BMC Research Notes

Technical Note

Genome-wide screens for effective siRNAs through assessing the size of siRNA effects
Xiaohua Douglas Zhang

Address: Biometrics Research, Merck Research Laboratories, West Point, PA 19486, USA

Email: Xiaohua Douglas Zhang - xiaohua_zhang@merck.com

Abstract

Background: RNA interference (RNAi) has been seen as a revolution in functional genomics and system biology. Genome-wide RNAi research relies on the development of RNAi high-throughput screening (HTS) assays. One of the most fundamental challenges in RNAi HTS is to glean biological significance from mounds of data, which relies on the development of effective analytic methods for selecting effective small interfering RNAs (siRNAs).

Findings: Based on a recently proposed parameter, strictly standardized mean difference (SSMD), I propose an analytic method for genome-wide screens of effective siRNAs through assessing and testing the size of siRNA effects. Central to this method is the capability of SSMD in quantifying siRNA effects. This method has relied on normal approximation, which works only in the primary screens but not in the confirmatory screens. In this paper, I explore the non-central *t*-distribution property of SSMD estimates and use this property to extend the SSMD-based method so that it works effectively in either primary or confirmatory screens as well as in any HTS screens with or without replicates. The SSMD-based method maintains a balanced control of false positives and false negatives.

Conclusion: The central interest in genome-wide RNAi research is the selection of effective siRNAs which relies on the development of analytic methods to measure the size of siRNA effects. The new analytic method for hit selection provided in this paper offers a good analytic tool for selecting effective siRNAs, better than current analytic methods, and thus may have broad utility in genome-wide RNAi research.

Findings
Background
Mean difference, fold change, percent inhibition, percent activity, percent viability, *Z*-score and their robust versions have been used to quantify effect size of an siRNA or a compound in HTS assays [1-7]. However, these metrics have issues in capturing data variability or being affected by sample size and hence cannot effectively assess the size of effect. The *p*-values from the *Z*-score method (or equivalently Mean ± *k* SD and its variant Median ± *k* MAD) and

classical *t*-test have widely been used to evaluate the chance of including siRNAs with no specific impact [1,2,5-7]. However, it is mean difference that these methods aim to test, and it is well-known that mean difference cannot effectively measure the magnitude of impact. In addition, the *p*-value from the *Z*-score method or *t*-test is affected by both sample size and the size of siRNA effect.

A recently proposed parameter, strictly standardized mean difference (SSMD) [8], measures the magnitude of impact

more effectively than any other currently used metrics. SSMD has been applied for quality control in genome-scale RNAi research [8-10]. Utilizing the fact that SSMD effectively measures the size of effect, Zhang proposes an SSMD-based hit selection method to maintain a balanced control of both FPR and FNR [11]. This method has also been applied to select hits in RNAi HTS primary experiments [12]. However, this method is based on normal approximation, which works only in the primary screens but not in the confirmatory screens. Here I construct a new analytic method for hit selection in HTS assays using non-central *t*-distribution property of SSMD estimates. This method works effectively whether sample size is small or large.

Issues of hit selection methods in primary screens

The size of siRNA effect is commonly assessed using percent inhibition/activation (i.e., $\frac{y - \bar{X}_-}{\bar{X}_+ - \bar{X}_-} \times 100$ where y is the measured intensity of an siRNA, $\bar{X}_+$ is the average intensity of a positive control and $\bar{X}_-$ the average intensity of a negative reference) and percent viability/activity (i.e., $\frac{y}{\bar{X}_-} \times 100$) or fold change (i.e., $\frac{y}{\bar{X}_-}$). Another commonly used method for hit selection in a primary HTS experiment is the *Z*-score method (i.e., $\frac{y - \bar{X}_-}{s_-}$ where s_- is sample standard deviation of a negative reference) along with its variants such as median ± *k* MAD method. The issues of *Z*-score method and its variants have been illustrated in [11,12]. The issues of percent inhibition and percent viability are illustrated in Figure 1. Here we consider the situation where the knockdown of a gene inhibits cell growth. It is clear that the magnitude of difference between the sample siRNA (represented by the black curve) and the negative control (represented by the green curve) is much less in Plate D than in Plates E and F. That is, the siRNA in Plate D has less inhibition effect than in Plates E and F. However, if using percent inhibition, we would conclude that the inhibition effect of the siRNA in Plate D (which has a percent inhibition of 27.3) is larger than the effect of the siRNA in Plate E (which has a percent inhibition of 20). If using percent viability, we would conclude that the inhibition effect of the siRNA in Plate D (which has a percent viability of 83.3) is larger than the effect of the siRNA in Plate E (which has a percent viability of 93.8). Therefore, both percent inhibition and percent viability produce misleading results.

Issues of hit selection methods in screens with replicates

In all confirmatory HTS screens and some primary screens, there are several sets of source plates. Each set is unique and has replicates (usually triplicates), thus each siRNA has replicates. Because plate-to-plate variability is usually higher than within-plate variability, a paired *t*-test is often used for hit selection in a confirmatory screen. That is, for each siRNA, we calculate the difference between the measured intensity of the siRNA and average intensity of a negative control in a plate, then calculate the corresponding *p*-value of the paired *t*-test in which the null hypothesis of zero mean difference is tested.

The strength of siRNA impact in a screen with replicates is represented by the magnitude of a paired difference between the measured intensity of an siRNA and average intensity of a negative reference. For one siRNA, a good metric for the assessment of siRNA effect should have one fixed population value and should have estimated values distributed around this population value if it is a statistical parameter. If it is not a statistical parameter, a good metric should have values distributed around a fixed value that can indicate effect size of the siRNA. The *t*-value, *Z*-score and their corresponding *p*-values are not statistical parameters. The *t*-values of the samples from siRNA A are not distributed around a fixed value, and actually go to infinity as sample size increases (blue points in Panel A3 of Figure 2). The corresponding *p*-values go to zero and thus cannot indicate the effect size of the corresponding siRNA (blue points in Panel A4). A similar situation occurs for the *t*-values and *p*-values corresponding to siRNA B (blue points in Panels B3 and B4 of Figure 2).

It is clear that the magnitude of paired difference for siRNA A is much larger than for siRNA B although the mean of the paired difference in Panel A1 (i.e., 2.5) is smaller than that in Panel B1 (i.e., 3). The black points in Panels A2 and B2 (i.e., random draws from the populations in Panels A1 and B1 respectively) also demonstrate that the magnitude of the paired difference in Panel A1 is larger than that in Panel B1. Therefore, a good metric for the assessment of siRNA impact should have a larger (or smaller) value for siRNA A than the value for siRNA B in the case where a large (or small) value of this metric indicates a large effect size. The *t*-value and *p*-value are both affected by sample size; thus we may obtain a larger *p*-value (or smaller *t*-value) corresponding to the samples from siRNA A than from siRNA B. For example, the *p*-values corresponding to the samples with 2 or 3 replicates from siRNA A are larger than the *p*-values corresponding to the samples with at least 10 replicates from siRNA B (Panels A4 and B4 of Figure 2). Therefore, the *p*-value from *t*-test or *Z*-score method cannot effectively measure the strength of siRNA impact.
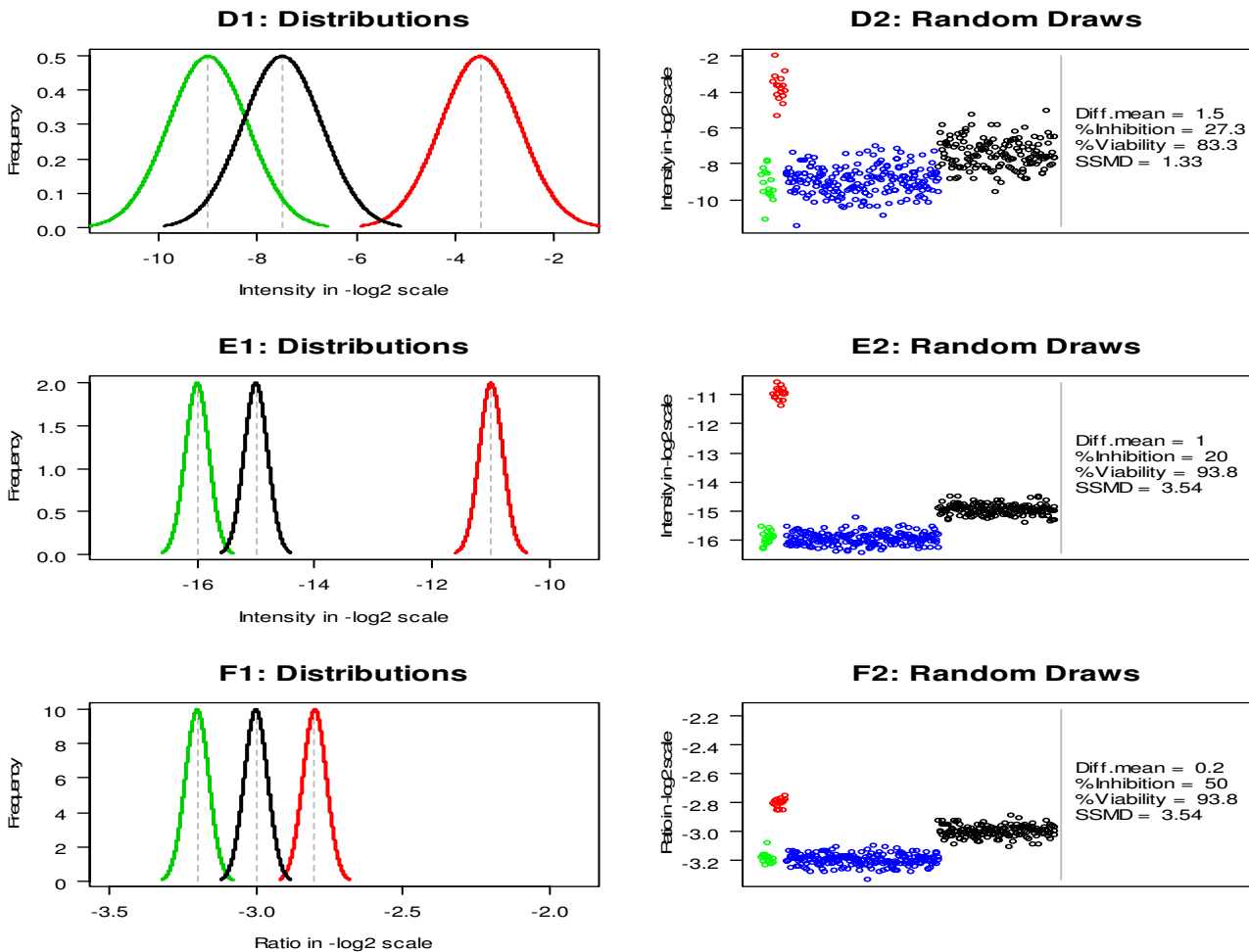
Figure 1
Comparison of mean difference (Diff.mean), percent inhibition (% Inhibition), percent viability (% Viability) and SSMD in three plates D, E and F. The population distributions of a positive control (red curve), a sample siRNA (black curve) and a negative control (green curve) in Plates D, E and F are displayed in Panels D1, E1 and F1 respectively. Panels D2, E2 and F2 show 16 random draws from the positive control (red points), 16 draws from the negative control (green points), another 200 draws from the negative control (blue points) and 152 draws from the sample siRNA (black points) in Plates D, E and F respectively.

Assessment of siRNA effects using SSMD
SSMD is a statistical parameter that measures the magnitude of both paired and unpaired differences and thus can be used to measure the magnitude of impact of siRNAs in both primary and confirmatory screens. For example, the values of *SSMD* between the siRNA and the negative control are 1.33, 3.54 and 3.54 in Plate D, E and F respectively, which appropriately indicates that the effect of the siRNA in Plate D is less than in Plates E and F and that the effect of the siRNA in Plate E is the same as in Plate F (Figure 1). The population values of SSMD for siRNA A and siRNA B are 3.54 and 0.71 respectively (Figure 2). The estimated SSMD values (denoted by the green points) all fall around the population values of SSMD (denoted by the green points)

orange lines) and do not have an increasing trend as sample size increases (Panels A3 and B3 of Figure 2). All these results indicate that SSMD appropriately indicates the effect size of an siRNA, better than percent inhibition/viability and *p*-value from *t*-test of testing no mean difference.

Based on both original and probability meanings of SSMD, an SSMD-based 1-2-3 rule [11], along with its extended version, has been proposed for classifying siRNA impact. The SSMD-based 1-2-3 rules provide a guideline for classifying the strength of siRNA impact. For example, in Figure 1, the siRNA in Plate D is classified as "moderate inhibition effect" and the siRNAs in Plates E and F are
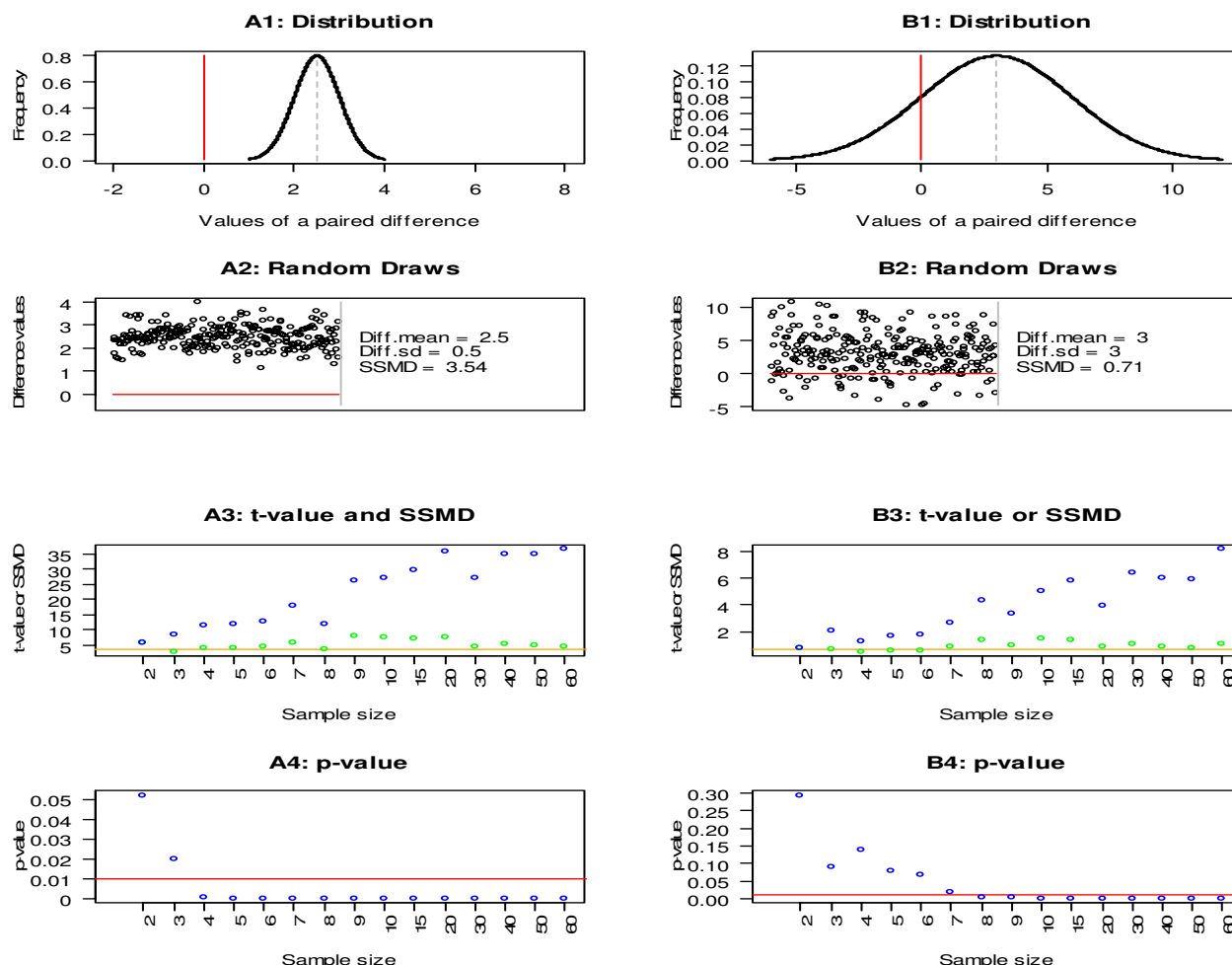
Figure 2
SSMD, Z-score and p-value of random samples of two paired differences corresponding to two siRNAs, siRNA A and siRNA B, respectively. Panels A1–A4 correspond to siRNA A and Panels B1–B4 correspond to siRNA B. Panels A1 and B1 display the population distributions of paired differences for siRNAs A and B respectively. Panels A2 and B2 show the appearance of random samples of the paired differences for siRNAs A and B respectively. "Diff.mean" and "Diff.sd" denote the mean and standard deviation of a paired difference respectively. In Panels A3 and B3, the blue (or green) points denote the *t*-values (or SSMD estimated values) of samples from siRNA A and siRNA B respectively; the orange lines denote the SSMD population values. In Panels A4 and B4, the blue points denote the *p*-values from t-test of testing paired difference mean and the red lines denote the cutoff of 0.01.

both classified as "strong inhibition effect". In Figure 2, siRNAs A and B are classified as "strong inhibition effect" and "weak inhibition effect" respectively. The 1-2-3 rule and extended 1-2-3 rule work in the situation where the population value of SSMD is known; they also work reasonably when sample size is large. In practice, the population value of SSMD is unknown and sample size is small especially in confirmatory RNAi HTS experiments. In such a case, we can provide a point estimate and a confidence interval of SSMD for each siRNA based on its estimated SSMD value [see additional file 1].

A balanced control of false positives and false negatives
Based on SSMD, we may maintain a flexible and balanced control of both the false negative rate (FNR), in which the siRNAs with strong effects are not selected as hits, and the restricted false positive rate (RFPR), in which the siRNAs with weak or no effects are selected as hits. The maximum RFPR and FNR in a decision rule are called restricted false positive level (RFPL) and false negative level (FNL), respectively. To use the SSMD-based method for selecting hits in the direction of positive values in HTS assays, we need to search for a cutoff β^* for the estimated SSMD so

that we can maintain a balanced control of both RFPR and FNR when we use the decision rule of declaring an siRNA as a hit if it has $\hat{\beta} \geq \beta^*$ and as a non-hit otherwise. The search of a cutoff can be achieved through error-cutoff plots in which we plot RFPLs and FNLs versus cutoffs of estimated SSDM values, as shown in Figure 3.

Based on Table 1, we calculate RFPL and FNL corresponding to each set of values for β^*, c_1 and c_2. In a primary screen, the majority of sample wells may be used as the negative reference in a plate. In such a case, n_2 is about 300. To select hits in the direction of positive values, we calculate RFPLs with respect to (w.r.t.) $c_2 = 0, 0.25, 0.5, 1$

and FNLs w.r.t. $c_1 = 2, 3, 4.5$ (Panel A1 of Figure 3). The commonly used error rates are 0.05, 0.025 and 0.01 in one direction. From Panel A1 of Figure 3, a cutoff between 1.4 and 1.8 can control RFPL w.r.t. $c_2 = 0$ to be less than 0.025, RFPL w.r.t. $c_2 = 0.25$ to be less than 0.051 and FNL w.r.t. $c_1 = 3$ to be less than 0.05. A cutoff between 1.9 and 2.1 can control RFPL w.r.t $c_2 = 0$ to be less than 0.005, RFPL w.r.t $c_2 = 0.25$ to be less than 0.01, and FNL w.r.t. $c_1 = 3$ to be less than 0.10. Therefore, any cutoff between 1.4 and 2.1 for SSMD is theoretically reasonable and maintains a balanced control of RFPR and FNR for selecting hits in the direction of positive values. Similarly, to select hits in the direction of negative values, a reasonable cutoff
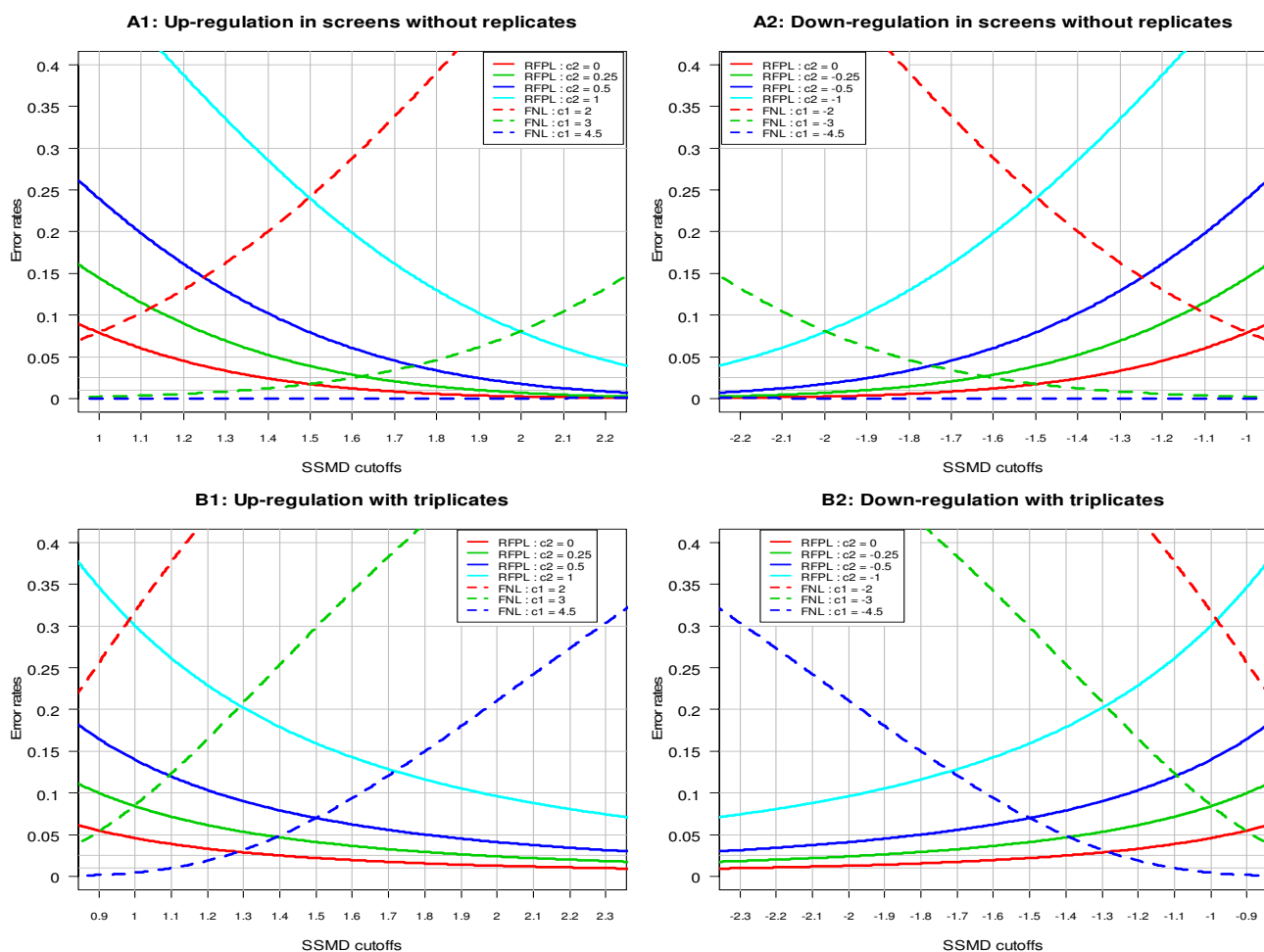

Figure 3
Error-cutoff plots for controlling both restricted false positive level (RFPL) and false negative level (FNL) in RNAi HTS assays. Panels A1 and A2 are for primary screens without replicates and Panels B1 and B2 are for confirmatory screens with triplicates. The unlabeled grey horizontal lines in the bottom of each panel indicate error rates of 0.025 and 0.01, respectively.

Table 1: SSMD-based decision rules and their false negative levels (FNLs) and restricted false positive levels (RFPLs) for hit selection in RNAi HTS experiments

I: Select up-regulated siRNAs ($c_1 \ge c_2 \ge 0$)		
Selection Criterion	FNL	RFPL
Ia: $\hat{\beta} \ge \beta^*$	$F_{t(v,bc_1)}\left(\dfrac{\beta^*}{k}\right)$	$1 - F_{t(v,bc_2)}\left(\dfrac{\beta^*}{k}\right)$
Ib: $\hat{\beta} \ge kQ_{t(v,bc_1)}(\alpha_1)$	α_1	$1 - F_{t(v,bc_2)}\left(Q_{t(v,bc_1)}(\alpha_1)\right)$
Ic: $\hat{\beta} \ge kQ_{t(v,bc_2)}(1-\alpha_2)$	$F_{t(v,bc_1)}\left(Q_{t(v,bc_2)}(1-\alpha_2)\right)$	α_2

II: Select down-regulated siRNAs ($c_1 \le c_2 \le 0$)		
Selection Criterion	FNL	RFPL
IIa: $\hat{\beta} \le \beta^*$	$1 - F_{t(v,bc_1)}\left(\dfrac{\beta^*}{k}\right)$	$F_{t(v,bc_2)}\left(\dfrac{\beta^*}{k}\right)$
IIb: $\hat{\beta} \le kQ_{t(v,bc_1)}(1-\alpha_1)$	α_1	$F_{t(v,bc_2)}\left(Q_{t(v,bc_1)}(1-\alpha_1)\right)$
IIc: $\hat{\beta} \le kQ_{t(v,bc_2)}(\alpha_2)$	$1 - F_{t(v,bc_1)}\left(Q_{t(v,bc_2)}(\alpha_2)\right)$	α_2

Notes:

(i) $\hat{\beta}$ is the estimate of SSMD and β^* is a cutoff of SSMD; $\hat{\beta} = kT$ where T has a noncentral t-distribution with degree of freedom v and non-central parameter $b\beta$, namely $T \sim t(v, b\beta)$; $F_{t(v,b\beta)}(\cdot)$ and $Q_{t(v,b\beta)}(\alpha)$ are the cumulative distribution function and the α quantile of $t(v, b\beta)$ respectively.

(ii) For an unpaired difference, $\hat{\beta} = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\dfrac{2}{K}\left((n_1-1)s_1^2 + (n_2-1)s_2^2\right)}}$ and $T = \dfrac{\left(\bar{X}_1 - \bar{X}_2\right)\Big/\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}{\sqrt{\left((n_1-1)s_1^2 + (n_2-1)s_2^2\right)\Big/(N-2)}}$ where

$K = 2 \cdot \left(\dfrac{\Gamma(\frac{N-2}{2})}{\Gamma(\frac{N-3}{2})}\right)^2 \approx N - 3.5$, $N = n_1 + n_2$, and n_1, $\bar{X}_1$, s_1, n_2, $\bar{X}_2$, s_2 are sample size, mean and standard deviation in two groups

respectively; $k = \sqrt{\dfrac{K}{2(N-2)}\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}$, $v = N - 2$, $b = \dfrac{\sqrt{2}}{\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}$.

(iii) For a paired difference, $\hat{\beta} = \dfrac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})}\sqrt{\dfrac{2}{n-1}}\dfrac{\bar{D}}{s_D}$ and $T = \dfrac{\sqrt{n}\bar{D}}{s_D}$ where n, $\bar{D}$ and s_D are sample size, sample mean and standard deviation of

a paired difference respectively; $k = \dfrac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n-2}{2})}\sqrt{\dfrac{2}{n(n-1)}}$, $v = n - 1$, $b = \sqrt{n}$.

is between -2.1 and -1.4 when we use the decision rule of declaring an siRNA as a hit if it has $\hat{\beta} \le \beta^*$ and as a non-hit otherwise (Panel A2 of Figure 3). Similarly from Panels B1 and B2 of Figure 3, in confirmatory screens with triplicates, a reasonable cutoff is between 1 and 1.4.

The choice of an exact cutoff between 1.4 and 2.1 (or between -2.1 and -1.4) in a real primary experiment relies on the refined tolerance of false positives and false negatives and the capacity of follow-up studies after that experiment. For example, if one has a low tolerance in missing hits with SSMD greater than 2 or 3 (or less than -2 or -3), one may choose a cutoff between 1.4 and 1.6 (or between -1.6 and -1.4). On the other hand, if follow-up studies have a low capacity of including selected hits, one may choose a cutoff between 1.8 and 2.1 (or between -2.1 and -1.8). These cutoffs may maintain a balanced control of

both RFPR for including siRNAs with weak or no effects and FNR for excluding siRNAs with strong effects.

Discussion

SSMD is usually applied to the measured intensity of each siRNA individually. In some screens, there may be a need to pool multiple measured values to a single value. For example, in the situations where there are two or more wells for each siRNA in a plate, we may use the mean or median of these replicates to represent the measured intensity of this siRNA. In screens where multiple siRNAs are designed to target the same gene to account for off-target effects, there may be a need to pool information across these siRNAs to form a single value for a gene. In those situations, SSMD can be applied to the pooled value for either an siRNA or a gene especially when the pooled value has a symmetric or nearly normal distribution.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

XDZ proposed all methods, derived all mathematical formulas, conducted all simulations and drafted the final manuscript.

Additional material

Additional file 1

SSMD estimate and its distribution. In this file, I provide statistical estimation and confidence interval for both unpaired and paired SSMD, derive non-central t-distribution property of SSMD estimates, and explore false positive and false negative rates when SSMD is used for hit selection in RNAi high-throughput screening experiments.

Click here for file
[http://www.biomedcentral.com/content/supplementary/1756-0500-1-33-S1.doc]

References

1. Chung NJ, Zhang XD, Kreamer A, Locco L, Kuan PF, Bartz S, Linsley PS, Ferrer M, Strulovici B: **Median absolute deviation to improve hit selection for genome-scale RNAi screens.** *Journal of Biomolecular Screening* 2008, **13**:149-158.
2. Espeseth AS, Huang Q, Gates A, Xu M, Yu Y, Simon AJ, Shi XP, Zhang XD, Hodor P, Stone DJ, Burchard J, Cavet G, Bartz S, Linsley P, Ray WJ, Hazuda D: **A genome wide analysis of ubiquitin ligases in APP processing identifies a novel regulator of BACE1 mRNA levels.** *Molecular and Cellular Neuroscience* 2006, **33**:227-235.
3. Gou D, Narasaraju T, Chintagari NR, Jin N, Wang PC, Liu L: **Gene silencing in alveolar type II cells using cell-specific promoter in vitro and in vivo.** *Nucleic Acids Research* 2004, **32**:.
4. Gou D, Zhang H, Baviskar PS, Liu L: **Primer extension-based method for the generation of a siRNA/miRNA expression vector.** *Physiological Genomics* 2007, **31**:554-562.
5. Malo N, Hanley JA, Cerquozzi S, Pelletier J, Nadon R: **Statistical practice in high-throughput screening data analysis.** *Nature Biotechnology* 2006, **24**:167-175.
6. Zhang XD, Yang XC, Chung NJ, Gates A, Stec E, Kunapuli P, Holder DJ, Ferrer M, Espeseth AS: **Robust statistical methods for hit selection in RNA interference high-throughput screening experiments.** *Pharmacogenomics* 2006, **7**:299-309.
7. Zuck P, Murray EM, Stec E, Grobler JA, Simon AJ, Strulovici B, Inglese J, Flores OA, Ferrer M: **A cell-based beta-lactamase reporter gene assay for the identification of inhibitors of hepatitis C virus replication.** *Analytical Biochemistry* 2004, **334**:344-355.
8. Zhang XD: **A pair of new statistical parameters for quality control in RNA interference high-throughput screening assays.** *Genomics* 2007, **89**:552-561.
9. Zhang XD: **Novel analytic criteria and effective plate designs for quality control in genome-wide RNAi screens.** *Journal of Biomolecular Screening* 2008, **13**:363-377.
10. Zhang XD, Espeseth AS, Johnson EN, Chin J, Gates A, Mitnaul LJ, Marine SD, Tian J, stec EM, Kunapuli P, Holder DJ, Heyse JF, Strulovici B, Ferrer M: **Integrating experimental and analytic approaches to improve data quality in genome-wide screens.** *Journal of Biomolecular Screening* 2008, **13**:378-389.
11. Zhang XD: **A new method with flexible and balanced control of false negatives and false positives for hit selection in RNA interference high-throughput screening assays.** *Journal of Biomolecular Screening* 2007, **12**:645-655.
12. Zhang XD, Ferrer M, Espeseth AS, Marine SD, Stec EM, Crackower MA, Holder DJ, Heyse JF, Strulovici B: **The use of strictly standardized mean difference for hit selection in primary RNA interference high-throughput screening experiments.** *Journal of Biomolecular Screening* 2007, **12**:497-509.