Latitudinal variation in plant nuclear DNA content at the global scale and its significance for climate adaptation: a case study of Asteraceae (Postprint)

Authors: dream interpretation, Yu Jing, Guo Shuiliang

Date: 2018-05-29T00:00:00+00:00

Abstract

Nuclear DNA content is an important biological concept that encompasses DNA C-value and genome size. Previous studies have reported contradictory findings regarding the variation patterns of plant nuclear DNA content along latitudinal gradients, with most describing the relationship between nuclear DNA content and factors such as latitude, altitude, and climate as linear. Whether nuclear DNA content holds significance for environmental adaptation remains controversial. These contradictory reports on the relationship between nuclear DNA content and environmental factors may be attributed to small sample sizes, narrow geographic ranges, and substantial differences in the genetic background of study subjects. If a taxonomic group with similar genetic background is sampled globally, nuclear DNA content may exhibit regular latitudinal gradient variation, potentially showing non-linear relationships with major climatic factors. Asteraceae (Compositae) is the largest family of angiosperms and a widely recognized natural taxonomic group. To reveal the variation patterns of plant nuclear DNA content along latitudinal gradients at a global spatial scale and to assess whether such variation has environmental adaptive significance, statistical analyses were conducted on the relationships between nuclear DNA content and latitude and bioclimatic factors using Asteraceae as the study system. Nuclear DNA content data for 822 Asteraceae species were retrieved from the "Plant DNA C-values Database" ; globally, 10 transects were established along the longitudinal direction, each spanning 15 degrees of longitude, with each transect further divided into 22 blocks spanning 7.5 degrees of latitude each; second, data for 14 bioclimatic factors for the period 1950-2000 were downloaded from the "WorldClim" website, and ArcGIS 9.3 was used to obtain the average values of the 14 bioclimatic factors for each block; based on records from the "Global Biodiversity Information Facility" (GBIF), the average nuclear DNA content data for Asteraceae plants in each block were calculated. To avoid the impact of multicollinearity among climatic variables on data analysis, principal component analysis was applied for dimensionality reduction. Mean temperature of the coldest quarter and precipitation of the driest quarter were found to be the factors with the highest loadings on the first and second principal components, respectively. After removing other climatic factors with correlations between -0.7 and +0.7 with them, three variables—mean temperature of the coldest quarter, precipitation of the driest quarter, and precipitation of the wettest month—were retained for further data analysis. The results revealed that nuclear DNA content of Asteraceae plants across the 10 global transects was closely related to latitude, and exhibited highly significant unimodal non-linear relationships with mean temperature of the coldest quarter, precipitation of the driest quarter, and precipitation of the wettest month, which could be fitted with binomial equations. Therefore, at a global spatial scale, plant nuclear DNA content shows regular non-linear variation along latitudinal gradients, and this variation holds strong climatic adaptive significance.

Full Text

Latitudinal Variation Patterns of Plant Nuclear DNA Amount on a Global Scale and Their Climatic Adaptation Significance: A Case Study with Asteraceae

Xie Meng, Yu Jing, Guo Shuiliang*

College of Life and Environmental Sciences, Shanghai Normal University, Shanghai 200234, China

Abstract

Nuclear DNA amount is an important biological concept that includes DNA C-value and genome size. There are conflicting reports about latitudinal variation patterns of plant nuclear DNA amount. Relationships of plant nuclear DNA amount with latitude, altitude, and climatic variables have been reported as linear. Disputes also exist as to whether nuclear DNA amount of plants is of environmental adaptation significance. We speculated that the conflicting reports were due to insufficient sampling of the taxa, limitations in the geographical range, and different genetic backgrounds of the samples. Asteraceae is not only the largest family within angiosperms, but also a widely accepted natural taxa. To clarify the latitudinal variation patterns of plant nuclear DNA amount on a global scale, and to explore the possible climatic adaptation significance of these patterns, we conducted a study with Asteraceae to analyze the relationships of nuclear DNA amount with latitude and bioclimatic variables. First, we obtained the data of nuclear DNA amounts of 822 species of Asteraceae from the Plant DNA C-value database. We selected ten global longitudinal transects, each with a span of 15 longitudinal degrees, and evenly divided each transect (from 82.5°N to 82.5°S) into 22 blocks. We obtained geographical records of the 822 species from the Global Biodiversity Information Faculty, together with fourteen bioclimatic factors within these blocks from the Worldclim Global Climate Database. We calculated the average DNA 1C-values and genome sizes of each species in each block. To avoid multicollinearity among the fourteen climatic variables, we performed principal component analysis (PCA) to reduce the dimensionality of the variables. We found that the mean temperature of the coldest quarter and the precipitation of the driest quarter had the highest loads in the first two principal components. The climatic variables with low correlation coefficients (-0.7 to 0.7) with the above two variables were included in the analyses. We found that the nuclear DNA amounts of Asteraceae had a regular latitudinal variation that could be expressed by polynomial functions. The relationships of nuclear DNA amount with the mean temperature of the coldest quarter, the precipitation of the driest quarter, and the wettest month were typically nonlinear with a unimodal pattern, and could be expressed by binomial equations. Therefore, on the global scale, plant nuclear DNA amount changes regularly on latitudinal gradients, which has distinct climatic adaptation significance.

Keywords: Asteraceae; nuclear DNA amount; DNA C-value; genome size; nonlinear relationship; latitude; bioclimatic variable

Introduction

Genome size [1], or the DNA C-value, refers to the amount of DNA in an unreplicated gamete nucleus of a species. The genome size of a species equals its C-value divided by its ploidy level. The traditional view holds that DNA C-value is highly constant for a species (with ‘C’ representing ‘constant’) [2]. However, with deeper research, it has become clear that DNA amount shows tremendous variation not only between species but also among closely related taxa [3], and holds significance for biogeography and ecological adaptation [4]. Some major worldwide weeds have smaller genomes than other plants, with malignant weeds having smaller genomes than general weeds [3]. Bennett and Guo Shuiliang et al. found that invasive plants within China’ s angiosperms often have smaller C-values than their non-invasive relatives [5-6]. DNA C-value shows positive correlation with altitude in some taxa, and regular variation with rainfall differences, suggesting adaptation to climatic conditions [7]. This is considered an important cause of niche differentiation. For example, compared with birds, amphibians show great variation in DNA content among species, which is related to their highly diverse habitats [8]. Similar patterns exist among higher plant taxa, with even greater differences in nuclear DNA content [9]. However, conclusions about relationships between nuclear DNA content and geographic/climatic factors are inconsistent. Early studies reported positive correlations between nuclear DNA content and latitude [9], while Grime and Mowforth found negative correlations between DNA C-value and latitude in the British flora [10]. Razafinarivo et al. found that DNA C-value in Madagascar coffee trees increased from north to southeast, and across the African continent from east to west [11]. Bottini et al. found that in *Berberis* diploid species, taxa with small DNA content tended to distribute in high-rainfall, low-productivity high-altitude regions, while those with high DNA content tended to distribute in mid-altitude forests with long growing seasons and high water availability [12]. Other studies found no correlation between nuclear DNA content and environment [13-14]. In most early literature, relationships between DNA content and latitude were described as linear. However, Knight and Ackerly, through quantile regression analysis of California flora, speculated that non-linear relationships might exist between nuclear DNA content and environmental factors [15]. Li Dandan et al. found that barnyard grass populations across China could be fitted with unimodal functions [16].

1. Data Sources

Nuclear DNA content data for 822 Asteraceae species were obtained from the Angiosperm DNA C-value database [18]. Genome size was calculated as C-value divided by ploidy level. Among the retrieved species, those with ploidy information were used to calculate genome size.

To quantitatively analyze the effects of latitude, longitude, and climatic factors on nuclear DNA content in Asteraceae, we established ten global longitudinal transects, each spanning 15° of longitude: 90°-75°W, 75°-60°W, 60°-45°W, 0°-15°E, 15°-30°E, 30°-45°E, 75°-90°E, 90°-105°E, 105°-120°E, and 120°-135°E. Transects 1, 2, and 3 covered Asia and Oceania; transects 4, 5, and 6 covered Europe and Africa; and transects 7, 8, and 9 covered North and South America. Each transect was evenly divided into 22 blocks spanning from 82.5°N to 82.5°S, with each block covering 7.5° of latitude.

Species geographic distribution data were obtained from the Global Biodiversity Information Facility (GBIF, www.gbif.org). Bioclimatic data were downloaded from the WorldClim Global Climate Database (http://www.worldclim.org) at 30 arc-seconds resolution (~86 km^2 per grid cell). Fourteen bioclimatic variables (Table 1) were extracted for each block using ArcGIS 9.3 and averaged across each block.

2. Statistical Analysis

First, we identified Asteraceae species with nuclear DNA content data in each block. Only blocks containing more than one species were used for statistical analysis. We calculated the mean DNA C-value, genome size, and their standard errors for Asteraceae species in each block.

Pearson correlation analysis was conducted with the mean bioclimatic factor values for each block as independent variables and DNA C-value as the dependent variable. We found non-linear relationships between DNA C-value and most climatic factors. Further calculation of correlations among the 14 climatic factors revealed high correlations between some factors, suggesting multicollinearity.

Therefore, conventional multiple stepwise linear regression was not used for variable selection.

To avoid the impact of multicollinearity among climatic factors on data analysis, we first applied principal component analysis (PCA) to reduce dimensionality. We selected the climatic factors with the highest loadings on the first two principal components, then removed climatic factors with Pearson correlation coefficients between -0.7 and 0.7 with these key factors. The best-fitting functions were applied to reflect relationships between climatic factors and DNA C-value. Before fitting, each climatic factor was range-standardized to make their variation ranges comparable. XLSTAT software was used for analyses.

3. Results

3.1 Variation Range of Nuclear DNA Content in Asteraceae Among the 822 Asteraceae species, the maximum DNA 1C-value was 26.15 pg (*Senecio pendulus*) and the minimum was 0.19 pg (*Leontodon*), with a variation range of 0.40-4.11 pg. The largest genome size was 8.5 pg and the smallest was 2.23 pg.

3.2 Latitudinal Variation Trends Across Ten Transects Nuclear DNA content showed a decreasing trend with increasing latitude northward from the equator. The latitudinal variation of nuclear DNA content across the ten transects is shown in Figure 2. The relationships between DNA 1C-value, genome size, and latitude could be fitted with polynomial functions, reaching significant or highly significant levels. In the first transect, the relationship between DNA C-value and latitude also conformed to a quadratic function.

3.3 Effects of Climatic Factors on Transect 5 PCA of the 14 climatic factors across the 22 blocks of transect 5 showed that the first two principal components explained 28.6% and 58.9% of the total variance, respectively (Table 2). The first principal component primarily reflected temperature factors, with mean temperature of the coldest quarter having the highest loading (C7). The second principal component primarily reflected precipitation factors, with precipitation of the driest quarter having the highest loading (C12).

We selected mean temperature of the coldest quarter, precipitation of the wettest month (C9), and precipitation of the driest quarter for further analysis. Fitting revealed that binomial equations better described the relationships between these three climatic factors and DNA C-value than linear relationships (Figure 3). All three factors significantly affected Asteraceae nuclear DNA content, showing significant unimodal non-linear relationships.

To compare the relative influence of the three climatic factors on DNA C-value, we standardized them to the same scale and fitted DNA 1C-value against the three factors. The results showed that mean temperature of the coldest quarter had the greatest influence on DNA C-value (range = 0.75, SD = 0.24), followed by precipitation of the driest quarter (range = 0.49, SD = 0.15), while precipitation of the wettest month had relatively smaller influence (range = 0.40, SD = 0.13). In terms of climatic conditions, blocks with moderate climate conditions tended to have higher DNA C-values, which decreased regularly as conditions moved toward extremes (Figure 4).

4. Discussion

While many studies have reported that nuclear DNA amount affects plant geographic distribution and ecological adaptation, the relationship between nuclear DNA amount and latitude/climatic factors remains controversial, with reports of positive correlation, negative correlation, or no correlation. Most studies have described these relationships as simple linear relationships. We speculate that contradictory reports may arise from: (1) study objects lacking consistent genetic backgrounds, (2) narrow sampling ranges, and (3) oversimplification of relationships as purely linear, which may also contribute to contradictory conclusions [15].

Li Dandan et al. found that barnyard grass populations across China could be fitted with unimodal functions [16]. Using the entire Asteraceae family with nuclear DNA content data, this study confirms on a global scale that nuclear DNA content shows regular non-linear variation with latitude [16]. Among the climatic factors, precipitation of the driest quarter and precipitation of the wettest month showed clear unimodal curves with DNA C-value. While DNA C-value appeared to decrease monotonically with decreasing mean temperature of the coldest quarter, binomial fitting of this relationship yielded predicted values significantly correlated with observed values at the 0.001 level, with much better fit than linear models ($R^2 = 0.638$ vs. 0.562), showing clear unimodal characteristics.

Levin and Funderburg compared DNA 4C-values of temperate and tropical plants, finding temperate species averaged 27.06 pg while tropical species averaged only 12.13 pg [9]. However, they argued this pattern was not universal and resulted from geographic replacement of lineages with different DNA content during evolution rather than environmental adaptation, suggesting DNA content variation lacks climatic adaptation significance. In contrast, this study found three bioclimatic factors significantly affected Asteraceae nuclear DNA content, with unimodal functions well explaining these relationships. Li Dandan et al. also found within-species DNA content variation was significantly affected by climatic factors [16], indicating this index has clear ecological adaptation significance.

However, this study found that blocks at or near the equator showed decreasing Asteraceae nuclear DNA content trends, with smaller genomes [19]. This pattern was particularly evident in certain transects and may relate to niche breadth. Bennett found tropical species often inhabit narrow climate ranges [19]. Stebbins argued that compared to temperate species, tropical species require fewer genes due to stable environments [20]. Temperate plants have wider niches due to duplicated structural genes, increasing nuclear DNA content. However, toward the poles, nuclear DNA content also decreases due to shortened growing seasons or uniform environments [21]. Unimodal models are commonly used to explain species distribution-environment relationships [22], and Vidic et al. argued large-genome plants are unsuitable for extreme environments [21].

A species distributed from temperate to tropical zones typically has optimal growth regions with corresponding climatic conditions, morphological traits, and nuclear DNA content, theoretically fitting unimodal models. How does nuclear DNA content affect plant morphology? Previous studies show small-DNA plants have higher seedling growth rates [23-24], increased leaf mass per area, higher cell division rates, and enhanced metabolic rates [25]. Nuclear DNA content affects cell size and weight through both genetic and physical mechanisms, with additive nuclear type effects extending to multicellular tissue structure size and development rates [26]. Strong positive correlations exist between genome size and cellular phenotypes such as guard cell length and epidermal cell area [26], giving theoretical significance to the ecological adaptation value of nuclear DNA content [16].

Asteraceae is a worldwide large family and a relatively natural taxonomic group with consistent genetic and evolutionary backgrounds. Data analysis across global transects gives our results certain universality. However, as a highly diverse plant group, whether these geographic variation patterns and climate relationships are universal across higher plants requires further research in more taxonomic groups. Due to data limitations, this paper mainly analyzed latitudinal variation in Asteraceae nuclear DNA content and its climatic factors. Relationships between nuclear DNA content and longitude/altitude factors globally are topics for future research.

References

[1] Leitch I J, Bennett M D. Genome downsizing in polyploid plants. Biological Journal of the Linnean Society, 2004, 82(4): 651-663. [2] Swift H. The constancy of deoxyribose nucleic acid in plant nuclei. Proceedings of the National Academy of Sciences of the United States of America, 1950, 36(11): 643-654. [3] Bennett M D. Plant genome values: how much do we know? Proceedings of the National Academy of Sciences of the United States of America, 1998, 95(5): 2011-2016. [4] Chen G Q, Guo S L, Yin L P. Applying DNA C-values to evaluate invasiveness of angiosperms: validity and limitation. Biological Invasions, 2010, 12(5): 1335-1348. [5] Guo S L, et al. DNA C-values and invasiveness of angiosperms: statistical analysis of 1200 species in China. Acta Ecologica Sinica, 2008, 28(8): 3698-3705. [6] Guo S L, et al. DNA C-value determination and its biological significance for 102 herbaceous plants in the Yangtze River Delta and adjacent regions. Acta Ecologica Sinica, 2015, 35(19): 6516-6529. [7] Bennett M D. DNA amount, latitude, and crop plant distribution. Environmental and Experimental Botany, 1976, 16(2/3): 93-108. [8] Gregory T R. Genome size and

developmental complexity. Plant Molecular Biology, 2005, 59(1): 1-5. [9] Levin D A, Funderburg S W. Genome size in angiosperms: temperate versus tropical species. The American Naturalist, 1979, 114(6): 784-795. [10] Grime J P, Mowforth M A. Variation in genome size—an ecological interpretation. Nature, 1982, 299(5879): 151-153. [11] Razafinarivo N J, et al. Geographical gradients in the genome size variation of wild coffee trees native to Africa and Indian Ocean islands. Tree Genetics & Genomes, 2012, 8(6): 1345-1358. [12] Bottini M C J, et al. Relationships among genome size, environmental conditions and geographical distribution in natural populations of Berberis species in NW Patagonia. Annals of Botany, 2000, 86(3): 565-573. [13] Teoh S B, Rees H. Nuclear DNA amounts in populations of Picea and Pinus species. Heredity, 1976, 36(1): 123-137. [14] Creber H M C, et al. Variation in DNA C-value in natural populations of Dactylis glomerata L. New Phytologist, 1994, 128(3): 555-561. [15] Knight C A, Ackerly D D. Variation in nuclear DNA content across environmental gradients: a quantile regression analysis. Ecology Letters, 2002, 5(1): 66-76. [16] Li D D, et al. Nuclear DNA contents of Echinochloa crus-galli and its Gaussian relationships with environments. Acta Oecologica, 2017, 79: 36-47. [17] Rayburn A L, Auger J A. Genome size variation in Zea mays adapted to different altitudes. Theoretical and Applied Genetics, 1990, 79(4): 470-474. [18] Bennett M D, Leitch I J. Angiosperm DNA C-values database [DB/OL]. http://www.kew.org/cvalues/. 2017-10-30. [19] Bennett M D. DNA amount, latitude, and crop plant distribution. Environmental and Experimental Botany, 1976, 16(2/3): 93-108. [20] Stebbins G L. Chromosomal variation and evolution. Science, 1966, 152(3728): 1463-1469. [21] Vidic T, et al. Selective significance of genome size in a plant community with heavy metal pollution. Ecological Applications, 2009, 19(6): 1515-1521. [22] Cui H J, et al. Influence of Mn(III) availability on the phase transformation from layered buserite to tunnel-structured todorokite. Clays and Clay Minerals, 2008, 56(4): 397-403. [23] Beaulieu J M, et al. Genome size evolution in relation to leaf strategy and metabolic rates revisited. Annals of Botany, 2007, 99(3): 495-505. [24] Knight C A, Beaulieu J M. Genome size scaling through phenotype space. Annals of Botany, 2008, 101(6): 759-766.

Table 1 Fourteen bioclimatic variables

Figure 1 [Figure 1: see original paper] Location of ten transects

Figure 2 [Figure 2: see original paper] Variation of nuclear DNA content of Asteraceae species across ten global latitudinal transects

Figure 3 [Figure 3: see original paper] Influences of three climatic factors on DNA 1C-values of Asteraceae on Section 5

Figure 4 [Figure 4: see original paper] Comparison of influences of three climatic factors on DNA 1C-values of Asteraceae on Section 5

Note: Figure translations are in progress. See original paper for figures.