Inference of RhoGAP/GTPase Regulation Using Single-cell Morphological Data from a Combinatorial Genetic Screen

Oaz Nir, Chris Bakal, Norbert Perrimon & Bonnie Berger

Supplementary figures and text:

Supplementary Figure 1	ROC curve for neural network based-alternative classification model
Supplementary Figure 2	Classification-based clustering algorithm for downstream TCs
Supplementary Table 1	Biologically validated RhoGAP/GTPase interactions and non-interactions
Supplementary Table 2	Classification of single-knockdown RhoGAP TCs to RhoGTPase overexpression TCs
Supplementary Table 3	Classification of single- and double-knockdown RhoGAP TCs to RhoGTPase overexpression TCs
Supplementary Table 4	Clustering of single-knockdown RhoGAP TCs
Supplementary Table 5	Classification of double-knockdown RhoGAP TCs to single-knockdown RhoGAP TCs
Supplementary Table 6	Alternative bootstrapping for mapping single-knockdown RhoGAP TCs to RhoGTPase overexpression TCs
Supplementary Table 7	Classification of the set of RhoGTPase overexpression TCs to itself
Supplementary Table 8	Robustness of classification to exclusion of data using jackknifing

Supplementary Methods

Data Normalization and Dimensionality Reduction

Normalization and dimensionality reduction was performed for the 273-TC dataset as follows. Each of the 145 raw features was normalized to have mean 0 and variance 1 across the full set of 12601 single cells. Normalization of the raw features was done to avoid inappropriately weighting some features over others (for example, because of arbitrary differences in unit measurements). Following normalization, dimensionality reduction was performed by computing principal components (PCs) for the full set of single-cell data, and then projecting each data point onto the first three PCs. Working in reduced feature space avoided inappropriately weighting particular morphological feature classes that are overrepresented in the set of raw features (for example, redundant measurements of nucleus shape).

Similarly, normalization was performed for the 90-TC dataset. Dimensionality reduction was performed using the first three PCs computed using the 273-TC dataset. We used principal components from the 273-TC dataset in order to readily compare RhoGAP knockdowns and RhoGTPase overexpression TCs; we used the larger dataset because it contained knockdown, overexpression, and control test data. We conducted robustness testing by varying the number of dimensions of reduced feature space and re-running the entire classification algorithm. Using greater than five principal components resulted in greatly decreased specificity for the classification algorithm, likely because the RhoGTPase overexpression TC point clusters are separated by large distances, causing the classifier to overfit upstream TCs. Using between two and five principal components resulted in the same predictions at optimal threshold, with the following (minor) differences from the 3-PC case. For two PCs, RhoGAP5A/Rac1 and

CdGAPr/Rac1 were also predicted; for four PCs, RhoGAP100F/Rac1 and RacGAP84C/Rho1 were not predicted; and for five PCs, RhoGAP5A/Rac1, CdGAPr/Rac1, and RhoGAPp190/Rac1 were also predicted, but RacGAP84C/Rho1 was not predicted. For definiteness, we chose to use three PCs for our final analysis and the exposition here, though it is possible in theory to optimize for the number of dimensions using the training data. In particular, the false positive rate was improved for the case of four or five PCs versus three PCs.

Mapping RhoGAP double-knockdowns to single-knockdowns to identify within-pathway genetic interactions

As an additional application of our classification model, we mapped the set of RhoGAP double-knockdowns (U) to RhoGAP single-knockdowns (D). Applying the model directly to the entire set, D, was not possible, because each element of D was not correctly mapped to itself. To remedy this, we clustered the single-knockdown TCs using a variant of EM designed to guarantee that, under the final clustering, all single-knockdown TCs would be correctly classified. The algorithm proceeds by iterating the following two steps, beginning with $k = 0$ and $D_0 = D$.

Iteration k:

(i) Map each element of D into D_k using the classification. If each element of D is mapped to the cluster containing it, set $\tilde{D} = D_k$ and exit.

(ii) Define D_{k+1} as follows. Let $f(D) = R_k = \{r_l\}$ denote the range of the classification of D mapped into D_k (i.e. the subset of D_k onto which elements of D were mapped in (i)). Then set

$$D_{k+1} = \left\{\cup f^{-1}(r_l)\right\}.$$

At each iteration, the elements of D mapped to the same target in D_k are grouped (by taking the union of single cells comprising each such element) into a single element of D_{k+1}. Note that upon termination of the algorithm, the clustering $\tilde{D}$ necessarily has the property that every element of D is classified to the cluster containing it. It is theoretically possible for the algorithm to enter a cycle (though this did not occur for our test data), in which case all elements forming a cycle are clustered together, thus allowing the algorithm to continue. In the worst case, the algorithm terminates by grouping all elements of D into a single element, which has to be mapped to itself.

For single-knockdown RhoGAP TCs, the clustering algorithm terminates with $\tilde{D} = D_2$, yielding a total of 5 clusters. (By way of comparison, for RhoGTPase overexpression TCs, the clustering algorithm terminates immediately, i.e. $\tilde{D} = D_0$.) The classification model was used to map all double-knockdown RhoGAP TCs into the set $\tilde{D}$ of clustered single-knockdown TCs.

Supplementary Results

Comparison with alternate methods

Mean scores and clustering-based approaches

It should be noted that using mean scores is isomorphic to certain clustering-based approaches. One can imagine constructing a classification scheme by defining cutoffs for linkage distances in a hierarchical clustering, for instance. Such an approach is equivalent to computing pairwise distances between mean TC feature scores and identifying the closest pairs. Clustering carries an added disadvantage, namely the possibility of conflating the classification of upstream TCs to downstream TCs; in clustering, all pairwise distances are considered and may influence the final clustering, whereas only pairwise distances between an upstream and downstream TC factor into the mean score method described above (and in our classification model).

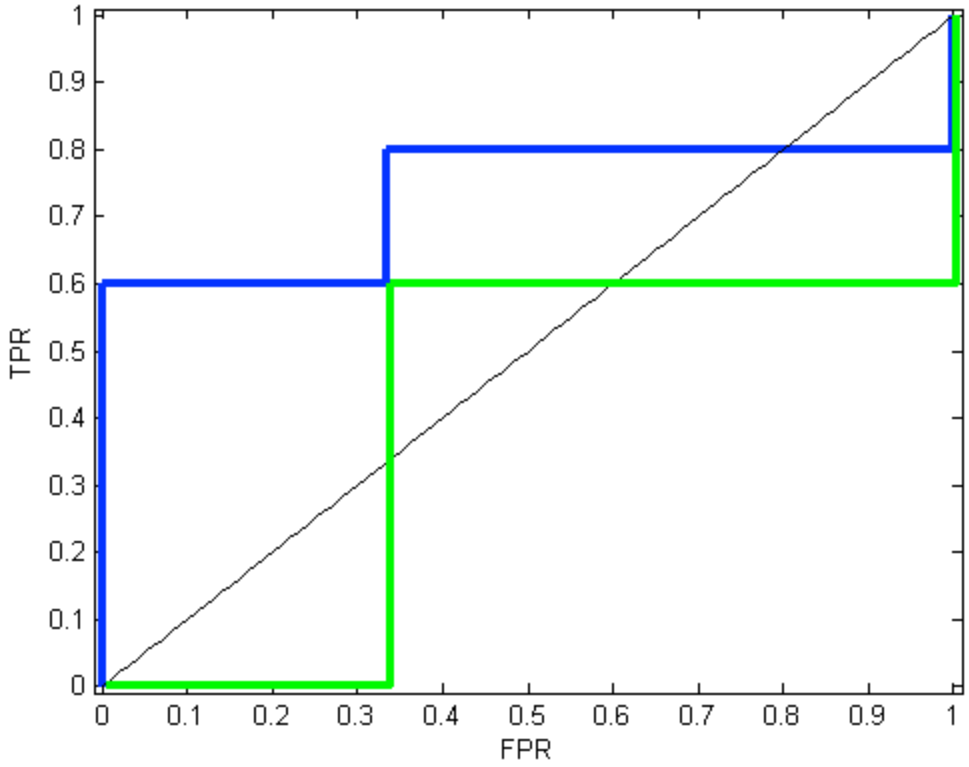
Incorporating other classifiers

In the main text, we described an alternative classification model based on computing Z-scores for each double RNAi TC for each neural network classifier. TCs with large Z-scores (as determined by Bonferroni correction at $p = .05$) for a given GTPase neural network were mapped to the GTPase under this classification model.

We also considered a more intricate model based on using neural network classifiers to perform a dimensionality reduction, as opposed to using PCA. In particular, using the two neural network classifiers for RacF28L and RhoF30L to represent single-cell morphology, we applied our classification model directly (with just Rac1 and Rho1 as potential downstream TCs); this was equivalent to performing a dimensionality reduction using scores from these two neural network classifiers as the basis for reduced-dimensional space, rather than principal components. However, at optimal threshold, this method attains 60% sensitivity and 67% specificity

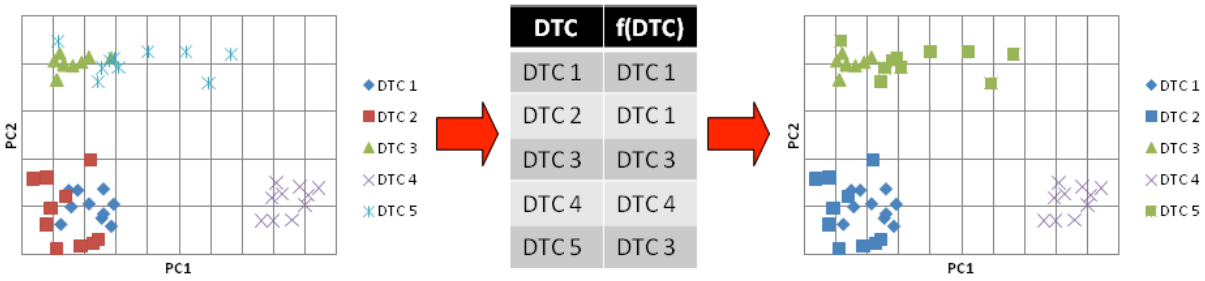
(**Supplementary Fig. 1**). As with the neural network Z-score method, this performance is poorer than that achieved by our classification model applied to PC-based data.

Supplementary Figure 1



Supplementary Fig. 1. ROC curve for neural network-based alternative classification model. An alternative classification model was constructed using neural network classifiers for RhoF30L and RacF28L to perform dimensionality reduction, in contrast to dimensionality reduction based on principal components, as in the main classification model proposed in this paper). The figure shows the ROC curve for the main classification model (blue) and for the alternative model (green). The neural network-based model cannot achieve greater than 60% sensitivity, and can only do so at 67% specificity. The PC-based model outperforms it and all other alternatives considered (see main text for discussion of alternative methods).

Supplementary Figure 2



Supplementary Fig. 2. Classification-based clustering algorithm for downstream TCs. The first iteration ($k = 0, D_0 = D$) of the algorithm is shown for hypothetical data consisting of a set, D, of five downstream TCs (DTCs). On the left, the single cells for all DTCs are shown in reduced feature space. In the middle (step (i) of the algorithm), the classification model is used to map each of the DTCs to the set D_0. In this example, each of DTC1, DTC3, and DTC4 is mapped to itself, but the other two DTCs are not. On the right (step (ii) of the algorithm), DTCs classified to the same target are consolidated into growing clusters. For the next iteration of the algorithm ($k = 1$), each of the five original DTCs is classified, but this time the target set consists of only three elements, as $D_1 = \{DTC1 \cup DTC2, DTC3 \cup DTC5, DTC4\}$.

Supplementary Table 1

Supplementary Table 1A:

RhoGAP	RhoGTPase
RacGAP50C	Rac1
RacGAP84C	Rac1
RhoGAP93B	Rac1
RhoGAPp190	Rho1
RacGAP50C	Cdc42

Supplementary Table 1B:

RhoGAP	RhoGTPase
RhoGAP5A	Rho1
RacGAP84C	Rho1
CdGAPr	Rho1

Supplementary Table 1. **(A)** Biologically validated RhoGAP/GTPase interactions among the 13 RhoGAPs and 3 RhoGTPases in our morphological datasets, taken from Flybase and BioGRID. **(B)** Biologically validated RhoGAP/GTPase non-interactions among the 13 RhoGAPs and 3 RhoGTPases in our morphological datasets taken from the literature.

Supplementary Table 2

RhoGAP	RhoGTPase	P-Score
RhoGAP92B	Rho1	0.000
RacGAP84C	Rho1	0.011
RhoGAPp190	Rho1	0.021
RhoGAP19D	Rho1	0.046
CdGAPr	Rac1	0.167
RhoGAP54D	Rac1	0.205
RacGAP50C	Rac1	0.231
RhoGAP93B	Rho1	0.437
RhoGAP5A	Cdc42	0.542
RhoGAP16F	Rho1	0.657
RhoGAP1A	Rho1	0.701
RhoGAP100F	Rac1	0.891
RhoGAP71E	Rho1	0.980

Supplementary Table 2. Classification of single-knockdown RhoGAP TCs into RhoGTPase overexpression TCs. The confidence scores shown here were computed using bootstrapping by drawing samples just from the 13 single-knockdown RhoGAP 13 TCs. Following Bonferroni correction, only the mapping of RhoGAP92B to Rho1 is significant at $p = .05$ (heavy shading). By considering the ROC curve (**Fig. 5C**), this model has optimal predictive power at a threshold of $p = .231$ (light shading), at which it correctly predicts 2/5 biologically validated interactions and 2/3 non-interactions. At this threshold, the model makes a total of 7 predictions. The probability of correctly predicting at least 2 out of 5 biologically validated interactions when making 7 predictions (out of 39 possibly) is $p = .21$.

Supplementary Table 3

Supplementary Table 3A

	Cd.r	100F	16F	p190	19D	1A	50C	54D	5A	71E	84C	92B	93B
Cd.r	Rac1	Rho1	Rho1	Rho1	Rho1	Rac1	Rac1	Cdc42	Rac1	Rho1	Rac1	Rho1	Rho1
100F		Rac1	Cdc42	Rho1	Rho1	Rho1	Rac1	Rho1	Rho1	**Rac1**	Rho1	Rho1	Rho1
16F			Rho1	Rho1	Rho1	Rac1	**Rac1**	Cdc42	Cdc42	Rac1	Rho1	Rho1	Rho1
p190				**Rho1**	Rho1	Rho1	Rac1	Rho1	Cdc42	Rho1	Rho1	Rho1	Rho1
19D					**Rho1**	Rho1	Rac1	n/a	Rac1	Rac1	Rho1	Rho1	Rho1
1A						Rho1	**Rac1**	Cdc42	Rho1	Rac1	Rho1	Rho1	Rho1
50C							Rac1	**Rac1**	Rac1	**Rac1**	**Rac1**	Rho1	**Rac1**
54D								Rac1	Rac1	Rac1	Rho1	Rho1	Rac1
5A									Cdc42	Rac1	Rho1	Rho1	Rho1
71E										Rho1	Rho1	Rho1	Rho1
84C											**Rho1**	Rho1	Rac1
92B												**Rho1**	Rho1
93B													Rho1

Supplementary Table 3B

	Cd.r	100F	16F	p190	19D	1A	50C	54D	5A	71E	84C	92B	93B
Cd.r	0.055	0.789	0.947	0.078	0.851	0.500	0.601	0.941	0.197	0.584	0.988	0.010	0.725
100F		0.749	0.103	0.869	0.112	0.405	0.026	0.822	0.360	**0.006**	0.349	0.003	0.312
16F			0.368	0.278	0.157	0.984	**0.000**	0.860	0.621	0.424	0.235	0.000	0.998
p190				**0.001**	0.854	0.886	0.047	0.410	0.743	0.316	0.174	0.000	0.681
19D					**0.012**	0.092	0.175	n/a	0.666	0.607	0.122	0.000	0.107
1A						0.505	**0.000**	0.541	0.635	0.109	0.274	0.003	0.846
50C							0.086	**0.000**	0.135	**0.000**	**0.000**	0.000	**0.023**
54D								0.091	0.087	0.468	0.239	0.000	0.643
5A									0.309	0.143	0.517	0.000	0.747
71E										0.965	0.186	0.146	0.313
84C											**0.003**	0.000	0.958
92B												**0.000**	0.000
93B													0.213

Supplementary Table 3. (A) Classification of single- and double-knockdown RhoGAP TCs into RhoGTPase overexpression TCs. **(B)** P-scores associated with classifications, as determined by bootstrapping with 1000 iterations. Classifications significant at $p = .0232$ are lightly shaded.

Double-knockdowns with RhoGAP92B as one of the knocked-down genes (the heavily shaded cells in the upper right of the matrix) were excluded, due to the fact that RhoGAP92B single-knockdown was mapped to Rho1 overexpression at Bonferroni-corrected $p = .05$ (see **Supplementary Table 2**). The threshold of $p = .0232$ was chosen based on ROC analysis to yield maximum sensitivity while simultaneously minimizing the false positive rate (see **Fig. 5C**).

Supplementary Table 4

Single-knockdown RhoGAP TC	Cluster Index
CdGAPr	1
RhoGAP92B	2
RhoGAPp190	3
RhoGAP19D	3
RacGAP84C	3
RhoGAP100F	4
RhoGAP16F	4
RhoGAP5A	4
RhoGAP93B	4
RhoGAP1A	5
RacGAP50C	5
RhoGAP54D	5
RhoGAP71E	5

Supplementary Table 4. Clustering of single-knockdown RhoGAP TCs. A clustering algorithm was designed was used to ensure that, following clustering, each single-knockdown RhoGAP TC would be mapped to the cluster containing it under the classification model. See text for details.

Supplementary Table 5

Supplementary Table 5A

	Cd.r	100F	16F	p190	19D	1A	50C	54D	5A	71E	84C	92B	93B
Cd.r	1	5	5	4	5	5	5	4	5	5	5	5	5
100F		4	**4**	4	3	5	4	5	4	5	3	2	5
16F			4	4	5	5	5	4	1	1	2	**2**	4
p190				3	4	4	5	4	4	5	4	**2**	4
19D					3	5	5	n/a	5	5	4	2	1
1A						5	5	4	4	5	5	2	5
50C							5	**5**	5	**5**	**5**	3	5
54D								**5**	4	5	4	2	5
5A									**4**	5	4	**2**	4
71E										5	5	5	**5**
84C											3	2	5
92B												**2**	3
93B													4

Supplementary Table 5B

	Cd.r	100F	16F	p190	19D	1A	50C	54D	5A	71E	84C	92B	93B
Cd.r	0.192	0.676	0.022	0.888	0.003	0.045	0.195	0.064	0.008	0.548	0.071	0.260	0.073
100F		0.277	**0.000**	0.795	0.502	0.778	0.059	0.820	0.798	0.052	0.756	0.911	0.485
16F			0.631	0.081	0.418	0.259	0.005	0.092	0.069	0.159	0.789	**0.000**	0.506
p190				0.300	0.491	0.429	0.446	0.792	0.097	0.031	0.640	**0.000**	0.092
19D					0.146	0.700	0.002	n/a	0.129	0.299	0.160	0.311	0.905
1A						0.597	0.014	0.005	0.538	0.001	0.824	0.283	0.334
50C							0.113	**0.000**	0.010	**0.000**	**0.000**	0.592	0.213
54D								**0.000**	0.012	0.464	0.489	0.888	0.002
5A									**0.000**	0.364	0.219	**0.000**	0.141
71E										0.013	0.673	0.791	**0.000**
84C											0.895	0.017	0.390
92B												**0.000**	0.265
93B													0.261

Supplementary Table 5. **(A)** Classification of double-knockdown RhoGAP TCs into single-knockdown RhoGAP TCs. The code (1-5) corresponds to the cluster index given in **Supplementary Table 4**. **(B)** P-scores associated with classifications, as determined by

bootstrapping with 1000 iterations. Classifications significant after Bonferroni correction at $p = .05$ are shaded. For discussion, see main text and **Table 3**.

Supplementary Table 6

RhoGAP	RhoGTPase	P-Score (Bootstrapping from single-knockdowns only)	P-Score (Bootstrapping from single- and double-knockdowns)
RhoGAP92B	Rho1	0.000	0.000
RacGAP84C	Rho1	0.011	0.003
RhoGAPp190	Rho1	0.021	0.001
RhoGAP19D	Rho1	0.046	0.012
CdGAPr	Rac1	0.167	0.055
RhoGAP54D	Rac1	0.205	0.091
RacGAP50C	Rac1	0.231	0.086
RhoGAP93B	Rho1	0.437	0.213
RhoGAP5A	Cdc42	0.542	0.309
RhoGAP16F	Rho1	0.657	0.368
RhoGAP1A	Rho1	0.701	0.505
RhoGAP100F	Rac1	0.891	0.749
RhoGAP71E	Rho1	0.980	0.965

Supplementary Table 6. Alternative bootstrapping for mapping single-knockdown RhoGAP TCs to RhoGTPase overexpression TCs. P-scores were calculated for classifications using two alternate methods: bootstrapping by drawing samples from single-knockdown TCs only (third column) and by drawing samples from both single- and double-knockdown TCs (fourth column). Sampling from both single- and double-knockdowns results in uniformly smaller p-scores, because including double-knockdowns increases the range of single-cell morphological phenotypes available. The ordering of the 13 RhoGAPs by p-score was maintained by shifting between the two bootstrapping schemes, save for a swap in the order of two adjacent pairs of interactions; the ROC curve is unchanged overall.

Supplementary Table 7

"Upstream" TC	Classification ("Downstream" TC)	P-Score
RhoF30L	RhoF30L	0.000
RacF28L	RacF28L	0.000
Cdc42Y32A	Cdc42Y32A	0.000

Supplementary Table 7. Classification of the set of RhoGTPase overexpression TCs to itself. Each of the three RhoGTPase overexpression TCs is correctly mapped. P-scores were calculated using bootstrapping with 1000 iterations, with samples drawn from the union (of single cells) of the three RhoGTPase overexpression TCs. The fact that $p < .001$ means that, for all three TCs, no bootstrapped samples of equal cell number were classified with greater confidence (mode frequency of the classification vector) than the actual set of cells. This provides significant confidence in the ability of the classification model to discriminate between the GTPase overexpression TCs. (In contrast, a preliminary clustering step was necessary when mapping (double-knockdown RhoGAP TCs) to the set of single-knockdown RhoGAP TCs).

Supplementary Table 8

Supplementary Table 8A

Group Definition	Group Size
Group 1 (p < .05/90)	14
Group 2 (.05/90 < p < .05)	10
Group 3 (.05 < p < .20)	19
Group 4 (.20 < p < .50)	16
Group 5 (.50 < p < .80)	16
Group 6 (.80 < p)	15
Total	**90**

Supplementary Table 8B

	90%	70%	50%	30%
Group 1 (p < .05/90)	0.999 (0.003)	0.999 (0.003)	0.998 (0.004)	0.984 (0.026)
Group 2 (.05/90 < p < .05)	0.997 (0.010)	0.984 (0.029)	0.959 (0.058)	0.928 (0.054)
Group 3 (.05 < p < .20)	0.954 (0.041)	0.927 (0.052)	0.907 (0.047)	0.823 (0.055)
Group 4 (.20 < p < .50)	0.856 (0.078)	0.813 (0.080)	0.784 (0.073)	0.735 (0.065)
Group 5 (.50 < p < .80)	0.724 (0.112)	0.707 (0.112)	0.679 (0.102)	0.598 (0.098)
Group 6 (.80 < p)	0.566 (0.116)	0.532 (0.096)	0.511 (0.077)	0.499 (0.053)

Supplementary Table 8. Robustness of classification to exclusion of data using jackknifing. For each single- and double-knockdown, 100 random samples consisting of X% (X = 30, 50, 70, 90) of the cells from that TC were selected and classified to the set of overexpression TCs. A consistency score was assigned based on the fraction of random samples correctly classified. **(A)** Single- and double-knockdowns were binned into groups depending on p-score of the true classification. **(B)** Mean consistency scores are shown here for all groups (standard deviations are shown in parentheses). Most importantly, classifications of the TCs that were classified at the optimal threshold of $p = .0232$ are extremely robust to data exclusion (top two lines in table). See also **Fig. 4D.**