EXCLI Journal 2008;7:151-153 – ISSN 1611-2156
Received: April 27, 2008, accepted: July 1, 2008, published: July 7, 2008

Original article:

NONRANDOM GENE DISTRIBUTION ON HUMAN CHROMOSOMES

Laleh Rafiee, Sasan Mohsenzadeh, Mostafa Saadat*

Department of Biology, College of Sciences, AND Institute of Biotechnology, Shiraz University, Shiraz 71454, Iran

* Corresponding author: Department of Biology, College of Sciences, Shiraz University, Shiraz 71454, Iran. E-mail addresses: saadat@susc.ac.ir AND msaadat41@yahoo.com

ABSTRACT

Human chromosomes are heterogeneous in structure and function. This is the reason for specific banding patterns produced by various chromosome staining techniques. The human genome is a mosaic of isochors and can be partitioned into five families, L1, L2, H1, H2 and H3, characterized by increasing GC level and gene concentrations. In this study we investigated the chromosome distribution of 22845 genes mapped at whole chromosomes reported in the Human Genome Data Base as of January 2007. Pearson correlation coefficient analysis showed that there is significant correlation between the number of mapped genes and percent of G-dark bands (r=-0.608, p=0.002). Also the correlation between the ratio (observed versus expected genes) and percent of G-dark bands was significant (r=-0.506, p=0.012). There was a significant difference between observed number of mapped genes and expected number of mapped genes on human chromosomes (χ^2=4842.7, df=23, p<0.00001). Taken together, these findings indicating the gene density in G-light bands is higher than that of the G-dark bands.

Keywords: Gene distribution, human chromosome

INTRODUCTION

Human chromosomes are characterized by different staining properties of their regions, namely by chromosomal bands. The most common methods are G, R, Q, T, and C banding. The banding patterns produced by these different techniques are related to each other. The Q banding patterns, produced by treatment with fluorochromes specific for AT-rich DNA regions are similar to G banding patterns, produced by staining with the Giemsa after proteolytic digestion. As a whole, G and Q banding produced a pattern complementary to R banding (Craig and Bickmore, 1993).

The human genome is a mosaic of isochors, long DNA segments, which can be partitioned into five families, namely two GC-poor families (L1 and L2), representatively about 62 % of the genome and three GC-rich families (H1, H2 and H3) (Bernardi, 1995). Gene distribution in chromosome correlated with the GC level of the chromosomal bands (Saccone et al., 1992, 1996; Bernardi et al., 1985; Mouchiroud et al., 1991; Zoubak et al., 1996).

Although it has long been assumed that gene distribution is heterogeneous and that G-dark bands have lower gene content compared to G-light bands, this conclusion was based on about 9,000 genes (Musio et al., 2002). Here we investigated the chromosome distribution of 22845 genes mapped at whole chromosomes reported in the Human Genome Data Base.

EXCLI Journal 2008;7:151-153 – ISSN 1611-2156
Received: April 27, 2008, accepted: July 1, 2008, published: July 7, 2008

METHODS

The data present in the Human Genome Data Base as of January 2007 was used for genes mapped to human chromosomes. Chromosomal lengths were measured as the base pairs of the chromosomes.

The percent of G-dark bands obtained using 800-band resolution. Expected number of genes on each chromosome was calculated based on the relative of the chromosomal length.

Correlations between the variables were determined using parametric Pearson's correlation coefficient analysis. Moreover the partial correlation coefficient analysis was done. To test the null hypothesis that the number of genes mapped on human chromosomes is associated with the chromosomal length, the chi-square test was conducted. Statistical analysis was performed using SPSS (version 13.0). A probability of p-value less than 0.05 was considered statistically significant.

RESULTS AND DISCUSSION

Table 1 shows the chromosomal distribution of 22845 genes reported in Human Genome Data Base and the percent of G-dark bands of the human chromosomes.

Pearson's correlation coefficient analysis showed that there were significant correlation between number of genes mapped to chromosomes and chromosomal length (r=0.656, df=22, p=0.001). There is significant correlation between number of genes mapped and percent of G-dark bands of each chromosome (r=-0.609, df=22, p=0.002). It should be noted that after controlling for chromosome length, partial correlation coefficient show that there is significant negative correlation between number of mapped genes and percent of G-dark bands (r=-0.601, df=22, p=0.002).

Table 1 also shows the ratio of observed versus expected genes on each human chromosome. This ratio is varying from 0.27 to 3.13 for chromosomes Y and 19, respectively. The present study shows significantly higher gene density for chromosomes 11, 12, 14, 16, 17, 19, 20 and 22. Also significantly lower gene density for chromosomes 2, 3, 4, 13, 18, 21 and Y. Previously Musio et al., (2002) reported the same correlation based on the about 8935 mapped genes. Our present data are in agreement with the above mentioned report (r=0.875, df=22, p<0.001) (Musio et al. 2002).

Table1: Chromosomal distribution of 22845 mapped genes

Human chromosomes	Percent of G-dark band	Observed number of genes mapped	Expected number of genes mapped	Observed/ expected ratio
1	47.1	2314	1833	1.26
2	46.9	1485	1801	0.82
3	48.9	1166	1479	0.79
4	53.3	889	1418	0.63
5	50.3	987	1341	0.74
6	50.1	1181	1267	0.93
7	46.2	1065	1177	0.90
8	48.4	795	1084	0.73
9	49.8	999	1040	0.96
10	50.1	878	1004	0.87
11	45.8	1448	997	1.45
12	51.7	1140	981	1.16
13	53.7	391	846	0.46
14	54.1	727	789	0.92
15	48.2	722	744	0.97
16	49.5	966	659	1.47
17	37.3	1293	584	2.21
18	54.4	313	564	0.55
19	50.1	1483	473	3.13
20	45.5	643	463	1.39
21	58.8	294	348	0.84
22	49.3	567	368	1.54
X	47.6	975	1148	0.85
Y	61.7	115	428	0.27

Interestingly, correlation between the ratio of observed/expected mapped genes and percent of G-dark bands is significant (r=-0.506, df=22, p=0.012). There was a significant difference between the observed number of mapped genes and expected number of mapped genes on human chromosomes (χ^2=4842.7, df=23, p<0.00001).

There are several reports concerning the distribution of genes on human chromosomes (Bickmore and Sumner, 1989; De La Torre et al., 1992; Von Kiel et al.,

1985). Oncogenes are scattered throughout the genome, but they tend to cluster at G-light chromosomes bands (Hecht 1988) and most of them have a tendency to be distributed near the telomeres (Lima-de-Faria et al., 1991; Lima-de-Faria and Mitelman 1988). It was reported that integration sites of human immunodeficiency virus type 1 (HIV-1) were distributed at a non-random manner (Saadat et al., 1998). These observations rejected the widely accepted view of genes being distributed randomly. The present findings indirectly indicating that G-light bands showed higher gene density compared with the G-dark bands.

Acknowledgments: This study was supported by Shiraz University.

REFERENCES

Bernardi G. The human genome: organization and evolutionary history. Annu Rev Genet 1995;29:445-76.

Bernardi G, Olofsson B, Filipski J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M, Rodier F. The mosaic genome of warm-blooded vertebrates. Science 1985;228: 953-8.

Bickmore WA, Sumner AT. Mammalian chromosome banding- and expression of genome organization. Trends Genet 1989; 5:144-8.

Craig J, Bickmore WA. Chromosome bands-flavours to savour. Bioessays 1993; 15:349-54.

De La Torre J, Sumner AT, Gosalves J, Stuppia L. The distribution of genes on human chromosomes as studies by in situ nick translation. Genome 1992;35:890-4.

Hecht F. Fragile sites, cancer chromosome breakpoints, and oncogenes all cluster in light G bands. Cancer Genet Cytogenet 1988;31:17-24.

Lima-de-Faria A, Mitelman F. The chromosome territory of human oncogenes. Bioscience Rep 1988;6:349-54.

Lima-de-Faria A, Mitelman F, Blomberg J, Pfeifer-Ohlsson S. Telomeric location of retroviral oncogenes in humans. Hereditas 1991;114:207-11.

Mouchiroud D, D'Onofrio G, Aissani B, Macaya G, Gautier C, Bernardi G. The distribution of genes in the human genome. Gene 1991;100:181-7.

Musio A, Mariani T, Vezzoni P, Frattini A. Heterogeneous gene distribution reflects human genome complexity as detected at the cytogenetic level. Cancer Genetics Cytogenet 2002;134:168-71.

Saadat M, Takahashi H, Bahmani MK, Ikuta K, Yoshida MC. Does human immunodeficiency virus type 1 integrate in chromosomes randomly? Chromosome Sci 1998;2:5-8.

Saccone S, De Sario A, Della Valle G, Bernardi G. The highest gene concentrations in the human genome are in T-bands of metaphase chromosomes. Proc Natl Acad Sci USA 1992;89:4913-7.

Saccone S, Caccio S, Kusuda J, Andreozzi L, Bernardi G. Identification of the gene-richest bands in human chromosomes. Gene 1996;174:85-94.

Von Kiel K, Hameister H, Somssich IE, Adolph S. Early replication banding reveals a strongly conserved functional pattern in mammalian chromosomes. Chromosoma 1985;93:69-76.

Zoubak S, Clay O, Bernardi G. The gene distribution of the human genome. Gene 1996;174:95-102.