Evolution of Polyploidy in Flowering Plants

Dissertation Thesis

2417024

Keywords

Polyploidy; Evolution; Angiosperm; Mimulus;

1 Introduction

Would some images of genome duplication be helpful?

Whole-genome duplication (polyploidisation) is an extreme mutation which occurs due to errors during cell division; such as non-disjunction where sister chromatids are not separated during anaphase. Failed disjunction leads to all the genetic material ending up in one of the two daughter cells; essentially doubling the genome within that cell. When 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 disadvantages such as triploid sterility, or limited mate-choice; the latter occurring via diploid pollen-swamping, or delayed flowering. On the other hand, polyploids are frequently linked with distinct traits such as "gigas effects", which describes their increased size and plant organs. Additionally, benefits of polyploidy can include enhanced capabilities for

buffering of deleterious mutation (due to increased heterozygosity), hybrid vigour (heterosis), and — somewhat contentiously — a reversal of selfing inhibition (Woodhouse et al. 2009; Ramsey and Ramsey 2014). These traits are thought to overcome the reproductive disadvantages of polyploidy and, in those rare cases, make this mutation key to the invasive and adaptive potential of plants. Let us examine these mechanisms in more detail:

Benefits

Gigas-effects

Genetic buffering

Hybrid Vigour

Costs

Are they realistic? How do they work? Why/when are they important? Consider the conditions.

Diploid Pollen-Swamping

Delayed Flowering

Triploid Sterility

So what are the core mechanisms?

Or the most suspicious?

How do these mechanisms link?

Does limited mate-choice set the scene for the evolution of selfing vs out-crossing?

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

Mimulus study system

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

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

Aims

Outcrossing VS cloning. Everyone can do both, but polyploids will be forced only to clone while mate options are reduced. Simplified polyploidy - on/off Cloning - if adjacent square empty, chance of cloning into it. Outcrossing set to maximum range (this exagerrates any pattern).

1. look at distribution of polyploids without cloning
2. look at distribution with cloning.
3. compare frequencies to natural populations.
4. possibly then disable/reduce cloning in diploids if sensible form reading.

2 Methods

In order to examine how changes in the rate of whole-genome duplication might affect the evolution of polyploidy in angiosperm systems, I created an individual-based model (IBM) which simulates the life cycle of the yellow monkey flower (*Mimulus guttatus*) as laid out by **?** and detailed in figure **??**. *Mimulus* are popular study organisms (**?**) so data was readliy available and whole-genome duplication has been well documented in the *M. guttatus* complex (**?**).

The model was created as an R package, and is available here along with simulation scripts and data output: https://github.com/rosemckeon/honours-project.

Model assumptions

The assumptions made by the model were as follows:

- All individuals began as diploid.
- Polyploids arose at a set rate of whole-genome duplication during all reproductive transitions.
- Diploids and polyploids could both out-cross, but never with each other. In addition, polyploids could not outcross with other polyploids of differing ploidy level, ie:- a tetraploid was not compatible with an octoploid.
- Offspring produced had the same ploidy level as their parents; two diploids produced a diploid, and so on.
- When whole-genome duplication occured, the ploidy level of an individual was doubled; in this case two diploid parents would produce a tetraploid.
- All plants were hermaphrodite; mothers and fathers were chosen via random sampling from the same pool of individuals.
- Pollen range spanned the entire landscape.
- Seed dispersal spanned the entire landscape.
- Asexual reproduction was disabled.
- There was no seedbank.
- Carrying capacity curbed exponential growth of the adult population at around 100000 (with some stochasticity).

The model was parameratised with known demographic data taken from **?** and **?**. For the purposes of this simulation, I chose data for an *M. guttatus* population that had the most stable properties; minimal growth that could be kept in check with carrying capacity and low elasticity. The chosen transition matrix also had no asexual reproduction which was useful so as not to over-complicate the initial simulations. **?** described the formula for computing transition data as laid out in figure **??**. The values

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

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

I modelled were taken from the Low-elevation perennial population in 2013, as shown in table 1, which produced the transition matrix shown in figure 2.

Transitions from seed to seedling were modelled using seed survival (D) and germination rates (G), rather than the calculated transtion values in figure 2. This enabled the model to ensure that ungerminated seeds had the opportunity to persist in the environment if a seedbank were to be turned on, without resampling seeds by some probability, and inevitably sampling some individuals that did infact transition.

The *rbinom* function in base R was used throughout the model–where transitions reduced cohort size–to choose individuals that transitioned successfully. This allowed for a natural amount of stochasticity around the transition probabilities, similar to that which we would expect in nature. Where transition values increased cohort size, the function *sample_n* from the package $\{Dplyr\}$ was used, instead, with replacement.

Simulations

Simulations were run for 1000 generations on a 100x100 landscape grid, with a starting population that contained 20 of each diploid life stage (seeds, seedlings and rosettes).

Control Simulations with the rate of whole-genome duplication set to zero confirmed the expected pattern of population growth and stability for a scenario where no polyploids can evolve.

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

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

Germination rate benefit Further simulations which inferred a germination benefit to polyploids were run using the same randomly sampled rates of polyploidisation as above. A randomly sampled range of germination rate modifiers were chosen from 100000 random uniform numbers between 1 and 1.5, where the modifier was used to multiply the germination rate (G = 0.652). The max value of 1.5 would cause polyploid seeds to germinate at very close to complete success (G = 0.978), simulating the observation that polyploid seeds can have

increased germination success due to germination over a wider temperature range (**?**).

References

Arrigo N., Barker M.S. (2012) Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology*, 15 (2), pp. 140–146.

Creasy M.R., Crolla J.A., Alberman E.D. (1976) A cytogenetic study of human spontaneous abortions using banding techniques. *Human Genetics*, 31 (2), pp. 177–196.

Hashimoto S., Igarashi L., Kobayashi M., Tsutsumi H. (2018) Long-term extrauterine survival in a triploid infant: A review of the clinical features of live-born infants with triploidy. *Hong Kong Journal of Paediatrics*, 23 (4), pp. 298–301.

Jiao Y., Wickett N.J., Ayyampalayam S., Chanderbali A.S., Landherr L., Ralph P.E., Tomsho L.P., Hu Y., Liang H., Soltis P.S., Soltis D.E., Clifton S.W., Schlarbaum S.E., Schuster S.C., Ma H., Leebens-Mack J., Depamphilis C.W. (2011) Ancestral polyploidy in seed plants and angiosperms. *Nature*, 473 (7345), pp. 97–100.

Mayrose I., Zhan S.H., Rothfels C.J., Magnuson-Ford K., Barker M.S., Rieseberg L.H., Otto S.P. (2011) Recently Formed Polyploid Plants Diversify at Lower Rates. *Science*, 333 (6047), p. 1257.

Ramsey J., Ramsey T.S. (2014) Ecological studies of polyploidy in the 100 years following its discovery. *Phil Trans R Soc B*, 369.

Svartman M., Stone G., Stanyon R. (2005) Molecular cytogenetics discards polyploidy in mammals. *Genomics*, 85, pp. 425–430.

Wood T.E., Takebayashi N., Barker M.S., Mayrose I., Greenspoon P.B., Rieseberg L.H. (2009) The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (33), pp. 13875–9.

Woodhouse M., Burkart-Waco D., Comai L. (2009) Polyploidy. *Nature Education*, 2 (1), p. 1.