

PHOCOS: inferring multi-feature phenotypic crosstalk networks

Yue Deng, Steven J. Altschuler* and Lani F. Wu*

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94158, USA

*To whom correspondence should be addressed

Abstract

Motivation: Quantification of cellular changes to perturbations can provide a powerful approach to infer crosstalk among molecular components in biological networks. Existing crosstalk inference methods conduct network-structure learning based on a single phenotypic feature (e.g. abundance) of a biomarker. These approaches are insufficient for analyzing perturbation data that can contain information about multiple features (e.g. abundance, activity or localization) of each biomarker.

Results: We propose a computational framework for inferring phenotypic crosstalk (PHOCOS) that is suitable for high-content microscopy or other modalities that capture multiple phenotypes per biomarker. PHOCOS uses a robust graph-learning paradigm to predict direct effects from potential indirect effects and identify errors owing to noise or missing links. The result is a multi-feature, sparse network that parsimoniously captures direct and strong interactions across phenotypic attributes of multiple biomarkers. We use simulated and biological data to demonstrate the ability of PHOCOS to recover multi-attribute crosstalk networks from cellular perturbation assays.

Availability and implementation: PHOCOS is available in open source at <https://github.com/AltschulerWu-Lab/PHOCOS>

Contact: steven.altschuler@ucsf.edu or lanifu@ucsf.edu

1 Introduction

A fundamental challenge in molecular biology is to understand how information flows through complex signal transduction networks (Collins *et al.*, 2007; Marbach *et al.*, 2012). Targeted perturbation experiments provide a powerful approach for identifying evidence of ‘crosstalk’ or influence between signaling components or pathways (Shmulevich *et al.*, 2002). Generally, these approaches allow one to assess whether perturbing one part of a signaling pathway alters the behavior of another part. Investigation of crosstalk could be at the level of individual components (e.g. specific genes or proteins) or collections of components (e.g. modules). Inferring crosstalk from experimental perturbation data conceptually requires two steps (Ku *et al.*, 2012): (i) quantification of perturbation-induced changes among selected biomarkers of the network; and (ii) application of statistical approaches that convert these observed changes into a crosstalk graph.

Most approaches for inferring crosstalk model interactions are based on a single attribute or ‘feature’ of a biomarker (Malioutov *et al.*, 2005; Markowitz *et al.*, 2007; Snijder *et al.*, 2013). This is sufficient for modeling interactions for many high-throughput assays, for example, RNA abundance in transcriptomics is commonly studied (Bader *et al.*, 2004). However, this approach is not adequate in cases where it is possible or desired to investigate

multiple different features per biomarker. In the case of proteomics, both abundance and level of modification (e.g. phosphorylation) could be simultaneously measured per protein species. In high-content imaging assays—the focus of our current study—hundreds of different features can be extracted for each biomarker per cell (Bakal *et al.*, 2007; Boland and Murphy, 2001; Carpenter *et al.*, 2006; Perlman *et al.*, 2004). Such features could include standard biomarker measurements, such as intensity (reflecting biomarker abundance or activity level) or localization (reflecting properties such as cytosolic versus nuclear levels, or unpolarized versus polarized states), as well as other measurements, such as texture or statistical properties of the brightest pixels. Thus, the challenge is that perturbation assays can cause simultaneous changes to multiple phenotypic features even on the same biomarker (Yang *et al.*, 2010). Information can be lost when modeling crosstalk for a single feature (Fig. 1 right cartoon): when only examining the intensity feature, a perturbation to biomarker ‘A’ may have no effect on biomarker ‘B’; however, the perturbation may alter the localization of B from wild type. For example, in the classic MAPK signaling cascade (Seeger and Krebs, 1995), activated MAPKKK protein may be observed to alter the activation level of its downstream MAPKK partner, whereas activated MAPKK protein may be observed to alter both the activity levels and localization of its downstream (transcription factor) MAPK partner.

Methods that capture and model the richness of network interactions are needed.

We propose a new framework to infer multi-feature phenotypic crosstalk (PHOCOS), which simultaneously monitors multiple phenotypic changes across biomarkers from targeted network perturbations. A critical application is to the analysis of high-content microscopy imaging data. Tools from computer vision can extract hundreds of phenotypic features per cell (Boland and Murphy, 2001). How do different observed properties of biomarkers relate to each other? PHOCOS searches for higher-resolution crosstalk not just between biomarkers, but between phenotypes of the biomarkers.

Below, we provide a description of our approach. We then show using synthetic data, where a ground truth is known, how the approach can be used to infer accurately multi-feature crosstalk networks from noisy data. Finally, we apply PHOCOS to predict crosstalk networks from the morphological and signaling changes of chemotaxing human neutrophils. PHOCOS provides a practical approach for inferring multi-feature crosstalk networks from assays that capture high-dimensional phenotypic descriptors of how different signaling components respond to perturbations.

2 Methods

2.1 An overview of PHOCOS

The computational framework of PHOCOS (Fig. 1, left boxes) has several main steps. In the first step (Section 2.2), we measured cellular responses to perturbations and extracted a number of single-cell phenotypic features to describe various properties of biomarker

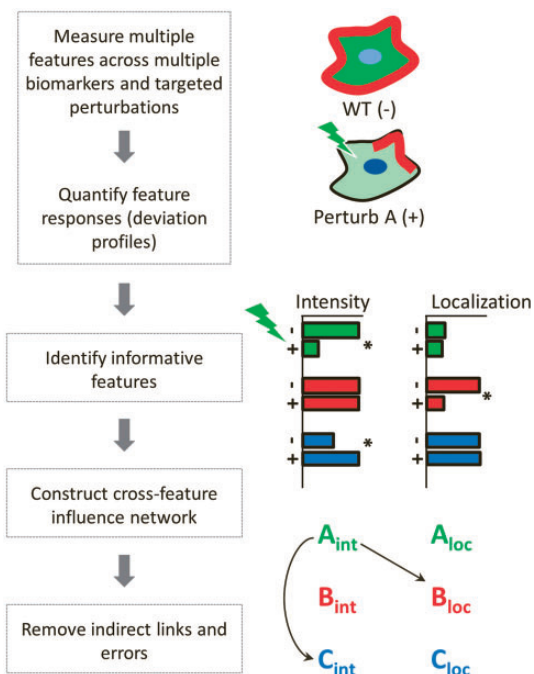


Fig. 1. Overview of PHOCOS. Left column: Flow chart. PHOCOS is used to infer multi-feature phenotypic crosstalk from high-content cellular images obtained from perturbation assay. The inputs to PHOCOS are cellular images with perturbation information; the output of PHOCOS is an inferred, sparse network of strong interactions between biomarkers based on observed phenotype. Right column: cartoon illustration of how perturbing the intensity of one biomarker (green) can affect either the localization (red) or the intensity (blue) of other biomarkers.

expression and localization patterns. Feature ‘deviation’ profiles were calculated by comparing non-perturbed versus perturbed phenotypes for all biomarkers and extracted features. In the second step (Section 2.2), clustering on the deviation profiles was performed to identify common phenotypic classes. The feature closest to each class center was selected as a surrogate of its phenotypic class. In the third step (Section 2.3), we inferred crosstalk from each observed perturbed biomarker to all other observed biomarkers. These influence links were merged across all perturbations to obtain a (possibly) dense and noisy ‘raw’ influence network. In the last step (Section 2.4), we performed sparse optimization as a heuristic to reduce indirect links and errors from the raw influence network. This final step produced a sparse crosstalk network capturing direct, strong influence links observed in the perturbed biological system.

2.2 Phenotypic deviation profiles and feature reduction

We applied PHOCOS to infer phenotypic crosstalk underlying the dynamic polarization processes of human neutrophils using a previously acquired microscopy image dataset (Ku *et al.*, 2012). Many molecular components involved in neutrophil polarity have been identified and organized into so-called ‘modules’ (Xu *et al.*, 2003) based on their function, localization and/or association with cytoskeletal structures, namely the ‘front’ (F) module (biomarker: actin), the ‘back’ (B) module (biomarker: phosphorylated Myosin II, pMyoII) and the ‘microtubule’ (MT) module (biomarker: tubulin). We note that below we interchangeably refer to these modules by their names (F, M, B) or by their biomarker readouts (actin, tubulin, pMyoII).

Briefly, in this dataset, human blood neutrophils were co-stained for biomarkers of actin, tubulin and pMyoII. Cells were stimulated with uniform f-Met-Leu-Phe (fMLP; 10 nM) at 37°C before formaldehyde fixation at different time points ranging from 0 to 600 s (control assay). Three sets of drug perturbations were used to disrupt the front (latrunculin A and jasplakinolide), back (Y27632 and calpeptin) and MT (nocodazole and Taxol) modules within the polarity-signaling network, as read out by actin, pMyoII and tubulin staining (respectively). Time course data were captured at 0, 15, 30, 45, 60, 90, 120, 180, 300, 450 and 600 s after stimulation with fMLP of both control and perturbation groups.

For each identified cell and biomarker, we extracted 76 features. We manually categorized features as ‘intensity’, ‘polarity’, ‘brightest pixel compactness’, ‘shape’, ‘Haralick’ or ‘Zernike’ based on the phenotype captured by the feature that was computed (Ku *et al.*, 2010; Loo *et al.*, 2007; Murphy, 2014).

For each drug and feature, ‘phenotypic deviation profiles’ were computed as previously described (Ku *et al.*, 2012). First, feature values were interpolated across time points to obtain continuous response curves. Consistent with previous work (Ku *et al.*, 2012), these time points were partitioned into either low-resolution (0-3, 3-7.5, 7.5-10 min) or high-resolution (0-1, 1-3, 3-5, 5-7.5, 7.5-10 min) periods as indicated in our studies below. These non-uniformly spaced periods were chosen to match the biphasic responses of neutrophils to chemoattractant. Second, for each of the periods, we computed ‘phenotypic deviation profiles’ $z_t^{(m,d,f)}$ by

$$z_t^{(m,d,f)} = \frac{A_t^{(m,d,f)} - A_t^{(m,e,f)}}{\text{std}(A_t^{(m,e,f)})},$$

where m , f and d represent the module, feature and drug perturbation, respectively, and $A_t^{(m,d,f)}$ (resp. $A_t^{(m,e,f)}$) is the area under the drug-perturbed (resp. control) response curve in period t . The

standard deviation $\text{std}(A_t^{(m,c,f)})$ in the denominator was calculated from 20 control replicates. The median phenotypic deviation profiles calculated from all drug replicates (nocodazole $n = 6$; Taxol $n = 3$; latrunculin A $n = 3$; jasplakinolide $n = 3$, Y27632 $n = 3$; calpeptin $n = 2$).

It is not to be expected that all 76 features would contain independent information. For instance, both the morphology and Zernike features, from the dataset above, could contain redundant information about the shape of a measured biomarker and could thus yield similar deviation profiles. Accordingly, unsupervised feature clustering was performed to identify common patterns of deviation profiles; this allowed us to reduce the collection of features by selecting a single representative from each cluster.

To achieve this goal, for each of the 76 features we created a long vector V^f ($\text{dim} = 90$) by stacking the deviation profiles obtained for each of the three modules, six drugs and five high-resolution periods (median values of replicates were used). We used k -means clustering (k was chosen using model-fit criteria) to identify common patterns among the deviation profiles; we refer to each cluster as a ‘phenotypic class’. The feature closest to the centroid was selected as the representative for the phenotypic class. We note that we chose to use the same set of features across all biomarkers; however, one could alternatively perform clustering on a per-module level to focus on module-specific biological phenotypes.

2.3 Learn multi-feature influence graph

We next learned an influence graph for each period based on the z-score profiles of representative features found in Section 2.2. The vertices of the graph are phenotypic nodes, which are a combination of biomarker and feature. The presence of an edge in the graph indicates that a perturbation to one node is observed to affect another, and the weight of the edge indicates the strength of this influence.

We represented the final phenotypic influence graph with $m \times k$ nodes, where m is the number of modules and k is the number of phenotypic classes. For each drug, feature and period, we inferred the strength of influence from the phenotypic node a^{f_1} (feature f_1 of module a) to b^{f_2} with the drug perturbation d_a (targeting module a) by:

$$W_t^{d_a}(a^{f_1}, b^{f_2}) = \frac{\gamma_t^{(a,d_a,f_1)}}{\sum_{i=1}^k \gamma_t^{(a,d_a,f_i)}} \gamma_t^{(b,d_a,f_2)}, \quad (1)$$

where $\gamma_t^{(a,d_a,f_1)} = |\mathcal{Z}_t^{(a,d_a,f_1)}| \delta(|\mathcal{Z}_t^{(a,d_a,f_1)}| > 1)$ and $\delta(g)$ is an indicator function whose value is 1 when the condition in the brackets is satisfied and zero otherwise. (The normalizing factor in the denominator reflects the fact that the observed change to phenotypic node b^{f_2} could only have come from features of a affected by perturbation d_a .)

The influence links discovered in each perturbation condition were then combined together in the form of a directed graph to represent influences among all the phenotypic nodes. Specifically, if a link was identified in multiple drug conditions, the maximal link strength across all conditions was adopted in the combined graph, which forms the matrix of observations:

$$D(a^{f_1}, b^{f_2}) = \max\{d_2\} W_t^{d_a}(a^{f_1}, b^{f_2}).$$

The rationale for taking the maximal link is that each module in this study was perturbed with two drugs of different mechanisms of action; we chose to accept any crosstalk revealed by either of the drugs. One multi-feature influence graph was inferred at each time point.

We note that we take into account the known targets of drugs in that we only search for a link from a^{f_1} to b^{f_2} if we know d_a is directly targeted to module a . There are a number of other computational approaches for graph inference, including the nested effects model (Markowitz et al., 2007) and, more generally, Bayesian inference approaches (Friedman et al., 2000). These approaches typically assume that there are a large number of biomarkers (e.g. gene expression levels) as well as a large number of perturbations, but the targets of the perturbations are generally not used explicitly in the inference of links. Our current approach is especially suitable for applications of high content imaging in which there are a relatively small number of observed biomarkers and perturbations can be selected to specifically target each biomarker. Incorporating knowledge of the target in the model provides higher confidence in directions of predicted crosstalk and greatly simplifies the computation.

2.4 PHOCOS graph reduction

We are interested in identifying how one observed node affects another observed node. For our purposes, we define a link as ‘direct’ if influence propagates directly from one observed node to a second observed node without passing through a third observed node.

We note that in biological networks, indirect effects could arise for multiple reasons, including perturbations that target multiple nodes, propagation of influence through unobserved nodes and propagation of effects through observed nodes. For the first case, we assume that perturbations were selected to specifically target a single network node (as is the case with our data). The second case is a general caveat of most network inference approaches (Barzel and Barabási, 2013), particularly when only limited numbers of nodes can be monitored (in fact, observed influence is likely to involve many unobserved intermediate molecular interactions). The focus of PHOCOS is on the third case, in which we develop an optimization approach to minimize the presentation of links found in Section 2.3 that arise via propagation through observed nodes (Fig. 2).

Indirect effects have been well described in a number of studies (Feizi et al., 2013; Weigt et al., 2009). Formally, indirect links may be described by the expansion

$$D_{\text{cmb}} = G + G_{\text{indr}} = G + G^2 + G^3 + \dots \quad (2)$$

where $G \in \mathbb{R}^{n \times n}$ is the ground truth direct-effect graph for all n nodes, and D_{cmb} is the combined graph containing both direct and indirect effects. We then make use of the closed form solution for

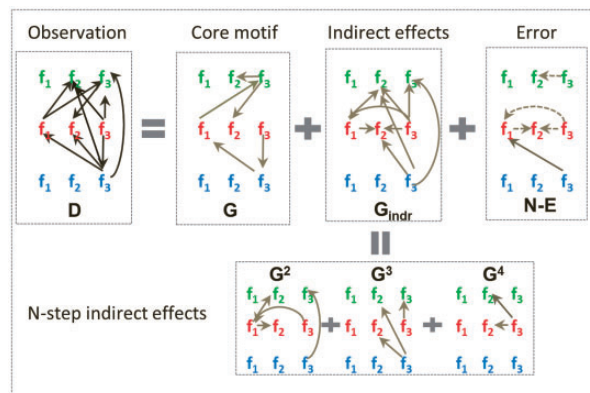


Fig. 2. An observed dense graph may be a combination of multiple effects, including N -step indirect effects and errors from missing (dashed arrows) or noisy links. These effects can mask the true, ‘core’ motif that underlies the network.

the infinite series observed in previous work (Feizi *et al.*, 2013; Spirtes *et al.*, 2000):

$$D_{\text{cmb}} = G(I - G)^{-1} \quad (3)$$

where $I \in \mathbb{R}^{n \times n}$ is the identity matrix. The convergence of this series requires that the absolute value of the largest absolute eigenvalue of G is less than 1 (discussed below).

The relationship between direct and indirect effects given above describes the situation in a theoretical setting. However, in practice, the situation can be much more complicated owing to missing links and noise. A ‘missing link’ problem arises when some links, E , on D_{cmb} cannot be measured directly. In experimental data, perturbations can reveal potential links between modules; however, it is difficult to deduce intra-module crosstalk (Fig. 2, nodes of same color). (Even a perturbation targeted to one module can affect multiple of its features, and it is difficult to determine influence of relationships among affected features.) Noise, N (e.g. white Gaussian noise), is also an inevitable factor making the practical influence of graph different from the theoretical case.

By taking these issues into consideration, we modeled connections between the observed graph D with D_{cmb} by: $D = D_{\text{cmb}} - E + N$. By replacing this term in Equation (3), we obtain

$$(D + E)(I - G) - G = N(I - G) \quad (4)$$

By using the L_2 norm on both sides of (4), the following inequality is obtained:

$$\|(D + E)(I - G) - G\|_2^2 = \|N(I - G)\|_2^2 \leq \|N\|_2^2 \|(I - G)\|_2^2 = \tau \quad (5)$$

It is easy to verify the problem in (5) is bounded above under the assumption that variance of noise is bounded. As mentioned above, there is a requirement that the largest absolute eigenvalue of G is less than 1. However, G is not directly observable from data. Therefore, in cases when this condition does not hold, heuristic approaches have been developed (Feizi *et al.*, 2013) to normalize the observed matrix D . In our analysis of the neutrophil data, D satisfied the constraint without the need for normalization.

A reasonable assumption (or hope) is that missing links only occupy a small portion of links on the true graph (i.e., E is sparse). Additionally, we seek to discover the major influences of information among observed nodes (i.e. G is sparse). Inspired by the field of compressed sensing (Baraniuk, 2007), we imposed a sparse prior on both G and E , which then allowed us to convert (5) to a dual-sparse optimization formalism in which we seek solutions to:

$$\begin{aligned} & \min(\|G\|_1 + \|E\|_1) \\ & \text{s.t.} \|(D + E)(I - G) - G\|_2^2 \leq \tau \end{aligned} \quad (6)$$

In the objective function, the L_1 norm is used to encourage sparse structures of G and E . (The L_1 norm $\|A\|_1 = \sum_{i,j} |A_{ij}|$ for a matrix A .)

To make the optimization process tractable, we made use of an approach (Chen *et al.*, 1998) to transform the objective function to

$$\min\left(\|G\|_1 + \|E\|_1 + \beta\|(D + E)(I - G) - G\|_2^2\right), \quad (7)$$

where the relaxation parameter β balances the reconstruction errors and the sparseness of the final reconstruction. This relaxation strategy is widely used in the signal processing field (Malioutov *et al.*, 2005; Yang *et al.*, 2010) to handle small noise and errors in the original data.

Problem in (7) falls into the sparse learning paradigm in machine learning. We used the prevalent Alternative Directional Minimization (ADM) method (Boyd *et al.*, 2011) to minimize the dual-sparse optimization problem. In detail, two auxiliary variables \hat{G} and \hat{E} are introduced to overcome the problem that the gradient of an L_1 norm is hard to derive. With the standard ADM formulation, the problem is converted to an unconstrained problem by minimizing L

$$\begin{aligned} L = & \|\hat{G}\|_1 + \|\hat{E}\|_1 + \beta\|(D + E)(I - G) - G\|_2^2 \\ & + \frac{\mu}{2}\left(\|G - \hat{G} + \mu^{-1}\Lambda_1\|_2^2 + \|E - \hat{E} + \mu^{-1}\Lambda_2\|_2^2\right), \end{aligned} \quad (8)$$

where Λ_1 and Λ_2 are two Lagrangian multipliers, $\mu > 0$ is the penalty parameter which will be incrementally increased during the optimization iterations. The four variables in (8) are sequentially updated in the iteration until convergence is achieved. In particular, we made use of the following numerical approach. A closed-form solution for the updating rules of \hat{G} and \hat{E} have been previously described (Boyd *et al.*, 2011; Deng *et al.*, 2013). The updating rules for G and E only involve the L_2 norm, which are easy to derive (Petersen and Pedersen, 2008). Lagrangian Multipliers, Λ_1 and Λ_2 , are updated according to dual ascending rules. The whole optimization was regarded as converged when the change of the objective function is less than a threshold, chosen empirically as 10^{-4} . After convergence, we further set links with weak strengths in G to zero to remove unexpected artifacts from numerical iterations.

3 Results

3.1 Test of graph reduction using simulated data

Before analyzing our experimental dataset, we performed *in silico* simulations to test the ability of PHOCOS to recover sparse direct effects from noisy observations when ground truth is known. First, we started with a graph structure containing three modules and five ‘common’ features per module (motivated by our experimental data used below), leading to a graph with 15 nodes (but no edges yet) and fraction S of edges randomly selected to be present. Here, the parameter S controls the sparsity of the graph. The strength of the edges was randomly drawn in the range of $(0,1]$. This graph, G , was considered as the ground truth (Fig. 3a, left).

Next, we generated simulated data via a process, inspired by a previously established data generative process (Feizi *et al.*, 2013). First, we added indirect effects as in Equation (2) to obtain D_{cmb} . As previously discussed (Section 2.4), we cannot infer intra-module links from our experimental data, and hence, all such simulated links were removed. Second, a noise matrix N was generated with $p\%$ (noise rate) of non-zero entries and edge strength chosen from a Gaussian distribution $N(0, 0.2)$. Then N was added to D_{cmb} , and all links with negative strength were removed. Third, a missing-link matrix E was computed by removing $m\%$ of inter-module links in D_{cmb} (Fig. 3a, same versus different colored nodes). The final ‘observed’ graph was given by $D = D_{\text{cmb}} + N - E$ (Fig. 3a, ‘Observed graph’).

We then applied PHOCOS to recover the ground truth (Fig. 3a, ‘Result using PHOCOS’). A major design of PHOCOS is its ability to robustly remove noise and identify missing links in the observed graph D . For comparison in our study, we compared the performance of PHOCOS with an existing approach for inferring direct-effect graphs (Feizi *et al.*, 2013). It is worth noting that this previous approach recovers direct effects by a closed-form solution (CF) without explicitly considering errors and missing links. Comparison of sample graphs showed that PHOCOS achieved much better

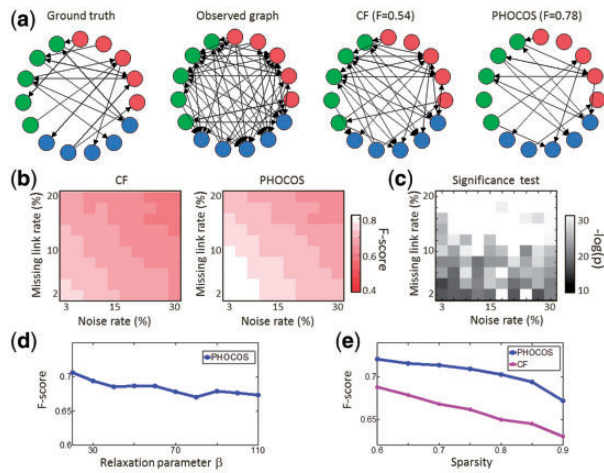


Fig. 3. Simulation results of PHOCOS graph reduction. (a) Shown is a simulated graph of three biomarkers and five features per marker. From left to right are ground truth graph (direct effect), noisy and dense experimental observation (input to PHOCOS), the direct effect inferred by the CF and the recovered cross-talk graph from PHOCOS. (b) Large-scale simulations of CF (left) and PHOCOS (right) graph recovery methods for varying noise rates and missing link ratios of graphs with 0.8 sparsity. (c) The P -value of the Wilcoxon rank-sum test for F-score results in b. (d) Performance of PHOCOS with different values of β . (e) Assessment of graph recovery for varying levels of sparsity.

recovery accuracy than CF (Fig. 3a). In the example shown, the ground truth graph contains 17 links. PHOCOS recovered 16 links, with 13 true positive, 3 false positives and 4 false negatives. The CF method recovered a dense graph of 35 links with 14 true positive, 21 false positives and 3 false negatives.

Next, we performed simulations to assess the performance of PHOCOS. First, we incrementally increased the missing link and noise rates. This allowed us to compare PHOCOS and CF for different error levels. Each entry in the heat-map (Fig. 3b) was obtained by averaging F-scores (which report accuracy; defined by the harmonic mean of precision and recall) calculated from 200 randomly generated graphs with $S = 0.8$. These simulations showed that PHOCOS has an improved ability to accurately recover ground truth networks (F-score = 0.71, compared with F-score = 0.64 for CF). We further performed a Wilcoxon rank-sum test to compare the F-score results of PHOCOS to CF (Fig. 3c) on each of the 200 simulations per entry in Figure 3b (i.e. each entry in the heat-map). The results suggested that PHOCOS significantly outperformed CF.

Second, the robustness of parameter selection was also examined for the β (see Equation (7)) in the optimization step. We computed the corresponding F-score for both the PHOCOS and CF approaches as we varied β across a realistic range for PHOCOS (Fig. 3d). (If β is too small, then the optimization favors a trivial solution of no links and if β is too high there is no penalty for sparsity of G and E.) We found the accuracy curve of PHOCOS is reliable across a range of parameter settings for β .

Third, we assessed the performance of PHOCOS based on the sparseness of the ground truth graph. Specifically, we varied the level of sparsity, S , and repeated the previous numerical simulations. Again, we found that PHOCOS consistently outperformed CF across levels of sparsity and did particularly well for the difficult case of dense graphs (Fig. 3e).

3.2 Feature selection for neutrophil polarization network

We applied our approach to the dataset of perturbed primary human neutrophils (see Fig. 4a and (Ku et al., 2012) for sample

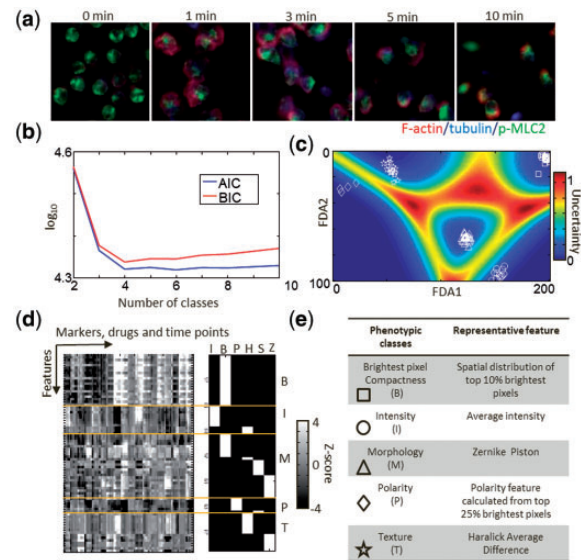


Fig. 4. Phenotypic class discovery from neutrophil polarization assays. (a) Sample images of polarized Neutrophil captured at different time points. (b) An ‘optimal’ number of classes was chosen based on two information-theoretic model selection criteria: AIC and BIC. Both indicators suggested five classes. (c) The visualization of class margins and corresponding projections of each feature vector in a 2D space suggested strong class separation (white symbols indicate feature class). (d) Unsupervised clustering showed common phenotypic attributes of grouped features in each of the five clusters, which provided intuitive cluster labels. (e) The feature nearest each cluster center was selected as cluster representative (symbols are as in c).

images). We measured 76 features that were extracted from each of three biomarkers. We next applied our feature reduction approach (Section 2.2) to identify archetypal phenotypes that showed distinct patterns of deviations in response to perturbations. In particular, we used k-mean clustering (with 20 replicates to avoid initialization artifacts). Two model selection criteria, Akaike information criterion (AIC) and Bayesian information criterion (BIC), were used to determine, k , the number of clusters (Burnham and Anderson, 2002). Both model fit criteria suggested $k = 4\sim 6$ as optimal cluster numbers (Fig. 4b). Therefore, we fixed $k = 5$ in the analysis. Clustering results were visualized with class separation boundaries and plotted in a projected 2D space using Fisher Discriminant analysis (Fig. 4c). This visualization served as a sanity check to show that our selected five clusters are well separated.

Which types of features were selected? In our feature extraction step, the attribute of each feature was annotated according to its implementation into one of six *a priori* classes (Fig. 4d, right panel, top axis): intensity (I), brightest pixel compactness (B), polarity (P), Haralick (H), shape (S) and Zernike (Z). We note that the intensity, shape, Haralick and Zernike features were computed as previously described (Boland and Murphy, 2001). The ‘brightest pixel compactness’ features measure properties of the spatial coverage of the smallest region covering the brightest pixels (Ku et al., 2010). ‘Polarity’ features also analyze the spatial distribution of brightest pixels, but additionally takes into account cell boundary properties (Ku et al., 2012).

Our unsupervised clustering revealed strong consistency in at least three of the five clusters in terms of feature-type composition (Fig. 4d). All the features in cluster (B) are from the hand-annotated brightest pixel compactness class. Cluster (I) contains a large portion of intensity features and a small portion of Haralick features

(Haralick features also contain intensity information). Most of the features in cluster (P) are polarity features. For the remaining two clusters: cluster (M) contains morphology, Zernike coefficients and brightest pixel compactness features, which mostly relate to the morphology of the biomarker; and cluster (T) contains Haralick and Zernike features, which relate to properties of texture. The feature closest to each cluster center mean was selected as a representative (Fig. 4e). The deviation curves of these representative features were used in subsequent steps to identify crosstalk.

3.3 Inference of multi-feature neutrophil crosstalk network

We applied PHOCOS to investigate the polarization networks of human neutrophils at different time intervals (the relaxation parameter β was empirically set to 100). Multi-feature influence graphs were computed (Section 2.3). Then, PHOCOS was applied to the influence graphs obtained in each time interval to recover time-varying crosstalk networks. We noted in the reduced graph, there were still some links with small weights, which were regarded as numerical artifacts from the optimization. We heuristically removed the link if its weight is $<60\%$ of the mean weight of the graph.

We followed previous work (Ku *et al.*, 2012) to divide the dynamic polarization process of neutrophil into three periods corresponding to the initialization (0~180s), establishment (180~450s) and maintenance (450~600s) phases of cell polarization. We then used PHOCOS to infer graphs for each period (Fig. 5, top graphs). To help visualize and interpret these results, we paired these results with graphs that summarized the total influence (across all features) observed between modules (Fig. 5, bottom graphs).

We calculated the fraction of all cross-feature links between any pair of modules. In detail, after PHOCOS reduction of the t^{th} temporal graph, we counted the total number of cross-feature links between biomarkers a and b by:

$$N_t(a, b) = \sum_{f_1, f_2} \text{sign}[w_t(a^{f_1}, b^{f_2})],$$

where $w_t(a^{f_1}, b^{f_2})$ is the weight between phenotypic nodes a^{f_1} and b^{f_2} . Then we normalized $N_t(a, b)$ by the total number of links (K_t) on the temporal graph at time t as $\frac{N_t(a, b)}{K_t}$. We drew a link between modules if the normalized number of cross-feature links exceeded a heuristically determined threshold of 0.2 (Fig. 5, bottom graphs).

From these results, it can be seen that patterns of multi-feature crosstalk networks dynamically evolve across different periods. This agrees with previous results on the same dataset from single-feature analysis (Ku *et al.*, 2012). More interestingly, the major information flow among these modules appears to evolve from an

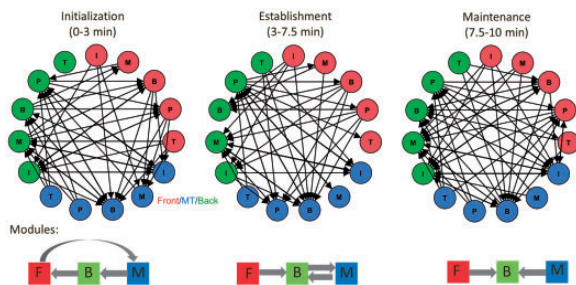


Fig. 5. Multi-feature graphs recovered by PHOCOS in different periods. In each column, the top panel shows the multi-feature graph and the bottom panel summarizes the major information flow in the corresponding period.

initial large feedback loop to a fan-in network motif (Liberali *et al.*, 2014).

As a different approach to interpret our time-varying multi-feature crosstalk graphs, we examined only the links that were persistent across all time intervals (Fig. 6a, left panel). We applied PHOCOS to this graph (Fig. 6a, right panel) and extracted single-feature, crosstalk sub-networks (Fig. 6b). We observed that the crosstalk diagrams of different features are different. The intensity graph exhibits a cascade topology from F to B to M. There is only one persistent link identified in the ‘brightest pixel’ network from F directly to M. The polarity graph also exhibits a cascade topology, but from M to B to F, which is in the reverse direction of the intensity graph. These intensity and polarity feature networks largely agree with previous work obtained by single-feature crosstalk inference (Ku *et al.*, 2012). There are also many new predictions about the presence or absence of information flow that are more difficult to assess directly from current literature. New experimental approaches, such as ontogenetic control of signaling components (Toettcher *et al.*, 2013), will make it more feasible in the future to design directed perturbations that alter one phenotype (e.g. spatial pattern of activation) of a biomarker while holding another phenotype (e.g. total level of activation) constant.

We also observed no persistent links for ‘texture’ and ‘morphology’ single-feature crosstalk diagrams. However, it is apparent that the texture and morphology features play important roles in the multi-feature network (Fig. 6a). As one example, reorganization of the MTs (node MT) was related to changes in the localization of p-MyosinII (node BP), which is consistent with current evidence that MTs transport signaling activators to the back of neutrophils (Wang *et al.*, 2013; Xu *et al.*, 2003). The observation that these types of interactions cannot be identified by single-feature analysis highlights the importance of conducting multi-feature network learning. Importantly, the inferred graphs serve as starting points for more targeted experimental validation studies that investigate how crosstalk regulates specific phenotypes, such as intensity versus localization, of biomarkers.

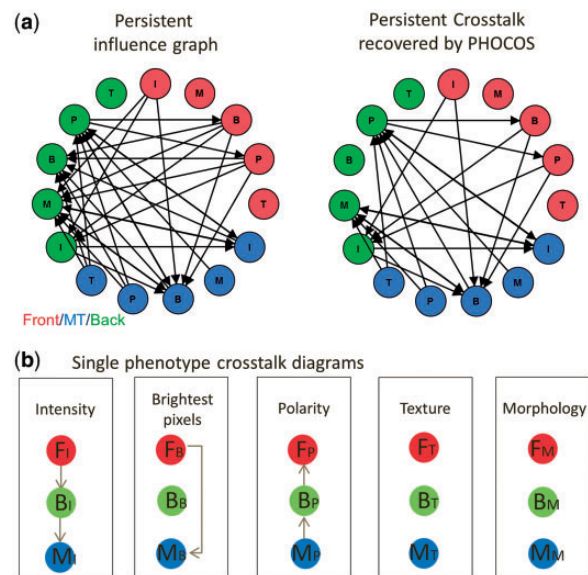


Fig. 6. Persistent crosstalk during neutrophil polarization. (a) Shown are links that appeared in all periods (Fig. 5) before and after graph reduction. (b) Single-feature persistent crosstalk diagrams identified from the multi-feature PHOCOS graph.

4 Robustness of PHOCOS

We investigated several aspects of robustness for graph inference using the neutrophil data. First, we considered an alternative division of periods. Based on the multiphasic responses of neutrophils to chemoattractant, we originally divided the polarization phases into the three periods. As comparison, we reanalyzed our data using a higher-resolution division into five periods (Fig. 7a; Section 2.2). We applied PHOCOS to obtain crosstalk networks in each period and visualized the major information flow as before (Fig. 5 bottom). Reassuringly, merged graphs based on their coarser periods (i.e. merging periods 1 and 2, and 3 and 4) are identical to the original coarse-grained graphs except for a single extra link from B to F during establishment (3-7.5 min). These experimental results support the intuition that dividing data into more precise time regions can provide more details about the graph evolution processes, while remaining consistent with more coarse-grained analysis of the dynamic processes. For the purposes of investigating robustness of PHOCOS, we chose to proceed with these higher-resolution periods.

Next, we considered using alternative features for PHOCOS inference. We sequentially changed the representative feature in each phenotypic cluster (Fig. 4d) and investigated how the graph inference results would be affected. In our previous analysis, the feature closest to the cluster center was chosen as a representative. Here, the j^{th} closest feature as the representative for each phenotypic class (Fig. 4d). Our investigations were performed independently on each previously identified phenotypic class. We applied PHOCOS to these new feature sets and inferred both the raw and the reduced (after PHOCOS reduction) graphs in each of the five periods. The Jaccard index was chosen to evaluate the similarity of graphs (based on the presence and absence of links) obtained with the different feature sets with the reference graphs (based on the first closet feature in each cluster). The Jaccard index scores (across the five periods) were averaged and summarized for both the raw (Fig. 7b, left) and reduced (Fig. 7b, right) graphs.

We noted that phenotypic classes I, B, P, T are quite robust. The Jaccard score curves are reliable even though we altered the representative feature in these clusters. However, crosstalk inferred using the morphological features M are less robust, reflecting the diversity of features in this class (Fig. 4d). Together, these investigations provide a general framework for studying robustness of PHOCOS to alternative choices of time resolution and feature choice.

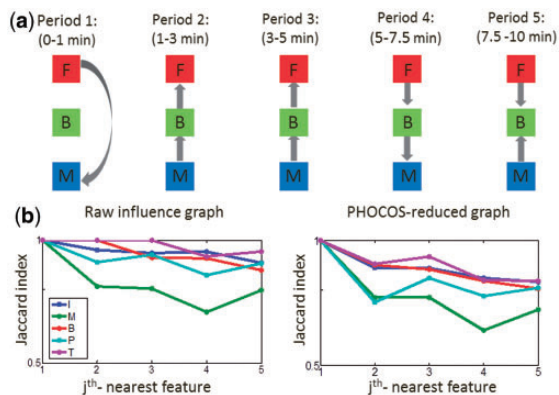


Fig. 7. Results of PHOCOS for different input choices. (a) Result based on increased time resolution. (b). Comparison of inferred crosstalk graphs for choosing representative features as the j^{th} -nearest feature to the cluster center of each phenotypic class.

5 Discussion

We propose PHOCOS as a framework for inferring multi-feature crosstalk networks from perturbation assays. Previous approaches have been developed to analyze gene regulatory networks from image readouts (Collinet et al., 2010; Graml et al., 2014; Nir et al., 2010). These approaches focused on single measurements on each biomarker. To the best of our knowledge, PHOCOS is the first crosstalk inference method that explicitly considers the potential interactions among different features (attributes) of biomarkers. The contributions of PHOCOS are summarized in the following three points.

First, from the view of data analysis, we conducted feature screening to identify common phenotypic classes from patterns of feature deviation profiles. In particular, for microscopy there are many possible features that could be extracted from images, and it is often not obvious which features offer novel insights into crosstalk. Our feature-screening step identified a small number of archetypical deviation patterns, which helped reduce the complexity of graph inference.

Second, from the view of graph inference, we proposed a dual-sparse optimization framework for direct-effect pursuit from experimental observations. The effectiveness of PHOCOS was demonstrated by its ability to identify core, direct-effect motifs from missing links, noise and indirect effects. From simulation studies, PHOCOS handled difficult cases (with noise and missing links) well for direct-effect graph learning. While the PHOCOS reduction method is motivated by the multi-feature network problem, it is based on a general framework that can be directly applied to other network inference problems, e.g. for gene or protein interaction networks.

Third, from the view of systems biology, we applied PHOCOS to study multiple phenotypic responses across multiple modules and time periods defined within the neutrophil polarization network. Consistent with previous work, we found that crosstalk can evolve dynamically and differently across different features. More importantly, we were able to identify novel cross-feature interactions that were missed by single-feature graph-learning approaches.

With the ability to infer multi-feature crosstalk influence networks comes both increased resolution and increased complexity. Depending on the features selected, some links may be immediately interpretable (e.g. activity levels of A affect localization of B), while others (e.g. activity levels of A affect the texture of B) may be difficult to interpret directly in terms of molecular mechanisms and interactions. As previously mentioned, new techniques, including optogenetics, will allow perturbation of specific features while holding others constant; these experimental approaches will both inspire the development of multi-feature crosstalk inference methods, such as PHOCOS, and be invaluable for testing predictions.

PHOCOS has the potential to help provide deeper insights into how biological networks transduce information. Content-rich approaches for monitoring cellular phenotypes, such as microscopy, provide multi-dimensional views into how molecular components respond to perturbations. PHOCOS can use these data to provide ‘high-resolution’ views of which phenotypic attributes support interaction among biomarkers. While PHOCOS was developed and tested on microscopy images, it can be extended to infer cross-feature interactions for other biological datasets that capture multi-attributes per biomarker.

Acknowledgments

The authors are grateful to all members of the Altschuler and Wu laboratory, particularly Jeremy Ku, Satwik Rajaram and Shan-Shan Liu.

Funding

This research was partially supported by the National Institute of Health grants R01GM112690 (to S.J.A.), R01CA184984 (to L.F.W.) and the Institute of Computational Health Sciences (ICHS) at UCSF (to S.J.A. and L.F.W.).

Conflict of Interest: none declared.

References

- Bader, J.S. *et al.* (2004) Gaining confidence in high-throughput protein interaction networks. *Nat. Biotechnol.*, **22**, 78–85.
- Bakal, C. *et al.* (2007) Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science*, **316**, 1753–1756.
- Baraniuk, R.G. (2007) Compressive sensing. *IEEE Signal Process. Mag.*, **24**, 12–13.
- Barzel, B. and Barabási, A.L. (2013) Network link prediction by global silencing of indirect correlations. *Nat. Biotechnol.*, **31**, 720–725.
- Boland, M.V. and Murphy, R.F. (2001) A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics*, **17**, 1213–1223.
- Boyd, S. *et al.* (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.*, **3**, 1–122.
- Burnham, K.P. and Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Berlin, Heidelberg: Springer Science & Business Media.
- Carpenter, A.E. *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.*, **7**, R100.
- Chen, S.S. *et al.* (1998) Atomic decomposition by basis pursuit. *SIAM J. Sci. Comput.*, **20**, 33–61.
- Collinet, C. *et al.* (2010) Systems survey of endocytosis by multiparametric image analysis. *Nature*, **464**, 243–249.
- Collins, S.R. *et al.* (2007) Functional dissection of protein complexes involved in yeast chromosome biology using a genetic interaction map. *Nature*, **446**, 806–810.
- Deng, Y. *et al.* (2013) Low-rank structure learning via nonconvex heuristic recovery. *IEEE Trans. Neural. Netw. Learn. Syst.*, **24**, 383–396.
- Feizi, S. *et al.* (2013) Network deconvolution as a general method to distinguish direct dependencies in networks. *Nat. Biotechnol.*, **31**, 726–733.
- Friedman, N. *et al.* (2000) Using Bayesian networks to analyze expression data. *J. Comput. Biol.*, **7**, 601–620.
- Graml, V. *et al.* (2014) A genomic multiprocess survey of machineries that control and link cell shape, microtubule organization, and cell-cycle progression. *Dev. Cell*, **31**, 227–239.
- Ku, C.J. *et al.* (2010) On identifying information from image-based spatial polarity phenotypes in neutrophils. In: *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Rotterdam, Netherlands, pp. 1029–1032.
- Ku, C.J. *et al.* (2012) Network crosstalk dynamically changes during neutrophil polarization. *Cell*, **149**, 1073–1083.
- Liberali, P. *et al.* (2014) A hierarchical map of regulatory genetic interactions in membrane trafficking. *Cell*, **157**, 1473–1487.
- Loo, L.H. *et al.* (2007) Image-based multivariate profiling of drug responses from single cells. *Nat. Methods*, **4**, 445–453.
- Malioutov, D. *et al.* (2005) A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE Trans. Signal Proc.*, **53**, 3010–3022.
- Marbach, D. *et al.* (2012) Wisdom of crowds for robust gene network inference. *Nat. Methods*, **9**, 796–804.
- Markowetz, F. *et al.* (2007) Nested effects models for high-dimensional phenotyping screens. *Bioinformatics*, **23**, i305–i312.
- Murphy, R.F. (2014) A new era in bioimage informatics. *Bioinformatics*, **30**, 1353.
- Nir, O. *et al.* (2010) Inference of RhoGAP/GTPase regulation using single-cell morphological data from a combinatorial RNAi screen. *Genome Res.*, **20**, 372–380.
- Pearlman, Z.E. *et al.* (2004) Multidimensional drug profiling by automated microscopy. *Science*, **306**, 1194–1198.
- Petersen, K.B. and Pedersen, M.S. (2008) The matrix cookbook. *Tech. Univ. Denmark*, **7**, 15.
- Seeger, R. and Krebs, E.G. (1995) The MAPK signaling cascade. *FASEB J.*, **9**, 726–735.
- Shmulevich, I. *et al.* (2002) Gene perturbation and intervention in probabilistic Boolean networks. *Bioinformatics*, **18**, 1319–1331.
- Snijder, B. *et al.* (2013) Predicting functional gene interactions with the hierarchical interaction score. *Nat. Methods*, **10**, 1089–1092.
- Spirites, P. *et al.* (2000) *Causation, Prediction, and Search*. Cambridge, Massachusetts, US: MIT Press.
- Toettcher, J.E. *et al.* (2013) Using optogenetics to interrogate the dynamic control of signal transmission by the Ras/Erk module. *Cell*, **155**, 1422–1434.
- Wang, Y. *et al.* (2013) Identifying network motifs that buffer front-to-back signaling in polarized neutrophils. *Cell Rep.*, **3**, 1607–1616.
- Weigt, M. *et al.* (2009) Identification of direct residue contacts in protein–protein interaction by message passing. *Proc. Natl. Acad. Sci.*, **106**, 67–72.
- Xu, J. *et al.* (2003) Divergent signals and cytoskeletal assemblies regulate self-organizing polarity in neutrophils. *Cell*, **114**, 201–214.
- Yang, J. *et al.* (2010) A fast alternating direction method for TVL1-L2 signal reconstruction from partial fourier data. *IEEE J. Sel. Top. Signal Process.*, **4**, 288–297.