Methods and Applications in Fluorescence

PAPER

Experimental validation of predicted cancer genes using FRET

Dimitri Guala¹, Kristoffer Bernhem², Hammou Ait Blal¹, Daniel Jans³, Emma Lundberg⁴, Hjalmar Brismar¹/²/³ and Erik I. L. Sonnhammer

¹ Science for Life Laboratory, Stockholm Bioinformatics Center, Department of Biochemistry and Biophysics, Stockholm University, Box 1031, 17121 Solna, Sweden
² Science for Life Laboratory, Department of Applied Physics, KTH—Royal Institute of Technology, Box 1031, 17121 Solna, Sweden
³ Science for Life Laboratory, School of Biotechnology, KTH—Royal Institute of Technology, Box 1031, 171 21 Solna, Sweden
⁴ Science for Life Laboratory, Department of Women’s and Children’s Health, Karolinska Institutet, Box 1031, 171 21 Solna, Sweden

E-mail: dimitri.guala@scilifelab.se

Keywords: fluorescent resonance energy transfer, experimental validation, protein-protein interactions, gene prioritization, cancer, MaxLink, functional association networks

Supplementary material for this article is available online

Abstract

Huge amounts of data are generated in genome wide experiments, designed to investigate diseases with complex genetic causes. Follow up of all potential leads produced by such experiments is currently cost prohibitive and time consuming. Gene prioritization tools alleviate these constraints by directing further experimental efforts towards the most promising candidate targets. Recently a gene prioritization tool called MaxLink was shown to outperform other widely used state-of-the-art prioritization tools in a large scale in silico benchmark. An experimental validation of predictions made by MaxLink has however been lacking. In this study we used Fluorescence Resonance Energy Transfer, an established experimental technique for detection of protein-protein interactions, to validate potential cancer genes predicted by MaxLink. Our results provide confidence in the use of MaxLink for selection of new targets in the battle with polygenic diseases.

Introduction

Many diseases that involve complex interplay between multiple genes and the environment, such as cancer, diabetes and Alzheimer’s disease, have huge implications for the public health. In order to combat these important issues, advances in high-throughput experimental methods, such as Genome Wide Association Studies, RNA inference screens, etc have lately been used to speed up the discovery of genes/proteins implicated in diseases with complex genetic causes. Hundreds (or thousands) of potential candidate genes involved in polygenic disorders have been uncovered by these experimental advances (Tranchevent et al 2011, Doncheva et al 2012). Follow up of all these leads with rigorous target validation is currently cost prohibitive and time consuming. Therefore prioritization of candidate disease genes is necessary and important in order to maximize the information from the high-throughput experiments, by focusing further experimental efforts on the most promising candidates. Current prioritization tools can usually provide a manageable list of genes, to focus on, and often rank the genes by their estimated importance for future investigations. Usually it’s not expected that a prioritization tool can provide a list that is 100% accurate but it’s important that the prioritized genes have a biologically meaningful relationship to the studied pathology (Piro and Di Cunto 2012).

MaxLink (Guala et al 2014) is a gene prioritization method that when provided with a set of genes known to be involved in a particular phenotype, e.g. cancer, produces a ranked list of genes/proteins tightly associated to the input set i.e. predicting novel genes/proteins potentially involved in the phenotype of interest. MaxLink works on a gene/protein association network such as FunCoup (Schmitt et al 2014) in order to make its predictions. MaxLink has recently been shown to outperform other state-of-the-art gene prioritization algorithms on a largest, to date, benchmark for gene prioritization tools (Guala and Sonnhammer 2017). Initial validations of MaxLink were performed using Gene Ontology (Ashburner et al 2000) and datasets of differential expression in
cancer versus normal tissue (Östlund et al 2010). MaxLink has also been tested on set of rare disease genes from Orphanet (Aymé 2003). However, a direct validation of predicted interactions is still lacking and would add additional credibility to this promising prioritization method.

There is a wealth of biochemical techniques to study protein-protein interactions (PPIs), but they usually require extraction of the studied proteins from their natural environment and are too laborious when many PPIs need to be tested. Fluorescence Resonance Energy Transfer (FRET) (Kenworthy 2001) is a widely used technique that allows the study of PPIs directly in fixed or live cells using fluorescence microscopy. FRET relies on transfer of energy between two fluorophores in close proximity (typically < 10 nm). The interaction partners of a PPI can be labeled with fluorophore conjugated antibodies that exhibit FRET when the labeled proteins are close enough for a possible interaction.

In this study we used FRET to experimentally validate a list of potential genes implicated in cancer, predicted by MaxLink. We attempted to test 18 predicted interactions between the known oncogene Janus kinase 2 (JAK2) and other proteins, and could confirm 15 of them experimentally. This substantially expands the list of verified JAK2 interactions that can be used to guide future functional experiments. Furthermore, the study demonstrates the high reliability of MaxLink predictions, and that a FRET based method is efficient for large-scale probing of protein-protein interactions.

Materials and methods

Functional association network

Network prioritization methods need a source of protein interactions in order to make their predictions. We used FunCoup 3.0 (Schmitt et al 2014) as the underlying protein interaction network, since the method for gene prioritization is seamlessly integrated into this network. FunCoup is one of the most comprehensive sources of gene/protein associations. It combines various different types of gene/protein interaction data e.g. Protein-Protein Interaction, Protein domain interaction, Protein co-expression, mRNA co-expression, Subcellular co-localization, co-miRNA regulation by shared miRNA targeting, shared transcription factor binding, phylogenetic profile similarity and genetic interaction. This data is collected from 11 species including Homo sapiens and 10 model organisms and transferred via orthologous associations in order to increase the amount of evidence and to make the inference of interactions more robust. FunCoup can be used with different confidence thresholds (0.1–1) marking the lowest allowed confidence of interaction between proteins. Confidence threshold of 0.75 was used in this study resulting in 1123 873 links between 12 391 genes.

Method for prediction of cancer genes

MaxLink (Guala et al 2014) was used for prediction of potential candidate genes. This gene prioritization method tries to exploit the notion of guilt-by-association (Altshuler et al 2000, Oliver 2000), i.e. the fact that proteins involved in the same or similar diseases tend to interact with each other (Zhu et al 2014). MaxLink uses the topology of an underlying protein association network e.g. FunCoup to identify genes/proteins that interact with a set of query genes/proteins involved in a phenotype of interest e.g. cancer. The mechanism by which MaxLink prioritizes candidate genes is depicted below (figure 1). MaxLink accepts a set of genes known to be involved in a phenotype as the query, and uses FunCoup to find all their direct neighbours. It then applies a connectivity filter corrected for multiple hypothesis testing, to select genes enriched in connections to the query. This step is done in order to filter out some of the hub genes that spuriously interact with the query due to their high number of connections. In the last step, MaxLink returns a list of potential candidates sorted by the number of connections to the query, called the MaxLink score (ML).

Validation pipeline

The overview of the validation pipeline is depicted below (figure 2). Its individual steps are described in detail in the following sections.
Selection of query genes
The query of genes known to be involved in cancer was compiled from Cancer Gene Sensus (CGC) (Forbes et al. 2011) and UniProt (The UniProt Consortium 2014). CGC (downloaded on 2012-MAR-15) genes and genes annotated with keywords related to cancer (Supp. table 1 is available online at stacks.iop.org/MAF/6/035007/mmedia) in the CC field of UniProt (downloaded on 2013-NOV-15) were combined (Östlund et al. 2010), and used as the query to MaxLink.

Prediction of cancer genes
We ran MaxLink with a connectivity filter cutoff of 0.5, meaning that in order to pass the filter a candidate gene needed to have more connections to the query genes than to the other genes in the network. FunCoup was used with confidence threshold of 0.75. One of the query cancer genes with the most predicted candidates, JAK2, was selected for experimental evaluation together with its predicted associations (Supp. table 2). The final selection of candidates for validation was further restricted to those that had antibodies in the Human Protein Atlas (HPA) (Uhlen et al. 2010) and were expressed in sufficient quantity ( Fragments per

Table 1. Tested FRET interactions. The list of predicted interactions intended for testing. The information supplied is the MaxLink score (ML), whether a direct interaction with JAK2 has been demonstrated before (Known) and the database or publication the evidence comes from: intAct (i), BioGrid (b), iRefIndex (w), or publication, respectively. The FPKM in the chosen cell line and the main HPA location of the protein are also presented.

| Symbol | ML | Known | Evidence | FPKM | Main HPA location |
|--------|----|-------|----------|------|------------------|
| BDP1   | 3  | no    |          | 6.96 | Nucleoplasm      |
| CASP8  | 45 | no    |          | 11   | cytoplasm        |
| DLG1   | 15 | no    |          | 30   | membrane (UniProt) |
| IKKβ   | 24 | no    |          | 3.96 | cytoplasm        |
| ITCH   | 11 | no    |          | 20   | nucleoplasm      |
| MAP3K1 | 11 | no    |          | 4.13 | cytoplasm        |
| NCO1   | 37 | no    |          | 21.2 | Nucleoplasm      |
| PLCG1  | 22 | indirectly | i | 26.2 | cytoplasm |
| PRKCD  | 35 | no    |          | 7.88 | Nucleoplasm, Golgi, Vesicles |
| PTN12  | 5  | indirectly | b | 46.5 | Plasma membrane, Cytoplasm |
| RFWD2  | 19 | no    |          | 10.6 | nucleoplasm      |
| SH2B2  | 6  | indirectly | i | 3.75 | cytoplasm        |
| STAT1  | 35 | yes | W        | 40.9 | cytoplasm        |
| STAT2  | 3  | yes | w        | 7.02 | Plasma membrane  |
| TNKS   | 4  | no    |          | 9.22 | nucleoplasm      |
| TRAF6  | 19 | indirectly | i | 5.25 | mitochondria     |
| WDR7   | 3  | no    |          | 3.41 | Nucleoplasm      |
| YES1   | 18 | indirectly | b | 61.2 | Plasma membrane, Cytoplasm |

Figure 2. Validation pipeline overview. The steps involved in the validation pipeline are 1. Compiling a set of query cancer genes. 2. Prediction of genes associated with the query using MaxLink. 3. The result of the prediction is a list of candidates with corresponding MaxLink score and p-value. 4. Experimental validation of predicted candidate gene interactions using FRET. 5. Analysis and display of results.
kilobase of exon per million reads mapped, FPKM > 3) in the chosen cell line (U-2 OS) (table 1).

Fluorescence resonance energy transfer (FRET)
FRET is a fluorescence based method that can be used to detect the close proximity (possible interaction) between two proteins. FRET is based on a nonradiative energy transfer mechanism between a donor and an acceptor fluorophore. The efficiency in the energy transfer is inversely proportional to the sixth power of the distance between the fluorophores and can typically occur if the distance is on the order of <10 nm (figure 3). The experimental setup includes labeling the potentially interacting pair of proteins with different fluorophores and measurement of fluorescence in wavelength bands corresponding to donor and acceptor fluorescence. FRET efficiency can then be calculated from the measured shift in emission from the donor to the acceptor wavelengths.

Experimental setup
In the experimental setup JAK2 was labeled with a commercially available primary antibody (mouse) and a secondary anti-mouse antibody carrying the donor fluorophore. Each candidate protein was labeled with an appropriate primary antibody (rabbit) from the HPA and a secondary anti-rabbit antibody carrying the acceptor fluorophore (figure 4(a)). Each interaction was studied in two wells of a 96-well plate by means of fluorescence imaging. Images were acquired in three different fluorescence channels e.g. donor, acceptor and FRET in 16 symmetrically allocated fields of view. The signal means of the FRET channel were compared to the signal means of the corresponding donor channel in order to establish presence of FRET.

In order to establish the maximum value of FRET in our experimental setup we used a positive control construct with one primary antibody for the target protein (e.g. mouse monoclonal anti-JAK2) and two different secondary antibodies (e.g. goat monoclonal anti-mouse and rabbit monoclonal anti-mouse) fused with donor and acceptor fluorophores respectively (figure 4(b)). This guaranteed that the donor and acceptor fluorophores were in sufficiently close vicinity of each other in order to produce a FRET signal detectable by fluorescence microscopy.

An absolute biological negative control can be difficult to establish in an experimental setup. If two proteins are expressed in different compartments of the cell (e.g. mitochondria and nucleoplasm), then there should be no measurable FRET. However, if the two proteins are in the same compartment, it is not possible to rule out that they could be physically or functionally interacting. We therefore used both an experimental and a theoretical negative controls. As an experimental negative control we used the mitochondrial inner membrane protein Mic60 as donor and the plasma membrane protein Na, K-ATPase α1 (NKA) as the acceptor. The theoretical negative control was based on calculation of random FRET in a volume with varying concentration of donor and acceptor molecules. To this end we use the equation for FRET efficiency according to Förster (Förster et al 1939, Schaufele et al 2005).

$$E = \frac{1}{1 + \left(\frac{R_{0}}{R_s}\right)^6}$$

Where $E$ is the FRET efficiency between the donor and acceptor fluorophores at distance $R_{DA}$, and $R_s$ is the Förster radius. In our study we use the dyes Alexa Fluor 488 and Alexa Fluor 555 as donor and acceptor with $R_0 = 7$ nm (Johnson and Spence 2011). Simulations were performed with donors and acceptors randomly distributed in a volume with typical protein concentrations ranging 10 nM – 1 μM. The pairwise distances $R_{DA}$ and the mean FRET efficiency $E$ were calculated for all donors and acceptors. The results of the simulation were used as a negative control together with the experimental measurements (figure 7).

For full details of the experimental setup including plate design, antibody details and concentrations please see below.

Cell cultivation
The U-S OS human osteosarcoma cell line (ATCC-LGC) used throughout this study was grown in 37 °C (humidified air, 5% CO2) in McCoy’s 5 A medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum (Sigma-Aldrich). Cells were seeded on to fibronectin coated 96-well glass bottomed plates (Greiner Sensoplate Plus, Cat# 655892, Greiner Bio-One, Germany) at a concentration of 15 000 cells/well and cultivated for 24 h to a confluency of 60%–70% before immunostaining was carried out.

Immunostaining
PBS-washed cells were fixed in 4% paraformaldehyde (PFA) in growth media supplemented with 10% FBS for 15 min, followed by permeabilization with 0.1% Triton X-100 in PBS for 3 × 5 min. After a washing step with PBS, cells were incubated with the primary antibodies overnight at 4 °C. The antibody targeting
JAK2 (ab37226, Abcam) was diluted to 1 μg/ml and rabbit polyclonal HPA antibodies targeting the proposed interaction partners were diluted to 2–4 μg/ml in blocking buffer (PBS with 4% fetal bovine serum). On the next day after 4 × 10 min washes with PBS, the cells were incubated for 90 min at room temperature with the following secondary antibodies (all from ThermoFisher Scientific) diluted to 1 μg/ml in blocking buffer: goat anti-mouse AlexaFluor 488 (ThermoFisher A11001/A11029), goat anti-rabbit AlexaFluor 555 (thermoFisher A21428/A21429). Cells were subsequently counterstained with DAPI for 10 min. After washing with PBS, the wells were completely filled with PBS containing 78% glycerol and sealed. The entire immunostaining procedure was carried out using Tecan Freedom EVO pipetting robotic systems.

Microscopy
Fluorescent images were acquired with a Leica SP5 confocal microscope (DM6000CS) equipped with a 40×/0.85NA objective (Leica Microsystems, Mannheim, Germany). The image acquisition was performed automatically using the MatrixScreener M3 in LAS AF software (Leica Microsystems, Mannheim, Germany). FRET experiments were performed by excitation at 488 nm (donor excitation) and 543 nm (acceptor excitation) with detection in wavelength bands 495–540 nm for the donor channel and 550–625 nm for the acceptor (FRET) channel. The settings for each image were as followed: Pinhole 1 Airy unit, 600 Hz scan speed, 16 bit acquisition, and an image size of 2048 × 2048 pixels with a pixel size of 0.1 μm. In order to reduce potential artifacts from Poisson and shot noise (Nagy et al 2014) the images were down-sampled to 1024 × 1024 pixels by a 2 × 2 averaging. In order to reduce the effect of variability in experimental factors, such as cell density, protein expression level, and staining efficiency, we captured 16 symmetrically allocated fields of view from each well and used triplicates for each studied interaction.

Analysis of FRET signal
In analysis of FRET, compensation for spectral cross-talk between donor and acceptor channels need to be performed. To achieve this, each experimental group consisted of a triplicate, labeled either with donor, acceptor or both fluorophores.

Compensation for donor bleedthrough ($B = 0.267$) was established by illuminating the donor-only sample with the donor excitation wavelength and measuring the background corrected mean intensities within a cell in both donor ($I_{DD}$) and acceptor ($I_{DA}$) channels.

$$B = \frac{I_{DA}}{I_{DD}}$$

Compensation for direct excitation of the acceptor ($D = 0.424$) was established using the acceptor-only sample by measuring the background corrected mean intensities within a cell in the acceptor channel using donor excitation wavelength ($I_{DA}$) and acceptor excitation wavelength ($I_{AA}$).

$$D = \frac{I_{DA}}{I_{AA}}$$

The samples with both donor and acceptor fluorophores were used to measure FRET. The background corrected intensity in donor ($I_D$) and acceptor channels ($I_A$) were recorded using donor excitation wavelength followed by recording the intensity in the acceptor channel ($I_A$) using acceptor excitation wavelength. The compensation factors $B$ and $D$ were used to calculate on a pixel by pixel basis the corrected FRET intensity ($FRET_{corr}$) and the FRET ratio ($FRET_{ratio}$) which was reported as a mean value for each analyzed cell and experimental group.
Results

The query submitted to MaxLink consisted of 787 unique genes (Supp. table 3) and yielded 706 candidates (Supp. table 2). We selected one of the cancer genes that had the most predicted candidates, JAK2, which is a non-receptor tyrosine kinase implicated in eosinophilic, lymphoblastic and myeloid leukemia. Experimental evaluation was performed together with 18 (figure 5, table 1) of its predicted interaction partners that remained after the initial 40 were restricted by availability of antibodies in the Human Protein Atlas (HPA) and sufficient intracellular expression of the protein in question.

Three of the studied interactions (Jak2 with PTPN12, TNKS and CASP8, respectively) were not possible to be tested due to failure in binding of the primary antibodies to their respective targets, as evident by the lack of staining for donor and/or acceptor antibodies. The remaining 16 interactions (the positive control and 15 predicted interactions) showed various levels of signal in the three studied channels as depicted below (figure 6).

Pixelwise analysis of the FRET signal in the positive control was able to clearly distinguish the background FRET intensities from the true signal (figure 7).

When FRET intensities were corrected for the intensities in the donor channel and presented as FRET\(_{\text{ratio}}\), i.e. the fraction of the intensity in the donor channel and the FRET channel, the ratio for the positive control marked the highest possible mean FRET\(_{\text{ratio}}\) of 0.619 (figure 7, Supp. table 4). The distribution of FRET\(_{\text{ratio}}\) for the positive control had a relatively narrow spread (std ≈ 0.015) (figure 8, Supp. table 4).

The results for predicted interactors exhibited mean FRET\(_{\text{ratio}}\) values between 0.055 and 0.086 (figure 8, Supp. table 4). The distributions for the predicted interactors were even tighter than that for the positive control and ranged between std:s of 0.006 and 0.014. The level of FRET\(_{\text{ratio}}\) thus seen for the predicted interactions corresponded to between 8% and 14% of maximum theoretically possible FRET\(_{\text{ratio}}\). Mean FRET\(_{\text{ratio}}\) values were also much higher than both the experimental and the simulated lower bounds, which had a mean FRET\(_{\text{ratio}}\) of 0.02 and 0.002, respectively, amounting to 3% and 0.3% of the theoretical maximum, respectively (figure 8, Supp. table 4).

Some of the interactions predicted by MaxLink have already seen support in the literature in terms of being already known associations in the FunCoup.

\[
\text{FRET}_{\text{corr}} = I_A - B \cdot I_D - D \cdot I_A
\]

\[
\text{FRET}_{\text{ratio}} = \frac{\text{FRET}_{\text{corr}}}{I_D}
\]
Figure 6. Fluorescence microscopy imaging. FRET images of JAK2 positive control experiment in Donor (green), Acceptor (red) and FRET (yellow) channels.

Figure 7. FRET image analysis. Calculated FRET ratio image of JAK2 positive control displayed using the ‘fire’ lookup table for FRET ratio from 0.0 to 1.3, the corresponding pixel histogram over FRET ratio. The peak in the histogram at 0.15 originates from the background in the FRET ratio image.

Figure 8. FRET analysis. Distributions of the intensity of FRET corrected for the intensity in the Donor channel, the FRET ratio for each of the successfully tested interactions including the positive control (+)CTRL, the biological negative control (−)CTRL and the simulated negative control (−)CTRL. Illustrated as boxplots, showing median and 1st, and 3rd quartiles of the data, and data points outside are shown with whiskers extending maximally to 1.5 times the inter-quartile range.
network e.g. JAK2-STAT1 or constituting part of other databases e.g. JAK2-PLCG1 in IntAct (Aranda et al 2010) (table 1). We included these interactions in our validation, both for completeness in terms of MaxLink’s predictions and also to see whether we could reproduce the previous findings using our experimental method of choice, i.e. FRET. However, we were particularly interested in interactions predicted by MaxLink which did not have any previous direct PPI evidence in the literature nor in any other of the major protein interaction databases, including FunCoup, UniProt, IntAct, BioGrid (Chatr-Aryamontri et al 2015) and iRefIndex (Razick et al 2008). One such example is the direct interaction of JAK2 and RFWD2 (table 1). RFWD2 is an E3 ubiquitin-protein ligase whose main function involves tagging of proteins by ubiquitin, which leads to subsequent degradation of the tagged protein in the proteasome (Uljon et al 2016). One of RFWD2’s targets is the prominent protein p53 involved in many different types of cancer (Muller and Vousden 2013). Among other RFWD2’s targets is the COP1 (Uljon et al 2016), which plays an important role in Chronic Lymphocytic Leukemia (CLL) (Fu et al 2015). There is evidence of JAK2 involvement in CLL (Todisco et al 2016). JAK2 localizes both to cytoplasm and the nucleoplasm as well as some other cell compartments (table 1) according to HPA, UniProt, GO and in our FRET experiments (figures 6 and 9). The primary location of RFWD2 in the cell is in the nucleoplasm according to HPA, which is also where we see most of the FRET signal stemming from (figure 9).

Another predicted interaction without previous direct PPI evidence is JAK2-PRKCD (table 1). Both proteins are kinases important for intra-cellular signaling (Aaronson 2002, Paul et al 2015) and most importantly they have both been shown to have key roles in the pathogenesis of breast cancer, and as potential treatment targets (Utrteger et al 2012, Paul et al 2015, Wang et al 2015); PRKCD is overexpressed in malignant mammary tissue where it acts by promoting cell growth and as a poor survival factor as well as enhancing resistance to apoptosis. JAK2 activation enhances proliferation and survival of aberrant mammary cells through its downstream targets (Rodriguez-Barrueco et al 2015) and it has been shown to increase breast cancer survival when targeted by microRNA (Wang et al 2015). Of the tested interactions, this pair exhibited one of the strongest FRET signals. Both JAK2 and PRKCD have nucleoplasm and cytoplasm as their primary locations according to both HPA and UniProt.

Discussion

Gene prioritization tools used for leveraging the wealth of information stemming from high-throughput attempts to map polygenic disorders require experimental validation besides in silico assessment of their performance. The gene prioritization tool MaxLink was recently shown to outperform other state-of-the-art prioritization tools. In this study, we have demonstrated that interactions predicted by MaxLink, could be validated using FRET, a sensitive technique for detection of protein interactions on the nanometer scale. In our experimental setup we were able to confirm some of the previously identified interaction partners of JAK2. We were also able to experimentally validate a number of new interactions predicted by MaxLink. These novel interactions may have implications for understanding Chronic Lymphocytic Leukemia and breast cancer pathogenesis as well as support their investigation as potential therapeutic targets.

The level of FRET measured in the tested interactions is subject to several factors, such as variation in the level of expression of both JAK2 and the interaction protein in the used cell line, the binding efficiency of secondary antibodies (labeled with donor and acceptor fluorophores, respectively) to the primary antibodies, and the steric conformation of antibody binding. The expression level of JAK2 (the donor) can be considered to be constant in the experiments. However, the expression level of the acceptor proteins varies by an order of magnitude for the different
predicted proteins (table 1, Supp. table 2). A low expression of the acceptor protein results in a reduced FRET efficiency. The steric conformation, i.e. how the antibodies bind and organize spatially in relation to the proteins, has a direct effect on the FRET efficiency, since FRET is inversely proportional to the sixth power of distance between donor and acceptor fluorophore. In the most unfavorable situation the antibodies point outwards from the interaction and then reduce the FRET efficiency significantly.

Those factors reduce the efficiency of FRET, i.e. the FRET$_\text{ratio}$ for the predicted interactions. In our positive control, most of these factors are not applicable and the FRET efficiency is higher than for the validations. Experimental conditions, such as unspecific background fluorescence and imperfections in the microscope and analysis setup can contribute to a baseline of background FRET noise. We estimated this both by simulations and experimentally and used it as a negative control. The experimentally measured FRET efficiency for predicted interactions were found to be at least one order of magnitude higher than both the experimental and simulated negative controls. Thus, despite all factors that reduce FRET efficiency, we were able to verify the predicted interactions with high confidence.

In conclusion, we have experimentally verified predictions of PPIs made with a state-of-the-art gene prioritization method, MaxLink. In the process, we have validated novel interactions with potential implications for cancerogenesis and drug targeting. This should provide confidence to users of gene prioritization methods in general and in particular for users of MaxLink in the identification and selection of new potential targets of relevance for e.g. polygenic diseases.

Acknowledgments

We want to acknowledge the Cell Profiling facility at Science for Life Laboratory for support with sample preparation and image acquisition, and the National Microscopy Infrastructure, NMI (VR-RFI 2016-00968) for support in analysis. Additional support was received from the Swedish Research Council (VR 2015-05342).

The authors have declared that no conflicting interests exist.

ORCID iDs

Dimitri Guala  @ https://orcid.org/0000-0003-2245-7557
Hjalmar Brismar  @ https://orcid.org/0000-0003-0578-4003

References

Aaronson D S 2002 A road map for those who don’t know JAK-STAT Science 296 1653–5
Altschuler D, Dalé M and Kruglyak L 2000 Guilt by association Nat. Genet. 26 135–7
Aranda B, Achuthan P, Alam-Faruque Y, Aranzen I, Bridge A, Derow C and Hermjakob H 2010 The IntAct molecular interaction database in 2010 Nucleic Acids Research 38 (Database issue) D525–31
Ashburner M, Ball C A, Blake J A, Botstein D, Butler H, Cherry JM and Sherlock G 2000 Gene ontology: tool for the unification of biology. The gene ontology consortium Nat. Genet. 25 25–9
Aynel S 2003 [Orphanet, an information site on rare diseases] Seins 672 46–7 PMID:12655825
Chatr-Aryamontri A, Breitkreutz B-J, Oughtred R, Boucher I, Heinicke S, Chen D and Tyers M 2015 The BioGRID interaction database: 2015 update Nucleic Acids Research 43 (Database issue) D470–8
Doncheva N T, Kacpourvi T and Albrecht M 2012 Recent approaches to the prioritization of candidate disease genes Wiley Interdisciplinary Reviews. Systems Biology and Medicine 4 429–42
Forbes S A, Bindal N, Bamford S, Cole C, Kok C Y, Beare D and Futreal P A 2011 COSMIC: mining complete cancer genomes in the catalogue of somatic mutations in cancer Nucleic Acids Research 39 (Database issue) D945–50
Fu C, Gong Y, Shi X, Shi H, Wan Y, Wu Q and Xu K 2015 Expression and regulation of COP1 in chronic lymphocytic leukemia cells for promotion of cell proliferation and tumorigenicity Oncology Reports 35 1493–500
Forster T, Energiewanderung Z and Van F 1939 Zwischenmolekule Energiewanderung und Fluoreszenz Ann. Phys., Lpz. 248 55–75
Guala D, Sjölund E and Sonnhammer E L L 2014 MaxLink: Network-based prioritization of genes tightly linked to a disease seed set Bioinformatics 30 2689–90
Guala D and Sonnhammer E L L 2017 A large-scale benchmark of gene prioritization methods Sci. Rep. 7 46598
Johnson L and Spence M T Z (ed) 2011 Fluorophores and their amine-reactive derivatives—section 1.3 The Molecular Probes Handbook. A guide to Fluorescent Probes and Labeling Technologies 11th edn (Carlsbad: Invitrogen Corp.) p 39 table 1.6
Kenworthy A K 2001 Imaging protein–protein interactions using fluorescence resonance energy transfer microscopy Methods 24 289–96
Muller P A I and Vosuden K H 2013 p53 mutations in cancer Nat. Cell Biol. 15 2–8
Nagy P, Szabó A, Váradi T, Kovács T, Batta G and Szilási L 2014 Maximum likelihood estimation of FRET efficiency and its implications for distortions in pixelwise calculation of FRET in microscopy Cytoometry Part A 85 942–53
Oliver S 2000 Guilt-by-association goes global Nature 403 601–3
Ostlund G, Lindskog M and Sonnhammer E L L 2010 Network-based Identification of novel cancer genes Molecular & Cellular Proteomics 9 648–55
Paul A, Darley M, Saha B, Tawfik O and Paul S 2015 PKCζ Promotes breast cancer invasion by regulating expression of E-cadherin and Zonula Occludens-1 (ZO-1) via NFκB-Bp65 Sci. Rep. 5 12520
Piro R M and Di Cunto F 2012 Computational approaches to disease-gene prediction: rationale, classification and successes The FEBS Journal 279 678–96
Razick S, Magklaras G and Donaldson I M 2008 iRefIndex: a consolidated protein interaction database with provenance BMC Bioinform. 9 405
Rodriguez-Barrueco R, Yu J, Saucedo-Cuevas L P, Olivier M, Llobet-Navas D, Putcha P and Silva J M 2015 Inhibition of the autocrine IL-6–JAK2–STAT3–calprotectin axis as targeted therapy for HR+HER2+ breast cancers Genes and Development 29 1631–48
Schaufele F, Demarco I and Day R N 2005 FRET Imaging in the wide-field microscope Molecular Imaging (San Diego: Elsevier) pp 72–94

Schmitt T, Ogris C and Sonnhammer E L L 2014 FunCoup 3.0: database of genome-wide functional coupling networks Nucleic Acids Research 42 D380–8

The UniProt Consortium 2014 Activities at the universal protein resource (UniProt) Nucleic Acids Research 42 (Database issue) D191–8

Todisco G, Manshouri T, Verstovsek S, Masarova L, Pierce S A, Keating M J and Estrov Z 2016 Chronic lymphocytic leukemia and myeloproliferative neoplasms concurrently diagnosed: clinical and biological characteristics Leukemia & Lymphoma 57 1054–9

Tranchevent L-C, Capdevila F B, Nitsch D, De Moor B, De Causmaecker P and Moreau Y 2011 A guide to web tools to prioritize candidate genes Briefings in Bioinformatics 12 22–32

Uhlen M, Oksvold P, Fagerberg L, Lundberg E, Jonsson K, Forsberg M and Ponten F 2010 Towards a knowledge-based human protein atlas Nat. Biotechnol. 28 1248–50

Uljon S, Xu X, Durzynska I, Stein S, Adelmant G, Marto J A and Blacklow S C 2016 Structural basis for substrate selectivity of the E3 Ligase COP1 Structure 24 687–96

Utrreger A J, Kazanietz M G and Bal De Kier Joffé E D 2012 Contribution of individual PKC isoforms to breast cancer progression IUBMB Life. 64 (1) 18–26

Wang X, Qiu W, Zhang G, Xu S, Gao Q and Yang Z 2015 MicroRNA-204 targets JAK2 in breast cancer and induces cell apoptosis through the STAT3/BCI-2/survivin pathway International Journal of Clinical and Experimental Pathology 8 5017–25 PMID:26191195

Zhu C, Wu C, Aronow B J and Jegga A G 2014 Computational approaches for human disease gene prediction and ranking Advances in Experimental Medicine and Biology 799 69–84