World Academy of Science, Engineering and Technology
International Journal of Bioengineering and Life Sciences
Vol:3, No:4, 2009

Statistics of Exon Lengths in Animals, Plants, Fungi, and Protists

Alexander Kaplunovsky, Vladimir Khailenko, Alexander Bolshoy, Shara Atambayeva, and Anatoliy Ivashchenko

***Abstract*—**Eukaryotic protein-coding genes are interrupted by spliceosomal introns, which are removed from the RNA transcripts before translation into a protein. The exon-intron structures of different eukaryotic species are quite different from each other, and the evolution of such structures raises many questions. We try to address some of these questions using statistical analysis of whole genomes. We go through all the protein-coding genes in a genome and study correlations between the net length of all the exons in a gene, the number of the exons, and the average length of an exon. We also take average values of these features for each chromosome and study correlations between those averages on the chromosomal level. Our data show universal features of exon-intron structures common to animals, plants, and protists (specifically, *Arabidopsis thaliana, Caenorhabditis elegans, Drosophila melanogaster, Cryptococcus neoformans, Homo sapiens, Mus musculus, Oryza sativa,* and *Plasmodium falciparum).* We have verified linear correlation between the number of exons in a gene and the length of a protein coded by the gene, while the protein length increases in proportion to the number of exons. On the other hand, the average length of an exon always decreases with the number of exons. Finally, chromosome clustering based on average chromosome properties and parameters of linear regression between the number of exons in a gene and the net length of those exons demonstrates that these average chromosome properties are genome-specific features.

***Keywords*—**Comparative genomics, exon-intron structure, eukaryotic clustering, linear regression.

Abbreviations—N_{ex} = number of exons in a gene; L_{ex} = net length of all exons in a gene; A_{ex} = average exon length in a gene; n_{ex} = average (over a chromosome) number of exons in a gene; l_{ex} = average (over a chromosome) net length of all exons in a gene; a_{ex} = average (over a chromosome) of the average exon length in a gene.

A. Kaplunovsky is with the Department of Evolutionary and Environmental Biology and Genome Diversity Center at the Institute of Evolution, University of Haifa, Israel and with the Department of Sciences, Holon Institute of Technology, Holon, Israel.

V. Khailenko is with the Department of Biotechnology, Biochemistry, and Plant Physiology at the Al-Farabi Kazakh National University, Kazakhstan.

A. Bolshoy is with the Department of Evolutionary and Environmental Biology and Genome Diversity Center at the Institute of Evolution, University of Haifa, Israel (corresponding author, phone/fax: +972-48240382; e-mail: bolshoy@research.haifa.ac.il).

S. Atambayeva is with the Department of Biotechnology, Biochemistry, and Plant Physiology at the Al-Farabi Kazakh National University, Kazakhstan.

A. Ivashchenko is with the Department of Biotechnology, Biochemistry, and Plant Physiology at the Al-Farabi Kazakh National University, Kazakhstan (corresponding author, phone/fax: +77272 490 164; e-mail: a_ivashchenko@mail.ru).

I. Introduction

ONE of the greatest enigmas of eukaryotic genome evolution is the widespread existence of introns. The introns have been detected in genes of both lower and higher eukaryotes, and also of their viruses, chloroplasts and mitochondria. There are several types of introns, and this study focuses on the most important type: the spliceosomal introns of nuclear-encoded protein genes. We study properties of exon-intron structure of these genes in selected eukaryotic genomes.

A putative link between the biological role of introns and the distribution of exon sizes in protein-coding genes was established soon after intron discovery [1]. Since then many studies – including statistical analysis – of the exon-intron structures of higher and lower eukaryote genes were performed [2-9]. The problem of intron length variability has a long history [8, 9], and it remains unsolved. We still do not know why intron lengths are so widely variable, both between different organisms and between different genes of the same organism.

Likewise, we do not understand the distribution of the intron densities (average numbers of introns per gene). At first, the intron density was thought to be related to the organismal complexity. The initial studies supported this hypothesis: *Homo sapiens* has 8.1 introns per gene in average [10], *Caenorhabditis elegans* – 4.7 [11], *Drosophila melanogaster* – 3.4 [12], and *Arabidopsis thaliana* – 4.4 [13]; while, by contrast, unicellular species were found to have less introns per gene [14]. However, further studies found pretty high intron densities in many single-celled species [15, 16], and intron densities in basidiomycete and zygomycete fungi are among the highest known among eukaryotes (4-6 per gene) [17, 18].

In this article, we focus on the exon features rather than those of introns. We study relations between the exon lengths, the protein lengths, the average exon sizes, and the numbers of exons per gene (exon densities). There is an interesting observation regarding distributions of exon lengths in different eukaryotes: exon sizes follow a lognormal distribution typical of a random Kolmogorov fractioning process [19, 20]. The evolutionary mechanisms of exon-intron structure formation are rather controversial. A theory suggesting that introns appeared as a result of insertion of transposons [21, 22] is currently quite popular. Frequently,

this point of view implicitly assumes that longer genes possess a higher probability of splitting since they are larger targets for transposons. Some of the present authors have showed [9] that the exon–intron organizations in *Arabidopsis thaliana,* in *Caenorhabditis elegans*, and in *Homo sapiens* have much in common. In particular, the net length of all exons in a gene correlates with the number of exons, while the average length of an exon decreases: there are fewer long exons (over 400 nucleotides) and more short exons (80 to 140 nucleotides). This observation seems to support the transposon hypothesis: longer exons appear as larger targets for insertion of mobile elements. Gudlaugsdottir *et al*. [19] found some arguments supporting both the intron-early theory [23] and the intron-late theory [24, 25] and proposed a mixture model. There is still much controversy and research on newly sequenced genomes should be continued. Here, we apply our efforts mainly for better visualization of new and old results, and application of clustering techniques to strengthen specific genomic properties of common exon-intron organization.

To avoid possible misunderstandings, we would like to clarify our terminology. By "gene" we mean a sequence of DNA nucleotides, which occupies a specific location along a chromosome and determines a particular characteristic in an organism. The structure of a typical protein-coding gene consists of a promoter, a transcription initiation site, a coding region including exons and introns, the polyadenylation signal, and a termination site. Exons are gene fragments that are transcribed in the functional mRNA. All coding sequences are either internal exons or parts of the first or the last exon, while there are non-coding exons, or partially non-coding exons. Introns are non-coding sequences. Some eukaryotic genes have no introns (intronless genes). There are structurally simple genes (two exons separated by one intervening sequence), and there are extremely complex genes whereby a very large number of exons form the final mRNA. For instance, the dystrophin gene comprises at least 70 exons and its length is more than one million base pairs of DNA.

II. Data and Methods

Nucleotide sequences of 76 chromosomes of 8 species (Table I) containing 5 chromosomes of *Arabidopsis thaliana* (AD), 6 chromosomes of *Caenorhabditis elegans* (CE), 5 chromosomes of *Drosophila melanogaster* (DM), 14 chromosomes of *Cryptococcus neoformans* (CN), 10 chromosomes of *Homo sapiens* (HS), 10 chromosomes of *Mus musculus* (M), 12 chromosomes of *Oryza sativa* (OS), and 14 chromosomes of *Plasmodium falciparum* (PF) have been obtained from GenBank http://www.ncbi.nlm.nih.gov.

Each gene was assigned 3 numbers: the net length L_{ex} of all its exons, the number N_{ex} of those exons, and an average exon length

$$A_{ex} = \frac{L_{ex}}{N_{ex}}$$

Linear regression for the number of exons in a gene as a function of the gene's net exon length $N_{ex}=a+b\cdot L_{ex}$ was performed using the program SPSS for every chromosome. For every chromosome, we also calculated the average net length l_{ex} of all the exons in a gene, the average number n_{ex} of such exons, and the average exon length a_{ex} - which is the mean of the A_{ex} values of individual genes,

$$a_{ex} = \frac{1}{n}\sum_{i=1}^{n} A_{ex},$$

where n is a number of genes in the chromosome. Note that the a_{ex} is different from the average length $\bar{a}_{ex}$ of all the exons in the chromosome, regardless of which gene(s) they belong to. (The $\bar{a}_{ex}$, is calculated as a total length of all exons in a chromosome divided by a total number of all exons in a chromosome, see [26]). The a_{ex} usually have significantly larger values than the $\bar{a}_{ex}$ because an average length of i-th exon exponentially decreases with an i (see [19]).

We also considered regression parameters a and b and a parameter of explained variation R^2. These data are compiled in Supplementary Material. Distance between each pair of chromosomes has based on these six parameters standardized in the interval $[-1 \div +1]$, and was calculated by SPSS as a Euclidean distance in a six-dimension space.

A matrix of distances for all 76 chromosomes was exported to the program *Neighbor* of *Phylip Package* (the University of Washington) http://evolution.genetics.washington.edu/phylip/doc/neighbor.html using Neighbor Joining Algorithm. Output file was viewed and drawn by the program *TreeView* of Prof. Rod Page http://taxonomy.zoology.gla.ac.uk/rod/treeview.html.

TABLE I
List of Processed Species and their Chromosomes

N	Name of the organism	Kingdom	Number of chromosomes	Processed chromosomes
1	*Arabidopsis thaliana*	Plant	5	1-5
2	*Caenorhabditis elegans*	Animal	6	1-6
3	*Cryptococcus neoformans*	Fungi	14	1-14
4	*Drosophila melanogaster*	Animal	4+X	2L,2R,3L,3R,X
5	*Homo sapiens*	Animal	22+XY	1-10
6	*Mus musculus*	Animal	19+XY	1-10
7	*Oryza sativa*	Plant	12	1-12
8	*Plasmodium falciparum*	Protists	14	1-14

III. Results and Discussion

A. Average Numbers of Exons and Net Exon Lengths in Different Chromosomes

For each of 76 chromosomes of eight species, we have calculated the average parameters l_{ex} (net length of gene's exons), n_{ex} (number of exons in a gene) and a_{ex} (average exon length). These averages turned out to be pretty similar for different chromosomes of the same species but rather distant for different species. Fig. 1 presents a scatter plot of the l_{ex} *vs* n_{ex}; it shows clear clustering of the chromosomes by species.

It also shows a wide separation between PF – a protist – and the other species (animals, fungi, and plants). The PF chromosomes have much longer average proteins (l_{ex}) and much lower exon density (n_{ex}) than all the other eukaryote chromosomes we have studied. Moreover, all species except PF have rather similar ranges of the l_{ex} parameter, but the n_{ex} fall into quite distinct regions on the plot for the DM (*D. melanogaster*) and CN, and more doubtful areas for plants (AD and OS) and mammals (*H. sapiens* and *M. musculus*).

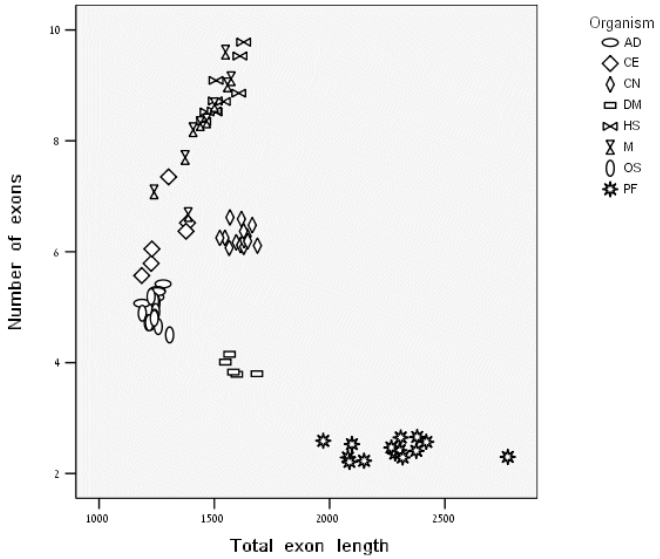

Fig. 1 Scatter-plot of the average net exon length per gene l_{ex} (x-axis) *vs* the average number of exons per gene n_{ex} (y-axis), for all 76 processed chromosomes of eight species

Fig. 2 is a scatter-plot of the average exon length a_{ex} *vs* the average number of exons in a gene n_{ex}; the outlier chromosomes of PF are not shown. This plot shows much better grouping of chromosomes belonging to the same species than Fig. 1 – all kingdoms are grouped separately. Still, the resolution is not sufficient and there is a slight overlapping between species from the same kingdom (M and HS, AD and OS). In addition, *C. elegans* chromosomes may be characterized by relatively short exons in average and rather big variation in intron density. To improve the resolution between the species, we are going to take a closer look at the relation between the average exon number and the average net exon length of a gene.

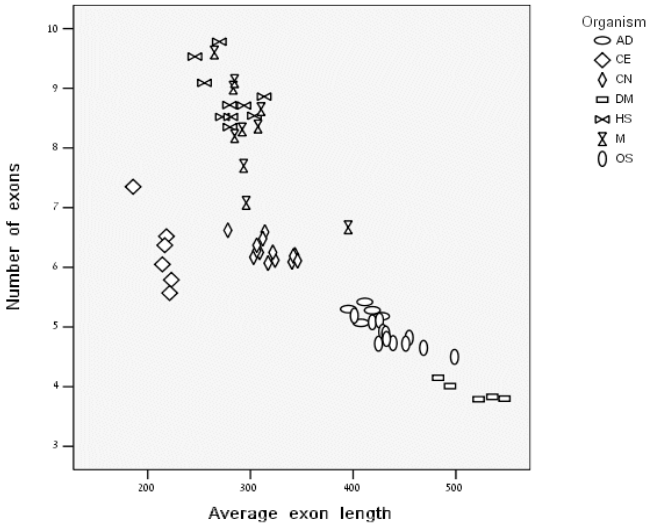

Fig. 2 Scatter-plot of the average exon length a_{ex} (x-axis) *vs* the number of exons n_{ex} (y-axis), for 62 processed chromosomes of seven species

B. Relations between the Average Exon Number and the Average Net Length of Exons in a Gene

It was already shown [8] that the average exon length in *A. thaliana*, *O. sativa*, *C. elegans*, and *Homo sapiens* genes decreases with an increasing number of introns. In addition, positive linear correlation was observed between the sum of exon lengths and the number of exons [8]. Fig. 3 shows the relation between the net length of exons and the number of exons in 12156 genes on ten chromosomes of *H. sapiens*. Parameters of linear regression $N_{ex}=a+b\cdot L_{ex}$ are a=1.118 and b = 0.005028. Explained variation of the regression $R^2=0.666$, significance $p< 0.001$.

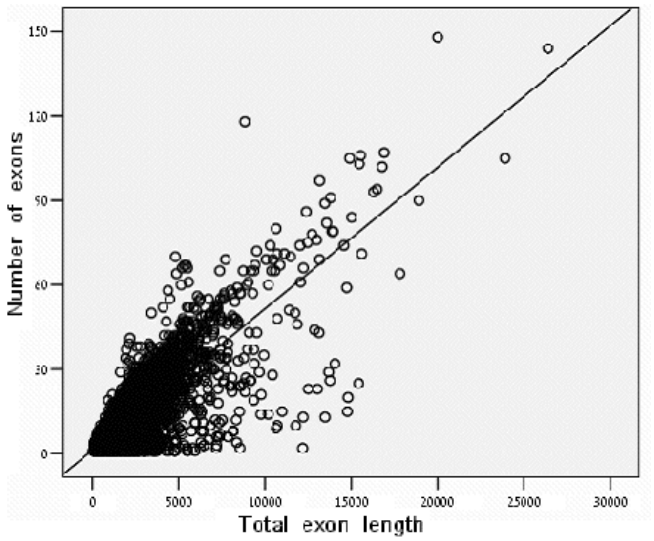

Fig. 3 Linear regression between the net length of exons of a gene (L_{ex}, x-axis) and the number of exons (N_{ex}, y-axis) in genes on all processed chromosomes of *H. sapiens*

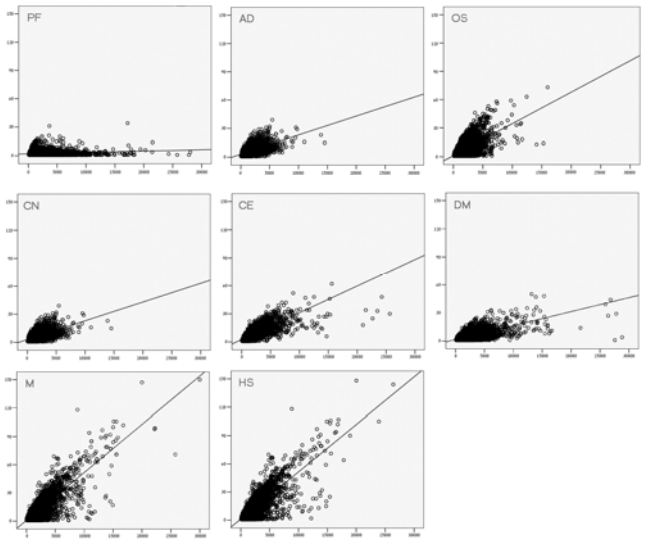

Fig. 4 Scatter plots of L_{ex} (x-axes) *vs* N_{ex} (y-axes) and lines of linear regression for chromosomes of *P. falciparum* (PF), *A. thaliana* (AD), *O. sativa* (OS), *C. neoformans* (CN), *C. elegans* (CE), *D. melanogaster*, *M. musculus* (M), and *H. sapiens* (HS)

Fig. 4 presents similar plots for all eight species. Each species is represented by a scatter-plot of L_{ex} *vs* N_{ex} with a linear regression. There are dramatic differences between average and maximal values of L_{ex} and N_{ex} for animals, plants, fungi, and protists, and especially between parameters a and b of the linear regression equation $y=a+bx$. In light of these differences, we decided to check if the regression parameters could be used in classification of genomes by their exon properties. We have calculated the linear regressions for all 76 processed chromosomes of all eight genomes. Our results show significant correlations between the protein lengths and the numbers of exons in all eight studied genomes. The Supplementary Material tabulates the parameters a and b of the linear regression $N_{ex}=a+b\cdot L_{ex}$; their values testify to high reliability of the correlation.

Fig. 5 presents the scatter-plot of parameters a and b of the linear regression $N_{ex}=a+b\cdot L_{ex}$ for all the processed chromosomes. One can recognize five clusters in the figure: (i) PF, (ii) DM, (iii) plants (AD + OS), (iv) mammals (M + HS), and (v) CE + CN. This means that clustering based on the linear regression parameters a and b follows the major differences between species from different kingdoms, and some reasonably observable differences between species from the same kingdom. There are some exceptions, and we would like to eliminate them by using the R^2 parameter - percent of the explained variation - of the regression analysis. It has negligible value for protists, medium values for plants and fungi, and relatively high values for animals.

Fig. 6 presents scatter plots for a *vs* R^2 (left) and b *vs* R^2 (right). It shows slightly improved resolution between the species: the CE and the CN chromosomes now belong to separate clusters, while the AD and the OS are almost (but not quite) separate. Hopefully, combining all the parameters together would give a better resolution than looking at any two parameters at a time.

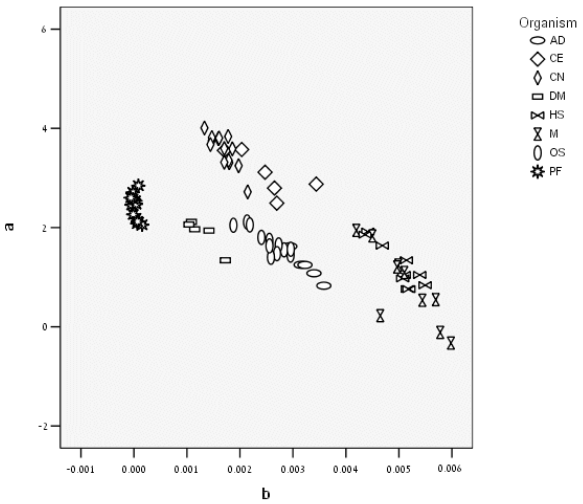

Fig. 5 Scatter-plot of parameters a (y-axis) and b (x-axis) of the linear regression $N_{ex}=a+b\cdot L_{ex}$

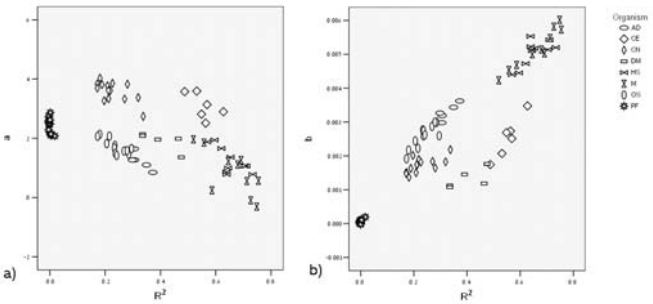

Fig. 6 Scatter-plots of parameters a, b, and R^2 of linear regression $N_{ex}=a+b\cdot L_{ex}$ for all the processed chromosomes. Left plot: a (y-axis) vs. R^2; right plot: b (y-axis) vs. R^2

B. Dendrogram of Chromosomes of All the Genomes

Our next goal is to visualize the chromosome classification using all of the parameters: n_{ex}, a_{ex}, a, b, and R^2 we have calculated (see Supplementary Material for the complete table of their values). We standardize each of the parameters to the interval $[-1 \div +1]$, and then calculate the Euclidean distances in six-dimensional parameter space between all pairs of chromosomes i and j according to

$$d_{ij}^2 = \left(n'_{i,ex} - n'_{j,ex}\right)^2 + \left(l'_{i,ex} - n'_{j,ex}\right)^2 + \left(l'_{i,ex} - n'_{j,ex}\right)^2 + \left(a'_i - a'_j\right)^2 + \left(b'_i - b'_j\right)^2 + \left(\left(R^2\right)_i - \left(R^2\right)_j\right)^2,$$

where $n'_{i,ex}$, $l'_{i,ex}$, $a'_{i,ex}$, a'_i, b'_i, and $R^2{}'_i$ are the standardized parameters of the chromosome i. Having calculated the distance matrix d_{ij}, we used the *Neighbor Joining Algorithm* to obtain the dendrogram of our chromosomes. The chromosomes of one species were grouped together but separately from other species. There is only one exception: the chromosomes of the two mammal species *M. musculus* and *H. sapiens* form a single mixed branch (Fig. 7).

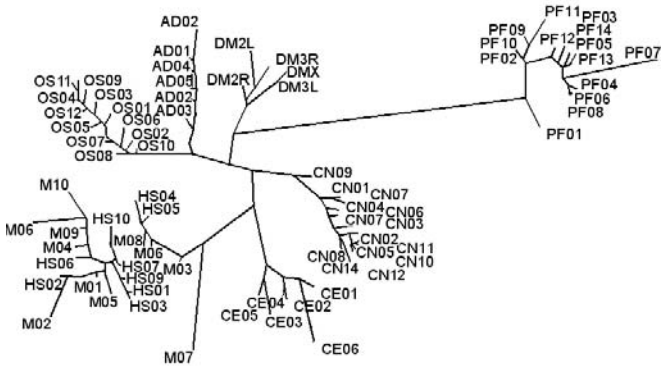

Fig. 7 Dendrogram of the 76 processed chromosomes of eight species based on weighted distances among parameters n_{ex}, l_{ex}, a_{ex}, a, b, and R^2 (a, b, and R^2 are parameters of the linear regression $N_{ex}=a+b\cdot L_{ex}$)

IV. Conclusion

Our results show both general and genome-specific features of the exon-intron organization of eukaryotic genes. The most general feature found in all genomes is the positive correlation between the number of introns in a gene and the corresponding protein's length (and equivalently, the net length of all the exons of the gene). In addition, in all the genomes we have studied, the average exon length in a gene decreases with the number of those exons. By while these laws of exon-intron statistics are quite general, the correlation parameters are genome-specific. For the first time, for our best knowledge, it was shown that they are specific to genomes rather than to individual chromosomes. Indeed, in the parameter space of average chromosome properties and linear regression parameters (between exon numbers and protein lengths), all chromosomes from the same genome form obvious clusters.

Clearly, the exon-intron structures of eukaryotic genes have many important parameters that we did not consider in this work; we have left them for the future research. The main goals of this article are to draw attention to the statistical properties of exon size distributions, and to visualize both the general laws of exon-intron organizations of genes and the genome-specific features.

Acknowledgment

A. K. thanks S. Hosid for the technical support.

References

[1] H. Naora and N.J. Deacon, "Relationship between the total size of exon and introns in the protein-coding genes of higher eukaryotes," *Proc. Natl. Acad. Sci.. USA*, vol. 79: pp. 6196–6200, 1982.

[2] J.D. Hawkins, "A survey on intron and exon lengths," *Nucleic Acids Res*., vol. 16: pp. 9893–9908, 1988.

[3] M. Deutsch and M. Long, "Intron-exon structures of eukaryotic model organisms," *Nucleic Acids Res*., vol. 27: p. 3219–3228, 1999.

[4] E.V. Kriventseva and M.S. Gelfand, "Statistical analysis of the exon-intron structure of higher and lower eukaryote genes," *J. Biomol. Struct. Dyn.*, vol. 17, pp. 281–288, 1999.

[5] A.A. Mironov and M.S. Gelfand, "Prediction and computer analysis of the exon-intron structure of human genes," *Mol. Biol*., vol. 38, pp. 70–77, 2004.

[6] A.T. Ivashchenko and S.A. Atambayeva, "Variation in lengths of introns and exons in genes of the Arabidopsis thaliana nuclear genome," *Russian Journal of Genetics*, vol. 40, pp. 1179–1181, 2004.

[7] S.W. Roy and D. Penny, "Intron length distributions and gene prediction," *Nucleic Acids Res*., vol. 35, pp. 4737–4742, 2007.

[8] S.A. Atambayeva, V.A. Khailenko, and A.T. Ivashchenko, "Intron and exon length variation in arabidopsis, rice, nematode, and human," *Mol. Biol*., vol. 42, pp. 312–320, 2008.

[9] A.T Ivashchenko,. V.A. Khailenko, and S.A. Atambayeva, "Variation of the lengths of exons and introns in Human Genome genes," *Russian Journal of Genetics*, vol. 45, pp.16–22, 2009.

[10] F.S. Collins *et al*., "Finishing the euchromatic sequence of the human genome. International Human Genome Sequencing Consortium," *Nature*, vol. 431, pp. 931–945, 2004.

[11] E.M. Schwarz *et al*., "WormBase: better software, richer content," *Nucleic Acids Res*., vol. 34 (Database), pp. D475–D478, 2006.

[12] R.A. Drysdale and M.A. Crosby, "FlyBase: genes and gene models," *Nucleic Acids Res*., vol. 33 (Database), pp. D390–D395, 2005.

[13] B.J. Haas *et al*., "Complete reannotation of the Arabidopsis genome: methods, tools, protocols and the final release," *BMC Biol*., vol. 3, p. 7, 2005.

[14] J.M.J. Logsdon, A. Stoltzfus, and W.F. Doolittle, "Molecular evolution: recent cases of spliceosomal intron gain?" *Curr. Biol*., vol. 8: pp. R560–R563, 1998.

[15] J.M Archibald,. C.J. O'Kelly, and W.F. Doolittle, "The chaperonin genes of jakobid and jakobid-like flagellates: implications for eukaryotic evolution," *Mol. Biol. Evol*., vol. 19, pp. 422–431, 2002.

[16] A.T Ivashchenko, M.I. Tauasarova, and S.A. Atambayeva, "Exon–Intron Structure of Genes in Complete Fungal Genomes," *Mol. Biol*., vol. 43, pp. 24–31, 2009.

[17] B.J. Loftus *et al*., "The genome of the basidiomycetous yeast and human pathogen Cryptococcus neoformans," *Science*, vol. 307, pp. 1321–1324, 2005.

[18] D. Martinez *et al*., "Genome sequence of the lignocellulose degrading fungus Phanerochaete chrysosporium strain RP78," *Nat. Biotechnol*., vol. 22, pp. 695–700, 2004.

[19] S. Gudlaugsdottir, D.R. Boswell, G.R. Wood, and J. Ma, "Exon size distribution and the origin of introns," *Genetica*, vol. 131, pp. 299–306, 2007.

[20] Y. Ryabov and M. Gribskov, "Spontaneous symmetry breaking in genome evolution," *Nucleic Acids Res*., vol. 36, pp. 2756–2763, 2008.

[21] G. Cho and R.F. Doolittle, "Intron distribution in ancient paralogs supports random insertion and not random loss," *J. Mol. Evol*., vol. 44, pp. 573–584, 1997.

[22] S.W. Roy, "The origin of recent introns: transposons?" *Genome Biol*., vol. 5, p. 251, 2004.

[23] W. Gilbert, "The exon theory of genes," in *Symp. Quant. Biol*., Cold Spring Harbor, vol.52, 1987, pp.901–905.

[24] T. Cavalier-Smith, "Selfish DNA and the origin of introns," *Nature*, vol. 315, pp. 283–284, 1985.

[25] J.M. Logsdon and J.D. Palmer, "Origin of introns – early or late?" *Nature,* vol. 369, pp. 526–528, 1994.

[26] M.K. Sakharkar, V.T. Chow, and P. Kangueane, "Distributions of exons and introns in the human genome," *In Silico Biol*., vol. 4, pp. 387–393, 2004.

SUPPLEMENTARY MATERIAL

AVERAGE CHROMOSOME CHARACTERISTICS AND REGRESSION PARAMETERS OBTAINED FOR 76 PROCESSED CHROMOSOMES

Chromo-some	l_{ex}	n_{ex}	a_{ex}	$N_{ex}=a+b\cdot L_{ex}$		
				a	$b\cdot 10^{-3}$	R^2
PF01	1972	2.59	1138	2.54	.03	.000
PF02	2079	2.29	1384	2.11	.09	.008
PF03	2308	2.65	1581	2.85	.08	.006
PF04	2375	2.41	1583	2.48	.03	.001
PF05	2284	2.36	1596	2.46	-.05	.005
PF06	2419	2.57	1578	2.47	.04	.002
PF07	2772	2.30	1835	2.19	.04	.003
PF08	2379	2.66	1532	2.71	-.02	.000
PF09	2096	2.53	1343	2.57	-.02	.000
PF10	2085	2.21	1357	2.08	.06	.007
PF11	2149	2.23	1472	2.27	-.01	.000
PF12	2303	2.42	1529	2.06	.16	.022
PF13	2267	2.47	1526	2.60	-.06	.005
PF14	2316	2.29	1543	2.12	.07	.008
AD01	1277	5.42	411	1.08	3.40	.354
AD02	1185	5.07	407	.83	3.58	.377
AD03	1246	5.18	427	1.25	3.15	.311
AD04	1252	5.30	395	1.62	2.94	.310
AD05	1252	5.28	418	1.25	3.23	.303
OS01	1246	5.08	418	1.55	2.83	.273
OS02	1241	5.11	425	1.45	2.96	.290
OS03	1225	5.19	401	1.56	2.96	.284
OS04	1245	4.92	428	1.74	2.56	.240
OS05	1187	4.89	431	1.65	2.73	.238
OS06	1241	4.82	454	1.47	2.70	.241
OS07	1213	4.72	451	1.80	2.41	.210
OS08	1225	4.73	438	2.11	2.13	.185
OS09	1218	4.72	424	2.05	2.19	.220
OS10	1256	4.65	468	1.40	2.59	.246
OS11	1305	4.50	498	2.05	1.88	.177
OS12	1239	4.80	432	1.63	2.56	.300
CN01	1594	6.17	303	3.31	1.80	.275
CN02	1613	6.12	323	3.75	1.47	.213
CN03	1545	6.25	308	3.79	1.59	.193
CN04	1642	6.28	306	3.35	1.79	.324
CN05	1617	6.59	314	3.59	1.86	.223
CN06	1664	6.48	312	3.81	1.61	.285
CN07	1627	6.09	340	3.32	1.70	.217
CN08	1628	6.21	343	3.81	1.47	.176
CN09	1564	6.07	317	2.72	2.15	.342
CN10	1644	6.19	341	4.01	1.33	.186
CN11	1686	6.11	345	3.67	1.44	.176
CN12	1523	6.25	321	3.25	1.98	.201
CN13	1567	6.62	277	3.84	1.78	.231
CN14	1626	6.37	306	3.59	1.70	.216
DM2L	1597	3.79	522	1.97	1.15	.469
DM2R	1565	4.15	482	1.94	1.42	.395
DM3L	1582	3.83	535	2.12	1.09	.340
DM3R	1547	4.01	494	1.34	1.72	.480
DMX	1684	3.80	547	2.06	1.04	.340
CE01	1383	6.52	218	3.12	2.47	.573
CE02	1225	5.79	222	2.49	2.70	.568
CE03	1377	6.37	216	3.57	2.03	.536
CE04	1229	6.05	214	2.80	2.65	.553
CE05	1185	5.57	221	3.56	1.71	.492
CE06	1301	7.35	185	2.88	3.44	.632
M01	1557	9.01	283	1.24	4.99	.696
M02	1548	9.60	264	.55	5.70	.759
M03	1373	7.70	293	1.94	4.20	.524
M04	1407	8.20	284	.54	5.44	.717
M05	1572	9.12	284	1.09	5.10	.684
M06	1238	7.08	295	-.33	5.98	.753
M07	1386	6.67	395	.22	4.65	.591
M08	1439	8.31	291	1.83	4.50	.562
M09	1500	8.65	310	1.21	4.97	.650
M10	1466	8.36	307	-.11	5.78	.731
HS01	1504	8.72	279	1.05	5.10	.708
HS02	1611	9.53	245	1.05	5.39	.715
HS03	1627	9.78	269	.84	5.50	.643
HS04	1538	8.71	293	1.92	4.41	.600
HS05	1605	8.86	313	1.86	4.36	.573
HS06	1503	8.52	272	.76	5.16	.738
HS07	1468	8.52	280	1.64	4.69	.625
HS08	1453	8.35	280	.98	5.07	.649
HS09	1499	8.54	303	.76	5.19	.644
HS10	1507	9.09	255	1.34	5.14	.658