



Published in final edited form as:

Nat Protoc. 2014 January ; 9(1): 90–103. doi:10.1038/nprot.2013.167.

## Protein interaction mapping with ribosome-displayed using PLATO ORF libraries

Jian Zhu<sup>1,4,9,10</sup>, H. Benjamin Larman<sup>1,2,3,4,9,10</sup>, Geng Gao<sup>1,4,9</sup>, Romel Somwar<sup>7</sup>, Zijuan Zhang<sup>8</sup>, Uri Laserson<sup>2,4,5</sup>, Alberto Ciccia<sup>1,4,9</sup>, Natalya Pavlova<sup>1,4,9</sup>, George Church<sup>2,4</sup>, Wei Zhang<sup>8</sup>, Santosh Kesari<sup>6</sup>, and Stephen J. Elledge<sup>1,4,9,\*</sup>

<sup>1</sup>Division of Genetics, Department of Medicine, Brigham and Women's Hospital, Boston, MA

<sup>2</sup>Harvard-MIT Division of Health Sciences and Technology, Cambridge, MA

<sup>3</sup>Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA

<sup>4</sup>Department of Genetics, Harvard University Medical School, Boston, MA

<sup>5</sup>Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA

<sup>6</sup>Division of Neuro-Oncology, Translational Neuro-Oncology Laboratories, Department of Neurosciences, U.C. San Diego, Moores Cancer Center, La Jolla, CA

<sup>7</sup>Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, New York, NY

<sup>8</sup>Department of Chemistry, University of Massachusetts Boston, Boston, MA

<sup>9</sup>Howard Hughes Medical Institute

### Abstract

Identifying physical interactions between proteins and other molecules is a critical aspect of biological analysis. Here we describe PLATO, an *in vitro* method for mapping such interactions by affinity enrichment of a library of full-length open reading frames displayed on ribosomes, followed by massively parallel analysis using DNA sequencing. We demonstrate the broad utility of the method by identifying known and new interacting partners of LYN kinase, patient autoantibodies and the small molecules gefitinib and dasatinib.

---

Users may view, print, copy, download and text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: [http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*Correspondence to: [selledge@genetics.med.harvard.edu](mailto:selledge@genetics.med.harvard.edu).

<sup>10</sup>These authors contributed equally to this work.

### Contributions

S.J.E. and H.B.L. conceived and supervised the project. pRD human ORFeome library was constructed by J.Z., and characterized by J.Z. and H.B.L. The PLATO protocol was developed by H.B.L. and J.Z. Clinical evaluations and patient sample acquisitions were performed by S.K. Statistical analysis was performed by U.L. under the supervision of G.M.C. R.S. provided gefitinib-conjugated beads. PLATO candidates were confirmed by J.Z. and G.G. A.C. provided support for the validation of LYN binding candidates. N.P. provided support for the validation of PND autoantigen candidates. Z.Z. and W.Z. provided biotin-dasatinib. The manuscript was prepared by H.B.L. and J.Z., and edited by S.J.E.

Several methods have been developed to characterize the specificities of protein-binding molecules. Display technologies are typically limited to shorter polypeptides and cDNA-based libraries suffer from highly non uniform clonal abundance distributions and incorrect reading frames.<sup>1</sup> Two-hybrid and split-reporter techniques,<sup>2</sup> are limited to analyses of bait molecules that can be presented within the cell, and are not suitable for drug or antibody target identification. More recently, protein microarrays have been used for these purposes,<sup>3</sup> but their construction typically requires individual proteins to be purified and arrayed, resulting in substantial costs and various degrees of protein denaturation.

To address these limitations, we developed PLATO (ParalleL Analysis of Translated ORFs), a method that combines *in vitro* display of full-length proteins with analysis by high-throughput DNA sequencing. We demonstrate the utility of PLATO by performing diverse interaction screens against the human ORFeome, a normalized collection of 15,483 cDNAs in the Gateway cloning system.<sup>4</sup>

To express an ORF library *in vitro*, PLATO employs ribosome display, a technique used to prepare a library of mRNA molecules that remain tethered to the proteins they encode by lacking a stop codon.<sup>5</sup> Ribosome display imposes minimal constraints on the length or composition of proteins that can be efficiently displayed.

We constructed a Gateway cloning-compatible ribosome display 'destination' vector (pRD-DEST; Supplementary Fig. 1), to be used as a recipient for a normalized pool of ORF 'entry' clones. After recombination, DNA is amplified by PCR, yielding linear templates lacking stop codons. Following *in vitro* transcription and translation, the ribosome-displayed ORFeome can be screened for binding to immobilized bait(s). Enrichment of candidate binding proteins can be rapidly assessed using quantitative real-time PCR (qPCR) with ORF-specific primers, or *en masse* by deep sequencing of the enriched mRNAs (Fig. 1a). Sequencing libraries can additionally be highly multiplexed, thereby reducing the cost of each screen. All steps required for PLATO are compatible with automation using standard liquid handling robotics.

Our strategy for deep sequencing of enriched display libraries employs recovery of the ORF 3' termini, which minimizes interference from RNA degradation and ensures stoichiometric correlation between tag counts and transcript abundance. To this end, we adopted the following protocol: (i) chemically fragment enriched mRNAs; (ii) reverse transcribe fragments using a common primer; (iii) polyadenylate cDNAs; (iv) add sample barcodes and sequencing adapters using two-stage PCR amplification (Fig. 1b). Subsequent multiplex deep sequencing analysis of pooled display libraries is reproducible and quantitative (Supplementary Fig. 2). Sequencing a sample of unenriched human pRD-ORFeome mRNA (input) detected the transcripts of 14,582 unique ORFs out of 15,483 total cDNAs in the entry clone library (94%, Fig. 1c).

To test the ability of PLATO to identify protein-protein interactions, we used LYN, which contains common structural components of the SRC family, including SH3, SH2 and kinase domains,<sup>6</sup> and has been extensively characterized for its interaction partners. After affinity enrichment of the human ORFeome using GST-LYN, GST alone or an unrelated GST-fused

protein (GST-Muted), we used Illumina sequencing to identify proteins specifically bound by GST-LYN (Fig. 2a, Supplementary Table 1, Supplementary Fig. 3a). A number of established LYN binding partners were among those identified, and we validated two by qPCR (Fig. 2b).<sup>7, 8</sup> We ranked candidate LYN interactors by their degree of enrichment on GST-LYN, and confirmed five of seven tested by western blot analysis (Fig. 2c). Of the two candidates not validated, one bound nonspecifically to GST, whereas the other was a true negative. Among the highly enriched ORFs, SH2 domain-containing proteins were overrepresented ( $P < 0.01$ , Fisher's test). Consistent with a role for LYN autophosphorylation in mediating these interactions, phosphatase treatment of immobilized GST-LYN abolished binding of SH2D1A and SH2D4A, but only partly diminished PIK3R3 binding, suggesting the presence of an additional interaction domain (Supplementary Fig. 3b). These proteins have not previously been reported to interact with LYN.

We next asked whether PLATO could be used to identify protein targets of antibodies from patients with autoimmune disease. We first examined target enrichment using to affinity purified P53 and PDCD4 antibodies immobilized on protein A/G beads for library immunoprecipitation. By qPCR, P53 and PDCD4 transcripts were robustly enriched by their cognate antibodies, but not by control antibodies (Supplementary Fig. 4).

In previous work, we synthesized an oligonucleotide library encoding a 36-residue overlapping human peptidome for display on bacteriophage T7 (T7-Pep). Deep sequencing of affinity-enriched T7-Pep using autoimmune cerebrospinal fluid (CSF) from three individuals with paraneoplastic neurological disorder (PND) uncovered known and novel autoantigens.<sup>9</sup> We screened these samples using PLATO. Unlike T7-Pep, the human ORFeome is an incomplete collection of full-length proteins, and our findings reflect the inherent complementarity of these libraries. For example, neuro-oncological ventral antigen 1 (NOVA1) is absent from the human ORFeome v5.1, and so PLATO was unable to detect this known autoreactivity in patient A, whereas it was robustly identified with T7-Pep. Conversely, PLATO identified numerous autoantigens for each patient that were missed in our peptidome screens (Supplementary Table 2). For example, PLATO analysis of patients A and B revealed immunoreactivity with known cancer autoantigens not detected with T7-Pep. Several of these reactive antigens were selected for confirmation via immunoprecipitation and western blotting (Fig. 2d, Supplementary Fig. 5a–d). In addition, we had previously established that antibodies from patient C recognized the tripartite motif containing proteins TRIM9 and TRIM67. PLATO considerably expanded the members of the TRIM family recognized by antibodies in this patient's CSF to include TRIM1/MID2, TRIM18/MID1, TRIM54 and TRIM55 (Fig. 2e). Notably, multiple sequence alignment results in tight clustering of this precise subset of the extended TRIM family, suggesting the presence of shared, conformational epitopes not represented in T7-Pep.<sup>10</sup> As an alternative PLATO readout, hybridization of autoantibody-enriched libraries to custom oligonucleotide microarrays revealed a similar list of autoantigens (Supplementary Fig. 6).

Discovering the targets of small molecules typically involves the use of cell extracts containing a wide distribution of protein abundances, which limits the accuracy of detection by mass spectrometry. Normalized ORF libraries and quantitative DNA sequencing might therefore offer greater power to detect protein-small molecule interactions. We tested this

idea with gefitinib, an inhibitor of epidermal growth factor receptor's (EGFR) tyrosine kinase domain. Gefitinib interacts with the ATP-binding pocket of EGFR and additional tyrosine kinases.<sup>11</sup> Analysis after ORFeome affinity enrichment on gefitinib-coupled beads revealed significant enrichment of 10 out of the 17 predicted targets tested (Fig. 2f). This experiment demonstrates the relative ease by which candidate protein interactions can be assayed with PLATO; the binding of any ORF can be rapidly assessed using qPCR without the need for cloning or western blotting. ORFeome libraries affinity enriched by the Src family tyrosine kinase inhibitor dasatinib exhibited overrepresentation of protein kinases ( $P = 0.0003$ ; Fisher's test), including the known target LCK and several targets not previously associated with this compound (Supplementary Table 3).

PLATO's limitations include incomplete ORFeome collections and a lack of protein post-translational modifications. However, the quality, completeness and availability of these libraries will continue to improve over time. In addition, very large ORF proteins may be displayed less efficiently and proteins containing membrane-spanning or aggregation-prone domains that normally require host cellular machinery for proper folding may aggregate; these factors may complicate data analysis. Finally, ribosome display imposes certain limitations on the conditions under which affinity enrichments can be performed (e.g. low temperature and absence of RNase contamination are essential), and using proteins containing nucleic acid-binding domains as baits may result in non-specific binding. When the required conditions for PLATO are met, however, this method provides three main advantages as a tool for proteomic investigations. First, it has minimal protein size and composition bias. Second, it has low cost and minimal instrument requirements. Finally, the rapidly declining cost of DNA sequencing will make PLATO an ideal platform for projects involving large numbers of samples, such as cohort-scale autoantibody profiling or structure-activity relationship analyses of small-molecule compounds.

## Online Methods

### Plasmids and antibodies

A ribosome display (RD) backbone vector<sup>5</sup> was modified by inserting a Gateway cassette (attR1-ccdB-CM<sup>R</sup>-attR2) to create the pRD-DEST destination vector according to the manufacturer's instructions (Invitrogen). Human ORFeome library v5.1 entry clones were pooled into eighteen super-pools (generally about ten 384-well entry plates per pool, based on plate serial number). For each super-pool, one LR reaction was performed to recombine the ORFs into the pRD-DEST vector. pDEST40 vector (Invitrogen) was used for transient expression of ORFs in 293T cells. pDEST15 vector (Invitrogen) was used for production of N-terminally GST-fused LYN (GST-LYN), MUTED (GST-MUTED) or control peptide (GST-Pep, DYKDDDDK) in BL21 E. coli. The antibodies used in this study include: rabbit IgG polyclonal antibody (Cat No. 2729, Cell Signaling), anti-p53 polyclonal antibody (Cat No. 9282, Cell Signaling), anti-PDCD4 polyclonal antibody (Cat No. A301-106A, Bethyl Lab), anti-RBM15 polyclonal antibody (Cat No. A300-821A, Bethyl Lab), and anti-V5 monoclonal antibody (Cat No. R960-25, Invitrogen).

## Patient samples

PND cerebrospinal fluid samples were described previously.<sup>9</sup>

## Ribosome display of human ORFeome v5.1

T7B and TolAk (0.2  $\mu$ M) primers were used to PCR amplify the pRD-ORFeome template (50 ng) by PrimeSTAR<sup>®</sup> HS PCR Kit (Takara). The following thermal cycles were used: 98°C, 2 min/98°C, 10 sec; 55°C, 10 sec; 72°C, 8 min; 10 cycles. The PCR product was purified using QIAquick<sup>®</sup> PCR Purification Kit (Qiagen). The purified PCR products of all eighteen super-pools were combined together in equal amounts (by weight). *In vitro* transcription was performed using T7 Ribomax Large *in vitro* Transcription Kit according to the manufacturer (Promega). RNA products were purified using RNeasy<sup>®</sup> column (Qiagen). *In vitro* translation was performed using RTS 100 E. coli HY Kit (5Prime). 7.5  $\mu$ g mRNA in a 50  $\mu$ l reaction containing 1  $\mu$ l RNaseOUT (Invitrogen) was subjected to *in vitro* translation performed on PCR machine at 30 °C for 15 min. 12.5  $\mu$ l aliquots of each translation reaction was diluted with 85.5  $\mu$ l ice cold RD selection buffer (RD wash buffer (50 mM Tris Acetate, 150 mM NaCl, pH to 7.5 50 mM Mg Acetate, 0.5% Tween 20), 2.5 mg/ml heparin, 1% BSA, 100  $\mu$ g/ml yeast tRNA with 2  $\mu$ l RNaseOUT. The reaction mixture was centrifuged at 14,000  $\times$  g for 5 min at 4 °C. The supernatant was then moved to a new, ice cold tube.

## ORFeome precipitation

GST-protein coated bead preparation: Expression of GST-Pep, GST-LYN and GST-MUTED was induced with 0.1 mM IPTG at 30 °C for 4 hours. Cells were pelleted and lysed in lysis buffer (50 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.5 mM EDTA, 5 mM MgCl<sub>2</sub>, 0.2% NP-40, 2 mM DTT, 0.2 mM PMSF, 1  $\mu$ g/ml pepstatin, 1  $\mu$ g/ml aprotinin, 1  $\mu$ g/ml leupeptin, 200  $\mu$ g/ml lysozyme) on ice for 1 hour. The lysate was sonicated for 1 min on ice (Branson; output 4.0, duty cycle 50%). The lysate was centrifuged and supernatant retained. MagneGST<sup>™</sup> glutathione particles (Promega) were coated with lysate at 4 °C for 4 hours. Beads were washed with buffer I (50 mM Tris pH 7.5, 500 mM NaCl, 1 mM EGTA, 10% Glycerol, 0.1% TritonX-100, 0.1% beta-mercaptoethanol, 1 mM PMSF) three times and buffer II (50 mM HEPES pH 7.5, 100 mM NaCl, 1 mM EGTA, 10% Glycerol, 0.1% beta-mercaptoethanol, 1 mM PMSF) three times. To assess LYN phosphotyrosine binding dependence, 20  $\mu$ l glutathione particles containing approximately 2  $\mu$ g protein were treated with 400 units of Lambda protein phosphatase (NEB) in 1 $\times$  Protein MetalloPhosphatases buffer (50 mM HEPES, 100 mM NaCl, 2 mM DTT, 0.01 % Brij 35 and 1 mM MnCl<sub>2</sub>) at 30 °C for 30 min with agitation. Patient antibody coated bead preparation: PND patient cerebrospinal fluid (CSF) containing 2.0  $\mu$ g of immunoglobulin or 2.0  $\mu$ g of rabbit IgG in 400  $\mu$ l 1 $\times$  PBST (3.2 mM Na<sub>2</sub>HPO<sub>4</sub>, 0.5 mM KH<sub>2</sub>PO<sub>4</sub>, 1.3 mM KCl, 135 mM NaCl, 0.05% Tween 20, pH 7.4) containing 1% acetylated bovine serum albumin (BSA) (Cat No. B2518, Sigma-Aldrich) were incubated with protein A/G magnetic beads (Invitrogen) at 4 °C overnight. Beads were washed with RD wash buffer five times. Drug coated bead preparation: gefitinib was immobilized on magnetic beads using a procedure previously described for covalently attaching the drug to sepharose 6B beads.<sup>13</sup> Biotinylated dasatinib (biotin-dasatinib, 500  $\mu$ M) was immobilized on 50  $\mu$ l of Dynabeads<sup>®</sup> MyOne<sup>™</sup> Straptavidin T1 beads (Invitrogen) by incubation in 1 $\times$  PBST containing 10% DMSO at 4°C overnight.

An equal amount of biotin (Sigma) was immobilized on beads as negative control. For all bead types: Beads were next blocked with RD selection buffer at 4°C for 2 hours. 100 µl of the ice cold RD selection buffer containing the translated ORFeome library was then incubated with the beads at 4 °C for 6 hours while rotating. For competition experiments, free biotin-dasatinib (100 µM) was pre-incubated with the translated ORFeome library at 4°C for 2 hours prior to the incubation of the library with biotin-dasatinib beads. Beads were then washed six times with 500 µl ice cold RD wash buffer. After the final wash, ribosomal complexes were disrupted by resuspension in 50 µl EB20 elution buffer (50 mM Tris Acetate pH 7.5, 150 mM NaCl, 20 mM EDTA) containing 1 µl RNaseOUT while rotating at 37°C for 10 min. The eluted mRNA was then purified using an RNeasy® column (Qiagen).

### RT-qPCR analysis of precipitated ORFs

Eluted mRNA samples were reverse transcribed using the primer PRDREV (0.1 µM) and SuperScript® III Reverse Transcriptase according to the manufacturer. The primer pair targeting the 3' end of the corresponding ORF (0.1 µM each) was used to measure its mRNA level with Platinum® SYBR® Green qPCR SuperMix (Cat No. 11744, Invitrogen) on a 7500 Fast PCR-System (Applied Biosystems) The following thermal cycles were used for qPCR: 95 °C, 1 min/95 °C, 5 sec; 60 °C, 30 sec; 40 cycles.

### mRNA sample preparation and Illumina sequencing

Recovered mRNA samples were fragmented using NEBNext® Magnesium RNA Fragmentation Module (NEB). The reaction was performed in a preheated thermal cycler for 3 min at 94 °C. Fragmented mRNA was cleaned up using Spin-50 mini-column (USA Scientific) and subjected to reverse transcription (RT) using SuperScript® III Reverse Transcriptase (Invitrogen) and the Tola100RT primer (0.1 µM). After RT, the primer was removed by Exonuclease I (TaKaRa) digestion at 37 °C for 30 min. The mRNA template was then removed by incubation with RNase H (Invitrogen) at 37 °C for 30 min. cDNA was purified using QIAquick® PCR Purification Kit (Qiagen). Polyadenylation of cDNA 3' end was performed using a TdT reaction kit (Invitrogen) according to the manufacturer. The TdT product was purified using QIAquick® PCR Purification Kit. The 1<sup>st</sup> PCR was performed using 0.25 µl Herculase II Fusion DNA Polymerase (Agilent), TdT product as the template, and 0.2 µM reverse primer Adaptor-(dT)<sub>24</sub> in 25 µl volume. A single thermal cycle was performed: 95°C, 2 min/50°C, 1 min/72°C, 7 min. The forward primer, 0.2 µM P5-PRDREV was then added in an additional 25 µl PCR mixture. The following thermal cycles were then performed: 95°C, 2 min/95°C, 20 sec; 55°C, 30 sec; 72°C, 1 min; 30 cycles/72°C, 5 min. The 1<sup>st</sup> PCR product was purified using QIAquick PCR Purification Kit. The 2<sup>nd</sup> PCR was composed of 0.5 µl Herculase II Fusion DNA Polymerase, 100 ng 1<sup>st</sup> PCR product as DNA template, and 0.2 µM primers (forward, P5-PRDREV; reverse, INDEX primer). The following thermal cycles were then performed: 95°C, 2 min/95°C, 20 sec; 55°C, 30 sec; 72°C, 1 min; 10 cycles/72°C, 5 min. The 2<sup>nd</sup> PCR product was purified using QIAquick PCR Purification Kit. The purified 2<sup>nd</sup> PCR product was quantified on a 7500 Fast PCR-System (Applied Biosystems) using Platinum® SYBR® Green qPCR SuperMix (Invitrogen), and 2 µl of 5 µM P5 and P7\_2 mix. The following thermal cycles were used: 50 °C, 2 min/95 °C, 10 min/95 °C, 15 sec; 60 °C, 2 min; 35 cycles. An equal amount of each

2<sup>nd</sup> PCR product was combined and sequenced on the Illumina HiSeq 2000 using a 50 cycle, multiplex single-end protocol with the custom primer, PRDREV-attB2-SP.

### Analysis of Illumina data

Sequences were aligned using the Bowtie software, version 0.12.7. An index file was constructed using the 50 3'-most nucleotides of each sequence from the ORFeome v5.1. Alignments were performed using the following parameters: -n 2 -l 30 --best --nomaqround --norc -k 1. A single mismatch was allowed in the 7 nucleotide barcode sequence which was used to assign each read to the appropriate sample library. We typically obtained between 5 and 10 million aligned reads per barcoded library. The alignments corresponding to each ORF were then aggregated, thus defining each library's read count vector. We considered only ORFs with an IP count greater than a certain threshold. Enrichment was then calculated by adding a pseudocount of 1 to each clone and then dividing the fractional abundance of each IP'ed clone by that in the appropriate negative control. For LYN IP's the negative control was GST-Pep or GST-MUTED, and for PND IP's the negative control was rabbit IgG. For biotin-dasatinib, the negative control was biotin.

### Validation of candidate interactions

pDEST40 plasmid harboring candidate ORFs was transiently transfected into 293T cells using Fugene<sup>®</sup> HD transfection reagent (Promega) according to the manufacturer. 48 hours post transfection, cells were harvested in 1 ml of 1× RIPA buffer (50 mM Tris-HCl pH 7.4, 1% NP-40, 0.25% Na deoxycholate, 150 mM NaCl, and 1 mM EDTA) containing protease inhibitor cocktail (1 mM PMSF, 1 μg/ml aprotinin, 1 μg/ml leupeptin, and 1 μg/ml pepstatin) and phosphatase inhibitor cocktail (Sigma). Precipitation was performed by incubating either bait-coated glutathione or protein A/G beads with the 293T cell lysates at 4 °C overnight with rotating. Beads were washed with ice cold 1x RIPA buffer six times and eluted in 2× Laemmli sample buffer. Samples were then separated on a SDS-PAGE gel. An anti-V5 antibody was used to detect the presence of candidate proteins after transfer onto a PVDF membrane.

### Microarray hybridization

Eluted mRNAs IP'ed from mixed PND sample (patient A:patient C = 1:1) was reverse transcribed using the primer PRDREV (as above). PCR was performed using cDNA as template, attB1 and attB2 (0.2 μM) primers by PrimeSTAR<sup>®</sup> HS PCR Kit. PCR products were recombined into entry vector pDONR223 by BP reaction. After transformation into DH5αE. coli, entry clones were recovered and then recombined into the destination vector pRD-DEST with the LR reaction. After transformation into DH5αE. coli, LR clones were recovered and subjected to PCR using T7B and TolAk primers. PCR products were purified for T7 *in vitro* transcription using MEGAscript<sup>®</sup> T7 Kit according to the manufacturer's instructions (Ambion). *In vitro* transcribed PND affinity purified sample mRNAs were labeled with Cy3 dye. The total input *in vitro* transcribed ORFeome mRNAs were labeled with Cy5 dye. Cy3 and Cy5-labeled RNAs were mixed (1:1) and hybridized on our custom human ORFeome microarrays (Agilent).<sup>14</sup>

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

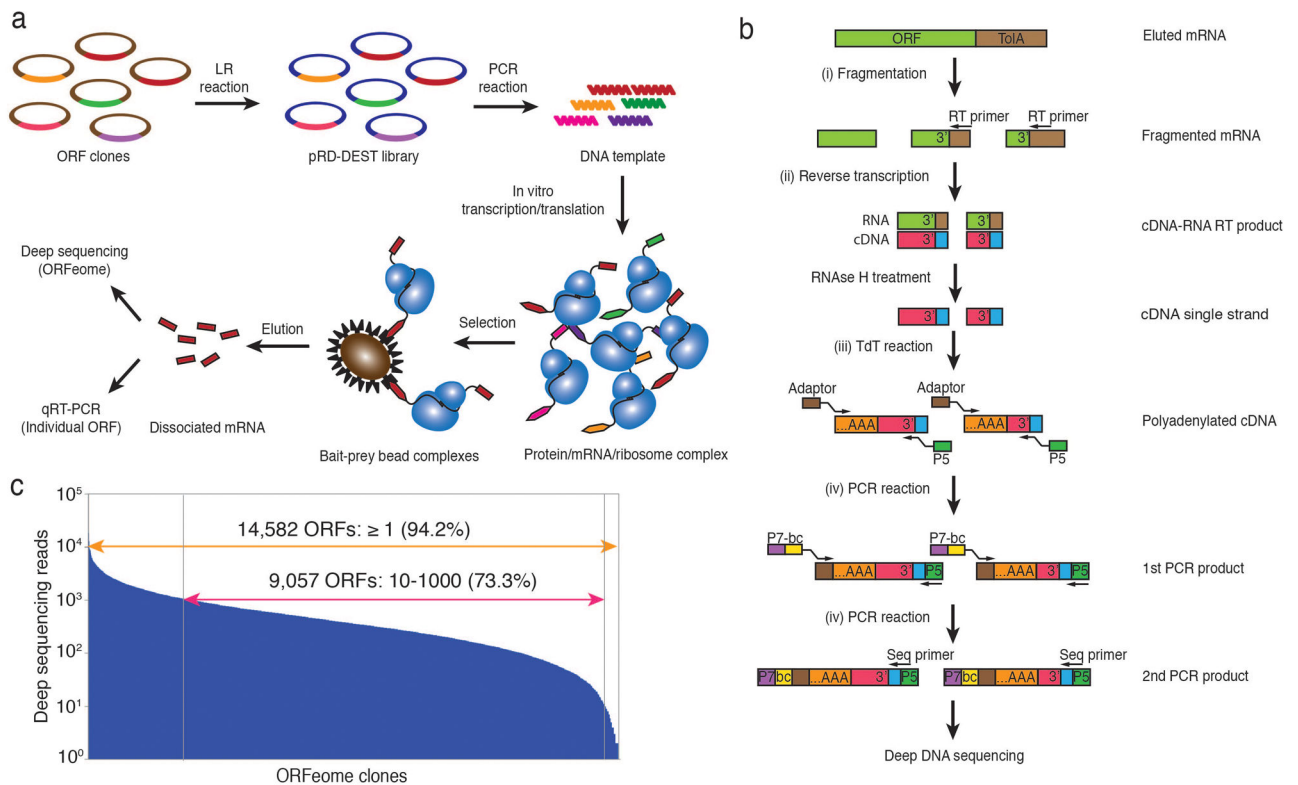
## Acknowledgments

We would like to thank K. Waraska, M. Cicero, S. Alian and A. Gagne for assistance with Illumina sequencing, and J. Laserson for statistical advice. Thanks to Dianrong Zhu for help with synthesis of biotin-dasatinib, which was partially supported by NIH U54 CA156734 to the University of Massachusetts Boston – Dana-Farber Harvard Cancer Center U54 Comprehensive Partnership (Project 3, Co-PIs: N. S. Gray, W. Zhang, and P. L. Yang). We also thank Nathanael Gray at Harvard Medical School for valuable advice, and Harold Varmus at the NCI for providing gefitinib reagents and advice. This work was supported in part by NIH grant 3P30CA023100-25S8 to S.K. S.J.E. is an investigator with the Howard Hughes Medical Institute.

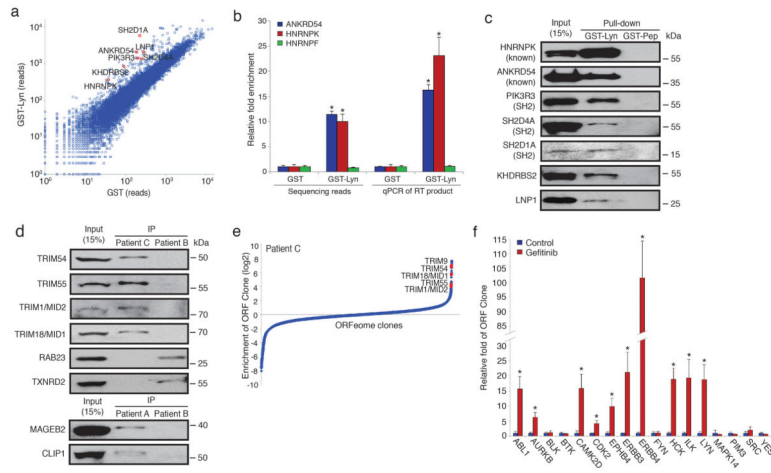
## References

1. Faix PH, et al. Phage display of cDNA libraries: enrichment of cDNA expression using open reading frame selection. *BioTechniques*. 2004; 36:1018–1022. 1024, 1026–1019. [PubMed: 15211753]
2. Zhu H, Bilgin M, Snyder M. Proteomics. *Annual review of biochemistry*. 2003; 72:783–812.
3. Jeong JS, et al. Rapid identification of monospecific monoclonal antibodies using a human proteome microarray. *Mol Cell Proteomics*. 2012; 11:O111, 016253. [PubMed: 22307071]
4. Lamesch P, et al. hORFeome v3.1: a resource of human open reading frames representing over 10,000 human genes. *Genomics*. 2007; 89:307–315. [PubMed: 17207965]
5. Amstutz P, Binz Zahnd C, Pluckthun A. Ribosome Display: In Vitro Selection of Protein-Protein Interactions. *Cell Biology, A Laborator Manual*. 2006; 1:13.
6. Boggon TJ, Eck MJ. Structure and regulation of Src family kinases. *Oncogene*. 2004; 23:7918–7927. [PubMed: 15489910]
7. Weng Z, et al. Identification of Src, Fyn, and Lyn SH3-binding proteins: implications for a function of SH3 domains. *Mol Cell Biol*. 1994; 14:4509–4521. [PubMed: 7516469]
8. Samuels AL, Klinken SP, Ingley E. Liar, a novel Lyn-binding nuclear/cytoplasmic shuttling protein that influences erythropoietin-induced differentiation. *Blood*. 2009; 113:3845–3856. [PubMed: 19064729]
9. Larman HB, et al. Autoantigen discovery with a synthetic human peptidome. *Nat Biotechnol*. 2011; 29:535–541. [PubMed: 21602805]
10. Carthagena L, et al. Human TRIM gene expression in response to interferons. *PLoS ONE*. 2009; 4:e4894. [PubMed: 19290053]
11. Brehmer D, et al. Cellular targets of gefitinib. *Cancer Res*. 2005; 65:379–382. [PubMed: 15695376]
12. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009; 37:1–13. [PubMed: 19033363]
13. Somwar R, et al. Superoxide dismutase 1 (SOD1) is a target for a small molecule identified in a screen for inhibitors of the growth of lung adenocarcinoma cell lines. *Proc Natl Acad Sci U S A*. 2011; 108:16375–16380. [PubMed: 21930909]
14. Yen HC, Xu Q, Chou DM, Zhao Z, Elledge SJ. Global protein stability profiling in mammalian cells. *Science*. 2008; 322:918–923. [PubMed: 18988847]



**Figure 1.**

Parallel analysis of *in vitro* translated ORFs (PLATO). **(a)** ORF display scheme. The pooled human ORFeome v5.1 entry vector library is attL-attR (“LR”) recombined into the pRD-DEST expression vector. Expression plasmids are PCR amplified to generate the DNA templates for *in vitro* transcription. Following *in vitro* translation, the protein-mRNA-ribosome complexes are incubated with protein, antibody or small-molecule bait immobilized on beads. The enriched mRNA library is recovered from bait-prey bead complexes for further analysis. **(b)** Processing of mRNA samples for deep DNA sequencing. After fragmentation and reverse transcription (RT) using a universal primer to recover the 3’ end of ORFeome transcripts, cDNA is polyadenylated with terminal deoxynucleotide transferase (TdT) and amplified for multiplex deep sequencing using primers containing a sample barcode and the P5 and P7 Illumina sequencing adaptors. **(c)** Sequencing reads of the unenriched human pRD-ORFeome mRNA library (the ‘input’ library). Most ORFs were sequenced at least once.



**Figure 2.** Identification of known and previously undescribed interactions using PLATO. **(a)** Interactions with LYN tyrosine-protein kinase. Scatter plot of each ORF's sequencing reads after enrichment on GST-LYN or GST. Several known and undescribed LYN binding candidates are highlighted in red. **(b)** Enrichment of two known interactors of LYN. Data were normalized to the GST enriched libraries (n=3, mean  $\pm$  s.d.; \*,  $P < 0.01$ ; t test). **(c)** Confirmation of known and predicted LYN binding proteins by affinity precipitation-western blotting of lysates from HEK 293T cells transiently overexpressing the individual V5-His-tagged candidate proteins. **(d)** Confirmation of previously unidentified autoantigens from a PND patient. **(e)** Interactions with autoantibodies. Enrichment ranking of PND autoantigens identified using CSF from patient C. **(f)** Interactions with a small molecule. Enrichment of previously identified targets of gefitinib. Data were normalized to the control-enriched libraries (n=3, mean  $\pm$  s.d.; \*,  $P < 0.05$ ; t test).