Title: Transposable element expression with variation in sex chromosome number supports a toxic Y effect on human longevity

Authors: Jordan Teoli[1,2,3,4], Marie Fablet[2,5], Claire Bardel[1,2], Anamaria Necsulea[2], Audrey Labalme[6], Hervé Lejeune[7], Jean-François Lemaitre[2], François Gueyffier[2], Damien Sanlaville[6], Cristina Vieira[2†*], Gabriel AB Marais[2,4,8,9†*], Ingrid Plotton[1,3†*]

Affiliations:

[1]Laboratoire de Biochimie et Biologie Moléculaire, Centre de Biologie et Pathologie Est, Hospices Civils de Lyon / Université Lyon 1; Bron, 69500, France.

[2]Laboratoire de Biométrie et Biologie Evolutive, UMR5558, CNRS / Université Lyon 1; Villeurbanne, F-69622, France.

[3]Institut Cellule Souche et Cerveau (SBRI), Unité INSERM 1208, Centre de Recherche INSERM; Bron, 69500, France.

[4]CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto; Vairão, 4485-661, Portugal.

[5]Institut universitaire de France.

[6]Service de Génétique, Hospices Civils de Lyon / Université Lyon 1; Bron, 69500, France.

[7]Service de médecine de la reproduction, Hôpital Femme-Mère-Enfant, Hospices Civils de Lyon; Bron, 69500, France.

[8]Departamento de Biologia, Faculdade de Ciências, Universidade do Porto; Porto, 4099-002, Portugal.

[9]BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão; Vairão, 4485-661, Portugal.

†Equal contributions as senior authors

*Corresponding author. Email: ingrid.plotton@chu-lyon.fr, gmarais@cibio.up.pt, cristina.vieira@univ-lyon1.fr

Abstract:
Lifespan differences between sexes is a puzzling question in human and evolutionary biology. Sex chromosomes have been proposed as playing a central role, and the Y male chromosome has been suspected of having a toxic genomic impact in this trait. As Y chromosomes are typically enriched in transposable elements (TE), this hypothesis suggests that in aged individuals, in which TE are less repressed, an increase of somatic mutations and acceleration of the aging process will be observed. Using an unprecedent expression dataset from humans with atypical karyotypes we show an increased TE expression, related to the Y chromosome. These findings fit with previous reports that the 47,XYY human population displays a shorter lifespan, and support the existence of a toxic Y effect on men lifespan.

One Sentence Summary:

A surge of transposable element expression due to the Y chromosome may explain why men die before women.

Main Text:

Life expectancy at birth is in average five years longer in women than men (*1*), a female-biased gap in longevity which is observed in almost all human populations (*2*) and which also explain why most supercentenarians are women (*3*). The reasons for this consistent female-biased gap in longevity in humans is a long-lasting conundrum. Both cultural and biological factors have been invoked (*2*, *4*). For the latter, sex-specific production of hormones (e.g. androgens) as well as male-specific deleterious mitochondrial mutations are biological factors that - to some extent - contribute to the sex differences in survival prospects (*5*, *6*). Recently, a new hypothesis has been proposed to explain them (reviewed in (*7*, *8*)). Differential survival among males and females are pervasive across the tree of life (*9*). In both vertebrates and invertebrates, the heterogametic sex (male in XY systems, female in ZW ones) tend to die first, which suggests sex chromosomes could contribute to establishing sex gaps in longevity (*10*, *11*). As the heterogametic sex carries one type of each sex chromosome, it is therefore more frequently affected by recessive X-linked (or W-linked) deleterious genetic mutations (unguarded X hypothesis), but theory and empirical data suggests this mechanism is probably not sufficient to explain the entire sex gap in longevity (*12*). Moreover, Y (or W) chromosomes are often less recombinant, and repeated sequences tend to accumulate in these chromosomes. With age, epigenetic regulation mechanisms are less effective, and transposable elements (TEs) can be reactivated. This was observed in aging fruit flies resulting in a surge of TE expression and an increase of the somatic mutation rate (*13*), higher in aged males than in aged females flies (*14*). Experiments uncoupling phenotypic sex and sex chromosomes content have shown that reduced-longevity is caused by the presence of the Y, not by phenotypic sex (*14*). This phenomenon, called the toxic Y effect, has been documented only in Drosophila species so far (*15*), and is still controversial (*16*).

The human Y chromosome is relatively small (about 57 Mb) compared to the X chromosome (about 156 Mb) but is an extremely repeat-rich chromosome compared to the X chromosome or autosomes (*17*, *18*). If a toxic Y effect exists in humans, we should observe that - everything else being equal - more TEs get reactivated in old 46,XY men than in old 46,XX

women, as in Drosophila (hereafter refereed as prediction 1). Moreover, in humans, 47,XXY and 47,XYY individuals have strikingly different longevities (*19*, *20*). 47,XYY is indeed associated with a 10-years-reduction in longevity while 47,XXY to a 2-years-reduction only compared to 46,XY individuals. This is observed despite the extra-Y having smaller observable effects on the 47,XYY individual's biology and health than the extra-X in the 47,XXY karyotype resulting in the Klinefelter syndrome (e.g. delays in first acquisitions, infertility, decreased bone mineral density ,…) (*19*, *20*). If a toxic Y effect explains the aforementioned observations and contributes to establishing the human sex gap in longevity, TE expression should be associated with the number of Y chromosome in a karyotype (hereafter referred as prediction 2).

We tested prediction 1 using published blood RNA-seq data from the Genotype-Tissue Expression (GTEx) project. We compared TE expression in 46,XX female and 46,XY male individuals belonging to different age groups from 20 to 70 years old. We found that only a few TE families have expression patterns consistent with our prediction and that TE expression seems to show a peak of expression at 30-40 years old in males (see **Supplementary Materials - Analysis of TE expression in the GTEx dataset** and **Fig. S1 to S8, Tables S1 to S6, Data S1**). However, it was difficult to draw strong conclusions because of the intrinsic nature of the data from GTEx database. We have many donors in different conditions and many different sequencing runs which made the data quite noisy (potential confounding factor, no longitudinal follow-up of same individual, survival bias, etc…). Moreover, the rather low sequencing depth and read length are not ideal for studying TE expression (see **Supplementary Materials - Materials and Methods**).

To test prediction 2, we generated RNA-seq data from blood samples of individuals with several gonosome composition on the karyotypes. We obtained blood samples from 25 individuals of Lyon university hospital of different karyotype compositions (46,XX, 46,XY, 47,XXY and 47,XYY) (**Table S7**). For each sample, we performed deep RNA-seq for a median of 174 million reads (min-max: 132 million – 206 million) *per* individual, which is an unprecedent dataset (**Table S7**). We first analyzed gene expression (**Fig. S3, Data S1**) to validate our wet and dry lab protocols (see **Supplementary Materials - Materials and Methods**) and to assess if some gene aging pathways (Gene Ontology (GO) - term analysis) were impacted by karyotype. The number of significant differentially expressed protein-coding genes depended on the compared pair of karyotypes (**Fig. S9, Table S8**). Comparing 46,XX and 46,XY samples (typical male and female karyotypes) returned 0.10% and 0.09% of over-expressed and under-expressed protein-coding genes, respectively (**Table S8**). Previous work using much larger sample sizes yielded more sex-biased genes (*21*); we were thus conservative here. We also found that most (13/20, 65%) male-biased protein coding genes are Y-linked genes, and that most female-biased protein coding genes are known X chromosome inactivation escapees (**Fig. S10, Tables S8 to S10, Data S2 and S3**), as expected from previous work (*21*). It is indeed well-known that the extra-X chromosome in 47,XXY or other karyotypes (e.g triple X) are inactivated; only the genes escaping inactivation exhibit expression changes (*22*, *23*). Furthermore, we found *XIST* and *TSIX* genes expressed only in karyotypes containing two X chromosomes (47,XXY and 46,XX) which was also expected as these genes are implicated in the X inactivation occurring only in the presence of two or more X chromosomes (*24–26*). Some autosomal protein coding genes were also differentially expressed when comparing 46,XX and 46,XY samples (**Table S8**).

Using pairwise karyotype comparisons, we found that 47,XYY and 47,XXY samples are the ones with the highest number of differentially expressed protein coding genes (~5-6%) compared to 46,XX samples (**Fig. S9, Table S8**). A GO-term analysis revealed that overexpressed genes in 47,XYY samples compared to 47,XXY samples were included in biological pathways (cell cycle, DNA damage/telomere stress induced senescence, senescence-associated secretory phenotype, cellular senescence, oxidative stress induced senescence) which may be related to a faster aging and cancer risk in 47,XYY individuals compared to 47,XXY (**Fig. S11, Data S4**).

Concerning TE expression (**Fig. S3, Data S1**), a principal component analysis (PCA) revealed that males and females had different profiles of TE expression as indicated by the sex clustering in the PCA plot (**Fig. S12**), a pattern that we also found for gene expression (**Fig. S13**). Similarly to protein-coding genes, TE families exhibiting expression changes were more numerous in 47,XYY/46,XX and 47,XXY/46,XX karyotype comparisons (**Fig. S14, Table S11**): 46 differentially expressed TE families in 47,XYY and 58 in 47,XXY compared to 46,XX. Most differentially expressed TE families were found to be significantly over-expressed in 47,XYY (35 TE families), 47,XXY (51 TE families) and 46,XY (5 TE families) compared to 46,XX karyotype, while few were under-expressed (11, 7 and 1 TE families respectively) (**Fig. S14, Table S11**). The differentially expressed TE families according to karyotype is shown in **Fig. 1-A and S15**. A similar pattern was found for a majority of differentially expressed TE families: a minimal expression in 46,XX, a maximal expression in 47,XYY and an in-between expression in 46,XY and 47,XXY karyotypes, pointing to a contribution of the Y chromosome in establishing this pattern. This pattern was also observed combining all the TEs (**Fig. 1-B**). Furthermore, when splitting individuals of a given karyotype in two age groups (lower and higher than the median for this group), we showed that TE expression tends to be higher later in life for 46,XY and 47,XYY individuals, which is in line with prediction 1 although limited sample size prevented us to firmly conclude (**Fig. S16**).

This pattern could be driven by the expression of TE insertions present in genes and in long non-coding RNA. We have then re-analyzed the data excluding insertions in exonic parts of all genes and in long non-coding RNA insertions. We also removed non-autonomous expression from introns of X and Y chromosomes (see **Supplementary Materials – Materials and Methods and Fig. S17**). The results exhibited fewer significant differentially expressed TE families, probably due to a reduction in statistical power, because of a reduction of counts per TE families. However, the pattern concerning overall expression of TEs remained (**Fig. S18 to S21, Table S12**).

The analysis of our own data revealed strong changes in TE expression in 47,XYY individuals, and a clear trend of low expression in 46,XX individuals, high expression in 47,XYY and intermediate expression in 46,XY and 47,XXY individuals for some TE families and for the global TE expression. Regardless the age of the individuals, our results showed that TE global expression is higher in 47,XYY individuals compared to other karyotypes, although the comparison was significant only with the 46,XX individuals (**Fig. 1-B**). Interestingly, similar observations were made in Drosophila where Y-linked TE expression is stronger in individuals with an additional Y chromosome (i.e. XYY) compared to XY individuals regardless their age (*14*). Our results suggest that a similar disruption of TE regulation is present in young individuals with 47,XYY karyotype in humans. Furthermore, in Drosophila, Y-linked TE expression is

stronger in aged flies than young flies for XY and XYY individuals but not for XXY individuals (*14*), which was in line with the trend we found in humans (**Fig. S16**).

Our analysis support both predictions of the toxic Y chromosome hypothesis, in particular with the data set that we have generated. We show an over-expression of TEs related to the number of Y-chromosomes in the karyotype, and suggest that a toxic Y effect may contribute to the sex gap in longevity in humans as showed in Drosophila ((*14*, *15*) but see (*16*)).

Sex gap in longevity is certainly a multifactorial trait and the Y toxic effect is just one of them. Our results indicate that relaxing the TE control with aging or adding extra Y chromosomes will contribute to an increase of TE expression and potentially to an increase of the somatic mutations. We also should keep in mind that an equilibrium is most certainly established in the interactions between chromosomes, and that adding Y chromosomes is disrupting it. This study also suggests that future research in aging treatments need to be sex related and most probably TE content related. The use of antiviral medication may have drastic effects on TE activity, and may be one clue for future research.

References and Notes:

1. World Health Organization, Life expectancy at birth (years), (available at https://www.who.int/data/gho/data/indicators/indicator-details/GHO/life-expectancy-at-birth-(years)).

2. T. L. Rochelle, D. K. Y. Yeung, M. H. Bond, L. M. W. Li, Predictors of the gender gap in life expectancy across 54 nations. *Psychology, Health & Medicine* **20**, 129–138 (2015).

3. B. J. Willcox, D. C. Willcox, L. Ferrucci, Secrets of Healthy Aging and Longevity From Exceptional Survivors Around the Globe: Lessons From Octogenarians to Supercentenarians. *The Journals of Gerontology: Series A* **63**, 1181–1185 (2008).

4. M. Luy, Causes of Male Excess Mortality: Insights from Cloistered Populations. *Population and Development Review* **29**, 647–676 (2003).

5. A. A. Maklakov, V. Lummaa, Evolution of sex differences in lifespan and aging: Causes and constraints: Prospects & Overviews. *BioEssays* **35**, 717–724 (2013).

6. E. Milot, C. Moreau, A. Gagnon, A. A. Cohen, B. Brais, D. Labuda, Mother's curse neutralizes natural selection against a human genetic disease over three centuries. *Nature Ecology & Evolution* **1**, 1400–1406 (2017).

7. G. Marais, J.-M. Gaillard, C. Vieira, I. Plotton, D. Sanlaville, F. Gueyffier, J.-F. Lemaitre, Sex gap in aging and longevity: can sex chromosomes play a role? *Biology of Sex Differences* **9**, 33 (2018).

8. G. Marais, J.-F. Lemaitre, C. Vieira, Y chromosome makes fruit flies die younger. *PCI Evol Biol*, 100105 (2020).

9. J.-F. Lemaître, V. Ronget, M. Tidière, D. Allainé, V. Berger, A. Cohas, F. Colchero, D. A. Conde, M. Garratt, A. Liker, G. A. B. Marais, A. Scheuerlein, T. Székely, J.-M. Gaillard, Sex differences in adult lifespan and aging rates of mortality across wild mammals. *Proceedings of the National Academy of Sciences* **117**, 8546–8553 (2020).

10. I. Pipoly, V. Bókony, M. Kirkpatrick, P. F. Donald, T. Székely, A. Liker, The genetic sex-determination system predicts adult sex ratios in tetrapods. *Nature* **527**, 91–94 (2015).

11. Z. A. Xirocostas, S. E. Everingham, A. T. Moles, The sex with the reduced sex chromosome dies earlier: a comparison across the tree of life. *Biol. Lett* **16**, 20190867 (2020).

12. T. Connallon, I. J. Beasley, Y. MDonough, F. Ruzicka, How much does the unguarded X contribute to sex differences in life span? *Evolution Letters* **6**, 319–329 (2022).

13. A. Tsurumi, W. Li, Global heterochromatin loss: A unifying theory of aging? *Epigenetics* **7**, 680–688 (2012).

14. E. J. Brown, A. H. Nguyen, D. Bachtrog, The Y chromosome may contribute to sex-specific ageing in Drosophila. *Nature Ecology & Evolution* **4**, 853–862 (2020).

15. K. H.-C. Wei, L. Gibilisco, D. Bachtrog, Epigenetic conflict on a degenerating Y chromosome increases mutational burden in Drosophila males. *Nat Commun* **11**, 5537 (2020).

16. R. Delanoue, C. Clot, C. Leray, T. Pihl, B. Hudry, Y chromosome toxicity does not contribute to sex-specific differences in longevity. *Nat Ecol Evol* (2023), doi:10.1038/s41559-023-02089-7.

17. H. Skaletsky, T. Kuroda-Kawaguchi, P. J. Minx, H. S. Cordum, L. Hillier, L. G. Brown, S. Repping, T. Pyntikova, J. Ali, T. Bieri, A. Chinwalla, A. Delehaunty, K. Delehaunty, H. Du, G. Fewell, L. Fulton, R. Fulton, T. Graves, S.-F. Hou, P. Latrielle, S. Leonard, E. Mardis, R. Maupin, J. McPherson, T. Miner, W. Nash, C. Nguyen, P. Ozersky, K. Pepin, S. Rock, T. Rohlfing, K. Scott, B. Schultz, C. Strong, A. Tin-Wollam, S.-P. Yang, R. H. Waterston, R. K. Wilson, S. Rozen, D. C. Page, The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825–837 (2003).

18. Human Genome Assembly GRCh38.p14, (available at https://www.ncbi.nlm.nih.gov/grc/human/data).

19. A. Bojesen, S. Juul, N. Birkebæk, C. H. Gravholt, Increased Mortality in Klinefelter Syndrome. *The Journal of Clinical Endocrinology & Metabolism* **89**, 3830–3834 (2004).

20. K. Stochholm, S. Juul, C. H. Gravholt, Diagnosis and mortality in 47,XYY persons: a registry study. *Orphanet Journal of Rare Diseases* **5**, 15 (2010).

21. M. Oliva, M. Muñoz-Aguirre, S. Kim-Hellmuth, V. Wucher, A. D. H. Gewirtz, D. J. Cotter, P. Parsana, S. Kasela, B. Balliu, A. Viñuela, S. E. Castel, P. Mohammadi, F. Aguet, Y. Zou, E. A. Khramtsova, A. D. Skol, D. Garrido-Martín, F. Reverter, A. Brown, P. Evans, E. R. Gamazon, A. Payne, R. Bonazzola, A. N. Barbeira, A. R. Hamel, A. Martinez-Perez, J. M. Soria, GTEx Consortium§, B. L. Pierce, M. Stephens, E. Eskin, E. T. Dermitzakis, A. V. Segrè, H. K. Im, B. E. Engelhardt, K. G. Ardlie, S. B. Montgomery, A. J. Battle, T. Lappalainen, R. Guigó, B. E. Stranger, The impact of sex on gene expression across human tissues. *Science* **369**, eaba3066 (2020).

22. K. Monkhorst, I. Jonkers, E. Rentmeester, F. Grosveld, J. Gribnau, X Inactivation Counting and Choice Is a Stochastic Process: Evidence for Involvement of an X-Linked Activator. *Cell* **132**, 410–421 (2008).

23. X. Zhang, D. Hong, S. Ma, T. Ward, M. Ho, R. Pattni, Z. Duren, A. Stankov, S. Bade Shrestha, J. Hallmayer, W. H. Wong, A. L. Reiss, A. E. Urban, Integrated functional genomic analyses of Klinefelter and Turner syndromes reveal global network effects of altered X chromosome dosage. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 4864–4873 (2020).

24. A. Loda, E. Heard, Xist RNA in action: Past, present, and future. *PLOS Genetics* **15**, e1008333 (2019).

25. F. M. Hassan, H. A. A. Razik, M. S. Wadie, D. S. Abdelfattah, XIST and RPS4Y1 long non-coding RNA transcriptome as sex biomarkers in different body fluids. *Egyptian Journal of Forensic Sciences* **9**, 16 (2019).

26. A. Skakkebæk, M. M. Nielsen, C. Trolle, S. Vang, H. Hornshøj, J. Hedegaard, M. Wallentin, A. Bojesen, J. M. Hertz, J. Fedder, J. R. Østergaard, J. S. Pedersen, C. H. Gravholt, DNA hypermethylation and differential gene expression associated with Klinefelter syndrome. *Scientific Reports* **8**, 13740 (2018).

27. J. Lonsdale, J. Thomas, M. Salvatore, R. Phillips, E. Lo, S. Shad, R. Hasz, G. Walters, F. Garcia, N. Young, B. Foster, M. Moser, E. Karasik, B. Gillard, K. Ramsey, S. Sullivan, J. Bridge, H. Magazine, J. Syron, J. Fleming, L. Siminoff, H. Traino, M. Mosavel, L. Barker, S. Jewell, D. Rohrer, D. Maxim, D. Filkins, P. Harbach, E. Cortadillo, B. Berghuis, L. Turner, E. Hudson, K. Feenstra, L. Sobin, J. Robb, P. Branton, G. Korzeniewski, C. Shive, D. Tabor, L. Qi, K. Groch, S. Nampally, S. Buia, A. Zimmerman, A. Smith, R. Burges, K. Robinson, K. Valentino, D. Bradbury, M. Cosentino, N. Diaz-Mayoral, M. Kennedy, T. Engel, P. Williams, K. Erickson, K. Ardlie, W. Winckler, G. Getz, D. DeLuca, D. MacArthur, M. Kellis, A. Thomson, T. Young, E. Gelfand, M. Donovan, Y. Meng, G. Grant, D. Mash, Y. Marcus, M. Basile, J. Liu, J. Zhu, Z. Tu, N. J. Cox, D. L. Nicolae, E. R. Gamazon, H. K. Im, A. Konkashbaev, J. Pritchard, M. Stevens, T. Flutre, X. Wen, E. T. Dermitzakis, T. Lappalainen, R. Guigo, J. Monlong, M. Sammeth, D. Koller, A. Battle, S. Mostafavi, M. McCarthy, M. Rivas, J. Maller, I. Rusyn, A. Nobel, F. Wright, A. Shabalin, M. Feolo, N. Sharopova, A. Sturcke, J. Paschal, J. M. Anderson, E. L. Wilder, L. K. Derr, E. D. Green, J. P. Struewing, G. Temple, S. Volpi, J. T. Boyer, E. J. Thomson, M. S. Guyer, C. Ng, A. Abdallah, D. Colantuoni, T. R. Insel, S. E. Koester, A. R. Little, P. K. Bender, T. Lehner, Y. Yao, C. C. Compton, J. B. Vaught, S. Sawyer, N. C. Lockhart, J. Demchok, H. F. Moore, The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580–585 (2013).

28. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

29. T. Wicker, F. Sabot, A. Hua-Van, J. L. Bennetzen, P. Capy, B. Chalhoub, A. Flavell, P. Leroy, M. Morgante, O. Panaud, E. Paux, P. SanMiguel, A. H. Schulman, A unified classification system for eukaryotic transposable elements. *Nature Reviews Genetics* **8**, 973–982 (2007).

30. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal RNA-Seq quantification. *Nature Biotechnology* **34**, 525-527 (2016).

31. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359 (2012).

32. E. Lerat, M. Fablet, L. Modolo, H. Lopez-Maestre, C. Vieira, TEtools facilitates big data expression analysis of transposable elements and reveals an antagonism between their activity and that of piRNA genes. *Nucleic Acids Research* **45**, e17 (2016).

33. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biology* **11**, R106 (2010).

34. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550 (2014).

35. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289–300 (1995).

36. M. Stephens, False discovery rates: a new deal. *Biostatistics* **18**, 275-294 (2016).

37. U. Raudvere, L. Kolberg, I. Kuzmin, T. Arak, P. Adler, H. Peterson, J. Vilo, g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* **47**, W191–W198 (2019).

38. S. Holm, A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics* **6**, 65–70 (1979).

39. F. Ansaloni, N. Gualandi, M. Esposito, S. Gustincich, R. Sanges, TEspeX: consensus-specific quantification of transposable element expression preventing biases from exonized fragments. *Bioinformatics* **38**, 4430–4433 (2022).

Acknowledgments:

We thank Yasmine Zerdoumi and Maxime Vallée and for their help in the data generation and data analysis.

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The GTEx data used for the analyses described in this manuscript were obtained from: dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000424.v8.p2.

Funding: this study was funded by an internal collaborative grant of LBBE and by ANR (grant number ANR-20-CE02-0015).

Author contributions: Conceptualization: GABM, CV, FG, IP, DS, JFL; Data curation: JT; Formal analysis: JT, MF, CB; Funding acquisition: GABM, CV, FG, IP, DS, JFL; Investigation: JT, MF, CB, AN, AL, FG, DS, CV, IP, GABM; Methodology: MF, CB, DS, CV, IP, GABM; Project administration: IP, CV, GABM; Resources: IP, DS, HL; Software: JT, MF, CB, AN; Supervision: IP, GABM, CV, MF, AN, CB, HL; Validation: JT, DS, CV, IP, GABM; Visualization: JT; Writing – original draft: GABM, CV, IP, JT; Writing – review & editing: all authors.

Competing interests: Authors declare no competing interests.

Data and materials availability: Fastq files of our dataset were submitted to …

Supplementary Materials:

Materials and Methods

Supplementary text

Figs. S1 to S21

Tables S1 to S12

References (*27-39*)

Data S1-S4

Figure captions

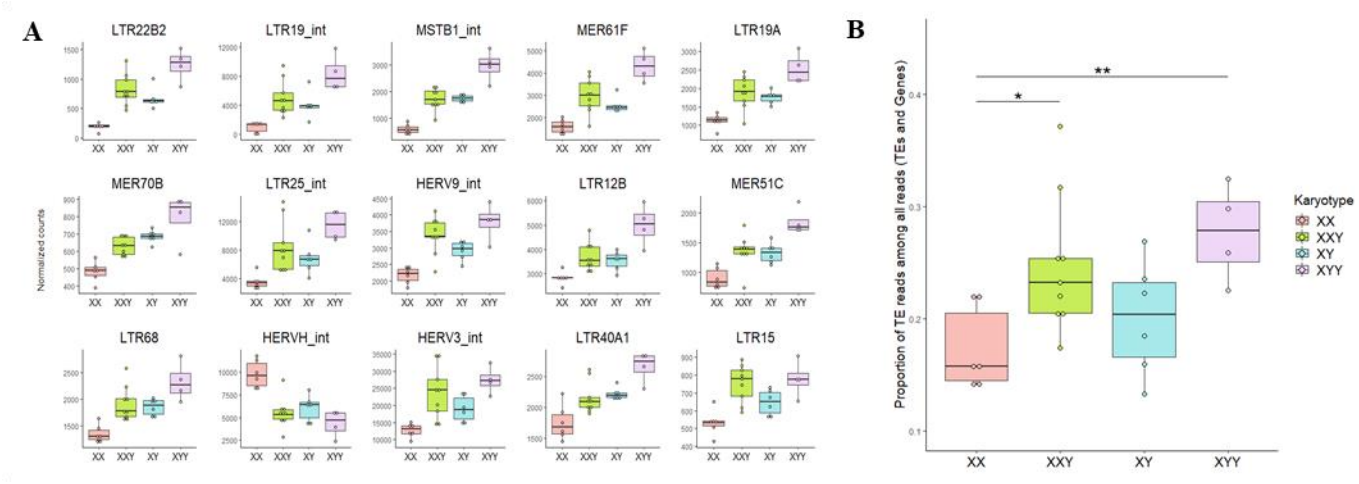

Fig. 1. TE expression in individuals with normal and atypical sex chromosome content in our dataset. (A) Boxplots of the 15 most significant differentially expressed TE families according to the karyotype using LR test, after removing batch effect. (B) Boxplots of normalized counts combining all TE families after removing batch effect regardless age. (**A**) Boxplots for the other significant differentially expressed TE families are in Supplemental material (**Fig. S15**). All of these 15 families had at least one copy on the Y chromosome and one copy on the X chromosome in the reference genome. (**B**) Raw p-values were less than 0.05 (*) for 47,XXY samples compared to 46,XX samples (P = 0.012), and less than 0.01 (**) for 47,XYY samples compared to 46,XY samples (P = 0.0095) and near 0.05 for 47,XYY samples compared to 46,XY samples (P = 0.067) using Wilcoxon test. Adjusted p-values using Holm method were no longer significant (P = 0.060, 0.057 and 0.267 respectively).