1   **Transposable Element Evolution in the Allotetraploid *Capsella bursa-pastoris* and**

2   **the Perfect Storm Hypothesis**

3   J. Arvid Ågren[1,2], Hui-Run Huang[3], Stephen I. Wright[1]

4

5   [1] Department of Ecology and Evolutionary Biology, University of Toronto

6   [2] Present address: Department of Molecular Biology and Genetics, Cornell University

7   [3] Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South

8   China Botanical Garden, the Chinese Academy of Sciences

9

10   Key words: C-value; genome size; mating system; polyploidy

11   Running head: *Capsella* transposons and the perfect storm hypothesis

12   Word count: 2009

13   Abstract word count: 248

14   Number of figures: 3

15   Number of tables: 1

16

17

18

19

20

21

22

23

1   **Abstract**

2   *Premise of the study*

3   Shifts in ploidy level will affect the evolutionary dynamics of genomes in a myriad of

4   ways. Population genetic theory predicts that transposable element (TE) proliferation may

5   follow because the genome wide efficacy of selection should be reduced and the increase

6   in gene copies may mask the deleterious effects of TE insertions. However, to date the

7   evidence of TE proliferation following an increase in ploidy is mixed, with some studies

8   reporting results consistent with this scenario and others signs of genome downsizing.

9   *Methods*

10   We used high-coverage whole genome sequence data to evaluate the abundance, genomic

11   distribution, and population frequencies of TEs in the self-fertilizing recent allotetraploid

12   *Capsella bursa-pastoris*, a species with prior evidence for genome-wide reductions in

13   selection at the amino acid level since the transition to selfing. We then compared the *C.*

14   *bursa-pastoris* TE profile with that of its two parental species, outcrossing *C. grandiflora*

15   and self-fertilzing *C. orientalis*.

16   *Key results*

17   We found no evidence that *C. bursa-pastoris* has experienced a large proliferation of

18   TEs. Instead, the abundance, both overall and near genes, as well as the population

19   frequencies of TEs, are intermediate to that of its two parental species *C. grandiflora* and

20   *C. orientalis*.

21   *Conclusions*

22   The lack of shift in TE profile beyond additivity expectations in *C. bursa-pastoris* can be

23   because of variety of factors. In general, we argue that allopolyploid lineages that retain

1   high outcrossing should provide a 'perfect storm' for TE proliferation, while highly

2   selfing polyploids may generally experience TE loss.

1 Introduction

2 A central goal of population and comparative genomics research is to understand what

3 factors drive the evolution of genome size and structure (Gregory, 2005; Lynch, 2007;

4 Alföldi and Lindblad-Toh, 2013; Koenig and Weigel, 2015). Coupled with decades of

5 information on genome size from flow cytometry across diverse lineages, the growing

6 wealth of whole genome sequence data has revealed how extensive and rapidly genome

7 size and structure can evolve, even among close relatives (Ungerer et al., 2006; Hawkins

8 et al., 2009; Wright and Ågren, 2011; Tenaillon et al., 2011; Leitch and Leitch, 2013;

9 Ågren and Wright 2015; Ågren et al., 2015). Yet our ability to explain this variation

10 remains in its infancy.

11   Whole genome duplication via polyploidization has long been considered to be a

12 major contributor to genome evolution (Adams and Wendel, 2005; Soltis and Soltis,

13 2012; Hollister 2015; Soltis et al. 2015). Most obviously, polyploidization will cause a

14 direct increase in the total amount of DNA per cell. This initial doubling of genome size

15 has often been followed by a process of diploidization, leading to a pattern of DNA loss

16 over time (Leitch and Bennet, 2004; Lysak et al., 2009; Renny-Byfield et al., 2013; Vu et

17 al., 2015). Moreover, although the direct role of recent polyploidy on genome size

18 evolution can be investigated and controlled for, most plant species have experienced a

19 history of whole genome duplication in their evolutionary history (Vision et al. 2000;

20 Jaillon et al., 2007; Jiao et al., 2011; Vanneste et al., 2014). A lack of complete

21 information on this history can therefore make it difficult to fully investigate the

22 importance of whole genome duplication events on the evolution of genome size and

23 structure.

1   In addition to the direct effect of ploidy on genome size, transposable element

2   (TE) proliferation may follow whole-genome duplications due to the masking of

3   deleterious insertions and a reduction of the efficacy of selection across the genome

4   caused by genome redundancy (recently reviewed in Parisod and Senerchia, 2012; Tayalé

5   and Parisod, 2013). Furthermore, host-mediated silencing of TEs may be disrupted in

6   allopolyploids (where an increase in ploidy is due to interspecific hybridization; Madlung

7   et al., 2002; 2005; Kraitshtein et al., 2010; Yaakov et al., 2011). This combination of

8   relaxed selection and a breakdown of silencing mechanisms could potentially drive

9   dramatic evolution of genome structure following whole genome duplication, from gene-

10   dense euchromatic regions with very low TE content, to major accumulation of TEs in

11   genic regions. Such a mechanism may explain, for example, the dramatic transposable

12   element expansion in the maize genome following whole genome duplication (Schnable

13   et al., 2009; Baucom et al., 2009; Diez et al., 2014). On the other hand, genome

14   downsizing in polyploids may lead to a net loss of transposable elements during the

15   process of diploidization (Parisod et al., 2010). Indeed, to date studies showing both

16   increases in TE copy number, e.g. in *Nicotiana tabacum* (Petit et al., 2007; 2010; but see

17   Renny-Byfield et al., 2011) as well as TE loss, e.g. in *Orobanche gracilis* (Kraitschtein et

18   al., 2010) following allopolyploidization have been reported. Thus, overall, the factors

19   driving proliferation vs. loss of transposable elements in polyploids remain poorly

20   understood.

21   Here, we use high-coverage whole genome sequence data to evaluate the

22   abundance, genomic distribution, and population frequencies of TEs in the self-fertilizing

23   allotetraploid *Capsella bursa-pastoris*. *Capsella bursa-pastoris* is a recently derived

1 allotetraploid, with population genomic evidence for genome-wide reduction in the

2 strength of selection on point mutations due to both gene redundancy and its selfing

3 mating system (Douglas et al. 2015), making it an interesting model to examine the early

4 fate of transposable elements following allopolyploidization. We look for evidence of TE

5 proliferation beyond what would be expected from additivity of its two parental diploid

6 species. We discuss our results in light of the literature on the association between

7 polyploidization, mating system, and TE abundance across plant species and use these

8 observations to put forward the Perfect Storm Hypothesis for how mating system and

9 ploidy may interact to allow TE-driven genome expansion.

10

11 **Material and Methods**

12 *Study system*

13 The genus *Capsella* consists of four species with varying mating system, ploidy, and

14 geographical distribution (Hurka et al. 2012; Figure 1). Selfing is believed to have

15 evolved multiple times from an ancient progenitor of the diploid ($2n = 2x = 16$) *Capsella*

16 *grandiflora*, a self-incompatible species restricted to Albania and northwestern Greece.

17 Most recently, *Capsella rubella* diverged from *C. grandiflora* within the last 100,000

18 years (Foxe et al. 2009; Guo et al. 2009; Slotte et al. 2013; Brandvain et al. 2013).

19 *Capsella orientalis* is believed to have evolved selfing prior to *C. rubella*, also from a *C.*

20 *grandiflora*-like ancestor. *C. orientalis* and *C. grandiflora* have been inferred to have

21 diverged approximately 930,000 years ago, providing the potential for a longer time

22 period of mating system divergence (Douglas et al. 2015). Whereas *C. rubella* has

23 expanded to a larger Mediterranean distribution, *C. orientalis* is now found in an area

1   spanning Eastern Europe to Central Asia (Hurka et al. 2012). The origin of the world-

2   wide distributed *Capsella bursa-pastoris* long remained elusive, but was recently

3   determined to be an allotetraploid ($2n = 4x = 32$) following a hybridization event between

4   *C. grandiflora* and *C. orientalis* within the last 100,000-300,000 years (Douglas et al.

5   2015). Consistent with this hybrid origin, a principal component analysis of the types of

6   TEs found in *C. bursa-pastoris* puts it as an intermediate between *C. grandiflora* and *C.*

7   *orientalis* and of all shared insertions found in two of the three species, the majority are

8   between *C. bursa-pastoris* and either *C. grandiflora* or *C. orientalis*, with very few

9   shared between *C. grandiflora* and *C. orientalis* to the exclusion of *C. bursa-pastoris*

10 (Douglas et al. 2015). In this study, we expand the TE analysis of *C. bursa-pastoris*, to

11 test for an accumulation of TEs following allopolyploid origins

12

13 *Identification and quantification of transposable elements*

14 We combined the datasets generated by Ågren et al. (2014) and Douglas et al. (2015). In

15 brief, these studies applied the PopoolationTE pipeline of Kofler et al. (2012) on 108-bp

16 paired-end Illumina reads on 8 *C. bursa-pastoris*, as well as 8 *C. grandiflora* and 10 *C.*

17 *orientalis* individuals. Since the approach is designed for pooled population data, we

18 adjusted it to use population frequencies to infer insertions as homo- or heterozygous, as

19 per Ågren et al. (2014) and Douglas et al. (2015). We used this approach to determine the

20 abundance, genomic locations, and population frequencies of TEs in the three species.

21

22 **Results and Discussion**

7

1   We quantified the abundance of four major categories of TEs: DNA, Helitrons, long

2   terminal repeat (LTR) retrotransposons and non-LTR retrotransposons. The three species

3   differ in their mean number of TEs (Kruskal-Wallis chi-squared $=$ 21.342, df $=$ 2,

4   p $<$ 0.00001), but all species show similar relative abundance across elements, with

5   LTR elements making up the bulk of the insertions (Figure 2).

6   To test whether *C. bursa-pastoris* has experienced an accumulation or loss of TEs

7   following its origin, we calculated the expected diploid TE copy number from a *C.*

8   *orientalis* $\times$ *C. grandiflora* hybrid and compared this number to the observed *C. bursa-*

9   *pastoris* abundance. We randomly paired up *C. orientalis* and *C. grandiflora* individuals

10   and calculated the average TE copy number of such a cross. We performed 1,000

11   replicates of *in silico* crosses, sampling with replacement. We then compared the

12   expected copy number to the observed abundance in *C. bursa-pastoris* and found that *C.*

13   *bursa-pastoris* harbours slightly more insertions than what would be expected under strict

14   additivity (Figure 2; Wilcoxon rank sum test, W = 6038, p = 0.01297).

15   Since TE insertions near genes will likely disrupt gene function, population

16   genetic theory predicts that selection will rapidly remove such insertions (Dolging and

17   Charlesworth 2008). Following a whole-genome doubling event, a tetraploid like *C.*

18   *bursa-pastoris* will carry twice as many gene copies as its diploid progenitors and the

19   fitness cost of an insertion should therefore be less. As a consequence, tetraploids may be

20   expected to accumulate more TEs near genes than diploids. To test this prediction, we

21   used the gene annotation from the reference genome of *C. rubella* (Slotte et al. 2013) to

22   calculate the distance to the closest gene for all TE insertions, in all three species. Again,

23   just like the overall abundance, we are interested in whether *C. bursa-pastoris* has more

1   insertions near genes than what would be expected by additivity from a *C. orientalis* × *C.*

2   *grandiflora* cross. Using the approach outlined above, we calculated the expected TE

3   copy number within 1000 bp of the closest gene from such a hybrid and compared it to

4   the observed abundance in *C. bursa-pastoris*. We find that *C. bursa-pastoris* harbours

5   slightly more insertions near genes, compared to what would be expected under strict

6   additivity (Figure 3; Wilcoxon rank sum test, $W = 7144.5$, $p = 0.000126$)

7   Ågren et al. (2014) previously reported that *C. grandiflora* showed an excess of

8   rare insertions relative to *C. orientalis*, consistent with higher TE activity in the former.

9   Here, we used the presence/absence of all TE insertions, across all individuals in the three

10   species to categorize insertions as either singletons (present in only one individual) or

11   non-singletons (present in more than one individual). We find that *C. grandiflora* has the

12   highest number of singletons, potentially suggesting a higher TE activity than *C.*

13   *orientalis* and *C. bursa-pastoris*, which both show similar proportions of singletons

14   (Table 1). However, differences in demographic history between the species are likely

15   also contributing to the frequency spectrum.

16

17   **Table 1 Number of singleton and non-singleton transposon insertions in three**
18   ***Capsella* species**

	Singletons	Non-singletons
C. orientalis	524 (20%)	2095 (80%)
C. grandiflora	3060 (29%)	7493 (71%)
C. bursa-pastoris	3768 (22%)	13358 (78%)

19

9

1   Overall, we found no evidence that the species is experiencing a large

2   proliferation of TEs. Instead, the abundance, both overall and near genes, as well as the

3   population frequencies of TEs, are intermediate that of its two parental species *C.*

4   *grandiflora* and *C. orientalis*. This is consistent with our analyses of genome structure

5   variation (Douglas et al. 2015) indicating that this allopolyploid has not experienced

6   large-scale 'genome shock' since its origins.

7

8   *The perfect storm hypothesis*

9   Despite a number of models predicting TE proliferation following allopolyploidization,

10   our results provide little evidence for expansion in *C. bursa-pastoris*. This is in contrast

11   with our results from genome-wide SNP patterns (Douglas et al. 2015), which suggest a

12   large-scale reduction in the efficacy of selection on amino acid and conserved noncoding

13   mutations. Consistent with this, recent reviews have suggested that TE proliferation is not

14   consistently observed in allopolyploids (Parisod and Sinerchia 2012; Tayalé and Parisod,

15   2013). One important consideration when predicting the effects of polyploidization on

16   genome evolution may be its association with mating system. Highly outcrossing species

17   should experience high rates of transposable element activity (Wright and Schoen, 1999;

18   Morgan, 2001; Charlesworth and Wright, 2001). Allopolyploidization events that are

19   associated with a retention of high rates of outcrossing could therefore represent a

20   'perfect storm', whereby TE activity remains high while genome redundancy enables

21   rapid proliferation. On the other hand, polyploidy is often associated with elevated rates

22   of self-fertilization compared with diploid relatives, and both asexual reproduction and

23   high rates of selfing are common in polyploid lineages (reviewed in e.g. Mable, 2004;

10

1   Husband et al., 2008; Robertson et al., 2011; Ramsey and Ramsey, 2014). Highly selfing

2   lineages such as *C. bursa-pastoris* may experience a reduction in copy number of

3   transposable elements and a reduction in the spread of new, active transposable elements

4   through populations, because of a loss of transmission via outcrossing.

5   It is notable that some of the most well-documented ancient TE expansion events,

6   including maize (Schnable et al., 2009; Baucom et al., 2009; Diez et al., 2014) and the

7   *Brassica* genus (Zhang and Wessler, 2004), are associated with ancient

8   allopolyploidization events involving outcrossing lineages. Whether this is simply

9   circumstantial or causal will require in-depth comparative analyses of the joint and

10 unique effects of polyploidy and mating system on genome size and TE proliferation. As

11 we gain increasingly detailed insights into the time since last whole genome duplication

12 event in many lineages, investigating how this interacts with mating system to structure

13 genome evolution is becoming increasingly feasible.

14

1 **References**

2 Adams, K.L., and J.F. Wendel. 2005. Polyploidy and genome evolution in plants. *Current*

3 *Opinion in Plant Biology* 8: 135–141.

4

5 Ågren, J.A., W. Wang, D. Koenig, B. Neuffer, D. Weigel, and S.I. Wright. 2014. Mating

6 system shifts and transposable element evolution in the plant genus *Capsella*. *BMC*

7 *Genomics* 15: 602.

8

9 Ågren, J.A., and S.I. Wright. 2015. Selfish genetic elements and plant genome size

10 evolution. *Trends in Plant Science* 20: 195–196.

11

12 Ågren, J.A., S. Greiner, M.T.J. Johnson, and S.I. Wright. 2015. No evidence that sex and

13 transposable elements drive genome size variation in evening primroses. *Evolution* 69:

14 1053–1062

15

16 Alföldi, J., and K. Lindblad-Toh. 2013.Comparative genomics as a tool to understand

17 evolution and disease. *Genome Research* 23: 1063–1068.

18

19 Baucom, R.S., J.C. Estill, C. Chaparro, N. Upshaw, A. Jogi, J. Deragon, R.P. Westerman,

20 et al. Exceptional diversity, non-random distribution, and rapid evolution of retroelements

21 in the B73 Maize genome. *PLoS Genetics* 5: e1000732.

22

1   Brandvain, Y., T. Slotte, K.M. Hazzouri, S.I. Wright, and G. Coop. 2013. Genomic

2   identification of founding haplotypes reveals the history of the selfing species *Capsella*

3   *rubella*. *PLoS Genetics* 9: e1003754.

4

5   Charlesworth, D., and S.I. Wright. 2001. Breeding systems and genome evolution.

6   *Current Opinion in Genetics and Development* 11: 685–690.

7

8   Diez, C.M., E. Meca, M.I. Tenaillon, and B.S. Gaut. 2014. Three groups of transposable

9   elements with contrasting copy number dynamics and host responses in the maize (*Zea*

10   *mays* ssp. *mays*) genome. *PLoS Genetics* 10: e1004298.

11

12   Douglas, G., G. Gos, K. Steige, A. Salcedo, K. Holm, J.A. Ågren, K. Hazzouri, et al.

13   2015. Hybrid origins and the earliest stages of diploidization in the highly successful

14   recent polyploid *Capsella bursa-pastoris*. *Proceedings of the National Academy of*

15   *Sciences USA* 112: 2806–2811.

16

17   Foxe, J.P., T. Slotte, E.A. Stahl, B. Neuffer, H. Hurka, and S.I Wright. 2009. Recent

18   speciation associated with the evolution of selfing in *Capsella*. *Proceedings of the*

19   *National Academy of Sciences USA* 106: 5241–5245.

20

21   Gregory, T.R. 2005. *The Evolution of the Genome*. Elsevier, San Diego, California, USA.

22

1   Guo, Y.L., J.S. Bechsgaard, T. Slotte, B. Neuffer, M. Lascoux, D. Weigel, and M.H.

2   Schierup. 2009. Recent speciation of *Capsella rubella* from *Capsella grandiflora*,

3   associated with loss of self-incompatibility and an extreme bottleneck. *Proceedings of the*

4   *National Academy of Sciences USA* 106: 5246–5251.

5

6   Hawkins, J.S., S.R. Proulx, R.A. Rapp, and J.F. Wendel. 2009. Rapid DNA loss as a

7   counterbalance to genome expansion through retrotransposon proliferation in plants.

8   *Proceedings of the National Academy of Science USA* 106: 17811–17816.

9

10 Hollister, J.D. 2015. Polyploidy: adaptation to the genomic environment. *New Phytologist*

11 205: 1034–1039.

12

13 Hurka, H., N. Friesen, D.A: German, A. Franzke, and B. Neuffer. 2012. 'Missing link'

14 species *Capsella orientalis* and *Capsella thracica* elucidate evolution of model plant

15 genus *Capsella* (Brassicaceae). *Molecular Ecology* 21: 1223–1238.

16

17 Husband, B.C., Ozimec, B., Martin, S.L., and L. Pollock. 2008. Mating consequences of

18 polyploid evolution in flowering plants: current trends and insights from synthetic

19 polyploids. *International Journal of Plant Sciences* 169: 195-206.

20

21 Koenig, K., and D. Weigel. 2015. Beyond the thale: comparative genomics and genetics

22 of *Arabidopsis* relatives. *Nature Reviews Genetics* 16: 285–298.

23

1   Kofler, R., A.J. Betancourt, and C. Schlötterer. 2012. Sequencing of pooled DNA

2   samples (Pool-Seq) uncovers complex dynamics of transposable element insertions in

3   *Drosophila melanogaster*. *PLoS Genetics* 8: e1002487.

4

5   Jaillon, O., J.M. Aury, B. Noel, B., A. Policriti, C. Clepet, A. Casagrande, N Choisne, et

6   al. 2007. The grapevine genome sequence suggests ancestral hexaploidization in major

7   angiosperm phyla. *Nature* 449: 463–467.

8

9   Jiao, Y., N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr, P.E. Ralph,

10 L.P. Tomsho, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. *Nature*

11 473: 97–100.

12

13 Kraitshtein, Z., B. Yaakov, V. Khasdan, and K. Kashkush. 2010. Genetic and epigenetic

14 dynamics of a retrotransposon after allopolyploidization of wheat. *Genetics* 186: 801–

15 812.

16

17 Leitch, I.J., and M.D. Bennett. 2004. Genome downsizing in polyploid plants. *Biological*

18 *Journal of the Linnean Society* 82: 651–663.

19

20 Leitch,I.J., and A.R. Leitch. 2013. Genome size diversity and evolution in land plants In

21 I.J. Leitch, J. Greilhuber, J. Doležel, J.F. Wendel (Eds.). *Plant Genome Diversity*, 307–

22 322. Springer,Vienna, Austria.

23

1   Lynch, M. 2007. *The Origins of Genome Architecture*. Massachusetts, USA.

2

3   Lysak, M.A., M.A. Koch, J.M. Beaulieu, A. Meister, and I.J. Leitch. 2009. The dynamic

4   ups and downs of genome size evolution in Brassicaceae. *Molecular Biology and*

5   *Evolution* 26: 85-98.

6

7   Mable, B.K. 2004. Polyploidy and self compatibility: is there an association? 162: 803–

8   811.

9

10   Madlung, A., R.W. Masuelli, B. Watson, S.H. Reynolds, J. Davison, and L. Comai. 2002.

11   Remodeling of DNA methylation and phenotypic and transcriptional changes in synthetic

12   *Arabidopsis* allotetraploids. *Plant Physiology* 129: 733–746.

13

14   Madlung, A., A.P. Tyagi, B. Watson, H. Jiang, T. Kagochi, R.W. Doerge, R.

15   Martienssen, et al. 2005. Genomic changes in synthetic *Arabidopsis* polyploids. *The*

16   *Plant Journal* 41: 221–230.

17

18   Morgan, M.T. 2001. Transposable element number in mixed mating populations.

19   *Genetics Research* 77: 261–275.

20

21   Parisod C., K. Alix, J. Just, M. Petit, V. Sarilar, C. Mhiri, M. Ainouche, et al. 2010.

22   Impact of transposable elements on the organization and function of allopolyploid

23   genomes. *New Phytologist* 186: 37–45.

1

2   Parisod C., and N. Senerchia. 2013. Responses of transposable elements to polyploidy. In

3   M.A. Grandbastien and J. Casacuberta (Eds.). *Plant Transposable Elements*, 147–168.

4   Springer, Heidelberg, Germany.

5

6   Petit, M., C. Guidat, J. Daniel, E. Denis, E. Montoriol, Q.T Bui, K.Y. Lim, et al. 2010.

7   Mobilization of retrotransposons in synthetic allotetraploid tobacco. *New Phytologist*

8   186: 135–147.

9

10   Petit, M., K.Y. Lim, E. Julio, C. Poncet, F.D. de Borne, A. Kovarik, A.R. Leitch, et al.

11   2007. Differential impact of retrotransposon populations on the genome of allotetraploid

12   tobacco (*Nicotiana tabacum*). *Molecular Genetics and Genomics* 278: 1–15.

13

14   Ramsey, J., and T.S Ramsey. 2014. Ecological studies of polyploidy in the 100 years

15   following its discovery. *Philosophical Transactions of the Royal Society of London B:*

16   *Biological Sciences* 369: 20130352.

17

18   Renny-Byfield S., M. Chester, A. Kovařík, S.C. Le Comber, and MA Grandbastien, M.

19   Deloger, R.A. Nichols et al. 2011. Next generation sequencing reveals genome

20   downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination

21   of paternally derived repetitive DNAs. *Molecular Biology and Evolution* 28: 2843–2854.

22

1   Renny☐Byfield, A., Kovarik, L.J. Kelly, J. Macas, P. Novak, M.W. Chase, R.A. Nichols

2   et al. 2013. Diploidization and genome size change in allopolyploids is associated with

3   differential dynamics of low☐and high☐copy sequences. *The Plant Journal* 74: 829–839.

4

5   Robertson, K. A., E.E. Goldberg, and B. Igic. 2011. Comparative evidence for the

6   correlated evolution of polyploidy and self-compatibility in Solanaceae. *Evolution* 65:

7   139–155.

8

9   Schnable, P.S., D. Ware, R.S. Fulton, J.C. Stein, F. Wei, S. Pasternak, C. Liang, et al.

10   2009. The B73 maize genome: complexity, diversity, and dynamics. *Science* 326: 1112–

11   1115.

12

13   Slotte, T., K.M. Hazzouri, J.A. Ågren, D. Koenig, F. Maumus, Y.L. Guo, K. Steige, et al.

14   2013. The *Capsella rubella* genome and the genomic consequences of rapid mating

15   system evolution. *Nature Genetics* 45: 831–835.

16

17   Soltis, P.S., and D.E. Soltis. 2012. *Polyploidy and Genome Evolution*. Springer, New

18   York, New York.

19

20   Soltis, P.S., D. Blaine Marchant, Y. Van de Peer, and D.S. Soltis. 2015. Polyploidy and

21   genome evolution in plants. *Current Opinion in Genetics and Development* 35: 199–125.

22

1   Tenaillon, M.I., M.B. Hufford, B.S Gaut, and J. Ross-Ibarra. 2011. Genome size and

2   transposable element content as determined by high-thoughput sequencing in maize and

3   *Zea luxurians*. *Genome Biology and Evolution* 3: 219–229.

4

5   Ungerer, M.C., S. Strakosh, and Y. Zhen. 2006. Genome expansion in three hybrid

6   sunflower species is associated with retrotransposon proliferation. *Current Biology* 16:

7   r872–r873.

8

9   Vanneste, K., G. Baele, S. Maere, and Y. Van de Peer. 2014. Analysis of 41 plant

10   genomes supports a wave of successful genome duplications in association with the

11   Cretaceous–Paleogene boundary. *Genome Research* 24: 1334–1347.

12

13   Vision, T.J., D.G. Brown, and S.D. Tanksley. 2000. The origins of genomic duplications

14   in *Arabidopsis*. *Science* 290: 2114–2117.

15

16   Vu, G.T.H., T. Schmutzer, F. Bull, H.X. Cao, J. Fuchs, T.D. Tran, G. Jovtchev, et al.

17   2015. Comparative genome analysis revealed divergent genome size evolution in a

18   carnivorous plant genus. *Plant Genome* 8: 3.

19

20   Wright, S.I, and D.J. Schoen DJ. 1999. Transposon dynamics and the breeding system.

21   *Genetica* 107: 139–148.

22

1   Wright, S.I., and J.A. Ågren. 2011. Sizing up *Arabidopsis* genome evolution. *Heredity*

2   107: 509–510.

3

4   Zhang, X., and S.R. Wessler. 2004. Genome-wide comparative analysis of the

5   transposable elements in the related species *Arabidopsis thaliana* and *Brassica oleracea*.

6   *Proceedings of the National Academy of Sciences USA* 101: 5589–5594.

7

8

1   Figure 1. Phylogenetic relationships within the *Capsella* genus. The evolution and natural

2   history of the genus are thoroughly discussed in Hurka et al. (2012), Slotte et al. (2013),

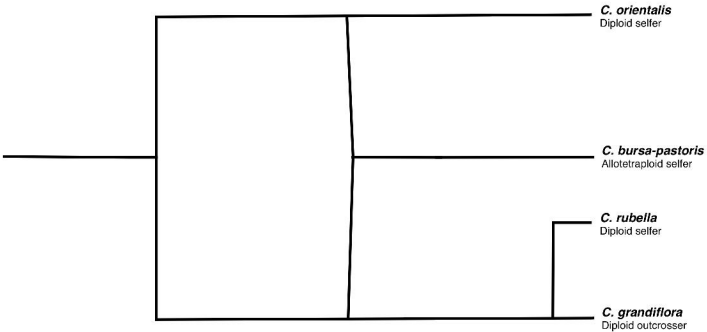3   and Douglas et al. (2015).

4

5   Figure 2. Boxplots of transposable elements (TEs) abundance in the three *Capsella*

6   species. The expected *C. bursa-pastoris* value was generated by performing 1,000

7   replicates of in silico crosses between *C. orientalis* × *C. grandiflora*, sampling with
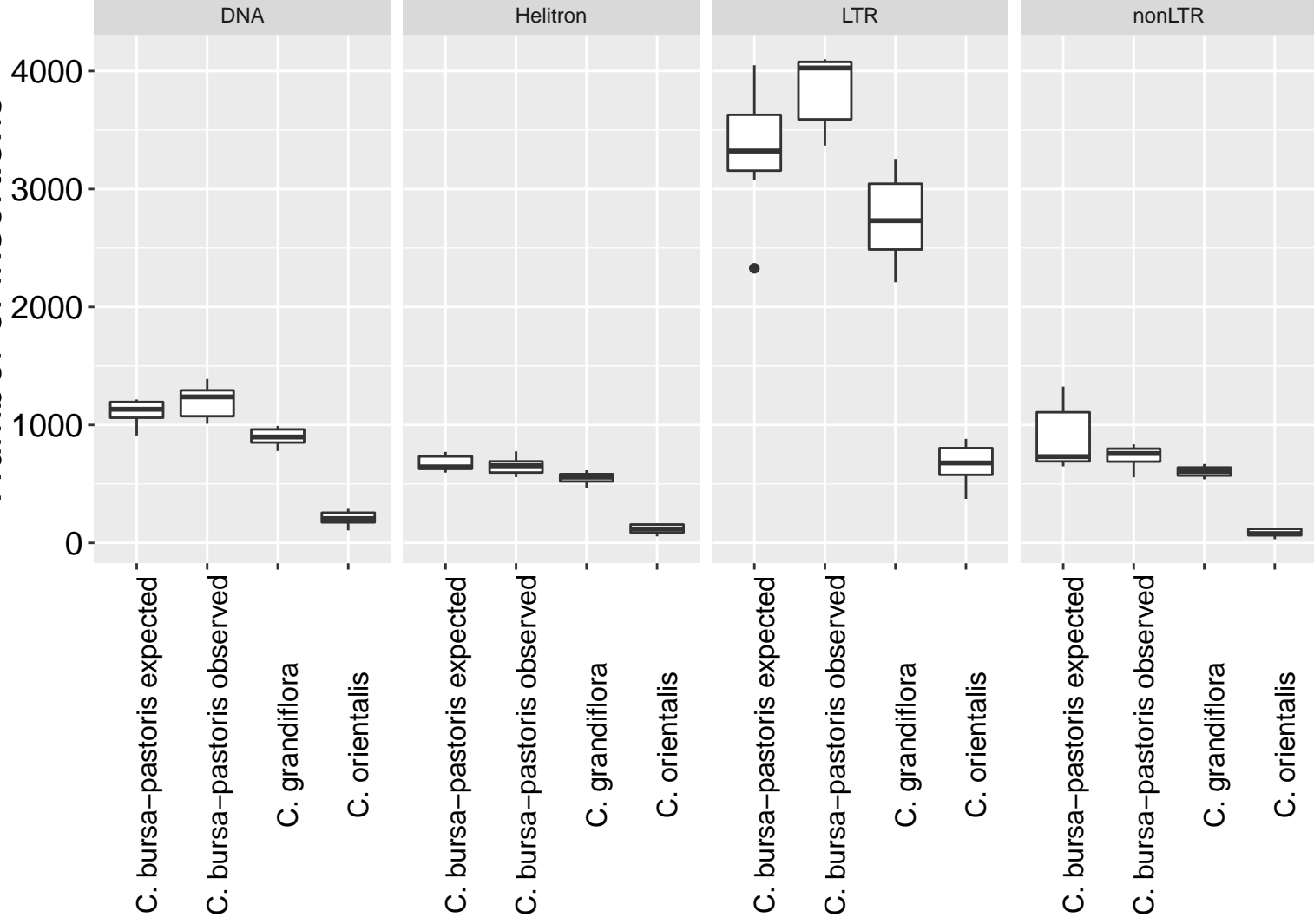
8   replacement.

9

10   Figure 3. Average abundance of transposable elements (TEs) in 100 bp bins near their

11   closest gene in the three *Capsella* species. Error bars are $\pm$ 1 standard error. The

12   expected *C. bursa-pastoris* value was generated by performing 1,000 replicates of in

13   silico crosses between *C. orientalis* × *C. grandiflora*, sampling with replacement.

14

15

C. orientalis
Diploid selfer

C. bursa-pastoris
Allotetraploid selfer

C. rubella
Diploid selfer

C. grandiflora
Diploid outcrosser