Distribution Patterns of Genetic Alleles: Towns vs. Cities

Brea Baker

Freeport High School, Freeport NY USA

Abstract. Relationships between population size and DNA inheritance patterns were studied with data from the Bioserver genetic database (Cold Spring Harbor Laboratory, Cold Spring Harbor NY USA) [1]. The Alu PV92 allele, a human specific Alu insertion on chromosome 16, was used as a marker. Five population centers were selected by size from samples of small, isolated towns and large cities. Chi Square values were found to increase with population size according to the following equation (y = 0.9161Ln(x) – 6.5221). Heterozygosity was also found to increase with population size according to the following equation (y = 0.063Ln(x) - 0.3539).

Keywords: molecular biology, bioinformatics, genetics, Hardy-Weinberg, heterozygosity.

INTRODUCTION

A study on human mating by Boyce (1988) examines the causes and consequences of different mating patterns in man with particular reference to biological, medical and demographic factors [7]. The study covers historical and demographic aspects; mate choice and assortative mating; social systems, religious rules and mating practices; medical and genetic issues. I believe that due to a smaller population, small towns will have more similar DNA inheritance patterns than those that reside in cities. This is due to the population differences. Cities not only have larger populations than small towns but also a higher migration rate which increases population diversity. Mark Lefers of Northwestern University once stated that heterozygosity is the measure of the genetic variation in a population [4]. According to Wikipedia, Hardy-Weinberg equilibrium is expressed with a simple equation

$$X^2 = \Sigma \frac{(\text{Observed frequency} - \text{Expected frequency})^2}{\text{Expected frequency}} \tag{1}$$

that can be used to discover the probable genotype frequencies in a population and to track their changes from one generation to another [2]. As a consequence, allele and genotype frequencies in a population remain constant from generation to generation unless specific disturbing influences are introduced. Those disturbing influences include non-random mating, mutations, selection, limited population size, random genetic drift and gene flow.

METHODS

Data about AluPV92 alleles was obtained from an online database (www.bioservers.org) maintained by Cold Spring Harbor Laboratory (Cold Spring Harbor NY USA). Five populations of various sizes were identified: 2 small populations (fewer than 10,000) and 3 large populations (greater than 40,000) (Table 1). The data set was analyzed for Chi Square and genetic drift using the database software (1).

TABLE I. The names, populations and percentages of population sampled. Population levels were obtained from a variety of online sources.

Populations That Were Analyzed	Population	Percentage of Population Sampled
Dallas, Texas	989,496	0.00001%
Funchal, Portugal	100, 847	0.00007%
Landau, Germany	43,048	0.0003%
Hedemora, Sweden	7,279	0.0008%
Ryssby, Sweden	689	0.004%

RESULTS

Figure 1 shows that chi square of the allele distribution in the particular population relative to an ideal distribution increases logarithmically with the population size. Using Microsoft Excel, the logarithm of the population was plotted along the horizontal axis and the chi square values were plotted along the vertical axis. A logarithmic trend line was fit to the data. The equation of the best fit line is:

$$y = 0.9161\mathrm{Ln}(x) - 6.5221 \tag{2}$$

Figure 2 shows that the heterozygosity of the allele distribution in the particular population increases logarithmically with the population size. Using the Microsoft Excel, the logarithm of the population was plotted along the horizontal axis and heterozygosity values were plotted along the vertical axis. A logarithm trend line was fit to the data. The equation of the best fit line:

$$y = 0.063\mathrm{Ln}(x) - 0.3539 \tag{3}$$

DISCUSSION

Despite the clear trend, it must be recognized that the sample sizes are quite small. For instance, in the best case (Ryssby, Sweden), merely 0.004% of the population is represented. And in the worst case (Dallas, Texas), only 0.00001% of the population is represented (Table 1).

CONCLUSIONS

As a result of studying Alu PV92 alleles from different populations, I observed that the small populations appear closer to the ideal equilibrium (Hardy Weinberg). Therefore, insofar as the Alu database is accurate, larger populations were introduced to specific disturbing influences such as gene flow or random genetic drift (or possibly mutation). I also observed that the smaller populations had a smaller heterozygosity. Thus, larger populations have a greater variety of genetic combinations. It is well known, from psychological studies, that people tend to choose mates similar to themselves. However, this study puts a limit on that similarity and demonstrates that people will occasionally take advantage of the greater variety available.

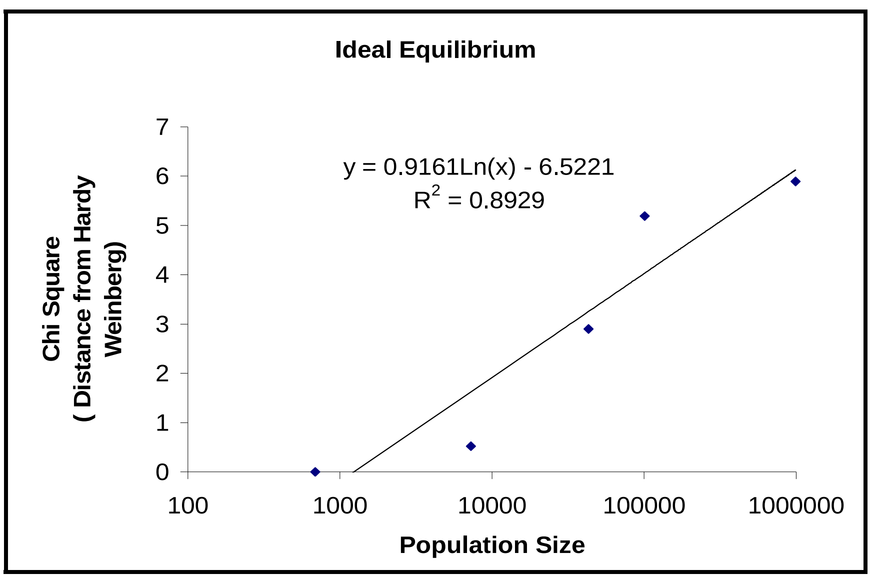

Figure 1. A graph exhibiting the relationship between chi square values and population size. As population size increases so does the distance from Hardy-Weinberg.

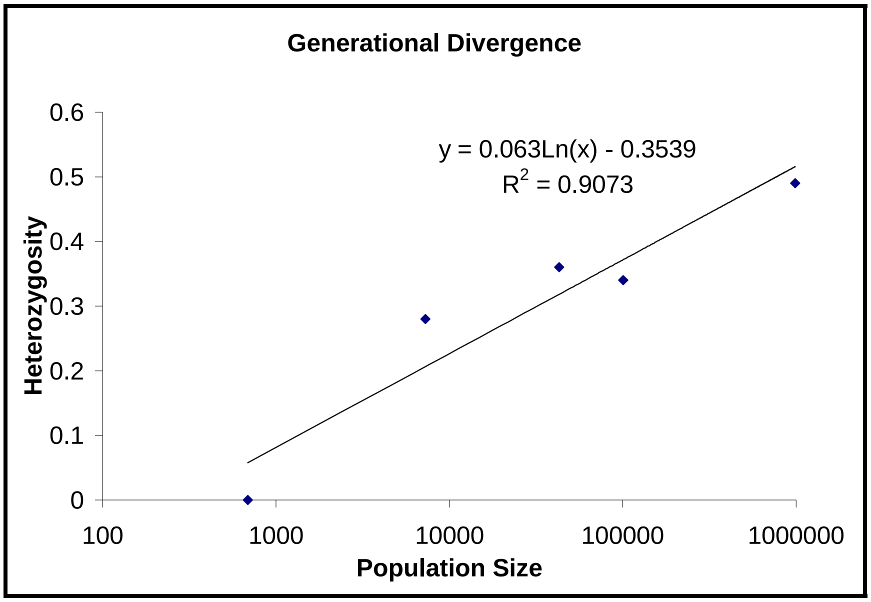

Figure 2. A graph exhibiting the relationship between heterozygosity and population size. As population size increases so does the heterozygosity.

ACKNOWLEDGMENTS

The author gratefully acknowledges the contributions of Dr. Robert Muratore and the support of Freeport Public Schools.

REFERENCES

1. Custom Workspaces and Educational Databases for Bioinformatics: Allele Server. Dolan DNA Learning Center, Cold Spring Harbor Laboratory.http://www.bioservers.org/html/sad/sad.html
2. Hardy-Weinberg Principle. Wikipedia. http://en.wikipedia.org/wiki/Hardy-Weinberg_principle Accessed 2009.
3. Alu Insertion Polymorphism: Theory. Genetic Origins. http://www.geneticorigins.org/pv92/aluframeset.htm
4. Lefers, Mark. Heterozygosity. Department of Biochemistry, Molecular Biology, and Cell Biology 2004. http://www.biochem.northwestern.edu/holmgren/Glossary/Definitions/Def-H/heterozygosity.html
5. Fisher R.A., Yates F. Chi Square Test. Statistical Tables for Biological Agricultural and Medical Research. 6th edition. http://www2.lv.psu.edu/jxm57/irp/chisquar.html
6. Moran, Laurence. Random Genetic Drift. The talk Origins Archive 1993. http://www.talkorigins.org/faqs/genetic-drift.html
7. Boyce, Anthony. Human Mating Patterns. Society for the Study of Human Biology Symposium Series (No. 28) 1988