deletions that cause or predispose to spermatogenic failure: *AZF_a*, *AZF_c* (or b2/b4), P5/distal-P1, P5/proximal-P1, and gr/gr deletions (61, 95-97, 113). Most of these deletions remove some or all copies of several multi-copy, testis-specific gene families. Knowledge of the repeat structure of the MSY also enabled the design of high-

throughput assays to investigate the frequency of deletions within the *AZFc* region in the general population. In >20,000 men unselected for spermatogenesis phenotype, 775 men were found to have a deletion involving the *AZFc* region, yielding an estimated prevalence of 1 in 27 (101). Much ongoing work in this field focuses on possible associations between spermatogenesis phenotypes and copy number variation in specific testis-expressed gene families (35, 36, 74, 75, 84).

Sequence of the mouse MSY: Fallout of X vs. Y battle

The mouse MSY's critical role in sperm production has been studied intensely in recent decades, aided by the identification of spontaneously arising MSY deletions and rearrangements in laboratory colonies. These studies have uncovered connections between specific gene deficits and distinct spermatogenesis defects, providing insights into gene function (8).

The short arm of the mouse MSY was shown to contain genes required early in spermatogenesis. The *Sxr* (Sex-reversed) strain, first identified in 1971 (12), carries most of the MSY short-arm, including *Sry* and 10 additional MSY genes, fused to an X chromosome (78). Mice carrying this X-Y fusion develop as males, with testes and germ (spermatogenic) cells. However, sperm produced by *Sxr* mice exhibit morphological defects. An *Sxr*-deletion variant was later discovered that retains *Sry* but is missing six of the additional genes (79). These *Sxr*-deleted mice have a more severe spermatogenesis phenotype compared to the original *Sxr* mice, with defects in spermatogonial proliferation, and meiotic arrest (114). Through transgene analyses, a single gene within the *Sxr*-deletion interval -- *Eif2s3y* -- was shown to be required for normal spermatogonial proliferation (77), and a more recent study indicated that *Eif2s3y* and *Sry*

are the only two Y-linked genes required for spermatogenesis to progress through the first meiotic division (130).

By contrast, genes on the long arm of the mouse MSY (Yq) appear to be required for the later stages of spermatogenesis, after meiosis has been completed. Yq-deleted mice have abnormal sperm morphology and impairment of sperm function (111, 112); the size of the deletion correlates with the severity of the defects (20, 117, 125). Yq deletions are also associated with a skewed sex ratio in offspring in favor of females (20).

The sequence of the mouse MSY, the largest and most ampliconic mammalian MSY known, was reported in 2014 (108), and the sequencing data have and continue to be instrumental to understanding the mouse MSY's striking biology. The euchromatic sequence of the mouse MSY is roughly 90 Mb in size, which is larger than the euchromatic portions of all three sequenced primate MSYs (human, chimpanzee, and rhesus macaque) combined. Less than 2% of the mouse MSY is demonstrably homologous to the primate MSYs or the ancestral autosomes from which the mammalian X and Y are derived. This ancestral sequence is confined to the short arm of the mouse MSY. The long arm of the MSY is comprised of mouse-specific ampliconic sequence (**Figure 4**). The basic element of the ampliconic sequence is a 0.5-Mb repeat unit, present in ~180 copies, and in aggregate, these repeats account for an impressive 3% of the mouse genome. Each repeat unit contains members of four distinct gene families that are relatively recent additions to the mouse Y chromosome. Consequently, the mouse MSY is gene-rich, with a total of ~700 genes compared to 78, 37, and 31 in human, chimpanzee, and rhesus macaque, respectively.

The amplified mouse Y gene families have non-allelic X homologs, which were also acquired and amplified in the rodent lineage (81, 98, 108). Both the X-linked and Y-linked gene

families are expressed specifically in testicular germ cells (108), raising the possibility that X-versus-Y antagonism in the male germ line may have driven the acquisition and massive amplification of these genes on the sex chromosomes. Gene knock-downs of the X- and Y-linked homologs of one multi-copy gene family – *Sly/Slx* – have been shown to distort the sex ratio in opposite directions. Sex ratio favors females when *Sly* expression is reduced, but it favors males when *Slx* expression is reduced, providing support for sex-chromosome driven intragenomic conflict (18, 19).

It is not known if parallel amplification of male germ-cell-expressed gene families on X and Y chromosomes is a peculiarity of the mouse lineage or if this phenomenon might be a widespread feature of mammalian sex chromosome evolution. One other notable example of X-Y co-amplification, albeit less dramatic, can be found in human. The male germ-cell-expressed genes *VCX* and *VCY* were acquired and amplified on the sex chromosomes of the human-chimpanzee ancestor (62). The functions of these gene families and the consequences, if any, of their disruption in human are unknown.

BIOLOGICAL THEME #3: BEYOND THE REPRODUCTIVE TRACT

The MSY's roles in testis determination and sperm production are firmly established, and the genes responsible for these phenotypes have been the focus of intense study. Far less is known about the functions of the remaining single-copy, broadly expressed genes in the MSY. These genes were once considered “lucky” survivors of the degeneration process that wiped out >600 genes on the MSY during the process of X-Y evolution and differentiation that spanned hundreds of millions of years. However, some intriguing studies point to the biological and medical importance of these single-copy genes.

MSY genes and Turner Syndrome

Turner syndrome occurs in about 1 in 2000 live female births, making it one of the most frequent chromosomal abnormalities in females. Turner phenotypes include short stature, ovarian insufficiency (oocyte depletion), congenital lymphedema, cardiovascular malformations, and specific neurocognitive deficits (40). Affected individuals are missing all or part of one sex chromosome, the most widely recognized form being monosomy X (45,X; sometimes referred to as “XO”). It is estimated that >99% of 45,X conceptuses die *in utero* (17), and that most individuals with Turner syndrome are mosaic for all or part of a second sex chromosome (44-46).

Even taking mosaicism into account, it is of note that Turner syndrome is the only human monosomy that even occasionally survives to birth, and this offers valuable clues regarding the etiology of Turner syndrome (134). The unique tolerance for monosomy of the X chromosome, compared to any of the autosomes, stems from the fact that one X chromosome is largely transcriptionally silenced, or inactivated, in the cells of 46,XX females. This mechanism of X-inactivation evolved to compensate for different dosages of X-linked genes in males and females. Why then is partial or complete sex-chromosome monosomy associated with any phenotype at all, if just one X chromosome is normally active in the cells of 46,XX females? Part of the answer lies in the finding that X-chromosome inactivation is far from complete (25). At least 15% of genes on the “inactive” X chromosome are actually transcribed in human (11). These genes that “escape” X inactivation may be required in two active copies for normal development, and the absence of the second copy could contribute to the phenotypic features of Turner syndrome. However, these arguments alone do not explain why 46,XY males do not exhibit Turner features. The most important group of candidate “Turner genes” on the X chromosome may be those that

also have single-copy, broadly expressed homologs on the Y chromosome (27, 55). These Y-linked genes are generally well conserved along with their X-linked homologs, consistent with their having shared functions. The pseudoautosomal regions (PARs) of the X and Y chromosomes may also contain candidate Turner genes, as 45,X individuals retain only one copy of each PAR. Indeed, the absence of a second copy of the pseudoautosomal gene *SHOX* contributes to the short stature seen in girls and women with Turner syndrome (92), but no other pseudoautosomal genes have been implicated thus far.

Further evidence of the involvement of Y-linked genes in Turner syndrome comes from studies of patients with X-Y translocations generated by ectopic X-Y crossovers (**Figure 2**). Several XY females arising from such translocations were reported to have certain phenotypic features of Turner syndrome, with lymphedema being the most commonly observed feature (6, 10, 26, 71). Notably, these XY Turner females retain two intact copies of the PAR. Therefore, the portion of the MSY that is missing seems likely to contain one or more genes whose deficiency contributes to the Turner phenotype. Molecular mapping studies narrowed down the Y-linked interval associated with Turner phenotypes in XY sex reversal cases to a 90-kb region that contains a single gene – *RPS4Y* (30) – which is a strong candidate because its X-linked homolog, *RPS4X*, escapes X-inactivation, and *RPS4Y* and *RPS4X* have been shown to be functionally interchangeable *in vitro* (126).

The case for the involvement of MSY genes in Turner phenotypes is a strong one, and other disorders involving sex chromosome anomalies provide additional clues regarding non-reproductive phenotypes associated with the human Y chromosome. Although the evidence is far from definitive, phenotypes as diverse as height, tooth size, and brain development have all been linked to the Y chromosome though studies involving individuals with sex chromosome

aneuploidies (XXY [Klinefelter's] and XYY males) and individuals with sex reversal (XY females and XX males) (2, 3, 7, 22, 80, 118).

Palindromes, isodicentric Y chromosomes, and Turner syndrome

One property of Turner syndrome that distinguishes it from other human aneuploidies is that most cases are not due to chromosome segregation errors during maternal meiosis. Instead, in about 75% of 45,X individuals, the single X chromosome is of maternal origin, implying that a paternally derived sex chromosome has gone missing (52). The human MSY's rearrangement-prone architecture may be responsible in part for this phenomenon (66). The large palindromes of the MSY undergo frequent gene conversion, a manifestation of intra-chromatid exchange, which maintains the high-degree of identity between palindrome arms. The MSY palindromes can also engage in inter-chromatid exchange, sometimes resulting in the generation of isodicentric Y chromosomes (**Figure 5**). The most common such rearrangement results in a symmetrical chromosome with two short arms (and, therefore, two copies of *SRY*), two centromeres, and two truncated long arms, fused together. This configuration is referred to as an idicYp chromosome.

The clinical consequences of idicYp formation can be diverse (66). One such consequence is spermatogenic failure: most idicYp chromosomes are missing a number of the Y chromosome's ampliconic genes, including those within the critical *AZFc* region. Some of these individuals are missing several single-copy genes as well, but phenotypes beyond the reproductive tract have not been well documented in these cases. A second, and less intuitive, consequence of idicYp formation is sex reversal and Turner syndrome. The presence of two centromeric regions makes isodicYp chromosomes mitotically unstable, and they are frequently

lost during cell division, leading to mosaicism for 45,X and 46,XidicYp cells. If a large proportion of a fetus' cells (or a particularly critical subset of fetal cells) are 45,X, the fetus will develop as an anatomic female (because the *SRY* gene was not present in the fetal gonad to trigger testis differentiation) with Turner phenotypic features (because the second sex chromosome was missing in certain somatic tissues during development). Therefore, what appears to be a paternal transmission error in Turner syndrome (41, 69) may actually be due, in some cases, to transmission of an isodicYp and its subsequent loss during embryonic or fetal development.

Evolutionary analysis reveals that Y-chromosome genes are a selected and specialized set

Clues regarding the functions of Y-chromosome genes beyond the reproductive tract have emerged from long-term studies of patients with sex chromosome anomalies. A second, complementary approach involves looking back tens or hundreds of millions of years in sex chromosome evolution. A comprehensive comparative genomic analysis of the Y chromosomes of multiple mammalian species has recently yielded novel insights into the extraordinary evolutionary longevity and fundamental biological functions of Y-linked genes (4).

Because of rampant genetic decay in the absence of recombination, the present-day human Y chromosome has lost all but 17 of the ~640 genes that it once shared with the X chromosome. Three of these ancestral Y-linked genes – *RBMY*, *TSPY*, and *HSFY* – ensured their survival by becoming ampliconic: they are now multicopy, are expressed exclusively in the testis, and likely play critical roles in spermatogenesis (107). From comparisons of three primate MSYs (human, chimpanzee, and rhesus macaque), it became evident that the remaining single-copy ancestral genes have been preserved quite effectively in the human lineage by natural selection (49-51). An expanded analysis including Y chromosome sequences from five additional

mammals – marmoset, mouse, rat, bull, and opossum – greatly increased our ability to reconstruct the evolutionary past – to look back at least 180 million years. This eight-species analysis not only strengthened previous findings that the MSY’s single-copy ancestral genes have impressive staying power, but also revealed that they form a functionally coherent group enriched for genes that are dosage-sensitive, are broadly expressed across the body, and encode general regulators of chromatin modification, transcription, translation, and protein stability (**Figure 6**) (4).

Multiple lines of evidence argue that the MSY’s single-copy ancestral genes function beyond the reproductive tract (4). First, most of these genes are broadly expressed across adult tissues and across developmental time (**Figure 6**). Second, the X homologs of the MSY’s broadly expressed genes appear to be highly dosage sensitive, in a manner that is consistent with their having vital roles in both sexes: the X homologs of all MSY broadly expressed genes escape X-inactivation, and they have a higher probability of being haploinsufficient (requiring a second copy on either the X or Y chromosome for normal organismal development or function) compared to other ancestral X-linked genes (**Figure 6**). X-linked intellectual disability syndromes in human provide further evidence for dosage sensitivity of these genes; heterozygous mutations in *UTX*, *KDM5C*, and *NLGN4X* – genes that have broadly expressed MSY homologs and escape X-inactivation – cause intellectual disabilities of varying degrees, implying haploinsufficiency (4, 67, 72, 102).

Across the genome, haploinsufficient genes have been reported to be enriched for regulatory proteins and transcription factors (21), and this same bias is observed in the X homologs of broadly expressed MSY genes. Functional annotations for many of these genes suggest regulatory functions: nuclear organization, DNA binding, and RNA binding (**Figure 6**). For example, the X homologs of *UTY* and *KDM5D* are histone lysine demethylases (65, 115), the

X homolog of *ZFY* is a transcriptional regulator of stem-cell self renewal (34), and the X homologs of *EIF1AY* and *DDX3Y* are translation initiation factors (24, 68). *DDX3X* also appears to be a component of the spliceosome (133). The haploinsufficiency model predicts that the broadly expressed Y-linked genes share functions with their X-linked counterparts. Functional interchangeability of X and Y homologs has been demonstrated experimentally for a handful of genes: *RPS4X/Y* and *DDX3X/Y* are functionally interchangeable *in vitro* (103, 126), and *Utx/y* are functionally redundant during mouse embryonic development (70, 104, 128).

Taken together, these medical and evolutionary analyses indicate that the MSY's broadly expressed genes have important biological functions throughout the body. Y-chromosome deletions and rearrangements that perturb these genes, or their copy number, will continue to be helpful in revealing the roles of these genes in development and physiology both within and beyond the reproductive tract.

LOOKING FORWARD: THE Y CHROMOSOME AND SEXUAL DIMORPHISM

A distinct and largely unexplored hypothesis regarding the biological function of MSY genes is that the presence or absence of the Y chromosome directly influences an individual's disease risks. A myriad of human diseases exhibit striking sex biases in susceptibility, incidence, or severity, as highlighted in the Institute of Medicine's report probing the pervasive biological influence of sex (129). To name just a few examples: lupus (6:1), rheumatoid arthritis (3:1), and unipolar depression (2:1) are strongly female biased, while autism (5:1), dilated cardiomyopathy (3:1), and ankylosing spondylitis (5:1) are strongly male biased. Historically, such biases have been attributed solely to hormonal, or extrinsic, factors. However, it is possible that intrinsic biochemical differences between XX and XY cells have biological consequences throughout the

body, and thus contribute to sexual dimorphism in human health and disease. A spate of recent studies examining large populations have implicated the human Y chromosome in numerous somatic diseases, including cancer (9, 28, 33), coronary artery disease (14), autism (69), and primary biliary cirrhosis (73). However, the genetic architecture of these diseases is complex, and the Y-linked gene or genes contributing to these disease phenotypes have not been identified.

As discussed above, evolutionary analyses of ancestral regions of mammalian Y chromosomes yielded strong evidence that X-Y homologous genes serve ubiquitous gene regulatory functions (4). Therefore, slight differences in the biochemical properties of protein isoforms encoded by X-Y gene pairs – or differences in the levels or patterns of expression of the X- and Y-linked genes – could have significant phenotypic consequences. For example, two human X-Y gene pairs encode histone demethylases. If the X- and Y-encoded isoforms have different expression levels, substrate specificities and/or enzyme affinities, this could contribute to genome-wide transcriptional regulatory differences between males and females. A complete understanding of the influence of sex chromosome complement in male and female health and disease will create a better toolkit for scientists, clinicians, and developers of novel therapies.

Methods to explore this intriguing hypothesis will likely involve analysis of large-scale molecular datasets from a variety of mammals, together with the generation of appropriate animal models, which is not a straightforward task when it comes to the Y chromosome. Access to the complete mouse Y chromosome sequence and recent advances in genome-editing technologies have finally opened the door to directed mutagenesis of Y-linked genes in mice (108, 121). The mouse, however, may ultimately prove to be of limited use as a model for the human sex chromosomes, because of the rapid evolution of the Y chromosome. Of the human MSY's 12 broadly expressed, single-copy genes, only five have orthologs in the mouse MSY (**Figure 6**).

By contrast, the rhesus macaque MSY carries an ortholog of each and every ancestral gene in the human MSY (49), and progress towards gene targeting in such nonhuman-primate models (13) may soon enable *in vivo* study of the function of all human MSY genes. New insights into MSY gene function can also guide systematic investigations of human organismal phenotypes associated with MSY gene deficits. Finally, thorough characterization of sex differences at the molecular and cellular level will provide a strong foundation for understanding the gene regulatory functions of X and Y homologs. Using this multifaceted approach, the field may come to understand the contribution of Y-linked genes to human biology and to medically consequential differences between the sexes at the molecular, cellular, and physiological levels. Ultimately, these insights will lead to a greater appreciation of the etiology of Turner syndrome and the underlying causes of sex biases in disease, enabling both medical research and practice to embrace and address the most fundamental polymorphism in our species – whether an individual has two X chromosomes, or one X plus one Y chromosome.

SUMMARY POINTS

1. Three biological themes have defined Y-chromosome research in the age of molecular genetics and genomics: testis determination, spermatogenesis, and biological roles beyond the reproductive tract.
2. The Y-chromosome's roles in testis determination and spermatogenesis are well defined and widely appreciated, especially in human and mouse, while the Y chromosome's roles beyond the reproductive tract are only beginning to be studied.
3. Complete MSY sequences from human, chimpanzee, rhesus macaque, and mouse have revealed that mammalian Y chromosomes vary widely in gene content, size, and structure.

4. The massive amplification of male germ-cell specific gene families on the mouse MSY was accompanied by parallel amplification of homologous gene families on the X chromosome, providing insight into molecular mechanisms of sex-linked meiotic drive in mammals.
5. Studies of patients with sex-chromosome anomalies combined with comprehensive comparative genomic analysis of multiple mammalian MSYs indicates that many MSY genes have fundamental biological roles throughout the body.
6. The future of Y-chromosome research lies in understanding how differences in sex chromosome complement (XX versus XY), and how the presence or absence of individual Y-chromosome genes, influence sex differences in health and disease, both within and beyond the reproductive tract.

FUTURE ISSUES

1. The X-Y co-amplification observed in mouse may be widespread among mammals. In order to explore this possibility, SHIMS sequencing of additional mammalian X and Y chromosomes will be required. Complete and accurate X- and Y-chromosome sequences are essential for unraveling the ampliconic repeats that are intrinsic to this phenomenon. Functional studies in human may also be possible, taking advantage of predictable and recurrent structural variants on the X and Y chromosomes.
2. Exploring the contribution of Y-linked genes to health and disease beyond the reproductive tract will require a multifaceted approach involving targeting Y-linked genes in nonhuman primates, systematic clinical investigations of patients with sex chromosome

anomalies, and large-scale molecular studies of male and female cells across tissues and species.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGEMENTS

We thank Daniel Bellott, Bluma Lesch, and Helen Skaletsky for careful reading of the manuscript and helpful discussions. Research in DCP's laboratory is funded by the National Institutes of Health, the Howard Hughes Medical Institute, the Whitehead Institute, and Biogen.

LITERATURE CITED

1. Aitken RJ, Marshall Graves JA. 2002. The future of sex. *Nature* 415: 963
2. Alvesalo L, Portin P. 1980. 47,XXY males: sex chromosomes and tooth size. *Am J Hum Genet* 32: 955-9
3. Alvesalo L, Varrela J. 1980. Permanent tooth sizes in 46,XY females. *Am J Hum Genet* 32: 736-42
4. Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, et al. 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* 508: 494-9
5. Berta P, Hawkins JR, Sinclair AH, Taylor A, Griffiths BL, Goodfellow PN. 1990. Genetic evidence equating *SRY* and the testis-determining factor. *Nature* 348: 448-50
6. Blagowidow N, Page DC, Huff D, Mennuti MT. 1989. Ullrich-Turner syndrome in an XY female fetus with deletion of the sex-determining portion of the Y chromosome. *Am J Med Genet* 34: 159-62
7. Brown WM. 1968. Males with an XYY sex chromosome complement. *J Med Genet* 5: 341-59
8. Burgoyne PS. 1998. The role of Y-encoded genes in mammalian spermatogenesis. *Semin. Cell. Dev. Biol.* 9: 423-32

9. Cannon-Albright LA, Farnham JM, Bailey M, Albright FS, Teerlink CC, et al. 2014. Identification of specific Y chromosomes associated with increased prostate cancer risk. *Prostate* 74: 991-8
10. Cantrell MA, Bicknell JN, Pagon RA, Page DC, Walker DC, et al. 1989. Molecular analysis of 46,XY females and regional assignment of a new Y-chromosome-specific probe. *Hum Genet* 83: 88-92
11. Carrel L, Willard HF. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 434: 400-4
12. Cattanach BM, Pollard CE, Hawker SG. 1971. Sex-reversed mice: XX and XO males. *Cytogenetics* 10: 318-37
13. Chan AW. 2013. Progress and prospects for genetic modification of nonhuman primate models in biomedical research. *ILAR J* 54: 211-23
14. Charchar FJ, Bloomer LD, Barnes TA, Cowley MJ, Nelson CP, et al. 2012. Inheritance of coronary artery disease in men: an analysis of the role of the Y chromosome. *Lancet* 379: 915-22
15. Charlesworth B, Charlesworth D. 2000. The degeneration of Y chromosomes. *Philos Trans R Soc Lond B Biol Sci* 355: 1563-72
16. Chumakov I, Rigault P, Guillou S, Ougen P, Billaut A, et al. 1992. Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* 359: 389
17. Cockwell A, MacKenzie M, Youings S, Jacobs P. 1991. A cytogenetic and molecular study of a series of 45,X fetuses and their parents. *J Med Genet* 28: 151-5
18. Cocquet J, Ellis PJ, Mahadevaiah SK, Affara NA, Vaiman D, Burgoyne PS. 2012. A genetic basis for a postmeiotic X versus Y chromosome intragenomic conflict in the mouse. *PLoS Genet* 8: e1002900
19. Cocquet J, Ellis PJ, Yamauchi Y, Mahadevaiah SK, Affara NA, et al. 2009. The multicopy gene Sly represses the sex chromosomes in the male mouse germline after meiosis. *PLoS Biol* 7: e1000244
20. Conway SJ, Mahadevaiah SK, Darling SM, Capel B, Rattigan AM, Burgoyne PS. 1994. *Y353/B*: a candidate multiple-copy spermiogenesis gene on the mouse Y chromosome. *Mammalian Genome* 5: 203-10
21. Dang VT, Kassahn KS, Marcos AE, Ragan MA. 2008. Identification of human haploinsufficient genes and their genomic proximity to segmental duplications. *Eur J Hum Genet* 16: 1350-7
22. de la Chapelle A. 1972. Analytic review: nature and origin of males with XX sex chromosomes. *Am J Hum Genet* 24: 71-105

23. de la Chapelle A, Tippet PA, Wetterstrand G, Page D. 1984. Genetic evidence of X-Y interchange in a human XX male. *Nature* 307: 170-1
24. Dever TE, Wei CL, Benkowski LA, Browning K, Merrick WC, Hershey JW. 1994. Determination of the amino acid sequence of rabbit, human, and wheat germ protein synthesis factor eIF-4C by cloning and chemical sequencing. *J Biol Chem* 269: 3212-8
25. Disteche CM. 2012. Dosage compensation of the sex chromosomes. *Annu Rev Genet* 46: 537-60
26. Disteche CM, Casanova M, Saal H, Friedman C, Sybert V, et al. 1986. Small deletions of the short arm of the Y chromosome in 46,XY females. *Proc Natl Acad Sci U S A* 83: 7841-4
27. Disteche CM, Filippova GN, Tsuchiya KD. 2002. Escape from X inactivation. *Cytogenet Genome Res* 99: 36-43
28. Dumanski JP, Rasi C, Lonn M, Davies H, Ingelsson M, et al. 2015. Mutagenesis. Smoking is associated with mosaic loss of chromosome Y. *Science* 347: 81-3
29. Ferguson-Smith MA. 1965. Karyotype-phenotype correlations in gonadal dysgenesis and their bearing on the pathogenesis of malformations. *Journal of Medical Genetics* 2: 142-55
30. Fisher EM, Beer-Romero P, Brown LG, Ridley A, McNeil JA, et al. 1990. Homologous ribosomal protein genes on the human X and Y chromosomes: Escape from X inactivation and possible implications for Turner syndrome. *Cell* 63: 1205-18
31. Foote S, Vollrath D, Hilton A, Page DC. 1992. The human Y chromosome: Overlapping DNA clones spanning the euchromatic region. *Science* 258: 60-6
32. Ford CE, Jones KW, Polani PE, De Almeida JC, Briggs JH. 1959. A sex-chromosome anomaly in a case of gonadal dysgenesis (Turner's syndrome). *Lancet* 1: 711-3
33. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, et al. 2014. Mosaic loss of chromosome Y in peripheral blood is associated with shorter survival and higher risk of cancer. *Nat Genet* 46: 624-8
34. Galan-Caridad JM, Harel S, Arenzana TL, Hou ZE, Doetsch FK, et al. 2007. Zfx controls the self-renewal of embryonic and hematopoietic stem cells. *Cell* 129: 345-57
35. Ghorbel M, Baklouti-Gargouri S, Keskes R, Chakroun N, Sellami A, et al. 2014. Deletion of CDY1b copy of Y chromosome CDY1 gene is a risk factor of male infertility in Tunisian men. *Gene* 548: 251-5
36. Giachini C, Nuti F, Turner DJ, Laface I, Xue Y, et al. 2009. TSPY1 Copy Number Variation Influences Spermatogenesis and Shows Differences among Y Lineages. *Journal of Clinical Endocrinology and Metabolism* 94: 4016-22
37. Girardi SK, Mielnik A, Schlegel PN. 1997. Submicroscopic deletions in the Y chromosome of infertile men. *Hum. Reprod.* 12: 1635-41

38. Graves JA. 2004. The degenerate Y chromosome--can conversion save it? *Reprod Fertil Dev* 16: 527-34
39. Graves JA. 2006. Sex chromosome specialization and degeneration in mammals. *Cell* 124: 901-14
40. Gravholt CH. 2004. Epidemiological, endocrine and metabolic features in Turner syndrome. *Eur J Endocrinol* 151: 657-87
41. Gropman A, Samango-Sprouse CA. 2013. Neurocognitive variance and neurological underpinnings of the X and Y chromosomal variations. *Am J Med Genet C Semin Med Genet* 163C: 35-43
42. Gubbay J, Collignon J, Koopman P, Capel B, Economou A, et al. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* 346: 245-50
43. Guellaen G, Casanova M, Bishop C, Geldwerth D, Andre G, et al. 1984. Human XX males with Y single-copy DNA fragments. *Nature* 307: 172-3
44. Hassold T, Benham F, Leppert M. 1988. Cytogenetic and molecular analysis of sex-chromosome monosomy. *Am J Hum Genet* 42: 534-41
45. Hook EB, Warburton D. 1983. The distribution of chromosomal genotypes associated with Turner's syndrome: livebirth prevalence rates and evidence for diminished fetal mortality and severity in genotypes associated with structural X abnormalities or mosaicism. *Human genetics* 64: 24-7
46. Hook EB, Warburton D. 2014. Turner syndrome revisited: review of new data supports the hypothesis that all viable 45,X cases are cryptic mosaics with a rescue cell line, implying an origin by mitotic loss. *Hum Genet*
47. Hughes IA, Davies JD, Bunch TI, Pasterski V, Mastroyannopoulou K, MacDougall J. 2012. Androgen insensitivity syndrome. *Lancet* 380: 1419-28
48. Hughes JF, Rozen S. 2012. Genomics and genetics of human and primate y chromosomes. *Annu Rev Genomics Hum Genet* 13: 83-108
49. Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Graves T, et al. 2012. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* 483: 82-6
50. Hughes JF, Skaletsky H, Pyntikova T, Graves TA, van Daalen SK, et al. 2010. Chimpanzee and human Y chromosomes are remarkably divergent in structure and gene content. *Nature* 463: 536-9
51. Hughes JF, Skaletsky H, Pyntikova T, Minx PJ, Graves T, et al. 2005. Conservation of Y-linked genes during human evolution revealed by comparative sequencing in chimpanzee. *Nature* 437: 100-3

52. Jacobs P, Dalton P, James R, Mosse K, Power M, et al. 1997. Turner syndrome: a cytogenetic and molecular study. *Ann. Hum. Genet.* 61: 471-83
53. Jacobs PA, Strong JA. 1959. A case of human intersexuality having a possible XXY sex determining mechanism. *Nature* 183: 302-3
54. Jäger RJ, Anvret M, Hall K, Scherer G. 1990. A human XY female with a frame shift mutation in the candidate testis-determining gene *SRY*. *Nature* 348: 452-4
55. Jegalian K, Page DC. 1998. A proposed path by which genes common to mammalian X and Y chromosomes evolve to become X inactivated. *Nature* 394: 776-80
56. Jones MH, Khwaja OS, Briggs H, Lambson B, Davey PM, et al. 1994. A set of ninety-seven overlapping yeast artificial chromosome clones spanning the human Y chromosome euchromatin. *Genomics* 24: 266-75
57. Kobayashi K, Mizuno K, Hida A, Komaki R, Tomita K, et al. 1994. PCR analysis of the Y chromosome long arm in azoospermic patients: evidence for a second locus required for spermatogenesis. *Human Molecular Genetics* 3: 1965-7
58. Koopman P, Gubbay J, Collignon J, Lovell-Badge R. 1989. *Zfy* gene expression patterns are not compatible with a primary role in mouse sex determination. *Nature* 342: 940-2
59. Koopman P, Gubbay J, Vivian N, Goodfellow P, Lovell-Badge R. 1991. Male development of chromosomally female mice transgenic for *Sry*. *Nature* 351: 117-21
60. Koopman P, Münsterberg A, Capel B, Vivian N, Lovell-Badge R. 1990. Expression of a candidate sex-determining gene during mouse testis differentiation. *Nature* 348: 450-2
61. Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, Cordum HS, et al. 2001. The *AZFc* region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. *Nat Genet* 29: 279-86
62. Lahn B, Page D. 2000. A human sex-chromosomal gene family expressed in male germ cells and encoding variably charged proteins. *Human Molecular Genetics* 9: 311-9
63. Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. *Science* 286: 964-7
64. Lahn BT, Page DC. 1997. Functional coherence of the human Y chromosome. *Science* 278: 675-80
65. Lan F, Bayliss PE, Rinn JL, Whetstone JR, Wang JK, et al. 2007. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 449: 689-94
66. Lange J, Skaletsky H, van Daalen SK, Embry SL, Korver CM, et al. 2009. Isodicentric Y chromosomes and sex disorders as byproducts of homologous recombination that maintains palindromes. *Cell* 138: 855-69

67. Lawson-Yuen A, Saldivar JS, Sommer S, Picker J. 2008. Familial deletion within NLGN4 associated with autism and Tourette syndrome. *Eur J Hum Genet* 16: 614-8
68. Lee CS, Dias AP, Jedrychowski M, Patel AH, Hsu JL, Reed R. 2008. Human DDX3 functions in translation and interacts with the translation initiation factor eIF3. *Nucleic Acids Res* 36: 4708-18
69. Lee NR, Wallace GL, Adeyemi EI, Lopez KC, Blumenthal JD, et al. 2012. Dosage effects of X and Y chromosomes on language and social functioning in children with supernumerary sex chromosome aneuploidies: implications for idiopathic language impairment and autism spectrum disorders. *J Child Psychol Psychiatry* 53: 1072-81
70. Lee S, Lee JW, Lee SK. 2012. UTX, a histone H3-lysine 27 demethylase, acts as a critical switch to activate the cardiac developmental program. *Dev Cell* 22: 25-37
71. Levilliers J, Quack B, Weissenbach J, Petit C. 1989. Exchange of terminal portions of X- and Y-chromosomal short arms in human XY females. *Proc Natl Acad Sci U S A* 86: 2296-300
72. Lindgren AM, Hoyos T, Talkowski ME, Hanscom C, Blumenthal I, et al. 2013. Haploinsufficiency of KDM6A is associated with severe psychomotor retardation, global growth restriction, seizures and cleft palate. *Hum Genet* 132: 537-52
73. Lleo A, Oertelt-Prigione S, Bianchi I, Caliarì L, Finelli P, et al. 2013. Y chromosome loss in male patients with primary biliary cirrhosis. *J Autoimmun* 41: 87-91
74. Lu C, Jiang J, Zhang R, Wang Y, Xu M, et al. 2014. Gene copy number alterations in the azoospermia-associated AZFc region and their effect on spermatogenic impairment. *Mol Hum Reprod* 20: 836-43
75. Lu C, Wang Y, Zhang F, Lu F, Xu M, et al. 2013. DAZ duplications confer the predisposition of Y chromosome haplogroup K* to non-obstructive azoospermia in Han Chinese populations. *Hum Reprod* 28: 2440-9
76. Magenis RE, Tochen ML, Holahan KP, Carey T, Allen L, Brown MG. 1984. Turner syndrome resulting from partial deletion of Y chromosome short arm: localization of male determinants. *J Pediatr* 105: 916-9
77. Mazeyrat S, Saut N, Grigoriev V, Mahadevaiah SK, Ojarikre OA, et al. 2001. A Y-encoded subunit of the translation initiation factor Eif2 is essential for mouse spermatogenesis. *Nat Genet* 29: 49-53
78. Mazeyrat S, Saut N, Sargent CA, Grimmond S, Longepied G, et al. 1998. The mouse Y chromosome interval necessary for spermatogonial proliferation is gene dense with syntenic homology to the human AZFa region. *Human Molecular Genetics* 7: 1713-124

79. McLaren A, Simpson E, Epplen JT, Studer R, Koopman P, et al. 1988. Location of the genes controlling H-Y antigen expression and testis determination on the mouse Y chromosome. *Proc Natl Acad Sci U S A* 85: 6442-5
80. Mittwoch U. 1992. Sex determination and sex reversal: Genotype, phenotype, dogma, and semantics. *Human Genetics* 89: 467-79
81. Mueller JL, Mahadevaiah SK, Park PJ, Warburton PE, Page DC, Turner JM. 2008. The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. *Nat. Genet.* 40: 794-9
82. Muller HJ. 1964. The relation of recombination to mutational advance. *Mutation Research* 106: 2-9
83. Nakahori Y, Kuroki Y, Komaki R, Kondoh N, Namiki M, et al. 1996. The Y chromosome region essential for spermatogenesis. *Hormone Research* 46 (Suppl. 1): 20-3
84. Noordam MJ, Westerveld GH, Hovingh SE, van Daalen SK, Korver CM, et al. 2011. Gene copy number reduction in the azoospermia factor c (AZFc) region and its effect on total motile sperm count. *Human molecular genetics* 20: 2457-63
85. Ohno S. 1967. *Sex Chromosomes and Sex-linked Genes*. Berlin: Springer-Verlag
86. Page DC. 1988. Is ZFY the sex-determining gene on the human Y chromosome? *Philosophical Transactions of the Royal Society of London* 322: 155-7
87. Page DC, Mosher R, Simpson EM, Fisher EM, Mardon G, et al. 1987. The sex-determining region of the human Y chromosome encodes a finger protein. *Cell* 51: 1091-104
88. Palmer MS, Sinclair AH, Berta P, Ellis NA, Goodfellow PN, et al. 1989. Genetic evidence that ZFY is not the testis-determining factor. *Nature* 342: 937-9
89. Pearks Wilkerson AJ, Raudsepp T, Graves T, Albracht D, Warren W, et al. 2008. Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. *Genomics* 92: 329-38
90. Quinn A, Koopman P. 2012. The molecular genetics of sex determination and sex reversal in mammals. *Semin Reprod Med* 30: 351-63
91. Qureshi SJ, Ross AR, Ma K, Cooke HJ, Intyre MA, et al. 1996. Polymerase chain reaction screening for Y chromosome microdeletions: a first step towards the diagnosis of genetically-determined spermatogenic failure in men. *Mol Hum Reprod* 2: 775-9
92. Rao E, Weiss B, Fukami M, Rump A, Niesler B, et al. 1997. Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nat Genet* 16: 54-63

93. Reijo R, Alagappan RK, Patrizio P, Page DC. 1996. Severe oligozoospermia resulting from deletions of azoospermia factor gene on Y chromosome. *Lancet* 347: 1290-3
94. Reijo R, Lee TY, Salo P, Alagappan R, Brown LG, et al. 1995. Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nat Genet* 10: 383-93
95. Repping S, Skaletsky H, Brown L, van Daalen SKM, Korver CM, et al. 2003. Polymorphism for a 1.6-Mb deletion of the human Y chromosome persists through balance between recurrent mutation and haploid selection. *Nature Genetics* 35: 247-51
96. Repping S, Skaletsky H, Lange J, Silber S, van der Veen F, et al. 2002. Recombination between palindromes P5 and P1 on the human Y chromosome causes massive deletions and spermatogenic failure. *American Journal of Human Genetics* 71: 906-22
97. Repping S, van Daalen SK, Korver CM, Brown LG, Marszalek JD, et al. 2004. A family of human Y chromosomes has dispersed throughout northern Eurasia despite a 1.8-Mb deletion in the azoospermia factor c region. *Genomics* 83: 1046-52
98. Reynard LN, Turner JM, Cocquet J, Mahadevaiah SK, Toure A, et al. 2007. Expression analysis of the mouse multi-copy X-linked gene Xlr-related, meiosis-regulated (Xmr), reveals that Xmr encodes a spermatid-expressed cytoplasmic protein, SLX/XMR. *Biol Reprod* 77: 329-35
99. Rice WR. 1996. Evolution of the Y sex chromosome in animals. *BioScience* 46: 331-43
100. Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, et al. 2005. The DNA sequence of the human X chromosome. *Nature* 434: 325-37
101. Rozen SG, Marszalek JD, Irenze K, Skaletsky H, Brown LG, et al. 2012. AZFc deletions and spermatogenic failure: a population-based survey of 20,000 Y chromosomes. *Am J Hum Genet* 91: 890-6
102. Rujirabanjerd S, Nelson J, Tarpey PS, Hackett A, Edkins S, et al. 2010. Identification and characterization of two novel JARID1C mutations: suggestion of an emerging genotype-phenotype correlation. *Eur J Hum Genet* 18: 330-5
103. Sekiguchi T, Iida H, Fukumura J, Nishimoto T. 2004. Human DDX3Y, the Y-encoded isoform of RNA helicase DDX3, rescues a hamster temperature-sensitive ET24 mutant cell line with a DDX3X mutation. *Exp Cell Res* 300: 213-22
104. Shpargel KB, Sengoku T, Yokoyama S, Magnuson T. 2012. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. *PLoS Genet* 8: e1002964
105. Simoni M, Gromoll J, Dworniczak B, Rolf C, Abshagen K, et al. 1997. Screening for deletions of the Y chromosome involving the DAZ (Deleted in AZoospermia) gene in azoospermia and severe oligozoospermia. *Fertil Steril* 67: 542-7

106. Sinclair AH, Berta P, Palmer MS, Hawkins JR, Griffiths BL, et al. 1990. A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* 346: 240-4
107. Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* 423: 825-37
108. Soh YQS, Alfoldi J, Pyntikova T, Brown LG, Graves T, et al. 2014. Sequencing the Mouse Y Chromosome Reveals Convergent Gene Acquisition and Amplification on Both Sex Chromosomes. *Cell* 159: 800-13
109. Stern C. 1957. The problem of complete Y-linkage in men. *American Journal of Human Genetics* 9: 147-66
110. Stuppia L, Mastroprimiano G, Calabrese G, Peila R, Tenaglia R, Palka G. 1996. Microdeletions in interval 6 of the Y chromosome detected by STS-PCR in 6 of 33 patients with idiopathic oligo- or azoospermia. *Cytogenet Cell Genet* 72: 155-8
111. Styrna J, Imai HT, Moriwaki K. 1991. An increased level of sperm abnormalities in mice with a partial deletion of the Y chromosome. *Genetical Research* 57: 195-9
112. Styrna J, Klag J, Moriwaki K. 1991. Influence of partial deletion of the Y chromosome on mouse sperm phenotype. *Journal of Reproduction and Fertility* 92: 187-95
113. Sun C, Skaletsky H, Rozen S, Gromoll J, Nieschlag E, et al. 2000. Deletion of azoospermia factor a (AZFa) region of human Y chromosome caused by recombination between HERV15 proviruses. *Human Molecular Genetics* 9: 2291-6
114. Sutcliffe MJ, Burgoyne P. 1989. Analysis of the testes of H-Y negative XOSr^b mice suggests that the spermatogenesis gene (*Spy*) acts during the differentiation of the A spermatogonia. *Development* 107: 373-80
115. Tahiliani M, Mei P, Fang R, Leonor T, Rutenberg M, et al. 2007. The histone H3K4 demethylase SMCX links REST target genes to X-linked mental retardation. *Nature* 447: 601-5
116. Tiepolo L, Zuffardi O. 1976. Localization of factors controlling spermatogenesis in the nonfluorescent portion of the Y chromosome long arm. *Human Genetics* 34: 119-24
117. Toure A, Szot M, Mahadevaiah SK, Rattigan A, Ojarikre OA, Burgoyne PS. 2004. A new deletion of the mouse Y chromosome long arm associated with the loss of *Ssty* expression, abnormal sperm development and sterility. *Genetics* 166: 901-12
118. Varrela J, Alvesalo L, Vinkka H. 1984. Body size and shape in 46,XY females with complete testicular feminization. *Ann Hum Biol* 11: 291-301

119. Vogt PH, Edelmann A, Kirsch S, Henegariu O, Hirschmann P, et al. 1996. Human Y chromosome azoospermia factors (AZF) mapped to different subregions in Yq11. *Hum Mol Genet* 5: 933-43
120. Vollrath D, Foote S, Hilton A, Brown LG, Beer-Romero P, et al. 1992. The human Y chromosome: A 43-interval map based on naturally occurring deletions. *Science* 258: 52-9
121. Wang H, Hu YC, Markoulaki S, Welstead GG, Cheng AW, et al. 2013. TALEN-mediated editing of the mouse Y chromosome. *Nat Biotechnol* 31: 530-2
122. Wang I, Weil D, Levilliers J, Affara NA, de la Chapelle A, Petit C. 1995. Prevalence and molecular analysis of two hot spots for ectopic recombination leading to XX maleness. *Genomics* 28: 52-8
123. Wang Q, Xue Y, Zhang Y, Long Q, Asan, et al. 2013. Genetic basis of Y-linked hearing impairment. *Am J Hum Genet* 92: 301-6
124. Wang QJ, Lu CY, Li N, Rao SQ, Shi YB, et al. 2004. Y-linked inheritance of non-syndromic hearing impairment in a large Chinese family. *J Med Genet* 41: e80
125. Ward MA, Burgoyne PS. 2006. The effects of deletions of the mouse Y chromosome long arm on sperm function--intracytoplasmic sperm injection (ICSI)-based analysis. *Biol Reprod* 74: 652-8
126. Watanabe M, Zinn AR, Page DC, Nishimoto T. 1993. Functional equivalence of human X- and Y-encoded isoforms of ribosomal protein S4 consistent with a role in Turner syndrome. *Nat Genet* 4: 268-71
127. Weil D, Wang I, Dietrich A, Poustka A, Weissenbach J, Petit C. 1994. Highly homologous loci on the X and Y chromosomes are hot-spots for ectopic recombinations leading to XX maleness. *Nat. Genet.* 7: 414-9
128. Welstead GG, Creighton MP, Bilodeau S, Cheng AW, Markoulaki S, et al. 2012. X-linked H3K27me3 demethylase Utx is required for embryonic development in a sex-specific manner. *Proc Natl Acad Sci U S A* 109: 13004-9
129. Wizemann TM, Pardue ML. 2001. *Exploring the Biological Contributions to Human Health: Does Sex Matter?* Washington, D.C.: National Academy Press. 288 pp.
130. Yamauchi Y, Riel JM, Stoytcheva Z, Ward MA. 2014. Two Y genes can replace the entire Y chromosome for assisted reproduction in the mouse. *Science* 343: 69-72
131. Yen PH. 1998. A long-range restriction map of deletion interval 6 of the human Y chromosome: a region frequently deleted in azoospermic males. *Genomics* 54: 5-12
132. Yen PH, Tsai SP, Wenger SL, Steele MW, Mohandas TK, Shapiro LJ. 1991. X/Y translocations resulting from recombination between homologous sequences on Xp and Yq. *Proc Natl Acad Sci U S A* 88: 8944-8

133. Zhou Z, Licklider LJ, Gygi SP, Reed R. 2002. Comprehensive proteomic analysis of the human spliceosome. *Nature* 419: 182-5
134. Zinn AR, Page DC, Fisher EM. 1993. Turner syndrome: the case of the missing sex chromosome. *Trends Genet* 9: 90-3

SIDEBARS

Y-linked hearing impairment – an exceptional case [placement: “Ignorance of Y chromosomes” section]

Only once has a Mendelian trait (apart from testis determination) been definitively linked to the Y chromosome: an unusual case of hearing impairment. In 2004, Y-linked hearing loss in a large Chinese family was reported, and in 2012 the molecular underpinnings of this curious phenomenon came to light (123, 124). Analysis of the family’s nine-generation pedigree revealed that hearing loss appeared spontaneously and then displayed Y-linked dominant inheritance, appearing in all male descendants of the original affected male. A combination of Y-chromosome resequencing, read-depth mapping, and fluorescence in situ hybridization (FISH) in affected and unaffected males revealed that hearing impairment was associated with a complex Y-chromosome rearrangement that included a ~160-kilobase transposition from chromosome 1. This chromosome 1 segment is located within a genetic interval previously implicated in hearing impairment, implying that an extra copy of a dosage-sensitive gene or genes contained within this region is the causal factor in this special case of Y-linked hearing impairment.

Traditional sequencing approaches do not work for Y chromosomes [placement: “Human Y chromosome sequence” section]

All standard approaches to genome sequencing, both BAC-based and whole-genome-shotgun (WGS), fall short in ampliconic regions. Take the human reference sequence as an example: it

was originally assembled from a patchwork of BAC clones derived from 13 different individuals representing 26 haplotypes. Ampliconic sequences consist of repeats that may differ from one another by less than 1 bp in 10,000, which is an order of magnitude less than the nearly 1 bp in 1,000 difference typically observed between alleles or haplotypes. In assembling the sequence of ampliconic regions, only occasional nucleotide differences will distinguish BACs that derive from different copies of an amplicon (as opposed to those that truly overlap). If an investigator employs BACs derived from multiple haplotypes, then the relatively frequent differences between haplotypes will mask the rare differences between amplicon copies. In the multi-haplotype assembly of the human genome, amplicons were thus frequently misassembled or mistakenly abandoned as redundant.

The WGS approach, which is routinely employed in both human medical resequencing and non-human *de novo* genome sequencing, has an even more profoundly negative effect on the representation of ampliconic sequences. WGS assemblies rely solely on either capillary-based or next-generation sequencing reads. Ampliconic regions are completely inaccessible by this strategy because of the disparity between the length of WGS reads (100 to 500 bp) and the length of amplicon-repeat units (10 kb to >1 Mb). Thus, reads deriving from different repeat units will invariably be collapsed into single contigs during assembly because WGS methods do not incorporate mapping information, which is necessary to identify and disentangle large amplicons.

ACRONYMS AND DEFINITIONS LIST

MSY (Male-Specific region of the Y chromosome)

Region of the Y chromosome that does not ordinarily engage in crossing-over with the X chromosome during male meiosis; excludes pseudoautosomal regions.

***SRY* (Sex-determining Region of the Y)**

Testis-determining gene found on all therian mammalian Y chromosomes; expression of *SRY* in the supporting cell lineage of the bipotential fetal gonad causes the gonad to develop as a testis rather than an ovary.

SHIMS (Single-Haplotype Iterative Mapping and Sequencing)

Sequencing methodology used to produce ultra-high quality reference sequence of ampliconic regions of the genome.

BACs (Bacterial Artificial Chromosomes)

Clones that contain large inserts (~200-kb) of target DNA; essential sequencing template for SHIMS approach.

***AZFc* (AZoospermia Factor c)**

Region of the human Y chromosome long arm whose deletion is the most frequent known genetic cause of spermatogenic failure; consists of ampliconic repeats that provide substrates for non-allelic homologous recombination, generating recurrent *de novo* deletions and rearrangements.

Ampliconic sequence

Genomic sequence comprised of repeat units displaying >99% identity over >10-kb. The term “amplicon” refers to a single repeat unit.

idicYp (Isodicentric Yp)

Mirror-image rearrangement of Y chromosome with two short arms, two centromeres, and two truncated, fused long arms; carries two copies of *SRY* gene.

RELATED RESOURCES LIST

1. Institute of Medicine's report "Exploring the Biological Contributions to Human Health: Does Sex Matter": <http://www.iom.edu/~media/Files/Report%20Files/2003/Exploring-the-Biological-Contributions-to-Human-Health-Does-Sex-Matter/DoesSexMatter8pager.pdf>
2. NIH Office of Research on Women's Health: <http://orwh.od.nih.gov>

FIGURE LEGENDS

Figure 1. Human sex chromosomes evolved from autosomes

Schematic representation of the evolution of human sex chromosomes, derived from an autosome pair present in the mammalian-avian ancestor. The same autosome pair contributed to the present-day chicken autosomes 1 and 4. Time scale at left indicates approximate branching times. mya = millions of years ago.

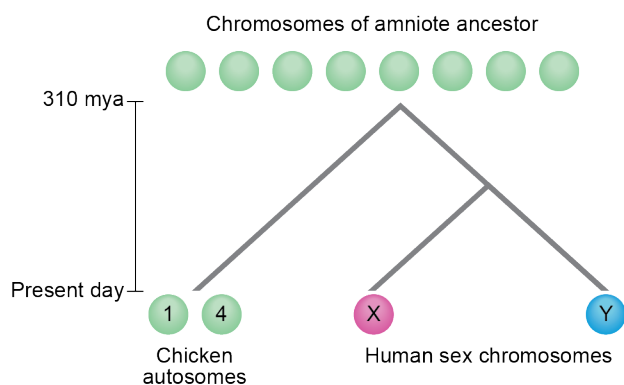


Figure 2. Ectopic X-Y recombination and the origins of human XX males and XY females

a. The short arm of the X chromosome and the entire Y chromosome are shown. Euchromatic X chromosome sequence shown in pink; all euchromatic MSY sequence classes are shown in blue for simplicity; pseudoautosomal regions, green; heterochromatin, gray. **b.** One of many possible sites of ectopic X-Y recombination is shown here. The crossover products are depicted below and the phenotypic consequences are indicated. cen = centromere, Mb = megabase.

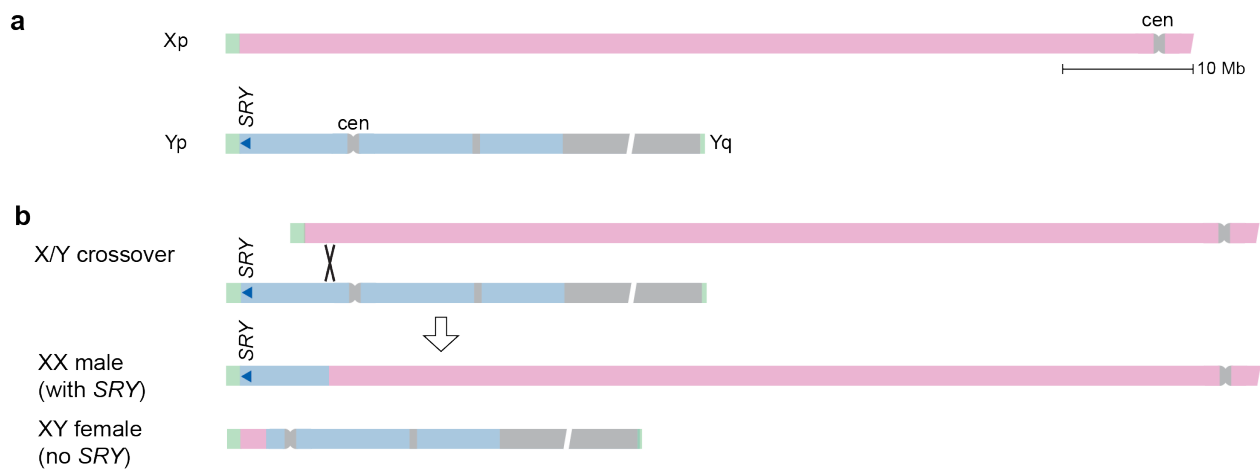


Figure 3. The human MSY and AZFc

a. Schematic representation of the entire human Y chromosome, including large heterochromatic region on Yq; color-coding indicates sequence class. “X-transposed” region arose from an X-to-Y transposition event that occurred ~2 million years ago. **b.** Repeat structure and gene locations within reference AZFc region. Different families of repeats are color-coded and arrows indicate repeat orientation. Each of the 5 gene families has multiple copies within AZFc; the location and orientation of each gene copy are indicated. cen = centromere, Mb = megabase, qter = long-arm terminus of Y chromosome.

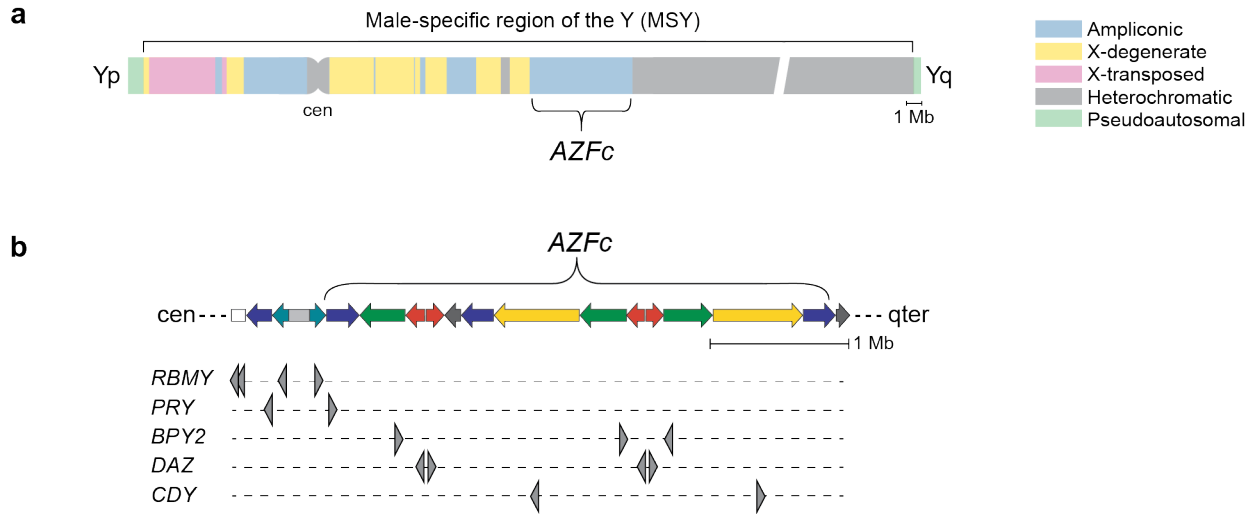


Figure 4. Comparison of ampliconic sequence content of mouse and human MSYs

Triangular dot-plots (drawn to scale) illustrate repeat structures of sequences. Each dot within the plot represents 100% nucleotide sequence identity within a 200-bp window. Direct repeats appear as horizontal lines, inverted repeats as vertical lines. Schematic representations of chromosomes are shown below plots; color-coding indicates sequence class. cen = centromere, Mb = megabase.

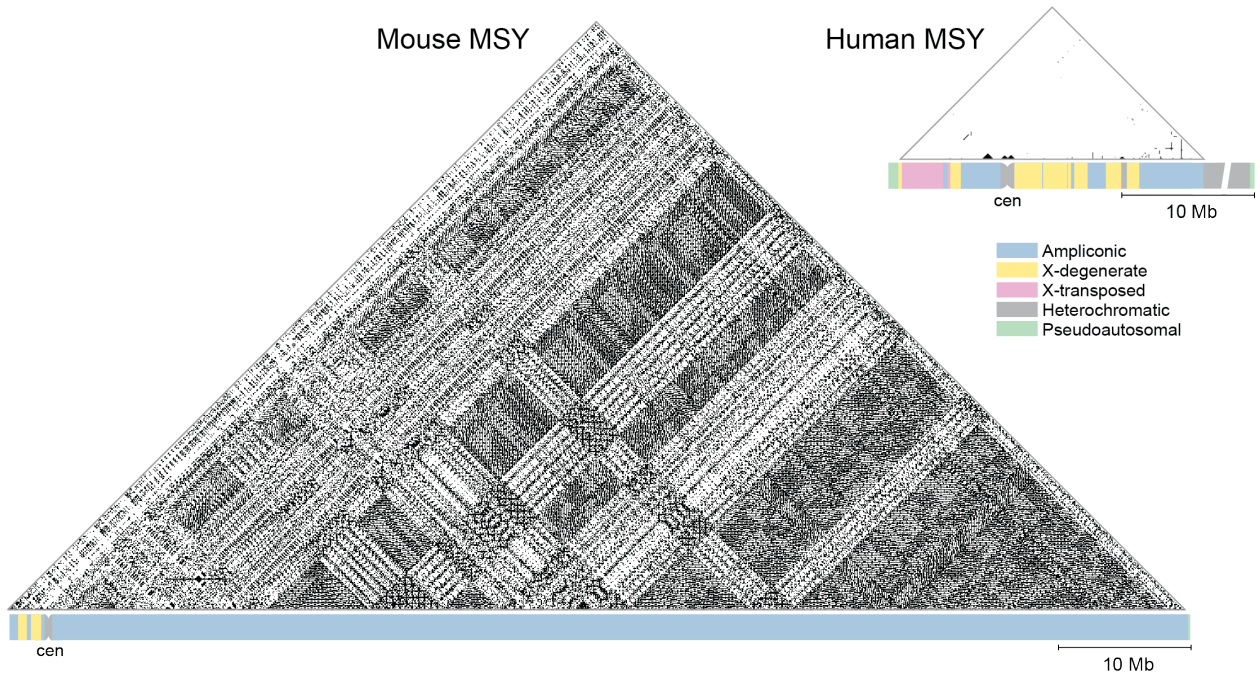


Figure 5. Isodicentric Yp (idicYp) formation

Ectopic recombination and crossing over between sister chromatids within human MSY palindrome P5 (shown as a pair of blue triangles) generates an idicYp chromosome (bottom). Most other Yq palindromes and inverted repeats have been found to mediate ectopic recombination in idicYp cases. cen = centromere.

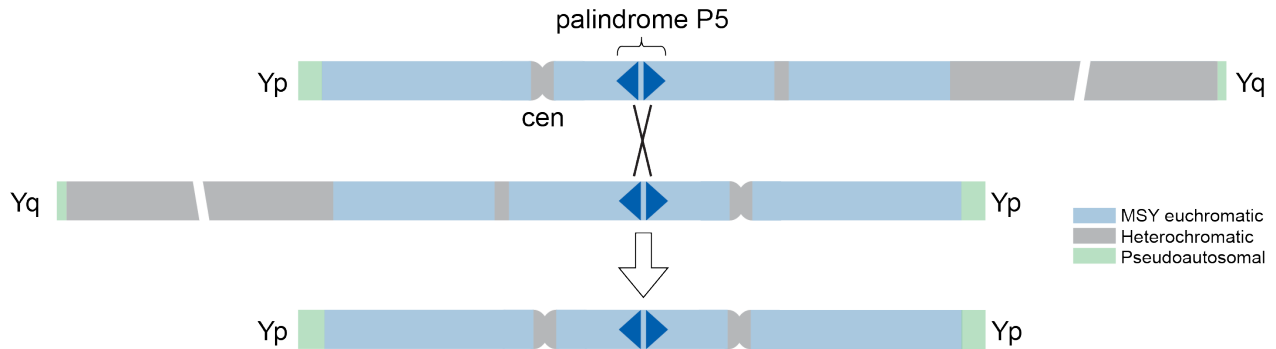


Figure 6. Comparative Y chromosome sequencing reveals long life spans and functional coherence of human MSY single-copy genes

a. Species tree indicating evolutionary relationships between the eight mammals with SHIMS-sequenced ancestral MSY sequences. Chicken is shown as an outgroup. Branch lengths are drawn to scale. **b.** Species distribution and features (expression breadth across tissues, expression in preimplantation embryos, haploinsufficiency probability, and predicted regulatory function) of human MSY single-copy ancestral genes, which are ranked according to evolutionary longevity. Total branch length for a given gene is the sum of branch lengths for each species possessing an intact homolog of that gene.

