Review Article

Repetitive Sequence and Sex Chromosome Evolution in Vertebrates

Tariq Ezaz and Janine E. Deakin

Institute for Applied Ecology, University of Canberra, Canberra, ACT 2601, Australia

Correspondence should be addressed to Tariq Ezaz; ezaz.tariq@gmail.com

Received 14 May 2014; Accepted 3 September 2014; Published 11 September 2014

Academic Editor: Paul Harrison

Sex chromosomes are the most dynamic entity in any genome having unique morphology, gene content, and evolution. They have evolved multiple times and independently throughout vertebrate evolution. One of the major genomic changes that pertain to sex chromosomes involves the amplification of common repeats. It is hypothesized that such amplification of repeats facilitates the suppression of recombination, leading to the evolution of heteromorphic sex chromosomes through genetic degradation of Y or W chromosomes. Although contrasting evidence is available, it is clear that amplification of simple repetitive sequences played a major role in the evolution of Y and W chromosomes in vertebrates. In this review, we present a brief overview of the repetitive DNA classes that accumulated during sex chromosome evolution, mainly focusing on vertebrates, and discuss their possible role and potential function in this process.

1. Introduction

Two major types of sex chromosome systems exist in vertebrates, XX female/XY male (e.g., human and salmon) and ZZ male/ZW female (birds and snakes). How these functionally important chromosomes evolve has been a topic for debate for more than a century, since the discovery of the first sex chromosomes in the late 1800s [1–3]. Despite the interest in this area, difficulties in sequencing highly repetitive Y and W chromosomes have hampered progress towards gaining a fuller understanding of the mechanisms involved in their evolution. This has resulted in the most detailed research on vertebrate sex chromosomes being carried out on species which have had at least part of the euchromatic region of their Y or W chromosome sequenced, for example, the evolutionary old sex chromosomes of eutherian mammals (three primates and two carnivores) [4–7] or the evolutionarily young sex chromosomes of fishes such as the half-smooth tongue sole [8], three-spine stickleback [9], and medaka [10]. The comparison between species of evolutionarily advanced Y chromosomes is not ideal for gaining insight into the mechanisms driving sex chromosome evolution as the chromosomes have undergone extensive changes and degeneration, perhaps even losing key clues required to unravel their evolution. In contrast, the sequencing of young sex chromosomes should help in elucidating these driving mechanisms [11], particularly if comparisons can be made to more highly diverged sex chromosomes that share a common ancestry.

Although we have some understanding of the molecular organization of sex chromosomes in model vertebrate species, it is largely unknown for the majority of the species where sex chromosomes have been identified cytologically. The conservation of sex chromosome gene content and the sex determining gene in most mammals does not reflect the diversities that exist in other vertebrate groups, where there has been rapid evolution of sex chromosomes in many lineages [12]. Despite the cytogenetic identification of non-homologous sex chromosomes among vertebrates, very few Y or W chromosomes have been sufficiently mapped and/or sequenced, for studies into their evolution, largely due to the abundant repetitive sequences on these chromosomes [13], and only a few sex determining genes have been identified. Nonetheless, the latest advances in molecular cytogenetics, DNA sequencing, and bioinformatics are making it possible to study the make-up of sex chromosomes in greater detail.
than ever before. Here we review the association of the amplification of repetitive sequences near the sex determining locus and discuss their possible role in the evolution of sex chromosomes and their potential function. As the technical advances in this field are relatively recent, we call upon data from organisms outside the vertebrate phylogeny to identify important areas for future research in vertebrates.

2. Sex Chromosome Evolution in Vertebrates

Sex chromosomes have evolved multiple times and independently throughout evolution. They have many unique features, including unique gene content (e.g., sex linked genes, including the master sex determining genes, such as SRY in most mammals) and existence of highly variable morphology among different taxa, often representing various evolutionary stages [14–16]. Such morphological variations have resulted due to a suppression of recombination leading to gene loss and the accumulation of repetitive sequence on one of the homologues (Y or W) [14, 17, 18]. They also experience special selective pressures compared to those of autosomes which, although debatable, are proposed to include both positive and negative selections, such as purifying selection to maintain sequences in the X-degenerate regions on the human Y chromosome [19] and even sex-specific selection, which has been demonstrated to impact on W chromosome gene expression in chicken [20]. These unique features make sex chromosomes the most dynamic entity in any genome [15, 17, 21], providing unique opportunities to study and understand genome evolution and organization.

Vertebrate sex chromosomes display enormous diversity in morphology and in gene content [12, 17, 22–27]. Such diversities not only imply multiple and independent origins of sex chromosomes, but also suggest evolution of a very specific molecular mechanism that is uniquely dynamic in performing a very specific task—sex determination. Despite such enormous diversities among taxa, two competing hypotheses have been put forward on sex chromosome evolution and degeneration based on the studies derived from mammals and birds (Y and W degeneration) and from nonamniotes, such as fish and frogs (fountain of youth) [15, 18, 28–31].

In most mammals, sex chromosomes are highly differentiated morphologically, usually represented by a large and gene rich X chromosome and a small, degenerated, and gene poor Y chromosome. This heteromorphism and subsequent in-depth molecular analysis led to the postulation of the most accepted theory for sex chromosome evolution. This hypothesis posits that sex chromosomes have evolved from an autosomal pair when the proto-Y chromosome acquired a sex determining gene/locus [28, 29]. Subsequently, chromosome rearrangements including accumulation of other sexually advantageous genes near the sex determining locus drove selection for suppression of recombination. This process in turn facilitated loss of active genes, deletions, and insertions, leading to degeneration of the proto-Y, making the sex chromosome pair morphologically differentiated [14, 18, 29, 32]. This hypothesis considers suppression of recombination as a result of chromosome rearrangements, which gradually expand into the nonrecombining regions leading to the gradual loss of genes from Y chromosomes, leaving only sets of genes that are required for development of testes and brain development [33]. It is well known that chromosome rearrangements (such as inversions, insertions, and deletions) suppress recombination between homologous chromosomes [34, 35]. The best example would include differentiation of mammalian sex chromosomes which has occurred through the genetic degradation of the Y chromosome following rearrangements [17, 23]. Although the chicken W chromosome appears to have evolved independently, it is proposed to have followed the same mechanisms [15]. Despite the tremendous morphological diversities of sex chromosomes observed in amniotes, the scenario is quite the opposite in nonamniotes, such as in fish and amphibians. Although most of the fish and amphibians (so far studied) display genotypic sex determination (GSD), in the majority of species, sex chromosomes are cryptic; that is, they are not morphologically differentiated [36–39]. Evolution of such homomorphic sex chromosomes in fish and amphibians has been attributed to the rapid turnover of sex chromosomes through acquisition of sex chromosome functions by native/existing genes, often unrelated to those that are part of the sex differentiating cascade [26, 30, 40, 41]. The mechanism behind this rapid turnover through de novo evolution of sex determining genes is not fully understood. However, Perrin [31] argued that such a mechanism is likely to be maintained by sex reversal and occasional recombination between sex chromosomes. The author’s observation is based on the empirical data that the recombination frequencies are associated with the phenotypes rather than the genotypes. Therefore, because the sex chromosomes are not strictly recombinationally isolated from each other (i.e., recombination in sex reversed individuals, which is common in fish and amphibians), the Y/W are protected from degradation (fountain of youth). Certainly, more evidence is required to support this hypothesis and is likely to be published in coming years from studies on fish to reptiles.

3. Chromosome Rearrangements and Repeat Accumulation

It is a well-observed phenomenon that chromosome rearrangements occurred particularly adjacent to the sex determination locus in most taxa. These include deletions, insertions, inversions, transpositions, and amplification of repetitive sequences [42–50]. This may suggest that it is likely that the sex determining locus or gene arises in a region of a chromosome which is unstable or fragile (e.g., contains common classes of fragile sites, such as AT-rich) or in a region which allows chromatin modification through cellular mechanisms (e.g., histone modifications). However, this is certainly not the case in mammalian sex chromosomes, as the human X chromosome contains only three fragile sites, while the Y contains none [46]. Nonetheless, the sex determining gene SRY lies very close to the pseudoautosomal region (PAR), which seems to be somewhat unstable [51]. However, the absence of fragile sites may represent advanced sex chromosomes, which have gone through the evolutionary
process of stabilization, but the presence of fragile sites may be a common phenomenon in many animals and plants with nascent sex chromosomes. It may also be likely that the sex determining locus or gene destabilizes the chromosome region in which it is located.

Simple repetitive sequences (e.g., microsatellites) are often accumulated in high copy numbers on the sex chromosomes in many taxa [41–45, 52–54], although those same motifs have low copy numbers distributed throughout the genome, implying preferential amplification. Even standard cytogenetic technique, such as C-banding, is able to detect heterochromatin accumulation on sex chromosomes (Figure 1). Why repetitive sequences preferentially amplify on sex chromosomes is still unknown, yet many theories have been put forward. One of the well-accepted theories is that the accumulation of repetitive sequences on one of the pair of sex chromosomes facilitates suppression of recombination between sex chromosome homologues, therefore, protecting the sexually beneficial mutations [14, 15]. On the other hand, it is equally plausible that chromosome rearrangements as well as repeat accumulation and amplification may occur near the sex determining locus as a result of suppression of recombination, rather than inducing it [55–57].

There have been a limited number of studies on relatively young sex chromosomes that have shown that the suppression of recombination, repeat accumulation, and chromosomal rearrangement can occur rapidly. In medaka (Oryzias latipes), the sex determining gene dmr1bY is derived from a duplicated fragment of an autosome that has inserted onto the proto-Y chromosome [10]. Sequence data indicates that repeats expanded in the sex determining region after the insertion onto the proto-Y [10]. Sequence comparisons and gene mapping of the homomorphic X and Y chromosomes of the three-spine stickleback (Gasterosteus aculeatus) have demonstrated that even these supposedly nascent sex chromosomes have diverged considerably, with higher repeat content on the Y than the X and chromosomal rearrangements such as inversions and deletion [9, 58]. However, in both of these cases, it is still unclear whether repeat accumulation led to suppressed recombination or vice versa.

While likely molecular mechanisms behind chromosome rearrangements are reasonably well known, the alternate mechanisms that initiate suppression of recombination near the sex determining locus are yet to be elucidated. It is also possible that a heritable epimutation, such as a change in DNA methylation, and not a genetic mutation in the sex determining locus may be the first step in sex chromosome evolution [59, 60]. If this epimutation could suppress recombination, it may result in the region being more susceptible to genetic mutation [59, 61] and the insertion of repetitive elements. One plausible and testable hypothesis behind suppression of recombination near the sex determining locus would lie within chromosome architecture. This primarily includes changes in chromatin structure (epigenetic such as DNA methylation and/or histone modification) inflicted by a sex-specific mutation, which is sufficient to suppress or reduce recombination between sex chromosomes. Investigation of chromosomal architecture around the region adjacent to a putative sex determining locus in nascent sex chromosomes from various taxa, including birds and snakes (e.g., ratite birds and boid snakes) would provide empirical evidence for unraveling any mechanisms which may be epigenetically driven.

Studies on the formation of the accumulation of repeats on the Y chromosomes of Drosophila species are perhaps more advanced than they are for vertebrates and provide valuable insight for future studies on the evolution of vertebrate sex chromosomes. By comparing Drosophila species with neo-sex chromosomes of different ages, it has been...
Table 1: An overview of repeats on sex chromosomes among divergent taxa.

<table>
<thead>
<tr>
<th>Satellite Telomere</th>
<th>Multigene rDNA</th>
<th>TE</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fish</td>
<td>Y/W</td>
<td>Y/W</td>
<td>[49]</td>
</tr>
<tr>
<td>Frog</td>
<td>Y/W</td>
<td>W</td>
<td>[62]</td>
</tr>
<tr>
<td>Snake</td>
<td>W</td>
<td>W</td>
<td>[38, 45]</td>
</tr>
<tr>
<td>Lizard</td>
<td>W</td>
<td>W</td>
<td>[52, 63]</td>
</tr>
<tr>
<td>Turtle</td>
<td></td>
<td>W</td>
<td>[64]</td>
</tr>
<tr>
<td>Bird</td>
<td>W</td>
<td>Y</td>
<td>[65–68]</td>
</tr>
<tr>
<td>Human</td>
<td>Y</td>
<td></td>
<td>[69]</td>
</tr>
<tr>
<td>Wallaby</td>
<td>Y</td>
<td></td>
<td>[70, 71]</td>
</tr>
<tr>
<td>Platypus</td>
<td></td>
<td></td>
<td>[72]</td>
</tr>
</tbody>
</table>

(ITS: interstitial telomere signal; TE: transposable elements).

possible to start tracing the events leading to the formation of a heterochromatic Y chromosome [73–76]. For instance, D. miranda, neo-sex chromosomes were formed approximately 1 million years ago (MYA) from the fusion of the ancestral Y chromosome with an autosome. This neo-Y chromosome presents a case of Y chromosome degradation and repeat sequence accumulation caught in the act, with only 1941 of the 2951 neo-X genes possessing intact open reading frames on the neo-Y and with almost half of these neo-Y genes expressed at lower levels than their neo-X counterparts [75]. In addition, nearly 50% of the neo-Y sequence consists of repeats [77, 78], demonstrating the rapid changes that have occurred in the evolution of sex chromosomes in just 1 million years of evolution [73]. Interestingly, the neo-Y is also enriched for epigenetic marks associated with heterochromatin, such as H3K9me2 and HP1α. Furthermore, neo-Y regions with higher repeat content have higher levels of H3K9me2 binding, which is consistent with models proposing that repeat accumulation enables the formation of heterochromatin on nascent Y or W chromosomes [76]. It is to be hoped that similar studies into the sequence and epigenetic features of vertebrate sex chromosomes will not be too far behind these Drosophila studies. Further studies involving high resolution sequence analysis of sex determining region in vertebrates, as well as investigation of chromosomal architecture surrounding the sex determining region, will reveal the true mechanisms that drive sex chromosome evolution after acquisition of a sexually advantageous gene or locus.

4. Is There a Particular Class (or Classes) of Repeats That Amplified Preferentially on Sex Chromosomes?

Large volumes of research papers have been published on physical mapping of various repetitive sequences in divergent taxa, from plants to mammals. These include amplification of satellite DNA (mini and micro), telomeric sequences (including megalotelomere in chicken), amplification of multi-gene families (rDNA and histones), taxon specific repeats, transposable elements (LINEs and SINEs), and multicopy genes. Table 1 presents a representative summary of the various classes of repetitive sequences which have been mapped in vertebrates. However, the majority of the mapping has been done on fishes as a tool for identifying sex chromosomes, as their sex chromosomes are often homomorphic. Amphibians also have a high frequency of homomorphic sex chromosomes; however, there have been only limited studies where repetitive sequences have been used to identify sex chromosomes [62]. The minisatellite repeat Bkm (branded krait minor) was characterized on the W chromosome of a snake species in the early 70s [79] and in other species including snakes [80–85]. Only two studies have so far been published on the repeat content of lizard sex chromosomes [52, 63]. It may not be the true scenario because of the unavailability of mapping information of all repetitive classes in representative taxa, but overwhelming data on the amplification of simple repeats near the sex-determining locus in divergent taxa suggest a common trend; that is, such accumulation of simple repeats may not be an artifact of the volume of the published literature but a true representation of the molecular mechanism. Information from sequenced Y and W chromosomes would support this as, for example, the half-smooth tongue sole W has more than double the repeat content of Z [8]. Perhaps this convergent amplification of simple repeats triggered by sex-linked mutation is the prime genomic driver that initiates sex chromosome evolution in many taxa.

5. Do Sex Chromosome Repeats Have a Function?

In the past, repetitive sequences were dismissed as part of the “junk DNA” [86], a term used by Ohno to describe the non-protein-coding regions of the genome [87]. We now know that at least some of these sequences are transcribed and obviously play a functional role in the genome [86]. When it comes to sex chromosomes, the idea of “junk DNA” is also gradually being rejected. Recent findings in taxonomically diverse species suggest that these sequences play an important role. However, evidence of a functional role of repeats on vertebrate sex chromosomes is limited at this stage, largely due to the difficulty in obtaining Y
or W chromosome sequences. Nonetheless, methods have been developed to achieve this challenging task [88, 89] and will undoubtedly lead to more studies in this area. Again, we are able to gain important insight into the function of sex chromosome repeats from species outside the vertebrate lineage.

One elegant study on the function of sex chromosome repeats was carried out on Schistosoma mansoni, a parasitic platyhelminth, with a 2W sex determination system. The Z and W chromosomes can be cytogenetically distinguished by their chromatin structure rather than by their size. Sequencing of male and female genomes led to the identification of W-specific sequences, which consisted of 36 W-specific repeats, including an SMalpha retroposon, LTRs, LINE2, and DNA transposons [89]. Of these repeats, transcription was detected for eight, and three of these were transcribed in the larval stages but not in immature or adult females. Using chromatin immunoprecipitation sequencing (ChIP-Seq), Lepesant and colleagues [89] were able to profile the level of an epigenetic mark associated with active chromatin, acetylated H3K9 (H3K9Ac), around repeats at various developmental stages. A gradual decrease in the level of H3K9Ac was observed for all W-specific repeats from larval to adult stages. Subsequent experiments performed on two transcribed and one nontranscribed repeats showed enrichment for another active mark, trimethylated H3K4 (H3K4Me3), in larval stages compared to adults. Repressive marks trimethylated H3K9 (H3K9Me3) and trimethylated H3K27 (H3K27Me3) associated with heterochromatin were enriched at these three repeats in cercariae, the larval stage capable of infecting a vertebrate host, but less abundant in miracidia (a different larval stage) and adults. It appears that the transcription of these repeats can be correlated with changes in chromatin structure. Lepesant and colleagues [89] suggest that it may even be changes in chromatin structure that may contribute to sex determination, perhaps even in the absence of a sex determining gene.

In Drosophila, the Y chromosome consists of megabase regions of repetitive sequences, such as microsatellites, transposable elements, and ribosomal DNA (rDNA), and it has been hypothesized that variation in the type, abundance, and distribution of these repetitive sequences can influence gene expression across the genome [90]. Most Drosophila Y chromosome polymorphisms are not located in protein-coding genes but in the heterochromatic regions where repetitive sequences are abundant. These polymorphisms affect the expression of many autosomal and X-linked genes [91–94], typically those that are located in repressive chromatin and are subject to tissue-specific expression [90]. It would appear that the Y chromosome acts as a giant regulator of gene expression through its global effects on chromatin dynamics [92, 95–97] and there is evidence from male and XXY female Drosophila lines, differing only in the origin of their Y chromosomes, that there is a yet to be ascertained underlying epigenetic mechanism involved [93]. This gene regulator is not acting by turning genes on or off but by working at a more subtle scale of gene regulation [98]. This effect on gene expression is also observed in XXY females where Y-linked genes are not expressed, providing additional support for the idea that the heterochromatic region of the Y acts as a gene expression modulator [93].

It would be interesting to know if the heterochromatic region of the human Y chromosome similarly plays a role in gene regulation. Such studies are yet to be conducted on the human Y, as most research efforts have focused on the protein-coding genes and not the heterochromatic region [98]. However, copy number variation of genes present in multiple copies on the mouse Y chromosome (Sly and Rnmy) is correlated with regulation of autosomal immune gene expression and, similar to the Drosophila Y, is most likely the result of a chromatin remodeling mechanism, suggesting that this regulatory function of the Y chromosome may be widespread [99]. Furthermore, epigenetic profiling of tandem repeat sequences within the euchromatic region of the human Y has shown specific patterns of histone modifications (active mark H3K9Ac, repressive marks H3K9me3 and H3K27me3) and CTCF binding (transcriptional repressor) associated with different repeats. These specific patterns of epigenetic marks suggest that tandem repeats on the Y chromosome may play a role in chromatin status and may act as regulatory elements [100].

The Bkm satellite repeat consists of tandem arrays of GATA nucleotides. Although not necessarily specific to the sex chromosomes, Bkm repeats are abundant on heterogametic sex chromosomes of divergent vertebrate species [81]. A sex and tissue-specific protein that binds specifically to Bkm repeats known as Bkm-binding protein (BBP) appears to be involved in coordinated decondensation of the heterogametic sex chromosome in germ cells [101, 102]. Subramanian et al. [103] reported that the GATA repeats on the human Y chromosome may play a role in marking the boundaries of chromatin domains. More recently, it has been discovered that GATA repeats have a conserved role, acting as insulators in both Drosophila and human cells [104].

Determining the function of repetitive sequences on Y or W chromosomes is in its infancy but it is clear that these sequences do play a functional role in gene regulation and chromatin structure. The demonstrated ability of the Drosophila and mouse Y chromosome to regulate gene expression across the genome is particularly exciting and is an area of research that needs to be pursued in more vertebrate species.

### 6. Conclusions/Future Directions

It is clear that repetitive sequences are a prominent feature of sex chromosomes across plant and animal kingdoms. These repeats have proven challenging for efforts to sequence Y and W chromosomes but bioinformaticians are rising to the challenge and are developing methods to obtain sequences from these unique chromosomes. This will enable many more Y or W-specific repeats to be identified and their functions to be determined. Typically, the heterochromatic regions of these chromosomes, where repeats are abundant, are ignored in favour of studying euchromatic regions containing protein-coding genes. However, even the limited number of studies that have been carried out on Y and W repeat function has shown that these sequences play an important role in gene
regulation, even beyond that of the sex chromosome itself. By determining the function of these repeats, perhaps we can then decipher whether the accumulation of repeats on sex chromosomes is a cause or consequence of recombination suppression and gain a better understanding of the steps involved in sex chromosome evolution.

Conflict of Interests
The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments
The authors would like to thank Ikuo Miura for his valuable comments on the paper. Tariq Ezaz and Janine E. Deakin are supported by ARC Future Fellowships (FT100100733, FT100100241).

References


A. J. Macdonald, N. Sankovic, S. D. Sarre et al., "Y chromosome microsatellite markers identified from the tammar wallaby (Macroopus eugenii) and their amplification in three other macropod species," *Molecular Ecology Notes*, vol. 6, no. 4, pp. 1202–1204, 2006.


