Supplementary Material

Heterogeneous Structure of Stem Cells Dynamics: Statistical Models and Quantitative Predictions

Paul Bogdan[1], Bridget Deasy[2,3], Burhan Gharaibeh[4,5], Timo Roehrs[3], Radu Marculescu[6]

[1]Department of Electrical Engineering, University of Southern California, Los Angeles, CA 90089-2560, USA.

[2]CellStock, Pittsburgh, PA 15237-1941, USA.

[3]McGowan Institute of Regenerative Medicine of UPMC and Department of Bioengineering, University of Pittsburgh, Pittsburgh, PA 15213, USA.

[4]Institute for Complex Engineered Systems, Carnegie Mellon University, Pittsburgh, PA15213, USA.

[5]Stem Cell Research Center (SCRC), University of Pittsburgh, Pittsburgh, PA 15219, USA.

[6]Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

Summary of Supplementary Material

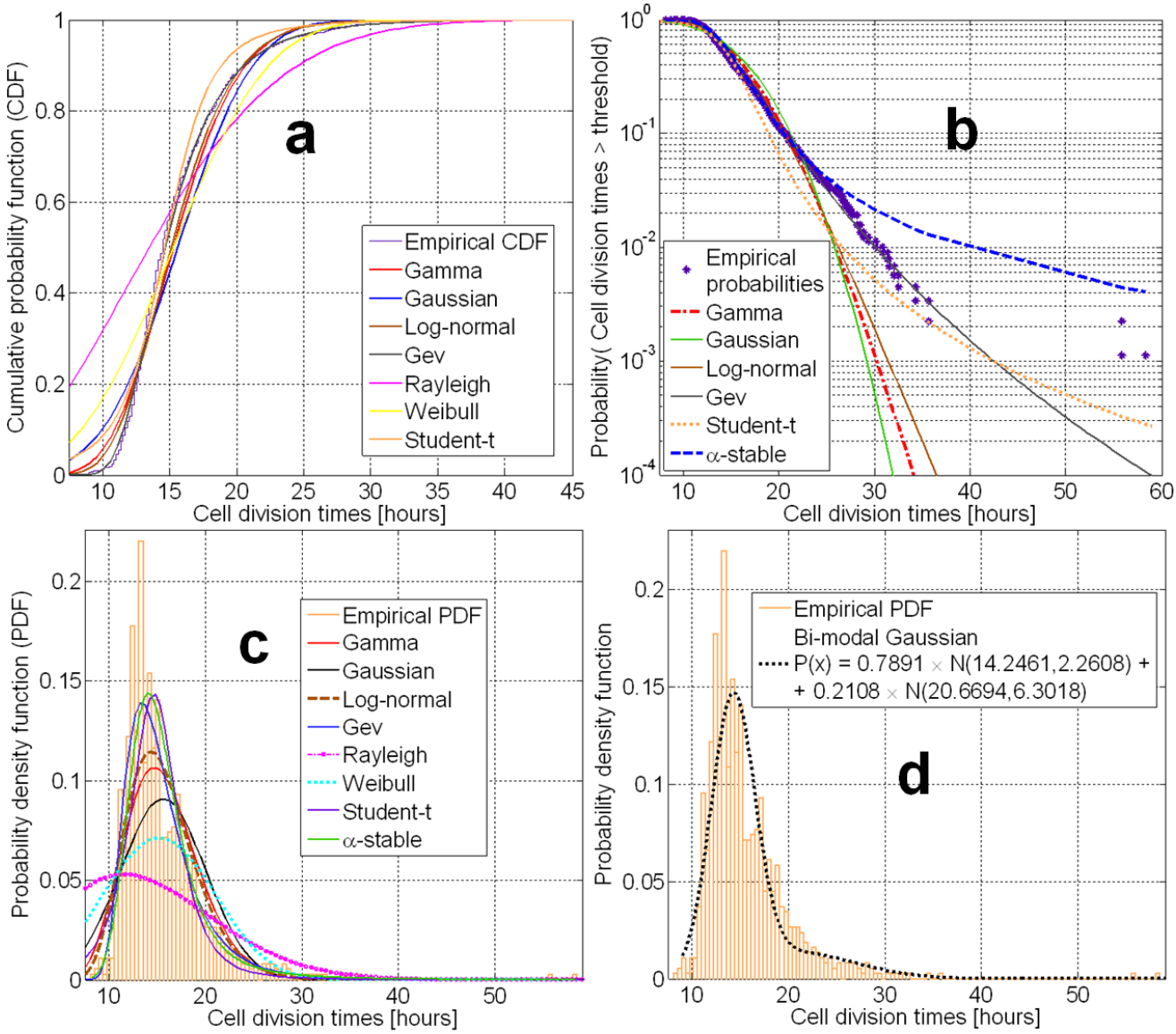


Figure 1. **Distribution Fittings for Mouse Stem Cell Data of Cell Division Times** suggests that data do not fit to several common uni-modal distributions or to Gaussian bi-modal. **a)** Cumulative probability function of the recorded mouse muscle stem cell division times and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev), Rayleigh, Weibull, Student-t). **b)** Probability that cell division times exceed a given threshold and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev), Student-t, α-stable distribution). **c)** Probability density function of the cell division times and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev) distribution, Rayleigh, Weibull, Student-t, α-stable distribution). **d)** Comparison between the empirical PDF of mouse stem cell division times and the bi-modal Gaussian distribution with estimated parameters summarized in the legend of this plot. By simply graphically analyze the plot we can conclude that bi-modal Gaussian distribution does not provide a good fitting.

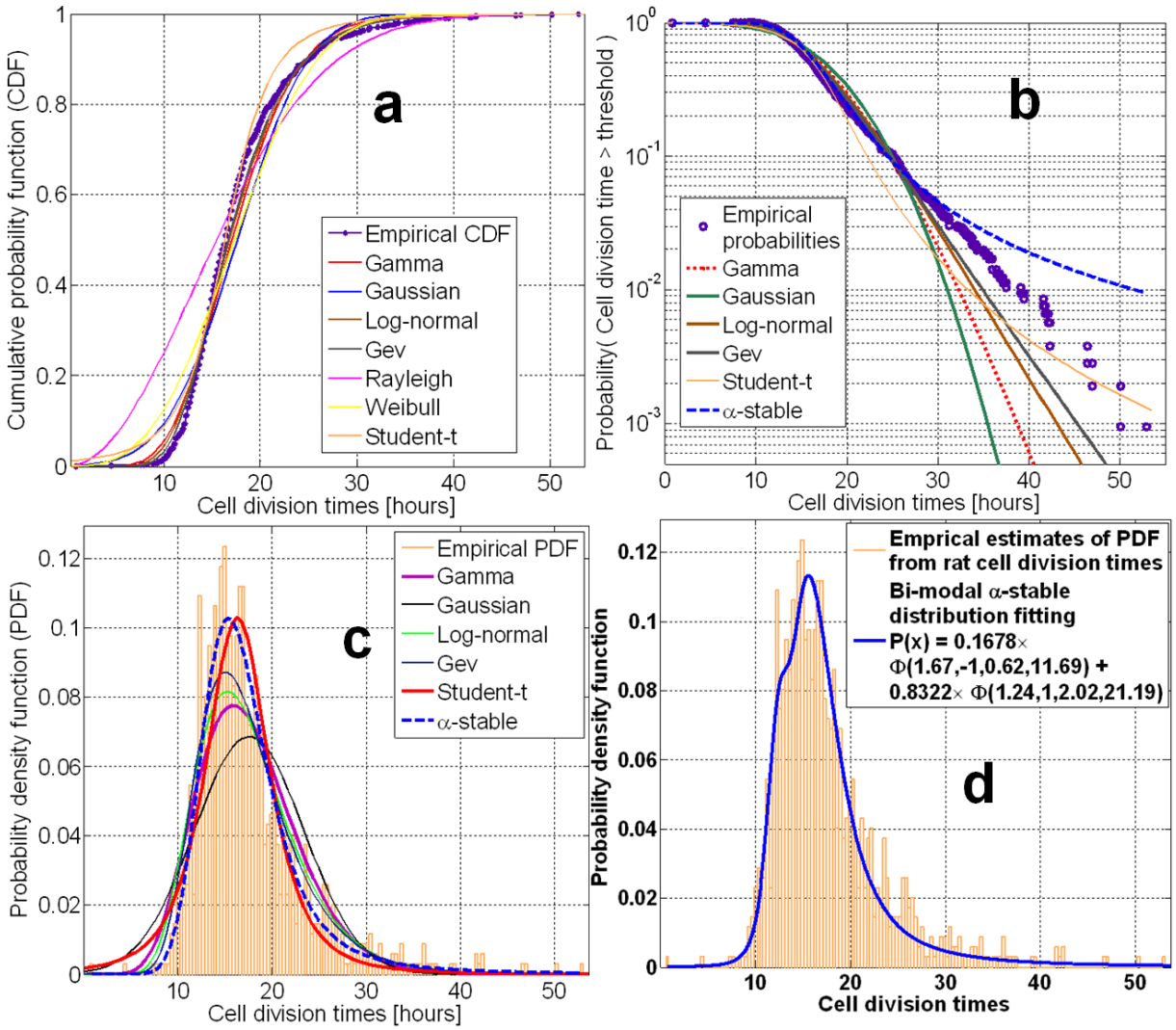


Figure 2.**a)** Cumulative probability function of the recorded rat musle stem cell division times and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev), Rayleigh, Weibull, Student-t). **b)**Probability that cell division times exceed a given threshold and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev), Student-t, α-stable distribution); for rat stem cell division times. **c)**Probability density function of the cell division times and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev) distribution, Rayleigh, Weibull, Student-t, α-stable distribution). **d)**Comparison between the empirical PDF of rat stem cell DTs and the bi-modal α-stable distribution for all 1055 samples. The p-value of 0.11 for the Kolmogorov-Smirnov (K-S) test shows that the postulated bi-modal α-stable distribution cannot be rejected as a model.

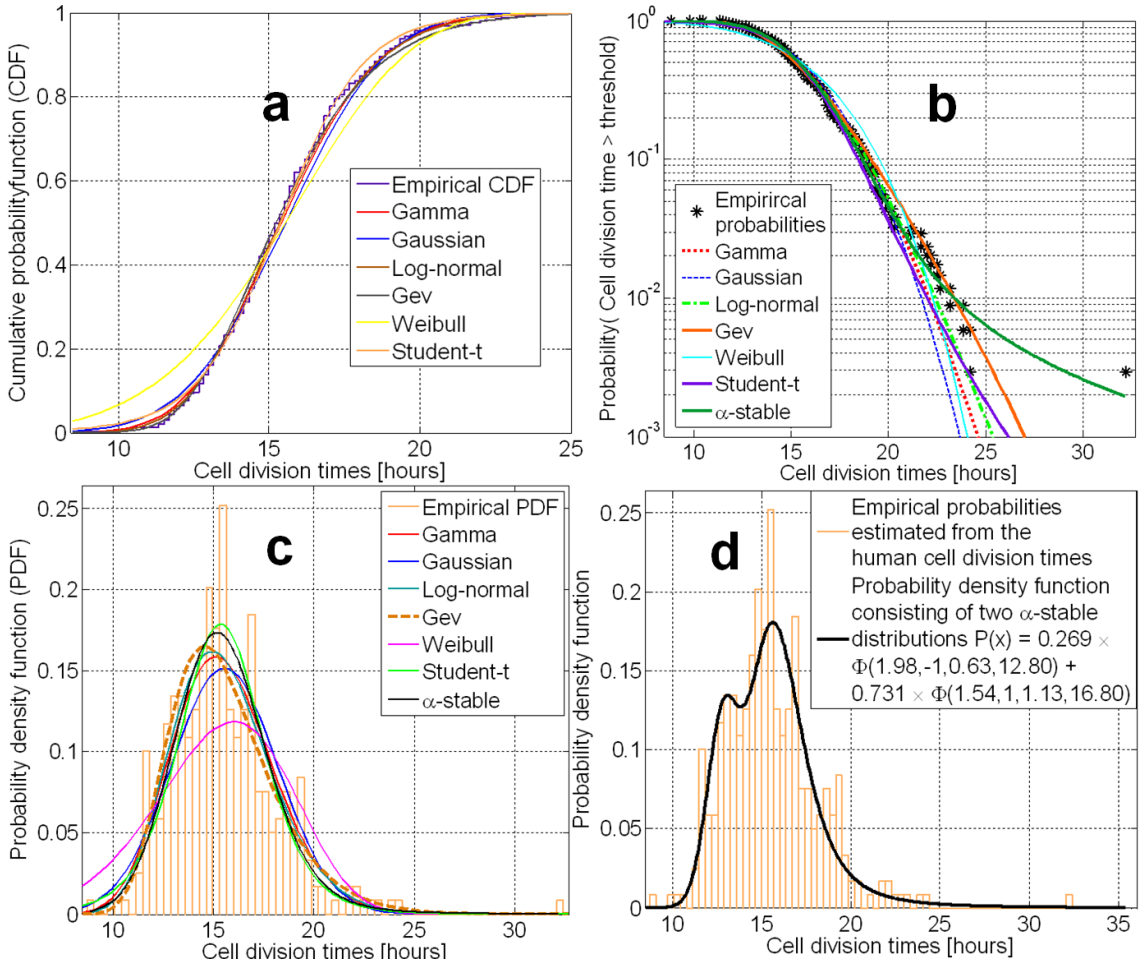


Figure 3.**a)**Cumulative probability function of the recorded human mesenchymal stem cell division times and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev), Rayleigh, Weibull, Student-t). **b)** Probability that cell division times exceed a given threshold and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev), Student-t, α-stable distribution); for human stem cell division times. **c)**PDF of the cell division times and several uni-modal maximum likelihood fittings (i.e., Gaussian, Gamma, Log-normal, Generalized extreme value (Gev) distribution, Rayleigh, Weibull, Student-t, α-stable distribution). d) Comparison between the empirical PDF of human MSC division times and the bi-modal α-stable distribution for all 350 samples. The p-value of 0.81 for the Kolmogorov-Smirnov (K-S) test shows that the postulated bi-modal α-stable distribution cannot be rejected as a model.

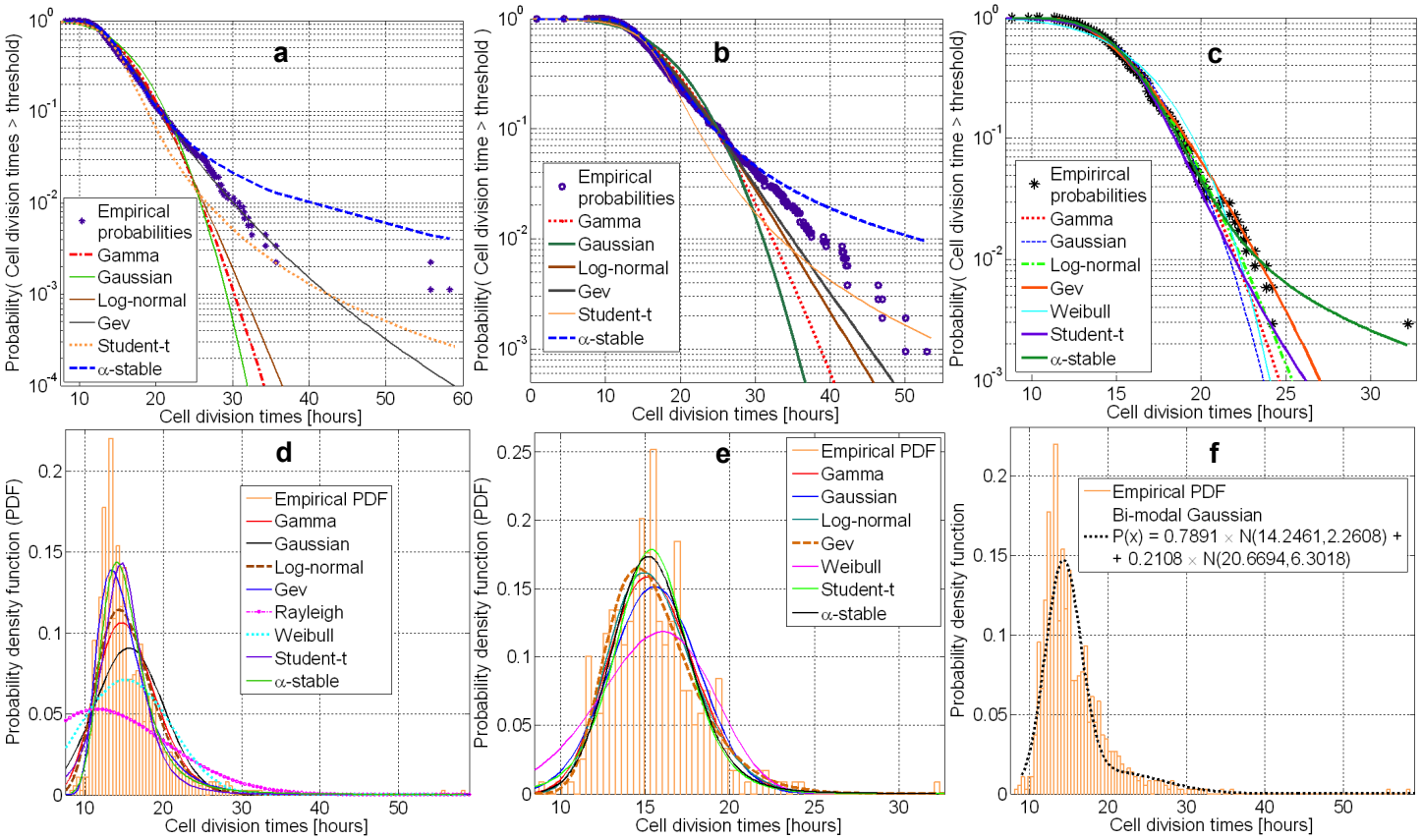


Figure 4. Statistical investigation and interpretation of tracked MDSC division times. Exceedance probability analysis (i.e., probability of observing a MDSC division time greater than a specific threshold) and its maximum likelihood fittings with a few known mathematical distributions (i.e., Gamma, Gaussian, Log-normal, generalized extreme value (Gev), Student-t, asymmetric alpha-stable) for (a) mouse MDSC division times, (b) rat MDSC division times and (c) human MSC division times. d) Maximum likelihood fittings of uni-modal bell-shaped probability density functions (PDF) cannot capture the complexity of mouse MDSC division times. Rejection of uni-modal mathematical modeling approaches for MDSC growth is motivated not only by the inconsistencies observed via graphical inspection of the empirical PDF against postulated PDF, but also by the very small p-value probabilities of the Kolmogorov-Smirnov (K-S) test. e) Maximum likelihood fitting of uni-modal PDFs does not offer a good characterization of human MSC division times.f) Plot showing the maximum likelihood fitting of a bi-modal Gaussian distribution. Although graphical analysis shows a better PDF fitting of empirical data, bi-modal Gaussian modeling should be rejected upon K-S test investigation. The high and narrow peaks together with existence of long tails suggest a modeling approach that goes beyond Gaussian assumption.

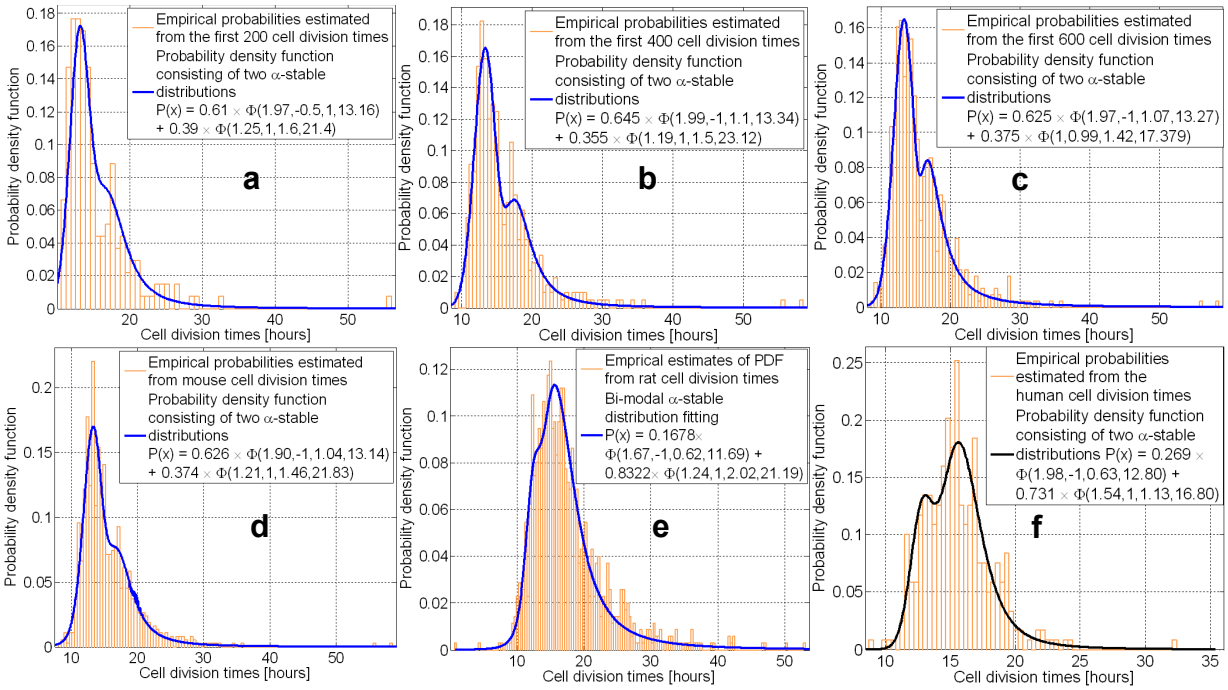


Figure 5. Statistical analysis of bi-modal asymmetric α-stable distribution fitting of stem cell division times. a) Comparison between the empirical PDF of mouse MDSC division times (DTs) and the bi-modal α-stable distribution (estimated parameters summarized in the legend) for the first 200 samples. The high p-value of 0.89 for the Kolmogorov-Smirnov (K-S) test shows that the postulated bi-modal α-stable distribution cannot be rejected as a model. b) Comparison between the empirical PDF of mouse MDSC DTs and the bi-modal α-stable distribution (estimated parameters summarized in the legend) obtained from first 400 samples. The K-S test shows that the postulated bi-modal α-stable distribution cannot be rejected (i.e., has a p-value of 0.93). c) Comparison between the empirical PDF of mouse MDSC DTs and the bi-modal α-stable distribution (estimated parameters summarized in the legend) for the first 600 samples. The p-value of 0.9 for the K-S test shows that the postulated bi-modal α-stable distribution cannot be rejected as a model. d) Comparison between the empirical PDF of mouse MDSC DTs and the bi-modal α-stable distribution for all 889 samples. The high p-value of 0.78 for the K-Stest shows that the postulated bi-modal α-stable distribution cannot be rejected as a model. e) Comparison between the empirical PDF of rat MDSC DTs and the bi-modal α-stable distribution for all 1055 samples. The p-value of 0.11 for the K-S test shows that the postulated bi-modal α-stable distribution cannot be rejected as a model. f) Comparison between the empirical PDF of human MSC DTs and the bi-modal α-stable distribution for all 350 samples. The p-value of 0.81 for the K-S test shows that the postulated bi-modal α-stable distribution cannot be rejected as a model.

Note 1.Goodness-of-fit Analysis for Uni-modal Distributions for Mouse Muscle Stem Cell division times.

To investigate whether the cell division times can be well characterized by uni-modal distribution, we performed the following steps:

- First, investigate the parameters of the postulated distribution (e.g., Gaussian, Gamma, Generalized extreme value, Log-normal, Rayleigh, Weibull, Alpha-stable) via maximum likelihood estimation method.
- Generate 100,000 samples from the postulated distribution.
- Perform the Kolmogorov-Smirnov (K-S) (a short description of the KS test is provided below) test to determine if cell division times and the newly generated samples are drawn from the same underlying distribution for a desired significance level (we used throughout the experiments 0.05 statistical significance value). More precisely, we performed a binary hypothesis testing problem characterized by two parameters: h and p-value. When $h = 1$ we reject the null hypothesis testing at predefined statistical significance level. When $h = 0$ we cannot reject the null hypothesis testing for the considered statistical significance level. The decision to reject the null hypothesis occurs when the considered significance level equals or exceeds the p-value.

Summary of KS test: The KS test measures the distance between the empirical and the postulated CDFs as the maximum value of their absolute differences:

$$KS\_dis\tan ce = \max_{x-sample\ \ domain} \left| P_{empirical}(x) - P_{postulated}(x) \right| \tag{1}$$

This KS distance is regarded as a random variable. At this point, we rely on the Glivenko-Cantelli theorem which states that there exists a uniform convergence of the empirical distribution ($F_{empirical}$) function to the true distribution (F_{true}) function from which the data samples originate; in mathematical terms this means that:

$$\sup_x \left| F_{empirical}(x) - F_{true}(x) \right| \xrightarrow{a.s.} 0 \tag{2}$$

Using the Glivenko-Cantelli theorem[5], the Kolmogorov-Smirnov test states that if the empirical observations characterized by the ($P_{empirical}$) come from the postulated distribution ($P_{postulated}$), then the KS distance converges to zero almost surely (denoted as "a.s." in equation (2) above). In addition, by central limit theorem, KS test states that:

$$\sqrt{n}\left| P_{empirical}(X_n \le x) - P_{postulated}(X \le x) \right| \xrightarrow{d} N\left(0, P_{postulated}(X \le x)\left[1 - P_{postulated}(X \le x)\right]\right) \tag{3}$$

or alternatively:

$$P\left(\sqrt{n}\left| P_{empirical}(X_n \le x) - P_{postulated}(X \le x) \right| \le u\right) = 1 - 2\sum_{i=1}^{\infty} (-1)^{i-1} e^{-2i^2u} \tag{4}$$

Both relations (3) and (4) help us determine whether to reject the null hypothesis (that the postulated distribution is a good fit for the empirical data) when this probability is below the significance level of 0.05.

The results concerning the appropriateness of modeling cell division times are summarized in Table 1. Note that because the data exhibited several peaks, we considered only uni-modal "bell-shaped" distribution (i.e., Gaussian, Gamma, Generalized extreme value, Log-normal, Weibull, Alpha-stable). Nevertheless, the analysis can be replicated for other types of distributions. Of note, to obtain these results we wrote several scripts in Matlab[4] and used advanced statistical techniques which are also described in [3].

Table 1.Goodness-of-fit Results for Uni-modal Distributions for Mouse Muscle Stem Cell division times.

Uni-modal Distribution	Parameters	Goodness-of-fit		
		h	*p*-value	Test statistic
Gaussian distribution	$\mu = 15.60 \pm 0.14$ $\sigma = 4.39 \pm 0.10$	1	2.9617×10^{-14}	0.13904668
Gamma distribution	$a = 16.46 \pm 0.77$ $b = 0.94 \pm 0.04$	1	1.6363×10^{-10}	0.11875231
Generalized extreme value distribution	$k = 0.12 \pm 0.021$ $\sigma = 2.66 \pm 0.07$ $\mu = 13.68 \pm 0.09$	1	$2.\ 7646 \times 10^{-3}$	0.05098503
Log-normal distribution	$\mu = 2.71 \pm 0.007$ $\sigma = 0.23 \pm 0.005$	1	1.2421×10^{-6}	0.09315231
Weibull distribution	$a = 17.18 \pm 0.19$ $b = 3.13 \pm 0.06$	1	5.6170×10^{-24}	0.18145422
α-stable distribution	$\alpha = 1.49 \pm 0.03$ $\beta = 1.00 \pm 0.00$ $\gamma = 1.93 \pm 0.05$ $\delta = 16.37 \pm 0.15$	1	2.1029×10^{-4}	0.13610798
Student-t distribution	$\mu = 14.63 \pm 0.11$ $\sigma = 2.55 \pm 0.11$ $\nu = 2.87 \pm 0.32$	1	3.2196×10^{-7}	0.09745309

As we can observe from third column, for all "bell-shaped" considered probability distributions, the assumption that the cell division times can be modeled through a unimodal PDF is rejected at 0.05 statistical significance test.

Note 2.Goodness-of-fit analysis for uni-modal distributions for Rat Muscle Stem Cell division times.

To investigate whether the cell division times can be well characterized by uni-modal distribution, we performed the following steps:

- First, investigate the parameters of the postulated distribution (e.g., Gaussian, Gamma, Generalized extreme value, Log-normal, Rayleigh, Weibull, Alpha-stable) via maximum likelihood estimation method.
- Generate 100,000 samples from the postulated distribution.
- Perform the Kolmogorov-Smirnov (K-S) test to determine if cell division times and the newly generated samples are drawn from the same underlying distribution for a desired significance level (we used throughout the experiments 0.05 statistical significance value). More precisely, we performed a binary hypothesis testing problem characterized by two parameters: h and p-value. When $h = 1$ we reject the null hypothesis testing at predefined statistical significance level. When $h = 0$ we cannot reject the null hypothesis testing for the considered statistical significance level. The decision to reject the null hypothesis occurs when the considered significance level equals or exceeds the p-value.

The results concerning the appropriateness of modeling cell division times are summarized in Table 1. Note that because the data exhibited several peaks, we considered only uni-modal “bell-shaped” distribution (i.e., Gaussian, Gamma, Generalized extreme value, Log-normal,Weibull, Alpha-stable). Nevertheless, the analysis can be replicated for other types of distributions.

Table 2: Goodness-of-fit results for uni-model distributions for Rat Muscle Stem Cell division times.

Uni-modal Distribution	Parameters	Goodness-of-fit		
		h	*p*-value	Test statistic
Gaussian distribution	$\mu = 17.57 \pm 0.17$ $\sigma = 5.81 \pm 0.12$	1	8.6207×10^{-16}	0.13560805
Gamma distribution	$a = 10.80 \pm 0.46$ $b = 1.62 \pm 0.07$	1	3.2798×10^{-8}	0.09652654
Generalized extreme value distribution	$k = 0.01 \pm 0.01$ $\sigma = 4.21 \pm 0.09$ $\mu = 15.12 \pm 0.13$	1	1.1136×10^{-3}	0.06240805
Log-normal distribution	$\mu = 2.81 \pm 0.01$ $\sigma = 0.30 \pm 0.006$	1	5.5998×10^{-4}	0.06520805
Weibull distribution	$a = 19.58 \pm 0.21$ $b = 2.96 \pm 0.06$	1	1.1036×10^{-13}	0.12596635
α-stable distribution	$\alpha = 1.48 \pm 0.02$ $\beta = 1.00 \pm 0.00$ $\gamma = 2.71 \pm 0.05$ $\delta = 18.70 \pm 0.18$	1	4.452×10^{-3}	0.05387772
Student-t distribution	$\mu = 16.31 \pm 0.14$ $\sigma = 3.54 \pm 0.15$ $\nu = 2.79 \pm 0.28$	1	1.3095×10^{-5}	0.07876635

As we can observe from third column, for all considered probability distributions, the assumption that the cell division times can be modeled through a uni-modal PDF is rejected at 0.05 statistical significance test except for the Alpha-stable distribution which shows a small p-value, but above the 0.05 significance level.

Note 3.Goodness-of-fit Analysis for Unimodal Distributions for Human Mesenchymal Stem Cell division times.

To investigate whether the cell division times can be well characterized by unimodal distribution, we performed the following steps:

- First, investigate the parameters of the postulated distribution (e.g., Gaussian, Gamma, Generalized extreme value, Log-normal, Rayleigh, Weibull, Alpha-stable) via maximum likelihood estimation method.
- Generate 100,000 samples from the postulated distribution.
- Perform the Kolmogorov-Smirnov (K-S) test to determine if cell division times and the newly generated samples are drawn from the same underlying distribution for a desired significance level (we used throughout the experiments 0.05 statistical significance value). More precisely, we performed a binary hypothesis testing problem characterized by two parameters: h and p-value. When $h = 1$ we reject the null hypothesis testing at predefined statistical significance level. When $h = 0$ we cannot reject the null hypothesis testing for the considered statistical significance level. The decision to reject the null hypothesis occurs when the considered significance level equals or exceeds the p-value.

The results concerning the appropriateness of modeling cell division times are summarized in Table 1. Note that because the data exhibited several peaks, we considered only uni-modal “bell-shaped” distribution (i.e., Gaussian, Gamma, Generalized extreme value, Log-normal, Weibull, Alpha-stable). Nevertheless, the analysis can be replicated for other types of distributions.

Table 3: Goodness-of-fit Results for Uni-model Distributions for Human Mesenchymal Stem Cell division times.

Uni-modal Distribution	**Parameters**	**Goodness-of-fit**		
		h	*p*-value	Test statistic
Gaussian distribution	$\mu = 15.54 \pm 0.14$ $\sigma = 2.63 \pm 0.10$	1	2.9540×10^{-2}	0.07928304093
Gamma distribution	$a = 37.40 \pm 2.84$ $b = 0.41 \pm 0.03$	0	1.9537×10^{-1}	0.05888304093
Generalized extreme value distribution	$k = -0.06 \pm 0.02$ $\sigma = 2.23 \pm 0.09$ $\mu = 14.41 \pm 0.13$	0	3.7814×10^{-1}	0.04973450292
Log-normal distribution	$\mu = 2.73 \pm 0.01$ $\sigma = 0.16 \pm 0.006$	0	5.6247×10^{-1}	0.04307777777
Weibull distribution	$a = 16.66 \pm 0.18$ $b = 5.26 \pm 0.17$	1	1.7593×10^{-4}	0.11800877192
α-stable distribution	$\alpha = 1.85 \pm 0.03$ $\beta = 0.99 \pm 0.002$ $\gamma = 1.63 \pm 0.03$ $\delta = 15.60 \pm 0.05$	0	5.2391×10^{-1}	0.04371304093
Student-t distribution	$\mu = 15.36 \pm 0.13$ $\sigma = 2.14 \pm 0.12$ $\nu = 6.37 \pm 1.90$	0	5.3407×10^{-1}	0.04402865497

As we can observe from third column, for all "bell-shaped" considered probability distributions, the assumption that the cell division times can be modeled through a uni-modal PDF is rejected at 0.05 statistical significance test except for the Student-t, Generalized extreme value, Log-normal, Alpha-stable distributions which show a p-value larger than the 0.05 significance level.

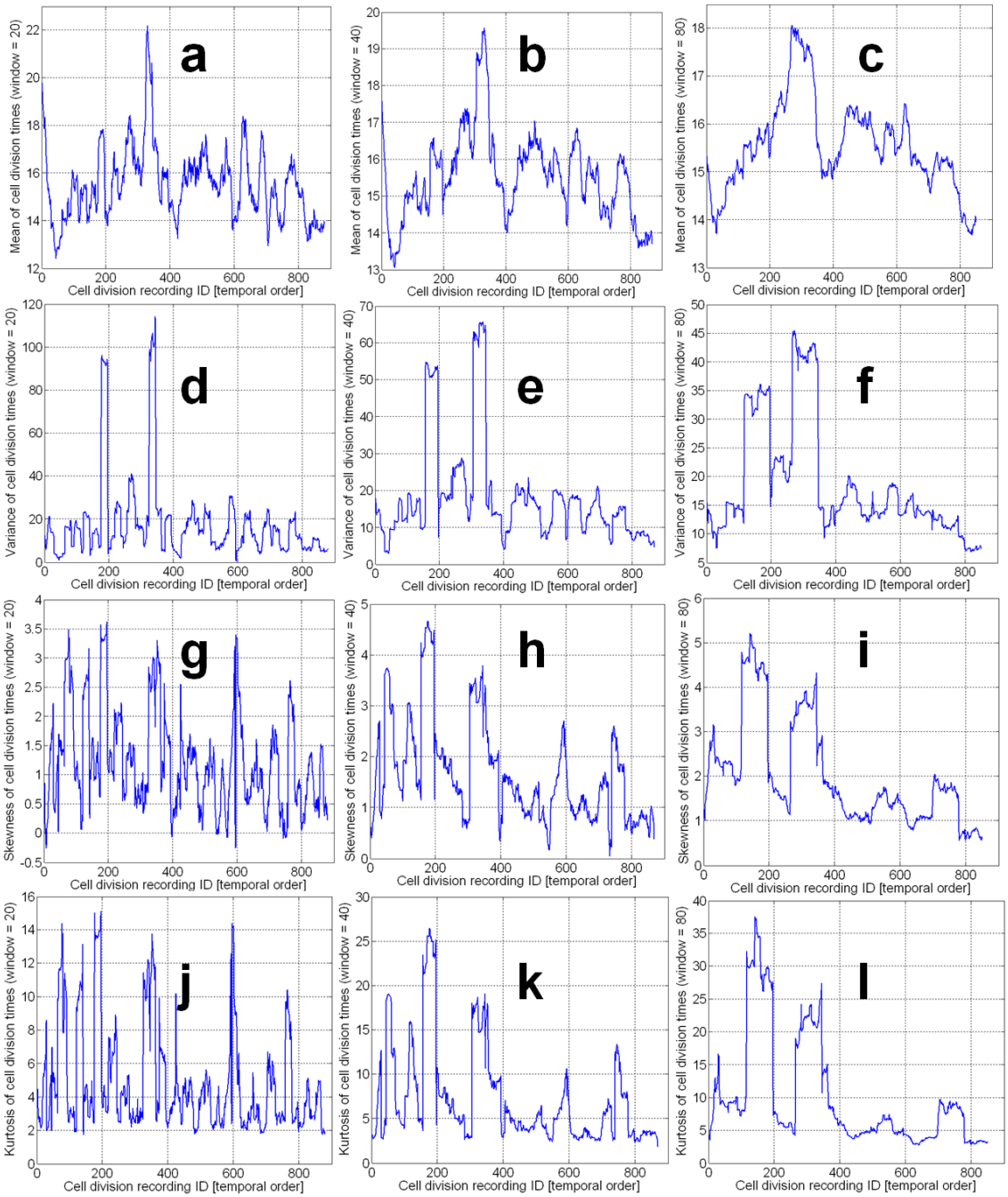


Figure 6. Mean of mouse stem cell division times computed as a function of time series and over a moving window of 20 (**a**), 40 (**b**) and 80 (**c**) samples. Variance of mouse cell division times over a moving window of 20 (**d**), 40 (**e**) and 80 (**f**) samples. Skewness of mouse cell division times over a moving window of 20 (**g**), 40 (**h**) and 80 (**i**) samples. Kurtosis of mouse cell division times over a moving window of 20 (**j**), 40 (**k**) and 80 (**l**) samples.

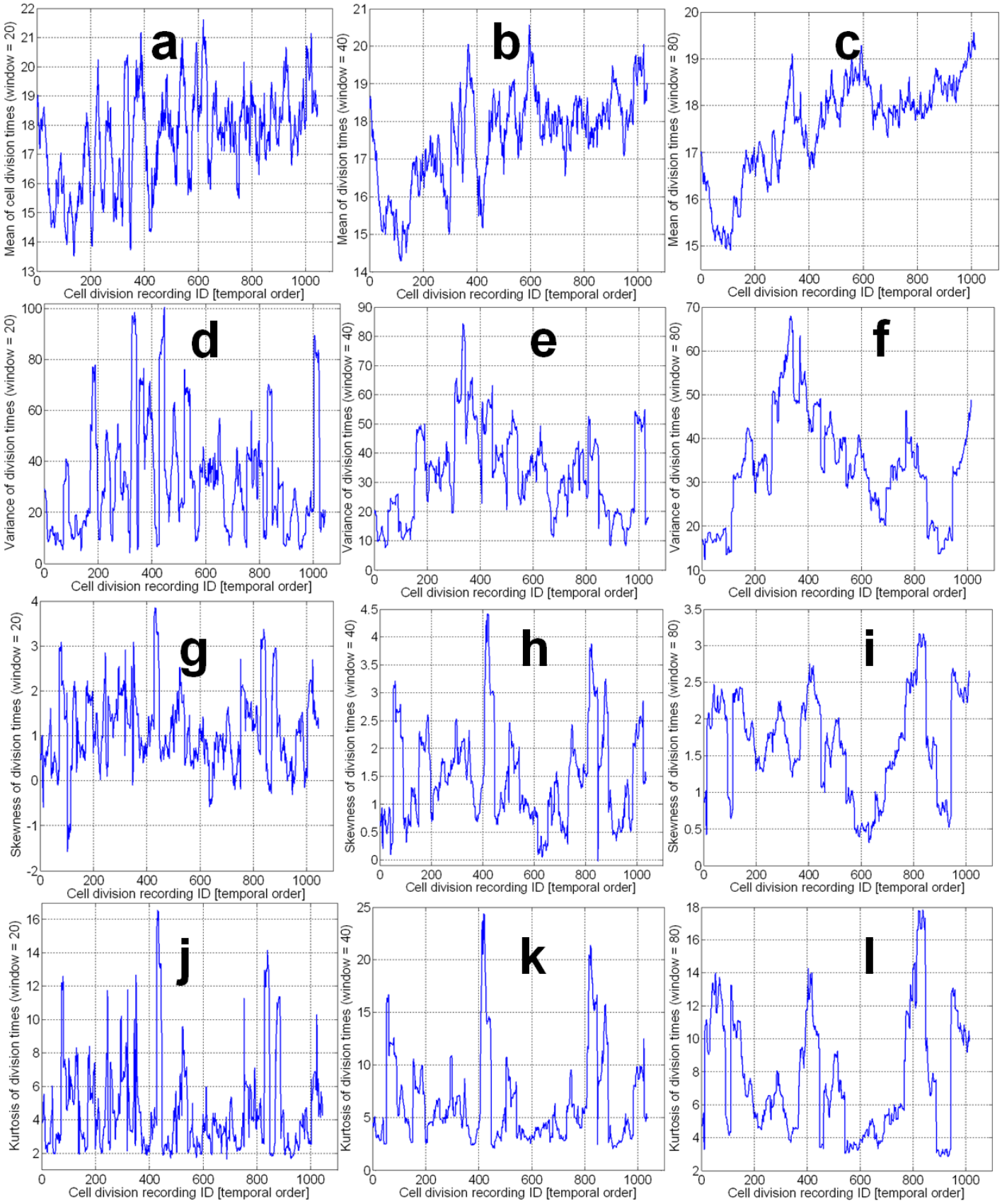


Figure 7. Mean of rat muscle stem cell division times computed as a function of time series and over a moving window of 20 (**a**), 40 (**b**) and 80 (**c**) samples. Variance of rat cell division times over a moving window of 20 (**d**), 40 (**e**) and 80 (**f**) samples. Skewness of rat celldivision times over a moving window of 20 (**g**), 40 (**h**) and 80 (**i**) samples. Kurtosis of rat cell division times over a moving window of 20 (**j**), 40 (**k**) and 80 (**l**) samples.

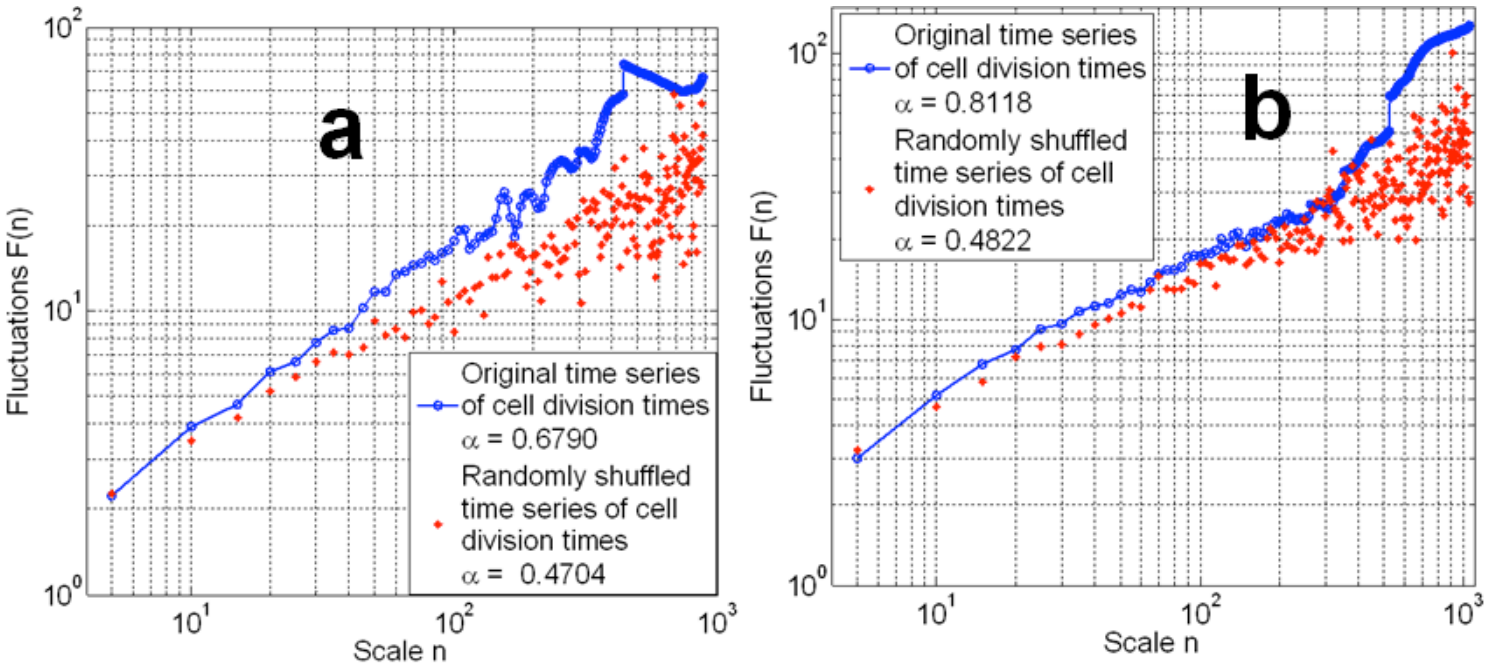


Figure 8.Detrended fluctuation analysis for the mouse (**a**) and rat (**b**) stem cell division time series shows that the cell division process or population growth is not a random process, but rather a positively correlated one. For the description of the detrended fluctuation analysis method the reader is referred to reference [5].

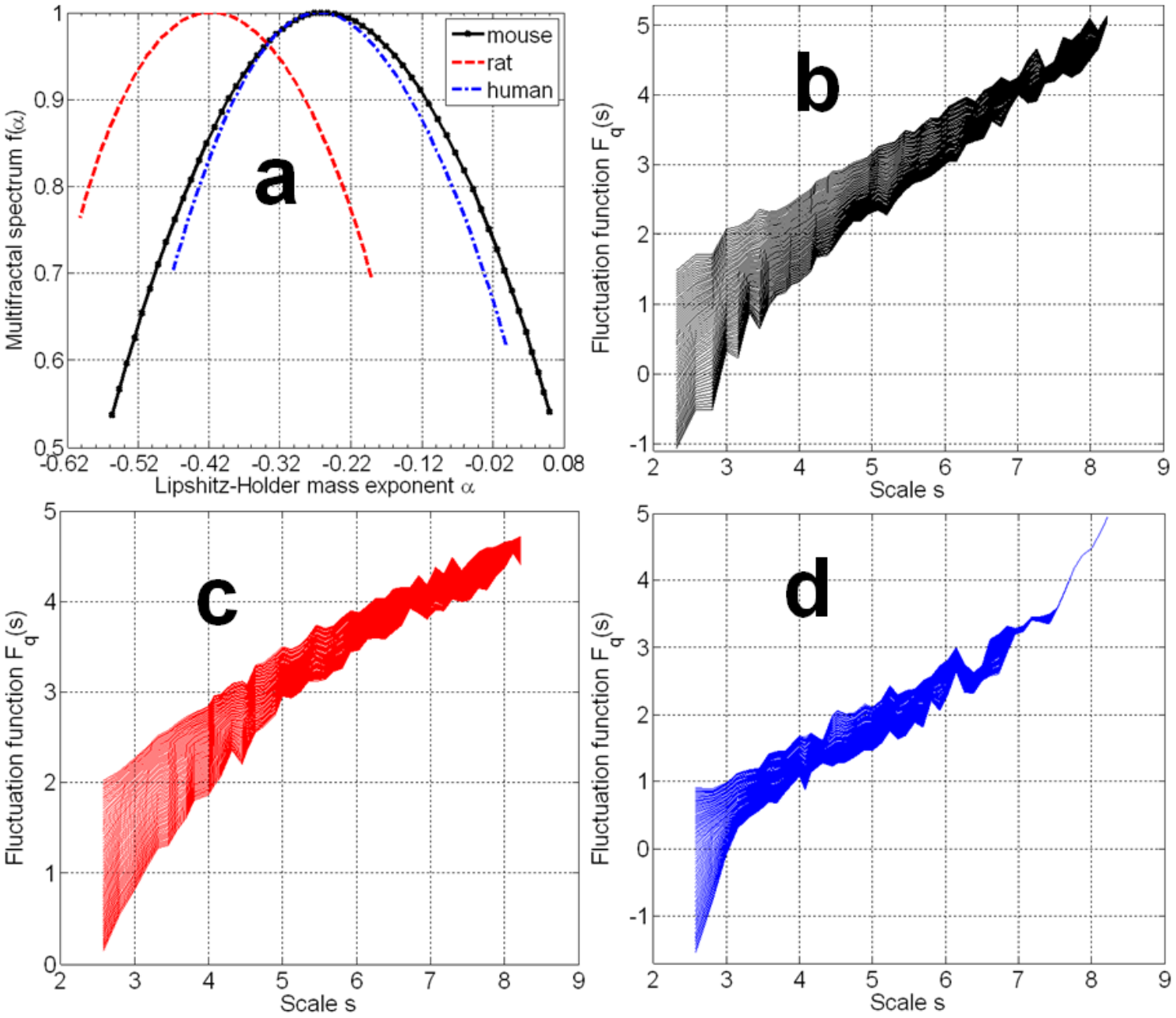


Figure 9.**a)**Multifractal spectrum as a function of the *Lipschitz-Holder* mass exponent for the mouse, rat and human stem cell division times.**b)**Fluctuation function Fq(s) as a function of scale s and for various q-th order moments for the mouse stem cells.**c)**Fluctuation function Fq(s) as a function of scale s and for various q-th order moments for the rat stem cells.**d)**Fluctuation function Fq(s) as a function of scale s and for various q-th order moments for thehuman MSCs. The fluctuations Fq(s) are estimated using the multifractaldetrended fluctuation analysis algorithm proposed in [1].

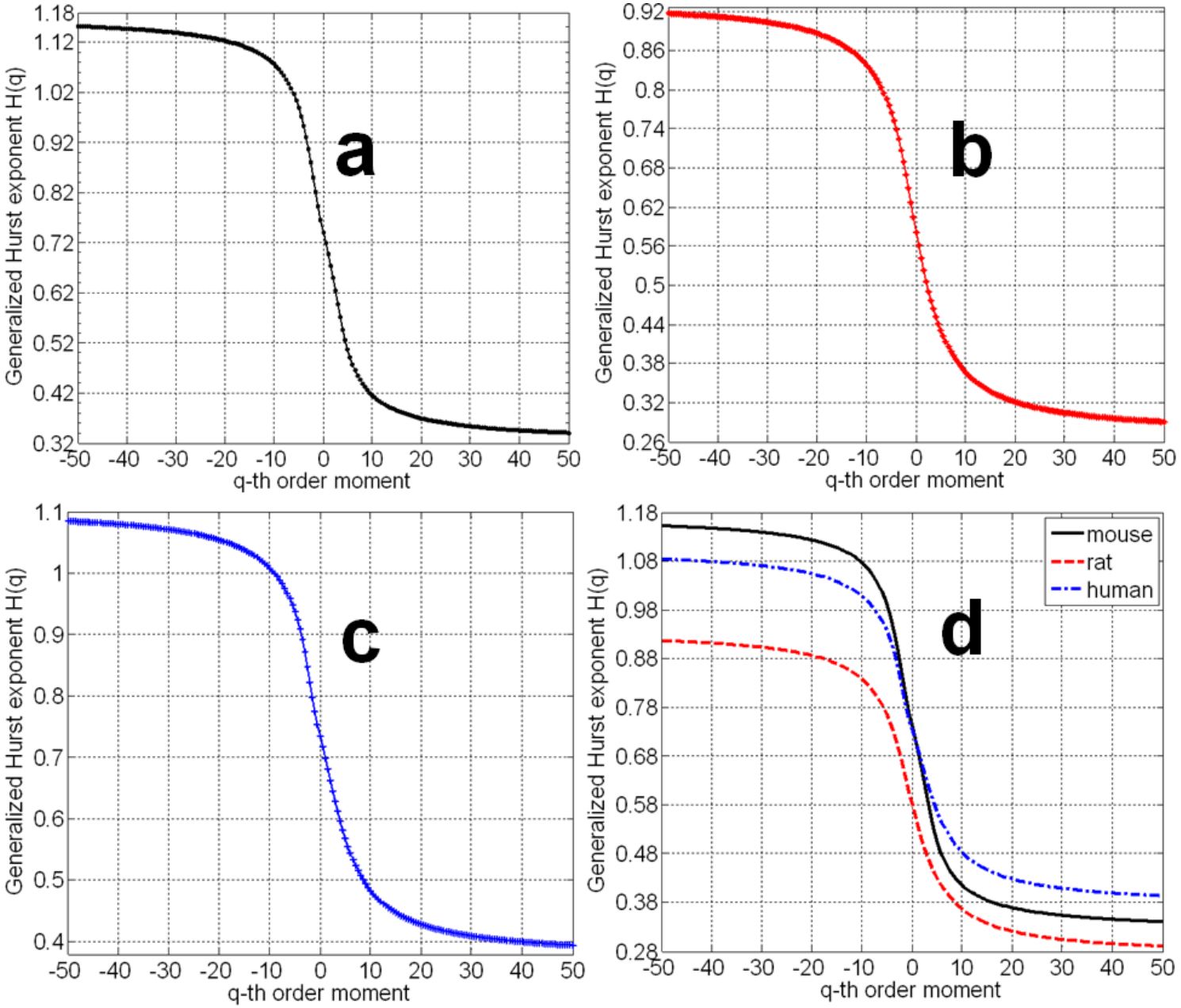


Figure 10.**a)** Generalized Hurst exponent as a function of the q-th order moment for the mouse stem cell division times.**b)** Generalized Hurst exponent as a function of the q-th order moment for the rat stem cells. **c)** Generalized Hurst exponent as a function of the q-th order moment for the human stem cell division times. **d)** Generalized Hurst exponent as a function of the q-th order moment for the mouse, rat, and human stem cell division times. The generalized Hurst exponent was computed using the algorithm in [1].

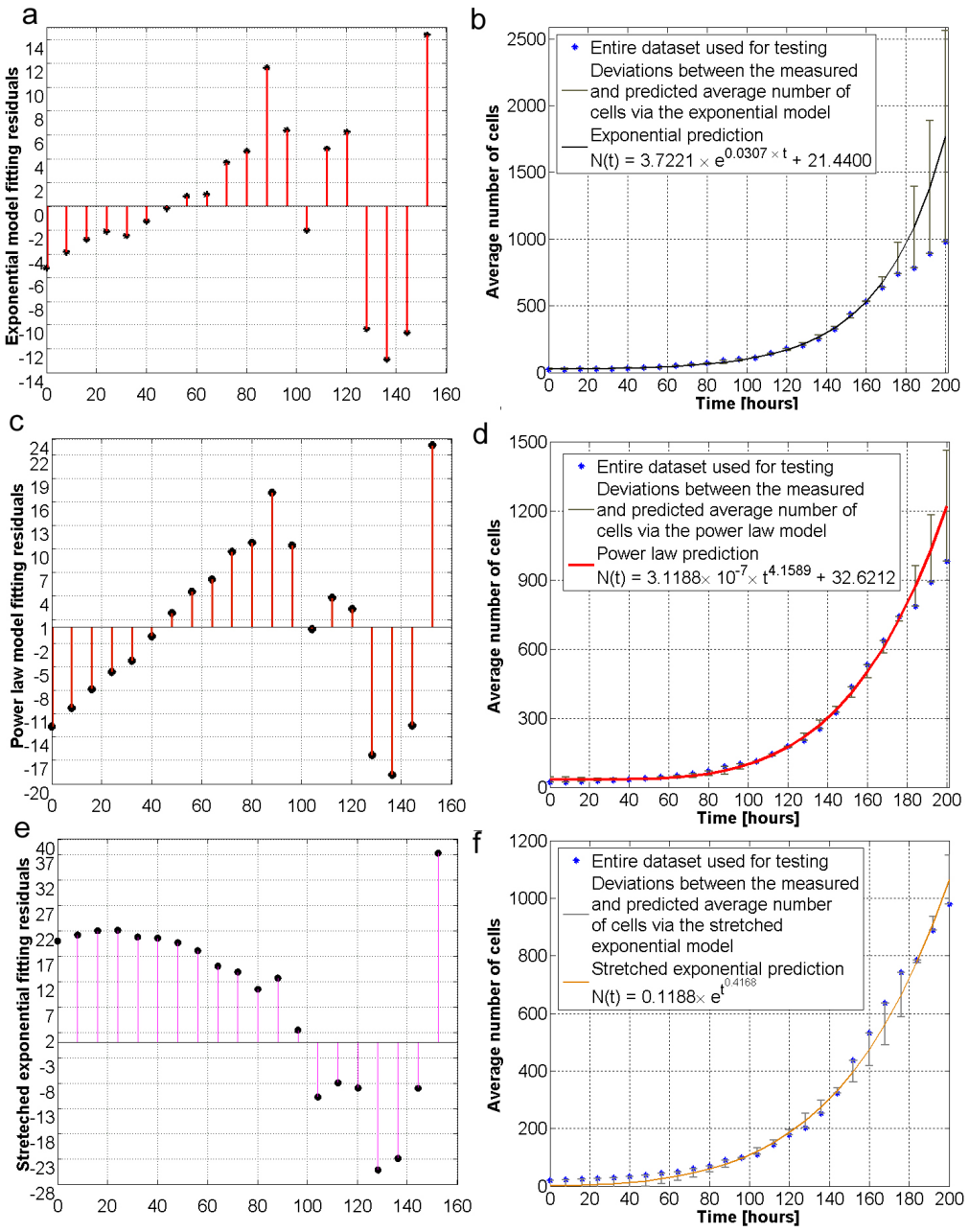


Figure 11. Empirical average number of cells as a function of time (blue stars) and three investigated growth models- (a,b) exponential (c,d) power law and (d,e) stretched exponential. Model fitting residuals are shown in a,c,e and fitted curves with standard errors are shown in b,d,e.The experimental setup is as follows: we collected data samples over a 10 day period, from 8 independent experimental setups, and acquired images at 10-minute intervals. Every 3 hours, we quantified the number of cells to generate growth curves. Once we have obtained distinct "trajectories" of the number of stem cells as a function of time, we have considered the first 75% of the data points (~150 hours) for fitting a model (i.e., an exponential, a power law, and a stretched exponential growth) and use the remaining 25% of the data (~50 hours) to test the validity of the predictions for each of the above-mentioned models. As one can notice from the above plot, the experimental measurements are better fitted by power law and stretched exponential rather than the exponential model. This is in agreement with our multi-fractal analysis which predicts that the intrinsic growth of stem cell population is characterized by non-exponential features.

Table 4: Approximate mean cell division time and fraction of each subpopulation of stem cells from mouse, rat and human data sets.

	Adult, Mouse muscle derived stem cell(MDSC) cell cycle time (fraction in subpopulation)	Adult, Rat muscle derived stem sell (MDSC) cell cycle time (fraction in subpopulation)	Newborn, Human mesenchymal stem cell (MSC) cell cycle time (fraction in subpopulation)
Subpopulation #1	13.14 hrs 0.626	11.7 hrs 0.168	12.8 hrs 0.27
Subpopulation #2	19.14 hrs 0.374	21.2 hrs 0.832	16.8 hrs 0.73

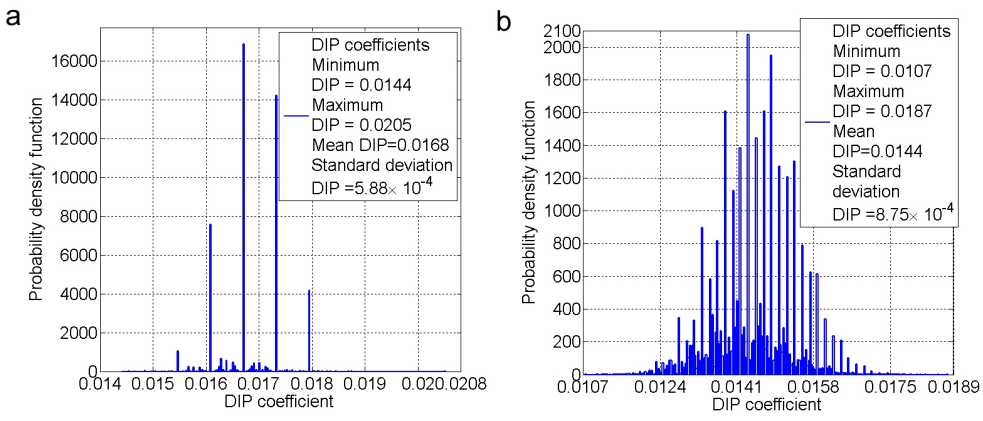
Note 4.Uni-modalgoodness-of-fit analysis via Hartigan's Dip Test.

We have also computed the Hartigan's Dip test[1]. The main idea of the Hartigan's Dip test is to measure the deviation of an empirical probability density function from the best fitting of a uni-modal model. If the empirical probability density function has a single peak and is uni-modal, then the Hartigan's Dip test is zero. Positive values of the Hartigan's Dip coefficient provide reasons to reject the null hypothesis that the empirical dataset is fitted by a uni-modal distribution. In essence, the Hartigan's Dip test performs a null-hypothesis that the samples come from anuni-modal. In addition to the Hartigan's Dip coefficient one could also compute the p-value for the null-hypothesis. If the p-value is smaller than 0.05 then the data has strong bimodality or multimodality features. If the p-value is greater than 0.05 but smaller than 0.10 then the data come from a bimodal distribution with a marginal significance. If the p-value is close to 1 then the data clearly comes from anuni-modal distribution. The Hartigan's DIP coefficient is obtained via bootstrapping selecting 90% of the samples for each iteration, computing the DIP coefficient and repeating this process for 400,000 times. The p-value is also computed via a bootstrapping method over 1,000,000 iterations. The Hartigan's Dip coefficient and p-values are summarized below.

Table 5: Hartigan's Dip test results for uni-modal distribution goodness-of-fit analysis.

Dataset	Hartigan's Dip coefficient	Hartigan's p-value
Mouse muscle stem cells (808 samples)	$1.68\times10^{-2}\pm0.05\times10^{-2}$	0.07
Rat muscle stem cells (975 samples)	$1.44\times10^{-2}\pm0.08\times10^{-2}$	0.08
Human mesenchymal stem cells (342 samples)	$2.14\times10^{-2}\pm0.01\times10^{-2}$	0.36

Figure 12. The probability density function of the Hartigan DIP coefficient estimated via a bootstrapping approach for the mouse (a) and rat (b) stem cell data sets. The minimum, maximum, mean and standard deviation of the Hartigan DIP coefficients are summarized in the legends of the two plots.



Note 5.Goodness-of-fit analysis for bi-modal Gaussian distribution.

To investigate whether the mouse, rat, human and stem cellsdivision times can be well characterized by uni- or bi-modal distributions, we performed the following steps:

- Estimate the parameters of an uni-Gaussian distribution via maximum likelihood estimation method and computed the Kolmogorov-Smirnov (K-S) test and the Akaike Information Criterion (AIC) to determine the goodness of fit.
- Estimate the parameters of a bi-modal Gaussian distribution via maximum likelihood estimation method with random initial conditions. We repeated the MLE estimation process for 10^9 times and record only the parameters for which the p-value of the Kolmogorov-Smirnov (K-S) exceeded the significance level (we used throughout the experiments 0.05 statistical significance value). More precisely, we performed a binary hypothesis testing problem characterized by two parameters: h and p-value. When $h = 1$ we reject the null hypothesis testing at predefined statistical significance level. When $h = 0$ we cannot reject the null hypothesis testing for the considered statistical significance level. The decision to reject the null hypothesis occurs when the considered significance level equals or exceeds the p-value. From the recorded values for each parameter of the bi-modal Gaussian distribution we computed the mean and standard deviation values.

The results concerning the appropriateness of modeling mouse, rat and human stem celldivision times are summarized below:

Table 6: Goodness-of-fit results for uni- and bi-model Gaussian distributions

Dataset		Mouse muscle stem cells	Rat musclestem cells	Human mesenchymal stem cells
Parameters of Uni-modal Gaussian distribution $N(\mu_1, \sigma_1)$		$\mu = 15.60 \pm 0.14$ $\sigma = 4.39 \pm 0.10$	$\mu = 17.57 \pm 0.17$ $\sigma = 5.81 \pm 0.12$	$\mu = 15.54 \pm 0.14$ $\sigma = 2.63 \pm 0.10$
Kolmogorov-Smirnov (KS) Goodness-of-fit	h	1	1	1
	p-value	0.29×10^{-15}	0.86×10^{-17}	0.29×10^{-3}
Akaike Information Criterion (AIC)		5157.10	6709.80	1635.31
Parameters of bi-modal Gaussian distribution $a*N(\mu_1, \sigma_1) + (1-a)*N(\mu_2, \sigma_2)$		$a = 0.66 \pm 0.04$ $\mu_1 = 13.77 \pm 0.15$ $\sigma_1 = 1.85 \pm 0.15$ $\mu_2 = 19.14 \pm 0.51$ $\sigma_2 = 5.14 \pm 0.46$	$a = 0.72 \pm 0.01$ $\mu_1 = 15.39 \pm 0.10$ $\sigma_1 = 2.85 \pm 0.08$ $\mu_2 = 23.32 \pm 0.40$ $\sigma_2 = 7.46 \pm 0.26$	$a = 0.87 \pm 0.10$ $\mu_1 = 15.08 \pm 0.21$ $\sigma_1 = 2.07 \pm 0.13$ $\mu_2 = 19.34 \pm 1.23$ $\sigma_2 = 3.94 \pm 0.94$
Kolmogorov-Smirnov (KS) Goodness-of-fit	h	0	0	**0**
	p-value	0.09	0.27	0.69
Akaike Information Criterion (AIC)		4790.37	6327.40	1603.45

Note 6.Goodness-of-fit analysis for uni- and bi-modal alpha-stable distributions.

To investigate whether the mouse, rat and human stem cell division times can be well characterized by uni- or bi-modal distributions, we performed the following steps:

- Estimate the parameters of anuni-modal α-stable distribution via maximum likelihood estimation method and computed the K-S test and the Akaike Information Criterion (AIC) to determine the goodness of fit.
- Estimate the parameters of a bi-modal α-stable distribution via a greedy Monte-Carlo approachwhich records the parameters for which the p-value of the KS test is the highest. The greedy Monte-Carlo approachperformsapproximately 10^7iterations in order to find the parameters for which the p-value of the Kolmogorov-Smirnov (K-S) exceeded the significance level (we used throughout the experiments 0.05 statistical significance value). More precisely, we performed a binary hypothesis testing problem characterized by two parameters: *h* and *p*-value. When $h = 1$ we reject the null hypothesis testing at predefined statistical significance level. When $h = 0$ we cannot reject the null hypothesis testing for the considered statistical significance level. The decision to reject the null hypothesis occurs when the considered significance level equals or exceeds the *p*-value. From the recorded values for each parameter of the bi-modal alpha-stable distribution we computed the mean and standard deviation values. The results concerning the appropriateness of bi-modal alpha stable distribution modeling are summarized below:

Table 7: Goodness-of-fit results for uni- and bi-modal alpha stable distributions

Dataset		Mouse muscle stem cells	Rat muscle stem cells	Human mesenchymal stem cells
Parameters of Uni-modal α-stable distribution $\Phi(\alpha, \beta,\gamma,\delta)$		$\alpha = 1.49$ $\beta = 1$ $\gamma = 1.93$ $\delta = 16.37$	$\alpha = 1.48$ $\beta = 1$ $\gamma = 2.71$ $\delta = 18.70$	$\alpha = 1.85$ $\beta = 0.99$ $\gamma = 1.63$ $\delta = 15.60$
Kolmogorov-Smirnov (KS) Goodness-of-fit	h	1	1	0
	p-value	2.10×10^{-4}	44.52×10^{-4}	5239.1×10^{-4}
Akaike Information Criterion (AIC)		1831.43	1586.63	2120.46
Parameters of bi-modal α-stable distribution $a^* \Phi(\alpha_1, \beta_1,\gamma_1,\delta_1) + (1-a)^* \Phi(\alpha_2, \beta_2,\gamma_2,\delta_2)$		$a = 0.626$ $\alpha_1 = 1.89$ $\beta_1 = -1.0$ $\gamma_1 = 1.03$ $\delta_1 = 13.14$ $\alpha_2 = 1.21$ $\beta_2 = 1.0$ $\gamma_2 = 1.45$ $\delta_2 = 21.83$	$a = 0.1678$ $\alpha_1 = 1.67$ $\beta_1 = -1.0$ $\gamma_1 = 0.62$ $\delta_1 = 11.69$ $\alpha_2 = 1.24$ $\beta_2 = 1.0$ $\gamma_2 = 2.02$ $\delta_2 = 21.19$	$a = 0.269$ $\alpha_1 = 1.99$ $\beta_1 = -1.0$ $\gamma_1 = 0.65$ $\delta_1 = 12.80$ $\alpha_2 = 1.55$ $\beta_2 = 1.0$ $\gamma_2 = 1.15$ $\delta_2 = 16.80$
Kolmogorov-Smirnov (KS) Goodness-of-fit	h	0	0	0
	p-value	0.78	0.11	0.81
Akaike Information Criterion (AIC)		1626.37	1466.89	1886.37

Note 7

Mouse and rat muscle stem cells were isolated by way of differential adhesion rates using a modification of methods as previously described [6,7]. Skeletal muscle was enzymatically digested using collagenase and dispase and serial pre-plating was performed to separate cell fractions preplate 1 (PP1), preplate 2 (PP2), preplate 3 (PP3), throough to preplate 6 (PP6). Cells which adhere to non-collegenated tissue culture plastic flasks within the first 30 minutes were PP1 cells. Medium and non-adhered cells from the PP1 flasks were then transferred to a fresh flask, and cells which adhered during 30 to 60 minutes were termed PP2 flasks. Again, media and non-adhered cells were transferred to fresh flasks and allowed to adhere such that PP3 cells were those that adhered 1 to 2 hours post plating. PP4 cells adhered after 2 to 24 hours, PP5 cells adhered between 24 to 48 hours. PP6 cels adhered between 48 to 120 hours. During the isolation process, cells were grown in Dulbecco's Modified Eagle Medium DMEM (Gibco) with 20% fetal bovine serum and 0.5% Chick Embryo Extract. In this study, we used mouse (PP2-PP3) and rat (PP6) cells, both of which have reported stem cell characteristics. Human mesenchymal stem cells were isolated from the human newborn umbilical cord. Full-term umbilical cords were received from Magee Women's Hospital (Pittsburgh, PA) under IRB No. 0606126. Whole UCs were first manually dissected into smaller sections then placed into collagenase (Sigma, 1mg/mL in PBS) enzyme and digested at 37 C for 6-18 hours with occasional shaking. Following digestion, the isolates were separated from the remaining tissue and washed to prevent further maceration by enzyme. A portion of freshly isolated cells was removed for flow cytometry analysis and the remainder was placed in culture; these methods are described in detail in [8].

References

[1] J.A. HartiganandP.M. Hartigan, “The dip test of unimodality,” Ann. Statist. 13, 70-84 (1985).
[2] J.W. Kantelhardt, S.A. Zschiegner, E. Koscielny-Bunde, S. Havlin, A. Bunde, H.E. Stanley, “Multifractaldetrended fluctuation analysis of nonstationary time series,” Physica A: Statistical Mechanics and its Applications, Volume 316, Issues 1–4, 15 December 2002, Pages 87-114.
[3] W.L. Martinez and A.R. Martinez, Computational Statistics Handbook with MATLAB, CRC Press, 2001.
[4] Matlab - The Language of Technical Computing, Mathworks Inc. homepage. www.mathworks.com
[5] C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger, “Mosaic Organization of DNA Nucleotides,”*Phys. Rev. E* 49, 1685 – Published 1 February 1994.
[6] T.A. Rando, and Blau, H.M. “Primary mouse myoblast purification, characterization, and transplantation for cell-mediated gene therapy,” J Cell Biol. 125, 1275-1287 (1994).
[7] B. Gharaibeh, A. Lu, J. Tebbets, B. Zheng, J.M. Feduska, M. Crisan, B. Peault, J. Cummins, and J. Huard, “Isolation of a slowly adhering cell fraction containing stem cells from murine skeletal muscle by the preplate technique,” Nat Protoc. 3, 1501-1509 (2008)
[8] R.C Schugar, S.M. Chirieleison, K.E. Wescoe, B.T. Schmidt, Y. Askew, J.J. Nance, J.M. Evron, B. Peault, and B.M. Deasy, “High harvest yield, high expansion, and phenotype stability of CD146 mesenchymal stromal cells from whole primitive human umbilical cord tissue,” J Biomed Biotechnol 2009:789526 (2009)