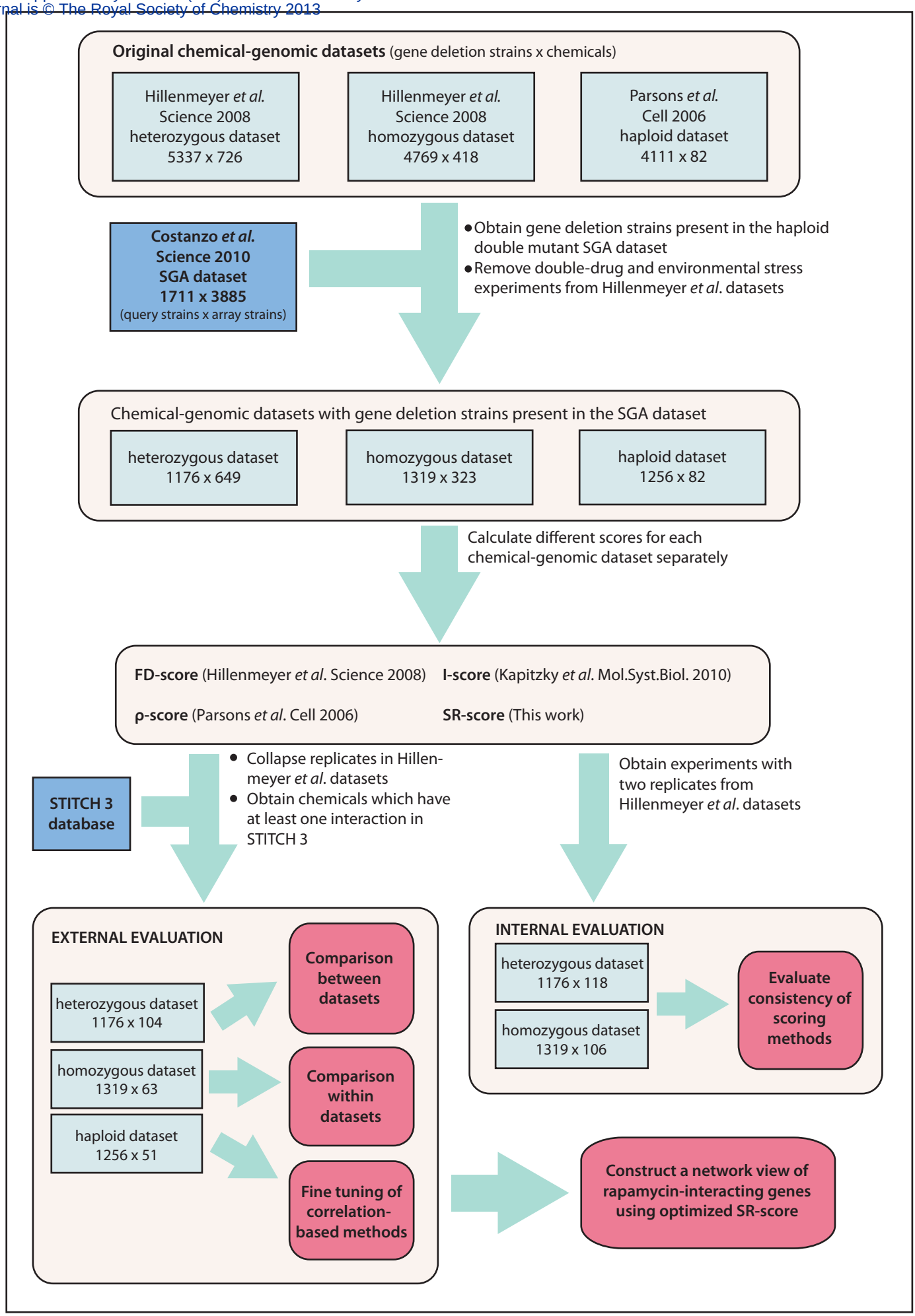
# Supplementary Information for

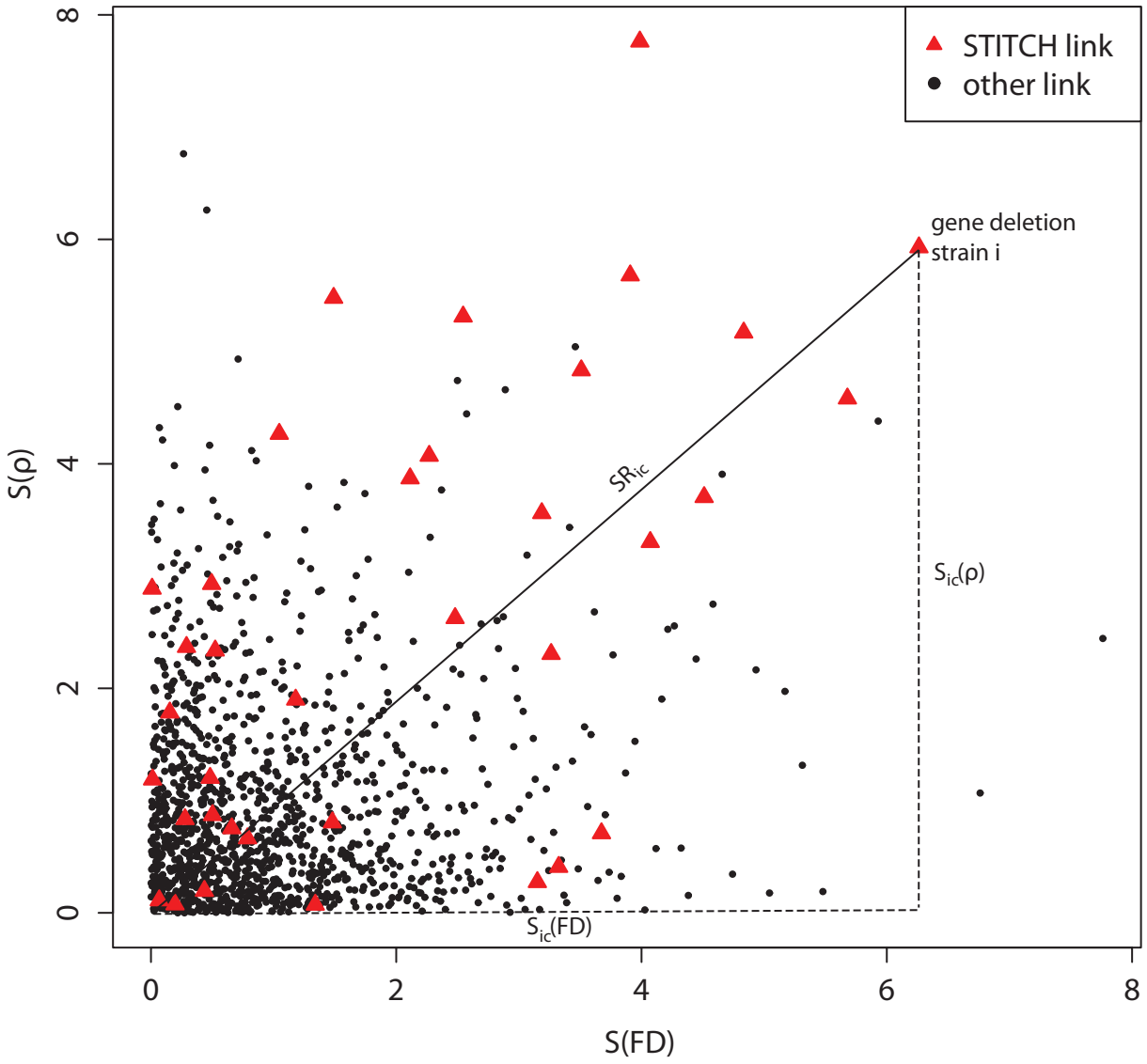
## *Predicting drug-target interactions through integrative analysis of chemogenetic assays in yeast*

**Marja A Heiskanen and Tero Aittokallio**

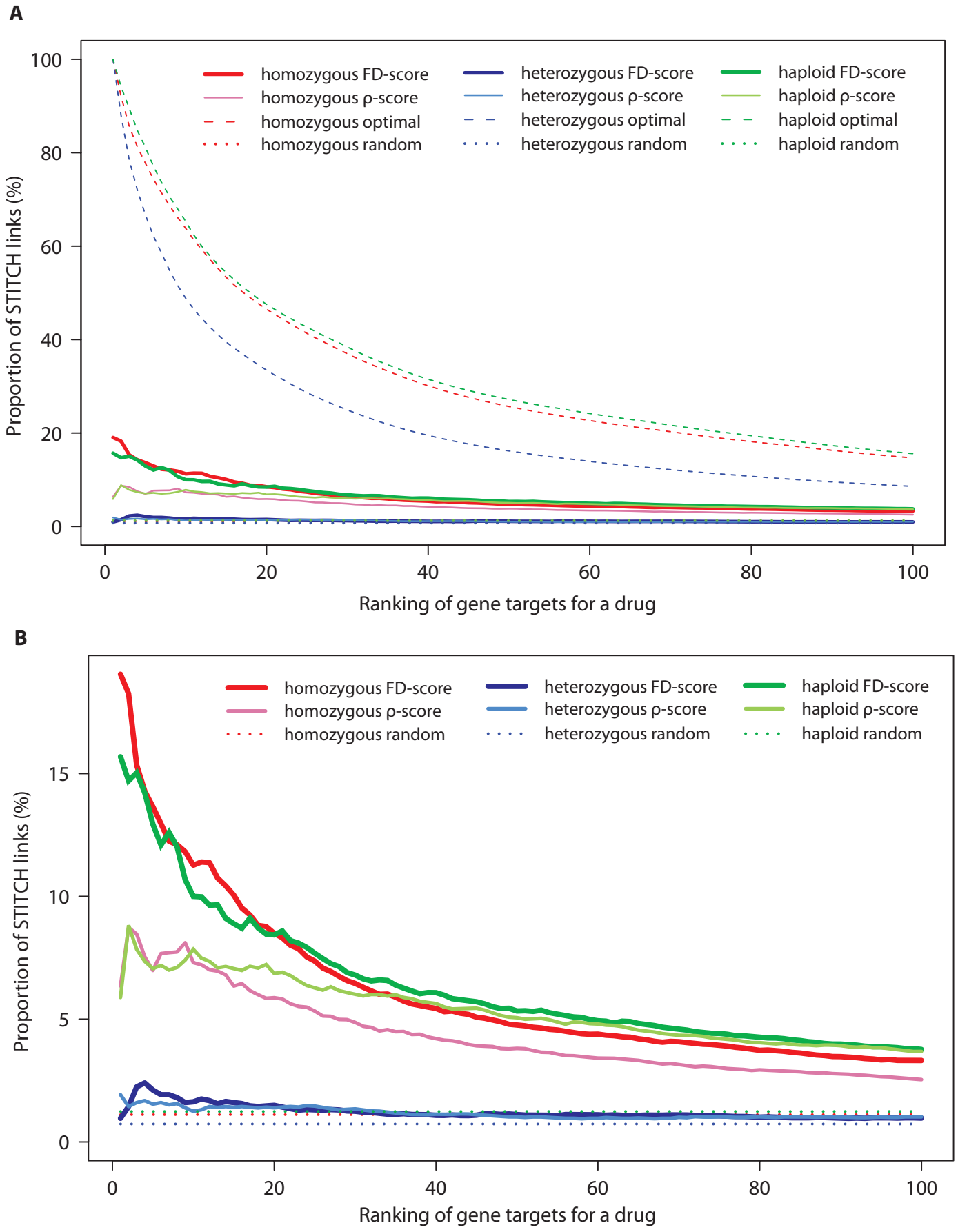
## Supplementary Information Table of Contents



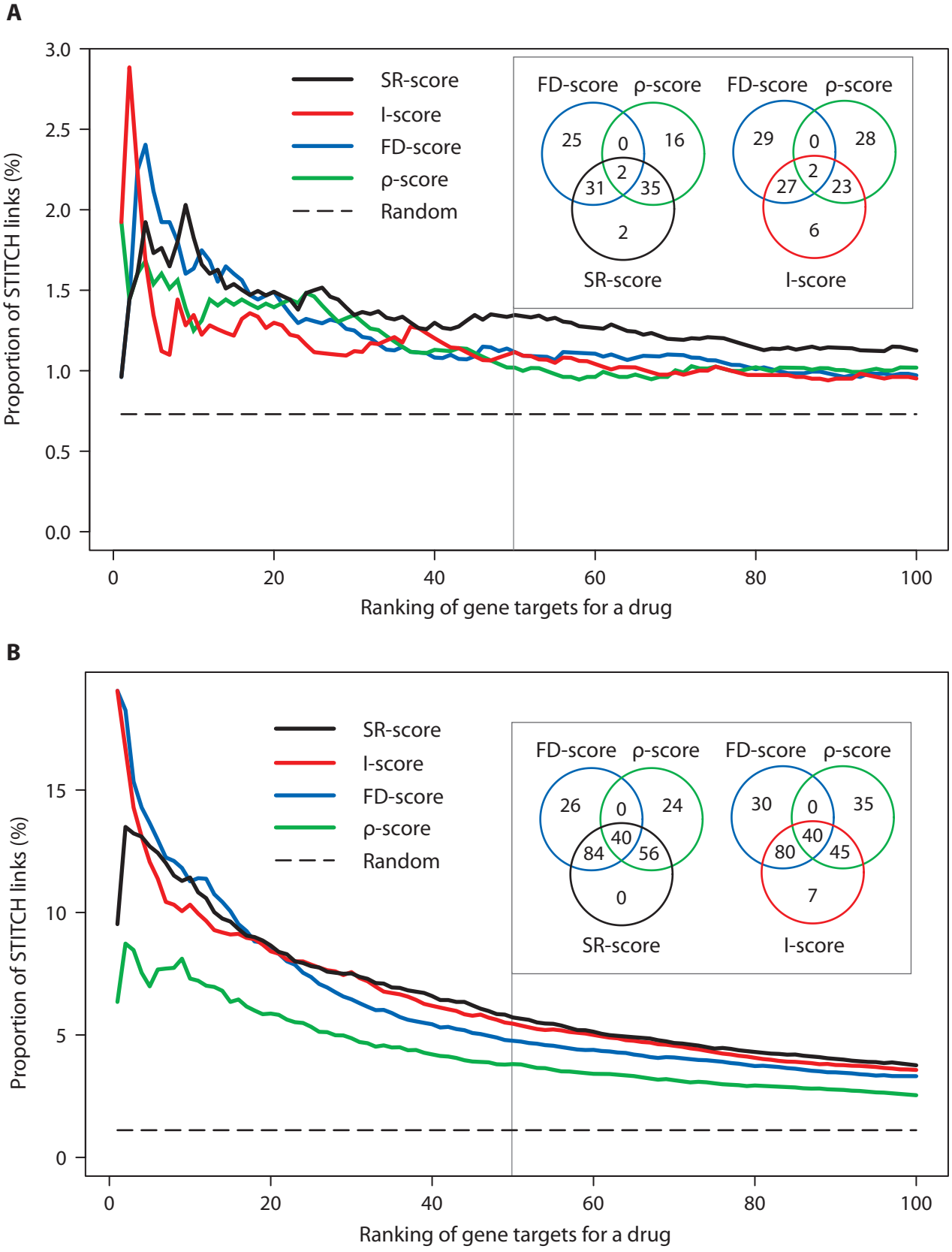

**Supplementary Figure S1. Workflow for the evaluation setup.**



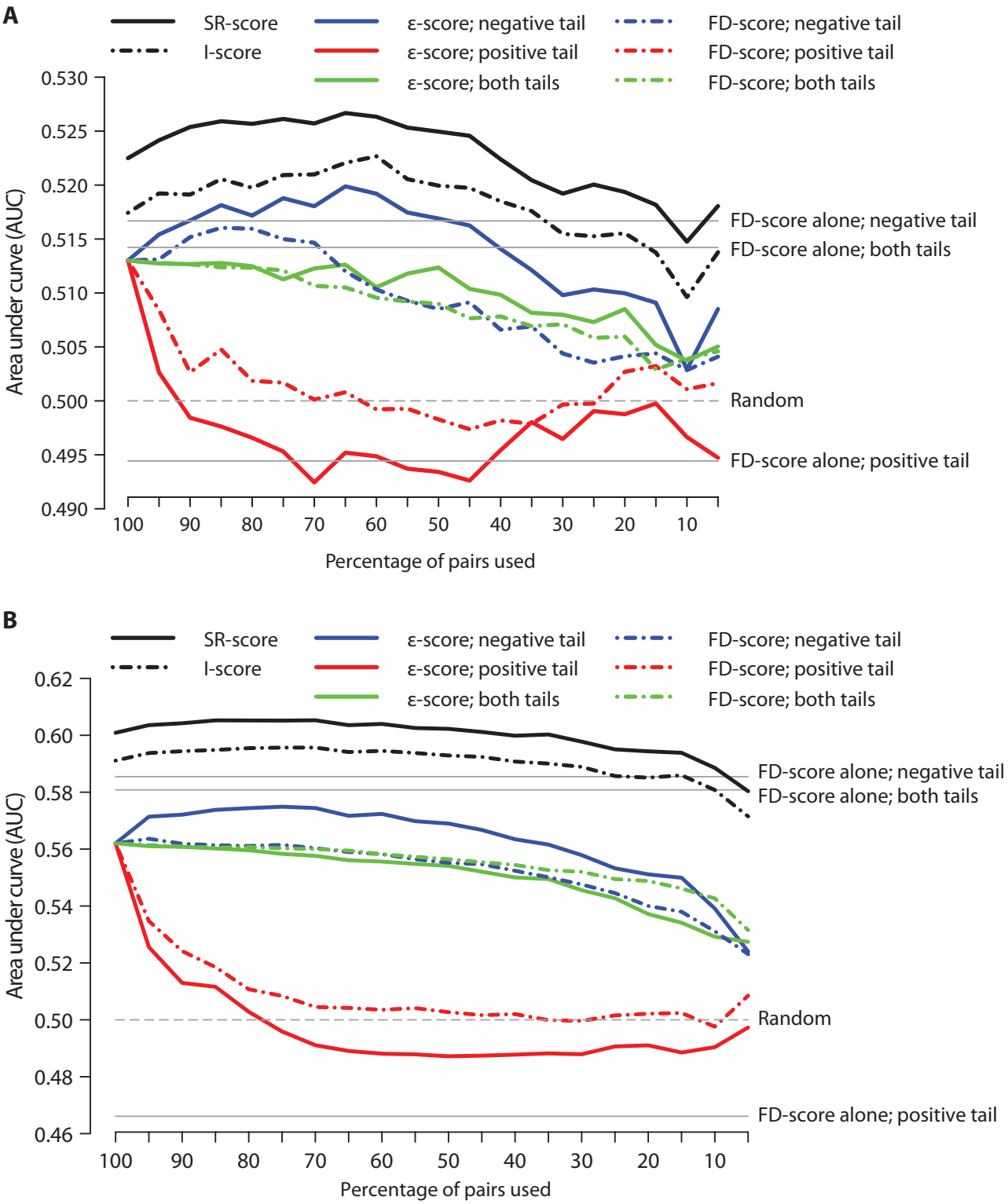

**Supplementary Figure S2. The operation of the SR-score.** For each chemical, gene deletion strains are first ranked according to their FD- and ρ-scores. Ranks are then transformed into the Savage-scores S(FD) and S(ρ), which emphasize early ranks. For a gene deletion strain *i* and chemical *c* (here, nocodazole), the SR-score is the distance of point ($S_{ic}$(FD), $S_{ic}$(ρ)) from the origin. The red triangles indicate those gene deletion strains which interact with nocodazole according to STITCH, and the black dots indicate non-interacting gene deletion strains. For example, the gene deletion strain *i* results in ranks 3 and 4 when using the FD- and ρ-scores, respectively, whereas the SR-score results in rank 1.



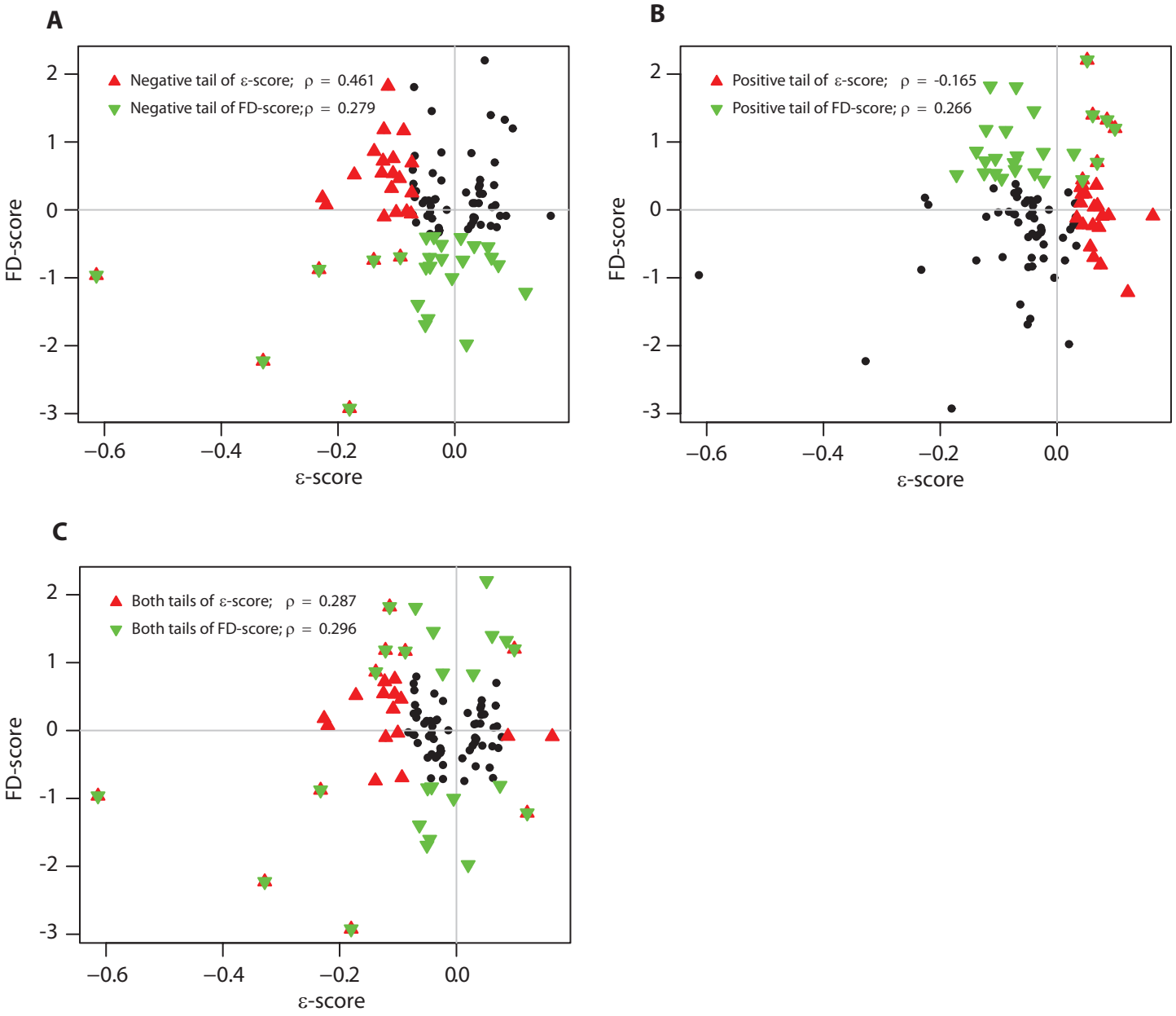

**Supplementary Figure S3. Comparison of the homozygous, heterozygous and haploid datasets when predicting drug-target interactions (A) with optimal curves shown, and (B) without optimal curves.** The performance curves illustrate the number of recovered STITCH links relative to total number of drug-gene pairs at varying ranks when the gene targets are ranked for each drug separately using different scoring methods.



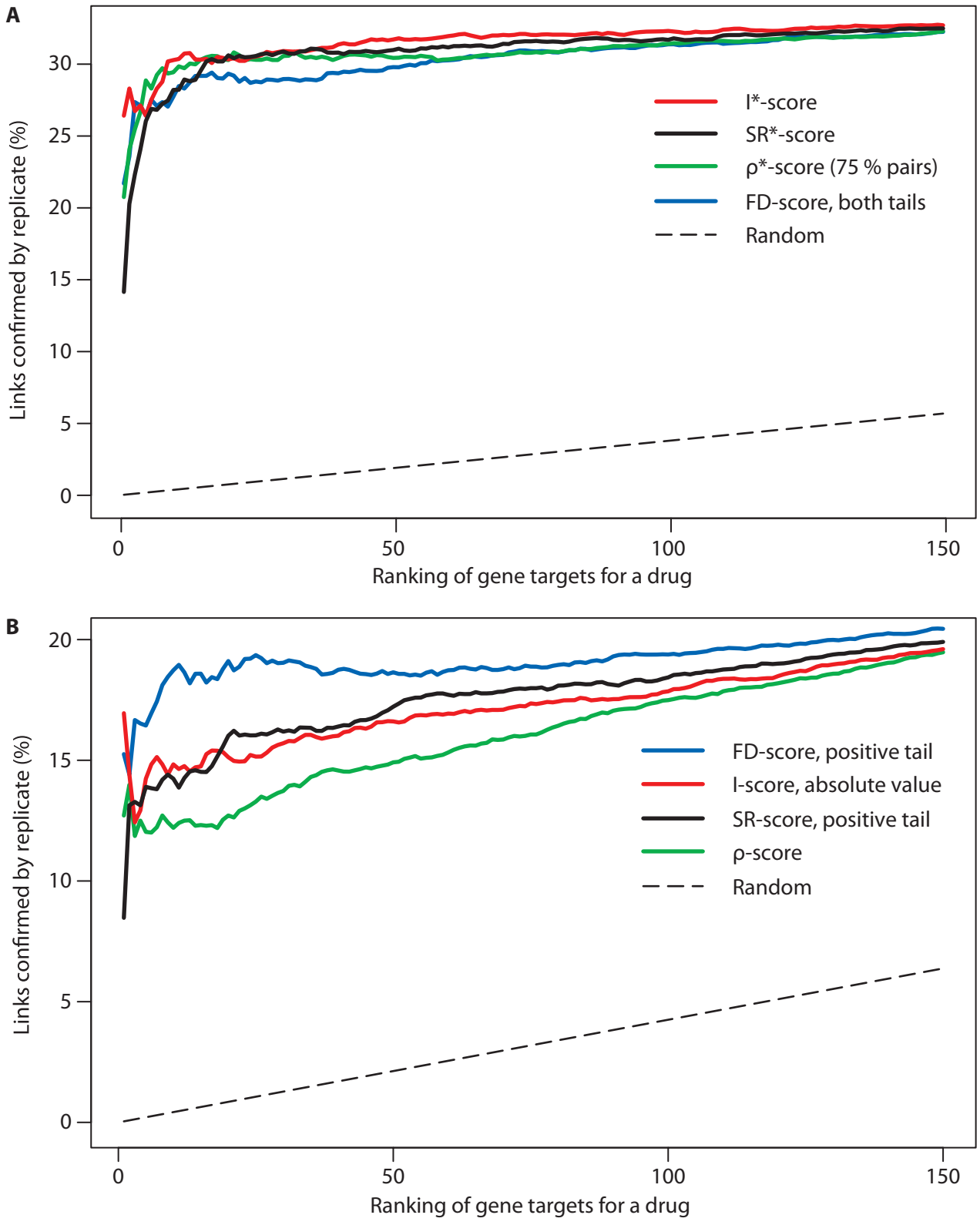

**Supplementary Figure S4. Performance of the different scoring approaches in (A) the heterozygous dataset and (B) the homozygous dataset.** The curves illustrate the number of recovered STITCH links relative to the total number of drug-gene pairs at varying ranks when the gene targets are ranked for each drug separately using different scoring methods. The Venn diagrams in the insets describe the number of overlapping STITCH links from the different scoring approaches at rank 50. Note: the y-axis scale is different in the **(A)** and **(B)** panels.



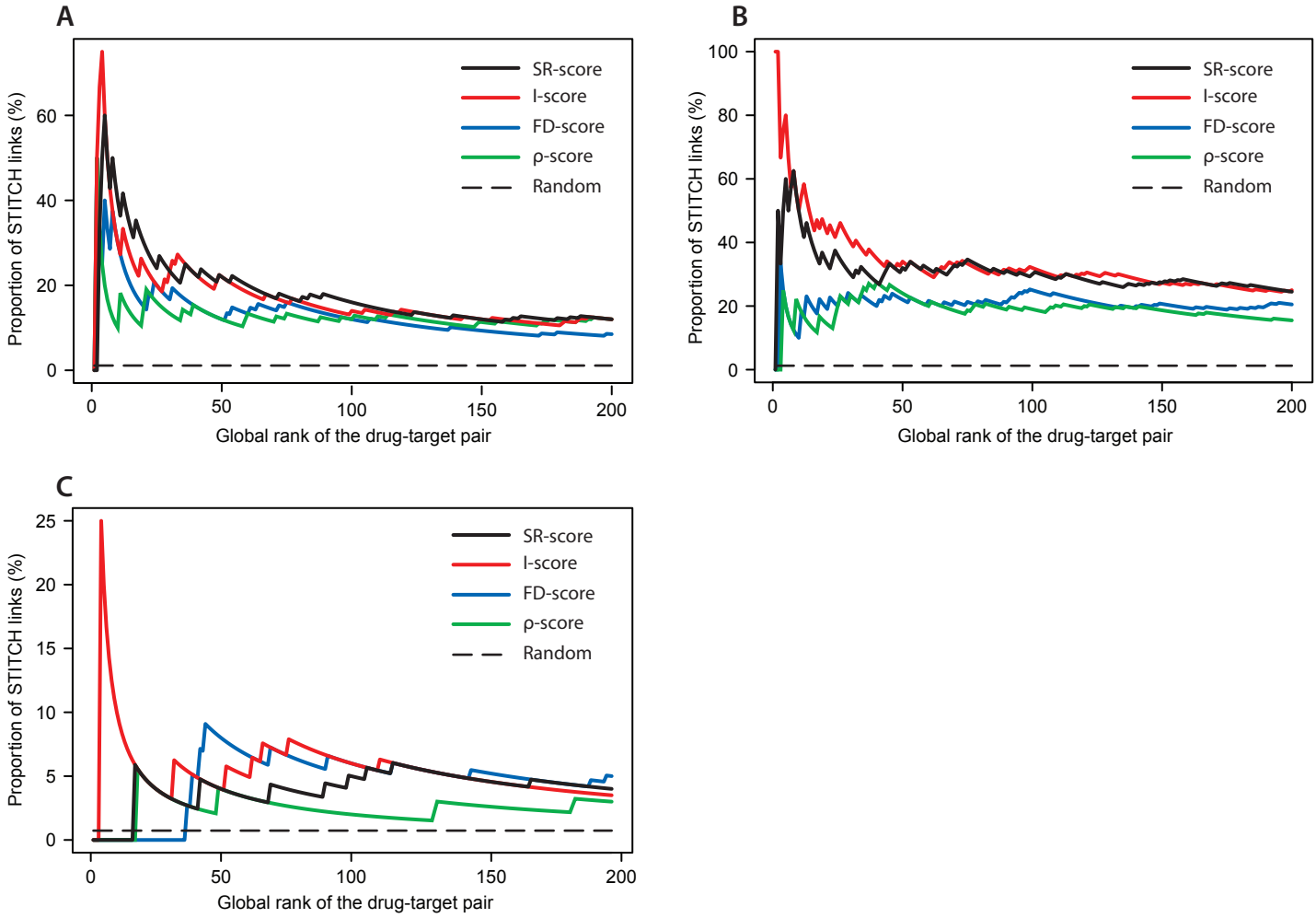

**Supplementary Figure S5. The effect of using negative, positive, or both tails of the genetic interaction profiles (ε-score) or chemical-genetic interaction profiles (FD-score).** The coloured traces describe the predictive accuracy for the ρ-score when using different percentages of non-missing FD- and ε-score pairs selected according to the respective tail of the SGA or **(A)** heterozygous or **(B)** homozygous datasets in the calculation of correlations. The traces for the I- and SR-scores are calculated using the negative tail of the SGA dataset, which shows the best performance in both datasets. Since the FD-scores do not rely on correlations, their AUC values are not affected by the percentages of the pairs used.



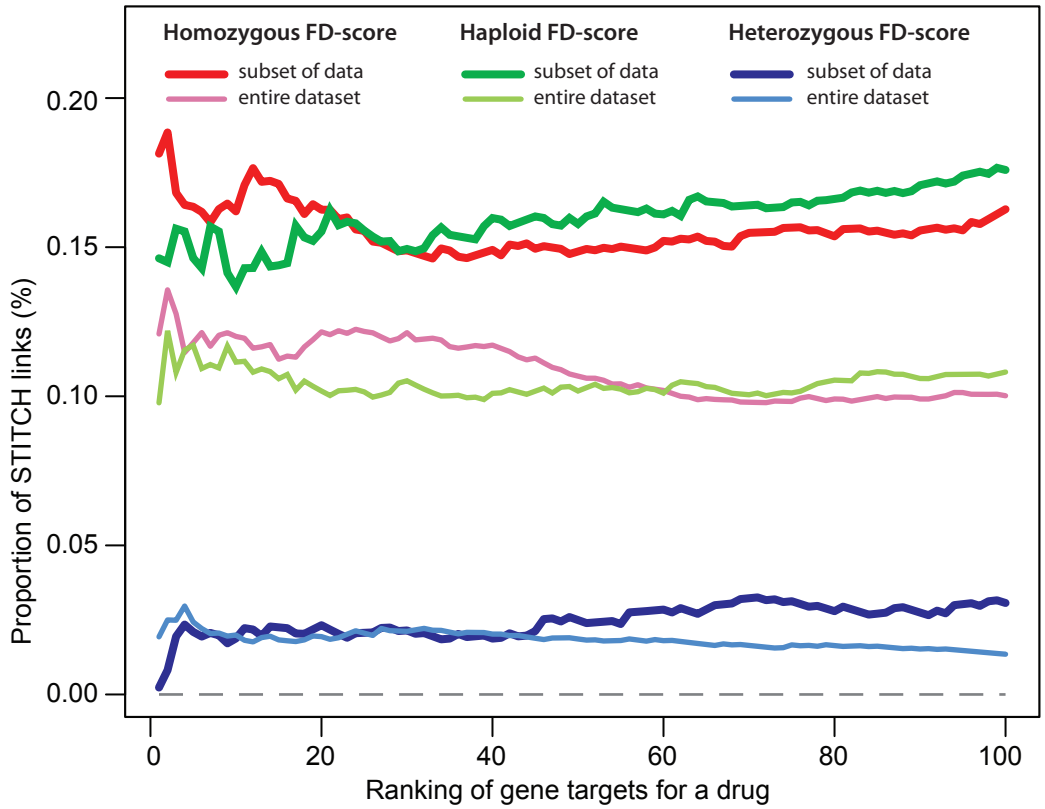

**Supplementary Figure S6. Scatterplots illustrating the pairs selected according to (A) negative, (B) positive, or (C) both tails of the genetic interaction profiles (red triangles) or chemical-genomic profiles (green triangles).** When calculating the ρ-score for gene deletion strain YJL047C and chemical $MnCl_2$ in the homozygous dataset the number of non-missing value in both profiles is 100, and the resulting correlation using all 100 pairs is 0.158. The scatterplots show an example of which pairs contibute to the calculation of the ρ-score when considering only the 25% pairs according to each tail and distribution. For instance, in (A), pairs selected according to 25 most negative values for the ε-score are shown as red triangles, and the resulting ρ-score is 0.461. The ρ-score of each case is shown in the figures. The example illustrates that concentrating only on the most negative synthetic lethal interactions lead to improved correlation beyond that obtained with other tails and chemical-genomic distribution, or when using all the non-missing pairs.




**Supplementary Figure S7. Performances of the best versions of the different scoring approaches when using the replicate measurements in (A) homozygous and (B) heterozygous datasets.** The curves illustrate the number of drug-target links recovered also by a replicate measurement relative to the total number of protein-gene pairs at varying ranks when the target genes are ranked for each drug separately using different scoring methods. In **(A)**, 75% of pairs according to the negative tail of the SGA dataset were used in the calculation of correlations in the ρ*-score. The ρ*-scores, along with the negative tail of the FD-score-distribution, were also used in the I*- and SR*-scores. In **(B)**, the best version of the I-score was obtained using the normal ρ-score and the absolute values of the FD-scores, whereas the best SR-scores were obtained using the positive tail of the FD-score-distribution along with the ρ-scores. The expected values for random scorings were obtained from the hypergeometric distribution.




**Supplemantary Figure S8. Global ranking of drug-target pairs in the (A) homozygous, (B) haploid, and (C) heterozygous dataset.** The curves illustrate the number of STITCH links relative to the total number of drug-gene pairs at varying ranks when ranking is done globally within the entire dataset using different scoring methods. The integration of two complentary scoring methods (FD- and ρ-scores) using either Z-score-based combination (I-score) or rank-based statistics (SR-score) improves the recovery rate of STITCH links in the homozygous **(A)** and haploid **(B)** datasets. The proportion of recovered STITCH links remains relatively high in the haploid datasets also at later ranks, suggesting the haploid dataset being the most accurate when predicting the potential drug-target interactions. On the contrary, only few of the globally best-ranking drug-target pairs are among the STITCH links in the heterozygous dataset **(C)**. These results are in line with the chemical-specific evaluations presented in the main text. Further, the detection accuracy provided by the conventionally used FD- or ρ-scores can be improved when applying the integrative scoring approaches (I- and SR-scores). Note: the y-axis scale is different in the three panels.




**Supplementary Figure S9. Effect of essential genes in the heterozygous dataset on the prediction of drug-target interactions.** The subset of the heterozygous dataset consisting of those genes overlapping with the genetic interaction dataset contains only 26 essential genes, whereas the entire heterozygous dataset contains 1065 out of 1110 essential genes (www.essentialgene.org). In order to find out whether the poor performance of the heterozygous dataset originates from the fact that almost all essential genes are absent in the overlapping dataset, we compared the relative performances of the subset and the entire heterozygous dataset in terms of recovering STITCH links. Since the FD-score is the only approach not relying on the genetic interaction dataset, the evaluation was performed using the FD-score alone. The corresponding curves for the entire datasets as well as for the subsets of genes overlapping with the genetic interaction dataset of the homozygous and haploid datasets are also shown for comparison. The curves are scaled between zero and one using optimal and random curves separately for each dataset. Hence, the scaled random curve is at zero (gray dashed line) and optimal curve is at one (not shown). Interestingly, in the homozygous and haploid datasets, the accuracies of recovering STITCH links decrease when the entire datasets are used, whereas the performance of the entire heterozygous dataset resembles that of the subset of data when considering early ranks. The decrease in accuracy when considering the entire dataset is expected as the STITCH links are more sparse and the ranks span wider ranges as the number of genes is increased compared to a subset of data. However, it may well be that the essential genes present in the entire heterozygous dataset partly account for the relatively good performance of the entire heterozygous dataset compared to the corresponding subset. Nevertheless, the performance of the heterozygous dataset remains relatively poor compared to the haploid and homozygous datasets.

**Supplementary Table S1.** AUC and partial AUC values using different confidence cut-offs of STITCH links

| | | AUC | | | | partial AUC; rank ≤ 150 | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Confidence cut-off | FD-score | p-score | I-score | SR-score | FD-score | p-score | I-score | SR-score | Number of STITCH links | Number of STITCH chemicals |
| Hillenmeyer et al. HETEROZYGOUS | 150 (low) | 0,515 | 0,511 | 0,515 | 0,520 | 0,510 | 0,510 | 0,509 | 0,513 | 893 | 104 |
| | 400 (medium) | 0,515 | 0,513 | 0,519 | 0,521 | 0,512 | 0,508 | 0,513 | 0,514 | 305 | 66 |
| | 500 | 0,511 | 0,511 | 0,512 | 0,516 | 0,509 | 0,506 | 0,508 | 0,510 | 212 | 48 |
| | 600 | 0,511 | 0,505 | 0,508 | 0,512 | 0,508 | 0,503 | 0,505 | 0,507 | 123 | 35 |
| | 700 (high) | 0,506 | 0,504 | 0,506 | 0,507 | 0,504 | 0,501 | 0,503 | 0,503 | 70 | 24 |
| | 800 | 0,505 | 0,503 | 0,506 | 0,506 | 0,504 | 0,500 | 0,503 | 0,503 | 33 | 15 |
| Hillenmeyer et al. HOMOZYGOUS | 150 (low) | 0,583 | 0,560 | 0,588 | 0,598 | 0,577 | 0,550 | 0,582 | 0,586 | 924 | 63 |
| | 400 (medium) | 0,594 | 0,562 | 0,600 | 0,606 | 0,583 | 0,550 | 0,588 | 0,591 | 333 | 47 |
| | 500 | 0,587 | 0,555 | 0,592 | 0,597 | 0,579 | 0,546 | 0,581 | 0,586 | 251 | 39 |
| | 600 | 0,570 | 0,550 | 0,578 | 0,582 | 0,563 | 0,542 | 0,567 | 0,571 | 160 | 29 |
| | 700 (high) | 0,546 | 0,533 | 0,549 | 0,551 | 0,541 | 0,525 | 0,544 | 0,543 | 93 | 18 |
| | 800 | 0,530 | 0,527 | 0,535 | 0,537 | 0,525 | 0,520 | 0,528 | 0,530 | 49 | 13 |
| Parsons et al. HAPLOID | 150 (low) | 0,591 | 0,576 | 0,597 | 0,605 | 0,578 | 0,565 | 0,587 | 0,589 | 795 | 51 |
| | 400 (medium) | 0,600 | 0,593 | 0,615 | 0,623 | 0,583 | 0,578 | 0,598 | 0,600 | 292 | 37 |
| | 500 | 0,595 | 0,582 | 0,602 | 0,612 | 0,576 | 0,568 | 0,586 | 0,589 | 203 | 29 |
| | 600 | 0,582 | 0,577 | 0,597 | 0,600 | 0,568 | 0,560 | 0,581 | 0,582 | 136 | 23 |
| | 700 (high) | 0,563 | 0,558 | 0,574 | 0,575 | 0,551 | 0,542 | 0,560 | 0,558 | 90 | 19 |
| | 800 | 0,541 | 0,537 | 0,548 | 0,549 | 0,532 | 0,527 | 0,538 | 0,538 | 50 | 15 |

**Supplemantary Table S2.** External evaluation of the scoring approaches in the different chemical-genomic datasets

| | | Hillenmeyer et al HETEROZYGOUS | | | | Hillenmeyer et al HOMOZYGOUS | | | | Parsons et al HAPLOID | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | **FD-score** | **ρ-score** | **I-score** | **SR-score** | **FD-score** | **ρ-score** | **I-score** | **SR-score** | **FD-score** | **ρ-score** | **I-score** | **SR-score** |
| **Hillenmeyer et al HETEROZYGOUS** | **FD-score** | 0,517 **p < 0,001** | 0,004 | -0,001 | -0,006 | -0,069 | -0,045 | -0,074 | -0,084 | -0,077 | -0,061 | -0,083 | -0,091 |
| | **ρ-score** | 0,582 | 0,513 **p = 0,001** | -0,004 | -0,009 | -0,072 | -0,049 | -0,078 | -0,088 | -0,081 | -0,065 | -0,087 | -0,095 |
| | **I-score** | 0,877 | 0,226 | 0,517 **p < 0,001** | -0,005 | -0,068 | -0,045 | -0,074 | -0,083 | -0,076 | -0,061 | -0,082 | -0,090 |
| | **SR-score** | 0,128 | **0,012** | 0,061 | 0,522 **p < 0,001** | -0,063 | -0,040 | -0,069 | -0,078 | -0,071 | -0,056 | -0,077 | -0,085 |
| **Hillenmeyer et al HOMOZYGOUS** | **FD-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | 0,585 **p < 0,001** | 0,023 | -0,006 | -0,015 | -0,008 | 0,007 | -0,014 | -0,022 |
| | **ρ-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | **0,006** | 0,562 **p < 0,001** | -0,029 | -0,039 | -0,032 | -0,016 | -0,038 | -0,046 |
| | **I-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | 0,251 | **< 0,001** | 0,591 **p < 0,001** | -0,010 | -0,002 | 0,013 | -0,009 | -0,017 |
| | **SR-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | **0,001** | 0,601 **p < 0,001** | 0,007 | 0,023 | 0,001 | -0,007 |
| **Parsons et al HAPLOID** | **FD-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | 0,470 | **0,003** | 0,830 | 0,514 | 0,594 **p < 0,001** | 0,015 | -0,006 | -0,014 |
| | **ρ-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | 0,502 | 0,118 | 0,232 | **0,035** | 0,061 | 0,578 **p < 0,001** | -0,022 | -0,030 |
| | **I-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | 0,207 | **< 0,001** | 0,439 | 0,932 | 0,236 | **< 0,001** | 0,600 **p < 0,001** | -0,008 |
| | **SR-score** | **< 0,001** | **< 0,001** | **< 0,001** | **< 0,001** | **0,049** | **< 0,001** | 0,141 | 0,543 | **0,001** | **< 0,001** | **0,007** | 0,608 **p < 0,001** |

Diagonal: Area under curve (AUC) and p-value compared to random within a dataset.
Above diagonal: Pair-wise differences of AUCs (row-column).
Below diagonal: p-values for the pair-wise score differences (Within a same dataset: pairwise permutation test. Between different datasets: pooled permutation test).

**Supplementary Table S3.** Ranks of the known drug-target interactions for different scoring approaches based on Cokol et al. Mol.Syst.Biol. 2011 and STITCH binding links

| HOMOZYGOUS DATASET | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Chemical** | **Gene** | **Gene name** | **FD-score** | **p-score** | **p*-score** | **I-score** | **I*-score** | **SR-score** | **SR*-score** | **Reference[1]** | **STITCH score[2]** | **STITCH action[3]** |
| amsacrine | YOL006C | TOP1 | 11 | 1239 | 1139 | 58 | 44 | 32 | 31 | STITCH | 822 | binding |
| atorvastatin | YML075C | HMG1 | | | | | | | | Hillenmeyer et al 2010 | 637 | binding |
| benomyl | YML124C | TUB3 | 3 | 1 | 1 | 1 | 1 | 2 | 1 | Hillenmeyer et al 2010 | 748 | |
| caffeine | YJR066W | TOR1 | | | | | | | | Kuranda et al 2006 | 459 | inhibition |
| camptothecin | YOL006C | TOP1 | 1055 | 146 | 148 | 187 | 342 | 311 | 304 | Pommier et al 1998 | 968 | binding, inhibition |
| cantharidin | YDL134C | PPH21 | 495 | 138 | 437 | 332 | 406 | 238 | 561 | STITCH | 449 | binding |
| carbendazine | YML124C | TUB3 | | | | | | | | Hillenmeyer et al 2010 | | |
| carmustine | YPL091W | GLR1 | | | | | | | | Hillenmeyer et al 2010 | 509 | inhibition |
| caspofungin | YLR342W | FKS1 | 1274† | 1283 | 1297 | 1274 | 1274 | 1317 | 1317 | Lesage et al 2004 | 708 | |
| colchicine | YML124C | TUB3 | | | | | | | | Hillenmeyer et al 2010 | 325 | inhibition, binding, activation |
| doxorubicin | YOL006C | TOP1 | | | | | | | | STITCH | 578 | binding |
| fenbendazole | YML124C | TUB3 | | | | | | | | Hillenmeyer et al 2010 | | |
| FK 506 | YLR433C | CNA1 | 902 | 725 | 1108 | 633 | 996 | 1031 | 1191 | STITCH | 979 | inhibition, binding |
| geldanamycin | YPL240C | HSP82 | | | | | | | | Mandal et al 2008 | 924 | binding |
| hydrogen peroxide | YDR256C | CTA1 | 291 | 172 | 95 | 248 | 128 | 242 | 159 | STITCH | 997 | binding, catalysis |
| hydrogen peroxide | YGR088W | CTT1 | 19 | 228 | 616 | 28 | 69 | 35 | 47 | STITCH | 996 | binding, catalysis |
| hygromycin | YJR075W | HOC1 | 95 | 66 | 42 | 57 | 32 | 59 | 45 | Kim et al 2006 | | |
| lithium | YHR046C | INM1 | 515 | 758 | 960 | 972 | 832 | 797 | 839 | Murray & Greenberg 2000 | 978 | inhibition |
| lovastatin | YML075C | HMG1 | 2 | 151 | 359 | 2 | 2 | 3 | 5 | Rine et al 1983 | 895 | inhibition, binding |
| mebendazole | YML124C | TUB3 | 11 | 1 | 1 | 1 | 1 | 1 | 1 | Laclette et al 1980 | | |
| nocodazole | YML124C | TUB3 | 27 | 5 | 4 | 3 | 4 | 7 | 7 | Laclette et al 1980 | 488 | inhibition, activation |
| pentamidine | YOR266W | PNT1 | | | | | | | | Luesch et al 2005 | | |
| radicicol | YPL240C | HSP82 | | | | | | | | Taldone et al 2009 | 635 | binding |
| rapamycin | YJR066W | TOR1 | 4 | 117 | 92 | 12 | 14 | 9 | 7 | Zheng & Schreiber 1997 | 988 | inhibition, binding |
| staurosporine | YGR040W | KSS1 | 914 | 1208 | 600 | 1017 | 716 | 1198 | 922 | STITCH | 200 | binding |
| thiabendazole | YML124C | TUB3 | 1068 | 1 | 1 | 2 | 7 | 2 | 2 | Chaklan et al 2008 | | |
| trichostatin A | YGL194C | HOS2 | | | | | | | | STITCH | 903 | binding |
| trichostatin A | YNL021W | HDA1 | | | | | | | | STITCH | 844 | binding |
| trichostatin A | YNL330C | RPD3 | | | | | | | | STITCH | 949 | binding |
| trichostatin A | YPR068C | HOS1 | | | | | | | | STITCH | 474 | inhibition, binding |
| vitamin K3 | YPL091W | GLR1 | | | | | | | | STITCH | 710 | binding |
| wortmannin | YJR066W | TOR1 | 769 | 1291 | 1235 | 1117 | 1016 | 1104 | 1111 | STITCH | 463 | binding |
| | | **MEAN OF RANKS** | **439** | **443** | **479** | **350** | **346** | **376** | **385** | | | |

[1] Reference according to Cokol et al Mol. Syst. Biol. 2011 or STITCH database.

[2] Specified if interaction exists in STITCH database.

[3] Specified if action is defined in STITCH database.

† Rank based on a missing value in the homozygous chemical-genomic dataset.

| HETEROZYGOUS DATASET | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Chemical** | **Gene** | **Gene name** | **FD-score** | **p-score** | **p*-score** | **I-score** | **I*-score** | **SR-score** | **SR*-score** | **Reference[1]** | **STITCH score[2]** | **STITCH action[3]** |
| amsacrine | YOL006C | TOP1 | 30 | 1143 | 1145 | 622 | 338 | 67 | 69 | STITCH | 822 | binding |
| atorvastatin | YML075C | HMG1 | 1 | 431 | 431 | 1 | 1 | 1 | 2 | Hillenmeyer et al 2010 | 637 | binding |
| benomyl | YML124C | TUB3 | 493 | 19 | 45 | 84 | 88 | 44 | 99 | Hillenmeyer et al 2010 | 748 | |
| caffeine | YJR066W | TOR1 | 1006 | 773 | 728 | 591 | 937 | 1050 | 1022 | Kuranda et al 2006 | 459 | inhibition |
| camptothecin | YOL006C | TOP1 | 99 | 741 | 962 | 401 | 358 | 212 | 221 | Pommier et al 1998 | 968 | binding, inhibition |
| cantharidin | YDL134C | PPH21 | 968 | 580 | 598 | 517 | 848 | 903 | 919 | STITCH | 449 | binding |
| carbendazine | YML124C | TUB3 | 1 | 306 | 72 | 2 | 1 | 1 | 1 | Hillenmeyer et al 2010 | | |
| carmustine | YPL091W | GLR1 | 5 | 779 | 758 | 80 | 35 | 9 | 12 | Hillenmeyer et al 2010 | 509 | inhibition |
| caspofungin | YLR342W | FKS1 | 8 | 880 | 157 | 16 | 1 | 17 | 14 | Lesage et al 2004 | 708 | |
| colchicine | YML124C | TUB3 | 535 | 5 | 2 | 37 | 12 | 13 | 5 | Hillenmeyer et al 2010 | 325 | inhibition, binding, activation |
| doxorubicin | YOL006C | TOP1 | 475 | 779 | 580 | 942 | 532 | 748 | 677 | STITCH | 578 | binding |
| fenbendazole | YML124C | TUB3 | 1 | 184 | 26 | 1 | 1 | 2 | 1 | Hillenmeyer et al 2010 | | |
| FK 506 | YLR433C | CNA1 | 341 | 453 | 774 | 598 | 622 | 492 | 598 | STITCH | 979 | inhibition, binding |
| geldanamycin | YPL240C | HSP82 | | | | | | | | Mandal et al 2008 | 924 | binding |
| hydrogen peroxide | YDR256C | CTA1 | | | | | | | | STITCH | 997 | binding, catalysis |
| hydrogen peroxide | YGR088W | CTT1 | | | | | | | | STITCH | 996 | binding, catalysis |
| hygromycin | YJR075W | HOC1 | 531 | 43 | 85 | 170 | 131 | 101 | 172 | Kim et al 2006 | | |
| lithium | YHR046C | INM1 | 1076 | 443 | 307 | 171 | 906 | 761 | 588 | Murray & Greenberg 2000 | 978 | inhibition |
| lovastatin | YML075C | HMG1 | 1 | 874 | 88 | 1 | 1 | 2 | 1 | Rine et al 1983 | 895 | inhibition, binding |
| mebendazole | YML124C | TUB3 | 5 | 77 | 9 | 2 | 1 | 7 | 3 | Laclette et al 1980 | | |
| nocodazole | YML124C | TUB3 | 12 | 7 | 14 | 1 | 1 | 4 | 6 | Laclette et al 1980 | 488 | inhibition, activation |
| pentamidine | YOR266W | PNT1 | | | | | | | | Luesch et al 2005 | | |
| radicicol | YPL240C | HSP82 | | | | | | | | Taldone et al 2009 | 635 | binding |
| rapamycin | YJR066W | TOR1 | 14 | 328 | 478 | 18 | 20 | 26 | 34 | Zheng & Schreiber 1997 | 988 | inhibition, binding |
| staurosporine | YGR040W | KSS1 | 1024 | 710 | 994 | 478 | 1081 | 1014 | 1137 | STITCH | 200 | binding |
| thiabendazole | YML124C | TUB3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | Chaklan et al 2008 | | |
| trichostatin A | YGL194C | HOS2 | 920 | 83 | 22 | 116 | 133 | 187 | 51 | STITCH | 903 | binding |
| trichostatin A | YNL021W | HDA1 | 7 | 102 | 101 | 8 | 6 | 12 | 11 | STITCH | 844 | binding |
| trichostatin A | YNL330C | RPD3 | 123 | 496 | 377 | 332 | 140 | 243 | 219 | STITCH | 949 | binding |
| trichostatin A | YPR068C | HOS1 | 892 | 689 | 404 | 668 | 632 | 973 | 708 | STITCH | 474 | inhibition, binding |
| vitamin K3 | YPL091W | GLR1 | 573 | 254 | 193 | 507 | 281 | 442 | 366 | STITCH | 710 | binding |
| wortmannin | YJR066W | TOR1 | 602 | 197 | 128 | 386 | 203 | 378 | 258 | STITCH | 463 | binding |
| | | **MEAN OF RANKS** | **361** | **421** | **351** | **250** | **271** | **286** | **266** | | | |

[1] Reference according to Cokol et al Mol. Syst. Biol. 2011 or STITCH database.

[2] Specified if interaction exists in STITCH database.

[3] Specified if action is defined in STITCH database.

| HAPLOID DATASET | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Chemical** | **Gene** | **Gene name** | **FD-score** | **p-score** | **p*-score** | **I-score** | **I*-score** | **SR-score** | **SR*-score** | **Reference[1]** | **STITCH score[2]** | **STITCH action[3]** |
| amsacrine | YOL006C | TOP1 | | | | | | | | STITCH | 822 | binding |
| atorvastatin | YML075C | HMG1 | | | | | | | | Hillenmeyer et al 2010 | 637 | binding |
| benomyl | YML124C | TUB3 | 13 | 1 | 1 | 2 | 2 | 2 | 2 | Hillenmeyer et al 2010 | 748 | |
| caffeine | YJR066W | TOR1 | 1 | 9 | 14 | 1 | 1 | 1 | 1 | Kuranda et al 2006 | 459 | inhibition |
| camptothecin | YOL006C | TOP1 | 1252 | 185 | 113 | 29 | 25 | 358 | 221 | Pommier et al 1998 | 968 | binding, inhibition |
| cantharidin | YDL134C | PPH21 | | | | | | | | STITCH | 449 | binding |
| carbendazine | YML124C | TUB3 | | | | | | | | Hillenmeyer et al 2010 | | |
| carmustine | YPL091W | GLR1 | | | | | | | | Hillenmeyer et al 2010 | 509 | inhibition |
| caspofungin | YLR342W | FKS1 | 146 | 12 | 10 | 34 | 35 | 25 | 24 | Lesage et al 2004 | 708 | |
| colchicine | YML124C | TUB3 | | | | | | | | Hillenmeyer et al 2010 | 325 | inhibition, binding, activation |
| doxorubicin | YOL006C | TOP1 | | | | | | | | STITCH | 578 | binding |
| fenbendazole | YML124C | TUB3 | | | | | | | | Hillenmeyer et al 2010 | | |
| FK 506 | YLR433C | CNA1 | 643 | 219 | 326 | 450 | 571 | 402 | 551 | STITCH | 979 | inhibition, binding |
| geldanamycin | YPL240C | HSP82 | 63 | 1251 | 1249 | 1118 | 1045 | 153 | 147 | Mandal et al 2008 | 924 | binding |
| hydrogen peroxide | YDR256C | CTA1 | 536 | 207 | 26 | 392 | 81 | 363 | 62 | STITCH | 997 | binding, catalysis |
| hydrogen peroxide | YGR088W | CTT1 | 250 | 508 | 693 | 589 | 738 | 405 | 454 | STITCH | 996 | binding, catalysis |
| hygromycin | YJR075W | HOC1 | 7 | 67 | 87 | 7 | 12 | 11 | 19 | Kim et al 2006 | | |
| lithium | YHR046C | INM1 | | | | | | | | Murray & Greenberg 2000 | 978 | inhibition |
| lovastatin | YML075C | HMG1 | | | | | | | | Rine et al 1983 | 895 | inhibition, binding |
| mebendazole | YML124C | TUB3 | | | | | | | | Laclette et al 1980 | | |
| nocodazole | YML124C | TUB3 | 6 | 1 | 1 | 1 | 2 | 1 | 2 | Laclette et al 1980 | 488 | inhibition, activation |
| pentamidine | YOR266W | PNT1 | 17 | 450 | 351 | 143 | 116 | 37 | 38 | Luesch et al 2005 | | |
| radicicol | YPL240C | HSP82 | 171 | 1256 | 1256 | 1161 | 1161 | 359 | 370 | Taldone et al 2009 | 635 | binding |
| rapamycin | YJR066W | TOR1 | 1 | 8 | 20 | 1 | 1 | 1 | 1 | Zheng & Schreiber 1997 | 988 | inhibition, binding |
| staurosporine | YGR040W | KSS1 | 598 | 482 | 366 | 606 | 477 | 684 | 541 | STITCH | 200 | binding |
| thiabendazole | YML124C | TUB3 | | | | | | | | Chaklan et al 2008 | | |
| trichostatin A | YGL194C | HOS2 | 99 | 10 | 10 | 26 | 28 | 17 | 20 | STITCH | 903 | binding |
| trichostatin A | YNL021W | HDA1 | 729 | 223 | 71 | 333 | 162 | 410 | 145 | STITCH | 844 | binding |
| trichostatin A | YNL330C | RPD3 | 251 | 27 | 27 | 81 | 67 | 51 | 52 | STITCH | 949 | binding |
| trichostatin A | YPR068C | HOS1 | 219 | 217 | 396 | 265 | 407 | 220 | 317 | STITCH | 474 | inhibition, binding |
| vitamin K3 | YPL091W | GLR1 | | | | | | | | STITCH | 710 | binding |
| wortmannin | YJR066W | TOR1 | 2 | 120 | 43 | 4 | 2 | 3 | 3 | STITCH | 463 | binding |
| | | **MEAN OF RANKS** | **263** | **276** | **266** | **276** | **260** | **184** | **156** | | | |

[1] Reference according to Cokol et al Mol. Syst. Biol. 2011 or STITCH database.

[2] Specified if interaction exists in STITCH database.

[3] Specified if action is defined in STITCH database.

**Supplementary Table S4.** Internal evaluation of the scoring approaches in the homozygous dataset

| **HOMOZYGOUS** | *FD-score neg* | FD-score pos | **FD-score abs** | *ρ-score* | **ρ*-score** | I-score neg | I-score pos | *I-score abs* | **I*-score** | *SR-score neg* | SR-score pos | SR-score abs | **SR*-score** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *FD-score neg* | 0,790 | 0,039 | -0,002 | 0,001 | -0,007 | -0,007 | 0,032 | -0,004 | -0,015 | -0,002 | 0,029 | -0,002 | -0,009 |
| FD-score pos | **0,000** | 0,751 | -0,041 | -0,038 | -0,046 | -0,047 | -0,007 | -0,044 | -0,054 | -0,041 | -0,010 | -0,042 | -0,048 |
| **FD-score abs** | 0,472 | **0,000** | 0,792 | 0,003 | -0,005 | -0,005 | 0,034 | -0,002 | -0,013 | 0,000 | 0,031 | 0,000 | -0,007 |
| *ρ-score* | 0,898 | **0,000** | 0,712 | 0,789 | -0,008 | -0,008 | 0,031 | -0,005 | -0,016 | -0,003 | 0,028 | -0,004 | -0,010 |
| **ρ*-score** | 0,411 | **0,000** | 0,564 | 0,077 | 0,797 | -0,001 | 0,039 | 0,002 | -0,008 | 0,005 | 0,036 | 0,004 | -0,002 |
| I-score neg | 0,091 | **0,000** | 0,275 | 0,188 | 0,933 | 0,797 | 0,039 | 0,003 | -0,008 | 0,005 | 0,037 | 0,005 | -0,001 |
| I-score pos | **0,000** | 0,246 | **0,000** | **0,000** | **0,000** | **0,000** | 0,758 | -0,036 | -0,047 | -0,034 | -0,003 | -0,034 | -0,041 |
| *I-score abs* | 0,401 | **0,000** | 0,626 | 0,364 | 0,701 | 0,324 | **0,000** | 0,794 | -0,011 | 0,002 | 0,034 | 0,002 | -0,004 |
| **I*-score** | **0,000** | **0,000** | **0,006** | **0,020** | 0,180 | **0,003** | **0,000** | **0,006** | 0,805 | 0,013 | 0,044 | 0,013 | 0,007 |
| *SR-score neg* | 0,668 | **0,000** | 0,995 | 0,514 | 0,404 | 0,075 | **0,000** | 0,503 | **0,001** | 0,792 | 0,031 | 0,000 | -0,007 |
| SR-score pos | **0,000** | **0,043** | **0,000** | **0,000** | **0,000** | **0,000** | 0,493 | **0,000** | **0,000** | **0,000** | 0,761 | -0,032 | -0,038 |
| SR-score abs | 0,647 | **0,000** | 0,932 | 0,461 | 0,450 | 0,180 | **0,000** | 0,486 | **0,003** | 0,839 | **0,000** | 0,792 | -0,006 |
| **SR*-score** | 0,066 | **0,000** | 0,201 | 0,108 | 0,702 | 0,727 | **0,000** | 0,302 | **0,015** | **0,031** | **0,000** | 0,085 | 0,799 |

Diagonal: Area under curve (AUC); all AUCs are better than random ($p < 0.001$).

Above diagonal: Pair-wise differences of AUCs (row-column).

Below diagonal: p-values for the score differences.

ρ*-score: correlation based on 75 % of pairs using the negative tail of the SGA dataset.

I* and SR*-scores: calculated using the negative tail of the FD-score and the ρ*-score.

Basic versions of each of the scoring approach are in italic; the best variations are bolded.

**Supplementary Table S5.** Internal evaluation of the scoring approaches in the heterozygous dataset

| **HETEROZYGOUS** | *FD-score neg* | **FD-score pos** | FD-score abs | ***ρ-score*** | ρ*-score | I-score neg | I-score pos | ***I-score abs*** | I*-score | *SR-score neg* | **SR-score pos** | SR-score abs | SR*-score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| *FD-score neg* | 0,640 | -0,029 | -0,023 | 0,003 | 0,008 | 0,003 | -0,007 | -0,009 | 0,006 | 0,004 | -0,013 | -0,008 | 0,005 |
| **FD-score pos** | **0,000** | 0,670 | 0,006 | 0,033 | 0,037 | 0,033 | 0,022 | 0,021 | 0,035 | 0,033 | 0,016 | 0,021 | 0,035 |
| FD-score abs | **0,000** | 0,193 | 0,663 | 0,026 | 0,031 | 0,026 | 0,016 | 0,014 | 0,029 | 0,027 | 0,010 | 0,015 | 0,028 |
| ***ρ-score*** | 0,569 | **0,000** | **0,000** | 0,637 | 0,005 | 0,000 | -0,011 | -0,012 | 0,002 | 0,000 | -0,016 | -0,011 | 0,002 |
| ρ*-score | 0,178 | **0,000** | **0,000** | 0,284 | 0,632 | -0,005 | -0,015 | -0,017 | -0,002 | -0,004 | -0,021 | -0,016 | -0,003 |
| I-score neg | 0,367 | **0,000** | **0,000** | 0,986 | 0,363 | 0,637 | -0,011 | -0,012 | 0,002 | 0,000 | -0,016 | -0,011 | 0,002 |
| I-score pos | 0,240 | **0,000** | **0,003** | **0,012** | **0,003** | 0,063 | 0,648 | -0,001 | 0,013 | 0,011 | -0,006 | -0,001 | 0,013 |
| ***I-score abs*** | 0,060 | **0,000** | **0,000** | **0,005** | **0,002** | **0,000** | 0,717 | 0,649 | 0,014 | 0,012 | -0,004 | 0,001 | 0,014 |
| I*-score | 0,123 | **0,000** | **0,000** | 0,663 | 0,589 | 0,436 | **0,029** | **0,000** | 0,635 | -0,002 | -0,019 | -0,014 | 0,000 |
| *SR-score neg* | 0,275 | **0,000** | **0,000** | 0,913 | 0,389 | 0,842 | **0,031** | **0,000** | 0,594 | 0,637 | -0,017 | -0,012 | 0,002 |
| **SR-score pos** | **0,035** | **0,000** | **0,049** | **0,000** | **0,000** | **0,002** | **0,026** | 0,245 | **0,001** | **0,000** | 0,653 | 0,005 | 0,018 |
| SR-score abs | 0,067 | **0,000** | **0,000** | **0,002** | **0,001** | **0,001** | 0,869 | 0,742 | **0,001** | **0,000** | 0,159 | 0,648 | 0,013 |
| SR*-score | 0,116 | **0,000** | **0,000** | 0,683 | 0,458 | 0,551 | **0,024** | **0,001** | 0,918 | 0,602 | **0,001** | **0,000** | 0,635 |

Diagonal: Area under curve (AUC); all AUCs are better than random ($p < 0,001$).

Above diagonal: Pair-wise differences of AUCs (row-column).

Below diagonal: p-values for the score differences.

ρ*-score: correlation based on 65 % of pairs using the negative tail of the SGA dataset.

I* and SR*-scores: calculated using the ρ*-score and the negative tail of the FD-score.

Basic versions of each of the scoring approach are in italic; the best variations are bolded.