Adaptation of the autosomal part of the genome on the presence of dioecy

Jitka Žlůvová*, Roman Hobza, Bohuslav Janoušek*

Department of Plant Developmental Genetics, Institute of Biophysics, Academy of Sciences of the Czech Republic, Královopolská 135, Brno, Czech Republic.

* corresponding authors

Email: jitka@ibp.cz, janousek@ibp.cz

Author Contributions: J.Z. and B.J. designed research; R.H. contributed new analysis tools; J.Z. and B.J. analyzed data; R.H. discussed the paper; and J.Z. and B.J. wrote the paper.

Competing Interest Statement: The authors declare no competing interest.

Keywords: genome evolution, adaptive evolution, dioecy

This file includes:

Abstract

We have attempted to answer the question of whether the presence of sex chromosomes in the genome can affect the evolution of the autosomal part of the genome. As a model, we used dioecious plants from the section *Otites* of the genus *Silene*. We have observed a rise in adaptive evolution in the autosomal and pseudoautosomal parts of the genome, which are associated with the evolution of dioecy. This rise is caused neither by the accumulation of sexually antagonistic genes in the pseudoautosomal region nor by the co-evolution of genes acting in mitochondria (in spite of the fact that the dioecy evolved in this case most likely from cytoplasmic male sterility). Thus, this rise in the amount of positively selected codons is most likely caused by the adaptive evolution of genes involved in the specialization of the autosomal part of the genome on the dioecy as described in sex-allocation theory.

Introduction

Separate sexes are very common in animals, but they also appear in plants (Renner & Müller, 2021). After sex determining gene(s) appear in the genome, many evolutionary processes are started on the sex chromosomes (Charlesworth, 2019). Moreover, many autosomal genes express in a sex-specific manner, and their expression is regulated by the presence of sex-determining gene(s). These sex-specific differences in gene expression lead to the formation of male and female individuals. But can the presence of separate sexes lead to evolutionary changes in the sequences of autosomal genes? Because males invest more in the quantity of their progeny, whereas the goal of the females is to invest in the quality of the progeny, the presence of separate sexes could lead to the adaptation of the autosomal genes to males' and females' different breeding strategies.

In animals, separate sexes and sex chromosomes are of very old origin, and thus it could be very difficult to search for signs of the adaptation of the autosomal part of the genome to the presence of separate sexes. On the other hand, many plant genera contain recently evolved dioecious clades. One of the plant genera with young sex chromosomes is the genus *Silene*. The dioecious *Silene* species are especially well-suited to this study because closely related non-dioecious species are known, transcriptomic data for both the dioecious species and their non-dioecious relatives are known, and genetic maps are available in several species (reviewed in (Balounova et al., 2019)). In the genus *Silene*, apart from the dosage compensation studies (Martin et al., 2019; Muyle et al., 2018), adaptation of the transcriptome to the dioecy was also studied from a quantitative point of view (Zemp, Widmer, & Charlesworth, 2018). However, it is not known whether the presence of sex chromosomes in the genome can affect the evolution of the coding regions in the autosomal part of the genome.

Results and discussion

In this study, we use the currently available RNAseq data obtained in the dioecious section *Otites* to study the influence of the rise of dioecy on the evolution of the sequences that are not sex-linked. Results of the phylogenetic analysis shown in Fig. 1 are mostly in accordance with the results from the previous study that did not include *S. nocturna* and *S. paradoxa*. A minor difference in topology is that the current results support the clade joining two members

of the group Cyri (defined previously (Balounova et al., 2019) according to phenotypic markers), while in our previous study, the group Cyri appeared as completely polyphyletic (Balounova et al., 2019). Current dating, based on the estimation of the synonymous divergence of the outgroup (S. nocturna) from the other species, suggests the age of dioecious section *Otites* cca 2.3 million years which is in the range of the previous estimate based on calibration by fossils (1.17–2.60 million years) (Slancarova et al., 2013).

The ω ratio (ω = dN/dS) is one of the most widely used statistical tests used to quantify selection pressures acting on protein-coding regions. This measure quantifies selection pressures by comparing the rate of substitutions at silent sites (dS), which are presumed neutral, to the rate of substitutions at non-silent sites (dN), which possibly experience selection. The comparison of the two- and three-ratios branch models in the CODEML program of PAML package (Yang, 2007) shows that the ω value of the branch preceding sex chromosome evolution (ω = 0.16; blue in Fig. 1) is lower than the ω value of the branch where dioecy and sex chromosomes evolved (ω = 0.29; red in Fig. 1). This difference is highly significant ($P < 10^{-99}$, likelihood ratio test; LRT). After the formation of the dioecy (green in Fig. 1), the ω value significantly decreased (ω = 0.26; $P < 10^{-99}$, LRT).

Because the increase of the ω values can be caused by the changes in the number of either positively or neutrally selected sites, we compared branch-site models of these branches. The results are summarized in Table 1. The branch where dioecy evolved showed a significant percentage of positively selected codons ($P < 10^{-99}$, LRT). On the other hand, we did not detect codons under positive selection in the branch preceding dioecy (blue in Fig. 1) ($P > 0.99$, LRT). Moreover, the ω values of sites under purifying selection did not increase significantly ($P = 0.95$, Wilcoxon test). To confirm these results, we performed tests for relaxed selection using RELAX program (Wertheim, Murrell, Smith, Kosakovsky Pond, & Scheffler, 2015) of HyPhy package (Kosakovsky Pond et al., 2020; Pond, Frost, & Muse, 2005). Transition from the gynodioecy to the dioecy was connected with the selection intensification ($P = 0.00$; $K = 1.55$). On the other hand, when the branch involving dioecy evolution and the branch after sex chromosome evolution were compared, non-significant relaxation was detected ($P = 0.339$, $K = 0.91$). Most of the identified adaptively evolved codons (63 out of 87 codons) are recruited from the neutrally evolved codons, which is significantly more than by chance ($P < 10^{-16}$, chi-squared goodness of fit test). Because the dioecy in the *Silene* section *Otites* most likely evolved from the gynodioecy, and because the gynodioecy is in the genus *Silene* usually of nucleo-cytoplasmic type, the dioecy evolved, in

this case, most likely by the genetic fixation of a male sterile cytoplasm and subsequent recruitment of a fertility restorer as sex determining gene, as discussed by (Zluvova et al., 2005). In this case, no drastic change in population size or rate of inbreeding is necessary to open the route to dioecy. The observed absence of any sign of selective pressure relaxation is in accordance with this hypothesis.

The rise of the amount of positively selected codons on the branch including sex chromosome evolution cannot be caused by the selection of sexually antagonistic genes on the sex chromosomes (Rice, 1996), because the analysed dataset does not contain any completely sex-linked gene. Moreover, it cannot be caused by the accumulation of sexually antagonistic genes in the pseudoautosomal region, because we found that only one of the studied ESTs was located in this part of the sex chromosomes.

Because all close relatives of the analyzed species are gynodioecious, most likely with cytoplasmic male sterility, one plausible explanation for the rise of adaptive evolution can be connected with the presumed nucleo-cytoplasmic origin of the sex-determining system. The fixation of a certain type of male-sterility causing cytoplasm could influence the evolution of nuclear genes involved in mitochondrial metabolism. However, among the 42 ESTs identified as having at least one adaptively evolved codon on the branch involving sex chromosome formation, we did not find an overrepresentation of proteins located in mitochondria (7 ESTs; $P = 0.77$, Pearson's chi-squared test).

We can hypothesize that this rise in the amount of positively selected codons is caused by the selection of genes involved in the specialization of the genome on the dioecy as described in the sex-allocation theory (Charnov, 1982), which states that female plants allocate more resources to the quality of seeds whereas male plants allocate resources into the amount of pollen. However, based on phenotypic data, it is possible to conclude that the adaptation of the genome to the dioecy is rather complex and includes a wide variety of genes (Geber, Dawson, & Delph, 1999). Indeed, the genes identified as adaptively evolving on the branch where dioecy evolved (see Supporting Information) show rather diverse characteristics. These results are in good accordance with the previous observations on the phenotypic level (Geber et al., 1999).

The discovered coincidence between the increased amount of adaptively evolved codons in autosomes and the evolution of dioecy cannot prove causality in this process. However, the

hypothesis that the changes in autosomes represent an adaptation to the sexually dimorphic phenotype of dioecious species appears most likely.

Materials and methods

The dataset from the previous work (Balounova et al., 2019) was supplemented by *S. nocturna* (SRR6040876) and *S. paradoxa* (SRR999296-SRR999299, publically available) samples. *S. pseudotites* has been excluded from the dataset based on its suspected hybrid origin. The reads of the *S. nocturna* and *S paradoxa* were assembled using Trinity (Haas et al., 2013). The assembly was treated with an Evigene pipeline to reduce the level of duplicates and used as a reference. The reads were mapped using Bowtie2 (Langmead & Salzberg, 2012), the SNPs were called via FreeBayes (Garrison & Marth, 2012) and phasing was performed with WhatsHap (Patterson et al., 2015). The regions with coverage lower than 2 masked using the maskfasta method in BEDTools (Quinlan & Hall, 2010). The phased sequences were added to the original alignments using the MAFFT aligner (Katoh & Standley, 2013) based on the results of best reciprocal blast hit search (Camacho et al., 2009). Phylogenetic analysis was done using the StarBeast2 module in BEAST 2 (Bouckaert et al., 2014; Ogilvie, Bouckaert, & Drummond, 2017). Two independent chains were run for 1000 000 000 states (Yule model, birth-death). For the dating, the calibration of the most recent common ancestor of *S. nocturna* and the other species was done similarly to Balounova et al. (Balounova et al., 2019) (based on estimated divergence per synonymous site dS (substitutions per synonymous site) and the divergence time estimates for several angiosperm lineages (Brassicaceae, Malvaceae, Euphorbiaceae, Fabaceae, Cucurbitaceae, Rosaceae, Solanaceae, and Poaceae), so as not to depend on a single fossil record or phylogenetic tree position, which yielded a mean substitution rate of 5.35×10^{-9} synonymous substitutions/site/year) (De La Torre, Li, Van de Peer, & Ingvarsson, 2017). Ks values for the distance of individual species of the section *Otites* to *S. nocturna* were estimated using KaKs calculator 2 program (Wang, Zhang, Zhang, Zhu, & Yu, 2010), and the mean value was used for the calibration of the tree. The resulting dataset did not contain any of the completely sex-linked genes identified previously (Balounova et al., 2019) (Martin et al., 2019).

PAML analyses were used to determine whether some ESTs evolved under selective pressure. The CODEML program of PAML (Yang, 2007) was used to estimate the ratio (ω) of the non-synonymous substitution rate (dN) to the synonymous substitution rate (dS). As

the reference tree, the phylogenetic tree constructed as described above was used. The equilibrium frequencies of codons were calculated from the nucleotide frequencies (CodonFreq=2) because it best fits the data as calculated by second-order AIC. All models in CODEML were run at five different initial ω values (ω = 0.1, 0.2, 0.5, 1, 2), and no problems with the convergence were observed. Both branch and branch-site models were applied to the branch preceding sex chromosome formation, to the branch including sex chromosome formation and to the branch following the sex chromosome formation. In the branch analyses, two-ratios models were compared to three-ratios models to reveal whether the ω values differ significantly. The resulting log likelihood values were evaluated using likelihood-ratio tests to determine any statistical significance of the difference. The chi2 program of PAML was used to estimate the P-values. The confidence intervals for proportion were computed online (http://vassarstats.net/prop1.html) using the method with continuity correction (Newcombe, 1998) that is derived from a procedure outlined in (Wilson, 1927). The results obtained in CODEML were further confirmed using a test for relaxed selection in RELAX program (Wertheim et al., 2015) of the HypHy package (Pond et al., 2005) (Kosakovsky Pond et al., 2020). Because the results did not converge well, the analysis was run 500-times and the results with the highest log likelihood were used.

The modified model A of the branch-site models was used to compute the percentages of purifying, neutral and adaptive codons in the analyzed ESTs, to compute the ω values, to compute the Bayes Empirical Bayes probability of codons being either purifying, neutral or adaptive and to identify ESTs with adaptively evolved codons. The Wilcoxon test computed online (https://www.socscistatistics.com/tests/signedranks/default2.aspx) was used to compare the ω values of the codons with $\omega < 1$ between the branches before and during sex chromosome and dioecy formation. The chi-squared goodness of fit test was used to calculate the P-value of adaptively evolving codons to evolve from neutrally evolving codons. The Pearson's chi-squared test was used to compute the P-value of a random presence of genes acting in mitochondria among the genes with positively selected codons.

Acknowledgments

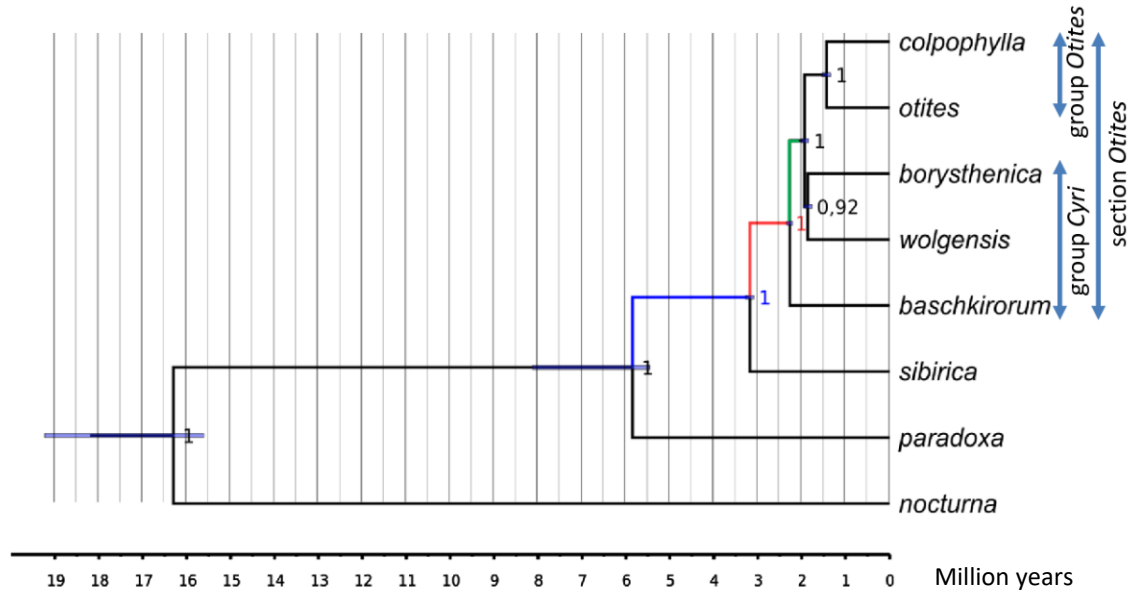
This research was supported by the Czech Science Foundation (grant 19-15609S). Computational resources were supplied by the project "e-Infrastruktura CZ" (e-INFRA LM2018140) provided within the program Projects of Large Research, Development and

187 Innovations Infrastructures and by the ELIXIR-CZ project (LM2018131), part of the
188 international ELIXIR infrastructure.
189

190 **Competing Interest Statement**

191 No financial and/or non-financial competing interests declared.

192 **Figures**
193



194
195
196
197 **Figure 1. Phylogenetic tree of the section *Otites* and the closely related species.** The
198 colors of the highlighted branches refer to analyses of adaptive analyses. For details, see
199 Results and Discussion.

Tables

	ω < 1		ω = 1		ω > 1	
branch	percentage (%)	95% CI of percentage difference (%)	percentage (%)	95% CI of percentage difference (%)	percentage (%)	95% CI of percentage difference (%)
preceding sex chromosome evolution	80.496		19.504		**0**	
		-0.19 - 0.31		-0.17 - 0.34		**0.02 - 0.03**
including sex chromosome evolution	80.557		19.419		**0.024**	
		-0.2 - 0.31		-0.17 - 0.32		**0.02 – 0.03**
following sex chromosome evolution	80.497		19.503		**0**	

Table 1. Branch-site analysis of *Silene* transcriptome before and after sex chromosome evolution. Confidence intervals (CI) are given for the difference of percentage between the neighboring rows. Statistically significant values are highlighted in bold.

206 **Bibliography**

207 Balounova, V., Gogela, R., Cegan, R., Cangren, P., Zluvova, J., Safar, J., … Janousek, B.
208 (2019). Evolution of sex determination and heterogamety changes in section Otites of the
209 genus Silene. *Scientific Reports*, *9*(1), 1045. doi: 10.1038/s41598-018-37412-x

210 Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J.
211 (2014). BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS*
212 *Computational Biology*, *10*(4), e1003537. doi: 10.1371/journal.pcbi.1003537

213 Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden,
214 T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics*, *10*, 421. doi:
215 10.1186/1471-2105-10-421

216 Charlesworth, D. (2019). Young sex chromosomes in plants and animals. *The New*
217 *Phytologist*, *224*(3), 1095–1107. doi: 10.1111/nph.16002

218 Charnov, E. L. (1982). The theory of sex allocation. *Monographs in Population Biology*, *18*,
219 1–355.

220 De La Torre, A. R., Li, Z., Van de Peer, Y., & Ingvarsson, P. K. (2017). Contrasting Rates of
221 Molecular Evolution and Patterns of Selection among Gymnosperms and Flowering
222 Plants. *Molecular Biology and Evolution*, *34*(6), 1363–1377. doi:
223 10.1093/molbev/msx069

224 Garrison, E., & Marth, G. (2012). *Haplotype-based variant detection from short-read*
225 *sequencing*.

226 Geber, M. A., Dawson, T. E., & Delph, L. F. (Eds.). (1999). *Gender and sexual dimorphism*
227 *in flowering plants*. Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-
228 662-03908-3

229 Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., …
230 Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the

231 Trinity platform for reference generation and analysis. *Nature Protocols*, *8*(8), 1494–
232 1512. doi: 10.1038/nprot.2013.084
233 Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version
234 7: improvements in performance and usability. *Molecular Biology and Evolution*, *30*(4),
235 772–780. doi: 10.1093/molbev/mst010
236 Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell,
237 B., … Muse, S. V. (2020). HyPhy 2.5-A Customizable Platform for Evolutionary
238 Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, *37*(1), 295–
239 299. doi: 10.1093/molbev/msz197
240 Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature*
241 *Methods*, *9*(4), 357–359. doi: 10.1038/nmeth.1923
242 Martin, H., Carpentier, F., Gallina, S., Godé, C., Schmitt, E., Muyle, A., … Touzet, P. (2019).
243 Evolution of Young Sex Chromosomes in Two Dioecious Sister Plant Species with
244 Distinct Sex Determination Systems. *Genome Biology and Evolution*, *11*(2), 350–361.
245 doi: 10.1093/gbe/evz001
246 Muyle, A., Zemp, N., Fruchard, C., Cegan, R., Vrana, J., Deschamps, C., … Marais, G. A. B.
247 (2018). Genomic imprinting mediates dosage compensation in a young plant XY system.
248 *Nature Plants*, *4*(9), 677–680. doi: 10.1038/s41477-018-0221-y
249 Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion:
250 comparison of seven methods. *Statistics in Medicine*, *17*(8), 857–872. doi:
251 10.1002/(sici)1097-0258(19980430)17:8<857::aid-sim777>3.0.co;2-e
252 Ogilvie, H. A., Bouckaert, R. R., & Drummond, A. J. (2017). Starbeast2 brings faster species
253 tree inference and accurate estimates of substitution rates. *Molecular Biology and*
254 *Evolution*, *34*(8), 2101–2114. doi: 10.1093/molbev/msx126
255 Patterson, M., Marschall, T., Pisanti, N., van Iersel, L., Stougie, L., Klau, G. W., &

256 Schönhuth, A. (2015). WhatsHap: Weighted Haplotype Assembly for Future-Generation
257 Sequencing Reads. *Journal of Computational Biology*, *22*(6), 498–509. doi:
258 10.1089/cmb.2014.0157
259 Pond, S. L. K., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: hypothesis testing using
260 phylogenies. *Bioinformatics*, *21*(5), 676–679. doi: 10.1093/bioinformatics/bti079
261 Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing
262 genomic features. *Bioinformatics*, *26*(6), 841–842. doi: 10.1093/bioinformatics/btq033
263 Renner, S. S., & Müller, N. A. (2021). Plant sex chromosomes defy evolutionary models of
264 expanding recombination suppression and genetic degeneration. *Nature Plants*, *7*(4),
265 392–402. doi: 10.1038/s41477-021-00884-3
266 Rice, W. R. (1996). Sexually antagonistic male adaptation triggered by experimental arrest of
267 female evolution. *Nature*, *381*(6579), 232–234. doi: 10.1038/381232a0
268 Slancarova, V., Zdanska, J., Janousek, B., Talianova, M., Zschach, C., Zluvova, J., …
269 Vyskot, B. (2013). Evolution of sex determination systems with heterogametic males and
270 females in silene. *Evolution*, *67*(12), 3669–3677. doi: 10.1111/evo.12223
271 Wang, D., Zhang, Y., Zhang, Z., Zhu, J., & Yu, J. (2010). KaKs_Calculator 2.0: a toolkit
272 incorporating gamma-series methods and sliding window strategies. *Genomics,*
273 *Proteomics & Bioinformatics / Beijing Genomics Institute*, *8*(1), 77–80. doi:
274 10.1016/S1672-0229(10)60008-3
275 Wertheim, J. O., Murrell, B., Smith, M. D., Kosakovsky Pond, S. L., & Scheffler, K. (2015).
276 RELAX: detecting relaxed selection in a phylogenetic framework. *Molecular Biology*
277 *and Evolution*, *32*(3), 820–832. doi: 10.1093/molbev/msu400
278 Wilson, E. B. (1927). Probable Inference, the Law of Succession, and Statistical Inference.
279 *Journal of the American Statistical Association*, *22*(158), 209–212. doi:
280 10.1080/01621459.1927.10502953

281 Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology*
282 *and Evolution*, *24*(8), 1586–1591. doi: 10.1093/molbev/msm088
283 Zemp, N., Widmer, A., & Charlesworth, D. (2018). Has adaptation occurred in males and
284 females since separate sexes evolved in the plant Silene latifolia? *Proceedings.*
285 *Biological Sciences / the Royal Society*, *285*(1883). doi: 10.1098/rspb.2017.2824
286 Zluvova, J., Lengerova, M., Markova, M., Hobza, R., Nicolas, M., Vyskot, B., … Janousek,
287 B. (2005). The inter-specific hybrid Silene latifolia x S. viscosa reveals early events of
288 sex chromosome evolution. *Evolution & Development*, *7*(4), 327–336. doi:
289 10.1111/j.1525-142X.2005.05038.x