A yeast BiFC-seq method for genome-wide interactome mapping

Limin Shang^{1,a}, Yuehui Zhang^{1,b}, Yuchen Liu^{1,c}, Chaozhi Jin^{1,d}, Yanzhi Yuan^{1,e},
Chunyan Tian^{1,f}, Ming Ni^{2,g}, Xiaochen Bo^{2,h}, Li Zhang^{3,i}, Dong Li^{1,j}, Fuchu He^{1,*k} &
Jian Wang^{1,*l}

¹State Key Laboratory of Proteomics, Beijing Proteome Research Center, National
Center for Protein Sciences (Beijing), Beijing Institute of Lifeomics, Beijing 102206,
China; ²Department of Biotechnology, Beijing Institute of Radiation Medicine,
Beijing 100850, China; ³Department of Rehabilitation Medicine, Nan Lou;
Department of Key Laboratory of Wound Repair and Regeneration of PLA, College
of Life Sciences, Chinese PLA General Hospital, Beijing 100853, China;
Correspondence should be addressed to F.H. (hefc@bmi.ac.cn) and J.W.
(wangjian@bmi.ac.cn).

* Corresponding authors.

E-mail : hefc@bmi.ac.cn (Fuchu He) , wangjian@bmi.ac.cn (Jian Wang)

^a ORCID: 0000-0002-6371-1956.

^b ORCID: 0000-0001-5257-1671

^c ORCID: 0000-0003-4691-4951

^d ORCID: 0000-0002-1477-0255

^e ORCID: 0000-0002-6576-8112

^f ORCID: 0000-0003-1589-293X

^g ORCID: 0000-0001-9465-2787

^h ORCID: 0000-0003-3490-5812

ⁱ ORCID: 0000-0002-3477-8860

^j ORCID: 0000-0002-8680-0468

^k ORCID: 0000-0002-8571-2351

^l ORCID: 0000-0002-8116-7691

Total word counts:4398 (from “Introduction” to “Conclusions”)

Total Keywords: 5

30 Total abstract word count: 130
31 Total figures: 6
32 Total Supplementary Figures: 10
33 Total Supplementary tables: 9
34 Total references: 44

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60 **Abstract**

61 Genome-wide physical protein-protein interaction (PPI) mapping remains a major
 62 challenge for current technologies. Here, we report a high-efficiency yeast
 63 bimolecular fluorescence complementation method coupled with next-generation
 64 DNA sequencing (BiFC-seq) for interactome mapping. We applied this technology to
 65 systematically investigate an intraviral network of Ebola virus (EBOV). Two-thirds
 66 (9/13) of known interactions of EBOV were recaptured and five novel interactions
 67 were discovered. Next, we used BiFC-seq method to map the interactome of the
 68 tumor protein p53. We identified 97 interactors of p53 with more than three quarters
 69 are novel. Furthermore, in more complex background, we screened potential
 70 interactors by pooling two BiFC-libraries together, and revealing a network of 229
 71 interactions among 205 proteins. These results show that BiFC-seq is a highly
 72 sensitive, rapid and economical method in genome-wide interactome mapping.

73

74 **KEYWORDS:** Bimolecular fluorescence complementation; Protein protein
 75 interaction; High-throughput; Genome-wide interactome; Next-generation sequencing

76

77

78

79

80

81

82

83

84

85

86

87

88

89 **Introduction**

90 Genome-wide yeast two hybrid (Y2H) screening and affinity-purification coupled
91 with mass spectrometry (AP-MS) have been extensively used for mapping the
92 interactomes of various species [1]-[4]. However, only limited coverage was obtained
93 for the interactome of most organisms [5],[6], which dues to low sensitivity, labor
94 intensive and high cost of the technologies.

95 The bimolecular fluorescence complementation (BiFC) assay is a powerful tool to
96 investigate the binary protein-protein interactions. A Venus-based BiFC method has
97 been used in large-scale interactome mapping of telomere signaling and SUMO [7][8]
98 which specifically detected transient or weak interactions that cannot be obtained by
99 Y2H or AP-MS. However, there is high background fluorescence intensity, which
100 leads to more artificial interactions [9].

101 To overcome these limitations, we developed a BiFC-seq method, which combines
102 the yEGFP-BiFC with NGS technology, expanding the BiFC application in
103 genome-wide interactome screening. Firstly, we have used yEGFP-BiFC method to
104 depict a small-scale intraviral PPI network of EBOV among its 9 encoding proteins,
105 which contains 14 interactions among 6 proteins. Then, we applied BiFC-seq in
106 high-throughput assay to screen p53 interactors from human universal library and
107 revealed 97 p53 interactors with 21 reported in literature. Finally, we carried out
108 genome-wide interactome screening by pooling two tagged libraries together using
109 BiFC-seq, generating an interaction network consists of 229 interactions with 12
110 reported in the BioGrid database.

111 **Methods**

112 **Data availability**

113 The NGS data were uploaded to iProx database (<http://www.iprox.org/>). The
114 sequencing data of p53 interactors are accessible with
115 <https://www.iprox.org/page/PSV023.html?url=1592296064595kJiq> and Password:
116 yjJp. The sequencing data of genome-wide PPI are accessible with

117 <https://www.iprox.org/page/PSV023.html?url=1592296389579vYEx> and Password:
118 U7Gw

119 **Construction of the yEGFP-BiFC vectors**

120 The yEGFP-BiFC constructs were generated using the pDBleu and pPC86 vectors
121 (Invitrogen, Grand Island, NY) as templates according to standard molecular
122 techniques. Briefly, the coding regions of the DNA activation or binding domains
123 were removed, and the DNA fragments coding for the N or C-terminal fluorescent
124 proteins and a linker coding a 15 amino acid spacer sequence (GGGS)₃ were
125 inserted. The remaining vectors were constructed with these backbone vectors using
126 standard molecular techniques.

127 **Construction of the cDNA library**

128 The cDNA library was constructed using Gateway recombination technology. Human
129 universal reference total RNA (Catalog No. 636538, Clontech, Mountain View, CA)
130 was used as a template to synthesize cDNA by reverse transcription. The cDNA was
131 ligated to adaptors with the *attB1* site. To construct the Gateway Entry vectors,
132 pDONR222 (Catalog No. 636538, Invitrogen, Grand Island, NY) was mixed with
133 purified cDNA and the BP Clonase™ II enzyme. The reactions were incubated at
134 25°C overnight and treated with Proteinase K at 37°C for 10 minutes to terminate the
135 reaction. The reaction products were transformed into *E. coli* DH10B competent cells,
136 colonies were grown in LB medium with kanamycin selection, and plasmids were
137 extracted with a Plasmid Mini Kit (Catalog No. 12125, Qiagen, Hilden, Germany). To
138 construct the cDNA library, pDONR222 entry vectors were mixed with
139 pPC86-YN157-CCDB vectors and the Gateway LR Clonase II enzyme. The reactions
140 were incubated at 25°C overnight and treated with Proteinase K at 37°C for 10
141 minutes to terminate the reaction. The reaction products were transformed into *E. coli*
142 DH10B competent cells, colonies were grown in LB medium with ampicillin
143 selection, and plasmids were extracted and stored at -80°C.

144 **yEGFP-BiFC assay**

145 The yeast strain *S. cerevisiae* AH109 (genotypes *MATa*, *trp1-901*, *leu2-3*, 112,
146 *ura3-52*, *his3-200*, *gal4 Δ* , *gal80 Δ* , *LYS2:GAL1_{UAS}-GAL1_{TATA}-HIS3*, *GAL2_{UAS}-*
147 *GAL2_{TATA}-ADE2*, *URA3:MEL1_{UAS}-MEL1_{TATA}-lacZ*) (Clontech, Mountain View, CA)
148 was used in the BiFC screening. A small-scale sequential transformation procedure
149 was performed using protocols from the manufacturer. After transformation, the yeast
150 cells were resuspended in 10 ml of liquid SD without tryptophan and leucine (SD-2)
151 medium, cultured in a 30°C shaker for 24 hours and incubated at 4°C for an additional
152 48 hours for fluorophore maturation. The yeast cells were collected by centrifugation
153 at 7000 rpm for 30 s and resuspended with PBS. The yeast cells with reconstituted
154 yEGFP were excited with a 488 nm laser and collected through a FACS Aria III (BD
155 Biosciences, Franklin Lakes, NJ) 530/30 nm bandpass filter. The sorted yeast cells
156 were collected with PBS, spread onto SD-2 plates and cultured in a 30°C incubator
157 for 2 days until yeast colonies reached 2-3 mm in diameter. Fluorescence at 488 nm
158 excitation and 519 nm emission were observed under an inverted fluorescence
159 microscope (Olympus, Tokyo, Japan).

160 **Yeast colony PCR**

161 The yeast colonies were picked and digested with 20 μ l 0.02 N NaOH. After 5
162 minutes of boiling, 2 μ l of the solution was used as a template to amplify the inserted
163 cDNA with the forward primer (5'-ATGGCTGACAAACAAAGATCTGGT-3') and
164 the reverse primer (5'- GCGGCCGCACCACTTTGTACAAGAAA-3'). The reaction
165 buffers were mixed and aliquoted by a Biomek FX Laboratory Automation
166 Workstation (Beckman Coulter, Brea, CA) into 384-well PCR plates. The PCR
167 reactions were performed with I-5TM 2 $\times$ High-Fidelity Master Mix (MCLAB,
168 TSINGKE Biological Technology, China) according to the manufacturer's
169 instructions. For each amplification, 3 μ l of the PCR products was mixed and purified
170 using the QIAquick PCR Purification Kit (Catalog No. 28104, Qiagen, Hilden,
171 Germany).

172 **NGS and BiFC-seq workflow**

173 The purified PCR products were sheared with a Covaris S220 system (Applied
174 Biosystems, Carlsbad, CA). To prepare the libraries, the sheared PCR products were
175 processed using a Nextera XT DNA Sample Preparation Kit according to the
176 manufacturer's instructions (Illumina, San Diego, CA). All the libraries were
177 sequenced at a 2×150 bp read length on an Illumina HiSeq 2000 platform. The images
178 generated by the HiSeq 2000 were converted to raw reads using base calling
179 (CASAVA version 1.8). After trimming the adaptor sequences from the raw reads, the
180 reads with $\geq 10\%$ unidentified bases (Ns) and $\geq 50\%$ low-quality bases (PHRED
181 quality scores ≤ 5) were also removed to implement quality control. The filtered reads
182 were mapped to the GRCh38 genome (Human GENCODE version 22) by TopHat
183 [10]. The raw counts of the sequencing reads for each transcript were calculated by
184 the HTSeq Python package [11]. DESeq2 was then used for different expression
185 analysis between p53 and control groups [12]. A Gene Ontology analysis was then
186 performed with the clusterProfiler R package
187 (<http://www.bioconductor.org/packages/devel/bioc/html/ClusterProfiler.html>).

188 **Co-IP assay**

189 HEK293 cells were cultured in Dulbecco's modified Eagle's medium (Invitrogen,
190 Grand Island, NY) with 10% (v/v) fetal bovine serum (FBS). After 48 hours of
191 transfection, the HEK293 cells with Flag-p53 and Myc-interactors were harvested and
192 lysed in HEPES lysis buffer [0.5% NP-40, 150 mM NaCl, 1% (v/v) Tween 20, 50 mM
193 Tris-HCl (pH 7.5), 0.1% 1 M DTT, and a protease inhibitor cocktail]. The
194 immunoprecipitations were performed using anti-Myc antibodies (Sigma-Aldrich, St.
195 Louis, MO) and protein A/G-agarose (ThermoFisher, San Jose, CA) at 4°C. The
196 lysates and immunoprecipitates were detected by the anti-Flag antibody, which was
197 followed by detection with the ECL substrate (ThermoFisher, San Jose, CA).

198

199

200 **Results and discussion**

201 **Development of yEGFP-based BiFC**

202 A codon-optimized yeast enhanced green fluorescent protein (yEGFP) [13] that was
 203 split into nonfluorescent halves at the 157 aa position to yield YN157 (1-157 aa) and
 204 YC157 (158-238 aa) fragments. The intact yEGFP had a strong fluorescence signal in
 205 yeast, whereas the split yEGFP produced undetectable fluorescence (Figure S1). To
 206 verify that this approach could be used for detecting PPIs, the fragments of yEGFP
 207 linked by (GGGS)₃ peptide sequence were fused to the N-terminal of b-Jun, b-Fos
 208 and ΔbFos as positive and negative indicators (**Figure 1A**). The co-transfection of
 209 YN157-bJun and YC157-bFos led to the reassembly of yEGFP and showed a strong
 210 fluorescence signal. However, the combination of YN157-bJun and YC157-ΔbFos
 211 showed slight background fluorescence (**Figure 1B**). The calculated signal to noise
 212 ratio (S/N) of the yEGFP-BiFC method was 33:1 for all the cells and 4:1 for the
 213 fluorescent-only cells, which allows for high-sensitivity detection of PPIs. The
 214 percentage of fluorescent cells in the positive group was 8-fold higher than that in the
 215 negative group under identical transfection conditions (**Figure 1C**). As expected, the
 216 fluorescent positive cell count was 16.5% by FACS, which was approximately 33-fold
 217 higher than the negative control (**Figure 1D**). The different fusion patterns of
 218 interacting proteins with the yEGFP fragments potentially affected the proximity of
 219 the split fluorescent protein fragments. Thus, we constructed b-Jun and b-Fos
 220 N-terminal or C-terminal fused proteins and tested these combinations in yeast
 221 (Figure 1a). We found that only the N-terminal fusion of YN157-bJun and the N- or
 222 C-terminal fusions of YC157-bFos and bFos-YC157 produced strong BiFC signals,
 223 whereas the other combinations produced slight fluorescent signals (Figure 1C).
 224 These results suggest that the yEGFP-based BiFC system is suitable for PPI screening
 225 in yeast and indicate that variation in the fusing of yEGFP fragments to targets can
 226 potentially affect screening results.

227 **An intraviral PPI network of Ebola virus by yEGFP-BiFC**

228 The recent West African epidemic of EBOV disease caused more than 11,000 deaths
 229 owing to its high infectivity and lethality. Although tremendous efforts have been
 230 made to the development of EBOV therapeutics, an effective treatment remains

231 elusive. The researches of pathogenic mechanisms of EBOV were mainly based on
232 the pathogen-host PPI features [14],[15]. However, the intraviral PPI network is still
233 unclear. Therapeutic strategy targeting interactions among internal EBOV proteins can
234 be promising. Thus, we carried out a comprehensive analysis of interactions of EBOV
235 proteins with each other using yEGFP-based BiFC. The EBOV genome encodes nine
236 proteins: nucleoprotein (NP), viral protein (VP24, VP30, VP35 and VP40),
237 glycoprotein (GP) and polymerase (L). The GP gene encodes for another two proteins
238 (sGP and GP2) through RNA editing [16].

239 To exhaustively explore the PPI networks of 9 encoding proteins of EBOV, we
240 fused each protein with YN157 or YC157 fragment, all in both orientations, yielding
241 36 fluorescent fragment tagged proteins. After validation of their expressions in yeast
242 cells (Figure S2), these tagged proteins were pairwise cotransformed into yeast cells
243 in an 18×18 matrix format in triplicate to screen intraviral PPIs of EBOV, bJun and
244 bFos/ΔbFos were used as controls. Transformed cells were analyzed by flow
245 cytometry and fluorescence microscope (Figure S3). The ratio of fluorescent cells in
246 each tested group was compared with the negative control (bJun and ΔbFos), and
247 those with significant difference ($P<0.05$) and the $\log_2$ ratio (test/control) greater than
248 2 were regarded as authentic PPIs. Finally, 51 interactions among 19 tagged proteins
249 were obtained (**Figure 2A**). After removing redundancy, the dataset contains 14
250 interactions among 6 EBOV proteins, of which 9 interactions were reported (**Figure**
251 **2B**). The interaction patterns of different orientations between fluorescent fragment
252 tags and EBOV proteins varied greatly by BiFC method. The VP35-NP interaction
253 can be detected by all pairwise combinations of tagged VP35 and NP proteins,
254 whereas the VP24-GP2 interaction can only be detected by VP24-YC157 and
255 GP2-YN157 combination (Figure S3), indicating the advantage of different fusion
256 combinations in overcoming steric hindrance to detect PPIs by BiFC technology.
257 Furthermore, by overexpressing Myc- and Flag- tagged EBOV proteins, we
258 confirmed 10 of 14 interactions (71.4%) by co-immunoprecipitation assay (co-IP) in
259 mammalian cells (**Figure 2C**). Our new findings are consistent with their biological
260 roles, for example, [the interactions of VP24-VP30 and VP24-GP2 suggest the role of](#)

261 VP24 in ribonucleoprotein complex (RNP) assembly and release process of GP2 out
 262 of the virus. The interactions of VP40 with VP24 and VP30 also confirmed that the
 263 VP40 involves in the process of releasing of viral particles from the host cells. By
 264 integrating our EBOV-BiFC protein networks with literature-curated datasets, a global
 265 intraviral EBOV protein networks containing 18 interactions among 7 proteins was
 266 obtained (Figure 2B) [17]-[23]. The comprehensive description of interplay among
 267 EBOV proteins provide a basis for studying protein functions in the aspect of viral
 268 replication and particle secretion as well as for advancing novel therapeutic
 269 approaches.

270

271 **Screening p53 interactors of human library by using BiFC-seq**

272 To further show that the BiFC-seq method is suitable for high-throughput screening,
 273 the well-known tumor suppressor p53 was used as a bait to screen its interactors
 274 against the human universal library fused with YN157 fragment (YN157-library). We
 275 used HDM2, a known partner of p53 as positive control [24]. The mutant form of
 276 HDM2 (Δ HDM2, with deletion of 1-139 aa), which lost its ability of binding to p53
 277 was used as negative control. Fluorescent cells was effectively distinguished between
 278 positive and negative controls by yEGFP-BiFC (Figure S4A and B), we also tested
 279 various linkers [25] between p53 or HDM2 and yEGFP halves by yEGFP-BiFC
 280 method, we found that the (GGGS)₃ linker produced the highest fluorescence ratio
 281 among the four tested linkers by flow cytometry (Figure S4C and D). To screen p53
 282 interactors against the library, the YN157-library was sequentially transformed into
 283 yeast competent cells containing p53 (YC157-p53 or p53-YC157) or their
 284 corresponding control groups with YC157 fragments (YC157-linker and
 285 linker-YC157), each transformation was performed in triplicate. After 72 hours of
 286 cultivation in SD liquid medium without tryptophan and leucine (SD-2),
 287 approximately 2,000 fluorescent cells out of 1×10^8 cells were obtained for each p53
 288 and control screening groups that were sorted by FACS (Figure S5), which is 10-100
 289 times the number of positive colonies obtained with traditional Y2H screening.

290 To determine the identities of the positive colonies, PCR products of yeast colonies

from each group were mixed and purified for NGS (Figure S6). The correlation coefficient of RNA-seq results from p53 groups among triplicate screenings was significantly higher than that of their corresponding controls (**Figure 3B**). After the different expression analysis by bioinformatics tools (**Figure 3A**), we identified 61 and 43 interactors for each of YC157-p53 and p53-YC157 screening groups. There are seven overlapped interactions between two groups (**Figure 3C** and Table S1). Totally 913 non-redundant interactions of p53 were recorded in BioGrid database [26]. To our knowledge this dataset is most comprehensive and we referred to these interactions as ‘known’ interactions. Twenty-one interactions of p53 were recaptured by this study. We estimated the sensitivity of BiFC-seq is ~2% (21/913), which is similar to the large-scale Y2H screening (the sensitivity is ~1%) [40]. The specificity of BiFC-seq was evaluated by co-IP assay. Totally 16 interactions of p53 were validated and 11 of them were re-confirmed. Thus the specificity of BiFC-seq is more than 60% (11/16) (**Figure 4C**). The sampling sensitivity was estimated by three repeat screens of YC157-p53 and p53-YC157. Totally 61 interactions were identified for the YC157-p53 group. Of these interactions, 45 interactions are commonly identified. More interactions were found with additional screens. The sampling sensitivity is 73.7% per screen. Similar result was obtained for the p53-YC157 data and the sampling sensitivity is 83.7%. Based on this data, we estimated at least six screens are needed to reach 90% saturation (Figure S7). In addition, the screening results were evaluated by PRINCESS tool, which was developed to assess the reliability of binary PPIs [27]. More than 64% of the interactions were of high confidence and had cutoff ratios greater than 2 (excluding the genes that were not recognizable) (Table S2). Then, we analyzed the function of p53 interactors by Gene Ontology, consistent with previous reports, the p53 partners participate in apoptosis, DNA damage response and translation *etc* (**Figure 4A** and Table S3). We further performed enrichment analysis of p53 by excluding 21 known p53 interactors. Most of the enriched GO terms are not changed, such as apoptosis, RNA catabolic process and translation *etc* (Figure S8 and Table S4). Four subnetworks of p53 interactors were enriched in the biological processes (**Figure 4B**), suggesting a mediator role of p53 among these subnetworks.

321 The interaction of p53 with 5 interactors involved in the cilia formation subnetwork
 322 indicates that the p53 might be associated with dividing of centrosomes [28], which is
 323 consisted with previous reports in mouse cells [29], indicating the reliability of our
 324 screening results. To estimate the reproducibility of BiFC-seq, we analyzed the data of
 325 using p53 as bait. For the three replicates of YC157-p53, 45 of 61 interactions are
 326 consistently identified. For the p53-YC157 groups, 36 of 43 interactions are
 327 consistently identified in three replicates. These results indicate that for the library
 328 screening of BiFC-seq, about 77.9% are reproducible. A part of interactions of p53
 329 were validated by co-IP and BiFC assay (Figure 4C and D).

330 The AP-MS method can be used to obtain protein complex information. Thus, we
 331 compared our data with a large-scale endogenous protein complex study [30]. A total
 332 of 50 high-confidence interactors of p53 were found by AP-MS, 13 of which were
 333 recorded in the BioGrid database (Table S5). Two of these interactions are overlapped
 334 with this study. AP-MS identifies members of stable complexes, whether they directly
 335 interact with the bait or not. In contrast to AP-MS method, BiFC-seq identifies binary
 336 interactions between pairs of proteins and tends to enrich transient interactions.
 337 Meanwhile, the stable binary interactions could be also identified by BiFC-seq
 338 method, such as interactions between ribosome proteins. The low overlap rate of the
 339 data from AP-MS and BiFC-seq might due to the differential features of technologies.

340 **A library to library interactome screening using BiFC-seq technology**

341 Based on the aforementioned results, we extended the BiFC-seq screen to a
 342 genome-wide scale. Thus, we fused the human universal library to the C-terminal of
 343 YC157 fragment, yielding YC157-library pools. YN157-library and YC157-library
 344 pools were co-transformed into AH109 yeast component cells. After cultivation for 72
 345 hours in SD-2 liquid medium followed by FACS analysis, fluorescent cells were
 346 divided into three groups according to their fluorescence intensities and
 347 approximately 3000 fluorescent cells were sorted out from each group (Figure S9).
 348 Yeast colonies were picked as PCR templates until their diameters reached 2-3 mm.
 349 To identify the interacting partners within the same fluorescent yeast cell from library

screening by a single NGS experiment, we joined the encoding sequences of interacting pair into one PCR amplicon using stitching PCR strategy, which contains of two rounds of PCR (**Figure 5A**). In the first round, yeast colony PCR was conducted to amplify encoding sequences of interactors from YN157-library and YC157-library individually. In the second round, amplicons from the first round were used as templates for the stitching PCR to combine the amplicons together by their corresponding complementary upstream primers (Figure S10). To exclude most non-specific PCR amplicons from the first round PCR, stitching PCR products with more than 1000 bp fragments were mixed and purified for the NGS analysis using Illumina Miseq platform. Approximately 1 million ~300 bp paired-end reads were generated from each group, the stitching amplicons that contain encoding sequences of interactors from both library pools linked by complementary sequences were extracted and mapped to the CDS regions of the human reference genome (GRCH38). Both of the gene encoding sequences from the same read that are in frame with the open reading frame (ORF) were considered to be interactors. After removing non-specific binding proteins that detected by screening interactors against YN157-library using YC157 fragments as baits, a network of 229 interactions among 205 proteins constituted by PPIs from three screening groups was generated and 12 interactions were reported in the BioGrid database (**Figure 6A** and Table S6). A minority of interactions were overlapped among three groups (**Figure 5B**), which indicated that a more wide range of PPI can be revealed by classifying the yeast cells with different fluorescent intensities. Totally, 137 interactions recognized by PRINCESS tool, 26 had high confidence with a cutoff ratio greater than 2 (Table S7). We classified these interactors by PANTHER [31]. Compared with all of proteins encoded in the human genome, these interactors have similar functional distribution, which suggest that the BiFC-seq method has no bias to obtain certain classes of PPI (**Figure 5C**). Furthermore, we found several clusters by superimposing PPI of this study onto BioGrid datasets (**Figure 6B-D**), such as gene expression, metabolic process, *etc.* For example, the FXYP domain containing ion transport regulator 2 (FXYP2), which was only found to be interacted with small molecule of

380 [cyclothiazide](#) in BioGrid database [32], participates in protein metabolic process
 381 through interacting with translation related partners (**Figure 6E**). [The resources of](#)
 382 [protein-protein interactions](#) are very useful for therapeutic discovery, such as drug
 383 repurposing. The potential drugs are identified by network-based methodologies. For
 384 example, dozens of candidate drugs were prioritized for SARS-CoV-2 by using the
 385 [interactome of virus and host](#) [33],[34]. For our interesting, the connections of
 386 ANXA2-DAXX and SNRPB2-PSMB6 were involved in the viral infection processes
 387 and regulation of immune process (**Figure 6F and G**). These results demonstrate the
 388 great potential of BiFC-seq technology in genome-wide interactome screening.

389 **Conclusion**

390 To our knowledge, the large-scale PPI studies using BiFC method were mainly based
 391 on optimized Venus protein, which is a well-worked system in mammalian cells.
 392 However, when applied in other organisms, such as yeast and plant, it has low
 393 signal-to-noise ratio and high background signal. Here, we developed a yEGFP-based
 394 yeast BiFC method with a higher sensitivity and specificity. Compared with
 395 traditional Y2H method, the baits are neither required to import into nucleus to
 396 activate reporters nor needed to test self-activation before screening, which has
 397 extended the scope of PPI screenings.

398 To validate the feasibility of using yEGFP-BiFC method in interactome research,
 399 we searched the interactors of the nine encoding proteins of EBOV. The results
 400 revealed a network of 14 interactions among 6 proteins with a high coverage of
 401 reported interactions (9/14) and 10 interactions were validated by co-IP assays. By
 402 integrating the PPIs of EBOV in this study with reported interactions, we generated a
 403 global EBOV PPI network, which provides new insights into pathogenesis from the
 404 perspective of the intraviral network. [Compared with genetics approaches, such as](#)
 405 [gene-trap insertional mutagenesis](#) [35], RNAi [36] or CRISPR-CAS9 screening [37],
 406 [which reveals gene-gene interactions](#) and most of them are genetics interactions, the
 407 [protein-protein interactions identified by BiFC-seq](#) are mostly physical and binary
 408 [interactions](#). The interactions from these two kinds of methods are complementary.

409 Next, we justified the feasibility of BiFC-seq in a high-throughput PPI screenings,
 410 p53 was employed to screen against a human cDNA-library. We obtained 21 reported
 411 and 76 novel interactors of p53 from 97 partners. Seven interactors were overlapped
 412 for the N-terminal and C-terminal labeling p53, which demonstrated the necessity of
 413 using different fusion constructs in PPI screening. For our interesting, we found four
 414 enriched subnetworks formed by p53 interactors. The involvement of p53 in the cilia
 415 formation suggests its role in restraining overgrowth of centrosomes, which is a
 416 prerequisite for cancer development in mice [38]. We also confirmed 11 interactions
 417 of p53 by co-IP assay, which suggests the high-confidence of these interactions.
 418 Moreover, we found some membrane proteins, such as calcium modulating ligand
 419 (CAMLG), secretory carrier membrane protein 3 (SCAMP3) and interferon induced
 420 transmembrane protein 3 (IFITM3) that interact with p53, which have low possibility
 421 to be detected by traditional Y2H method.

422 Compared with the widely used Y2H and AP-MS technologies for genome-wide
 423 interactome mapping, the BiFC-seq approach is highly sensitive, has excellent
 424 screening capability, and incurs relatively low financial and labor costs. We obtained a
 425 total of 97 p53 interactions, of which 21 were previously known, whereas only 51
 426 interactions were obtained and recorded in the BioGrid database for three
 427 proteomic-scale Y2H screenings using p53 as bait (Table S8) [39]-[41]. Thus, the
 428 BiFC-seq approach had a higher sampling sensitivity than the traditional Y2H library
 429 or array screening approaches.

430 Furthermore, we extended the BiFC-seq method to a genome-wide interactome
 431 screening by pooling YN157 and YC157 tagged human cDNA libraries together. By
 432 analyzing 9000 fluorescent cells from three screening groups with different
 433 fluorescence intensities, we obtained 229 interactions of 205 proteins in a single
 434 BiFC-seq screening. The interactions participate in various biological processes, such
 435 as gene expression, metabolism and viral infection, which suggests that the BiFC-seq
 436 technology is high-sensitive to delineate genome-wide interactome in future
 437 application. We have compared the interactions identified by BiFC-seq with
 438 non-interacting protein pairs from Negatome dataset [42]. The transcriptomics data of

439 74 tissues were from human protein atlas database (HPA) [43]. The interactions with
440 absolute value of spearman correlation coefficient (SCC) greater than 0.5 were
441 considered positive co-expression pairs. Compared with Negatome data, the p53
442 interactors were more likely to be co-expressed. The ratio of co-expressed interactions
443 of p53 is 20.8%, while for Negatome data the value is 11.4%. For the interactions of
444 library screening, the ratio (13.1%) is slightly greater than that of Negatome data
445 (Table S9).

446 Compare with traditional approaches, such as Y2H and AP-MS technologies, the
447 BiFC-seq has its limitations. Most of the Y2H method employs more than one
448 reporters while BiFC-seq depends on EGFP reporter only, which might lead to more
449 false positives. Moreover, in contrast to AP-MS, BiFC-seq is amenable in yeast cells,
450 which results in the interactions that depends on post-translational modification or
451 mediated by other proteins couldn't be detected. AP-MS is performed *in vivo*, which
452 investigates protein complexes under physiological conditions.

453 Recently, the CrY2H-seq technology was introduced, which uses a Cre
454 recombinase reporter and next-generation DNA sequencing to identify interactions
455 [44]. CrY2H-seq is a highly scalable screening approach to detect binary
456 protein-protein interactions in yeast. Similar to BiFC-seq, CrY2H-seq is
457 cost-effective and time-efficient, which greatly increases interaction detection rates.
458 On the other hand, since CrY2H-seq is based on yeast transcription factor that the
459 self-activation of considerable baits is inevitable (~16% baits are self-activated).
460 CrY2H-seq uses auxotrophic rescue reporter, which usually takes more time (2~5 day)
461 to allow the growth of yeast. BiFC-seq uses a fluorescent reporter, and the signal can
462 be detected within 2 days in yeast. Compare with CrY2H-seq, a single reporter of
463 BiFC-seq might increase the false positives.

464 It should be noted that only limited interactions were identified for the human
465 interactome by a library-library screening in this study. The major reasons are as
466 following. First, the quality of cDNA library is limited. The high abundance cDNA,
467 such as ribosomal genes, accounts for the top list the identified interactions. The
468 variation of cDNA isoforms and truncations also restrict the number of unique

469 interactions after removing the redundant interactions. Second, it remains difficult to
 470 clone all of the human genes to BiFC-seq vectors, which might greatly reduce the
 471 redundant interactions and false positives from truncations or out-of-frame
 472 expressions. They cloned 1956 genes to CrY2H-seq vectors and 8577 interactions
 473 were obtained, which partially reflects the importance of cloning the full-length genes.
 474 Third, we sorted 9000 positive colonies, which represent about 200 interactions. To
 475 obtain more interactions by library to library screening, more colonies need to be
 476 identified in future.

477 In summary, we have developed a BiFC-seq method, which is amenable in
 478 high-throughput PPI screening by using FACS. The sensitivity of BiFC-seq was
 479 substantially increased by coupling with NGS. This technology is highly sensitive for
 480 screening physical binary PPIs and can be easily applied to other large-scale
 481 interactome mapping studies.

482

483 **Authors' contributions**

484 L.S. performed BiFC screens. Y.Z. performed the co-IP assays. Y. L. performed the
485 PCR assays. L.S., C.J., Y.Y., C.T., D.L., L.Z., and J.W. contributed to the data
486 analysis. M.N. and X.B. performed the NGS experiment. L.S. and J.W. wrote the
487 manuscript. F.H. and J.W. conceived and supervised the study.

489 **Competing interests**

490 The authors declare no competing financial interests.

492 **Acknowledgments**

493 We thank Professor Alistair from University of Aberdeen to kindly provide
494 pUC19-YGFP plasmid. This work was supported by grants from the the National
495 Key Research and Development Program (No. 2017YFA0505700) and the grant
496 from the National Key Lab of Proteomics (SKLP-K201805, SKLP-K201804).

499 **References**

- 500 [1] Huttlin EL, Bruckner RJ, Paulo JA, Cannon JR, Ting L, Baltier K, et al.
501 Architecture of the human interactome defines protein communities and disease
502 networks. *Nature* 2017;545:505-9.
- 503 [2] Shah PS, Link N, Jang GM, Sharp PP, Zhu T, Swaney DL, et al. Comparative
504 Flavivirus-Host Protein Interaction Mapping Reveals Mechanisms of Dengue
505 and Zika Virus Pathogenesis. *Cell* 2018;175:1931-45 e18.
- 506 [3] Choi SG, Richardson A, Lambourne L, Hill DE, Vidal M. Protein Interactomics
507 by Two-Hybrid Methods. *Methods Mol Biol* 2018;1794:1-14.
- 508 [4] Vidal M, Fields S. The yeast two-hybrid assay: still finding connections after 25
509 years. *Nat Methods* 2014;11:1203-6.
- 510 [5] Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, et al.
511 Disease networks. Uncovering disease-disease relationships through the

512 incomplete interactome. *Science* 2015;347:1257601.

513 [6] Venkatesan K, Rual JF, Vazquez A, Stelzl U, Lemmens I, Hirozane-Kishikawa T,
514 et al. An empirical framework for binary interactome mapping. *Nat Methods*
515 2009;6:83-90.

516 [7] Sung MK, Lim G, Yi DG, Chang YJ, Yang EB, Lee K, et al. Genome-wide
517 bimolecular fluorescence complementation analysis of SUMO interactome in
518 yeast. *Genome Res* 2013;23:736-46.

519 [8] Miller KE, Kim Y, Huh WK, Park HO. Bimolecular Fluorescence
520 Complementation (BiFC) Analysis: Advances and Recent Applications for
521 Genome-Wide Interaction Studies. *J Mol Biol* 2015;427:2039-55.

522 [9] Lee OH, Kim H, He Q, Baek HJ, Yang D, Chen LY, et al. Genome-wide YFP
523 fluorescence complementation screen identifies new regulators for telomere
524 signaling in human cells. *Mol Cell Proteomics* 2011;10:M110 001628.

525 [10] Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, et al. Differential
526 gene and transcript expression analysis of RNA-seq experiments with TopHat
527 and Cufflinks. *Nat Protoc* 2012;7:562-78.

528 [11] Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with
529 high-throughput sequencing data. *Bioinformatics* 2015;31:166-9.

530 [12] Love MI, Huber W, Anders S. Moderated estimation of fold change and
531 dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.

532 [13] Cormack BP, Bertram G, Egerton M, Gow NA, Falkow S, Brown AJ.
533 Yeast-enhanced green fluorescent protein (yEGFP): a reporter of gene
534 expression in *Candida albicans*. *Microbiology* 1997;143 (Pt 2):303-11.

535 [14] Gustin JK, Bai Y, Moses AV, Douglas JL. Ebola Virus Glycoprotein Promotes
536 Enhanced Viral Egress by Preventing Ebola VP40 From Associating With the
537 Host Restriction Factor BST2/Tetherin. *J Infect Dis* 2015;212 Suppl 2:S181-90.

538 [15] Garcia-Dorival I, Wu W, Dowall S, Armstrong S, Touzelet O, Wastling J, et al.
539 Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus
540 biology through targeted inhibition of host-cell protein function. *J Proteome Res*
541 2014;13:5120-35.

- 542 [16]Baseler L, Chertow DS, Johnson KM, Feldmann H, Morens DM. The
543 Pathogenesis of Ebola Virus Disease. *Annu Rev Pathol* 2017;12:387-418.
- 544 [17]Groseth A, Charton JE, Sauerborn M, Feldmann F, Jones SM, Hoenen T, et al.
545 The Ebola virus ribonucleoprotein complex: a novel VP30-L interaction
546 identified. *Virus Res* 2009;140:8-14.
- 547 [18]Huang Y, Xu L, Sun Y, Nabel GJ. The assembly of Ebola virus nucleocapsid
548 requires virion-associated proteins 35 and 24 and posttranslational modification
549 of nucleoprotein. *Mol Cell* 2002;10:307-16.
- 550 [19]Reid SP, Cardenas WB, Basler CF. Homo-oligomerization facilitates the
551 interferon-antagonist activity of the ebolavirus VP35 protein. *Virology*
552 2005;341:179-89.
- 553 [20]Watanabe S, Noda T, Kawaoka Y. Functional mapping of the nucleoprotein of
554 Ebola virus. *J Virol* 2006;80:3743-51.
- 555 [21]Hartlieb B, Muziol T, Weissenhorn W, Becker S. Crystal structure of the
556 C-terminal domain of Ebola virus VP30 reveals a role in transcription and
557 nucleocapsid association. *Proc Natl Acad Sci U S A* 2007;104:624-9.
- 558 [22]Liu Y, Stone S, Harty RN. Characterization of filovirus protein-protein
559 interactions in mammalian cells using bimolecular complementation. *J Infect*
560 *Dis* 2011;204 Suppl 3:S817-24.
- 561 [23]Johnson RF, McCarthy SE, Godlewski PJ, Harty RN. Ebola virus VP35-VP40
562 interaction is sufficient for packaging 3E-5E minigenome RNA into virus-like
563 particles. *J Virol* 2006;80:5135-44.
- 564 [24]Kruse JP, Gu W. Modes of p53 regulation. *Cell* 2009;137:609-22.
- 565 [25]Petschnigg J, Groisman B, Kotlyar M, Taipale M, Zheng Y, Kurat CF, et al. The
566 mammalian-membrane two-hybrid assay (MaMTH) for probing
567 membrane-protein interactions in human cells. *Nat Methods* 2014;11:585-92.
- 568 [26]Chatr-Aryamontri A, Breitkreutz BJ, Oughtred R, Boucher L, Heinicke S, Chen
569 D, et al. The BioGRID interaction database: 2015 update. *Nucleic Acids Res*
570 2015;43:D470-8.
- 571 [27]Li D, Liu W, Liu Z, Wang J, Liu Q, Zhu Y, et al. PRINCESS, a protein

572 interaction confidence evaluation system with multiple data sources. *Mol Cell*
573 *Proteomics* 2008;7:1043-52.

574 [28]Gupta GD, Coyaud E, Goncalves J, Mojarad BA, Liu Y, Wu Q, et al. A Dynamic
575 Protein Interaction Landscape of the Human Centrosome-Cilium Interface. *Cell*
576 2015;163:1484-99.

577 [29]Bazzi H, Anderson KV. Acentriolar mitosis activates a p53-dependent apoptosis
578 pathway in the mouse embryo. *Proc Natl Acad Sci U S A* 2014;111:E1491-500.

579 [30]Malovannaya A, Lanz RB, Jung SY, Bulynko Y, Le NT, Chan DW, et al.
580 Analysis of the human endogenous coregulator complexome. *Cell*
581 2011;145:787-99.

582 [31]Mi H, Muruganujan A, Thomas PD. PANTHER in 2013: modeling the evolution
583 of gene function, and other gene attributes, in the context of phylogenetic trees.
584 *Nucleic Acids Res* 2013;41:D377-86.

585 [32]Imming P, Sinning C, Meyer A. Drugs, their targets and the nature and number
586 of drug targets. *Nat Rev Drug Discov* 2006;5:821-34.

587 [33]Zhou Y, Hou Y, Shen J, Huang Y, Martin W, Cheng F. Network-based drug
588 repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov*
589 2020;6:14.

590 [34]Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A
591 SARS-CoV-2 protein interaction map reveals targets for drug repurposing.
592 *Nature* 2020.

593 [35]Cheng F, Murray JL, Zhao J, Sheng J, Zhao Z, Rubin DH. Systems
594 Biology-Based Investigation of Cellular Antiviral Drug Targets Identified by
595 Gene-Trap Insertional Mutagenesis. *PLoS Comput Biol* 2016;12:e1005074.

596 [36]Laufer C, Fischer B, Billmann M, Huber W, Boutros M. Mapping genetic
597 interactions in human cancer cells with RNAi and multiparametric phenotyping.
598 *Nat Methods* 2013;10:427-31.

599 [37]Shen JP, Zhao D, Sasik R, Luebeck J, Birmingham A, Bojorquez-Gomez A, et al.
600 Combinatorial CRISPR-Cas9 screens for de novo mapping of genetic
601 interactions. *Nat Methods* 2017;14:573-6.

602 [38]Coelho PA, Bury L, Shahbazi MN, Liakath-Ali K, Tate PH, Wormald S, et al.
603 Over-expression of Plk4 induces centrosome amplification, loss of primary cilia
604 and associated tissue hyperplasia in the mouse. *Open Biol* 2015;5:150209.

605 [39]Wang J, Huo K, Ma L, Tang L, Li D, Huang X, et al. Toward an understanding
606 of the protein interaction network of the human liver. *Mol Syst Biol* 2011;7:536.

607 [40]Rual JF, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, et al.
608 Towards a proteome-scale map of the human protein-protein interaction network.
609 *Nature* 2005;437:1173-8.

610 [41]Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, et al. A
611 human protein-protein interaction network: a resource for annotating the
612 proteome. *Cell* 2005;122:957-68.

613 [42]Smialowski P, Pagel P, Wong P, Brauner B, Dunger I, Fobo G, et al. The
614 Negatome database: a reference set of non-interacting protein pairs. *Nucleic
615 Acids Res* 2010; 38:D540-4.

616 [43]Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A,
617 et al. Proteomics. Tissue-based map of the human proteome. *Science* 2015;
618 347:1260419.

619 [44]Trigg SA, Garza RM, MacWilliams A, Nery JR, Bartlett A, Castanon R, et al.
620 CrY2H-seq: a massively multiplexed assay for deep-coverage interactome
621 mapping. *Nat Methods* 2017; 14:819-25.

622

623 **Figure legends**

624 **Figure 1** Development of yEGFP-based BiFC. (A) **Diagram of different fusion**
 625 **patterns of yEGFP splitted fragments of control groups.** (B) Yeast cells cotransformed
 626 with control groups as indicated were detected under fluorescence microscope, the
 627 nuclei were labeled with DAPI. DIC: differential interference contrast microscopy
 628 scale bar: 12.5 μ m. (C) Quantitative results of the fluorescence intensity of the yeast
 629 fluorescent cells. Total cell signal to noise ratio (S/N): 200 transformed yeast cells
 630 from positive and negative groups were randomly picked up, fluorescent intensities
 631 were analyzed by ImageJ and sumed up from each group to calculate Total S/N.
 632 Fluorescent cell S/N: the fluorescent intensities of 200 fluorescent cells from
 633 positive and negative groups were sumed up from each group to calculate
 634 Fluorescent cell S/N; and fluorescent cell ratio: the percentage of
 635 fluorescence-positive cells out of the total cells. Data are presented as means $\pm$ SD
 636 (n=3). Error bars: s.d. ** $P < 0.01$, Student's t-test. (D) Approximately 20,000 yeast
 637 cells from positive and negative controls were analyzed by flow cytometry, the
 638 fluorescence-positive cell ratios were 16.5% and 0.5% respectively.

639

640 **Figure 2** **An intra-EBOV PPI network revealed by BiFC** (A) An 18 $\times$ 18
 641 cotransformation of yEGFP fragments tagged EBOV proteins were carried out as
 642 indicated, interactors are marked with a black circle, bJun and bFos/ Δ bFos were
 643 used as control indicators. Data are presented as means $\pm$ SD (n=3), Student's t-test.
 644 (B) A global intra-PPI network of EBOV combined by our BiFC screening result
 645 with literature-curated datasets. (C) Validation of interactions among EBOV proteins
 646 by co-IP. EBOV proteins fused with Falg- or Myc- tags were expressed in HEK293
 647 cells. Immunoprecipitations were performed using anti-Myc or anti-Flag antibodies.

648

649 **Figure 3** **Identification of p53 interactors by BiFC-seq.** (A) Flowchart of the
 650 BiFC-seq screening. The fluorescence-positive cells were used as templates to
 651 amplify the encoding genes of p53 binding partners. PCR products were sequenced

652 by NGS and analyzed with bioinformatics tools to identify the interactors of p53. (B)
653 Correlation analysis of p53 and their corresponding control screening groups. Data
654 are presented as means $\pm$ SD (n=3), * P <0.05, ** P <0.01, Student's t-test. (C) Number
655 of screening interactors of p53 from YC157-p53 and p53-YC157 groups was
656 presented.

657

658 **Figure 4 Analysis and validation of p53 interactors obtained by BiFC-Seq.** (A)
659 Selection of enriched biological processes for p53 screening results. (B)
660 Subnetworks constituted by p53 interactors from the BiFC-seq results, blue and red
661 dots represent interactors of p53 that reported and unreported from BioGrid database
662 individually. (C) Validation of the p53 interactions by co-IP assays, wherein p53 was
663 expressed as a Flag-tagged fusion and its partners were expressed as Myc-tagged
664 fusions. HEK293 cells were transfected with the indicated constructs,
665 immunoprecipitated and detected with anti-Flag antibody. (D) Validation of p53
666 interactions with novel screening partners by yEGFP-BiFC. p53 were cotransformed
667 with interactors into yeast component cells as indicated, fluorescence were detected
668 under microscope 72 hours after cotransformation, scale bar: 10 μ m.

669

670 **Figure 5 Library to library screening by BiFC-seq.** (A) Experimental workflow of
671 genome-wide interactome screening. (B) Venn diagram of interaction overlaps
672 among three screening groups. (C) Protein class distribution of genome-wide
673 interactors along with that of human genome.

674

675 **Figure 6 A PPI networks based on BiFC-seq.** (A) A global network constituted by
676 library to library screening results by BiFC-seq. (B-G) Various biological processes
677 are presented by integrating BiFC-seq interactions with BioGrid dataset, grey and
678 red lines represent interactions from BioGrid database and BiFC-seq results
679 individually.

680

681 **Supplementary material**

682 **Figure S 1**

683 **Expression of yEGFP in yeast cells.**

684 Intact yEGFP or YN157/YC157 fragments were transformed into yeast cells that
685 were then cultured at 30°C for 24 hours in SD-2 medium and incubated at 4°C for
686 approximately 48 hours for fluorophore maturation before observation, scale bar: 10
687 μm. WL: whole length.

688

689 **Figure S 2.**

690 **Expression of EBOV proteins fused with fluorescent fragments.**

691 N-terminal flag tagged EBOV proteins fused with fluorescent fragments as indicated,
692 expression of fused protein are detected by anti-flag antibody.

693

694 **Figure S 3.**

695 **Validation of intra EBOV PPIs by yEGFP-BiFC.**

696 To validate the intra EBOV PPIs detected by flowcytometry, EBOV proteins with
697 yEGFP fragment tags were cotransformed into yeast cell, the fluorescence of cells
698 were observed under fluorescent microscope

699

700

701 **Figure S 4.**

702 **Validation of the interaction of HDM2 and p53 by BiFC.**

703 (A) p53 and HDM2/ΔHDM2 were cotransformed into yeast cells as indicated, the
704 fluorescent yeast cells were observed under fluorescnece microscope, scale bar: 10
705 μm (B) fluorescent cells from (a) were analyzed by flow cytometry. (C,D) Various
706 peptides that used to link YN157-HDM2 and YC157-p53 as indicated were check
707 for their fluorescent cell ratio 72 hours after cotransformation by flow cytometry.
708 Data are presented as means ± SD (n=3). Error bars, standard deviation. *, P<0.05,
709 n.s., non significant.

710

711

712 **Figure S 5**

713 **The fluorescent yeast cells from p53 screening were sorted by FACS.**

714 YN157-tagged human universal library was sequentially transformed into yeast cells
715 containing p53 (YC157-p53 and p53-YC157) and control (YC157-linker and
716 liker-YC157) groups. The fluorescent cells were sorted 72 hours after transformation.
717 Approximately 2000 fluorescent cells were sorted from the P2 gate of each group.

718

719 **Figure S 6**

720 **Yeast colony PCR.**

721 (A) Fluorescent yeast colonies were used as templates to amplify the cDNA inserts.
722 A portion of the PCR products were used for electrophoresis as indicated. (B) The
723 electrophoresis results of the purified PCR products for each group. Mr: DNA2000
724 marker.

725

726

727 **Figure S 7**

728 **Number of interactions detected after each screen using p53 as bait by BiFC-seq**

729

730

731 **Figure S 8**

732 **The enriched Gene Ontology terms (biological processes) for unreported**
733 **interactors of p53.**

734

735 **Figure S 9**

736 **The fluorescent yeast cells from library to library screening were sorted by**
737 **FACS.**

738 The fluorescent yeast cells were sorted out into three groups according to their
739 fluorescence intensities.

740

741

742 **Figure S 10**

743 **Electrophoresis of stitching PCR products.**

744 Gel electrophoresis of stitching PCR products. N-library and C-library represent
745 yeast colony PCR products derivated from YN157-library and YC157-library
746 controls. N-library +C-library represents stitching PCR products from test groups.
747 Mr: DNA15000 marker.

748

749

750 **Table S 1.**

751 **The screening results of using p53 as bait by BiFC -seq.**

752 **Gene Symbol:** NCBI ENTREZ Symbols of the interactors.

753 **Group:** Specific highly expressed transcripts in each group according of p53-YC157
754 and YC157-p53. p53 indicates the overlap interactors using p53-YC157 and
755 YC157-p53 as baits.

756

757

758 **Table S 2.**

759 **Evaluation of the reliability of the p53 interactions by PRINCESS.**

760 Confidence scores were measured by a Bayesian approach that combines biological
761 evidence from multiple sources. #interaction: the number of interactions in each
762 category; LR: likelihood ratio.

763

764 **Table S3.**

765 **The Gene Ontology analysis of p53 interactors.**

766 **GO ID:** Gene Ontology ID.

767 **Description:** description of the GO term.

768 **GeneRatio:** ratio of the number of genes in a GO category to all the genes in the
769 sample.

770 **BgRatio:** ratio of the number of genes in a GO category to the background genes.

771 **Gene ID:** NCBI ENTREZ GeneID of the genes belonging to a GO category.

772 **Count:** number of all genes belonging to a GO category.

773

774 **Table S4.**

775 **The Gene Ontology analysis of unreported interactors of p53.**

776 **GO ID:** Gene Ontology ID.

777 **Description:** description of the GO term.

778 **GeneRatio:** ratio of the number of genes in a GO category to all the genes in the
779 sample.

780 **BgRatio:** ratio of the number of genes in a GO category to the background genes.

781 **Gene ID:** NCBI ENTREZ GeneID of the genes belonging to a GO category.

782

783

784 **Table S 5.**

785 **The reported interactions in three proteomic-scale Y2H screenings using p53 as**
786 **bait.**

787 Official Symbol_a/Official Symbol_b: NCBI ENTREZ Symbol of a or b gene; Gene
788 ID_a/Gene ID_b: NCBI ENTREZ GeneID of a or b gene.

789

790

791 **Table S 6.**

792 **A genome-wide BiFC-seq PPIs screening results.**

793 The reads of the same amplicon derived from YN157-library and YC157-library
794 were mapped to the genome CDS regions individually. The gene symbols of the
795 interactors from three screening groups were shown.

796

797 **Table S 7.**

798 **Evaluation of the reliability of the genome-wide PPI screening results by**
799 **PRINCESS.**

800 Confidence scores were measured by a Bayesian approach that combines biological
801 evidence from multiple sources. #interaction: the number of interactions in each
802 category; LR: likelihood ratio

803

804

805

806 **Table S 8.**

807 **The p53 interactors in a large-scale AP-MS screening using p53 as bait.**

808 Official Symbol_a/Official Symbol_b: NCBI ENTREZ Symbol of a or b gene

809 Gene ID_a/Gene ID_b: NCBI ENTREZ GeneID of a or b gene

810

811 **Table S 9.**

812 Co-expression analysis of PPI pairs identified by BiFC-seq.

813 BiFC-seq identified PPI pairs were compared with human negatome dataset, which
814 is a collection of protein and domain pairs that are unlikely engaged in direct
815 physical interactions. Those with absolute value of spearman correlation coefficient
816 (SCC) greater than 0.5 were considered co-expression pairs.

A

B

C

D

Figure 1

A

B

C

Figure 2

Figure 3

A

B

C

D

Figure 4

A

2,000 colonies

Interactors

B

C

Figure 5