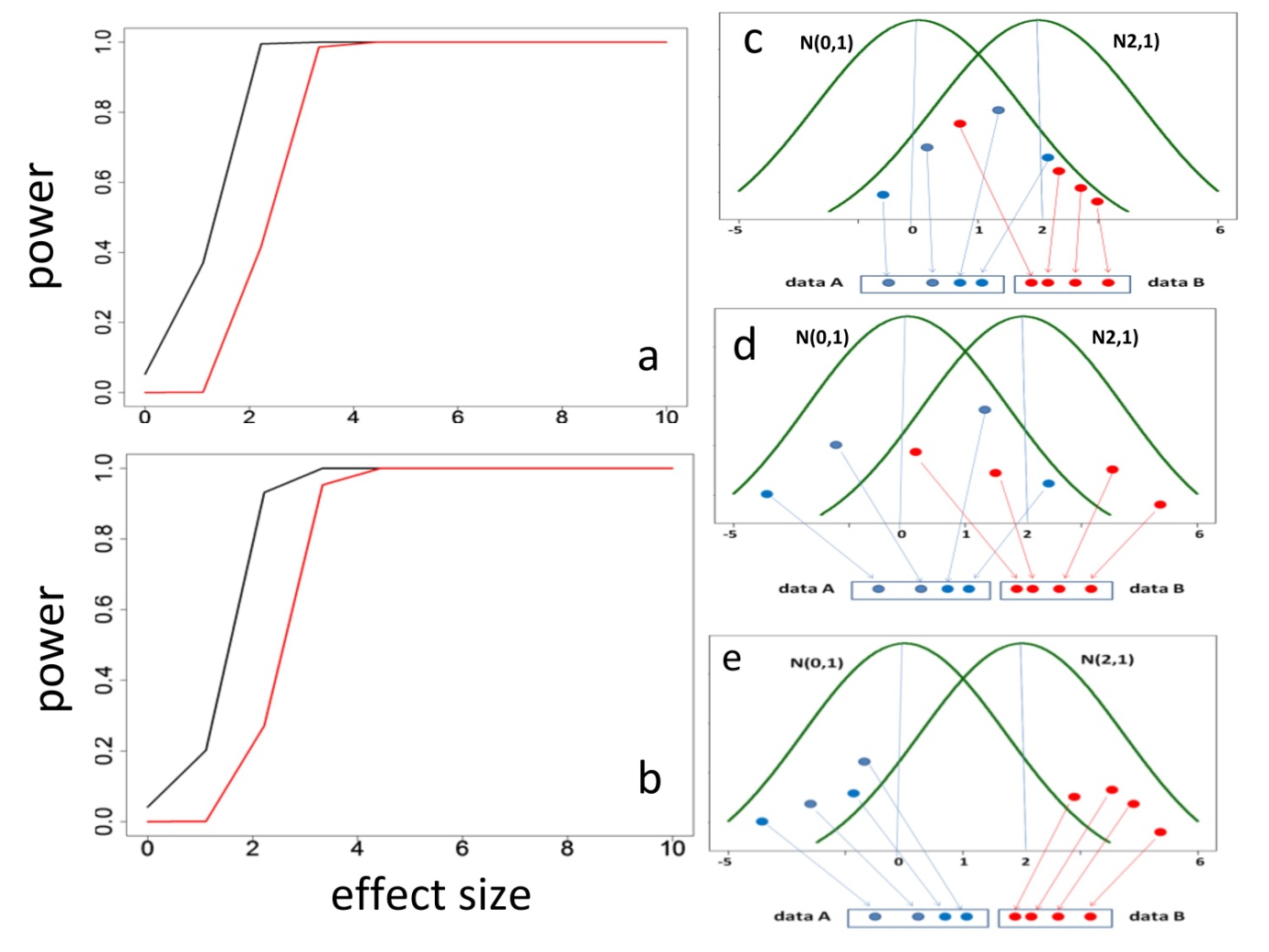
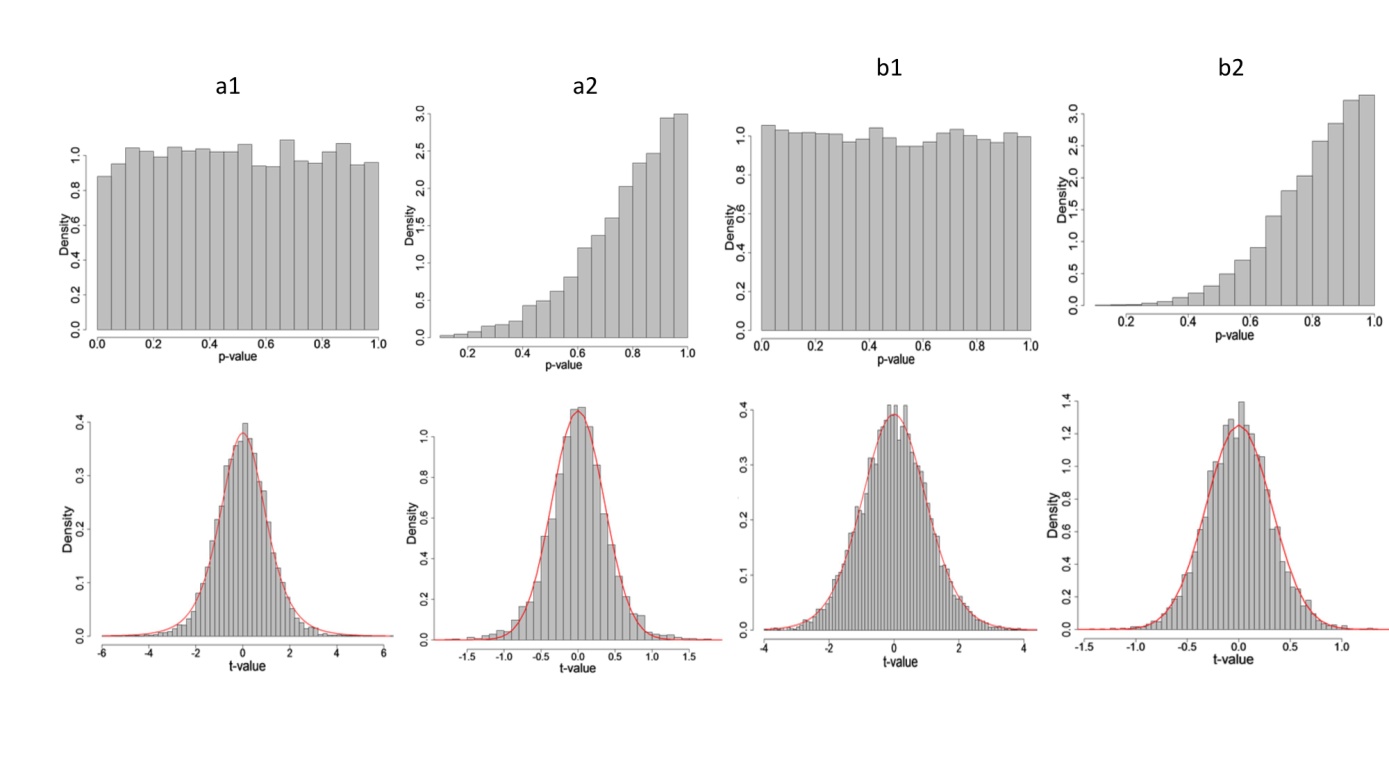
**Supplementary Figures**

****

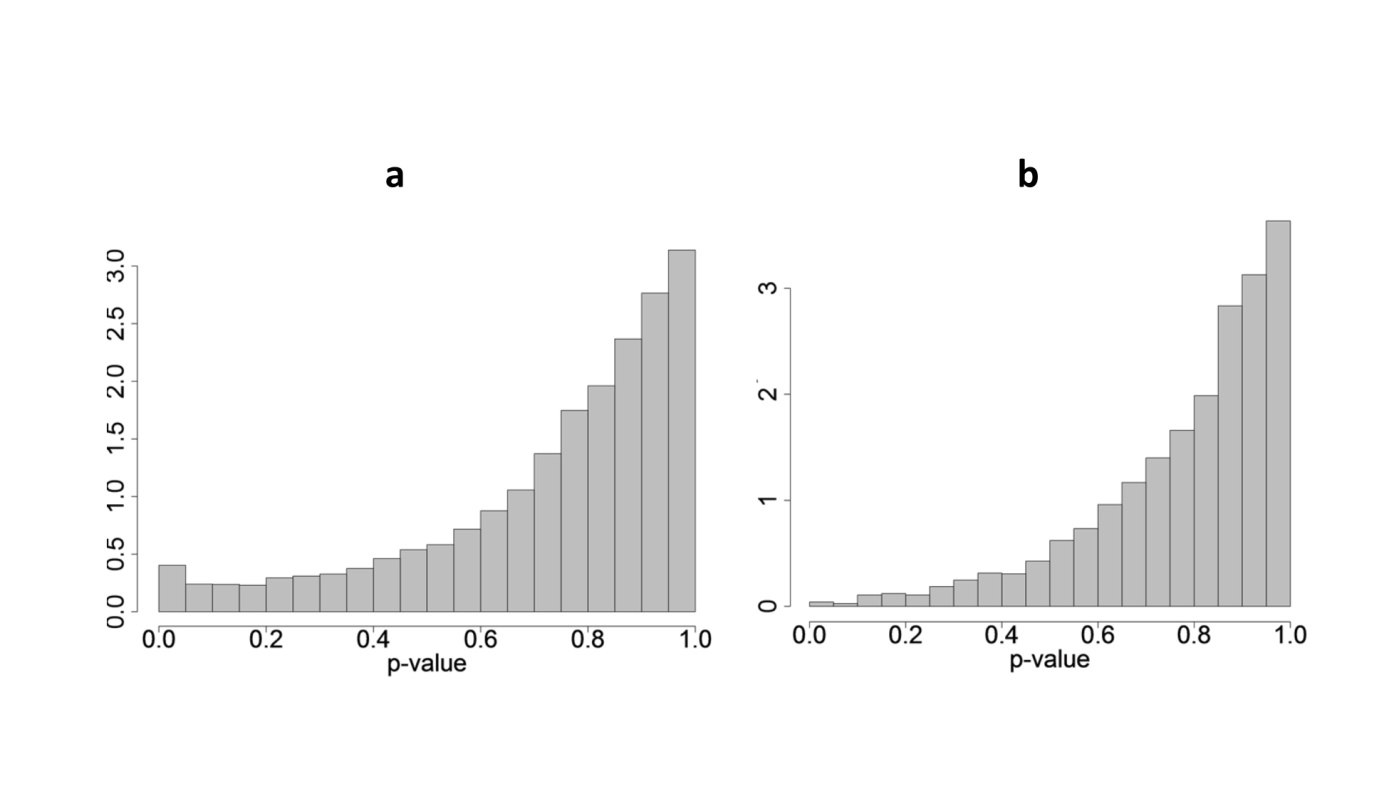
**Figure S1. Power and distribution overlap**

The empirical power was calculated byperforming a statistical method to identify significant differences between two samples under p = 0.05 on a set of 10,000 two-sample experimental datasets with equal sample size = 9 replicates (a) or 15 replicates (b) that were generated from standard normal distributions N(0,1) and N(E,1) where 100% experiments had effect size E where E=0,1,..,10. Three cases possibly occurring in two overlapped distributions are used to explain why t-test has higher power than -test when E < 5. Case 1(c): two datasets A and B with sample size = 4 are sampled from overlap (common) area. Case 2(d): two datasets A and B with sample size = 4 are partially sampled from overlap (common) area and case 3(e): two datasets A and B with sample size = 4 are separately sampled from separated N(0,1) and N(2,1). Differences between two datasets in three cases are detected to be significant by t-test at p<0.05 but only in case 3(e) difference can be tested to be significant by -test. Positives in cases 1 (c) and 2(d) are false, only positives in case 3(e) are true. The proportion of cases 1 and 2 is the largest when E=1 and reduced as E increases while proportion of case 3 increases as E increases and become 100% when E 5.

****

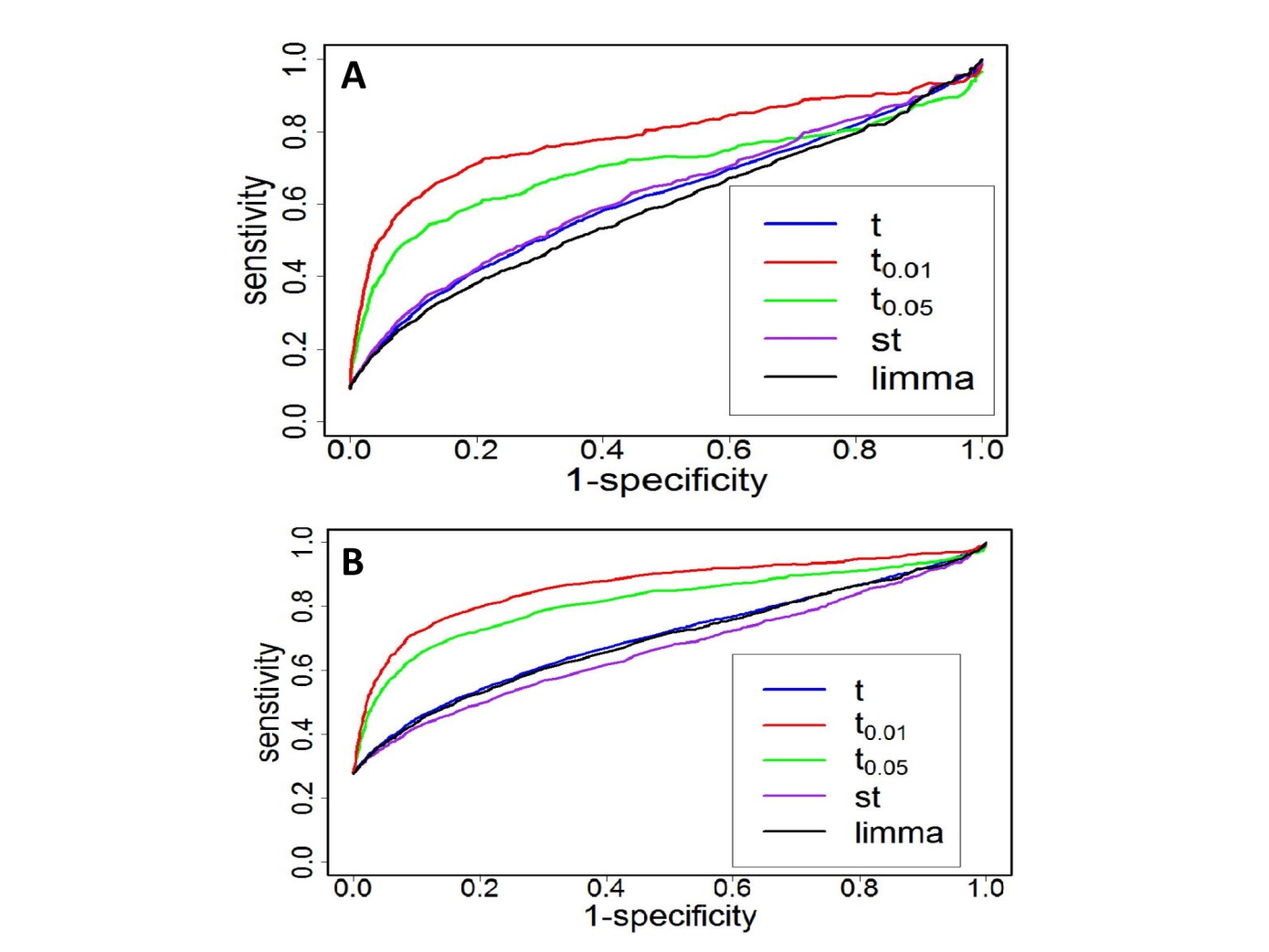
**Figure S2. Null t- and p-value distributions**

Density distributions of t-statistics and p-values of a t-test were constructed on 10,000 independent null experiments where two samples with equal sample size = 6 (a) and 15 (b) were independently sampled from a standard normal distribution N(0,1). 1: t-test. 2: -test. In t-test, theoretical distribution (red curve) is completely matched with the observed t-distribution, while, in -test, the theoretical curve (red curve) given from density distribution function of -test also very well fits the observed t-distribution. Similarly, in t-test, density of p-value are fluctuated about 1 along with p-value, while in -test p-value density is distributed from 0 up to 3.5 as p-value increases from 0 to 1.



**Figure S3. P-value distributions**

a: P-value density distribution of -test constructed by performing -test on 10,000 independent null experiments where two samples with equal sample size = 6 were independently sampled from normal distribution N(0,13) is 0.5, not 0 at p-value but increases from 0.25 up to 3.0 as p-value increases from 0.2 to 1.0. b: p-value density distribution of -test was constructed by performing -test on 10,000 independent null experiments by two-step random sampling: one sample with sample size = 12 was randomly sampled from the same normal distribution N(0,13) at the first step and then randomly split into two samples with equal sample size = 6. The distribution is from 0 up to 3.0 along with p-value, which is the same with p-value density distribution in supplementary Figure S2a and S2b.

****

**Figure S4. ROC comparison among 4 statistical methods for identifying differentially expressed genes**

Microarray data of 7129 genes in two conditions each having 4 replicates were simulated from gene-wise normal distributions, where and where and were come from the original microarray dataset(Tusher, Tibshirani and Chu, 2001), E=5U >1.2 or E=5U <0.8 was assigned to 10% (A) or 30% (B) of genes where.

**Supplementary Tables**

**Table S1**. **16 simulation scenarios**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Scenario # | Sample size | | Proportion | variance | Paired or unpaired |
| Condition A | Condition B |
| 1(A) | 3 | 3 | 10% | 156.062 | unpaired |
| 2(B) | 4 | 4 | 10% | 156.062 | unpaired |
| 3(C) | 3 | 5 | 10% | 156.062 | unpaired |
| 4(D) | 10 | 10 | 10% | 156.062 | unpaired |
| 5(E) | 3 | 3 | 10% | 11916.3 | unpaired |
| 6(F) | 3 | 3 | 10% | 156.062 | paired |
| 7(G) | 15 | 15 | 10% | 156.062 | unpaired |
| 8(H) | 10 | 10 | 10% | 156.062 | paired |
| 9(A) | 3 | 3 | 30% | 156.062 | unpaired |
| 10(B) | 4 | 4 | 30% | 156.062 | unpaired |
| 11(C) | 3 | 5 | 30% | 156.062 | unpaired |
| 12(D) | 10 | 10 | 30% | 156.062 | unpaired |
| 13(E) | 3 | 3 | 30% | 11916.3 | unpaired |
| 14(F) | 3 | 3 | 30% | 156.062 | paired |
| 15(G) | 15 | 15 | 30% | 156.062 | unpaired |
| 16(H) | 10 | 10 | 30% | 156.062 | paired |

Table S2. False findings of *t*-test,-test, and Wilcoxon-test testing for 10,000 null hypotheses in each of 30 null simulation two-sample experiments with equal sample sizes (n= 3) re-sampled from a normal distribution N( =82.05, =13).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Simulation  experiment |  | | |  | | |
| parametric test | | non-parametric test | parametric test | | non-parametric test |
| *t*-test | -test | Wilcoxon test | *t*-test | -test | Wilcoxon test |
| 1 | 512 | 376 | 0 | 102 | 57 | 0 |
| 2 | 497 | 361 | 0 | 90 | 60 | 0 |
| 3 | 467 | 331 | 0 | 92 | 57 | 0 |
| 4 | 473 | 332 | 0 | 96 | 56 | 0 |
| 5 | 502 | 374 | 0 | 105 | 59 | 0 |
| 6 | 524 | 386 | 0 | 99 | 52 | 0 |
| 7 | 496 | 356 | 0 | 120 | 80 | 0 |
| 8 | 508 | 377 | 0 | 117 | 63 | 0 |
| 9 | 512 | 362 | 0 | 89 | 61 | 0 |
| 10 | 487 | 341 | 0 | 99 | 60 | 0 |
| 11 | 470 | 344 | 0 | 99 | 62 | 0 |
| 12 | 497 | 360 | 0 | 114 | 60 | 0 |
| 13 | 478 | 341 | 0 | 96 | 59 | 0 |
| 14 | 493 | 355 | 0 | 109 | 71 | 0 |
| 15 | 514 | 359 | 0 | 99 | 52 | 0 |
| 16 | 498 | 359 | 0 | 102 | 62 | 0 |
| 17 | 533 | 383 | 0 | 120 | 77 | 0 |
| 18 | 490 | 363 | 0 | 118 | 71 | 0 |
| 19 | 477 | 336 | 0 | 82 | 49 | 0 |
| 20 | 508 | 370 | 0 | 105 | 61 | 0 |
| 21 | 470 | 332 | 0 | 91 | 53 | 0 |
| 22 | 471 | 332 | 0 | 80 | 49 | 0 |
| 23 | 463 | 336 | 0 | 89 | 56 | 0 |
| 24 | 477 | 338 | 0 | 120 | 65 | 0 |
| 15 | 493 | 358 | 0 | 94 | 58 | 0 |
| 26 | 502 | 347 | 0 | 98 | 59 | 0 |
| 27 | 478 | 344 | 0 | 98 | 58 | 0 |
| 28 | 493 | 351 | 0 | 88 | 44 | 0 |
| 29 | 500 | 345 | 0 | 100 | 66 | 0 |
| 30 | 487 | 362 | 0 | 109 | 72 | 0 |
| **mean** | **492.3333** | **353.7** | **0** | **100.6667** | **60.3** | **0** |
| **SD** | **17.6153** | **16.0369** | **0** | **11.2198** | **8.0693** | **0** |

Table S3. False findings of t-test, -test, and Wilcoxon-test testing for 1000 null hypotheses in 30 null simulation experiments with equal sample sizes (n=6) re-sampled from a normal distribution N( =0, =1).

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Simulation experiment |  |  | | |  | | |
|  | parametric test | | non-parametric test | parametric test | | non-parametric test |
|  | t-test | -test | Wilcoxon test | t-test | -test | Wilcoxon test |
| 1 | 47 | | 2 | 51 | 4 | 0 | 10 |
| 2 | 44 | | 2 | 40 | 13 | 0 | 6 |
| 3 | 46 | | 1 | 30 | 11 | 0 | 13 |
| 4 | 67 | | 0 | 41 | 11 | 0 | 11 |
| 5 | 50 | | 2 | 41 | 18 | 0 | 8 |
| 6 | 37 | | 0 | 51 | 8 | 0 | 5 |
| 7 | 44 | | 2 | 50 | 13 | 0 | 9 |
| 8 | 63 | | 2 | 38 | 6 | 0 | 10 |
| 9 | 46 | | 1 | 35 | 9 | 0 | 8 |
| 10 | 51 | | 2 | 51 | 8 | 0 | 8 |
| 11 | 49 | | 0 | 39 | 11 | 0 | 10 |
| 12 | 45 | | 2 | 37 | 9 | 0 | 10 |
| 13 | 43 | | 1 | 33 | 16 | 0 | 5 |
| 14 | 50 | | 2 | 36 | 8 | 0 | 10 |
| 15 | 58 | | 0 | 47 | 12 | 0 | 10 |
| 16 | 48 | | 1 | 38 | 11 | 0 | 7 |
| 17 | 67 | | 3 | 33 | 8 | 0 | 7 |
| 18 | 50 | | 2 | 33 | 7 | 0 | 6 |
| 19 | 55 | | 1 | 47 | 5 | 0 | 12 |
| 20 | 51 | | 1 | 36 | 10 | 1 | 15 |
| 21 | 54 | | 2 | 37 | 15 | 0 | 11 |
| 22 | 47 | | 0 | 36 | 3 | 0 | 9 |
| 23 | 52 | | 1 | 39 | 11 | 0 | 8 |
| 24 | 46 | | 0 | 38 | 8 | 0 | 12 |
| 25 | 51 | | 1 | 44 | 8 | 0 | 9 |
| 26 | 46 | | 1 | 42 | 8 | 0 | 11 |
| 27 | 47 | | 0 | 42 | 8 | 0 | 5 |
| 28 | 45 | | 0 | 47 | 8 | 0 | 11 |
| 29 | 47 | | 1 | 45 | 11 | 0 | 7 |
| 30 | 54 | | 2 | 42 | 8 | 0 | 6 |
| **mean** | **50** | | **1.1667** | **40.6** | **9.5** | **0.0333** | **9.0** |
| **SD** | **6.6333** | | **0.8595** | **5.8867** | **3.3397** | **0.1825** | **2.5118** |

Table S4. Findings, powers, and type I error rates of t-test, -test, and Wilcoxon-test for testing 7000 null hypotheses and 3000 alternative hypotheses under in each of 30 simulation two-sample experiments with equal sample sizes (n=3) re-sampled from normal distributions N( =82.05, =13) and N( =82.05E, =13) where E=5U is effect size, U=[0,1] is uniform variables.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Parametric test | | | | | | | | Non-parametric test | | | | |
| experiment | t-test | | | | -test | | | |  | Wilcoxon-test | | | | |
| findings | true findings | power  (%) | type I error  rate | findings | true findings | power  (%) | type I error  rate | findings | | true findings | power  (%) | type I error  rate | | |
| 1 | 2799 | 2408 | 80.3 | 0.039 | 2753 | 2499 | 83.3 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 2 | 2760 | 2391 | 79.7 | 0.034 | 2733 | 2526 | 84.2 | 0.021 | 0 | | 0 | 0 | 0 | | |
| 3 | 2789 | 2371 | 79.0 | 0.042 | 2716 | 2491 | 83.0 | 0.023 | 0 | | 0 | 0 | 0 | | |
| 4 | 2911 | 2524 | 84.1 | 0.039 | 2871 | 2617 | 87.2 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 5 | 2801 | 2427 | 80.9 | 0.037 | 2757 | 2527 | 84.2 | 0.023 | 0 | | 0 | 0 | 0 | | |
| 6 | 2820 | 2472 | 82.4 | 0.035 | 2835 | 2592 | 86.4 | 0.024 | 0 | | 0 | 0 | 0 | | |
| 7 | 2770 | 2406 | 80.2 | 0.036 | 2789 | 2534 | 84.5 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 8 | 2777 | 2425 | 80.8 | 0.035 | 2772 | 2536 | 84.5 | 0.024 | 0 | | 0 | 0 | 0 | | |
| 9 | 2832 | 2465 | 82.2 | 0.037 | 2789 | 2566 | 85.5 | 0.022 | 0 | | 0 | 0 | 0 | | |
| 10 | 2802 | 2435 | 81.2 | 0.037 | 2791 | 2531 | 84.4 | 0.026 | 0 | | 0 | 0 | 0 | | |
| 11 | 2848 | 2461 | 82.0 | 0.039 | 2787 | 2532 | 84.4 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 12 | 2821 | 2417 | 80.6 | 0.040 | 2836 | 2565 | 85.5 | 0.027 | 0 | | 0 | 0 | 0 | | |
| 13 | 2842 | 2434 | 81.1 | 0.041 | 2776 | 2547 | 84.9 | 0.023 | 0 | | 0 | 0 | 0 | | |
| 14 | 2810 | 2454 | 81.8 | 0.036 | 2737 | 2527 | 84.2 | 0.021 | 0 | | 0 | 0 | 0 | | |
| 15 | 2826 | 2440 | 81.3 | 0.039 | 2780 | 2546 | 84.9 | 0.023 | 0 | | 0 | 0 | 0 | | |
| 16 | 2850 | 2447 | 81.6 | 0.040 | 2791 | 2547 | 84.9 | 0.024 | 0 | | 0 | 0 | 0 | | |
| 17 | 2768 | 2427 | 80.9 | 0.034 | 2785 | 2550 | 85.0 | 0.024 | 0 | | 0 | 0 | 0 | | |
| 18 | 2756 | 2385 | 79.5 | 0.037 | 2757 | 2501 | 83.4 | 0.027 | 0 | | 0 | 0 | 0 | | |
| 19 | 2949 | 2540 | 84.7 | 0.041 | 2855 | 2644 | 88.1 | 0.021 | 0 | | 0 | 0 | 0 | | |
| 20 | 2748 | 2379 | 79.3 | 0.037 | 2727 | 2480 | 82. 7 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 21 | 2768 | 2412 | 80.4 | 0.036 | 2749 | 2501 | 83.4 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 22 | 2838 | 2446 | 81.5 | 0.039 | 2785 | 2553 | 85.1 | 0.023 | 0 | | 0 | 0 | 0 | | |
| 23 | 2746 | 2411 | 80.4 | 0.034 | 2741 | 2499 | 83.3 | 0.024 | 0 | | 0 | 0 | 0 | | |
| 24 | 2737 | 2364 | 78.8 | 0.037 | 2737 | 2481 | 82.7 | 0.026 | 0 | | 0 | 0 | 0 | | |
| 25 | 2745 | 2392 | 79.7 | 0.035 | 2765 | 2516 | 83.9 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 26 | 2787 | 2438 | 81.3 | 0.035 | 2774 | 2541 | 84.7 | 0.023 | 0 | | 0 | 0 | 0 | | |
| 27 | 2756 | 2391 | 79.7 | 0.037 | 2733 | 2483 | 82.8 | 0.025 | 0 | | 0 | 0 | 0 | | |
| 28 | 2880 | 2477 | 82.6 | 0.040 | 2793 | 2590 | 86.3 | 0.020 | 0 | | 0 | 0 | 0 | | |
| 29 | 2757 | 2404 | 80.1 | 0.035 | 2788 | 2526 | 84.2 | 0.026 | 0 | | 0 | 0 | 0 | | |
| 30 | 2812 | 2431 | 81.0 | 0.038 | 2780 | 2548 | 84.9 | 0.023 | 0 | | 0 | 0 | 0 | | |
| mean | **2803.50** | **2429.13** | **80.9** | **0.037** | **2776.07** | **2536.53** | **84.6** | **0.024** | **0** | | **0** | **0** | **0** | | |
| SD | **50.87** | **40.68** | **0.0136** | **0.0022** | **37.125** | **38.949** | **0.0129** | **0.0017** | **0** | | **0** | **0** | **0** | | |

**Appendix A: Theorem and proof**

**Theorem**: Let be treatment effect in an experiment and be noise (residual value) in replicate observation *j* in condition *i*. For a given two-condition experiment with and replicates producing two datasets  and where minimum of two datasets is larger than 0. If (1) treatment effect is larger than noise variances, or more exactly, and , (2) difference between two noise variances is smaller than difference between overall variance and sum of two noise variances, then

(1) or  ,

(2) and

(3).

**The proof of the theorem**

1. In the case of equal sample sizes
2. If, , and , we have , , and . Since ,  or if, , , , and , we then have .









.

From condition (1) showing  and , we therefore have >0.

(3)





Similarly, .



It shows 

which is equivalent to . For , , (3) holds.

When , according to condition (2) in the theorem, we have



Since , we have.  respectively timed by and  on the right and left hands results in , leading to

. (3) also holds in the case of .

1. In the case of unequal sample sizes
2. If, , and , we have , , and. Since , we have  or if, , , , and, we still have , leading to >0.







According to condition (1),  and , we got .

(3)





Similarly, .



Equation above shows



.

For ,, leading to the fact that , (3) holds. For , according to condition (2) in the theorem, we have



Likewise, because of, we have, which leads to.

For the case of , (3) still holds.

**Appendix B. Derivation of -distribution**

Our derivation of - density function begins with standard normal distribution:

where is a variable in standard normal space *U*, , and *v* is a variance variable in *V* following withdegrees of freedom . Transformation of U to T is

and

where T is space of variable t and where where is a measure of gap and variation between two samples and is a null . , and also , which leads to . Thus, we have

,

which leads to

and

.

A jointed probability of variables *v* and *t* is

where and . Integration of over *v* is

Set , then and . We get

With , we have

Using , we obtain

.

This is density function of . is related to degree of freedom (*n*). When, suggesting that t-value is shrunken to center point and density at center area increases, as shown by Dirac delta function. However, when 28, 1, = , that is, distribution is reduced to t-distribution (Fig.1).

**Appendix C. Probability cumulative function of**

In deriving the cumulative function for the -distribution we begin with estimating the integral for a symmetric region. Let , then we have

where and , so . leads to .Then we have . Thus, .

From one to the other-sided results, we have

Since , that is, , then cumulative function of distribution is also written as

**Supplementary Simulation**

**Single-experiments**

We used a set of real data to calculate mean and variance for parameters of normal distribution. This set of data (78.08525, 99.3769, 82.9608, 69.7071) is retrieved from RPPA (reverse phase protein arrays) data (personal communication) and the mean =82.5 and variance 156.06. Therefore, our simulation was conducted in an experimental field. Two sets of data were simulated by using normal distribution with =82.5E and where E=5U or E=1 where 0 < U. E value is effect size and randomly assigned to one of two normal distributions with equal probability. E>1.2 or E<0.8 (equivalent to fold change of 1.2 and 0.8) is randomly assigned to P of N independent experiments. E>1.2 means that experiments have positive (up) effect and E < 0.8 indicates negative (down) effect. No effect occurs in experiments if 0.8E 1.2. Therefore, E is different from experiment to experiment. To make sufficient comparisons between - and t-tests, we designed 13 scenarios: two levels for proportion of true positive experiments: P=10% and 30%; 4 levels for equal sample sizes: 3, 4, 10, 15; 2 levels for variances: small variance (156.06) and big variance (11916.25); two levels for paired statues: unpaired and paired samples. The scenario 13 is standard normal distribution. We simulated a set of datasets of experiments with equal sample size of 6 from standard normal distributions with or = E and variance =1. For each scenario, we simulated 10,000 independent experiments. These 13 scenarios were used to make ROC ([Receiver operating characteristic](https://en.wikipedia.org/wiki/Receiver_operating_characteristic)) curves to compare performances of statistical methods.

To compare important statistical properties (for example, type I error and power) of and t-tests at a given significance level, we designed a set of 16 scenarios in Supplementary Table S1. Simulation schedule for these 16 scenarios is the same with that for 13 scenarios above but each scenario was repeated 30 times for calculation of mean and standard deviation (SD). In addition, we also designed a set of treatment experiments in which two group with 3 replicates are respectively sampled from two standard normal distributions N(0,1) and N(E,1) where E = 0, 1, 2, …, 10. We simulated 5000 such experiments for each treatment effect.

**Microarray experiments**

For the multiple experiments, we simulated two microarray datasets. Simulation schedule is that expression data of 7129 genes were created from gene-wise normal distributions with = and where and were respectively mean and variance of gene g derived from a real microarray dataset(Tusher et al., 2001). Likewise, effect size we set was and assigned to 10% and 30% of 7129 genes in one of two conditions with equal probability where *U* is uniform variable (). The genes with >1.2 or <0.8 are defined as differentially expressed (up- and down-regulated) genes. Replicate number was set to be 4 for each condition.

**Null experiments**

In simulation experiments, two datasets with given replicate numbers are generated by randomly sampling from a given distribution. We designed 6 simulation schedules. Schedule 1: two datasets in a null experiment with equal sample size of 3 are independently and randomly sampled from normal distribution N(82.5, ). We simulated 10, 000 such null experiments and each experiment repeated 30 for calculating average and SD of type I error rates. Schedule 2: two datasets in a null experiment with equal sample size of 6 are independently and randomly sampled from standard normal distribution N(0, 1). We here simulated 1, 000 such null experiments but also repeated 30 for calculating average and SD of type I error rates. Schedule 3: 5 sets of 5,000 null experiments with sample sizes = 3, 6, 9, 12, and 15 are independently and randomly simulated from normal distribution N(0, 13). Schedule 4: 6 sets of 5,000 null experiments with given equal sample size of 6 ware randomly sampled from a set of normal distributions N(0,) where = 3, 9, 12, 50, and 100. Each set of null experiments was also repeated 30 times for calculating average and standard deviation of type I error rate. Schedule 5: two datasets in a null experiment with equal sample sizes of 3, 6, and 15 are independently and randomly sampled from standard normal distributions N(0, 1) and N(0,13). We here simulated 10, 000 such null experiments to build t- and p-value density distributions. Schedule 6: 12 replicate values are randomly sampled from normal distribution N(0,13) and then are randomly assigned to two groups with equal sample size. This is called two-step random simulation way. We also simulated 10, 000 such null experiments to generate completely random data without gap between groups to build t- and p-value density distributions.

**Supplementary Biological Data**

**Single-experiment data**

The data for testing inhibition effect of *CSF1* (colony-stimulating-factor-1) on human breast cancer cell proliferation were come from Qin et al(Qin et al., 2014). *CSF1* mRNA was knocked down to block NCOA1-promoted macrophage recruitment by synthetic siRNA (Dharmacon) or lentivirus-mediated shRNA expression in human MDA-231-LM3.3 cell line. Non-targeting shRNAs in MDA-231-LM3.3 were used as control. Since MDA-231-LM3.3 is metastatic cancer cells developed from human breast cancer cell line MDA-MB-231, MDA-MB-231 without shCSF-1 knockdown was used as negative control for testing difference in proliferation between metastatic and non-metastatic cancer cell lines. This experiment was repeated for 6 times (n=6 mice).

**Null experimental data**

Null experiment is such an experiment that there is no statistical difference between two conditions because the data are randomly sampled from the same distribution. So, null experiments can be used to directly estimate type I errors of statistical Methods. In the real biological null experiments, data distribution is unknown, and hence the results of tests are independent of a specified distribution. Fornage et al(Fornage et al., 2008) provided data of ideal biological null experiment: A microarray data of 8799 genes were generated from two male rattus strains: stroke-resistant (SR) and stroke-prone (SP) rats and two dietary regimens (*N*=6 in each strain-diet group): a “stroke-permissive diet” high in sodium (HS) (0.63% potassium, 0.37% sodium) and 1% NaCl drinking solution and a “stroke-protective diet” low in sodium and high in potassium (LS) (1.3% potassium, 0.35% sodium) and water, which construct four groups: LSSP, HSSP, LSSR, and HSSR, each group having 6 biological replicates. The null experiments were conducted within groups with equal sample size of 3 replicates. In the simulated experiments, two datasets with given replicate numbers were generated by randomly sampling from a given distribution.

**Real microarray data**

Microarray data (GEO access number: GSE9691) was downloaded from gene expression omnibus (GEO). The data contain 22277 gene probes and four groups: shEcad, shctrl, DNEcad and shEcad+shBcat, three biological replicates per group and were used to study gene expression changes induced by knockdown of E-cadherin with shRNA-mediated (shEcad), expressions of dominant-negative E-cadherin (DNEcad) and double knockdown of E-cadherin and Beta-catenin (shEcad+shBcat) in immortalized human mammary epithelial cells(Onder et al., 2008). For the convenience, only two groups shEcad and shctrl in this microarray datasets were used to compare five statistical methods st-test(Opgen-Rhein and Strimmer, 2007), , t-test, limma(Ritchie et al., 2015) and SAM (Tusher et al., 2001).

**References**

Fornage, M., Chiang, Y. A., O'Meara, E. S.*, et al.* (2008). Biomarkers of Inflammation and MRI-Defined Small Vessel Disease of the Brain: The Cardiovascular Health Study. *Stroke* **39**, 1952-1959.

Onder, T. T., Gupta, P. B., Mani, S. A., Yang, J., Lander, E. S., and Weinberg, R. A. (2008). Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. *Cancer Res* **68**, 3645-3654.

Opgen-Rhein, R., and Strimmer, K. (2007). Accurate ranking of differentially expressed genes by a distribution-free shrinkage approach. *Stat Appl Genet Mol Biol* **6**, Article9.

Qin, L., Wu, Y. L., Toneff, M. J.*, et al.* (2014). NCOA1 Directly Targets M-CSF1 Expression to Promote Breast Cancer Metastasis. *Cancer Res* **74**, 3477-3488.

Ritchie, M. E., Phipson, B., Wu, D.*, et al.* (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47.

Tusher, V. G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* **98**, 5116-5121.