

Research

## Neo-sex chromosomes in the black muntjac recapitulate incipient evolution of mammalian sex chromosomes

Qi Zhou<sup>✉\*</sup>, Jun Wang<sup>✉‡§¶</sup>, Ling Huang<sup>✉†¥</sup>, Wenhui Nie<sup>¥</sup>, Jinhuan Wang<sup>¥</sup>, Yan Liu<sup>†¥</sup>, Xiangyi Zhao<sup>\*</sup>, Fengtang Yang<sup>#</sup> and Wen Wang<sup>\*</sup>

Addresses: <sup>\*</sup>CAS-Max Planck Junior Research Group, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences (CAS), 32# Jiao-chang Road, Kunming, Yunnan 650223, People's Republic of China. <sup>†</sup>Graduate School of Chinese Academy Sciences, 19# Yu-quan Road, Beijing 100039, People's Republic of China. <sup>‡</sup>The Institute of Human Genetics, University of Aarhus, Nordre Ringgade 1, DK-8000 Aarhus C, Denmark. <sup>§</sup>Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230, Odense M, Denmark. <sup>¶</sup>Beijing Genomics Institute, Bei-shan Road, Shenzhen 518083, People's Republic of China. <sup>¥</sup>Kunming Cell Bank, State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, 32# Jiao-chang Road, Kunming, Yunnan 650223, People's Republic of China. <sup>#</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton Road, Hinxton, Cambridge, CB10 1SA, UK.

✉ These authors contributed equally to this work.

Correspondence: Fengtang Yang. Email: fy1@sanger.ac.uk. Wen Wang. Email: wwang@mail.kiz.ac.cn

Published: 14 June 2008

Genome **Biology** 2008, **9**:R98 (doi:10.1186/gb-2008-9-6-r98)

The electronic version of this article is the complete one and can be found online at <http://genomebiology.com/2008/9/6/R98>

Received: 9 December 2007

Revised: 4 March 2008

Accepted: 14 June 2008

© 2008 Zhou *et al.*; licensee BioMed Central Ltd.

This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The regular mammalian X and Y chromosomes diverged from each other at least 166 to 148 million years ago, leaving few traces of their early evolution, including degeneration of the Y chromosome and evolution of dosage compensation.

**Results:** We studied the intriguing case of black muntjac, in which a recent X-autosome fusion and a subsequent large autosomal inversion within just the past 0.5 million years have led to inheritance patterns identical to the traditional X-Y (neo-sex chromosomes). We compared patterns of genome evolution in 35-kilobase noncoding regions and 23 gene pairs on the homologous neo-sex chromosomes. We found that neo-Y alleles have accumulated more mutations, comprising a wide variety of mutation types, which indicates cessation of recombination and is consistent with an ongoing neo-Y degeneration process. Putative deleterious mutations were observed in coding regions of eight investigated genes as well as cis-regulatory regions of two housekeeping genes. *In vivo* assays characterized a neo-Y insertion in the promoter of the *CLTC* gene that causes a significant reduction in allelic expression. A neo-Y-linked deletion in the 3'-untranslated region of gene *SNX22* abolished a microRNA target site. Finally, expression analyses revealed complex patterns of expression divergence between neo-Y and neo-X alleles.

**Conclusion:** The nascent neo-sex chromosome system of black muntjacs is a valuable model in which to study the evolution of sex chromosomes in mammals. Our results illustrate the degeneration scenarios in various genomic regions. Of particular importance, we report - for the first time - that regulatory mutations were probably able to accelerate the degeneration process of Y and contribute to further evolution of dosage compensation.

## Background

It is believed that in human and other eutherian mammals, the heteromorphic sex chromosomes evolved from a pair of ordinary autosomes between 166 and 148 million years ago [1-4]. After the birth of the sex-determining gene, extensive recombination suppression evolved between proto-X and proto-Y to prevent the sexual reversal along the entirety of the chromosome pair, with the exception of a short 'pseudoautosomal region' (PAR) [5]. Theoretical models predict that proto-Y chromosomes would further suffer a rapid accumulation of deleterious mutations and be subjected to drastic gene loss [6], consistent with only 45 genes surviving on the human Y chromosome as compared with 1,000 functional genes on the X [1,3,7]. Y degeneration is expected to be driven by multiple forces, including Muller's ratchet, background selection, the Hill-Robertson effect with weak selection, and the 'hitchhiking' of deleterious alleles by favorable mutations [5]. These factors can work on both ancient Y and heterosynaptic autosomes with suppressed recombination because of the diminished efficiency of natural selection along them. Newly evolved sex chromosome systems are required to test hypotheses about Y degeneration because ancient Y chromosomes bear few remaining traceable ancestral sequences. Such cases include the recently originated sex determination system, like that of *Silene* [8], or 'neo-sex chromosomes', formed through a recent fusion or translocation between an autosome and a sex chromosome followed by extensive recombination suppression, and thus showing inheritance patterns like that of the better known ancient sex chromosomes [9,10].

Our current knowledge of neo-sex chromosome evolution is primarily due to extensive work in *Drosophila* and plants [8,10-12]. Most of them focused on degeneration patterns in protein-coding regions of neo-Y alleles. However, given the elaborate regulation of gene expression [13,14], the scenarios of Y degeneration involving both gradual loss of functions of protein products as well as regulatory disorders remain to be elucidated. Furthermore, sex chromosome systems have evolved independently in different phyla many times, creating a need to examine their evolution in different taxa [15]. Sex chromosomes in mammals, including our own, have different autosomal origins from those of other organisms, involving entirely distinct gene sets in the process of sex determination and dosage compensation [15,16]. The inhibition of recombination between the mammalian proto-X and proto-Y was achieved through chromosomal inversions on the Y chromosome, whereas it was achieved through zero crossover in male germline cells in *Drosophila* [5,17]. Finally, different generation time and population size between mammals and other species would have great impact on the rate and patterns of mammalian Y-chromosome degeneration [18].

Thus, direct investigation of a neo-sex chromosome system in a mammal promises new insights into several fundamental

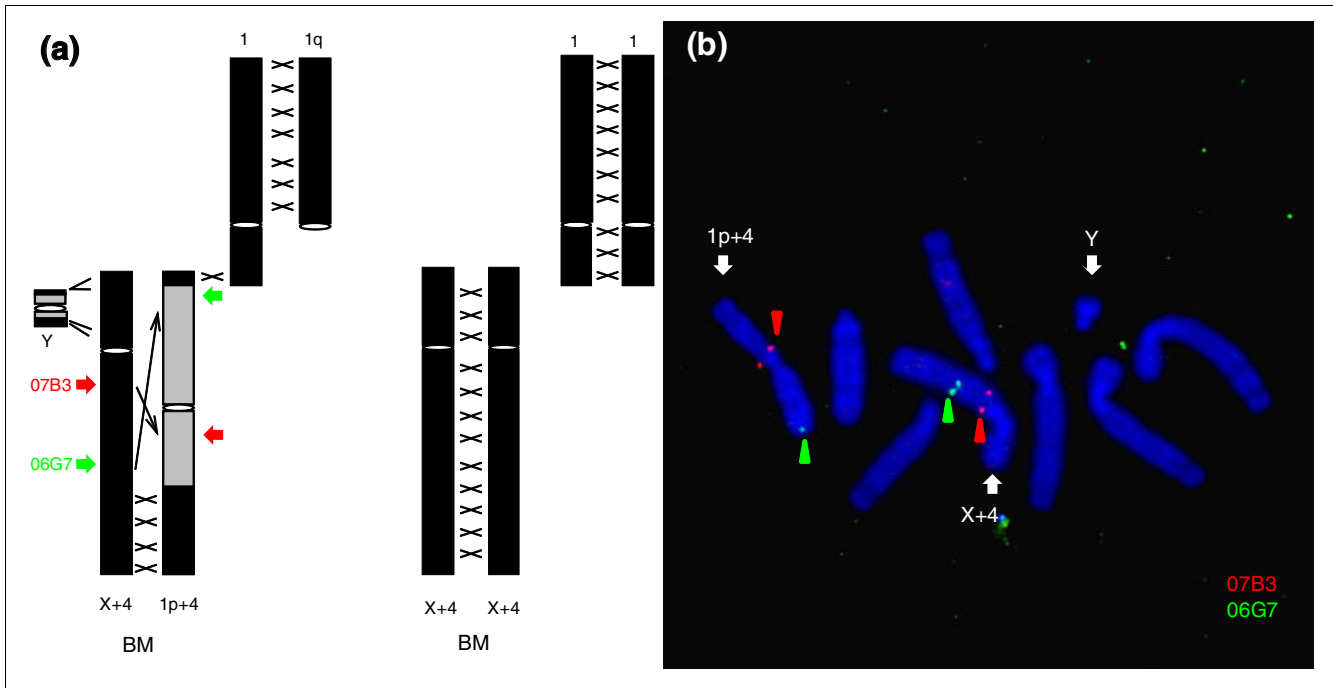
issues, including both mode and tempo of mammalian Y degeneration, and how mammals cope with degenerated Y alleles before the creation of *Xist*-dependent dosage compensation [19]. In this study we sought to address these questions using the black muntjac (*Muntiacus crinifrons*), an Asian barking deer, as the model [20]. In this species, a male-specific extensive chromosome inversion on autosome 4 and fusion of its homolog to the ancient X [21,22] has led to the very recent creation (within the past approximately 0.5 million years) [23] of a neo-sex chromosome system, similar to the inferred creation of the ancient mammalian X-Y system. Such a rare nascent mammalian neo-sex chromosome system provides an unprecedented opportunity to study sex chromosome evolution in mammals.

## Results and discussion

### Neo-sex chromosomes in the black muntjac

The black muntjac is endemic to a narrow region of southeastern China [20]. Habitat disruption in recent decades has rendered the species one of the most endangered mammals in the world. Cytogenetic analysis revealed a compact karyotype, with  $2n = 8\text{♀}/9\text{♂}$ , and half of the genome forms a pentavalent (Figure 1a) during male meiosis [24]. The chromosome 4 pair has been shaped by several large chromosomal events. First, one copy experienced a centric fusion to the regular X chromosome, forming a new 'X+4' (see Additional data file 1). In addition, the short arm of chromosome 1 has undergone a male-specific translocation to the homologous chromosome 4, creating a primitive '1p+4' chromosome [21,22]. Surprisingly, subsequent inversions involving large part of the primitive 1p+4 were proposed to take place in male black muntjacs (regions from '22a' to '17a'; Additional data file 1) [21,22]. We further confirmed this chromosomal rearrangement using dual-color fluorescence *in situ* hybridization (FISH) with two bacterial artificial chromosome (BAC) clones of the Indian muntjac (*Muntiacus muntjac vaginalis*) as probes (Figure 1b). The green signal in the middle of X+4 has switched its location to the distal end of 1p+4 as a result of the inversion (Figure 1b). Because no inversion loop has been detected during meiosis of male black muntjacs [24], the whole 1p+4 should remain heterosynaptic except for the two distal ends, which are not involved in the inversion. These two ends are analogous to those of ancient human Y chromosome, which can still synapse with each other and can thus be regarded new 'PAR's. The X+4 chromosome can thus be regarded as a 'neo-X', and the 1p+4 as a 'neo-Y' chromosome, because of probable lack of recombination in the inverted region.

The black muntjac X+4:1p+4 neo-sex chromosome pair comprises approximately one-fifth of the entire genome and thus bears thousands of neo-sex gene pairs [21,24]. Although X-autosome fusions were also detected in other muntjac species, such 1p+4 inversion is found exclusively in male black muntjacs [25]. The absence of such a system within *Muntia-*



**Figure 1**  
 Neo-sex chromosomes of *Muntiacus crinifrons*. **(a)** Paring patterns in male and female black muntjacs during meiosis. The black areas on chromosomes represent homosynapsis regions, whereas gray areas represent heterosynapsis regions. Cross lines between chromosomes represent homologous recombination. A pentavalent involving five chromosomes will form during meiosis of male black muntjacs, whereas such structures are absent in female black muntjac. **(b)** Inversion on the neo-Y chromosome revealed by fluorescence *in situ* hybridization using two Indian muntjac bacterial artificial chromosome (BAC) clones, 06G7 (green) and 07B3 (red), as probes.

*cus* genus, from which black muntjac is estimated to have diverged within the past 0.5 million years [23], indicates a very recent origin, and makes this system the youngest known mammalian neo-sex systems.

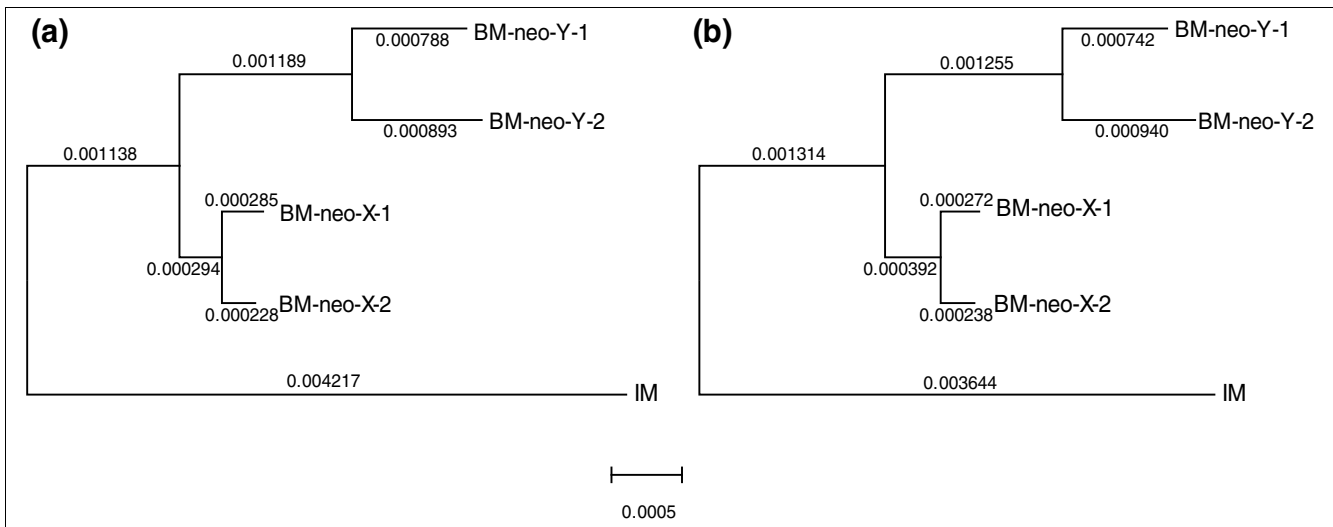
**Variation pattern in neo-Y noncoding regions**

To seek molecular evidence for inhibition of recombination and to compare the evolutionary patterns of neo-X and neo-Y chromosomes, we sequenced the intergenic and intronic fragments from the inverted region composing a total of 35.1 kilobases (kb), as well as 12.6 kb from the new 'PAR's and autosomal regions for two male and one female black muntjacs, as well as the orthologous segments from one male Indian muntjac, which is taken as the outgroup.

Under neutrality, intraspecific genetic diversity ( $\theta$ ) is expected to be proportional to the effective population size ( $N_e$ ) times the mutation rate ( $\mu$ ; specifically,  $\theta = 4N_e\mu$ ) [26]. Therefore, if recombination has authentically ceased between neo-Y and neo-X, then we would expect a reduced DNA polymorphism caused by reduced  $N_e$  for the neo-Y regions [6,27]. Otherwise, the polymorphism level of neo-Y should be similar to that of 'PAR' regions or autosomes. Consistent with the former expectation, we found a significantly lower number of segregating sites ( $\chi^2$  test,  $P < 0.01$ ; see Materials and methods [below]) and a lower polymorphism level of neo-Y regions ( $0.00168 \pm 0.00120$  versus  $0.00190 \pm 0.00029$ ) compared

with regions that undergo homologous meiotic pairing (Table 1), indicating suppression of recombination between neo-Y and neo-X. Of course, other factors including the Hill-Robertson effect, background selection, hitchhiking effect, or Muller's ratchet could have all reduced the polymorphism level within the investigated noncoding regions [6]. In contrast, the 'male-driven evolution' effect, which proposes that the neo-Y would undergo more rounds of cell division per generation in male than in female germ lines, would increase the mutation rate and polymorphism level of neo-Y alleles compared with those of neo-X and autosomes [28]. Our phylogenetic analysis of neo-X and neo-Y sequences in the male black muntjac confirms that the neo-Y alleles have accumulated far more mutations than neo-X (Figure 2; statistically significant, by Tajima's relative rate tests,  $P < 0.001$ ).

A direct estimate of the degree of male-driven evolution effect ( $\alpha$ ) on neo-Y can be derived from the comparison of branch lengths of homologous neo-Y and neo-X. We calculated the ratios of neo-Y to neo-X mutations separately from internal, external, and summed branch lengths of neo-Y and neo-X, and we further estimated the  $\alpha$  values based on the method proposed by Miyata and coworkers [29] (Table 2; see Materials and methods [below]). Because all of the ratios are higher than 3, the  $\alpha$  values estimated from different branches all approach to infinity [29]. This indicates a strong male-driven effect on the neo-Y chromosome in the male black muntjac,



**Figure 2**  
Dendrogram constructed using noncoding sequences of neo-Y and neo-X fragments. A total of 35.1-kilobase noncoding sequences of neo-Y and neo-X alleles in the black muntjac and the orthologous sequences in the Indian muntjac as the outgroup were used to construct the dendrogram. 'BM' stands for 'male black muntjac'; 'IM' stands for 'Indian muntjac'; '1' and '2' represent the two male black muntjac individuals separately. **(a)** Tree constructed by neighbor-joining method. Branch lengths calculated using Kimura's two-parameter method are shown above the corresponding branches. **(b)** Tree constructed by maximum likelihood method. Branch lengths calculated by baseml in PAML package using 'HKY85' method were shown above the corresponding branches.

and a small effect of ancestral polymorphism on the estimate [29-31]. It has been reported that great variation in mutation rates in different genomic regions greatly affect the estimation of  $\alpha$  value [32]. However, such effects may be limited in our case because the investigated regions are randomly sampled along the neo-Y chromosome and then concatenated together to represent a chromosome-wide estimate (see Materials and methods [below]). Male-specific methylation effects in the germline may also make little contribution to such differences in mutation rate between neo-X and neo-Y alleles. It predicts a greater GC content on neo-X alleles and an elevated rate of C  $\rightarrow$  T transitions on neo-Y alleles [33]. However, we detected a similar GC content and an almost identical ratio of GC  $\rightarrow$  AT versus AT  $\rightarrow$  GC substitutions between the neo-Y and neo-X alleles (Additional data file 2).

**Table 1**

**Segregating sites and polymorphism in noncoding sequences of different genomic regions in two male black muntjacs**

	neo-Y regions	PAR/autosomal regions
Sequenced length (bp)	35,156	12,653
Segregating sites <sup>a</sup>	59	44
Segregating sites/kb	1.68	3.48
Polymorphism ( $\theta_w$ )	0.00168 $\pm$ 0.00120	0.00190 $\pm$ 0.00029

<sup>a</sup>Segregating sites means total sites with substitutions in the investigated alleles. Standard deviations calculated based on variance (see Materials and methods) were shown for polymorphism data. bp, base pairs; kb, kilobases; PAR, pseudoautosomal region.

The net effect of all of these factors explains why the polymorphism of neo-Y region (0.00168) is one-quarter higher than that of recombining regions (0.00190/4; Table 1). Finally, other factors, including sexual selection and population subdivisions, could have effects on the polymorphism pattern [34]. These findings indicate that it is highly likely that recombination has ceased between the inverted neo-Y region and its homolog.

**Degeneration in coding regions of neo-Y-linked genes**

Because of the limited sample size of neo-Y alleles, the polymorphism data exhibit a high degree of variance, and differences between recombining and nonrecombining noncoding regions may only be indicative rather than statistically significant (Table 1). That stated, we continued to seek for other typical evidence of recombination inhibition (accumulation of deleterious mutations due to reduced efficiency of natural selection) [6]. To evaluate the degeneration effects of such mutations on neo-Y-linked genes, we characterized 23 gene

**Table 2**

**Estimation of  $\alpha$  (male:female ratio of mutation rate) from different branch lengths in the neo-sex system of black muntjac**

	neo-Y/neo-X	$\alpha$	95% CI of $\alpha$
Internal branches	4.0442 $\pm$ 1.2633	$\infty$	25.3848 to $\infty$
External branches	3.2768 $\pm$ 1.2356	$\infty$	4.2578 to $\infty$
Summed branches	3.6857 $\pm$ 0.8867	$\infty$	27.8507 to $\infty$

Note that we calculated neo-Y:neo-X ratios of mutation rate from branch lengths designated in Figure 2a. We further calculated their confidence intervals (CIs) and estimated  $\alpha$  values following [30,31].

**Table 3****Summary of neo-Y specific mutations occurred in the cDNA of investigated genes**

Gene symbol	Positions on neo-Y <sup>a</sup>	Shared mutation	Polymorphic mutation
<i>MYO1D</i>	17a		One synonymous substitution
<i>RPL19</i>	17a	One synonymous substitution	One synonymous substitution
<i>CACNB1</i>	17a		<b>A to S<sup>b</sup></b>
<i>CLTC</i>	17a	Two synonymous substitutions	
<i>ZNF24</i>	17a	One synonymous substitution	
<i>AKAP7</i>	3b	Three synonymous substitutions; V to M	
<i>SYNE1</i>	3b	One synonymous substitution; I to V; <b>V to E<sup>b</sup></b>	Two synonymous substitutions
<i>SNX22</i>	8	46 bp deletion in 3'-UTR	
<i>RIC8B</i>	1a	<b>S to P<sup>b</sup></b>	
<i>SCN1A</i>	3a		<b>D to G<sup>b</sup></b>
<i>MPPE1</i>	4a	One synonymous substitution	

<sup>a</sup>Positions correspond to chromosomal regions depicted in Figure 1a. <sup>b</sup>Missense mutations with drastic change of biochemical property of amino acid (polarity or charge), and those in bold represent sites that are highly conserved in other eutherian mammals with published sequences. We defined 'shared mutations' as those mutations detected in neo-Y alleles in both male black muntjacs, whereas 'polymorphic mutation' are those detected in only one male black muntjac.

pairs that are located in the inverted region on the neo-Y chromosome (Additional data file 1), one pair in the neo-Y 'PAR' region, and six pairs on autosomes. These genes were selected based on the annotations of BAC sequences of the Indian muntjac and cow genomic information in order to represent a diversity of regions across the neo-sex chromosomes [35]. In total, we obtained about 25 kb cDNA sequences in the inverted region and 6.3 kb in noninverted regions from two male black muntjacs, one female black muntjac, and one male Indian muntjac.

Among the 23 gene pairs located in the inverted region, we observed 14 mutations in eight genes shared by the two males in the protein-coding or untranslated regions (UTRs) of the neo-Y alleles (Table 3). In contrast, we detected no variation in genes located in the neo-Y 'PAR' region. For the autosomal region, only one out of six genes is heterozygous at one site in both the two male black muntjacs. Such an elevated rate of gene evolution exhibited by neo-Y alleles relative to their neo-X homologs and other autosomal genes is consistent with reduced efficiency of natural selection against deleterious mutations in the neo-Y. Combined with the variation patterns in intronic and intergenic regions (see the section above), these findings confirmed the recombination inhibition on the neo-Y chromosome.

Specifically, four nonsynonymous substitutions shared by the two male black muntjacs and two polymorphic nonsynonymous substitutions in one of the males were observed (Table 3). Of these six changes, four are highly conserved in amino acids across six other eutherian mammals (human, chimpanzee, dog, mouse, rat, and cow; Table 3). These four mutations change the polarity or charge of the amino acid, which might severely affect the protein function of the neo-Y copies. We also detected a neo-Y-linked deletion in 3'-UTR of the gene *SNX22* in both males investigated. This deletion abolishes a

microRNA target site (see below) and thus may affect stability of the neo-Y allele's mRNA [36,37]. By taking into account the above-described four possible deleterious nonsynonymous mutations and the deletion in the 3'-UTR of *SNX22*, we conservatively estimate the rate of accumulation of deleterious mutations in exon regions of neo-Y genes to be approximately 0.4 mutations/kb per million years (5 mutations/25 kb per 0.5 million years). We also estimated that only about 1.3 nonsense mutations would be expected within 25 kb of investigated cDNA sequences, assuming equal chance of mutation at each site (see Materials and methods [below]). Such a low probability of nonsense mutation plus the effect of possible purifying selection give credence to our observation that there were no coding frame-disrupting mutations among the investigated muntjac neo-Y alleles.

This finding differs from that of a recent investigation conducted in the neo-sex system of *Drosophila miranda* [38], which found that the neo-Y alleles of 24 out of 64 genes contain premature stop codons and/or frameshift mutations, suggesting an average loss rate of 20 genes per million years in the proto-Y of *Drosophila*. Such a drastic difference in the degeneration rate between neo-Y chromosomes of mammals and *Drosophila* could be attributable to the great differences in their mutation rates ( $2.2 \times 10^{-9}$  versus  $3.1 \times 10^{-7}$  per base per year for mammal and *Drosophila*, respectively) [39,40], generation time, DNA repair efficiencies [41], and effective population size [18]. It also suggests that, during the nonlinear process of mammalian Y degeneration [3], the rate of gene loss might be very slow in the early stage of mammalian Y chromosome evolution.

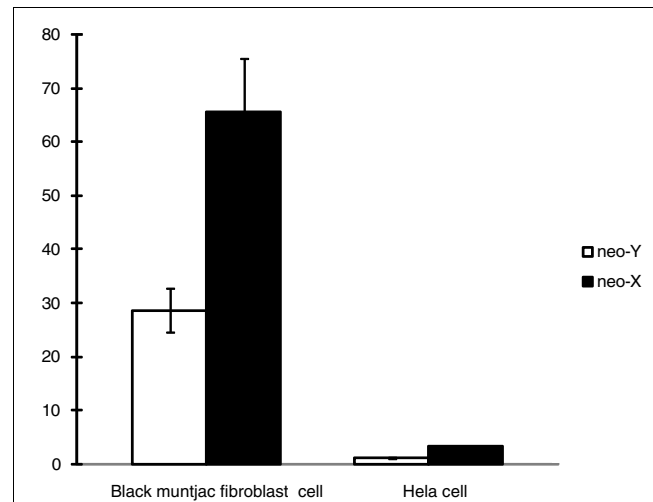
Multiple factors can fix the above neo-Y specific mutations in populations in both coding and noncoding regions, such as Muller's ratchet, background selection, hitchhiking effects on linked deleterious mutations, and faster mutation rate in

males [6,28,42,43]. Muller's ratchet must have had strong effect on accumulation of these mutations, given that the number of mutant-free chromosomes in a population is positively correlated with the effective population size ( $N_e$ ) [6,18]. Because of the usual small  $N_e$  values of mammals, mutant-free neo-Y alleles should be vulnerable to an irreversible loss by random drift by the 'ratchet' process, which further leads to accumulation of deleterious mutations. It is less likely that the observed neo-Y variations were from either ancestral polymorphism or gene conversion between neo-X and neo-Y, because we used both female black muntjac and Indian muntjac genes as the references to define neo-Y mutations [42]. It is also less likely that there were recent selective sweeps on the neo-Y of the black muntjac, based on the DNA polymorphism data. Recent strong positive selection would homogenize the neo-Y chromosomes among individuals [6], but we observed many polymorphic sites existing in both noncoding and coding regions of the muntjac neo-Y chromosome (Tables 1 and 3). Regarding other processes, because of the difficulty associated with collecting population data for such a rare species, we cannot currently test the existence of either a slower rate of adaptive evolution [44] or background selection [6], as previously proposed for *D. miranda* [10].

### Degeneration in cis-regulatory regions of neo-Y-linked genes

Recent studies have revealed that cis-regulatory regions more often underlie the expression divergence between species and evolution of morphologic diversity [14,45,46]. Apart from the direct evidence of degeneration in coding regions presented above, we sought to survey degeneration in regulatory regions on the neo-Y chromosome. We analyzed putative promoters expanding 1.5 kb around transcriptional start sites (TSSs) of eight neo-Y genes, which have mutations shared by the two male black muntjacs in their protein-coding/UTR regions (Table 3).

We detected neo-Y specific mutations shared by both males present in the investigated regions in three out of eight genes (see Additional Data File 3). In order to assess the effect of these mutations on gene expression, allelic promoters were subsequently cloned into pGL3-Basic plasmid in front of a firefly luciferase reporter gene. Each plasmid was mixed with a plasmid (pRLs-TK) containing a constitutive promoter driving sea-pansy luciferase reporter gene and co-transfected into HeLa cells and male black muntjac fibroblast cells. The ratios of firefly luciferase to sea-pansy luciferase, representing normalized promoter activity, were compared between neo-Y-linked promoters and neo-X-linked promoters. We found a significant decrease in neo-Y promoter activity in both *HeLa* and black muntjac cell lines for the *CLTC* gene (Figure 3). We did not detect promoter activities for the cloned fragments of other two genes; specifically, there is no detectable expression of the reporter genes using the cloned fragments as the promoters. The only detected neo-Y specific mutation is a 1 base pair (bp) insertion 55 base pairs upstream of the putative TSS



**Figure 3**

Mutation in the promoter of *CLTC* gene severely causes downregulation of the neo-Y copy. Dual-reporter assay of promoter activities of *CLTC*. Standard error among triplicates was shown on the bar.

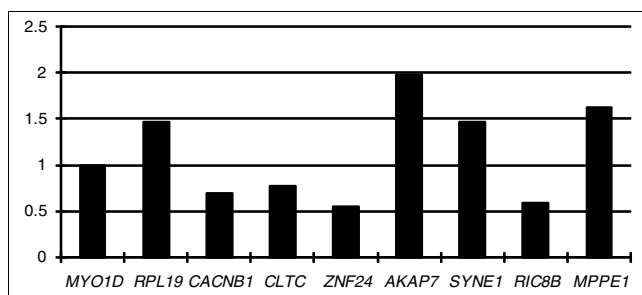
of the *CLTC* gene shared by both males (Additional data file 3). It is probably the mutation responsible for reducing the transcription level of neo-Y, given that this region was proposed to contribute positively to the core promoter activity by a recent comprehensive analysis of 387 promoter structures in humans [47].

In addition to mutations in promoter regions, we also detected a neo-Y specific deletion shared by both male black muntjacs in the 3'-UTR of the gene *SNX22* as mentioned above (Table 3). The abundance of microRNA targeted sequences in this region drove us to investigate whether this deletion would affect such motifs [36,37]. Using human microRNA database as a reference, we found that the deleted region on the neo-Y is in fact predicted to be a binding motif targeted by the microRNA hsa-mir-210 [48], which also has orthologs in cow and mouse with validated expression. Intriguingly, we found evidence that this same microRNA targeted motif is also deleted because of male-specific RNA editing in intact neo-X alleles (see Additional data file 4); 23 bp overlapping the target sequence are absent from transcripts in males (the edited form is much less frequent in females; see Additional data file 5). The diminutive size of the involved sequence and lack of splicing boundaries suggest that it is a true RNA editing event rather than a splicing event. It could be a result of selection against allelic expression imbalance between neo-X and neo-Y or a sign of degeneration of *trans* factors on neo-Y in males, which control the RNA-editing degree in the 3'-UTR of *SNX22*.

One of the fundamental issues in the process of Y-chromosome degeneration is whether it is mainly driven by direct degradation of the protein products (for example, in *D. miranda*) or reducing gene expression levels on an evolving Y

[6]. Our results suggest that these two processes may work in concert after the immediate recombination suppression on the neo-Y chromosome. It is proposed that deleterious mutations with mild fitness effects have a higher rate of fixation on a degenerating Y [6], as illustrated in Table 3. After accumulation of such mutations in the protein coding regions, a subsequent reduced expression of Y would be favored by natural selection to prevent the production of defective Y-linked products [6,18]. Alternatively, reduced expression of Y may take place before the accumulation of deleterious mutations in protein products. In the case of *CLTC* and *SNX22*, both are broadly expressed genes in most of the human tissues, with orthologs present in all vertebrates whose genomes have been sequenced. *CLTC* plays a critical role through coating membrane vesicles during endocytosis [49] and *SNX22* is a member of the protein family responsible for protein trafficking [50]. The protein coding regions of these housekeeping genes must be under strong selective constraints, in which no amino acid replacement substitution was detected (Table 3). On the other hand, recent studies showed that *cis*-regulatory regions - especially microRNA and transcription factor binding sites - usually have lower interspecies sequence conservation compared with protein coding regions, suggesting that regulatory mutations can be more easily fixed [46,51].

This suggests a scenario involving two independent steps of degeneration. First, mutations in regulatory regions (for example, the promoter of *CLTC*) would decrease or even turn off the gene expression of the Y alleles. Second, decreased expression could reduce selective constraints acting on the protein products of proto/neo-Y-linked genes [52]. We note that both these processes could in fact be favored to maintain the optimal gene dose between males and females after the establishment of dosage compensation in X [53]. Also, both processes would further accelerate the degeneration process in the Y-linked coding regions, which is supported by the recent finding that lower expressed neo-Y genes appear to have a faster accumulation rate of deleterious mutations in *D. miranda* [54].



**Figure 4**  
Expression divergences between nine neo-Y and neo-X gene pairs. All expression assays were done in duplicate and double checked in both male individuals. Mean expression ratios of neo-Y to neo-X are shown. The genes are arranged following the order from the centromere to the distal region of the 1p+4 chromosome.

### Transcription divergence between neo-Y and neo-X alleles

We semi-quantified and compared the mRNA abundance of neo-Y and neo-X alleles in 11 genes with mutations listed in Table 3. Consistent with our promoter assay, *CLTC* exhibits a lower expression level on neo-Y. We found significant expression difference between neo-Y and neo-X allele only in the gene *SNX22* (Fisher's exact test,  $P < 0.01$ ). However, this gene was further excluded from the result, together with *SCN1A*, because of their effects of allelic-biased amplification (see Materials and methods [below]). Eight of the remaining nine genes exhibited differential expression; interestingly, as many showed higher expression on neo-Y as on neo-X (4 versus 4; Figure 4). This finding suggests that neo-Y alleles can be distorted from normal expression level, either downregulated or upregulated, as a result of degenerating control of gene expression. Together with the mutation analysis results presented above, the expression patterns of the nine investigated genes exhibit a similar random inactivation mode of gene evolution on a degenerating Y chromosome proposed for *D. miranda* [12]. This model predicts that neo-Y genes are randomly inactivated, regardless of their level of adaptation. As shown in Figure 4, the regulation direction of neo-Y genes fluctuates along the neo-Y chromosome, and between some adjacent genes it is even opposite, suggesting that there is no large segment inactivation in the black muntjac's neo-Y chromosome. Two genes (*CLTC* and *ZNF24*) with only synonymous neo-Y mutations exhibited lower expression level, whereas two genes (*SYNE1* and *AKAP7*) with nonsynonymous neo-Y mutations exhibited higher expression level. In addition, the human orthologs of all nine investigated genes are widely expressed housekeeping genes. Therefore, whether a neo-Y allele is subjected to expression alteration or the direction of alteration may be not strongly correlated with the characteristics of the gene.

This random inactivation pattern indicates that the *Xist*-dependent dosage compensation [19] has not spread across the whole neo-X chromosome in females within the past 0.5 million years. This could be attributed to a lack of sufficient time for the establishment of action of *Xist* on neo-X regions. Concordantly, there is also clear molecular evidence that neo-X chromosome in male *D. miranda* is only partially dosage compensated [55]. A recent study [56] showed that even in the 'ancient' sex determination system, such as that of human, 15% of X-linked genes would escape the X inactivation, suggesting that the mammalian dosage compensation mechanism is leaky and its evolution process might be slow. As suggested by our results in male black muntjacs described above, not all the neo-Y alleles have been subjected to the reduction of gene expression. A nonspecific whole-set dosage compensation mechanism is likely to be seriously deleterious for genes that are transcriptionally active or are undergoing degeneration. It is also possible that a major regulator of dosage compensation, such as *Tsix* gene, counteracted *Xist* RNA on neo-X regions to prevent such deleterious effects [57,58].

However, such delicate regulation is less likely to have been established within such a short evolutionary time. Overall, we propose here that in the early stages, dosage compensation in males might have evolved gradually in a gene-by-gene fashion. This model is also consistent with the conjecture that evolution is accomplished more through small steps than large changes [59].

Charlesworth [18] has proposed two models to describe evolution of dosage compensation in manners similar to the two paths of Y-chromosome degeneration mentioned above. The common features of both models involve the upregulation of gene expression from the proto-X chromosome at the initial stage of Y degeneration [18], which has been suggested by recent global expression analyses conducted in *Drosophila*, worm, and mammals [60,61]. One model proposed by Charlesworth suggests that there would be selection promoting upregulation of the transcription of X alleles to compensate for defective products, with mutations accumulated in the coding regions of Y alleles. Alternatively, mutations could arise to make the Y alleles less responsive to the regulatory molecules, as exemplified by the case of *CLTC* in this study. It would consequently be a passive process, upregulating the X alleles' expression as a result of the excess of regulatory molecules from the Y alleles [18]. Further study into the paths of Y chromosome degeneration and comparison of neo-X expression between males and females using more gene pairs would be able to uncover which model mainly contributes to the Y degeneration and evolution of dosage compensation.

## Conclusion

Here we characterize the recently formed neo-sex system of the black muntjac. This unprecedented system is valuable for studying mammalian Y-chromosome degeneration and evolution of dosage compensation. Our results provided molecular evidence for recombination suppression in the neo-Y chromosome. As a result, excess of putatively deleterious mutations were observed in the coding regions of the investigated genes, probably because of the Muller's ratchet effect or background selection. Most importantly, we report here the first study of the role of regulatory mutations during the degeneration process of mammalian Y, and we provide empirical data showing their degenerative effect on gene expression. Such mutations might further accelerate the degeneration and give rise to the evolution of dosage compensation in a gene-by-gene manner. These results demonstrate that mammalian Y degeneration is a complex gradual process spanning diverse genomic structures.

## Materials and methods

### FISH, separation of 1p+4 by cell Sorter, PCR, and sequence analysis

Frozen kidney from a male black muntjac, and fibroblast cell lines of another male (KCB82001) and a female black munt-

jac (KCB81002E) were provided by Kunming Cell Bank of the Chinese Academy of Sciences. Total genomic DNA was extracted using the PURRGENE® DNA Isolation Kit (Gentra Systems Inc., Minneapolis, MN USA), and total RNA was extracted using RNeasy® Mini Kit (Qiagen, Valencia, CA USA). Based on the syntenic relationship among black muntjac, Indian muntjac, and cow and annotation results of Indian muntjac [35], we selected genes randomly distributed along the neo-Y chromosome, PAR region, and autosomal regions for primer designs (Table 3). The investigated noncoding regions are introns or flanking intergenic regions of these selected genes. We prepared cell suspensions of the male black muntjac for further FISH and flow sorting of 1p+4 chromosomes. The sorting procedure using a FACStar Plus flow sorter (Becton Dickinson, Franklin Lakes, NJ USA) and FISH analysis using Indian muntjac BAC clones (o6G7 and o7B3) as probes were performed as previously described [21].

We used two strategies to discriminate neo-X and neo-Y alleles. First, we compared genomic DNA PCR products of Indian muntjac, and female and male black muntjacs, and we inferred neo-Y-specific alleles. We also used flow-sorted 1p+4 chromosomes as PCR templates to confirm the neo-Y-specific mutations. The products were subject to sequencing with BigDye Terminators v3.0 (Applied Biosystems) after purification (QIAquick® PCR Purification Kit; Qiagen, Valencia, CA USA). Trace data were manually trimmed and aligned from both directions using Lasergene suite (DNASTAR Inc., Madison, WI USA). We concatenated sequences of noncoding regions together for further analysis. Segregating sites in these regions were counted and polymorphism data were analyzed using DnaSP 4.0 [62]. Confidence intervals for polymorphisms were calculated by variance [63].

If we assume the 1p+4 chromosome can recombine with its homologous chromosome, then its polymorphism level is expected to be similar to that of other autosomes or 'PAR' regions. Under such a null hypothesis, we test the significance of difference for polymorphism between 1p+4 and other chromosome regions with  $\chi^2$  test. Sequences of neo-Y and neo-X alleles were subjected to construction of phylogenetic trees and distance calculation with MEGA 3.1 [64]. After removing regions with indels, gene tree was constructed using the neighbour joining and maximum likelihood method, and distances were calculated with Kimura's two-parameter and HKY85 model, respectively [65].

We used internal, external, and summed branch lengths in Figure 2a to estimate  $\alpha$ . Comparison of  $\alpha$  estimates from different branch lengths can be used to test whether the ancestral polymorphism have affected the calculation [31]. According to the method proposed by Miyata and coworkers [29],  $Y/X = 3\alpha/(2 + \alpha)$ , where Y and X stand for mutation rate for Y-linked and X-linked sequences, respectively. The variance of Y is  $V(Y) = Y(1 - Y)/(L [1 - 4Y/3]^2)$  and the variance of X is  $V(X) = Y(1 - X)/(L [1 - 4X/3]^2)$ , where L is the length of the



sequence [31]. The variance of  $Y/X$  is  $V(Y/X) = V(Y)/E(X)^2 + E(Y)^2V(X)/E(X)^4$ . The 95% confidence interval of  $\alpha$  was estimated, following the method proposed by Huang and coworkers [30]. Because  $Y/X(0.001189/0.000294$ , internal branches) is greater than 3 based on branch lengths of neo-Y and neo-X in this study, the  $\alpha$  value approaches infinity. To estimate expected stop codon number  $E(S)$  in the investigated coding regions, we assume that only one substitution would arise per codon, given the extremely young age of this neo-sex system. Assuming the chances of mutation are equal across the sequences,  $E(S)$  would be given by the following:

$E(S) = 3\mu t \sum_{i=1}^n p_i$ , where  $\mu$  is the estimated mammalian mutation rate measured per base [39],  $t$  is the age of the neo-sex system,  $n$  is the number of the codon in the investigated region and  $p_i$  is the probability of a certain codon becoming a stop codon.

### Promoter assay

We predicted the TSS information for black muntjac genes with ClustalW [66], using orthologous human, mouse, and cow sequence as reference. Regions expanding 1 kb upstream and 500 bp downstream of the predicted TSS were amplified as putative promoter regions in black muntjacs. The PCR products were subject to TA cloning using pMD19-T vector (TaKaRa, Dalian, Shangdong China) for screening products without PCR artifacts and further enzyme digestion. The digested PCR product was cloned into pGL3-Basic firefly luciferase vector (Promega, Madison, WI USA) using T4 ligase (Invitrogen, Carlsbad, CA USA). Then the vector was transformed into competent cell (TOP10; Invitrogen, Carlsbad, CA USA) and purified by QIAprep® Spin Miniprep Kit (Qiagen, Valencia, CA USA).

Hela and male black muntjac fibroblast cell lines were respectively seeded into 24-well plates to reach densities of  $10^4$  to  $10^5$  cells per well 24 hours before transfection. Generally, 0.2  $\mu$ g pGL3-Basic vector with cloned promoter and 0.04  $\mu$ g *Renilla* (sea-pansy) luciferase vector (pRL-TK; Promega, Madison, WI USA) were mixed in Opti-MEM® I Reduced Serum Medium (Invitrogen, Carlsbad, CA USA) and co-transfected into cells with Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA USA). We used pGL-Basic vector as negative control and assayed the luciferase signal with the Dual-Luciferase® Reporter 1000 Assay System (Promega, Madison, WI USA). The ratios of firefly luciferase to sea pansy luciferase were compared between neo-Y-linked promoters and neo-X-linked promoters. All of the assays for each promoter were done in triplicate.

### Expression abundance analysis

We amplified coding regions containing neo-Y-linked mutations from genomic DNA and total RNA (One Step RNA PCR Kit, TaKaRa, Dalian, Shangdong China), respectively. The genomic PCR was used as a control to exclude PCR biases caused by allelic preferential amplification. Out of 11 genes,

we observed biased amplification for only two genes, namely *SCN1A* and *SNX22*; the extent of the amplification bias is similar to the copy abundance difference observed in the mRNA assay. The reaction was done for 25 to 30 cycles, and the PCR products were cloned into pMD19-T vector (TaKaRa, Dalian, Shangdong China). Twenty to 30 white clones were randomly picked and were subject to further sequencing. The expression ratio of neo-Y and neo-X alleles was determined by counting the number of separate alleles. All assays were conducted in duplicate and double checked in both male individuals. We further compared the mean expression ratio of duplicates for different genes.

### Abbreviations

BAC, bacterial artificial chromosome; bp, base pair; FISH, fluorescence *in situ* hybridization; kb, kilobases;  $N_e$ , effective population size; PAR, pseudoautosomal region; PCR, polymerase chain reaction; TSS, transcriptional start site; UTR, untranslated region.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

WW and FY designed the project, analyzed the data, and wrote the paper. QZ, JW, and LH performed the PCR experiments, and expression and promoter assays. WN, JW, YL, and XZ cultivated the fibroblast cells and performed the FISH experiments. QZ and JW analyzed the data and wrote the paper. All authors read and approved the final manuscript.

### Additional data files

The following additional data are available with this paper. Additional data file 1 is a schematic ideogram showing the neo-sex system of *Muntiacus crinifrons*. Additional data file 2 is a description with a table showing substitution patterns on noncoding sequences of neo-Y and neo-X. Additional data file 3 is a table showing neo-Y linked mutations in promoter regions and their effects on allelic expression. Additional data file 4 is a figure showing the RNA editing process of neo-X allele in 3'-UTR of *SNX22*. Additional data file 5 is a table showing the RNA editing degree of *SNX22* neo-X allele in male black muntjacs is higher than that in female.

### Acknowledgements

We thank B Charlesworth, D Charlesworth, D Bachtrog, S Roy, K Zeng, and X Li for incisive discussions during preparing the manuscript. We thank W Su and H Yu for assistance during cultivating muntjac fibroblast cells. We are also thankful to two anonymous reviewers for their constructive comments on the manuscript. This work was supported by a CAS-Max Planck Society Fellowship, a NSFC award (No. 30325016), a NSFC key grant (No. 30430400), a 973 Program (No. 2007CB815703-5), and a key project of the

State Key Laboratory of Genetic Resources and Evolution of China to WW.

## References

- Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, Chinwalla A, Delehaunty A, Delehaunty K, Du H, Fewell G, Fulton L, Fulton R, Graves T, Hou SF, Latrielle P, Leonard S, Mardis E, Maupin R, McPherson J, Miner T, Nash W, Nguyen C, Ozersky P, Pepin K, Rock S, et al: **The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes.** *Nature* 2003, **423**:825-822.
- Graves JA: **Evolution of the testis-determining gene: the rise and fall of SRY.** *Novartis Found Symp* 2002, **244**:86-97. discussion 97-101, 203-106, 253-107.
- Graves JA: **Sex chromosome specialization and degeneration in mammals.** *Cell* 2006, **124**:901-914.
- Rens W, O'Brien P, Grutzner F, Clarke O, Graphodatskaya D, Tsend-Ayush E, Trifonov V, Skelton H, Wallis M, Johnston S, Veyrunes F, Graves J, Ferguson-Smith M: **The multiple sex chromosomes of platypus and echidna are not completely identical and several share homology with the avian Z.** *Genome Biol* 2007, **8**:R243.
- Lahn BT, Page DC: **Four evolutionary strata on the human X chromosome.** *Science* 1999, **286**:964-967.
- Charlesworth B, Charlesworth D: **The degeneration of Y chromosomes.** *Philos Trans R Soc Lond B Biol Sci* 2000, **355**:1563-1572.
- Ross MT, Graffham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP, Frankish A, Lovell FL, Howe KL, Ashurst JL, Fulton RS, Sudbrak R, Wen G, Jones MC, Hurler ME, Andrews TD, Scott CE, Searle S, Ramser J, Whittaker A, Deadman R, Carter NP, Hunt SE, Chen R, Cree A, Gunaratne P, et al: **The DNA sequence of the human X chromosome.** *Nature* 2005, **434**:325-337.
- Nicolas M, Marais G, Hykelova V, Janousek B, Laporte V, Vyskot B, Mouchiroud D, Negrutiu I, Charlesworth D, Moneger F: **A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants.** *PLoS Biol* 2005, **3**:e4.
- Steinemann M, Steinemann S: **Enigma of Y chromosome degeneration: neo-Y and neo-X chromosomes of *Drosophila miranda* a model for sex chromosome evolution.** *Genetica* 1998, **102-103**:409-420.
- Bachtrog D, Charlesworth B: **Reduced adaptation of a non-recombining neo-Y chromosome.** *Nature* 2002, **416**:323.
- Liu Z, Moore PH, Ma H, Ackerman CM, Ragiba M, Yu Q, Pearl HM, Kim MS, Charlton JW, Stiles JL, Zee FT, Paterson AH, Ming R: **A primitive Y chromosome in papaya marks incipient sex chromosome evolution.** *Nature* 2004, **427**:348-352.
- Bachtrog D: **Expression profile of a degenerating neo-Y chromosome in *Drosophila*.** *Curr Biol* 2006, **16**:1694-1699.
- Birney E, Stamatoyanopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhani P, Dillon SC, Dorschner MO, Fiegler H, et al: **Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project.** *Nature* 2007, **447**:799-816.
- Wittkopp PJ, Haerum BK, Clark AG: **Evolutionary changes in cis and trans gene regulation.** *Nature* 2004, **430**:85-88.
- Ohno S: *Sex Chromosomes and Sex Linked Genes* Berlin, Germany: Springer Verlag; 1967.
- Zarkower D: **Establishing sexual dimorphism: conservation amidst diversity?** *Nat Rev Genet* 2001, **2**:175-185.
- Gethmann RC: **Crossing over in males of higher Diptera (*Brachycera*).** *J Hered* 1988, **79**:344-350.
- Charlesworth B: **Model for evolution of Y chromosomes and dosage compensation.** *Proc Natl Acad Sci USA* 1978, **75**:5618-5622.
- Plath K, Mlynarczyk-Evans S, Nusinow DA, Panning B: **Xist RNA and the mechanism of X chromosome inactivation.** *Annu Rev Genet* 2002, **36**:233-278.
- Ohtaishi N, Gao YT: **A review of the distribution of all species of deer (*Tragulidae*, *Moschidae* and *Cervidae*) in China.** *Mammal Rev* 1990, **20**:125-144.
- Yang F, O'Brien PC, Wienberg J, Ferguson-Smith MA: **Evolution of the black muntjac (*Muntiacus crinifrons*) karyotype revealed by comparative chromosome painting.** *Cytogenet Cell Genet* 1997, **76**:159-163.
- Huang L, Chi J, Wang J, Nie W, Su W, Yang F: **High-density comparative BAC mapping in the black muntjac (*Muntiacus crinifrons*): molecular cytogenetic dissection of the origin of MCR 1p+4 in the X1X2Y1Y2Y3 sex chromosome system.** *Genomics* 2006, **87**:608-615.
- Wang W, Lan H: **Rapid and parallel chromosomal number reductions in muntjac deer inferred from mitochondrial DNA phylogeny.** *Mol Biol Evol* 2000, **17**:1326-1333.
- Ma K, Shi LM: **Comparative studies on synaptonemal complexes in spermatocytes of Chinese muntjac *Muntiacus reevesi*, black muntjac *M. crinifrons* and Indian muntjac *M. muntjak*.** *Yi Chuan Xue Bao* 1988, **15**:282-289.
- Yang F, Carter NP, Shi L, Ferguson-Smith MA: **A comparative study of karyotypes of muntjacs by chromosome painting.** *Chromosoma* 1995, **103**:642.
- Kimura M: *The Neutral Theory of Molecular Evolution* Cambridge, UK: Cambridge University Press; 1983.
- Caballero A: **On the effective size of populations with separate sexes, with particular reference to sex-linked genes.** *Genetics* 1995, **139**:1007-1011.
- Li WH, Yi S, Makova K: **Male-driven evolution.** *Curr Opin Genet Dev* 2002, **12**:650-656.
- Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T: **Male-driven molecular evolution: a model and nucleotide sequence analysis.** *Cold Spring Harb Symp Quant Biol* 1987, **52**:863-867.
- Huang W, Chang BH, Gu X, Hewett-Emmett D, Li W: **Sex differences in mutation rate in higher primates estimated from AMG intron sequences.** *J Mol Evol* 1997, **44**:463-465.
- Makova KD, Li WH: **Strong male-driven evolution of DNA sequences in humans and apes.** *Nature* 2002, **416**:624-626.
- Berlin S, Brandstrom M, Backstrom N, Axelsson E, Smith NG, Ellegren H: **Substitution rate heterogeneity and the male mutation bias.** *J Mol Evol* 2006, **62**:226-233.
- Huttley GA, Jakobsen IB, Wilson SR, Easteal S: **How important is DNA replication for mutagenesis?** *Mol Biol Evol* 2000, **17**:929-937.
- Laporte V, Charlesworth B: **Effective population size and population subdivision in demographically structured populations.** *Genetics* 2002, **162**:501-519.
- Zhou Q, Huang L, Zhang J, Zhao X, Zhang Q, Song F, Chi J, Yang F, Wang W: **Comparative genomic analysis links karyotypic evolution with genomic evolution in the Indian Muntjac (*Muntiacus muntjak vaginalis*).** *Chromosoma* 2006, **115**:427-436.
- Xie X, Lu J, Kulbokas EJ, Golub TR, Mootha V, Lindblad-Toh K, Lander ES, Kellis M: **Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals.** *Nature* 2005, **434**:338-345.
- Kuersten S, Goodwin EB: **The power of the 3' UTR: translational control and development.** *Nat Rev Genet* 2003, **4**:626-637.
- Bachtrog D: **Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in *Drosophila*.** *Genome Res* 2005, **15**:1393-1401.
- Kumar S, Subramanian S: **Mutation rates in mammalian genomes.** *Proc Natl Acad Sci USA* 2002, **99**:803-808.
- Haag-Liautaud C, Dorris M, Maside X, Macaskill S, Halligan DL, Charlesworth B, Keightley PD: **Direct estimation of per nucleotide and genomic deleterious mutation rates in *Drosophila*.** *Nature* 2007, **445**:82-85.
- Wood RD: **DNA repair in eukaryotes.** *Annu Rev Biochem* 1996, **65**:135-167.
- Bartolome C, Maside X, Yi S, Grant AL, Charlesworth B: **Patterns of selection on synonymous and nonsynonymous variants in *Drosophila miranda*.** *Genetics* 2005, **169**:1495-1507.
- Vicoso B, Charlesworth B: **Evolution on the X chromosome: unusual patterns and processes.** *Nat Rev Genet* 2006, **7**:645-653.
- Orr HA, Kim Y: **An adaptive hypothesis for the evolution of the Y chromosome.** *Genetics* 1998, **150**:1693-1698.
- Wray GA: **The evolutionary significance of cis-regulatory mutations.** *Nat Rev Genet* 2007, **8**:206-216.
- Borneman AR, Gianoulis TA, Zhang ZD, Yu H, Rozowsky J, Seringhaus MR, Wang LY, Gerstein M, Snyder M: **Divergence of transcription factor binding sites across related yeast species.**

- Science 2007, **317**:815-819.
47. Cooper SJ, Trinklein ND, Anton ED, Nguyen L, Myers RM: **Comprehensive analysis of transcriptional promoter structure and function in 1% of the human genome.** *Genome Res* 2006, **16**:1-10.
  48. Rusinov V, Baev V, Minkov IN, Tabler M: **MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence.** *Nucl Acids Res* 2005, **33**:W696-W700.
  49. Dodge GR, Kovalszky I, McBride OW, Yi HF, Chu ML, Saitta B, Stokes DG, Iozzo RV: **Human clathrin heavy chain (CLTC): partial molecular cloning, expression, and mapping of the gene to human chromosome 17q11-qter.** *Genomics* 1991, **11**:174-178.
  50. Song J, Zhao KQ, Newman CL, Vinarov DA, Markley JL: **Solution structure of human sorting nexin 22.** *Protein Sci* 2007, **16**:807-814.
  51. Chen K, Rajewsky N: **The evolution of gene regulation by transcription factors and microRNAs.** *Nat Rev Genet* 2007, **8**:93-103.
  52. Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH: **Why highly expressed proteins evolve slowly.** *Proc Natl Acad Sci USA* 2005, **102**:14338-14343.
  53. Bachtrog D: **A dynamic view of sex chromosome evolution.** *Curr Opin Genet Dev* 2006, **16**:578-585.
  54. Bartolome C, Charlesworth B: **Evolution of amino-acid sequences and codon usage on the *Drosophila miranda* neo-sex chromosomes.** *Genetics* 2006, **174**:2033-2044.
  55. Marin I, Franke A, Bashaw GJ, Baker BS: **The dosage compensation system of *Drosophila* is co-opted by newly evolved X chromosomes.** *Nature* 1996, **383**:160-163.
  56. Carrel L, Willard HF: **X-inactivation profile reveals extensive variability in X-linked gene expression in females.** *Nature* 2005, **434**:400-404.
  57. Lee J, Davidow LS, Warshawsky D: **Tsix, a gene antisense to Xist at the X-inactivation centre.** *Nat Genet* 1999, **21**:400-404.
  58. Sado T, Wang Z, Sasaki H, Li E: **Regulation of imprinted X-chromosome inactivation in mice by Tsix.** *Development* 2001, **128**:1275-1286.
  59. Fisher RA: *The Genetical Theory of Natural Selection* Oxford, UK: Clarendon Press; 1930.
  60. Gupta V, Parisi M, Sturgill D, Nuttall R, Doctolero M, Dudko OK, Malley JD, Eastman PS, Oliver B: **Global analysis of X-chromosome dosage compensation.** *J Biol* 2006, **5**:3.
  61. Nguyen DK, Disteché CM: **Dosage compensation of the active X chromosome in mammals.** *Nat Genet* 2006, **38**:47-53.
  62. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R: **DnaSP, DNA polymorphism analyses by the coalescent and other methods.** *Bioinformatics* 2003, **19**:2496-2497.
  63. Tajima F: **Measurement of DNA polymorphism.** In *Mechanisms of Molecular Evolution* Sunderland, UK: Sinauer Associates. Inc; 1993:37-59.
  64. Tamura K, Dudley J, Nei M, Kumar S: **MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0.** *Mol Biol Evol* 2007, **24**:1596-1599.
  65. Yang Z: **PAML: a program package for phylogenetic analysis by maximum likelihood.** *Comput Appl Biosci* 1997, **13**:555-556.
  66. Thompson JD, Higgins DG, Gibson TJ: **CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice.** *Nucleic Acids Res* 1994, **22**:4673-4680.