Footprints of Inversions at Present and Past Pseudoautosomal Boundaries in Human Sex Chromosomes

Claire Lemaitre,[1,2] *Marilia D. V. Braga,*[1,3] *Christian Gautier, Marie-France Sagot, Eric Tannier, and Gabriel A. B. Marais*

Université de Lyon; Université Lyon 1; Centre National de la Recherche Scientifique; Institut National de Recherche en Informatique et en Automatique; UMR5558; Laboratoire de Biométrie et Biologie évolutive; Villeurbanne, F-69622 cedex, France

The human sex chromosomes have stopped recombining gradually, which has left five evolutionary strata on the X chromosome. Y inversions are thought to have suppressed X–Y recombination but clear evidence is missing. Here, we looked for such evidence by focusing on a region—the X-added region (XAR)—that includes the pseudoautosomal region and the most recent strata 3 to 5. We estimated and analyzed the whole set of parsimonious scenarios of Y inversions given the gene order in XAR and its Y homolog. Comparing these to scenarios for simulated sequences suggests that the strata 4 and 5 were formed by Y inversions. By comparing the X and Y DNA sequences, we found clear evidence of two Y inversions associated with duplications that coincide with the boundaries of strata 4 and 5. Divergence between duplicates is in agreement with the timing of strata 4 and 5 formation. These duplicates show a complex pattern of gene conversion that resembles the pattern previously found for *AMELXY*, a stratum 3 locus. This suggests that this locus—despite *AMELY* being unbroken—was possibly involved in a Y inversion that formed stratum 3. However, no clear evidence supporting the formation of stratum 3 by a Y inversion was found, probably because this stratum is too old for such an inversion to be detectable. Our results strongly support the view that the most recent human strata have arisen by Y inversions and suggest that inversions have played a major role in the differentiation of our sex chromosomes.

Introduction

Chromosomal inversions are very common in animal, fungal, and plant genomes (Murphy et al. 2005; Yogeeswaran et al. 2005; Fischer et al. 2006; Ranz et al. 2007; Bhutkar et al. 2008). Although inversions have received long-standing attention in genetics (Dobzhansky 1950; Charlesworth B and Charlesworth D 1973; Charlesworth 1974; Sperlich and Pfriem 1986), the forces that establish inversions and their evolutionary significance remain poorly understood. An important characteristic of inversions is that recombination is suppressed at the inverted regions in chromosomal heterozygotes (Navarro et al. 1997; Andolfatto et al. 2001). This makes inversions particularly prone to accumulating mutations involved in local adaptation. Based on this, it has been suggested that inversions could have a significant role in speciation (Noor et al. 2001; Otto and Barton 2001; Rieseberg 2001; Navarro and Barton 2003; Kirkpatrick and Barton 2006). Another possible case of the evolutionary importance of inversions is the evolution of sex determination and sex chromosomes.

Well-differentiated sex chromosomes such as the human XY chromosomes do not recombine except in small regions (called pseudoautosomal regions, PARs). Theory predicts that recombination between newly formed sex chromosomes should be suppressed at male-determining genes so that they are genetically linked to the Y and no neutral or hermaphroditic recombinants are formed (Nei 1969; Charlesworth D and Charlesworth B 1978). It is also predicted that later in the evolution of the sex chromosomes, the accumulation of antagonistic genes (beneficial for males and deleterious for females) should gradually suppress recombination between X and Y, making the Y chromosome a fully or almost fully nonrecombining chromosome (Charlesworth et al. 2005).

In a pioneer work, Lahn and Page (1999) found that synonymous divergence between X–Y homologous gene pairs correlated with gene position on the X chromosome in a stair-like shape. They took this as evidence that the human XY were originally recombining autosomes that gradually stopped recombining, forming "evolutionary strata" (i.e., groups of genes which, because they do not recombine anymore, start diverging at the same time). Because gene order for those homologous gene pairs was found to be completely different on the X and on the Y chromosomes, they suggested that large Y inversions might have caused the evolutionary strata. The *XG* gene that spans the current pseudoautosomal boundary on the human X but is truncated on the human Y fits well with this idea.

Iwase et al. (2003) looked at Amelogenin, a gene that they believed to be located on an ancient pseudoautosomal boundary (strata 3/4) because the X–Y divergence drops from 30% to 10% within this gene. However, this gene was thought not to be involved in a Y inversion because its Y copy (*AMELY*) is not truncated. The sequencing of the euchromatic part of the human Y made the picture even more blurred instead of clarifying it. Skaletsky et al. (2003) reanalyzed *XY* gene pairs as in Lahn and Page (1999) with more data but they did not find four well-defined strata as in Lahn and Page (1999). The limits between strata seemed to overlap, especially those of the most recent strata (3 and 4). Following this work, doubts were raised about the Y inversion model put forward by Lahn and Page Charlesworth et al. 2005). Chromosomal rearrangements are known to accumulate at a faster rate in regions of reduced recombination. The rearrangements between X and Y could well have postdated the formation of the strata (and not predated it as expected in the Lahn and Page model).

[1] Equal contribution to this work.

[2] Present address: Université de Bordeaux, Centre de Bioinformatique–Génomique Fonctionnelle Bordeaux, Bordeaux, France.

[3] Present address: Universität Bielefeld, Technische Fakultät, AG Genominformatik, Bielefeld, Germany.

Key words: inversion, duplication, recombination, sex chromosomes, evolutionary strata.

E-mail: marais@biomserv.univ-lyon1.fr, Eric.Tannier@inria.fr.

Genome. Biol. Evol. Vol. 2009:56–66.
doi:10.1093/gbe/evp006
Advance Access publication April 30, 2009

In the paper reporting the complete sequencing of the human X, Amelogenin was dismissed as evidence against the Y inversion model (Ross et al. 2005). Indeed, *AMELX* and *AMELY* can undergo gene conversion, which could explain the Amelogenin peculiar X–Y divergence pattern (Marais and Galtier 2003; Ross et al. 2005). Using the GRIMM software, Ross et al. (2005) reported the first attempt to reconstruct the X–Y chromosomal rearrangements and found a scenario consistent with strata 4 and 5, a new stratum that they defined. Inferring a scenario of inversions is known to be a very challenging task when analyzing a sequence with relatively few markers and many inversions as in the case of the XY chromosomes. GRIMM uses an algorithm known to efficiently infer the minimum number of inversions from one sequence to another. However, GRIMM does not include any framework to find and analyze all the optimal scenarios. It just gives one arbitrarily drawn optimal scenario, which considerably weakens the conclusions of Ross et al. (2005). Thus, evidence for the model of the formation of the strata by Y inversions proposed by Lahn and Page remains dubious.

The gradual loss of recombination that formed the strata had a profound impact on the human Y degeneration. The strata have different levels of degeneration (the most recent ones being the least degenerate) and the level of dosage compensation, which is known to be a response to Y degeneration, is correlated with the strata (the most recent ones having fewer genes showing dosage compensation) (Carrel and Willard 2005). The gradual loss of recombination between our sex chromosomes has enhanced some processes of degeneration (genetic hitchhiking) over others (Muller's ratchet) compared with what would have happened if recombination had been stopped once (Bachtrog 2008). Evolutionary strata are not a bizarre feature of our sex chromosomes. They have been found in other organisms such as mice (Sandstedt and Tucker 2004), cats (Pearks Wilkerson et al. 2008), cattle (Van Laere et al. 2008), chicken (Handley et al. 2004; Nam and Ellegren 2008), and the plant *Silene latifolia* (Nicolas et al. 2005; Bergero et al. 2007) and seem to be a general phenomenon in heteromorphic sex chromosomes. Despite the importance of the strata for the biology and evolution of the sex chromosomes in general, we still know very little on how they are formed.

Here, our goal was to test whether the reconstruction of the human X–Y chromosomal rearrangements fits with the currently defined evolutionary strata in our species. We focused on the X-added region (XAR, it is located on the X p-arm and comprises PAR and strata 5, 4, and 3 that show about 5%, 10%, and 30% of X–Y divergence, respectively) because gene order is conserved between human and chicken (human XAR matches with chicken chromosome 1q with almost no rearrangements; see Ross et al. 2005). Detected rearrangements between X and Y in that region should have occurred on the Y. We first used the same 12 markers as in Ross et al. (2005) and evaluated the number of possible scenarios of Y inversions given the gene orders in XAR and in its Y homolog using a method that we developed (Braga et al. 2008). We found that there are many scenarios in which Y inversions coincide with strata boundaries. Using simulations with randomly distributed inversions on Y, we found that it is unlikely that this pattern has emerged just by chance. Another set of simulations with Y inversions occurring only among strata indicates that recent strata have arisen by inversions on the Y.

If Y inversions have formed the strata, by finding and analyzing the Y regions homologous to the strata boundaries on the X, we should be able to find footprints of those inversions. To do that, we used a method designed to detect and analyze genomic regions with breakpoints (Lemaitre et al. 2008) and we found clear evidence of inversions associated with duplications at the PAR/stratum 5 and the strata 5/4 boundaries. Analysis of the divergence between duplicates for both boundaries allowed us to date the inversions and is in agreement with stratum 5 being more recent than stratum 4. This strongly suggests that in humans, recent strata have arisen by inversions on the Y and provides support to the idea Lahn and Page (1999) first put forward that recombination between X and Y stopped because of inversions on the Y. A pair of duplicates associated with an inversion shows evidence for gene conversion, and this suggests that gene conversion in *AMEL* is in fact consistent with *AMEL* spanning the strata 3/4 boundary. However, we could not find either clear footprints of inversions for stratum 3 or any optimal scenarios consistent with the formation of stratum 3 by a single Y inversion (including all the available markers for this stratum and not just two as in Ross et al. 2005), but this may simply be because stratum 3 is too old, footprints of inversions have been erased, and X–Y are too rearranged. We discuss the case of ancient human strata.

Materials and Methods

Identification of New Markers

Already known markers are from Skaletsky et al. (2003) and Ross et al. (2005). New markers were identified from an alignment of the human X–Y chromosomes (from the National Center for Biotechnology Information version 36, hg18) using BlastZ (Schwartz et al. 2003). Local similarities found with BlastZ were concatenated if they had the same order and orientation and if they were less than 30 kb apart on both chromosomes. Concatenates smaller than 30 kb and those located in known ampliconic regions and in the pericentromeric regions were discarded. We thus obtained three new markers (see table 1).

Analysis of Optimal Scenarios for X–Y Rearrangements

To analyze the formation of strata 4 and 5, we ran the software BaobabLuna (Braga et al. 2008) on the following sequences: X = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12) and Y = (−12, 11, −2, −1, −10, −9, 8, −5, 7, 6, −4, 3) with numbers corresponding to markers in Ross et al. (2005) and minus indicating a change in orientation. BaobabLuna computes the minimum distance of inversions using the same algorithm GRIMM uses, which is based on an analysis of the breakpoint graph (Tesler 2002). In addition, it enumerates a set of inversions scenarios representative of all possible scenarios that realize this distance (grouped in classes of equivalence). This way it is possible to derive some properties of the whole set of optimal inversion

Table 1
List of Markers on X and Y Chromosomes in Humans in This Study

Marker Name[a]	X Position[b]	Y Position[b]	X Order[c]	Y Order[c]	Strata	References[d]
PAR	0–2709520	0–2709520	0	0	PAR	Lahn and Page (1999)
*GYG**, *ARSD**, *ARSE**, *ARSF***, *ADLICAN***(1)	2672359–3346731	12492110–13139179	1	−6	5	*Lahn and Page (1999) and **Skaletsky et al. (2003)
PRK (2)	3345018–3848954	7068601–7506089	2	−5	5	Lahn and Page (1999)
Anonymous	3662755–3909738	19743211–19851471	3	−20	4	This article
Anonymous (3)	4110549–4490406	17583377–18076812	4	19	4	Ross et al. (2005)
Anonymous (4)	4602689–5384111	16706206–17570219	5	−18	4	Ross et al. (2005)
Around NLGN4 (5)	5384848–6313029	14981290–15805945	6	−15	4	Skaletsky et al. (2003)
Anonymous (6)	6594680–6624810	16664668–16691008	7	17	4	Ross et al. (2005)
Around STS (7)	6625496–7448677	15807027–16376778	8	16	4	Lahn and Page (1999)
Anonymous (8)	7449397–7646086	14794314–14971778	9	14	4	Ross et al. (2005)
Around VC (9)	7731889–7952770	14681748–14772569	10	−13	4	Skaletsky et al. (2003)
Around KAL1 (10)	8388775–8678660	14456224–14455780	11	−12	4	Lahn and Page (1999)
TBL1 (11)	9367582–9694004	6818075–7040054	12	4	3 (or 4) [e]	Skaletsky et al. (2003)
APXL	9803943–9836714	13139984–13177590	13	7	3 (or 4) [e]	Skaletsky et al. (2003)
Anonymous	9925052–10026743	2935524–6736276	14	2	3 (or 4) [e]	This article
AMEL (12)	11221454–11228802	6756180–6804332	15	−3	3	Lahn and Page (1999)
TMSB4	12893995–12914689	14259652–14336452	16	11	3	Lahn and Page (1999)
TXNLG	16713573–16773411	20187740–20234258	17	22	3	Skaletsky et al. (2003)
EIF1A	20052557–20069887	21146999–21164428	18	−23	3	Lahn and Page (1999)
ZF	24071318–24144376	2855296–2922379	19	1	3	Lahn and Page (1999)
MAP3/TAB3	30755480–30819301	13771944–13828537	20	9	3	This article
BCoR	39795364–39917376	20076630–20184596	21	21	3	Skaletsky et al. (2003)
CRSP2P-CASK	40392502–41667660	13240309–13592325	22	8	3	Lahn and Page (1999)
UT	44617701–44856791	13869035–14101947	23	−10	3	Lahn and Page (1999)

[a] In brackets are indicated marker numbers in Ross et al. (2005).
[b] Positions on X and Y from National Center for Biotechnology Information version 36, hg18.
[c] Order on the X and Y chromosomes of each marker is indicated.
[d] Reference mentioning the markers for the first time is shown. The three new markers that we found are mentioned.
[e] See section on Gene Conversion between *XG* Copies.

scenarios that possibly differentiate the X and Y gene order. Markers 1 and 2 belong to stratum 5, markers 3 to 10 belong to stratum 4, and markers 11 and 12 belong to stratum 3 (see table 1). To analyze the formation of stratum 3, we ran the same program on the following sequences: X = (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23) and Y = (19, 14, −15, 12, −2, −1, 13, 22, 20, −23, 16, −11, −10, 9, −6, 8, 7, −5, 4, −3, 21, 17, −18) with stratum 3 being extended from marker 11 to marker 23. We also did the analysis without marker 23 (see Discussion).

Simulation of Inversions

Simulations were conducted to create a sequence Y from a sequence X with a given number of inversions d. A simulated Y sequence was made in d steps. At step i, two positions were selected at random, and an inversion between these two positions was produced. The inversion was accepted only if going back from Y_i (Y after i inversions) to Y_0 (Y without inversion = X) implied i inversions (parsimony criterion). We stopped the process after d parsimonious inversions. We repeated this 1,000 times and obtained the free inversion simulations. The set of strata-constrained simulations was obtained by adding another constraint in the simulation process. The first inversion had to form stratum 3 (markers 11, 12). Then, two inversions had to take place in the following order: one inversion including all stratum 4 markers (with the possibility of including some markers of stratum 3) and a second inversion including all stratum 5 markers (with the possibility of including some markers of strata 4 and 3). The other inversions occurred in a nondetermined order within the already formed strata. All inversions had to fulfill the parsimony criterion (see above). Different values of d were tested and gave very similar results. Results shown in the paper (table 2, fig. 2) are with $d = 8$ (results for other values of d are shown in the supplementary data [Supplementary Material online]).

Analysis of Breakpoints

We used the method by Lemaitre et al. (2008). Briefly, this method works on breakpoint regions. A breakpoint is defined by two adjacent markers on one sequence (here X) that are not adjacent on the homologous sequence (here Y). By aligning an X region with a breakpoint with its two corresponding Y regions, it is possible to locate precisely the breakpoint and to analyze it. Alignments were performed with BlastZ (Schwartz et al. 2003) on repeat-free (using RepeatMasker) sequences. The minimum breakpoint interval is obtained by looking at the distribution of hits using a partitioning algorithm.

Levels of Divergence Between Duplicates

Duplicates found with the method described in the previous section were aligned on the entire Y chromosome and no other copies were found. The level of divergence

Table 2
Analysis of All the Optimal Scenarios for X–Y Rearrangements

Classes of Equivalence	Curr. Three Strata	Alt. Three Strata	Two Strata	One Stratum	No Strata	Total
1	420	0	2,520	0	7,140	10,080
2	0	0	0	1,260	8,820	10,080
3	0	0	0	1,260	8,820	10,080
4	0	120	0	0	216	336
5	0	0	0	0	336	336
6	0	0	0	0	840	840
Total	420	120	2,520	2,520	26,172	31,752

NOTE.—The 12 markers from Ross et al. (2005) have been used (see also table 1). Classes of equivalence group scenarios with the same inversions but in different orders (Braga et al. 2008). Curr. Three Strata are the strata defined by Ross et al. (2005). Alt. Three Strata are alternative strata with the following definition: stratum 5 = {1, 2, 3}; stratum 4 = {4, . . ., 10}; and stratum 3 = {11, 12}. Two strata: {1, 2} and {3, . . ., 12}. One stratum: {1, . . ., 12}. Subtotals and grand total are indicated. All the classes of equivalence are described in the supplementary data (Supplementary Material online).

between duplicates was obtained using exact pairwise alignment tools (the Water and Matcher programs from the EMBOSS tools suite; see Rice et al. 2000).

For the analysis of *XG*, we ran codeml on coding sequences to get d_N and d_S values (Yang 1997, 2007, and PAML on the Web: http://coot.embl.de/pal2nal/; Suyama et al. 2006). The percentage of similarity for total (coding + noncoding) DNA was obtained by aligning sequences using BlastZ with the chaining option (Schwartz et al. 2003). Only blocks of more than 70% similarity were kept. To compute the percentage of similarity, we summed all the identical sites and divided by the size of the X sequence.

Results and Discussion

Comparing Scenarios for X–Y Chromosomal Rearrangements and Evolutionary Strata 4 and 5

Ross et al. (2005) presented a scenario of inversions between the human X and Y chromosomes consistent with the evolutionary strata 3, 4, and 5. They obtained it using GRIMM (Tesler 2002), a software package that uses marker order in two sequences to propose a scenario of inversions minimizing the total number of inversion events, and applied it to 12 pairs of X–Y markers spanning a small part of stratum 3, and the whole of strata 4, 5, and PAR (and totaling 11 Mb of sequences). In the proposed scenario (shown in fig. 1), there are two large inversions that coincide with strata 4 and 5 and could have formed them and five small inversions that are included in the strata and do not affect strata boundaries. They also found an inversion consistent with stratum 3, but because only a small part of this stratum was investigated, no conclusion could be drawn. This showed for the first time that it was possible to find a scenario of inversions consistent with the human strata, at least the most recent ones (i.e., strata 4 and 5). However, there is an ongoing debate about the relevancy of the scenario proposed by GRIMM and other similar programs. GRIMM uses the algorithm of Hannenhalli and Pevzner (1999) and we know that it is an accurate way of getting the minimum number of inversions between two sequences with different marker orders. The scenario proposed by GRIMM, however, is only one possible optimal scenario (i.e., with the minimum number of inversions). There may be many other such optimal scenarios, and GRIMM does not offer the possibility of identifying and analyzing these scenarios. This, of course, weakens the conclusion on the evolution of human strata from the GRIMM results. The proposed scenario is consistent with strata formation, but there may be equally good scenarios not consistent with strata formation. To solve this problem in general, it was suggested to enumerate all the possible optimal scenarios (Siepel 2003). The result of this showed that the number of optimal scenarios is often huge. Such an enumeration method is, however, too demanding in terms of computational time. A mathematical formalism was proposed to accelerate this process (Bergeron et al. 2002). Briefly, the idea is to group optimal scenarios into classes of equivalence. All scenarios in any one of the classes are composed of the same inversions but in different order. Finding an efficient algorithm to enumerate all classes of equivalence remained, however, an open problem for a few years, until Braga et al. (2008) designed and implemented one, which is efficient when the number of rearranged markers is not too large.

We applied this method to the same 12 markers used by Ross et al. (2005) to evaluate how their conclusion was affected by analyzing all the optimal scenarios (see table 1 and fig. 1 for more information about these 12 markers). We found six classes of equivalence (i.e., groups of optimal scenarios with the same inversions but in different order). The solution proposed by GRIMM is in one of them, but there are five other classes with different inversions for a total of 31,752 optimal scenarios. We then counted the number of optimal scenarios consistent with strata 3, 4, and 5 with strata boundaries as defined in Lahn and Page (1999) and Skaletsky et al. (2003) and corrected by Ross et al. (2005). We expected that in the scenarios consistent with the strata, there were three ordered inversions affecting, first, markers in stratum 3 (markers 11, 12); then, markers in stratum 4 (markers 3 to 10); and finally, markers in stratum 5 (markers 1, 2). A given inversion forming a stratum had to comprise all the markers of the stratum and could comprise additional markers from the previous stratum (for instance, markers 11 and 12 could be involved in a large inversion forming stratum 4, as in fig. 1). The remaining inversions were small-scale ones occurring within already formed strata. With these criteria, we found that only one class of equivalence—the one including the GRIMM scenario—agrees with the currently defined strata 3, 4, and 5 (see table 2). Inside this class, 420 scenarios were found consistent with these strata. We also looked at scenarios consistent with three strata with different boundaries than the ones proposed in Ross et al. (2005), with two strata, with one stratum only, and without strata (see table 2). In each case, we found a number of optimal scenarios consistent with the tested strata structure depending on the classes (see table 2). When pooling results for all classes, we found more optimal scenarios for the currently defined three strata (420) than for the alternative three strata (120), but there are more scenarios for two strata (2,520) and one stratum (2,520), and even more

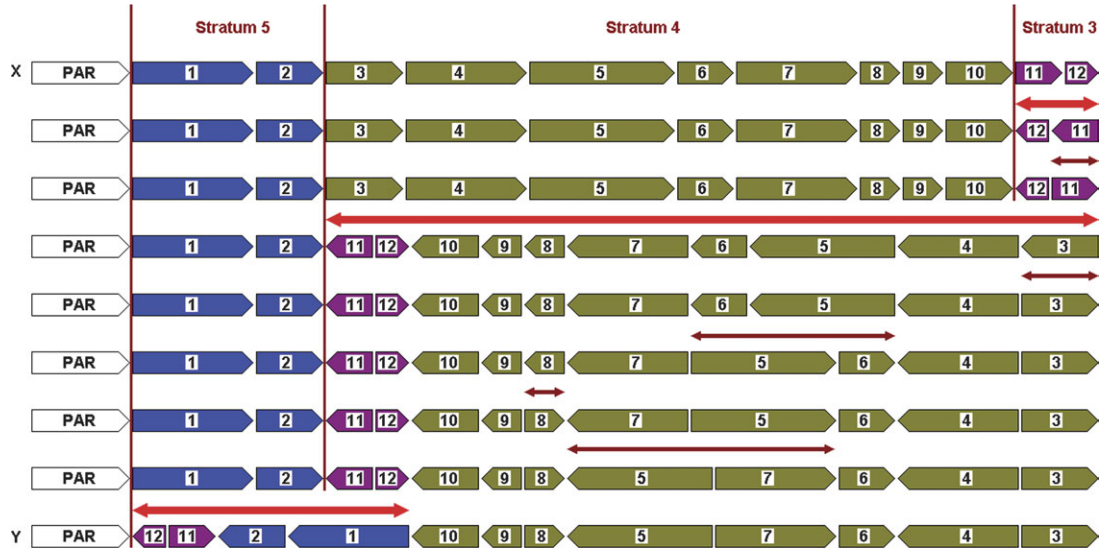


FIG. 1.—A possible scenario for the X–Y rearrangements and the evolution of recent human strata (adapted from Ross et al. 2005). This scenario has been obtained by GRIMM using 12 markers covering PAR, stratum 5, stratum 4, and the beginning of stratum 3. Strata definitions are from Ross et al. (2005). See the list of markers in table 1. Inversions that coincide with strata and that could have formed them are indicated in red. Other inversions are in brown.

for no stratum at all (26,172), than for the currently defined three strata.

We then performed simulations with random inversions in order to help us interpret the data. We did a first round of simulations with free inversions (see Materials and Methods) among 12 markers with the number of inversions being equal to the observed number for the human X–Y data of Ross et al. (2005). In the second round of simulations, the idea was to model the formation of strata by inversions. We simulated inversions that were constrained by three strata with boundaries mimicking the currently defined strata 3, 4, and 5 (see Materials and Methods). In the process of simulations, inversions had to be ordered (starting with the oldest strata markers and finishing with the youngest) and had to respect strata boundaries. Small inversions could occur in already formed strata only. We called this set of simulations the "strata-constrained inversions." We then ran the method of Braga et al. (2008) on the simulated sequences (comparing each time the initial sequence and one sequence with simulated inversions). This generated distributions of parameters such as total number of optimal scenarios, total number of classes of equivalence, and others. The mean values of these parameters are shown in table 3. We compared the distributions of free versus strata-constrained simulations by using nonparametric statistics. For all of them, the distributions for free and strata-constrained simulations are significantly different (see table 3). The most relevant parameter is probably the number of optimal scenarios consistent with the three strata over the total number of optimal scenarios (#strata_scen/#tot_scen). We found that this parameter is significantly smaller for free than for strata-constrained simulations ($P < 10^{-16}$). Figure 2 shows the distributions of this parameter for the two sets of simulations. The observed value for the human XY data (from table 2) is a clear outlier in the "free" distribution. We get a P value of 0.009 from this comparison, which means that it is very unlikely that the inversions between X and Y (in the region studied by Ross et al. 2005) occurred freely along the Y sequence without any constraints. The comparison of the observed value of #strata_scen/#tot_scen for the human XY data with the distribution of the same parameter for the strata-constrained simulations resulted in a nonsignificant P value, which means that the process used to simulate this set may be the same for the human sex chromosomes. Although the two sets of simulations that we generated are extreme cases and we did not investigate intermediary cases, these results suggest that the hypothesis that evolutionary strata 4 and 5 (no conclusion can be drawn for stratum 3, see above) have been formed by Y inversions is a likely hypothesis.

Looking at Chromosomal Breakpoints Near Strata 4 and 5 Boundaries

A strong piece of evidence for the model of Y inversions put forward by Lahn and Page (1999) would be to find a region spanning a strata boundary on the X chromosome matching with two broken bits on the Y chromosome. At the pseudoautosomal boundary, the Y copy of the *XG* gene is truncated. If stratum 5 had been formed by a Y inversion, we should be able to find the missing bit of *XG* at the end of the putative inversions having formed stratum 5 on the Y chromosome (upstream of marker 1 on the Y sequence on fig. 1). The same rationale can be applied to stratum 4. To check this, we used a method to precisely detect and analyze chromosomal breakpoints in sequences. This method starts with a BlastZ (Schwartz et al. 2003) comparison between two sequences where there is a breakpoint (Lemaitre et al. 2008). It maps all the hits on the sequences and uses the distribution of hits to find the minimum interval where the breakpoint is (see Materials and Methods). It gives a precise picture of the similarities at breakpoints.

Table 3
Comparison of Free and Strata-Constrained Simulations and True X–Y Sequences

	#tot_ scen	#strata_ scen	#strata_ scen/#tot_scen	#tot_ class	#strata_ class	#strata_ class/# tot_class
Observed XY	31,752	420	0.0132	6	1	0.167
Free simulations	36,100	12	0.0004	173	0.323	0.0019
Strata-constrained simulations	41,700	406	0.0135	191	6.91	0.0544
P values, free versus strata-constrained	$<10^{-5}$	$<10^{-16}$	$<10^{-16}$	$<10^{-3}$	$<10^{-16}$	$<10^{-16}$

NOTE.—Values are medians of the distribution of the different parameters. Parameters are #tot_scen = total number of optimal scenarios; #strata_scen = number of scenarios consistent with the three currently strata 3, 4, 5; #tot_class = total number of classes of equivalence; #strata_class = number of classes of equivalence consistent with the three currently defined strata 3, 4, 5. Statistical tests are nonparametric tests (Wilcoxon) for comparing medians of two distributions. Observed values for XY are from table 2.

Figure 3*A* shows the results for PAR/stratum 5. We found similarities between X and Y at the end of the PAR region (from 2.67 to 2.71 Mb on both X and Y sequences). This includes the part of the *XG* gene found both in X and Y that defines the human pseudoautosomal boundary. Interestingly, we also found similarities between PAR and a sequence around position 13 Mb of the Y chromosome that happens to be one of the ends of stratum 5 on that chromosome. These similarities extend over 40 kb and mirror the similarities found in the PAR, which clearly indicates an inverted duplication. Instead of finding only the missing bit of the *XG* region when looking at the end of stratum 5 in the Y chromosome, we thus found an inverted duplication of almost the entire *XG* region. Analysis of the strata 4/5 boundary gave very similar results (see fig. 3*B*). Again, the dot plot indicates duplications with similarities between the X region from 3.66 to 3.74 Mb (end of stratum 5) and both Y regions from 7.06 to 7.18 Mb (end of stratum 5) and from 19.72 to 19.8 Mb (beginning of stratum 4). Duplicated sequences flanking stratum 4 suggest again a single inversion event spanning the whole stratum 4. In figure 3*C*, we show the duplicates and their orientations on the X and Y sequences for PAR/stratum 5 and stratum 5/stratum 4.

Inversions are often found associated with duplications (Casals and Navarro 2007; Kehrer-Sawatzki and Cooper 2008). This association has been interpreted first as the result of nonallelic homologous recombination between duplicates. Indeed, there are well-documented cases of such a mechanism. In humans for instance, there is a polymorphic inversion on chromosome Xq28 that includes the *FLNA* and *EMD* loci and that is flanked by inverted duplicates. It was shown that these inverted duplicates are present in all placental mammals and that there is a recurrent inversion of the segment between these duplicates in several mammalian lineages (Cáceres et al. 2007). However, some inversions come from another mechanism (Casals and Navarro 2007; Kehrer-Sawatzki and Cooper 2008). Human and chimp genomes differ by several chromosomal rearrangements. One of them is a pericentric inversion in the chromosome 10 of chimpanzee in the region around the *SLCO1B3* gene. A comparison between humans and chimps revealed that this pericentric inversion produced a duplication found only in chimps, which suggests that the inversion generated the duplication and not the contrary (Kehrer-Sawatzki et al. 2005). More recently, a multigenome comparison in *Drosophila* revealed that 60% of the inversions produced duplications and were formed by a mechanism called isochromatid model with staggered single-strand breaks (Ranz et al. 2007). In this mechanism, two pairs of staggered single-strand breaks result in long 5′ overhangs, which can then be filled in by DNA synthesis. When followed by a repair pathway called nonhomologous end joining, this results in an inversion flanked by inverted duplications of the sequences between the paired single-strand breaks (Ranz et al. 2007).

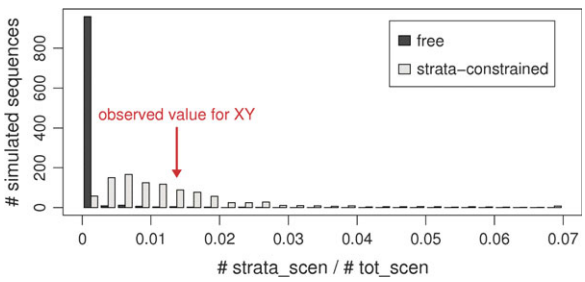
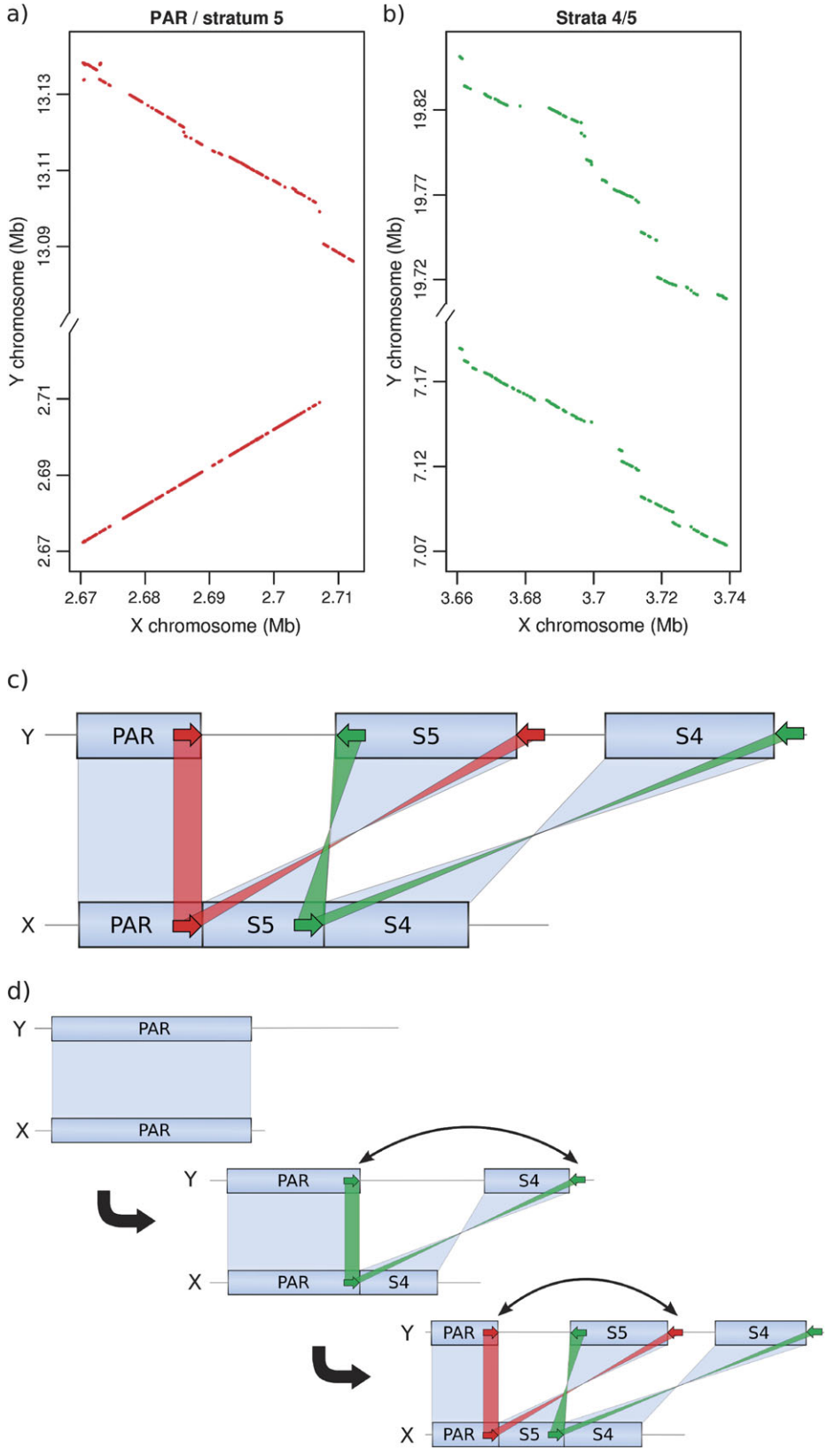


FIG. 2.—Distributions of the number of optimal scenarios consistent with the human strata 4 and 5 over the total number of optimal scenarios (#strata_scen/#total_scen) for free and strata-constrained simulated sequences. Free simulations have been obtained by random inversions (black boxes). Strata-constrained simulations have been obtained by simulating formation of strata by inversions (using currently defined strata 3, 4, and 5 for humans) with additional small inversions occurring within strata after their formation (white boxes). See main text (Materials and Methods, Results, and Discussion) for more details. #strata_scen/#total_scen values for simulated sequences have been obtained using the Braga et al. (2008) program. The value observed for the true X–Y sequences is indicated by a red arrow.

Both models could explain our results. A Y-specific duplication could have occurred and then an inversion between the duplicates in agreement with the nonallelic recombination model. The Y inversion itself could have produced the duplicates as in the isochromatid model. We tend to favor the latter because it is more parsimonious (one event—the inversion—in the isochromatid model with staggered single-strand breaks instead of two events—a duplication and an inversion—in the nonallelic recombination model). In figure 3*D*, we show a scenario of the formation of strata 4 and 5 consistent with the isochromatid model, but in any case, our results clearly point toward a single inversion event spanning the whole stratum 5 and another similar event for stratum 4. An important point is that the orientation of the duplicates is fully compatible with two large inversions giving rise to stratum 4 first and then to stratum 5 (see fig. 3*D*). Indeed, if we invert stratum 5 back in place on the Y chromosome, the duplicates for stratum 4

are in inverted orientation. We also compared the X and Y duplicated sequences found at PAR/stratum 5 and strata 4/5 (see Materials and Methods). The level of divergence between duplicated segments is 30% for PAR/stratum 5 and 50% for strata 4/5, which is fully consistent with the fact that the formation of stratum 5 is more recent than the formation of stratum 4. However, the divergence is clearly higher than the reported divergence for stratum 5



(5%) and stratum 4 (10–15%) (see Iwase et al. 2003; Skaletsky et al. 2003; Ross et al. 2005). This may be because these estimates of divergence and ours have been obtained by different methods. In previous work, the estimates mainly came from the comparison of synonymous sites of coding regions (d_S values). Our estimates have been obtained on X and Y regions that include nonhomologous sequences (i.e., DNA repeats and other inserted/deleted sequences), which decreases the quality of the global alignment and increases divergence. Nevertheless, our results strongly suggest that there are footprints of inversions at the recent strata boundaries and that these inversions have produced the strata.

Evidence for Gene Conversion Between *XG* Copies

A more careful analysis of the divergence pattern between the pair of duplicates (namely the *XG* gene) associated with the stratum 5 inversion gave unexpected results. *XG* has three copies—the X copy (*XG-X*) and two Y copies: the one in the PAR (*XG-Y1*) and the entire copy in the nonrecombining Y, hereafter called NRY (*XG-Y2*). *XG-Y1* and a part of *XG-X* are in the PAR and are nearly identical. We obtained from ENSEMBL v49 the sequences of *XG-X* and *XG-Y2*. *XG-Y2* is clearly a pseudogene because it has two premature stop codons compared with *XG-X* (data not shown). We first computed the d_N and d_S using PAML (Yang 1997, 2007) for the exons of these copies (see fig. 4). Surprisingly, we found very different results for the part common to *XG-X*, *XG-Y1*, and *XG-Y2* (hereafter called *XG-5′*) and for the other part (found only in *XG-X* and *XG-Y2*, hereafter called *XG-3′*). *XG-5′* has a lower d_S value (0.060) than *XG-3′* (0.091), and *XG-5′* has a much lower d_N/d_S ratio (0.019) than *XG-3′* (1.01). The results for *XG-3′* are in agreement with *XG-Y2* being a pseudogene (d_N/d_S ratio of 1), but the lower d_S value and the much lower d_N/d_S ratio for *XG-5′* suggests genetic exchanges from the functional *XG-X* to the nonfunctional *XG-Y2*. This is striking because *XG-Y2* is in the NRY and is not expected to recombine. The analysis of total DNA (including exons and introns) confirmed the results for the exons only (see fig. 4). We found a higher percentage of similarity for *XG-5′* (83.26%) than for *XG-3′* (67.32%) when comparing *XG-X* and *XG-Y2*. Genetic exchanges from *XG-X* to *XG-Y2* are not expected. This could have happened by X–Y gene conversion involving *XG-X* and *XG-Y2* or by Y–Y gene conversion involving *XG-Y1* and *XG-Y2*. We know that Y–Y gene conversion exists in humans because evidence of strong gene conversion has been found among Y genes of the same multigene families (Rozen et al. 2003; Bhowmick et al. 2007). The drop in divergence between *XG-X* and *XG-Y2* seems to lie in a region of about 8 kb where both copies share more than 95% of identity, which is consistent with a single very recent event of gene conversion.

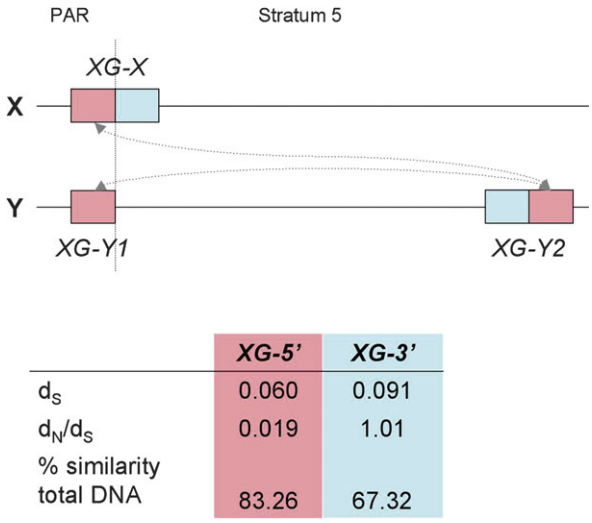


FIG. 4.—Divergence patterns among *XG* copies. The three copies of *XG* are shown—the X copy (*XG-X*) and two Y copies: the one in the PAR (*XG-Y1*) and the entire copy in the NRY (*XG-Y2*). *XG-Y1* and *XG-Y2* are the inverted duplicates shown in red in figure 3. *XG-X* and *XG-Y2* have been compared. Two regions were defined: *XG-5′* (1–29,500) and *XG-3′* (29,500–63,854). d_S and d_N were estimated using PAML and the percentage of similarity for total DNA was obtained using BlastZ (see Material and Methods). Gray arrows indicate possible events of gene conversion.

These results have important implications for *AMEL*. The boundary between strata 3 and 4 was thought to lie within *AMEL* because of a drop of divergence from 30% to 10% within this gene (Iwase et al. 2003). But *AMELY* is not truncated and this was considered as evidence that there was no inversion affecting *AMEL*. *AMEL* was clearly a problem for the Lahn and Page (1999) model of strata formation by Y inversion until it was found that *AMELX* and *AMELY* showed evidence for gene conversion, which could explain the peculiar divergence pattern of *AMEL* (Marais and Galtier 2003; Ross et al. 2005). The strata

←

FIG. 3.—Analysis of breakpoints at strata and PAR boundaries in humans. (*A*) Dot plot for the PAR/stratum 5 boundary. This shows the similarities between the X region at the PAR/stratum 5 boundary with two "broken" regions on the Y. It clearly shows that the X region is duplicated on the Y (with one duplicate being inverted). Total length of the region = 45 kb. (*B*) Dot plot for the strata 4/5 boundary. This shows the similarities between the X region at the strata 4/5 boundary with two broken regions on the Y. It clearly shows that the X region is duplicated on the Y (both Y duplicates are in inverted orientation compared with the X homologous region). Total length of the region = 110 kb. See main text (Materials and Methods, Results, and Discussion) for more details. (*C*) Picture showing the location and orientation of the duplications on the chromosomes X and Y. On the Y chromosome, stratum 5 is flanked by duplication of the PAR/stratum 5 region of the X (shown in red), which indicates a large inversion spanning the entire stratum 5. Duplicates of the strata 4/5 region of the X are found at the ends of stratum 5 and stratum 4 (shown in green). This defines a large inversion spanning the whole stratum 4. Importantly, duplicates are in an orientation consistent with two large inversions that have formed stratum 4 first and then stratum 5. (*D*) Sketch showing the scenario with two inversions for the formation of strata 4 and 5 under the isochromatid model with staggered single-strand breaks (see Ranz et al. 2007). The first inversion reduces the size of the PAR and forms stratum 4 with two inverted duplicates flanking the inversion. The second inversion reduces further the size of the PAR and forms stratum 5 with two inverted duplicates flanking the inversion. Note that duplicates associated with the formation of stratum 4 are no longer inverted because one of them is involved in the inversion that has formed stratum 5. Lines with arrows indicate inversions. In (*C*) and (*D*): Blocks of similarities are indicated by blue boxes and shadows. Black lines indicate large stretches of nonhomologous sequences. Sizes of boxes and lines are not in scale.

3/4 boundary has been put between *KAL1X* and *TBL1X* by Ross et al. (2005), but this boundary is still debated. Our results on *XG* suggest that the strata 3/4 boundary could be in *AMEL*. An inversion could have formed stratum 3 and produced two copies of *AMEL* on the Y chromosome with one copy being complete. The same kind of configuration as for XG could have existed with a part of *AMELX* and a truncated *AMELY* in the PAR and an entire *AMELY* in the NRY with a possibility of gene conversion between these copies. This would have produced the divergence pattern that we now observed when comparing *AMELX* and *AMELY*, which resembles that of *XG-X* and *XG-Y2*. We looked for a truncated *AMELY* by BlastZ search (*AMELY* against the whole Y chromosome) and found no other hit than *AMELY* itself. However, the formation of stratum 3 is an ancient event and the truncated *AMELY*, which was made a pseudogene at that time, may well no longer be recognizable or may have been deleted.

Discussing the Case of the Ancient Human Strata

Following Ross et al. (2005), we focused mainly on strata 4 and 5 in all previous sections of Results and Discussion. In the first section, we had some markers from stratum 3, but they have been mainly used to delimit stratum 4, and stratum 3 was not analyzed entirely. We only included two stratum 3 markers as in Ross et al. (2005). In this section, we address the question whether stratum 3 was formed by a single inversion on the Y chromosome. We used all the markers available in the literature for stratum 3 plus 3 new markers that we found when looking for similarities between the X and Y sequences. In total, we had 23 markers (stratum 5: 2, stratum 4: 7 + 1 new, stratum 3: 11 + 2 new) that with the PAR covered the first 45 Mb of the X (see table 1). We ran our method on the set of 23 markers. We could still find optimal scenarios consistent with the formation of strata 4 and 5 by Y inversions, which shows that the conclusions obtained with 12 markers remain unchanged by adding more markers. We could not, however, find any scenario consistent with the formation of stratum 3 by a Y inversion (using strata 3/4 boundary as in Ross et al. 2005).

There are several explanations for that. One is that the strata 3/4 boundary is not well defined and this may affect the results. When we put the strata 3/4 boundary at *AMELX* (by removing the *AMEL* marker and putting the boundary at its place), the results remain unchanged. The second possible explanation is that stratum 3 was formed by a mechanism different from Y inversions. Consistent with this idea, we analyzed stratum 3 with our method (Lemaitre et al. 2008) but we could find no clear evidence (no duplications) of large inversions as we found for strata 5 and 4. In *S. latifolia*—a dioecious plant with recently evolved (<10 mya) sex chromosomes—three strata have already evolved (Nicolas et al. 2005; Bergero et al. 2007). Maps for the X and Y chromosomes are being developed, and preliminary data suggest that stratum 2 may have been formed by a large paracentric inversion but for strata 1 and 3 the mechanism is less clear (Bergero et al. 2008). Stratum 3 may have been formed by translocation in some *S. latifolia* populations but not all, which suggests that other type of chromosomal rearrangements than inversion could form strata. However, the only translocation that has affected the human stratum 3 seems to be the translocation of the entire XAR (see Ross et al. 2005), and it is not clear why this would have stopped recombination at some part of XAR (e.g., stratum 3) and not others (strata 5, 4, and the current PAR). Moreover, absence of footprints of an inversion spanning stratum 3 may simply be due to a too high level of divergence between X and Y sequences for this stratum.

A third explanation is that stratum 3 is not just one stratum, but several. It is interesting to notice that more ancient strata get larger and larger. Stratum 1 alone covers the whole q arm of the X chromosome (Lahn and Page 1999; Skaletsky et al. 2003). This, of course, is surprising and may be simply due to the difficulty in identifying distinct strata when these strata are ancient. Stratum 3 may well include several strata that were formed successively in a short period of time and that are no longer distinguishable simply because the levels of divergence between the X and Y sequences in these strata are very similar. Another line of evidence supporting this hypothesis is that a similar number of strata has been observed in recent and ancient heteromorphic sex chromosomes. Ancient sex chromosomes such as that of humans and chicken have five and three strata, respectively (Ross et al. 2005; Nam and Ellegren 2008), and recently evolved but already differentiated XY such as that of *S. latifolia* also have three strata (Nicolas et al. 2005; Bergero et al. 2007, 2008). This suggests that strata can accumulate at a fast rate and that in old sex chromosomes, old strata may not be distinguishable.

Interestingly, it is possible to find an optimal scenario consistent with the markers in stratum 3 up to *CASK*, which suggests that stratum 3 could comprise a first stratum from the strata 3/4 boundary (either between *KAL1X* and *TBL1X* or at *AMELX*) to *CASK* and another one including *UTX*. We know that heterozygotes for an inversion have recombination suppressed at the inversion. Inhibition of the molecular mechanism of recombination or selection against gametes with chromosomal rearrangements could explain this suppression. The suppression is complete for pericentric inversions, but for paracentric inversions, it is stronger at the inversion breakpoints (Navarro et al. 1997; Andolfatto et al. 2001). For sufficiently large paracentric inversions, recombination is not suppressed in the middle of the inversion. Indeed, for such inversions, the probability of two crossovers occurring within the inversion (and avoiding chromosomal rearrangements) can be high. All this seems to be true also for inversions affecting sex chromosomes in *Drosophila americana* (McAllister 2003; Evans et al. 2007) and the black muntjac (Zhou et al. 2008). Recombination between sex chromosomes will be suppressed efficiently with relatively small inversions only, which reinforces the idea that the ancient strata that we have defined based on the X–Y divergence, such as stratum 3, are in fact a mosaic of smaller strata that are no longer distinguishable.

Concluding Remarks

Our results strongly suggest that strata 4 and 5—the most recent strata in humans—were formed by Y inversions, which give support to the model proposed by Lahn and Page (1999) that X–Y recombination has been

suppressed by Y inversions. For stratum 3—an older stratum—we could not find evidence for such a Y inversion. This may be due to a wrong definition of the strata 3/4 boundary, to the formation of stratum 3 by a mechanism different from Y inversions or more likely to the existence of several strata (maybe two) within stratum 3 (see section on discussing the case of the ancient human strata). This illustrates the difficulty of working on an ancient stratum; stratum 3 was formed before the radiation of the principal placental mammalian orders (Lahn and Page 1999). In such an ancient stratum, high levels of X–Y divergence, paucity of Y markers, Y deletions in great number, and a multitude of X–Y rearrangements make reliable inferences difficult. The task is even more difficult for strata 1 and 2. These strata were formed at the very early stages of the XY evolution probably before the placental–marsupial split (Lahn and Page 1999, for the age of the human XY; see Rens et al. 2007; Potrzebowski et al. 2008; Veyrunes et al. 2008). These strata are extremely differentiated, with only a handful of markers still detectable on both sex chromosomes and with a massive gene loss on the Y chromosome (stratum 1: X = 588 genes, Y = 3 genes; stratum 2: X = 151, Y = 2 genes, from ENSEMBL v49 data). Moreover, the X region with strata 1, 2, and its homolog in chicken are rearranged (see Ross et al. 2005) and the absence of outgroups makes almost impossible the reconstruction of the X–Y rearrangements for these strata. We probably need to look at more recently evolved sex chromosome systems such as *S. latifolia* XY to obtain more data on the formation of strata, especially the first ones.

Supplementary Material

Supplementary data are available at Genome Biology and Evolution online (http://www.oxfordjournals.org/our_journals/gbe/).

Acknowledgments

We thank Laurent Duret, Mark Kirkpatrick, Eduardo Rocha, and David Sankoff for helpful comments on this article. This work was partially supported by the Agence Nationale de la Recherche (ANR) (REGLIS project NT05-3_45205 and GENOMICRO project ANR-05-JCJC-0139-01 and SiXY project ANR-08-JCJC-0109). M.D.V.B. is supported by the Programme AlBan, the European Union Programme of High Level Scholarships for Latin America, scholarship no. E05D053131BR.

Literature Cited

Andolfatto P, Depaulis F, Navarro A. 2001. Inversion polymorphisms and nucleotide variability in *Drosophila*. Genet Res. 77(1):1–8.

Bachtrog D. 2008. The temporal dynamics of processes underlying Y chromosome degeneration. Genetics. 179(3):1513–1525.

Bergero R, Charlesworth D, Filatov DA, Moore RC. 2008. Defining regions and rearrangements of the *Silene latifolia* Y chromosome. Genetics. 178(4):2045–2053.

Bergero R, Forrest A, Kamau E, Charlesworth D. 2007. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia:* evidence from new sex-linked genes. Genetics. 75:1945–1954.

Bergeron A, Chauve C, Hartmann T, St-Onge K. 2002. On the properties of sequences of reversals that sort a signed permutation. JOBIM. p. 99–108.

Bhowmick BK, Satta Y, Takahata N. 2007. The origin and evolution of human ampliconic gene families and ampliconic structure. Genome Res. 17(4):441–450.

Bhutkar A, et al. 2008. Chromosomal rearrangement inferred from comparisons of twelve *Drosophila* genomes. Genetics. 179:1657–1680.

Braga MDV, Sagot M-F, Scornavacca C, Tannier E. 2008. Exploring the solution space of sorting by reversals with experiments and an application to evolution, transactions on computational biology and bioinformatics. IEEE/ACM Trans Comput Biol Bioinform. 5:348–356.

Cáceres M. Sullivan RT, Thomas JW. National Institutes of Health Intramural Sequencing Center Comparative Sequencing Program. 2007. A recurrent inversion on the eutherian X chromosome. Proc Natl Acad Sci USA. 104(47):18571–18576.

Carrel L, Willard HF. 2005. X-inactivation profile reveals extensive variability in X-linked gene expression in females. Nature. 434:400–404.

Casals F, Navarro A. 2007. Chromosomal evolution: inversions: the chicken or the egg? Heredity. 99(5):479–480.

Charlesworth B. 1974. Inversion polymorphism in a two-locus genetic system. Genet Res. 23:259–280.

Charlesworth B, Charlesworth D. 1973. Selection of new inversion in multi-locus genetic systems. Genet Res. 21:167–183.

Charlesworth D, Charlesworth B. 1978. A model for the evolution of dioecy and gynodioecy. Am Nat. 112:975–997.

Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity. 95(2):118–128.

Dobzhansky T. 1950. The genetics of natural populations. XIX. Origin of heterosis through natural selection in populations of *Drosophila pseudoobscura*. Genetics. 35:288–302.

Evans AL, Mena PA, McAllister BF. 2007. Positive selection near an inversion breakpoint on the neo-X chromosome of *Drosophila americana*. Genetics. 177(3):1303–1319.

Fischer G, Rocha EP, Brunet F, Vergassola M, Dujon B. 2006. Highly variable rates of genome rearrangements between hemiascomycetous yeast lineages. PLoS Genet. 2(3):e32.

Handley LJ, Ceplitis H, Ellegren H. 2004. Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. Genetics. 167:367–76.

Hannenhalli S, Pevzner PA. 1999. Transforming cabbage into turnip: polynomial algorithm for sorting signed permutations by reversals. Journal of the ACM. 46(1):1–27.

Iwase M, et al. 2003. The amelogenin loci span an ancient pseudoautosomal boundary in diverse mammalian species. Proc Natl Acad Sci USA. 100(9):5258–5263.

Kehrer-Sawatzki H, Cooper DN. 2008. Molecular mechanisms of chromosomal rearrangement during primate evolution. Chromosome Res. 16(1):41–56.

Kehrer-Sawatzki H, Sandig CA, Goidts V, Hameister H. 2005. Breakpoint analysis of the pericentric inversion between chimpanzee chromosome 10 and the homologous chromosome 12 in humans. Cytogenet Genome Res. 108(1–3): 91–97.

Kirkpatrick M, Barton N. 2006. Chromosome inversions, local adaptation and speciation. Genetics. 173(1):419–434.

Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. Science. 286:964–967.

Lemaitre C, Tannier E, Gautier C, Sagot M-F. 2008. Precise detection of rearrangement breakpoints in mammalian chromosomes. BMC Bioinformatics. 9:286.
Marais G, Galtier N. 2003. Sex chromosomes: how X-Y recombination stops. Curr Biol. 13(16):R641–R643.
McAllister BF. 2003. Sequence differentiation associated with an inversion on the neo-X chromosome of *Drosophila americana*. Genetics. 165(3):1317–1328.
Murphy WJ, et al. 2005. Dynamics of mammalian chromosome evolution inferred from multispecies comparative maps. Science. 309(5734):613–617.
Nam K, Ellegren H. 2008. Scrambled eggs: the chicken (*Gallus gallus*) Z chromosome contains at least three non-linear evolutionary strata. Genetics. 180:1131–1136.
Navarro A, Barton NH. 2003. Accumulating postzygotic isolation genes in parapatry: a new twist on chromosomal speciation. Evolution. 57:447–459.
Navarro A, Betrán E, Barbadilla A, Ruiz A. 1997. Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. Genetics. 146(2):695–709.
Nei M. 1969. Linkage modifications and sex difference in recombination. Genetics. 63(3):681–699.
Nicolas M, et al. 2005. A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. PLoS Biol. 3:e4.
Noor MA, Grams KL, Bertucci LA, Reiland J. 2001. Chromosomal inversions and the reproductive isolation of species. Proc Natl Acad Sci USA. 98:12084–12088.
Otto SP, Barton NH. 2001. Selection for recombination in small populations. Evolution. 55:1921–1931.
Pearks Wilkerson AJ. 2008. Gene discovery and comparative analysis of X-degenerate genes from the domestic cat Y chromosome. Genomics. 92:329–338.
Potrzebowski L, et al. 2008. Chromosomal gene movements reflect the recent origin and biology of therian sex chromosomes. PLoS Biol. 6(4):e80.
Ranz JM, et al. 2007. Principles of genome evolution in the *Drosophila melanogaster* species group. PLoS Biol. 5(6):e152.
Rens W, et al. 2007. The multiple sex chromosomes of platypus and echidna are not completely identical and several share homology with the avian Z. Genome Biol. 8(11):R243.
Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 16(6):276–277.
Rieseberg LH. 2001. Chromosomal rearrangements and speciation. Trends Ecol Evol. 16:351–358.
Ross MT, et al. 2005. The DNA sequence of the human X chromosome. Nature. 434:325–337.
Rozen S, et al. 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature. 423(6942):873–876.
Sandstedt SA, Tucker PK. 2004. Evolutionary strata on the mouse X chromosome correspond to strata on the human X chromosome. Genome Res. 14:267–272.
Schwartz S, et al. 2003. Human-mouse alignments with BLASTZ. Genome Res. 13:103–107.
Siepel A. 2003. An algorithm to enumerate sorting reversals for signed permutations. J Comput Biol. 10:575–597.
Skaletsky H, et al. 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature. 423:825–837.
Sperlich D, Pfriem P. 1986. Chromosomal polymorphism in natural and experimental populations. In: Ashburner M, Carson HL, Thomson JN, editors. The genetics and biology of Drosophila, 3e. New York: Academic Press. p. 257–309.
Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res. 34:W609–W612.
Tesler G. 2002. GRIMM: genome rearrangements web server. Bioinformatics. 18(3):492–493.
Van Laere AS, Coppieters W, Georges M. 2008. Characterization of the bovine pseudoautosomal boundary: documenting the evolutionary history of mammalian sex chromosomes. Genome Res. 18:1884–1895.
Veyrunes F, et al. 2008. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. Genome Res. 18(6):965–973.
Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13(5):555–556.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24(8):1586–1591.
Yogeeswaran K, et al. 2005. Comparative genome analyses of *Arabidopsis* spp.: inferring chromosomal rearrangement events in the evolutionary history of *A. thaliana*. Genome Res. 15(4):505–515.
Zhou Q, et al. 2008. Neo-sex chromosomes in the black muntjac recapitulate incipient evolution of mammalian sex chromosomes. Genome Biol. 9(6):R98.

Kateryna Makova, Associate Editor

Accepted April 12, 2009