

Ecological Correlates of Genome Size in Plants

Anne E. Roberts
Department of Biology
Hartwick College
Oneonta, NY 13280

This thesis is submitted in partial satisfaction of the requirements for the degree of
Bachelor of Arts from the Department of Biology, Hartwick College

Thesis advisor

Date

Chair, Biology Department

Date

Ecological Correlates of Genome Size in Plants

Anne E. Roberts

Dr. Stanley K. Sessions, Advisor

ABSTRACT

A wide range in genome size (C-value, mass of DNA per haploid nucleus) is found within and between families of plants. Large genome sizes appear to impose physiological constraints on plants which could potentially limit their evolution, distribution, life-cycle, and geographic range. However, little is known about how genome size affects the composition of a single plant community. In this study, I analyzed the population dynamics of over forty species of vascular plants from a long term (48 year) study of old field succession in regards to genome size. My results suggest that genome size variation may be found along a temporal gradient in a single biotic community.

INTRODUCTION

Succession, or the sequential change in the relative abundance of plants in a community, is often attributed as the consequence of the particular biological characteristics of the component species, such as life cycles or physiological traits (Peet 1992). Other characteristics of component plants that may drive succession include competition between individuals for resources, the alteration of the environment induced by the component plants' use of resources, and various "physiological and energetic constraints" that may interfere with a species' competitive ability (Huston and Smith 1982). While many patterns and mechanisms involved in succession have been recognized, a general theory of successional change has not been established (Peet 1992).

Over the past 50 years, there has been a growing interest in a peculiar characteristic of plants: extensive variation in the mass of genetic material within the nucleus or genome size (C-value, the mass of DNA per haploid nucleus; Bennet and Leitch 2005). Numerous genome size estimates have shown that there is over a 1000-fold variation of C-values in angiosperms alone, and that the distribution is skewed heavily to the right, meaning there are very few plants with large genomes (Knight et al 2005). As cohesion in the field of plant genomics continues to develop, the biological importance of genome size becomes more apparent as a useful tool to understand many aspects of ecological vegetation. Studies have shown that plants with large genomes carry constraints on a number of life and population traits, including length of the mitotic cycle (3), developmental rate and life cycle (Bennett 1972, Bennet 1987), extinction rates (Vinogradov 2003), distribution at a global level (Ohri 2005), seed size, photosynthetic rate, and diversification (Knight et al 2005).

As genome size appears to have an effect on many of the processes which facilitate succession, it seems reasonable to hypothesize that there may be a connection between genome size and ecological vegetative change. The purpose of this study is to test the hypothesis that genome size may play a role in the succession of a single biotic community.

MATERIALS AND METHODS

Genome size is conventionally expressed as C-value, or the mass of DNA in a haploid nucleus. Since plant genome sizes are actually measured from diploid nuclei, and often reported as diploid (2C), I will use genome size to refer to the amount of DNA measured in a diploid (2C) nucleus. First, I made an attempt to measure genome sizes of plants with recorded densities and frequencies in different aged stands by Billings(1936) in an old field successional study. A few of the plants I selected, based on availability, included *Solidago canadensis*, *Morus rubra*, *Acer rubrum*, and *Jugulans nigra*. I used Feulgen image analysis densitometry, a recent accepted method of quantifying DNA using an image-analysis computer program to measure the amount of DNA stained by the Feulgen reaction to estimate genome size (Hardie *et al* 2002). However, since published C-values are readily available for many other plant species, I shifted my focus to conducting a meta-analysis of another study of succession that had many plants with known genome size.

For the meta-analysis, population dynamics of 42 vascular plants with published genome sizes were selected from the Buell-Small Succession study (BSS). The BSS is comprised of ten old-field successional plots, located in Somerset County, New Jersey, USA, 40°30' N, 74°34,' in the Hutchinson Memorial Forest (S.J. Meiners, pers. comm.). Genome size estimates were taken from the Plant DNA C-values database, accumulated at the Royal Botanical Gardens at Kew by Bennett *et al* (2004) (Table 1). Most estimates were measured using either flow cytometry or Feulgen microdensitometry (Bennet *et al* 2004).

For over 48 years, population dynamics had been collected from each of the ten plots at the Buell-Small Successional years by systematic estimations of plant growth during July. These surveys estimated proportion of the plots occupied (frequency) and the percentage of average plot cover (cover) for each species present. From these data, frequency and cover growth rates, years of peak cover and frequency, year of dominance, and span of 50% cover and frequency had been calculated (S.J.Meiners, pers. comm.; Figure1). These parameters were chosen as indicators of succession for this study.

I used a log transformation to remove non-linear relationships within the dataset and then linear regression analyses to look for trends among population dynamics and genome size. Correlations between genome size and population dynamics with regression lines of $R^2 < 0.65$ were evaluated using a Pearson Rank Correlation test. As 10 out of the 42 plants surveyed were from the Poaceae family, this process was repeated within this family.

Table 1. 2C-values and families of representative species.

Family	Species	2C value
Amaranthaceae	<i>Amaranthus retroflexus</i>	1.7
Asteraceae	<i>Hieracium piloselloides</i>	2.15
Asteraceae	<i>Taraxacum officinale</i>	2.55
Asteraceae	<i>Solidago canadensis</i>	3.15
Asteraceae	<i>Centaurea dubia</i>	3.6
Asteraceae	<i>Lactuca serriola</i>	3.7
Asteraceae	<i>Hieracium caespitosum</i>	9.05
Asteraceae	<i>Achillea millefolium</i>	15.3
Asteraceae	<i>Leucanthemum vulgare</i>	21.3
Brassicaceae	<i>Raphanus sativus</i>	2.9
Caryophyllaceae	<i>Silene latifolia</i>	5.4
Caryophyllaceae	<i>Cerastium fontanum</i>	5.85
Chenopodiaceae	<i>Chenopodium album</i>	1.53
Convolvulaceae	<i>Calystegia sepium</i>	1.6
Cyperaceae	<i>Carex sp.</i>	0.77 ^a
Fabaceae	<i>Trifolium pratense</i>	1.3
Fabaceae	<i>Trifolium repens</i>	1.4
Fabaceae	<i>Trifolium hybridum</i>	1.6
Fabaceae	<i>Melilotus sp.</i>	2.45 ^b
Junaceae	<i>Juncus sp.</i>	1.98 ^c
Lamiaceae	<i>Prunella vulgaris</i>	1.3
Liliaceae	<i>Allium vineale</i>	36
Onagraceae	<i>Oenothera biennis</i>	2.44
Plantaginaceae	<i>Plantago lanceolata</i>	2.4
Poaceae	<i>Digitaria sanguinalis</i>	2.4
Poaceae	<i>Setaria faberii</i>	3.25
Poaceae	<i>Sorghum halepense</i>	3.3
Poaceae	<i>Lolium perenne</i>	4.15
Poaceae	<i>Dactylis glomerata</i>	6.6
Poaceae	<i>Agrostis stolonifera</i>	7
Poaceae	<i>Phleum pratense</i>	8.3
Poaceae	<i>Poa pratensis</i>	8.48
Poaceae	<i>Bromus racemosus</i>	21.2
Poaceae	<i>Elytrigia repens</i>	25.95
Polygonaceae	<i>Polygonum aviculare</i>	1.7
Polygonaceae	<i>Rumex acetosella</i>	3.35
Polygonaceae	<i>Rumex crispus</i>	8.8
Portulacaceae	<i>Portulaca oleracea</i>	3.35
Rosaceae	<i>Fragaria virginiana</i>	0.8
Rosaceae	<i>Prunus serotina</i>	1
Rosaceae	<i>Rosa multiflora</i>	1.65
Umbrelliferae	<i>Daucus carota</i>	2

^A Average of 31 values for *Carex sp.*, ^B Average of 3 values for *Melilotus sp.*,

^C Average of 4 values for *Juncus sp.*

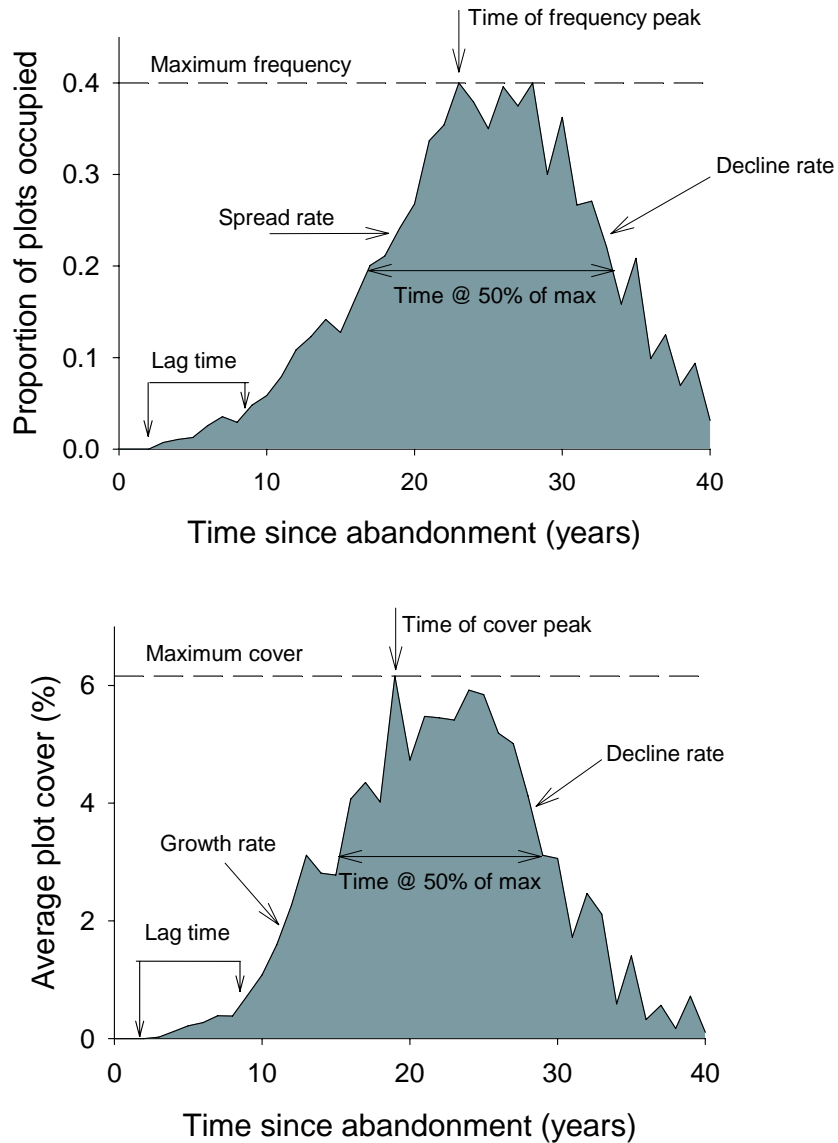


Figure 1. Example of the population dynamics measured for each species. Frequency and cover of *Solidago canadensis* (S.J. Meiners, pers. comm.).

RESULTS

Eighteen families were represented by the 42 species in this study (Table 1). Poaceae and Asteraceae were the two the most common families with 10 representative species in the Poaceae family and eight in the Asteraceae family (Table 1). The other fifteen families were represented by four species or less. Within the entire dataset, there was a wide range in genome size from 0.77 to 36 pg (2C) (Table1). There were only five species with genome sizes over 15 pg (Figure 2). Wide ranges were also recorded within individual families. Within the Poaceae family, genome size ranged from 2.4 to 25.95 pg (2C), and within Asteraceae I found a range from 2.15 to 21.3 pg (2C) (Table 1).

The linear regression analyses did not show any significant trends between genome size and population dynamics in the entire data set (Figure 3). However, six dynamics (year of peak cover, year of peak frequency, span of 50% frequency, year of

maximum dominance, and maximum cover) all showed similar positive slopes in regards to genome size (Figure 2). Within the single Poaceae family, significant correlations were found between genome size and the year of peak cover, the span of 50% frequency, and the year of peak frequency (Figure 4). Again, the relationships showed a positive slope. These trends, although not significant, were mirrored between the same variables within the larger dataset (Figure 5).

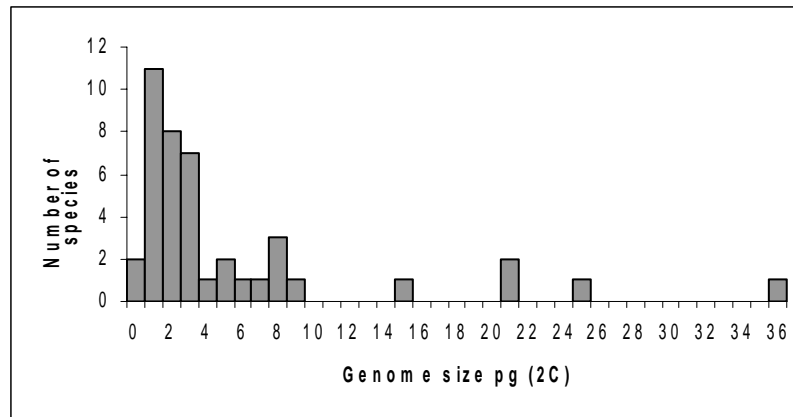


Figure 2. A bar graph representing distribution of genome size within the total dataset.

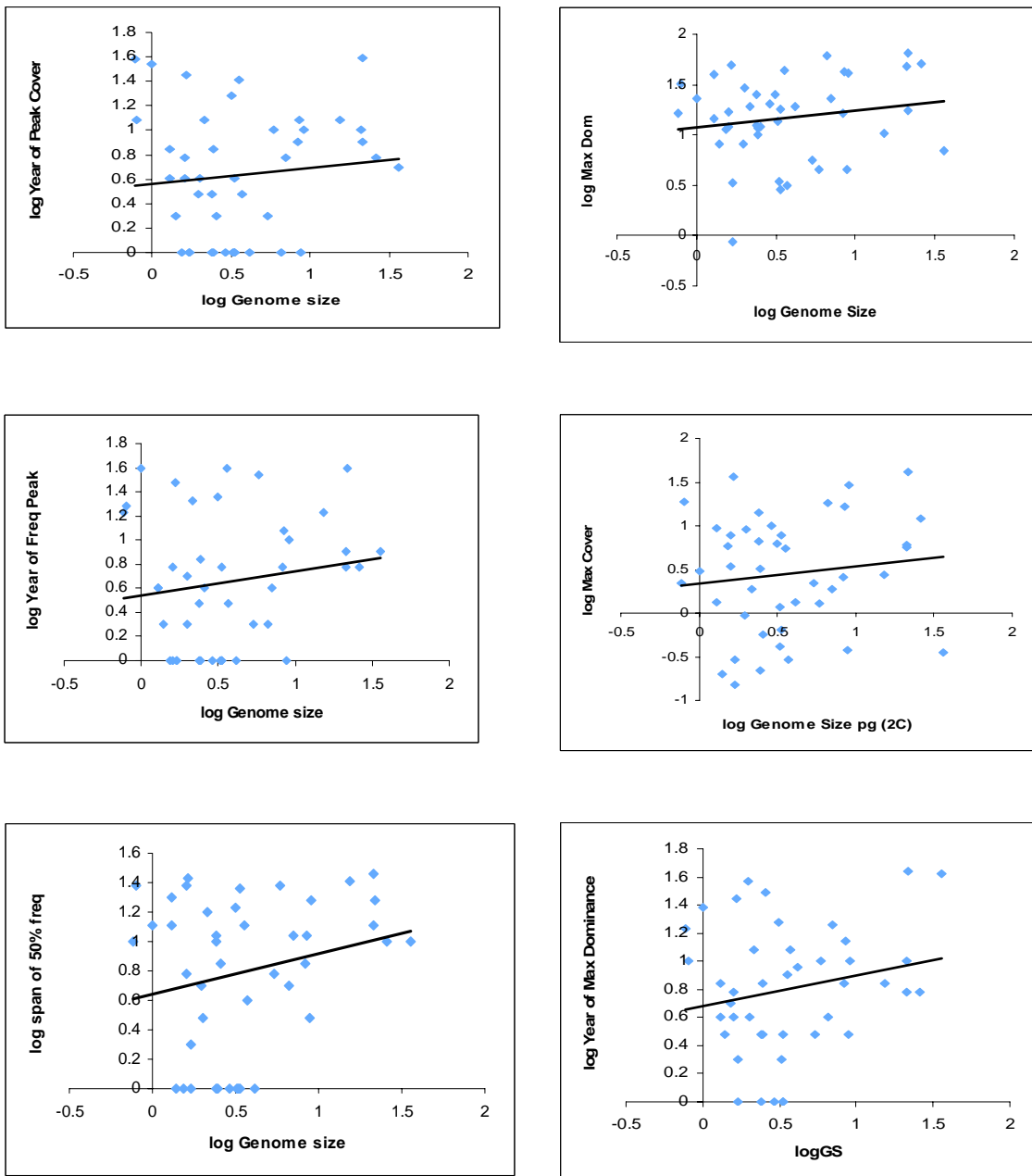


Figure 3. General increasing trends seen in year of peak cover, year of freq peak, span of 50% freq, max dominance, year of maximum dominance and maximum cover in the linearized dataset in relationship to the log transformed genome size.

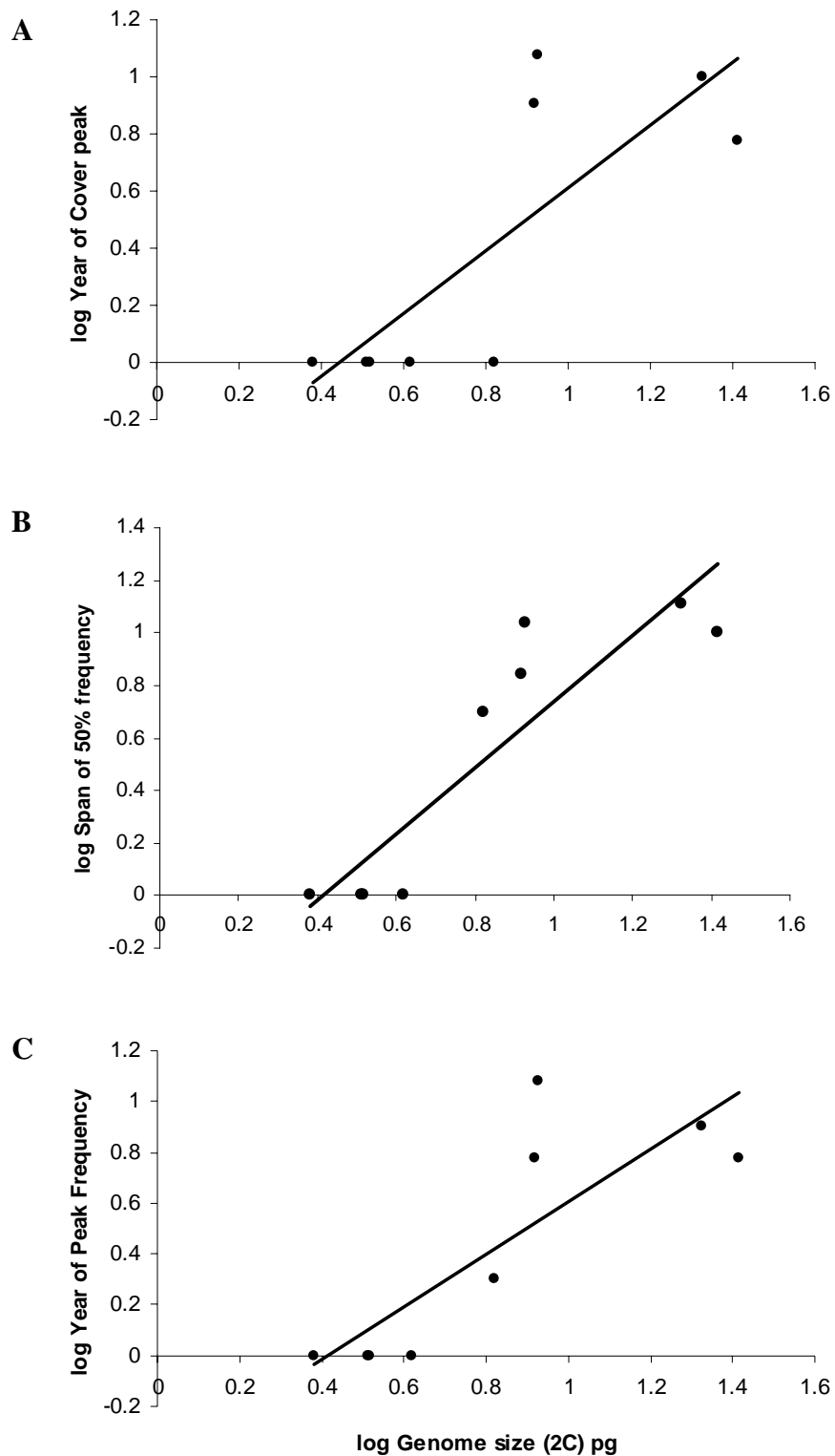


Figure 4. Relationships between genome size and population dynamics for 10 species of grass. Linear regression line shown for year of peak cover and genome size (A, $R^2=0.66$), span of 50% frequency (B, $R^2=0.79$), and year of peak frequency (C, $R^2=0.74$).

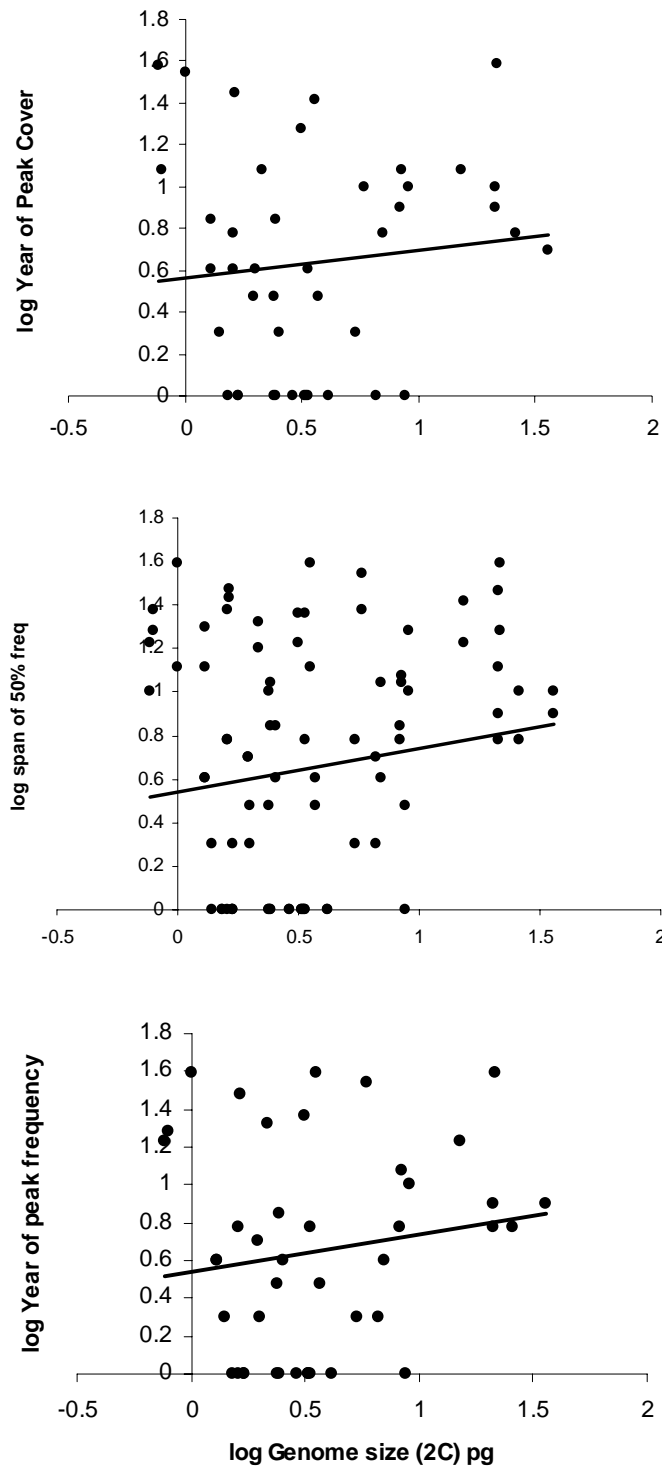


Figure 5. The general trends of year of cover, span 50% frequency, and year of peak frequency in the entire dataset.

DISCUSSION

The results of this study suggest some fascinating ideas concerning spatial and temporal distribution of genome size variation in plants. First, the wide range in genome size recorded for the species used in this study support the research indicating that there is an extensive variation of nuclear genetic material between and within plant families. It also supports the idea that wide variation of genome size can be found in species located in the same environment and that species with large genomes are relatively uncommon. The similar positive linear trends found in year of peak cover, year of peak frequency, span of 50% frequency, year of maximum dominance and maximum cover in relationship to genome size suggest that that variation in genome size can be found across a temporal gradient as well. Simply put, species with small genomes may be more likely to cover larger areas of a plot, and more plots at an earlier date in secondary succession than species with large genomes.

Within the Poaceae family, it appears that species with large genomes have significantly later years of peak cover and peak frequency and longer spans of 50% frequency. This may indicate that, while grass species with larger genomes peak at later dates, they are found in more plots for a longer duration. While plants with large genomes have been suggested to have several constraints affecting their growth, development, and distribution, from a successional standpoint, many of these constraints may be interpreted as evolved traits suitable for filling a certain temporal niche. It is possible that plants with large genomes evolved in order to survive in a different temporal niche to compete with plants with small genomes, or found it possible to survive only in a later temporal niche because of their large genomes.

These significant positive relationships found in the Poaceae family may also bring some validation to the interpretation of temporal variation on a larger scale mentioned above. However, if all the phylogenetic relationships of the plants used in this study are determined, it is possible these results may shift in a number of ways. First, phylogenetic noise may be one of the reasons that significant relationships were not found within the whole dataset. If the sample size could be reduced by clustering closely related species, perhaps a significant trend could be seen in the whole dataset. On the flip side, reducing the sample size could also reduce the chances of finding significant relationships, especially within the individual families. As our understanding of phylogenetic relationships between plants are rapidly changed and rearranged on a regular basis, due to advances in phylogenetic systematics and the expanding field of plant genomics, it may be only a matter of time before data for determining the relationships between the species used in this study become available.

Until that time, however, a great deal more work can be accomplished using the data from the Buell-Small's Successional study. First, I would still suggest measuring genome size of the remaining 40 species analyzed by the BSS, specifically focusing on families where more representatives are found in the dataset, such as the Asteraceae family. It is also possible that another method can be used to analyze more specifically the rates of change in the larger dataset, such as quantile regression. This could pin-point significant changes in the data which might be skewed by the large number of species with small genome sizes when analyzed linearly. It would also be interesting to see these trends of genome sizes in relationship to temporal change are found in other successional studies or if the results of this study are site specific.

If this pattern is seen in other successional sites, it may be that genome size can be used as tool to predict ecological vegetative change. This could be particularly useful in reforestation or other physical applications, such as successional agriculture. However, the implication that the results of this study have on the evolutionary significance of genome size is probably more important. While it still remains in question if genome size evolution is maladaptive, adaptive, or neutral, it seems that plants with large genomes have found alternative ways to be successful in a community, disregarding their success on a larger evolutionary scale, by dominating at later dates and for a longer duration.

ACKNOWLEDGEMENTS

I would like to thank Dr. Scott Meiners (EIU) for his collaboration, Dr. Stanley Sessions for his guidance, Rob Hunt, Deb Woodham, and everyone in the Biology, Math, and Economics department at Hartwick College who helped with the statistical evaluations of my project.

LITERATURE CITED

- Billings, W. D. 1938. The structure and development of old field shortleaf pine stands and certain associated physical properties of the soil. *Ecological Monographs*. 8(3): 437-499.
- Bennett, M.D. 1972. Nuclear DNA content and minimum generation time in herbaceous plants. *Proceedings of the Royal Society of London, Series B* 178: 259-275
- Bennet, M.D. 1987. Variation in genomic form in plants and its ecological implications. *New Phytologist* 106: 177-200.
- Bennett, M.D., and I.J. Leitch. 2005. Nuclear DNA amounts in Angiosperms: Progress, problems, and prospects. *Annals of Botany* 95: 45-90.
- Bennet, M.D., A.V. Cox, and I.J. Leitch. 2001. Angiosperm DNA C-values database. <http://www.rbgekew.org.uk/cval/database1.html>. Accessed 4/2007.
- Hardie, D.C., T.R. Gregory, and P.D.N Herbert. 2002. From pixels to picograms: A beginners' guide to genome quantification by Feulgen image analysis densitometry. *The Journal of Histochemistry & Cytochemistry* (50)6: 735-749.
- Huston, M. and T. Smith. 1986. Plant Succession: life history and competition. *The American Naturalist* 130(2): 168-198
- Knight, C.A., N.A. Molinari, D.A. Petrov. 2005. The large genome constraint hypothesis: Evolution, ecology, and phenotype. *Annals of Botany* 95: 177-190.
- Ohri, D. 2005. Climate and growth form: The consequences for genome size in plants. *Plant Biology* 7: 449-458.
- Peet, R.K. 1992. Community structure and ecosystem function. *In: Plant Succession: Theory and predictions.* (Glenn-Lewin D.C., R.K. Peet, and T.T. Veblen Edt.) Chapman & Hall, London.
- Vinogradov AE. 2003. Selfish DNA is maladaptive: evidence from the plant Red list. *Trends in Genetics* 19: 609-614.