Original Article

The study of gene expression statistics

Vyacheslav V. Saenko[1]

[1]PhD, Professor, Senior Researcher at S.P. Kapitsa Research Technological University, Ulyanovsk State University

Email: mr.fominan@yandex.ru

Abstract

The paper studies the issue of the law of the gene expression distribution obtained with the use of the next-generation sequencing technology (NGS). It has been shown that gene expression has the form of a shift-scale mixture of distributions. One of the components of this mixture is fractionally stable distribution with characteristic indicators varying within the range $0.91 \leqslant \alpha \leqslant 1.24$, $0.17 \leqslant \beta \leqslant 0.75$. This component describes the distribution of gene expression at values $FPKM > 1$. The use of Fisher's goodness-of-fit test χ^2 does not reject the the hypothesis of fractionally stable distribution. Another component of the mixture appears at expression values $FPKM < 1$ and may be associated with errors in the sequencing process using the technology of NGS.

Keywords: Gene expression, next-generation sequencing technology, fractionally stable distribution, sequencing data processing.

1. Introduction

Gene expression studies carried out over the past two decades have shown the universal nature of these distributions, that the distribution of gene expression levels has a power-law character and, the Zipf-Pareto distribution is well described at high expression levels.

$$p(x) \propto x^{-\alpha-1}, x \to \infty. \quad (1)$$

This behavior is typical for the expressions of various organisms and tissues and was obtained both when using the DNA microarray technology [1, 2, 3, 4, 5, 6], and with the application of next-generation sequencing (NGS) technology [7, 8].

The studies of the results of gene expression obtained using DNA microarrays, carried out in the work [1] for human, mouse and yeast tissue samples, as well as samples of cancer and healthy human cells showed that the best approximation among all examined was the discrete Pareto distribution p(m)=(m+b)^(-α-1)/z, where the value α changes in the range from 0.974 to 1.88.

Similar studies were conducted in the works [2, 3, 9]. In these works, the authors also drew the conclusion that the distribution of gene expression was well described by the law (1) with a power value α∈[0.86,1] in [2], α∈[0.69,1.09] in [3] and α∈[0.7,1.0] in the work [9]. The work [4] is devoted to the study of the type of distribution. The authors of this work also came to the conclusion that the best results among all the studied distributions were obtained by using the double Pareto-lognormal distribution.

Similar behavior is shown by the study of gene expression obtained with the use of NGS. The question of the type of statistical distribution of transcriptomic data was studied in the paper [7]. In this work, the authors conclude that the distribution of gene expression is described by the distribution (1). Moreover, on different scales, different power relationships are observed.

Address for correspondence: Vyacheslav V. Saenko
S.P. Kapitsa Research Technological University, Ulyanovsk State University
Email: mr.fominan@yandex.ru

Access this article online

Quick Response Code:

Website: www.pnrjournal.com

DOI: 10.47750/pnr.2022.13.03.130

How to cite this article: Vyacheslav V. Saenko, The study of gene expression statistics, J PHARM NEGATIVE RESULTS 2022;13: 869-874.

The article [8] is devoted to the study of the relationship between gene expression and the size of the cell population. It shows that regardless of the size of the cell population, the expression of all genes obeys the Pareto distribution.

Another universal property of the gene expression distribution is that the distribution of the logarithm of gene expression can be described by shift-scale mixtures of normal distributions. This fact is confirmed by the results of the works [9, 10, 11, 12]. At the same time, the number of components in this mixture of distributions can be different. In the work [9] it was found that the number of components was three. In the work [11] the authors drew the conclusion that the number of components of a mixture was two, and in the work [10] it is shown that the number of components can vary from 1 to 5.

The issue of the type of distribution is of interest in connection with the need to normalize experimental data on gene expression. The need for normalization arises due to the fact that the initial tissue samples can be prepared under different conditions, and the experimental data themselves can be taken from various sources. In addition, various errors can occur during the sequencing process. For example, the substitution error in NGS is one of the typical mistakes. It plays an important role in detecting low-frequency genetic variants and can be corrected both experimentally and through calculation [13]. In addition, weakly expressing genes can be nullified as a result of restrictions on the length of the sequence to be sequenced and the level of sensitivity of the technology used. Therefore, the use of an appropriate data normalization method can minimize such a kind of errors and allow making a comparison between the results of gene expression obtained in different experiments. Using the assumption about the power-law distribution of gene expression in the works [14, 15] the appropriate methods of normalization were developed.

However, the Zipf-Pareto distribution (1) is somewhat artificial. It is known that the Zipf distribution, which is discrete, appears in the problem of studying the distribution of the frequency of words in a text. This distribution describes the frequency of occurrence of words in the text depending on the rank of the word. Apart from this, the distribution (1) is not a limited distribution. In this regard, in our opinion, the most suitable class of distributions to describe the distribution of gene expression is the class of fractionally stable distributions [16, 17]. The question of the possibility of using stable and fractionally stable distributions to describe the distribution of gene expression obtained with the help of the DNA microarray technology has already been studied in the works [5, 6]. Although it was shown in these works that Fisher's Chi-square test rejected the the hypothesis of a stable and fractionally stable character of experimental distributions, but the theoretical distributions themselves well described the structure and shape of the experimental distributions. Another fact that testifies in favor of using fractionally stable distributions rather than the Pareto distribution is that these distributions are limiting distributions for the sum of random variables and have asymptotics of the form (1) at $x \to \infty$. All this makes distributions from this class good candidates for approximating gene expression profiles.

However, there are several factors that make it difficult to work with this class of distributions. The main difficulty here is the lack of explicit expressions for the distribution densities. Therefore, to calculate the densities, one must resort to numerical methods, in particular, to the Monte Carlo method, which increases the calculation efforts and is not always convenient. Nevertheless, methods for calculating densities have been developed and tested, and this allows them to be used to solve various problems. This paper discusses the issue of the possibility of using fractional-stable distributions to describe the distribution of gene expression obtained with the use of the NGS technology.

2. Results of using Next-Generation Sequencing

An open database Gene Expression Omnibus (http://www.ncbi.nlm.nih.gov/geo/) was used as a source of experimental data on NGS. An open database was used as a source of experimental data on NGS. Several experimental series were selected from this database which were obtained using NGS, the expression in which is given in units FPKM. In particular, the experimental data were processed for human tissues (series: GSE44875 and GSE50760), macaque tissues (series GSE53690), mouse tissues (series: GSE53690 and GSE53110) and drosophila ones (series: GSE54600 and GSE41487). The processing results and processing technique are given below. It was found that the experimental distributions were fractionally stable in the range of values 1÷1000 units. FPKM. Outside this range, a deviation of the behavior of the experimental distribution from the density of the fractionally stable law is observed.

The experimental data processing technique is as follows. The entire range of expression change values was divided into three non-overlapping intervals: $R=\{x:E\_min \leqslant x \leqslant E\_max\}$, $R\_1=\{x:x \leqslant E\_min\}$ and $R\_2=\{x:x \geqslant E\_max\}$. It was visually established that the range of values corresponding to the domain R, empirical density has behavior characteristic of a fractionally stable law. Therefore, to approximate the empirical density inside the domain R the expression values $X \in R$ were selected from the experimental data. The sample obtained in this way was considered as a sample of independent identically distributed random variables $X\_i,\ldots,X\_n$. Next, on the basis of the sample obtained $X\_i,i=1,\ldots,n$ the parameters of the fractionally stable distribution are estimated and the hypothesisH_0 is put forward. To estimate the parameters of fractional-stable distributions statistically, we used the minimum distance method described in the work [18]. To test this the hypothesis, Fisher's goodness-of-fit test χ^2 was used.

The results of approximation of the experimental gene

expression densities within the domain R for human, mouse, macaque, and drosophila tissues are shown in Fig. 1. For the processed experimental data, it was found that the boundaries of the domain R are determined by the values E_min≳1,E_max≲1000. It is clearly seen from the figures presented that fractionally stable densities describe well the experimental distributions in this domain. Fisher's test χ^2 also confirms this conclusion. For the data shown in the figures, this criterion does not reject the hypothesis H_0 at the level of significance 1%. More complete results of the parameter estimation and verification of the hypothesis H_0 are given in table 5 (see Appendix 5). The column H_0 contains the results of Fisher's test χ^2 (0 – the test rejects the hypothesis H_0, 1 – the test does not reject the hypothesis H_0). From the table it becomes clear that practically for all processed data the hypothesis H_0 is not rejected. Knowing the boundaries of the domain R one can calculate the probability that is contained inside: p=P{E_min⩽X⩽E_max}=(1/N)∑_(i=1)^N▒ I(E_min⩽X_i ⩽E_max), where I(A) is the event indicator A, N is the total number of values contained in the original sample. Probability values p are also given in table 5. It is clear that from 36% to 56% of values in the initial experimental data have a fractionally stable law of distribution. As we can see from the table the characteristic indicators of the fractionally stable law α and β, for all studied samples vary within 0.91⩽α⩽1.24, 0.17⩽β⩽0.75.

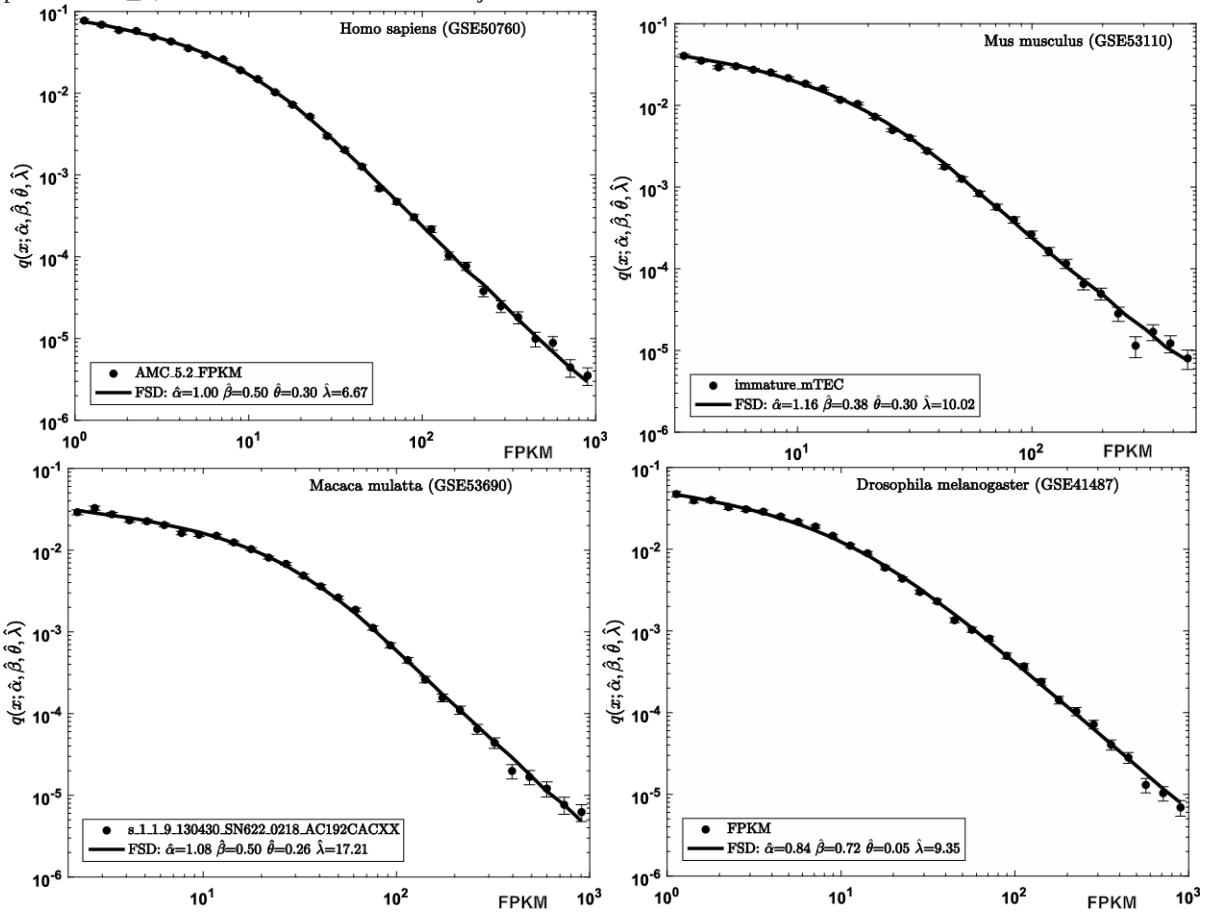

Figure 1: Probability density distribution of gene expression for human tissues (Homo Sapiens), mouse tissues (Mus Musculus), macaque tissues (Macaca Mulatta), drosophila tissues (Drosophila melanogaster). Experimental data points, the solid curve is the fractional-stable density for the estimated parameter values α̂,β̂,θ̂,λ̂. Parameter values are given in Figures

Let us now consider the behavior of the experimental density in the domains x∉R. In Fig. 2 the distribution of the probability density in the entire range of expression values is shown. It is clear from this figure that at x∈R_2 the deviation of the empirical density from the theoretical relationship is slight. Indeed, in the domain R_2 the experimental density has a power law x^(-α-1) which is described well by a fractional-stable law extended into this domain. However, the attempts to verify the hypothesis H_0 for the domains R+R_2 lead to the necessity to reject this the hypothesis. The reason for this lies in the fluctuations of the empirical and theoretical (since the theoretical density was calculated by the Monte Carlo method) densities, caused by the statistical error of the results. The increase in statistical errors in R_2 is caused by a small volume of data contained in it. In fact, in table 5 the probability is given p_2=P{X∈R_2}. As we can see, no more than 2% of data are contained in this area of values. Therefore, an increase in fluctuations leads to an increase in the value of the calculated

Fisher's statistics χ^2 which can lead to a false-negative result of testing the hypothesis H_0. Therefore, without loss of generality, it can be asserted that in the domain R_2 the experimental data also obey a fractionally stable law, but only in the case of their power-law relationship.

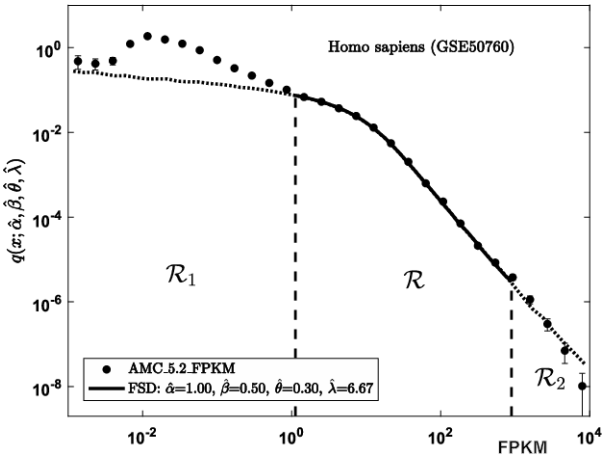

Figure 2: The distribution of the probability density of the expression of genes obtained by using NGS over the entire range of their values. Points are experimental data for human tissue (series of experiments GSE50760), the solid curve is the density of the fractionally stable law for the indicated values of the parameters in the domain R (dashed curves show the boundaries of this domain E_min and E_max), the dotted curve is the fractional-stable density for the same parameter values and x∉R

In the domain R_1 the empirical distribution density has a completely different behavior than the fractional-stable law (see Fig. 2). One of the possible reasons for this behavior may be the appearance of an additional component in the distribution of experimental data. This means that in this domain one part of the data has a fractionally stable distribution, the other part of the data is distributed according to some unknown distribution law f(x;Φ), where Φ is the vector of parameters of this law. Such models are well known and are called shift-scale mixtures of distributions. In the general case, such shift-scale mixtures have the form

$$F(x) = \sum_{j=1}^{M} w_j f_j((x-\mu_j)/\lambda_j; \mathbf{\Phi}_j), \qquad (2)$$

where w_j - weight coefficients (w_1+w_2+⋯+w_M=1), f_j (x;Φ_j) - components of this mixture of distributions, M - the number of components of the mixture, Φ_j - the parameters of the jth component of the mixture, μ_j,λ_j - the parameter of the shift and scale of the jth component of the mixture.

Based on this, we can conclude that in the domain R_1+R_2+R experimental data on gene expression obtained with the use of NGS are quite likely to be described by a shift-scale mixture with two components

$$F(x) = w_1 q(x;\alpha,\beta,\theta,\lambda) + w_2 f(x;\mathbf{\Phi})$$

The first component q(x;α,β,θ,λ) is a fractionally stable law of distribution (3), the second component f(x;Φ) is a distribution law unknown at present. At the same time, in the domain R_1 the influence of the component f(x;Φ) is dominant and in the domain R+R_2 the influence of this component is negligible. This leads to the fact that in the domain R+R_2 the data are described well through a fractionally stable law.

3. Conclusion

The issue of the nature of the probabilistic distribution of gene expression was considered in this work. Since the experimental distributions have power-law asymptotics of the form (1), then this made it possible to put forward an assumption about the fractionally stable character of the experimental distributions. It was possible to establish that the distribution of gene expression has the form of a shift-scale mixture of distributions, the general form of which is determined by the formula (2). In this mixture one component is determined by a fractionally stable law (3). It was possible to determine the domains where each of the components dominates. In particular, a fractionally stable distribution law dominates in the domain R+R_2. The influence of other components of the mixture in this domain is negligible. It can be asserted that in this range of values, gene expression is described by a fractionally stable law. This statement is confirmed by the criterion of consent. This assertion is confirmed by Fisher's goodness-of-fit test χ^2 which does not reject this assumption. In the domain R_1 the influence of the remaining components of the mixture begins to affect, which is expressed by the appearance of a "hump" in the experimental distribution (see Fig. 2). At present, we could not determine the type of these components. However, considering that in the domain R+R_2 the influence of these components is negligible; it can be asserted that these components rapidly decrease with an increase of x.

It is possible to make some assumptions about the origin of the component f(x;Φ). One of the possible reasons for its appearance may be errors in alignment to the base gene. In this case, when aligning readings to the base gene, the nucleotide sequence may be prematurely terminated, an excessively long sequence may be assembled, or a sequence may be assembled that does not match any of the base genes. All this leads to an erroneous determination of the gene, and, consequently, to an erroneous determination of the amount of gene expression. This the hypothesis is supported by the fact that the component f(x;Φ) has a rapidly decreasing "tail" of the distribution. As it is known, all data associated with errors of various kinds have this property. This the hypothesis also suggests one of the possible ways to apply the results obtained. Indeed, since with expressions <1 FPKM errors begin to play a significant role, then for further data processing it is necessary to select only genes with expressions ⩾1 FPKM. Another area of application of the results obtained is the development of methods for

normalizing experimental data taking account of the fractionally stable nature of distributions. However, all of the foregoing requires further studies in this direction and verification of the assumptions made.

Appendix

4. Fractionally stable laws and estimates of their parameters

Fractionally stable distributions are limiting distributions of sums of independent identically distributed random variables. This class of distributions was first introduced in the work [19]. It received its name in the work [16]. The density of a fractionally stable distribution is described by the Melin transform of two stable densities

$$q(x;\alpha,\beta,\theta,\lambda) = \int_0^\infty g(xy^{\beta/\alpha};\alpha,\theta,\lambda)g(y;\beta,1,1)y^{\beta/\alpha}dy. \quad (3)$$

Here g(x;α,θ,λ) and g(y;β,1,1) are the densities of a stable and one-sided stable law [20] described by a characteristic function

$$\hat{g}(k;\alpha,\theta,\lambda) = \exp\{-\lambda|k|^\alpha \exp\{-i\alpha\theta(\pi/2)\mathrm{sign}(k)\}\}. \quad (4)$$

Characteristic indicators vary within 0<α⩽2 and 0<β⩽1. As one can see, fractional stable distributions are a four-parameter class of distributions fully described by their parameters. The first two parameters are characteristic parameters and vary within the range α∈(0,2] and β∈(0,1], θ this is the parameter of asymmetry (|θ|⩽min(1,2/α-1)) and λ this is the parameter of scale λ>0. From the definition (3) it follows that this class of distributions has power asymptotics q(x;α,β,θ,λ)∝x^(-α-1),x→∞. From the expression (3) it also follows that at β=1 the class of fractionally stable distributions becomes the class of stable laws. Indeed, if β=1 and θ=1, then the strict-stable density g(y,1,1,1) is a singular distribution in the point y=1. In view of this property from (3) we get ∫_0^∞ g(xy^(β/α);α,θ,λ)δ(y-1)y^(β/α) dy=g(x;α,θ,λ), where δ(y-1) is the Dirac function. If we assume α=2,β=1,θ=0, then we obtain a normal distribution. Thus, the class of fractionally stable laws contains the class of stable laws as a subclass at β=1.

5. Results of testing the hypothesis H_0 for next-generation sequencing data

Table 1: The results of approximation of gene expression obtained with the use of next-generation sequencing, by fractional-stable distributions and testing the hypothesis H0 by using Fisher's test χ2. The results were given according to the series of experiments: GSE50760 (human), GSE53110 (mouse), GSE53690 (macaque), GSE54600 (drosophila)

Organism	Channel name					H_0	p	p_2
Human	AMC_2.2_FPKM	1.01	0.51	0.30	6.72	1	0.543	0.002
	AMC_3.2_FPKM	0.98	0.46	0.30	6.91	1	0.536	0.003
	AMC_5.2_FPKM	1.00	0.50	0.30	6.67	1	0.540	0.002
	AMC_6.2_FPKM	0.99	0.40	0.35	6.52	0	0.541	0.002
	AMC_7.2_FPKM	0.98	0.50	0.30	6.82	1	0.545	0.002
	AMC_8.2_FPKM	0.97	0.40	0.31	7.33	0	0.540	0.003
	AMC_9.2_FPKM	0.98	0.50	0.30	6.87	1	0.545	0.002
	AMC_10.2_FPKM	1.00	0.45	0.31	6.98	1	0.542	0.003
	AMC_12.2_FPKM	0.99	0.44	0.30	7.25	1	0.546	0.003
	AMC_13.2_FPKM	0.95	0.39	0.30	7.42	1	0.537	0.003
	AMC_17.2_FPKM	1.01	0.36	0.34	7.89	1	0.567	0.003
	AMC_18.2_FPKM	0.96	0.39	0.30	7.43	0	0.539	0.003
	AMC_19.2_FPKM	0.99	0.54	0.31	7.21	0	0.555	0.002
	AMC_20.2_FPKM	0.99	0.49	0.30	7.25	1	0.546	0.003
	AMC_21.2_FPKM	0.99	0.38	0.30	7.35	1	0.544	0.003
Mouse	cTEC	1.20	0.54	0.15	12.00	1	0.440	0.001
	mTEC	0.91	0.23	0.65	7.70	0	0.515	0.001
	immature_mTEC	1.16	0.38	0.30	10.02	1	0.515	0.001
	mature_mTECE	1.06	0.17	0.45	9.53	1	0.470	0.001
	Aire_neg_mTEC	1.05	0.35	0.32	9.62	1	0.508	0.001
	Aire_pos_mTEC	0.94	0.21	0.60	8.61	0	0.500	0.001
	Aire_knockout_mTEC	1.05	0.37	0.32	9.79	1	0.462	0.001
Macaque	s_1_1_7_130430	1.10	0.38	0.28	17.85	1	0.554	0.005
	s_1_1_9_130430	1.08	0.50	0.26	17.21	1	0.548	0.005
	s_2_1_20_130607	0.97	0.57	0.40	16.32	1	0.599	0.005
	s_1_1_12_130430	1.07	0.66	0.27	16.58	0	0.556	0.006

	s_1_1_22_130607	1.10	0.61	0.30	17.27	0	0.557	0.005
	s_2_1_25_130607	1.11	0.49	0.28	17.31	1	0.558	0.006
	s_1_1_11_130430	1.13	0.52	0.29	17.76	1	0.559	0.005
	s_1_1_10_130430	1.13	0.59	0.25	17.70	1	0.555	0.006
	s_1_1_16_130430	1.11	0.63	0.25	18.19	1	0.554	0.005
	s_1_1_18_130430	1.19	0.51	0.23	19.10	1	0.550	0.005
	s_1_1_15_130430	1.16	0.63	0.19	19.31	1	0.535	0.004
	s_1_1_19_130430	1.12	0.36	0.30	18.84	1	0.552	0.006
	s_1_1_5_130430	1.14	0.46	0.21	20.34	1	0.541	0.005
Drosophila	D_Yki_1	1.04	0.46	0.27	18.68	1	0.415	0.010
	D_Yki_2	1.00	0.37	0.39	18.38	1	0.369	0.011
	D_Yki_3	1.10	0.59	0.20	21.95	1	0.374	0.012
	M_Yki_1	1.07	0.41	0.30	20.11	1	0.364	0.011
	M_Yki_2	0.93	0.47	0.30	16.27	0	0.361	0.011
	M_Yki_3	1.10	0.60	0.26	21.80	1	0.407	0.011
	M_Yki_T_1	1.15	0.48	0.20	22.60	1	0.398	0.008
	M_Yki_T_2	1.08	0.48	0.28	24.72	0	0.383	0.012
	M_Yki_T_3	1.13	0.56	0.20	22.21	1	0.510	0.008
	Tr_Yki_13	1.22	0.69	0.15	23.00	1	0.512	0.008
	Tr_Yki_23	1.14	0.74	0.15	22.61	1	0.509	0.008
	Tr_Yki_33	1.24	0.75	0.11	25.54	1	0.431	0.008

Acknowledgments

This research study was carried out on the basis of the infrastructure of the Ulyanovsk State University with the support of the Ministry of Science and Higher Education of the Russian Federation. The authors declare there is no conflict of interests related to this article.

References

[1] V. A. Kuznetsov, G. D. Knott, and R. F. Bonner. General statistics of stochastic process of gene expression in eukaryotic cells. Genetics, 161(3):1321–1332, jul 2002.

[2] Chikara Furusawa and Kunihiko Kaneko. Zipf's Law in Gene Expression. Physical Review Letters, 90(8):8–11, feb 2003.

[3] Hiroki R Ueda, Satoko Hayashi, Shinichi Matsuyama, Tetsuya Yomo, Seiichi Hashimoto, Steve A Kay, John B Hogenesch, and Masamitsu Iino. Universality and flexibility in gene expression from bacteria to human. Proceedings of the National Academy of Sciences, 101(11):3765–3769, mar 2004.

[4] Chuan Lu and Ross D King. An investigation into the population abundance distribution of mRNAs, proteins, and metabolites in biological systems. Bioinformatics (Oxford, England), 25(16):2020–7, aug 2009.

[5] Viacheslav Saenko and Yurij Saenko. Approximation of Microarray Gene Expression Profiles by the Stable Laws. International Journal of Environmental Engineering, 2(1):98–102, 2015.

[6] Viacheslav Saenko and Yury Saenko. Application of the fractional-stable distributions for approximation of the gene expression profiles. Statistical Applications in Genetics and Molecular Biology, 14(3):295–306, jan 2015.

[7] Silvia Lazzardi, Filippo Valle, Andrea Mazzolini, Antonio Scialdone, Michele Caselle, and Matteo Osella. EMERGENT STATISTICAL LAWS IN SINGLE-CELL TRANSCRIPTOMIC DATA. pages 1–28, dec 2021.

[8] Vincent Piras and Kumar Selvarajoo. The reduction of gene expression variability from single cells to populations follows simple statistical laws. Genomics, 105(3):137–144, 2015.

[9] Mark D. Alter, Daniel B. Rubin, Keri Ramsey, Rebecca Halpern, Dietrich A. Stephan, L. F. Abbott, and Rene Hen. Variation in the large-scale organization of gene expression levels in the hippocampus relates to stable epigenetic variability in behavior. PLoS ONE, 3(10), 2008.

[10] Minou Nowrousian. Fungal gene expression levels do not display a common mode of distribution. BMC Research Notes, 6(1):559, dec 2013.

[11] Daniel Hebenstreit, Miaoqing Fang, Muxin Gu, Varodom Charoensawan, Alexander Van Oudenaarden, and Sarah A. Teichmann. RNA sequencing reveals two major classes of gene expression levels in metazoan cells. Molecular Systems Biology, 7(497):1–9, 2011.

[12] Daniel Hebenstreit and Sarah A. Teichmann. Analysis and simulation of gene expression profiles in pure and mixed cell populations. Physical Biology, 8(3), 2011.

[13] Xiaotu Ma, Ying Shao, Liqing Tian, Diane A. Flasch, Heather L. Mulder, Michael N. Edmonson, Yu Liu, Xiang Chen, Scott Newman, Joy Nakitandwe, Yongjin Li, Benshang Li, Shuhong Shen, Zhaoming Wang, Sheila Shurtleff, Leslie L. Robison, Shawn Levy, John Easton, and Jinghui Zhang. Analysis of error profiles in deep next-generation sequencing data. Genome Biology, 20(1):50, dec 2019.

[14] Bin Wang. A Zipf-plot based normalization method for high-throughput RNA-seq data. PLoS ONE, 15(4):1–15, 2020.

[15] Matteo Borella, Graziano Martello, and Davide Risso. PsiNorm : a scalable normalization for single-cell RNA-seq data Banchmarked methods. pages 1–13, 2021.

[16] V. N. Kolokoltsov, V. Yu. Korolev, and Vladimir V. Uchaikin. Fractional Stable Distributions. Journal of Mathematical Sciences, 105(6):2569–2576, 2001.

[17] V. E. Bening, V. Yu. Korolev, T. A. Sukhorukova, G. G. Gusarov, Viacheslav V. Saenko, Vladimir V. Uchaikin, and V. N. Kolokoltsov. Fractionally stable distributions. In V. Yu. Korolev and N. N. Skvortsova, editors, Stochastic Models of Structural Plasma Turbulence, pages 175–244. Brill Academic Publishers, Utrecht, 2006.

[18] Viacheslav V. Saenko. Estimation of the Parameters of Fractional-Stable Laws by the Method of Minimum Distance. Journal of Mathematical Sciences, 214(1):101–114, feb 2016.

[19] Marcin Kotulski. Asymptotic distributions of continuous-time random walks: A probabilistic approach. Journal of Statistical Physics, 81(3-4):777–792, nov 1995.

[20] Vladimir M. Zolotarev. One-dimensional stable Distributions. Amer. Mat. Soc., Providence, RI, 1986.