
1

The evolution of polyploidy and the extreme cost of pollen-swamping

Dissertation Thesis

2 2417024

Abstract

3 Amazing summary of project here...

Keywords

4 polyploidy; evolution; angiosperm; *mimulus*; *guttatus*; pollen-swamping; mate-choice;

5 1 Introduction

6 Would some images of genome duplication be helpful?

7 Whole-genome duplication (polyploidisation) is an extreme mutation
8 which occurs due to errors during cell division; such as non-disjunction
9 where sister chromatids are not separated during anaphase. Failed
10 disjunction leads to all the genetic material ending up in one of the two
11 daughter cells; essentially doubling the genome within that cell. When
12 this occurs during gametogenesis, it can lead to whole organisms with

more than two complete sets of chromosomes (polyploids). According to ?, the estimated rate of polyploidisation is one order of magnitude higher than that of standard mutation (10^{-5}). In humans, it coincides with around 5% of miscarriages (Creasy et al. 1976). Moreover, while there have been some live birth cases of triploid human infants, they usually die within hours. In irregular cases, they survive for several months, but none have survived more than six years (Hashimoto et al. 2018).

Newer data for miscarriage rate?

No polyploid mammal species has ever been found (Svartman et al. 2005), and the same is also usually true for birds (?). However, while in these taxa, genome-doubling is not well tolerated, for amphibians, fish, fungi, reptiles, and plants, the story is quite different. Amongst these, polyploidy is exceptionally well tolerated in plants. Jiao et al. (2011) found that genome-doubling is a ubiquitous feature in the lineages of almost all flowering plants (angiosperms), and Wood et al. (2009) also found that it has occurred even more recently (post genera formation) for 35% of all vascular plants. Plant surveys of natural populations over the last century further support this, reporting mean polyploid frequencies of 25% (?).

The Role of Genome-Doubling in Plant Diversification

Genome-doubling is thought to be fundamental in the diversification of plant species; having been found to coincide with 15% of angiosperm speciation events, and 31% in ferns (Wood et al. 2009). Furthermore, Jiao et al. (2011) showed that whole-genome duplication could even be linked with major biological innovations that have lead to diversification of the plant phyla. Their phylogenetic analysis brought to light two ancient groups of duplication events: the first grouping of duplications (around 319 millennia ago) influenced the diversification of regulatory genes that

were integral to seed development and the more recent grouping (around 192 millennia ago) affected genes that enabled flower development. The authors appropriately conclude that these events must have played a significant role in the rise of seed plants, as well as in the later diversification of angiosperms. Seeds and flowers were both crucial innovations that allowed plants to transition from an aquatic life-cycle to a terrestrial one by removing their dependence on water for reproduction — an event, broadly considered as one of the most significant in the history of our planet.

The full story, however, is not quite that clear-cut. According to [Mayrose et al. \(2011\)](#), genome-doubling significantly increases speciation rates of diploids, but new polyploid lines do not further speciate by that same mechanism. So, polyploid speciation rates are far smaller in comparison. Furthermore, they found polyploid extinction rates to be far higher than those of diploids because polyploids are clumped around the tips of phylogenies. This Likelihood-based analysis of vascular plants provided clear, and quantitative, support for the traditionally popular view that polyploidy most often leads to evolutionary dead ends.

First proposed by? Stebbins 71?

So, what is it then that allows this extreme mutation to succeed and become fixated? What scenarios and mechanisms combine to allow rare polyploid lines to succeed? Are the majority of the polyploids we see in natural populations doomed to extinction? Furthermore, can we predict which conditions will lead to their survival?

Despite having been studied for over a century, the factors that drive the success of polyploid establishment in the face of high extinction rates are still unclear. [Arrigo and Barker \(2012\)](#) conclude that polyploids tend to become extinct at the establishment phase due to reproductive

70 disadvantages such as triploid sterility, or limited mate-choice; the latter
71 occurring via diploid pollen-swamping, or delayed flowering. On the
72 other hand, polyploids are frequently linked with distinct traits such as
73 “gigas effects”, which describes their increased size and plant organs.
74 Additionally, benefits of polyploidy can include enhanced capabilities for
75 buffering of deleterious mutation (due to increased heterozygosity),
76 hybrid vigour (heterosis), and a reversal of selfing inhibition ([Woodhouse et al. 2009](#); [Ramsey and Ramsey 2014](#)). These traits are thought to
77 overcome the reproductive disadvantages of polyploidy and, in those rare
78 cases, make this mutation key to the invasive and adaptive potential of
79 plants.
80

81 *The cost of being polyploid (limited mate-choice)*

82 *Diploid Pollen-Swamping*

83 *Delayed flowering*

84 Delayed flowering (cost) is associated with gigas-effects of
85 increased size (benefit). TRADE-OFF.

86 *Mimulus study system*

87 Polyploidisation events have been well documented in *Mimulus guttatus*
88 (?).

89 ? details 3-stage *M. guttatus* life-cycle and transition matrices.

90 *Aims*

91 Here we aimed to quantify the cost of pollen-swamping by testing how
92 high the rate of whole-genome duplication would need to be in order to
93 overcome the strength of negative selection. We asked:

- 94 • Will increasing the rate of whole-genome duplication lead to the
95 fixation of polyploidy?

- 96 ● And, if so, at what threshold point does this occur?

97 In addition, we aimed to discover how varying the starting population
98 size affected these results:

- 99 ● Does an established population react in the same way to increased
100 duplication rates as a pioneering population entering an empty
101 landscape?

102 **2 Methods**

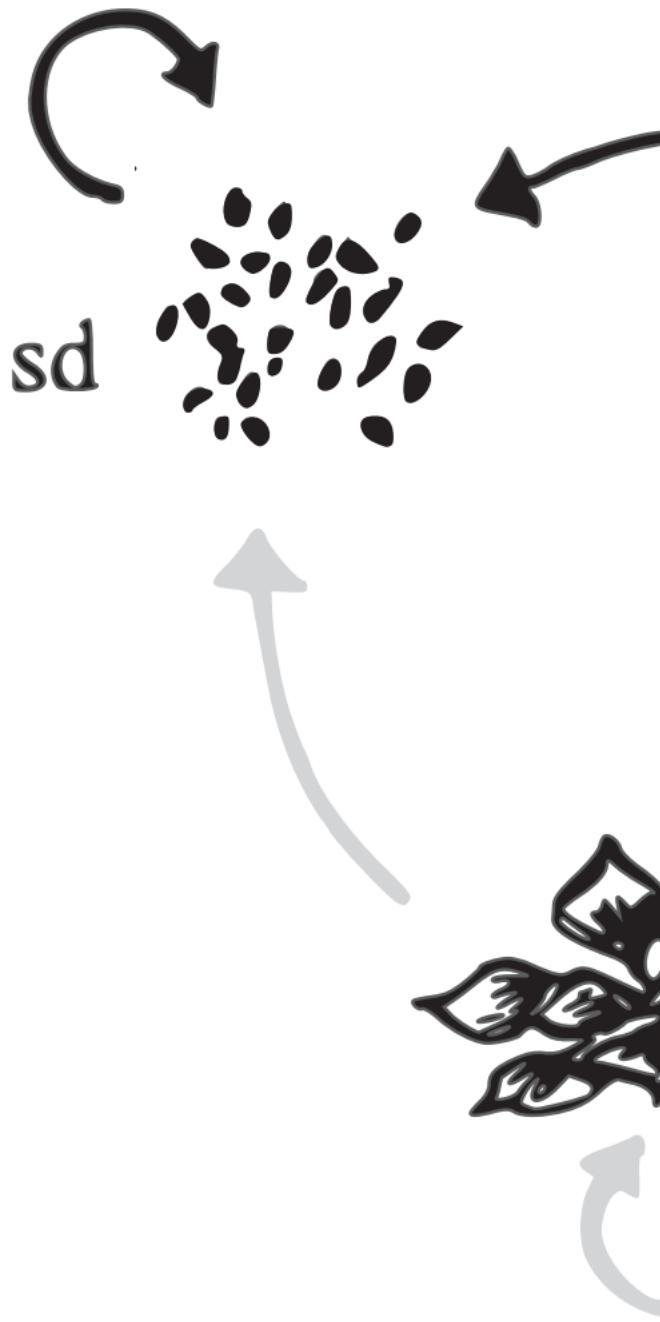
103 An individual-based modelling approach was used to examine these
104 aims. The model was created, from scratch, in R (?) to simulate the life
105 cycle of the yellow monkey flower (*Mimulus guttatus*) as laid out by ?
106 and detailed in figure ???. The *M. guttatus* complex was a suitable choice
107 due to whole-genome duplication being well documented throughout (?),
108 and data being readily available.

knitr::include_graphics ("_images/m-guttatus-lifecycle-pe

109 *The model*

110 The model was created as an open source R package *{sploidy}*, and is
111 available along with the simulation scripts used for this work and the raw
112 data, on Github (?).

113 **Assumptions** To test our hypotheses the model needed to make
114 particular assumptions about the reproductive interactions between
115 individuals, as well as how particular parameters (like whole genome
116 duplication rate) were applied mechanistically. To see if the rate of
117 duplication could overcome the effects of limited mate-choice, mate
118 choice had to first be limited in some manner. This was achieved using a
119 simple matching rule that defined whether attempts to fertilise were



⁶
Fig. 2 Life cycle graph for *Mimulus* g

120 successful, based on parental ploidy level. Simulating, therefore, the cost
121 of pollen-swamping.

122 The model assumptions were as follows:

- 123 • All individuals began as diploid.
- 124 • Polyploids arose at a set rate of whole-genome duplication during
125 all sexually reproductive transitions.
- 126 • Diploids and polyploids could both out-cross, but never with each
127 other. In addition, polyploids could not outcross with other
128 polyploids of differing ploidy level, ie:- a tetraploid was not
129 compatible with an octoploid.
- 130 • All individuals could self-fertilise.
- 131 • Offspring produced had the same ploidy level as their parents; two
132 diploids produced a diploid, and so on.
- 133 • When whole-genome duplication occurred, the ploidy level of an
134 individual was doubled; in this case two diploid parents would
135 produce a tetraploid.
- 136 • All plants were hermaphrodite; mothers and fathers were chosen via
137 random sampling from the same pool of individuals.
- 138 • Pollen range spanned the entire landscape.
- 139 • Seed dispersal spanned the entire landscape.

140 The *rbinom* function in base R was used throughout the model—where
141 transitions reduced cohort size—to choose individuals that transitioned
142 successfully. This allowed for a natural amount of stochasticity around
143 the transition probabilities, similar to that which we would expect in
144 nature. Where transition values increased cohort size, the function
145 *sample_n* from the package *{Dplyr}* (?) was used, instead, with
146 replacement.

Table 1. Components of Low-elevation perennial **Mimulus guttatus** life-cycle transition. Seed survival (D) is a general value taken from ?. All other components taken from ? who recorded values specific to that **M. guttatus** population in the year 2013 (N = 77). G = germination rate, O = ovule number per flower, F = flower production, S = winter survival, R = rosette production, A = proportional recruitment success of ovules relative to rosettes.

D	G	O	F	S	R	A
0.534	0.652	494	5.34	0	0	0.00067

147 *Simulation strategy*

148 Rounds of simulations were run using the model function *sploidy* in
 149 order to home-in on the parameter ranges to include in the experiment
 150 design. These initial simulations were run for 1000 generations on a
 151 100x100 landscape grid (to allow the default carrying capacity to
 152 maintain the adult population size at around one hundred thousand). The
 153 starting population contained sixty diploid individuals (twenty of each
 154 life stage), to simulate a pioneering species entering an empty space and
 155 the **ploidy_rate** parameter was varied between zero and 0.5. Seed
 156 dormancy was disabled by setting **seed_longevity** to zero. As was clonal
 157 reproduction, via the transition matrix provided to the **trans** parameter.
 158 Transition parameters—the transition matrix (**trans**), germination rate (**G**),
 159 and seed survival rate (**D**)—were taken from known demographic data (??).

160 *Transition data* Data for an *M. guttatus* population that had the most
 161 stable properties were used; minimal growth that could be kept in check
 162 with carrying capacity and low elasticity. ? described the formula for
 163 computing transition data as laid out in figure ?? and provided the values
 164 for a low-elevation perennial *M. guttatus* population as shown in table 1.
 165 These data produced the transition matrix (figure 2) provided to the
 166 model, along with the values of G (0.652) and D (0.534) shown in figure
 167 1.

```
knitr::include_graphics ("images/m-guttatus-transformula")
```

$$\begin{matrix}
 & \text{Seed}_t & \text{Seedling}_t & \text{Rosette}_t \\
 \text{Seed}_{t+1} & D(1 - G) & \text{FOA}(1 - G) & \text{FOA}(1 - G) \\
 \text{Seedling}_{t+1} & DG & \text{FOAG} & \text{FOAG} \\
 \text{Rosette}_{t+1} & 0 & SR & SR
 \end{matrix}$$

Figure 2. How the transition matrix (table 2) was calculated from the data provided in table 1, according to this formula taken from ?.

Table 2. Transition matrix used for the model, based on the formula shown in ?? and the data from figure 1.

	Seed (t)	Seedling (t)	Rosette (t)
Seed (t+1)	0.186	0.615	0.615
Seedling (t+1)	0.348	1.152	1.152
Rosette (t+1)	0.000	0.000	0.000

168

169 *Varying the rate of whole-genome duplication* In order to quantify how
 170 the rate of whole-genome duplication affected fixation of polyploid
 171 genotypes, a randomly sampled range of polyploidisation rates were
 172 chosen from an exponential distribution of 100000 random uniform
 173 numbers ranging between 0.001 and 0.010. Using an exponential
 174 distribution gave higher resolution around the lower parameter rates
 175 which we expected to be most interesting.

176 Diagnostic plotting of these data revealed that increased rates needed to
 177 be examined. Further simulations were then run using randomly sampled
 178 rates from 100000 random uniform numbers between 0.010 and 0.500 in
 179 an attempt to determine how high the mutation rate needed to be in order
 180 for polyploids to break free from the inherent selection against them due
 181 to mate-choice limitation.

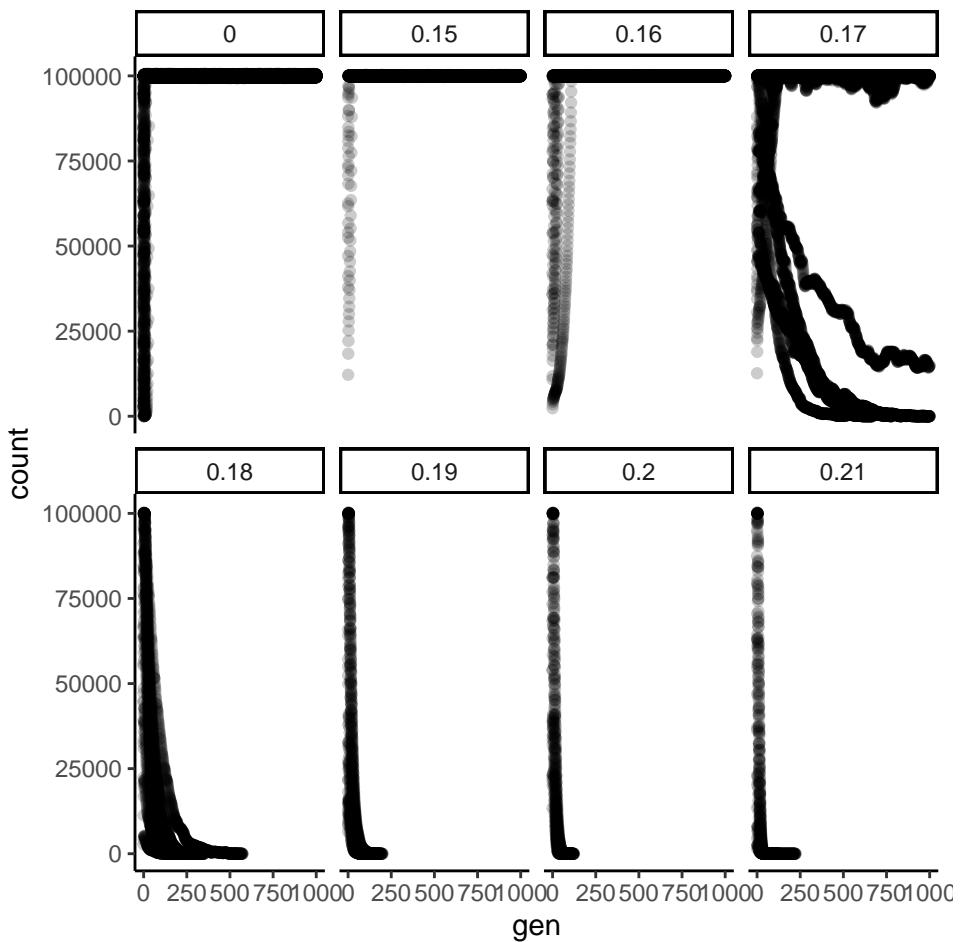
182 *Germination rate benefit* Further simulations which inferred a
 183 germination benefit to polyploids were run using the same randomly

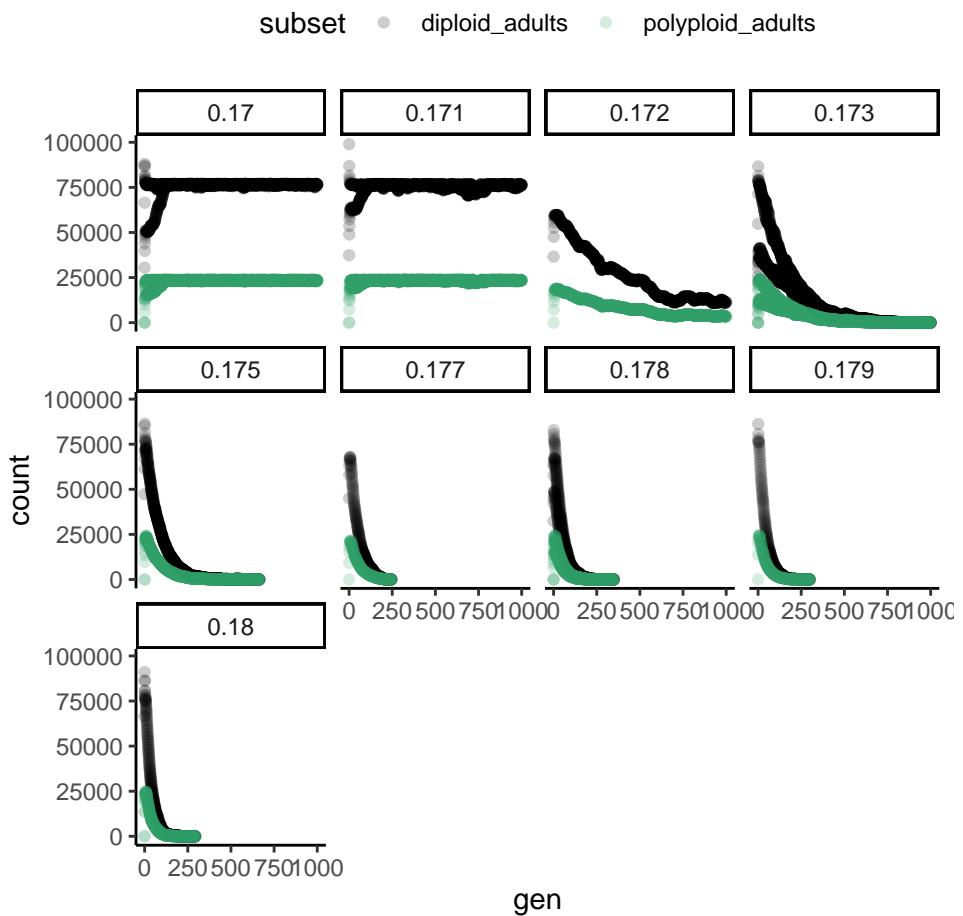
184 sampled rates of polyploidisation as above. A randomly sampled range of
185 germination rate modifiers were chosen from 100000 random uniform
186 numbers between 1 and 1.5, where the modifier was used to multiply the
187 germination rate ($G = 0.652$). The max value of 1.5 would cause
188 polyploid seeds to germinate at very close to complete success ($G =$
189 0.978), simulating the observation that polyploid seeds can have
190 increased germination success due to germination over a wider
191 temperature range (?).

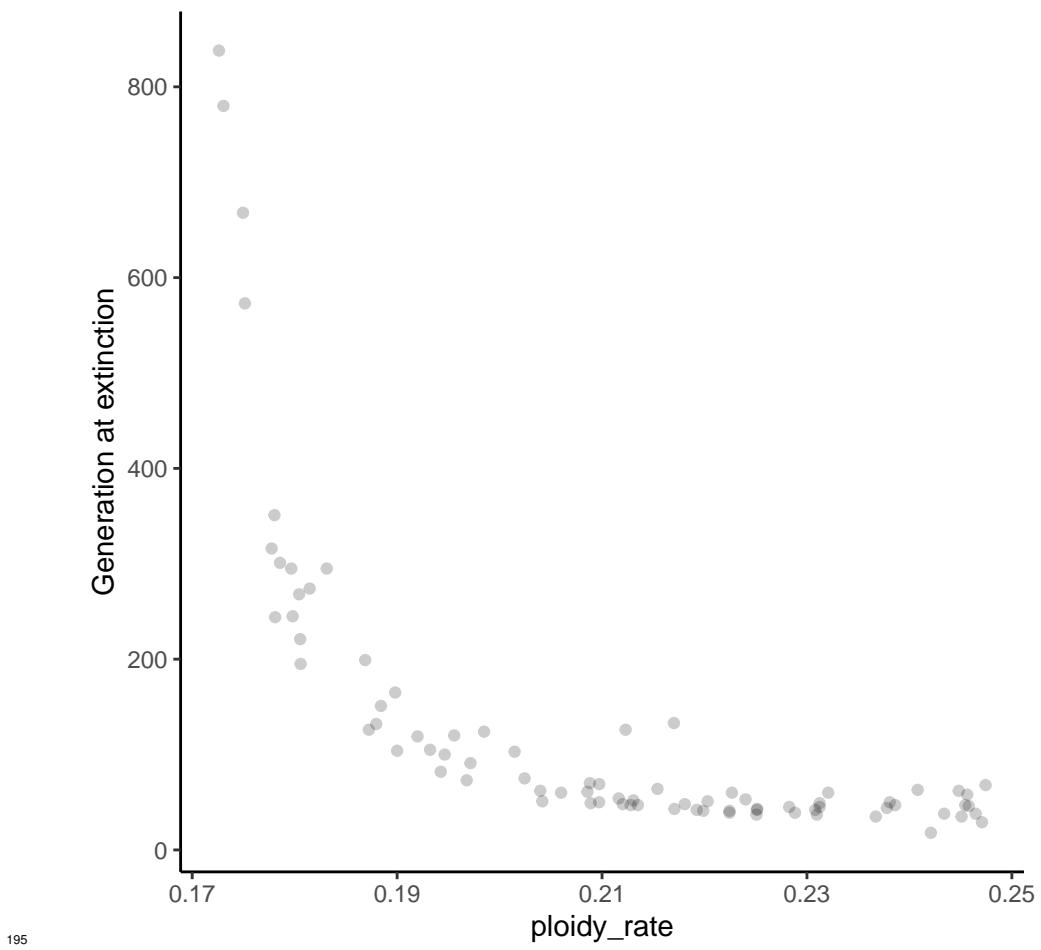
192 3 Results

```
#library(sploidy)
#count_combine("data/results", "thesis/_data/results")
counts <- readRDS("_data/results.rds") %>%
  dplyr::select(-diploid_seeds, -polyploid_seeds, -rosette)
  tidyverse::gather("subset", "count", seeds:total)

ploidylvl_colours = c("#000000", "#30a068")
ploidy_gradient = c("#10442a", "#3bbf7d")
```







196 **4 Discussion**

197 Does limited mate-choice set the scene for the evolution of beneficial
198 traits like selfing?

199 **5 Conclusion**

200 Does limited mate-choice set the scene for the evolution of beneficial
201 traits like selfing?

202 References

- 203 Arrigo N., Barker M.S. (2012) Rarely successful polyploids and their legacy in plant
204 genomes. *Current Opinion in Plant Biology*, 15 (2), pp. 140–146.
- 205 Creasy M.R., Crolla J.A., Alberman E.D. (1976) A cytogenetic study of human
206 spontaneous abortions using banding techniques. *Human Genetics*, 31 (2), pp.
207 177–196.
- 208 Hashimoto S., Igarashi L., Kobayashi M., Tsutsumi H. (2018) Long-term extrauterine
209 survival in a triploid infant: A review of the clinical features of live-born infants with
210 triploidy. *Hong Kong Journal of Paediatrics*, 23 (4), pp. 298–301.
- 211 Jiao Y., Wickett N.J., Ayyampalayam S., Chanderbali A.S., Landherr L., Ralph P.E.,
212 Tomsho L.P., Hu Y., Liang H., Soltis P.S., Soltis D.E., Clifton S.W., Schlarbaum S.E.,
213 Schuster S.C., Ma H., Leebens-Mack J., Depamphilis C.W. (2011) Ancestral
214 polyploidy in seed plants and angiosperms. *Nature*, 473 (7345), pp. 97–100.
- 215 Mayrose I., Zhan S.H., Rothfels C.J., Magnuson-Ford K., Barker M.S., Rieseberg L.H.,
216 Otto S.P. (2011) Recently Formed Polyploid Plants Diversify at Lower Rates.
217 *Science*, 333 (6047), p. 1257.
- 218 Ramsey J., Ramsey T.S. (2014) Ecological studies of polyploidy in the 100 years
219 following its discovery. *Phil Trans R Soc B*, 369.
- 220 Svartman M., Stone G., Stanyon R. (2005) Molecular cytogenetics discards polyploidy
221 in mammals. *Genomics*, 85, pp. 425–430.
- 222 Wood T.E., Takebayashi N., Barker M.S., Mayrose I., Greenspoon P.B., Rieseberg L.H.
223 (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the
224 National Academy of Sciences of the United States of America*, 106 (33), pp.
225 13875–9.
- 226 Woodhouse M., Burkart-Waco D., Comai L. (2009) Polyploidy. *Nature Education*,
227 2 (1), p. 1.