Ancient Male Recombination Shaped Genetic Diversity of Neo-Y Chromosome in *Drosophila albomicans*

Kazuhiro Satomura[1] and Koichiro Tamura*[,1,2]

[1]Department of Biological Sciences, Tokyo Metropolitan University, Tokyo, Japan

[2]Research Center for Genomics and Bioinformatics, Tokyo Metropolitan University, Tokyo, Japan

***Corresponding author:** E-mail: ktamura@tmu.ac.jp.

Associate editor: Yoko Satta

Abstract

Researchers studying Y chromosome evolution have drawn attention to neo-Y chromosomes in *Drosophila* species due to their resembling the initial stage of Y chromosome evolution. In the studies of neo-Y chromosome of *Drosophila miranda*, the extremely low genetic diversity observed suggested various modes of natural selection acting on the nonrecombining genome. However, alternative possibility may come from its peculiar origin from a single chromosomal fusion event with male achiasmy, which potentially caused and maintained the low genetic diversity of the neo-Y chromosome. Here, we report a real case where a neo-Y chromosome is in transition from an autosome to a typical Y chromosome. The neo-Y chromosome of *Drosophila albomicans* harbored a rich genetic diversity comparable to its gametologous neo-X chromosome and an autosome in the same genome. Analyzing sequence variations in 53 genes and measuring recombination rates between pairs of loci by cross experiments, we elucidated the evolutionary scenario of the neo-Y chromosome of *D. albomicans* having high genetic diversity without assuming selective force, i.e., it originated from a single chromosomal fusion event, experienced meiotic recombination during the initial stage of evolution and diverged from neo-X chromosome by the suppression of recombination tens or a few hundreds of thousand years ago. Consequently, the observed high genetic diversity on the neo-Y chromosome suggested a strong effect of meiotic recombination to introduce genetic variations into the newly arisen sex chromosome.

Key words: **meiotic recombination, achiasmy, sex chromosome, founder effect.**

Introduction

The effect of meiotic recombination on genetic diversity has been well characterized by a number of studies on Y chromosome evolution, as meiotic recombination is suppressed between X and Y chromosomes to maintain their functional diversification on sex determination (Nei 1969). These studies have shown that genetic diversity of Y chromosome is much lower than that of X chromosome in various species (Whitfield et al. 1995, Berset-Brändli et al. 2007, Rozen et al. 2009, Muir et al. 2011, Larracuente and Clark 2013, Wilson Sayres et al. 2014). A variety of evolutionary scenarios involving natural selection under the suppression of meiotic recombination have been invoked to explain the low genetic diversity of Y chromosome (e.g., Bachtrog 2013 for review). Among these studies, an extremely low genetic diversity was explored in the neo-Y chromosome of *Drosophila miranda*, suggesting recent positive selection acting on the nonrecombining genome (Bachtrog 2004). The neo-Y chromosome was originated by a fusion of ancestral Y chromosome and an autosome, resulting in a situation where a huge number of genes from the fused autosome are shared by its gametologous neo-X chromosome. This situation gives us an excellent opportunity to compare evolutionary changes between neo-X and neo-Y chromosomes in the same genes.

However, we should notice that neo-Y chromosome evolution does not necessarily mirror Y chromosome evolution in achiasmatic species, for example, the *Drosophila* species, in which meiotic recombination is completely suppressed in males. In this case, meiotic recombination is shut out immediately after fusion due to the switch of inheritance mode from biparental to paternal, resulting in a situation where a newly generated neo-Y chromosome has no chance to acquire preexisting genetic variations from its gametologous chromosome by meiotic recombination and remains monomorphic until newly occurring mutations accumulate over time. This is a big difference from typical Y chromosome evolution, where meiotic recombination is suppressed gradually so that young Y chromosome has a chance to acquire preexisting genetic variations from the original autosome during the initial stage of evolution. Therefore, the observed low genetic diversity of the young neo-Y chromosome may be largely attributable to the founder effect caused by a single chromosomal fusion event. Additionally, the smaller effective population size may be partly responsible for the lower genetic diversity of neo-Y chromosome. Because a pair of parents has one Y chromosome, three X chromosomes and four individual autosomes, the genetic diversity of Y chromosome is expected to eventually become one-third of that of X chromosome and one-fourth of that of autosome. On the other hand, this smaller population size may relax synonymous codon usage bias to increase synonymous substitutions. Furthermore, if Y chromosome evolves faster than X chromosome and autosomes due to differences in germline cell production between males and females as known as

Article

"male-driven evolution" (Miyata et al. 1987), this works for higher genetic diversity of neo-Y chromosome. Therefore, we should take all these factors into account to examine the effect of natural selection on the genetic diversity of neo-Y chromosome.

To examine how natural selection acts on genetic diversity during Y chromosome evolution, it is necessary to study a neo-Y chromosome that was recombining initially but suppressed recombination gradually during the course of evolution. In this article, we report the case of neo-Y chromosome evolution in *Drosophila albomicans* as such an example. *D. albomicans* is a member of the *Drosophila nasuta* subgroup of the *Drosophila immigrans* species group in the subgenus *Drosophila*. The ancestor of *D. albomicans* is thought to have canonical X and Y sex chromosomes and three pairs of autosomes (second, third, and fourth chromosomes) as do *D. nasuta* and other species in the *D. nasuta* subgroup at present (fig. 1*A*). However, the neo-X and neo-Y chromosomes that current *D. albomicans* has are thought to have originated by two independent fusions of X and third chromosomes and of Y and third chromosomes, respectively (fig. 1*B*, Wilson et al. 1969, Wakahama et al. 1983). Although the male recombination is perfectly suppressed at present in *D. albomicans*, we show that it likely occurred in the ancestral population, which resulted in high genetic diversity of the neo-Y chromosome in this species.

Results and Discussion

Genetic Diversity of Neo-Y Chromosome in *D. albomicans*

The neo-Y chromosome of *D. albomicans* is known to be much younger than that of *D. miranda* (Bachtrog and Charlesworth 2002, Bachtrog 2006). This allows us to expect that the *D. albomicans* neo-Y chromosome harbors less genetic diversity than the *D. miranda* neo-Y chromosome under the assumptions that the neo-Y chromosomes in both species had a single origin and have not recombined. In this condition, only a way of neo-Y chromosome's acquiring genetic diversity is mutation and the younger neo-Y chromosome has had shorter time to accumulate mutations. To examine this expectation, we determined the nucleotide sequences of 27 genes on the neo-sex chromosomes and 26 genes on the second chromosome for 16 wild strains of *D. albomicans* as well as for 16 wild strains of *D. nasuta*, a sibling species of *D. albomicans* having no neo-sex chromosome but the canonical sex chromosomes and the third chromosome separately. The result showed an unexpectedly high genetic diversity of the neo-Y linked genes in *D. albomicans* (table 1, see supplementary tables S1–S5, Supplementary Material online, for details). The level of genetic diversities by means of nucleotide diversity, π (Nei and Tajima 1981) and Watterson's θ (Watterson 1975), were very similar between the homologous genes on the neo-X and neo-Y chromosomes in *D. albomicans* as well as between those and the genes on the second chromosome at either synonymous or intronic sites. The differences are statistically nonsignificant, except for the difference in θ for intronic sites between

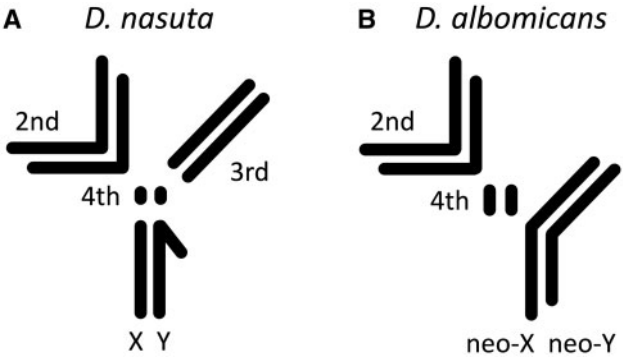

FIG. 1. Schematic representation of male karyotypes of *D. nasuta* (*A*) and *D. albomicans* (*B*).

second and neo-Y chromosomes ($P < 0.05$). The situation is virtually the same in *D. nasuta*, in which the genetic diversities of the genes on the unfused third chromosome were very similar to those on the second chromosome and the differences are all statistically nonsignificant.

Given that genes on Y-chromosome are expected to be degraded during the course of evolution, the observed high genetic diversity on the neo-Y chromosome may be explained by relaxed functional constraints on nonfunctionalized genes. However, this is implausible in the present case because of the following two reasons. First, there are more synonymous segregating sites (*n*S) than nonsynonymous segregating sites (*n*N) in spite of much fewer synonymous sites than nonsynonymous sites in all the 27 genes on the neo-Y chromosome (supplementary table S4, Supplementary Material online). The *n*N/*n*S ratio for the neo-Y linked genes is no greater than that for the neo-X linked genes in 16 out of 27 genes (supplementary tables S3 and S4, Supplementary Material online). The result does not change if we use the ratio between synonymous and nonsynonymous π's instead of *n*S and *n*N. Second, analysis of variance (ANOVA) for the effective number of codons (ENCs) (Wright 1990) indicates no significant difference in synonymous codon usage bias between the neo-X and neo-Y chromosomes, whereas the variation among genes is highly significant with $P < 0.001$ (table 2, see supplementary tables S6 and S7, Supplementary Material online, for details). These results suggest that there is no difference in the functional constraints between the neo-X and neo-Y linked genes, which is consistent with the inference based on a transcriptome analysis that only 1–2% of genes on the neo-Y chromosome of *D. albomicans* have been nonfunctionalized (Zhou and Bachtrog 2012).

Another candidate to explain the high genetic diversity is a higher rate of point mutation on the neo-Y chromosome than on the neo-X chromosome. To test this hypothesis, we compared the average number of synonymous substitutions per site between the neo-X and neo-Y linked genes with their homologous genes on the *D. nasuta* third chromosome (table 3). The average values are 0.0337 ± 0.0035 and 0.0356 ± 0.0042 for the neo-X and neo-Y genes, respectively, where the value for the neo-Y chromosome is higher than that for the neo-X chromosome in 17 out of 27 genes. Although the neo-Y linked genes show slightly faster

Table 1. Nucleotide Diversity (π), Watterson's θ, and Tajima's *D* at Synonymous Sites and Intronic Sites in 53 Nuclear Genes[a] of *D. albomicans* and *D. nasuta*.

Chromosome		No. of Genes	No. of Sites	No. of Segregating Sites	$\pi \pm$ SE ($\times 100$)	$\theta \pm$ SE ($\times 100$)	Tajima's $D \pm$ SE
D. albomicans							
Second	Synonymous sites	26	4,169	361	2.21 ± 0.33	2.51 ± 0.34	-0.47 ± 0.15**
	intron	23	3,680	203	2.24 ± 0.39	2.34 ± 0.33	-0.38 ± 0.20
Neo-X	Synonymous sites	27	4,511	280	1.77 ± 0.25	1.87 ± 0.25	-0.17 ± 0.20
	intron	15	1,819	125	1.59 ± 0.33	1.98 ± 0.45	-0.50 ± 0.22*
Neo-Y	Synonymous sites	27	4,512	233	1.80 ± 0.25	1.54 ± 0.21	0.63 ± 0.22**
	intron	15	1,855	85	1.40 ± 0.34	1.36 ± 0.33	0.26 ± 0.29
D. nasuta							
Second	Synonymous sites	26	4,167	294	1.91 ± 0.32	1.96 ± 0.30	-0.21 ± 0.18
	intron	23	3,678	178	1.87 ± 0.33	1.85 ± 0.29	-0.11 ± 0.18
Third	Synonymous sites	27	4,513	324	2.13 ± 0.34	2.14 ± 0.33	-0.08 ± 0.19
	intron	15	1,850	127	1.69 ± 0.56	1.86 ± 0.65	-0.23 ± 0.24

[a]*Adh, Apf, Arr1, bur, CaBP1, Crc, Ddc, Droj2, eIF-3, p40, fbp, Fdh, Gpdh, Oscp, porin, Rga, RpL4, Set, SPE, Spn88Ea, srp, Surf4, Tps1, und, CG3609, CG5028, CG6255* on the second chromosome and *aay, Amyrel, blw, ERp60, GstS1, Idh, Jheh3, Mpcp, Mys45A, NHP2, Nurf-38, Pdh, Pdi, Pgm, rept, RnrS, Rpn1, SdhB, Srp68, Tango7, Thiolase, Uch-L3, wal, CG5068, CG7430, CG8446, CG33138* on the neo-sex (*D. albomicans*) or the third (*D. nasuta*) chromosomes. *The value is different from zero with a statistical significance ($P < 0.05$). **The value is different from zero with a statistical significance ($P < 0.01$).

Table 2. Results of ANOVA for ENC among 27 Neo-X and Neo-Y Linked Genes from 16 *D. albomicans* Strains.

Source of Variation	DF	SS	MS	*F*	*P*
neo-X vs. neo-Y	1	0.0266	0.0266	0.0472	0.828
Among genes	26	35527.97	1366.460	2422.445	<0.001
Among strains	15	4.773	0.318	0.564	0.903
Residual	821	463.112	0.564		
Total	863	35995.880	41.710		

evolutionary rate than the neo-X linked genes, the difference is not significant ($P < 0.31$), indicating that the higher rate of point mutation is not a major factor of the high genetic diversity on the neo-Y chromosome.

The observed high genetic diversity on the neo-Y chromosome of *D. albomicans* is in clear contrast to the case of *D. miranda*, where the level of nucleotide variability on the neo-Y chromosome is 30 times lower than that in highly recombining portions of the genome (Bachtrog 2004). The observed high genetic diversity of the neo-Y linked genes suggests that at least either the assumptions of a single origin or no recombination does not hold for the neo-Y chromosome evolution in *D. albomicans*. If we assume multiple origins without recombination to explain the observed neo-Y diversity at a very similar level to the autosome diversity, we need to assume that most of the neo-Y chromosomes were produced by independent chromosome fusions. This is unrealistic because chromosomal fusions have been obviously rare events in the evolutionary history of *Drosophila* species, as suggested by well-conserved karyotypes among closely related species (Ashburner 1989, chapter 5-1).

Meiotic Recombination in *D. albomicans* and *D. nasuta*

The possibility of meiotic recombination in males was examined by cross experiments. The recombination was examined between *Arr1* and *Fdh* and between *Surf4* and *CaBP1* on the second chromosome and between *GstS1* and *Pgm* on the neo-sex (*D. albomicnas*) or third (*D. nasuta*) chromosomes, because the genotypes of these genes were distinguishable by single nucleotide polymorphisms (SNPs) used as a genotype marker. The occurrence of meiotic recombination in the male germline was examined by presence or absence of recombinants at the backcross first generation (BC_1) in three *D. albomicans* and three *D. nasuta* chromosomal inversion-free strain pairs, respectively.

The results showed that male recombination did not occur in *D. albomicans* but did in *D. nasuta* either in the second or neo-sex/third chromosome (table 4). Between *Arr1* and *Fdh* on the second chromosome, there was no recombinant from male F_1 hybrids in *D. albomicans* irrespective of the combination of parental strain pairs. Because there were only a few recombinants from the female F_1 hybrids, we additionally examined the recombination between *Surf4* and *CaBP1*. The result was the same: there were recombinants from the female F_1 hybrids but none from the male F_1 hybrids. In contrast, there were many recombinants between *Arr1* and *Fdh* from the male F_1 hybrids in *D. nasuta*, although the recombination rate was lower than that from the female F_1 hybrids. The male recombination in *D. nasuta* was recognized between *GstS1* and *Pgm* on the third chromosome as well. Again, there was no recombinant by male recombination in *D. albomicans*. These results consistently suggest that male meiotic recombination occurs in *D. nasuta* but not in *D. albomicans*. It should be noted that the observed differences in the recombination rate among different strain pairs may arise from the differences in gene arrangement among different chromosomal inversion types (see Materials and Methods for details).

Ancestral Male Recombination in *D. albomicans*

Although male recombination was not detected in *D. albomicans*, whether or not male recombination occurred in the ancestral *D. albomicans* population and/or the most recent common ancestor of *D. albomicans* and *D. nasuta* remains

Table 3. Number of Synonymous Substitutions Per Site and d_A Distance for 27 Genes among Neo-X and Neo-Y Chromosomes in *D. albomicans* and Third Chromosome in *D. nasuta*.

Gene	d_{XY}		d_A		
	Neo-X — nas	Neo-Y — nas	Neo-X — nas	Neo-Y — nas	Neo-X — neo-Y
aay	0.03444	0.04103	−0.00008	0.00612	0.00414
Amyrel	0.02453	0.01398	0.01793	0.00570	0.02766
blw	0.01958	0.01456	0.00387	0.00354	0.00874
ERp60	0.00830	0.00828	0.00135	0.00130	0.00120
GstS1	0.03505	0.02906	0.00691	0.00446	0.00336
Idh	0.04369	0.04843	0.01533	0.01858	−0.00081
Jheh3	0.04804	0.04839	0.01558	0.02023	0.00154
Mpcp	0.00895	0.00754	0.00359	0.00300	0.00001
Mys45A	0.00698	0.00711	−0.00183	−0.00120	0.00014
NHP2	0.01928	0.02260	0.00059	0.00474	0.00061
Nurf-38	0.04331	0.04377	0.03365	0.03627	−0.00001
Pdh	0.06347	0.09267	−0.00110	0.01723	0.02827
Pdi	0.02481	0.02605	0.00703	0.00645	−0.00095
Pgm	0.02337	0.02499	−0.00001	0.00425	0.00590
rept	0.03210	0.03070	0.01551	0.01526	−0.00049
RnrS	0.00844	0.00803	0.00531	0.00433	−0.00007
Rpn1	0.04685	0.04806	0.03100	0.03110	−0.00127
SdhB	0.05040	0.05462	0.01323	0.02218	0.00094
Srp68	0.04168	0.05306	0.01742	0.03173	0.00915
Tango7	0.04134	0.02696	0.03221	0.00190	0.02230
Thiolase	0.01911	0.04019	0.00452	0.01852	0.00524
Uch-L3	0.08157	0.08363	0.02716	0.04326	0.03007
wal	0.05834	0.06122	0.00889	0.01602	−0.00279
CG5068	0.03966	0.03538	0.02386	0.02251	−0.00069
CG7430	0.03008	0.02829	0.01230	0.01158	−0.00043
CG8446	0.01581	0.01650	0.00592	0.00094	0.00686
CG33138	0.04137	0.04652	−0.00269	0.00117	−0.00692
Average ± SE	0.03372 ± 0.00352	0.03562 ± 0.00423	0.01102 ± 0.00211	0.01301 ± 0.00234	0.00525 ± 0.00192

Table 4. Numbers of Recombinants and Nonrecombinants and Recombination Rates (%) Observed at Backcross First Generation in *D. albomicans* and *D. nasuta*.

		Numbers of Recombinants and Nonrecombinants and % Recombination Rates								
		Second						Neo-sex/Third		
Parent strains		*Arr1–Fdh*			*Surf4–CaBP1*			*GstS1–Pgm*		
♀	♂	rec	non-rec	%	rec	non-rec	%	rec	non-rec	%
D. albomicans										
HRS402	F_1(NG3 × HRS402)	0	32	0.0	0	32	0.0	0	32	0.0
F_1(NG3 × HRS402)	HRS402	3	29	9.4	5	27	15.6	17	15	53.1
TSN7	F_1(SZ43 × TSN7)	0	32	0.0	0	32	0.0	0	32	0.0
F_1(SZ43 × TSN7)	TSN7	2	30	6.3	1	31	3.2	16	16	50.0
KM7	F_1(HNZ12 × KM7)	0	32	0.0	0	32	0.0	0	32	0.0
F_1(HNZ12 × KM7)	KM7	0	32	0.0	5	27	15.6	15	17	46.9
D. nasuta										
VNS23	F_1(TNR35 × VNS23)	5	27	15.6	—	—	—	12	20	37.5
F_1(TNR35 × VNS23)	VNS23	10	22	31.3	—	—	—	9	23	28.1
C3	F_1(K33 × C3)	6	26	18.8	—	—	—	3	29	9.4
F_1(K33 × C3)	C3	11	21	34.4	—	—	—	10	22	31.3
MS77	F_1(BZV1 × MS77)	3	29	9.4	—	—	—	1	31	3.1
F_1(BZV1 × MS77)	MS77	8	24	25.0	—	—	—	12	20	37.5

unknown. To consider the possibility of the ancestral male recombination in *D. albomicans*, we examined the nucleotide sequence data in the following two ways.

First, we analyzed the shared SNP sites between the neo-X and neo-Y chromosomes of *D. albomicans* as well as between those and the third chromosome of *D. nasuta*. If the variations on the neo-Y chromosome were transmitted by meiotic recombination, many SNP sites are expected to be shared by the neo-Y, neo-X, and third chromosomes at present. Indeed, this is the case: 139 out of 280 and 233 synonymous SNP sites on the neo-X and neo-Y chromosomes, respectively, were common, which was much more than the amount expected from random point mutations occurred independently on both chromosomes, that is, $2 \times 280 \times 233/(4{,}511 + 4{,}512) = 14.5$ (fig. 2). In addition, the neo-Y chromosome shares as many SNP sites with the *D. nasuta* third chromosome as does the neo-X chromosome (107/233 vs. 117/280). This suggests that the neo-Y chromosome as well as the neo-X chromosome recombined with the unfused third chromosome before it reached fixation. It should be noted that the observed shared SNP sites did not come from the ancestral polymorphism without recombination, given that each of the current neo-X and neo-Y chromosome populations was descended from a single fused chromosome.

The second approach is based on the following expectation. If the neo-Y chromosome recombined with the neo-X and/or ancestral unfused third chromosomes in the initial stage of evolution, the effective population size of the neo-Y chromosome (together with its gametologous chromosomes) was four for a male and female pair of flies, but it has been reduced to be one by the male recombination suppression that isolated the population of neo-Y chromosomes from the population of neo-X chromosomes. In contrast, if male recombination has not occurred during the entire course of the neo-Y evolution, the effective population size should be increased as the neo-Y chromosome increased from a single (or at least few) origin to eventually reach fixation. To infer the change in the effective population size, we computed Tajima's *D* for the synonymous sites and the intronic sites (table 1). When population size is reduced, Tajima's *D* is expected to be positive and vice versa in the absence of the influence of natural selection (Tajima 1989). In *D. nasuta*, the value of Tajima's *D* is not significantly different from zero either for synonymous or intronic sites (table 1). In *D. albomicans*, the genes on the neo-X chromosome and the second chromosome showed negative values, in which the value from synonymous sites on the second chromosome was statistically significant from zero ($P < 0.01$). In contrast, the genes on the neo-Y chromosome showed positive values, in which the value from synonymous sites was statistically significant ($P < 0.01$). This positive value of Tajima's *D* is hardly explained by the effect of natural selection, because their homologous genes on the neo-X chromosome under the same functional constrains gave negative values. However, since Ohsako et al. (1994) reported a substantial population subdivision between the China-Southeast Asia population and the Japan–Taiwan population in *D. albomicans*, the obtained Tajima's *D* values might be largely influenced by

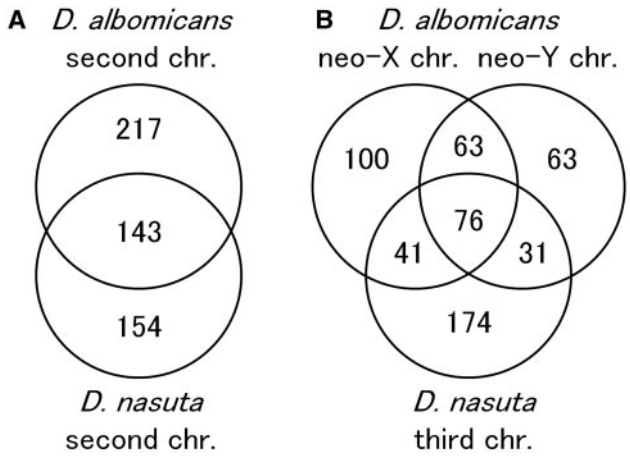

FIG. 2. Venn diagrams showing overlap of synonymous SNP sites between the *D. albomicans* and *D. nasuta* second chromosomes (*A*) and among the *D. albomicans* neo-X and neo-Y chromosomes and the *D. nasuta* third chromosome (*B*).

population structuring. To control this effect, we computed Tajima's *D* as well as π and θ for the strains from China and Southeast Asia and those from Japan and Taiwan separately (supplementary tables S8–S14, Supplementary Material online) and tested the difference in the Tajima's *D* values by ANOVA. Although the statistical significance was declined by the smaller sample size due to the population splitting, the tendency of Tajima's *D* value obtained from the entire population (negative for the neo-X and second and positive for the neo-Y) did not change. The ANOVA clearly shows that the difference between the neo-X and neo-Y linked genes is statistically significant as well as the difference among genes is, whereas the difference between the China–Southeast Asian population and the Japan–Taiwan population is nonsignificant (table 5). The significant difference among genes is attributed to the differences in functional constraints among genes as observed in the significant difference in ENC among genes (table 2). In contrast, the significant difference between the neo-X and neo-Y is not attributable to the difference in functional constraints, because there is no difference in functional constrains between the neo-X and neo-Y linked genes as discussed earlier. Therefore, it is concluded that the obtained positive values of Tajima's *D* for the neo-Y linked genes indicates a population size reduction of the neo-Y chromosome during the course of evolution most probably due to the suppression of male recombination.

To examine the sequence divergence between the neo-X and neo-Y chromosomes occurred after the suppression of male recombination, we computed the net number of substitutions per site (d_A, Nei 1987, pp. 276–277) for the 27 genes between the neo-X and neo-Y chromosomes from 16 *D. albomicans* strains and those and the third chromosome from 16 *D. nasuta* strains (table 3). As a result, the average net sequence divergence between the neo-X and neo-Y chromosomes was computed to be 0.00525 ± 0.00192, which is statistically significant from zero ($P < 0.01$), suggesting sequence divergence accumulated between these chromosomes after the suppression of male recombination. Assuming that the

Table 5. Results of ANOVA for Tajima's *D*.

Source of Variation	DF	SS	MS	*F*	*P*
neo-X vs. neo-Y	1	4.588	4.588	8.206	0.005
CS vs. JT	1	1.768	1.768	3.163	0.079
Among genes	26	34.710	1.335	2.388	0.002
Residual	79	44.169	0.559		
Total	107	85.236	0.797		

CS, China-Southeast Asia; JT, Japan-Taiwan

rate of neutral substitution to be 1.1×10^{-8} per year among *Drosophila* species (Tamura et al. 2004), we estimate the divergence time of the neo-X and neo-Y chromosomes as the time of recombination suppression to be ~240,000 years (with 95% confidence interval ranging from 70,000 to 410,000 years). Similarly, we estimate the time of *D. albomicans* and *D. nasuta* speciation to be about 546,000 years (with 95% confidence interval ranging from 344,000 to 748,000 years), which is significantly older than 120,000 years (with 95% confidence intervals ranging from 15,000 to 290,000 years) estimated by Bachtrog (2006). However, because the variation of time estimates among genes is considerably large (from 0 to 1.5 My) in both studies, the obtained difference in the average is attributed to the difference in the genes used. At any rate, the time of recombination suppression between the neo-X and neo-Y chromosomes was most probably much later than the speciation of *D. albomicans* and *D. nasuta*, supporting the scenario that the neo-Y chromosome recombined with the neo-X chromosome for a certain period of time and then ceased the recombination tens or a few hundreds of thousand years ago.

Conclusion

All the results presented in this article support the scenario that the neo-Y chromosome of *D. albomicans* originated from a fusion of an ancestral Y chromosome and a third chromosome, recombined with its gametologous chromosome during the initial stage of evolution and then ceased meiotic recombination tens or a few hundreds of thousand years ago (fig. 3). The currently observed high genetic diversity on the neo-Y chromosome suggests that the diversity was introduced by the ancient meiotic recombination effectively and has not been altered significantly after the cessation of meiotic recombination. This is a clear contrast to the case of *D. miranda*, where the neo-Y chromosome showed an extremely low genetic diversity. Although Bachtrog (2004) suggested a selective sweep causing the low genetic diversity, we alternatively propose a founder effect due to the peculiar generation by a chromosomal fusion event and the subsequent immediate cessation of meiotic recombination by achiasmy. In this study, we did not find any evidence of selective force acting on the genetic diversity on the neo-Y chromosome in *D. albomicans*. In contrast, the observed high genetic diversity clearly indicates a strong effect of meiotic recombination to introduce genetic diversity into the newly generated neo-Y chromosome. Yet, we still do not know about the driving force for the fixation of the neo-X and neo-Y chromosomes. We speculate that the fused chromosomes had some selective advantage from the facts that both neo-X and neo-Y have reached fixation within a short evolutionary time and that fusions of Muller's elements have occurred independently many times during the evolution of *Drosophila* species (Ashburner 1989, chapter 5-1). If this is the case, there must be a footprint of selective sweep detectable among the genomic nucleotide sequences of *D. albomicans*, which remains open to further genome-wide studies. Especially, data for the complete genome sequence of *D. albomicans* are indispensable to investigate the stratum structure accompanying gradual suppression of meiotic recombination during Y chromosome evolution.

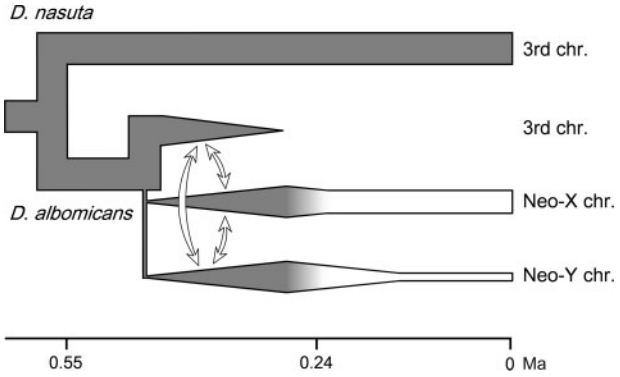

Fig. 3. Schematic diagram of the evolution of neo-X and neo-Y chromosomes in *D. albomicans*. The width of branches represents the population size and the gray color represents the meiotic recombination in males. After the speciation of *D. albomicans* and *D. nasuta* occurred 0.55 (0.34–0.75) million years ago (Ma), in *D. albomicans* the neo-X and neo-Y chromosomes emerged via chromosomal fusion events between the canonical third chromosome and the canonical X and Y chromosomes, respectively. Then, the neo-X and neo-Y chromosomes have gradually substituted for the preexisting third chromosome with recombining each other and with the third chromosome (indicated by crescent-shaped arrows) to share the polymorphisms at the sequence level. Finally, the male recombination was ceased 0.24 (0.07-0.41) Ma, resulting in the divergence of neo-X and neo-Y chromosomes and the reduction in the effective population size especially for neo-Y chromosome.

Materials and Methods

Drosophila Strains

Sixteen *Drosophila albomicans* and sixteen *D. nasuta* isofemale strains were used for the comparative sequence analyses. The *D. albomicans* strains were collected in six geographic locations in Japan, three in China, three in Taiwan, and four in Southeast Asia. The *D. nasuta* strains were collected in five geographic locations in India, two in Sri Lanka, two in Seychelles, two in Mauritius islands, two in Reunion islands, and three in Africa. The collection locality and year of these strains were listed in supplementary table S15, Supplementary Material online. These strains were cultured in the standard cornmeal medium at 20 °C.

DNA Sequences

We selected 53 genes (*aay, Adh, Amyrel, Apf, Arr1, blw, bur, CaBP1, CG3609, CG5028, CG5068, CG6255, CG7430, CG8446, CG33138, Crc, Ddc, Droj2, eIF-3p40, ERp60, fbp, Fdh, Gpdh,*

GstS1, *Idh*, *Jheh3*, *Mpcp*, *Mys45A*, *NHP2*, *Nurf-38*, *Oscp*, *Pdh*, *Pdi*, *Pgm*, *porin*, *rept*, *Rga*, *RnrS*, *RpL4*, *Rpn1*, *SdhB*, *Set*, *SPE*, *Spn88Ea*, *srp*, *Srp68*, *Surf4*, *Tango7*, *Thiolase*, *Tps1*, *Uch-L3*, *und*, *wal*), which were single copy and orthologous in the 12 Drosophila genomes (Drosophila 12 Genomes Consortium 2007), to determine the nucleotide sequences for the 16 *D. albomicans* strains and the 16 *D. nasuta* strains. Because many of these strains were heterozygous for many of the genes, we crossed a single male from each strain to a single female of an inbred *D. albomicans* strain, NG3, for which nucleotide sequences of all the genes were predetermined, and determined the nucleotide sequences of the neo-X and neo-Y linked genes from a single female and a single male of the F_1 hybrid, respectively.

The DNA fragments of each target gene were amplified by PCR with Ex Taq polymerase (Takara Co. Ltd.) from total DNA extracted from a whole body by a silica-gel based method following Boom et al. (1990). We newly designed 101 PCR primers and used them together with two (Gpdh F and Gpdh R) and seven (Amyrel F, Amyrel R, Apf F, Apf R, NHP2 F, NHP2 R and Pdh F) sets designed by Kwiatowski et al. (1997) and Chang et al. (2008), respectively, to amplify the DNA fragments for the 53 genes in total (supplementary table S16, Supplementary Material online). The resultant PCR fragments were purified by the silica-gel based method and the nucleotide sequences were determined by using BigDye Terminator Ready Reaction Cycle Sequencing Kit v1.1 or v3.1 and a 3130*xl* Genetic Analyzer (Applied Biosystems).

The nucleotide sequences obtained have been deposited to GenBank (accession numbers LC057726–LC059853).

Chromosomal Assignment of Genes

Whether each gene is located on the neo-sex or second chromosome was determined by the correspondence of chromosomal arms between *D. albomicans* and *Drosophila melanogaster*. According to Chang et al. (2008), the second chromosome of *D. albomicans* corresponds to Muller's elements B and E and the neo-Y chromosome corresponds to Muller's elements C and D. Therefore, genes on Muller's element B, that is, chromosomal arm 2L in *D. melanogaster*, were assumed to be on the second chromosome in *D. albomicans*. In this way, the chromosomal assignment was made for all the 53 genes. To confirm the chromosomal assignment, we further examined whether the pattern of inheritance was sex-linked or autosomal in 32 F_2 hybrids of X108 males and NG3 females for 22 genes using SNP makers obtained from the nucleotide sequences determined. As the result, we confirmed that ten genes (*Amyrel*, *blw*, *GstS1*, *NHP2*, *Pdi*, *Pgm*, *SdhB*, *CG5068*, *CG7430*, *CG33138*) on the third chromosome in *D. nasuta* or its corresponding part of the neo-X and neo-Y chromosomes in *D. albomicans* and 12 genes on the second chromosome (*Adh*, *Arr1*, *CaBP1*, *Crc*, *eIF-3p40*, *Fdh*, *Gpdh*, *RpL4*, *Surf4*, *Tps1*, *CG3609*, *CG5028*) (supplementary table S17, Supplementary Material online). For all the 22 genes, their chromosomal assignments were consistent with those expected from the correspondence to the chromosomal arms in *D. melanogaster*. Therefore, we concluded that the assignments for the rest of the genes based on Muller's element were valid.

Measurement of Recombination Rates

To measure the recombination rate in *D. albomicans* and *D. nasuta*, we counted the number of recombinants resulted from cross experiments. These two fly species were known to have chromosomal inversion polymorphisms (Ranganath and Krishnamurthy 1975, Lin et al. 1977). Because crossing over is suppressed in inversion heterozygotes (Sturtevant 1926), it is critical to use a strain pair that have the same inversion type as the parents. Therefore, we observed salivary grand chromosomes from 25 or more individuals of the F_1 hybrid for 112 and 41 combinations of strain pairs in *D. albomicans* and *D. nasuta*, respectively, to confirm the absence of inversion loop indicating inversion heterozygotes (supplementary fig. S1, Supplementary Material online). On the basis of the observation, we selected three strain pairs of *D. albomicans* (NG3 female and HRS402 male, SZ43 female and TSN7 male, and HNZ12 female and KM7 male) and three strain pairs of *D. nasuta* (TNR35 female and VNS23 male, K33 female and C3 male, and BZV1 female and MS77 male) as the parent strains for the cross experiments (supplementary tables S18 and S19, Supplementary Material online).

To count the recombinants, we used the pairs of genes that can distinguish between the parental strains in all the six strain pairs. Two pairs of genes on the second chromosome (*Arr1* and *Fdh*, and *Surf4* and *CaBP1*) and a pair of genes (*GstS1* and *Pgm*) on the third chromosome (*D. nasuta*) or the neo-sex chromosomes (*D. albomicans*) met this condition. Although the strains in the same strain pair are homosequential, there are chromosomal inversions among different species pairs (supplementary tables S18 and S19, Supplementary Material online). Therefore, the chromosomal position of marker genes and the distance between them may not be the same among strain pairs.

For each cross, we conducted a pair mating on the standard cornmeal medium at 25 °C. Then, we backcrossed a single female F_1 hybrid obtained to a single male of the paternal strain and counted the recombinants detected by the SNP markers among 16 males and 16 females in the backcross first generation (BC_1) to measure the female recombination rate. Similarly, to measure the male recombination rate, we backcrossed a single male F_1 hybrid to a single female of the paternal strain and counted the number of recombinants in the BC_1 generation.

Comparative Sequence Analyses

Nucleotide sequence alignments for the 55 individual genes in *D. albomicans* and *D. nasuta* were obtained by using Clustal W (Thompson et al. 1994) with local manual adjustments. The genetic diversity was quantified by means of Watterson's θ (Watterson 1975) and nucleotide diversity, π (Nei and Tajima 1981) under the condition of complete gap site deletion by using DnaSP5.1 (Librado and Rozas 2009).

The number of synonymous substitutions per sites was estimated by the Nei–Gojobori method (Nei and Gojobori

1986) with Jukes–Cantor multiple-hit correction (Jukes and Cantor 1969) by using MEGA5 software (Tamura et al. 2011). Using the synonymous distances obtained, we computed the number of nucleotide substitutions per site (d_{XY}) and the net number of nucleotide substitutions per site (d_A) between the neo-X and neo-Y chromosomes in *D. albomicans* and the third chromosome in *D. nasuta*.

Supplementary Material

Supplementary figure S1 and tables S1–S19 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors thank Aya Takahashi for stimulating discussion and valuable comments. Further thanks to the associate editor and two anonymous reviewers for their valuable comments and suggestions on earlier versions of this article. This work was partially supported by Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research Project (23370096) to K.T.

References

Ashburner M. 1989. Drosophila: A laboratory handbook. New York: Cold Spring Harbor Laboratory Press.

Bachtrog D. 2004. Evidence that positive selection drives Y-chromosome degeneration in *Drosophila miranda*. *Nat Genet.* 36:518–522.

Bachtrog D. 2006. The speciation history of the *Drosophila nasuta* complex. *Genet Res.* 88:13–26.

Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat Rev Genet.* 14:113–124.

Bachtrog D, Charlesworth B. 2002. Reduced adaptation of an evolving neo-Y chromosome. *Nature* 416:323–326.

Berset-Brändli L, Jaquiéry J, Perrin N. 2007. Recombination is suppressed and variability reduced in a nascent Y chromosome. *J Evol Biol.* 20:1182–1188.

Boom R, Sol CJ, Salimans MM, Jansen CL, Wertheim-van Dillen PM, van der Noordaa J. 1990. Rapid and simple method for purification of nucleic acids. *J Clin Microbiol.* 28:495–503.

Chang TP, Tsai TH, Chang HY. 2008. Fusions of Muller's elements during chromosome evolution of *Drosophila albomicans*. *Zool Stud.* 47:574–584.

Drosophila 12 Genomes Consortium. 2007. Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature* 450:203–218.

Jukes TH, Cantor CR. 1969. Evolution of protein molecules. In: Munro HN, editor. Mammalian protein metabolism. Academic Press, New York. pp. 21–132.

Kwiatowski J, Krawczyk M, Jaworski M, Skarecky D, Ayala FJ. 1997. Erratic evolution of glyserol-3-phosphate dehydrogenase in *Drosophila*, *Chymomyza*, and *Ceratitis*. *J Mol Evol.* 44:9–22.

Larracuente AM, Clark AG. 2013. Surprising differences in the variability of Y chromosomes in African and cosmopolitan populations of *Drosophila melanogaster*. *Genetics* 193:201–214.

Librado P, Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.

Lin FJ, Tseng HC, Chiang W. 1977. Chromosomal polymorphism in *Drosophila albomicans*. *Drosoph Inf Serv.* 52:153.

Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T. 1987. Male-driven molecular evolution: a model and nucleotide sequence analysis. *Cold Spring Harb Symp Quant Biol.* 52:863–867.

Muir G, Bergero R, Charlesworth D, Filatov DA. 2011. Does local adaptation cause high population differentiation of *Silene latifolia* Y chromosomes? *Evolution* 65:3368–3380.

Nei M. 1969. Linkage modification and sex difference in recombination. *Genetics* 63:681–699.

Nei M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol.* 3:418–426.

Nei M, Tajima F. 1981. DNA polymorphism detectable by restriction endonucleases. *Genetics* 97:145–163.

Ohsako T, Aotsuka T, Kitagawa O. 1994. The origins of the Japanese mainland population of *Drosophila albomicans*. *Jpn J Genet.* 69:183–194.

Ranganath HA, Krishnamurthy NB. 1975. Chromosomal polymorphism in *Drosophila nasuta*. III. Inverted gene arrangements in South Indian populations. *J Hered.* 66:90–96.

Rozen S, Marszalek JD, Alagappan RK, Skaletsky H, Page DC. 2009. Remarkably little variation in proteins encoded by the Y chromosome's single-copy genes, implying effective purifying selection. *Am J Hum Genet.* 85:923–928.

Sturtevant AH. 1926. A crossover reducer in *Drosophila melanogaster* due to inversion of a section of the third chromosome. *Biol Zentralbl.* 46:697–702.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 28:2731–2739.

Tamura K, Subramanian S, Kumar S. 2004. Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Mol Biol Evol.* 1:36–44.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.

Wakahama KI, Shinohara T, Hatsumi M, Uchida S, Kitagawa O. 1983. Metaphase chromosome configuration of the *immigrans* species group of *Drosophila*. *Jpn J Genet.* 58:315–326.

Watterson GA. 1975. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol.* 7:256–276.

Whitfield LS, Sulston JE, Goodfellow PN. 1995. Sequence variation of the human Y chromosome. *Nature* 378:379–380.

Wilson FD, Wheeler MR, Harget M, Kambysellis M. 1969. Cytogenetic relations in the *Drosophila nasuta* subgroup of the *immigrans* group of species. *Univ Texas Pub.* 6918:207–253.

Wilson Sayres MA, Lohmueller KE, Nielsen R. 2014. Natural selection reduced diversity on human Y chromosomes. *PLoS Genet.* 10:e1004064.

Wright F. 1990. The "effective number of codons" used in a gene. *Gene* 87:23–29.

Zhou Q, Bachtrog D. 2012. Chromosome-wide gene silencing initiates Y degeneration in *Drosophila*. *Curr Biol.* 22:522–525.