CRISPR screening identifies M1AP as a new MYC regulator with a promoter-reporter system

Akiko Yamamoto [Equal first author, 1] , **Morito Kurata** [Corresp., Equal first author, 1] , **Iichiroh Onishi** [1] , **Keisuke Sugita** [1] , **Miwa Matsumura** [1] , **Sachiko Ishibashi** [1] , **Masumi Ikeda** [1] , **Kouhei Yamamoto** [1] , **Masanobu Kitagawa** [1]

[1] Department of Comprehensive Pathology, Graduate School of Medical and Dental Sciences, Tokyo Medical and Dental University, Tokyo, Japan

Corresponding Author: Morito Kurata
Email address: kurata.pth2@tmd.ac.jp

Background. *MYC* is one of the proto-oncogenes contributing to tumorigenesis in many human cancers. Although the mechanism of *MYC* regulation is still not fully understood, learning about the comprehensive mechanism controlling the transcriptional activity of *MYC* will lead to therapeutic targets. The CRISPR/Cas9 library system is a simple and powerful screening technique. This study aims to identify new transcriptional upstream activators of *MYC* using the CRISPR activation library with new promoter-reporter systems.

Methods and Results. The *MYC* promoter-reporter system was developed with a photoconvertible fluorescent protein, Dendra2, and named "p*MYC*-promoter-Dendra2." This *MYC* promoter-reporter system was designed to harbor a proximal *MYC* promoter at −3.1 kb. Both the CRISPR activation library and p*MYC*-promoter-Dendra2 were induced to HEK 293T cells, and Dendra2-positive cells, that are supposed that *MYC* should be upregulated, were collected individually by a cell sorter. Among the 169 cells collected, 12 clones were successfully established. Then, p*MYC*-promoter-Dendra2 was transfected again into these 12 clones, and 2 of 12 clones showed Dendra2 positivity. In this procedure, the cells with non-specific autofluorescence were correctly distinguished by utilizing the photoswitchable character of Dendra2. Using extracted guide DNA of these two Dendra2 positive clones, polymerase chain reaction (PCR) was performed to amplify the guide RNA (gRNA) containing region, which was introduced by the CRISPR activation library. Eventually, *PLEKHO2*, *MICU*, *MBTPS1*, and *M1AP* were identified, and these gRNAs were transfected individually into HEK 293T cells again using the CRISPR activation system. Only *M1AP* gRNA transfected cells showed Dendra2-positive fluorescence. Then, the overexpression vector for *M1AP* with a doxycycline-inducible vector confirmed that *M1AP* induced high *MYC* expression by real-time quantitative PCR and western blot. Furthermore, the dual-luciferase assay showed a significant increase of promoter activity, and *MYC* mRNA was higher in *M1AP*- overexpressing cells. *M1AP* is highly expressed in several cancers, though, a positive correlation between *M1AP* and *MYC* was observed only in human acute myeloid leukemia.

Conclusion. The present study confirmed that the experimental method using the CRISPR library technology functions effectively for the identification of molecules that activate endogenous *MYC*. This method will help elucidate the regulatory mechanism of *MYC* expression, as well as supporting further drug research against malignant tumors.

CRISPR Screening Identifies M1AP as a New MYC Regulator with a Promoter-Reporter System

5 **Akiko Yamamoto[1, 2], Morito Kurata[1, 2]**, Iichiroh Onishi[1], Keisuke Sugita[1], Miwa**

6 **Matsumura[1], Sachiko Ishibashi[1], Masumi Ikeda[1], Kohei Yamamoto[1], Masanobu Kitagawa[1]**

8 1. Department of Comprehensive Pathology, Graduate School of Medical and Dental Sciences,

9 Tokyo Medical and Dental University, Tokyo, JAPAN

10 2. These authors contributed equally to this work.

13 **Corresponding Author:

14 Morito Kurata,

15 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, JAPAN

16 E-mail address: kurata.pth2@tmd.ac.jp

18
19
20
21
22
23
24

26 **Abstract**

27 **Background.** *MYC* is one of the proto-oncogenes contributing to tumorigenesis in many human

28 cancers. Although the mechanism of *MYC* regulation is still not fully understood, learning about

29 the comprehensive mechanism controlling the transcriptional activity of *MYC* will lead to

30 therapeutic targets. The CRISPR/Cas9 library system is a simple and powerful screening

31 technique. This study aims to identify new transcriptional upstream activators of *MYC* using the

32 CRISPR activation library with new promoter-reporter systems.

33 **Methods and Results.** The *MYC* promoter-reporter system was developed with a

34 photoconvertible fluorescent protein, Dendra2, and named "p*MYC*-promoter-Dendra2." This

35 *MYC* promoter-reporter system was designed to harbor a proximal *MYC* promoter at −3.1 kb.

36 Both the CRISPR activation library and p*MYC*-promoter-Dendra2 were induced to HEK 293T

37 cells, and Dendra2-positive cells, that are supposed that *MYC* should be upregulated, were

38 collected individually by a cell sorter. Among the 169 cells collected, 12 clones were

39 successfully established. Then, p*MYC*-promoter-Dendra2 was transfected again into these 12

40 clones, and 2 of 12 clones showed Dendra2 positivity. In this procedure, the cells with non-

41 specific autofluorescence were correctly distinguished by utilizing the photoswitchable character

42 of Dendra2. Using extracted guide DNA of these two Dendra2 positive clones, polymerase chain

43 reaction (PCR) was performed to amplify the guide RNA (gRNA) containing region, which was

44 introduced by the CRISPR activation library. Eventually, *PLEKHO2*, *MICU*, *MBTPS1*,

45 and *M1AP* were identified, and these gRNAs were transfected individually into HEK 293T cells

46 again using the CRISPR activation system. Only *M1AP* gRNA transfected cells showed

47 Dendra2-positive fluorescence. Then, the overexpression vector for *M1AP* with a doxycycline-

48 inducible vector confirmed that *M1AP* induced high *MYC* expression by real-time quantitative

49   PCR and western blot. Furthermore, the dual-luciferase assay showed a significant increase of

50   promoter activity, and *MYC* mRNA was higher in *M1AP-* overexpressing cells.

51   *M1AP* is highly expressed in several cancers, though, a positive correlation between *M1AP* and

52   *MYC* was observed only in human acute myeloid leukemia.

53   **Conclusion.** The present study confirmed that the experimental method using the CRISPR

54   library technology functions effectively for the identification of molecules that activate

55   endogenous *MYC*. This method will help elucidate the regulatory mechanism of *MYC* expression,

56   as well as supporting further drug research against malignant tumors.

57

58   **Introduction**

59   *MYC* is one of the most well-known proto-oncogenes, contributing to tumorigenesis in many

60   human cancers. It can function as a transcriptional factor in both cancerous and non-cancerous

61   cells, and it is involved in cell growth, proliferation, metabolism, and transformation (*Dang,*

62   *2012*). The Wnt-β-catenin signaling pathway is one of the most well-described pathways

63   involved in the regulation of *MYC*. Due to mutations in the *adenomatous polyposis coli* (*APC*)

64   gene or abnormalities in *glycogen synthase kinase-3β*, Wnt-signal is stimulated and the excessive

65   accumulation of β-catenin occurs. Then, active β-catenin moves into the nucleus and binds to T

66   cell factor (TCF), followed by inducing *MYC* expression (Liu et al., 2002) (*Novak, Dedhar,*

67   *1999*). Mitogen-activated protein kinase (MAPK), which is a serine/threonine kinase, and its

68   subfamily extracellular signal-regulated kinase (ERK), regulate *MYC* transcriptional activity by

69   inducing phosphorylation of MYC (MAPK/ERK signaling pathway). Furthermore, as non-

70   coding RNA, microRNAs like *miR-34* (*Kong et al., 2008*) and *let-7* (*Sampson et al., 2007*)

71   suppress *MYC* expression by binding to the 3′-untranslated region of *MYC* mRNA. As for long

72 noncoding RNAs, *PVT1* regulates the expression of *MYC* by inhibiting the degradation of MYC

73 protein (*Tseng et al., 2014*). Thus, it has been reported that various regulatory mechanisms

74 function in a complex fashion regarding *MYC* expression. Though, other molecules and elements

75 like crosstalk between each pathway are not fully understood.

76   Concerning the high *MYC* expression rate of multiple malignant tumors, elucidating the

77 mechanism controlling the transcriptional activity of *MYC* is a crucial issue because it will lead

78 to therapeutic targets. Using the CRISPR/Cas9 library system, with its collection of more than

79 70,290 different guide RNAs (gRNAs), enables us to identify molecules that relate knockout or

80 activation of endogenous genes. To date, the CRISPR library has been utilized for various

81 studies. *Shalem et al.* (2014) identified genes whose loss is involved in resistance to vemurafenib

82 by introducing the knockout library into a human melanoma cell line followed by treatment with

83 *BRAF* inhibitor vemurafenib. In contrast, *Konnermann et al.* (2015) identified molecules that

84 exhibit resistance to PLX-4720 by introducing an activation library. In this way, the CRISPR

85 library has been used to identify the certain genes involved in drug sensitivity or resistance, toxin

86 resistance, and hypoxia response (*Konermann et al., 2015; Gilbert et al., 2014; Kampmann,*

87 *2018; Kurata et al., 2018; Shalem et al., 2014; Wang et al., 2014; Koike-Yusa et al., 2014;*

88 *Doench et al., 2016; Jain et al., 2016*); however, a screening of molecules that act on the

89 promoter region of a specific gene and control transcriptional activity has not yet been tested. In

90 addition, autofluorescence is highly problematic for collecting fluorescent cells in the screening.

91 Therefore, a photoconversion protein, Dendra2, was used in this system. Dendra2 is derived

92 from the octocoral *Dendronephthya sp. T.* and can change color with UV-violet light (405 nm

93 laser). This study aims to utilize the CRISPR activation library for the detection of new

94 transcriptional activators of *MYC*. For this purpose, we developed an original tool for evaluating

95 *MYC* activity and establishing screening methods.

96

97 **Materials & Methods**

98 ***MYC Promoter-Reporter System and Lentivirus Plasmids***

99   The Polymerase chain reaction (PCR) primers were designed across the proximal *MYC*

100 promoter at -3.1 kb from the MYC transcriptional start site (TSS), and PCR was performed with

101 KOD-FX (Toyobo). The primers were 5′-TTAGCTAGCGAGGGTTTTCTTTGAGGGGC-3′

102 (forward) and 5′-AACGGATCCCGGAGATTAGCGAGAGAGGA-3′ (reverse).

103   A plasmid of Dendra2 (pDendra2) was purchased from Takara-Clontech. The

104 cytomegalovirus immediate early promoter (PCMVIE) was removed with restriction enzymes

105 (Afl III and Nhe I), and the multi-cloning site was replaced with Nhe I and BamH I. Amplified

106 fragments of −3.1 kb of the *MYC* promoter region with Nhe I and BamH I restriction enzyme

107 sites were integrated into pDendra2.The new reporter system with Dendra2 and the −3.1 kb *MYC*

108 promoter region (p*MYC*-promoter-Dendra2) was used for further experiments. The double-

109 stranded DNA fragments of *M1AP* with attB sites were purchased from gBlocks® (IDT) and

110 generated into the pENTR221 vector using Gateway® BP clonase. Then they were transferred to

111 the TripZ-TRE-DEST-IRES-GFP-EF1A-rtTA vector (*Kurata et al., 2016*) using Gateway® LR

112 clonase.

113

114 ***Cell Culture and SAM Library Screening***

115   Human embryonic kidney cells (HEK 293T) were maintained in Dulbecco's modified Eagle's

116 medium (DMEM, Fuji Film) with 10% fetal bovine serum and penicillin/streptomycin with 5%

117   CO_2 at 37°C as previously used (*Kurata et al., 2016*). To transduce the lentivirus, HEK 293T

118   cells (5 x 10^5 cells/well) were seeded in a six-well cell culture plate one day before the

119   transfection. The next day, 3 μg of lentivirus plasmid was transfected with 1 μg of pMDG and 2

120   μg of pCMV using Lipofectamine™ 3000 Reagent (Invitrogen) according to previous work

121   (*Kurata et al., 2016*). Twelve hours after transfection, the medium was changed to fresh DMEM.

122   The virus supernatant was harvested at 48 h after transfection, then filtered with Millex™-HP

123   0.45 μm (Millipore). The target HEK 293T cells (5 x 10^5 cells/well) were seeded in a six-well

124   plate one day prior to transduction and were transduced with this lentivirus supernatant with 5

125   μg/ml polybrene (Sigma). The used lentiviral plasmids were lentiMPHv2 (Addgene, #89308),

126   lentidCAS-VP64_blast (Addgene, #61425), and lentiSAMv2 (Addgene, #61597). At first,

127   lentiMPHv2 was transduced into HEK 293T and treated with Hygromycin B for 2 weeks. Then,

128   lentidCAS9-VP64 was induced in MPH-expressing HEK 293T cells and treated with Blasticidin

129   for 2 weeks. The MPH-dCas9-VP64-expressing HEK 293T applied to further experiments. The

130   control and *MYC* activation gRNA (Fig. S1C) were induced in MPH-dCas9-VP64-expressing

131   HEK 293T cells. The CRISPR activation library, lentiSAMv2, was also transduced into MPH-

132   dCas9-VP64-expressing HEK 293T cells and then treated with Zeocin (300 μg/ml, Invivogen)

133   for 2 weeks.

134

135   ***Collection of Dendra2-Positive Cells and Identification of gRNAs***

136   After 2 weeks of selection with Zeocin, p*MYC*-promoter-Dendra2 was transfected into SAM

137   library lentivirus-transduced cells with Lipofectamine™ 3000 Reagent (Invitrogen). HEK 293T

138   cells transfected with p*MYC*-promoter-Dendra2 were evaluated for Dendra2-positive cells by

139   fluorescence microscope (Keyence). Cells were collected at 72 h and washed with phosphate-

140   buffered saline (PBS) and applied to the cell sorter. The Dendra2-positive cells were collected

141   into a 96-well plate with MoFlo XDP (Beckman Coulter).

142   A total of 169 cells were collected, and 12 clones were successfully established. Collected

143   clones were centrifuged and washed with PBS solution and pelletized. Lysis buffer solution with

144   proteinase K (20 mg/ml) was added to cells and left for 30 min at 60°C. After treatment at 94°C

145   for 10 min, the lysate was centrifuged for 15 min, and the supernatant was collected as genome

146   DNA (gDNA) containing fractions. Using extracted gDNA of each collected clone, PCR was

147   conducted with GoTaq® qPCR Master Mix (Promega) with the following primers: 5′-

148   GAGGGCCTATTTCCCATGAT -3′ (forward) and 5′-GAGGGCCTATTTCCCATGAT-3′

149   (reverse). PCR product containing the gRNA region was employed for TA cloning with

150   pGEM®-T Vector Systems (Promega) and gRNAs were sequenced by Sanger-sequencing.

151

152   ***Preparation of CRISPR backbone plasmid with target gRNAs and Transcriptional activation***

153   ***for MYC using the CRISPR Activation System***

154   The oligonucleotides for each target gRNA sequence were designed using the list from the

155   SAM library and integrated into the pENTR221/U6/stuffer/MS2 vector (*Kurata et al., 2018*).

156   The MPH-dCas9-VP64-expressing HEK 293T cells (1 x 10^5 cells/ml) were seeded in a 24-well

157   cell culture plate. The next day, 2.5 μg of individual gRNA vector and p*MYC*-promoter-Dendra2

158   were transfected with Lipofectamine ™ 3000 Reagent. At 48 hours after the transfection,

159   the *MYC* transcriptional activation was evaluated by Dendra2-positivity by Keyence.

160

161   ***Evaluation of MYC expression***

162   A pTRIPZ-*M1AP*-*GFP* was transduced into the HEK 293T cells. These transduced cells were

163   selected with puromycin (1 µg/ml, Invitrogen) for 2 weeks. Seventy-two hours after the

164   induction of Doxycycline (DOX) (1 µg/ml), the collected cells were employed for RNA

165   isolation.

166   RNA was isolated using an RNeasyR Mini Kit (Qiagen) according to the manufacturer's

167   instructions. Complementary DNA (cDNA) was generated from RNA with ReverTra AceR

168   qPCR RT Master Mix (Toyobo). *Beta-ACTIN* was used as an endogenous control. Using the ABI

169   Prism 7900HT (Applied Biosystems), quantitative PCR (qPCR) analysis was performed to

170   quantify the RNA level using SYBR Mix. The sequences for the PCR primers used in gene

171   expression were as follows: *MYC*: 5′-CGACTCTGAGGAGGAACAAGAA-3′ (forward) and 5′-

172   CAGCAGAAGGTGATCCAGACT-3′ (reverse), *β-ACTIN*: 5′-CACAGAGCCTCGCCTTTGCC-

173   3′ (forward) and 5′-CACAGAGCCTCGCCTTTGCC-3′ (reverse). The mRNA level of the

174   targeted gene was analyzed by comparison with the standard calibration curve. Western blot

175   analysis was performed according to standard procedures(Kurata et al., 2016). The primary

176   antibodies used were anti-c-Myc (Abcam ab32072) and anti-GAPDH (Santa Cruz sc47724),

177   each at a dilution of 1:1000. The primary antibodies were detected with horseradish peroxidase

178   (HRP)-conjugated secondary antibodies at 1:50000. Protein bands were visualized with

179   Clarity™ Western ECL Substrate (Bio-Rad).

180

181   ***Dual-Luciferase Reporter Assay***

182   The fragments of the *MYC* promoter was integrated into a pGL4 vector (Promega). A pT3.5-

183   CAG-*M1AP* was generated with pENTR221-*M1AP* and pT3.5-CAG-DEST (*Kurata et al.,*

184   *2018*) using Gateway® LR clonase. The luciferase reporter vectors, the pT3.5-CAG-*M1AP*

185   vector, and the pRL Renilla luciferase reporter vector were co-transfected in the HEK 293T cells.

186   The samples were harvested 48 and 72 hours after the transfection. Each assay was biologically

187   triplicated and repeated. Luciferase activities were measured using the dual-luciferase reporter

188   assay system (Promega) and Lumat LB9507 (Perkin Elmer). RNA was harvested at the same

189   time points and the expression of *MYC* mRNA was analyzed by qPCR.

190

191   ***M1AP and MYC expressions in databases***

192   For the purpose of comparing gene expression between cancer and non-cancerous tissue, the

193   GEPIA web server was employed (Tang et al., 2017) (http://gepia2.cancer-pku.cn/).

194   Datasets for AML samples were used in BloodSpot database, that database was used to retrieve

195   gene expression data from curated human and murine microarray and RNA-sequencing (RNA-

196   Seq) datasets (*Bagger et al., 2016)*. GSE42519,

197   GSE13159, GSE15434, GSE61804, GSE14468, and The Cancer Genome Atlas (TCGA) datasets

198   were used for analysis. In addition, following database were used; The Human Protein Atlas

199   (https://www.proteinatlas.org/) and cBioPortal (https://www.cbioportal.org) using TCGA data

200   (*Cancer Genome Atlas Research Network et al., 2013; Gao et al., 2013; Cerami et al., 2012*)

201

202   ***Statistical analysis***

203   The data were statistically analyzed using the EZR version 1.36 software. A value of $p < 0.05$

204   was considered statistically significant for all analyses.

205

206   **Results**

207   ***Development of the MYC Promoter-Reporter System for Library Screening***

208   Using PCR primers designed across the proximal *MYC* promoter at −3.1 kb (−3134 to −40)

209   from the *MYC* TSS, PCR was performed. The *MYC* promoter region was integrated upstream of

210   *Dendra2*. The pDendra2 harboring the 3.1 kb *MYC* promoter region was called "p*MYC*-

211   promoter-Dendra2" for further experiments (Fig. 1SA). A plasmid p*MYC*-promoter-Dendra2 and

212   gRNA of *MYC* and *HPRT* (as a negative control) were transfected in HEK 293T cells using the

213   CRISPR activation system. We confirmed that Dendra2 showed positivity only in the cells with

214   gRNA for *MYC* (Fig. 1SB, C). Therefore, p*MYC*-promoter-Dendra2 was utilized as a *MYC*

215   promoter-reporter analysis tool in this study.

216

217   ***MYC Transcriptional Activator Screening Using the CRISPR Activation Library***

218   To screen transcriptional activators working on the *MYC* promoter region, the CRISPR

219   activation library was employed. The SAM library lentivirus was infected into MPH-dCas9-

220   VP64-expressing HEK 293T cells (1 X 10^7 cells). After selection, p*MYC*-promoter-Dendra2 was

221   transfected into these HEK 293T cells. The multiplicity of infection (MOI) ranged from 1.4 to

222   5.2.

223   At 72 h after transfection, cells were proceeded to Dendra2-positive cell collection using flow

224   cytometry (positive ratio: 0.017%). Sorted single cells were collected into a 96-well plate and

225   incubated for 2 weeks. Among the collected 169 cells, 12 clones were successfully established

226   (Fig. 1).

227
228   ***Second Screening to Identify Candidate Genes as MYC Activators***

229   For the second screening, p*MYC*-promoter-Dendra2 was again transfected into the 12 clones.

230   Using fluorescence microscope, clones 3 and 5 were shown to be truly Dendra2 positive (Fig.

231   2A-L), and the photoswitchable attribute of Dendra2 was useful for distinguishing it from cells

232   with non-specific autofluorescence (Fig. 2S). Using extracted gDNA of clones 3 and 5, PCR was

233   performed to amplify the gRNA containing region, which was introduced by the CRISPR

234   activation library. Then, the PCR product was employed for TA cloning, followed by colony

235   PCR. By this method, four candidate genes—*PLEKHO2, MICU, MBTPS1,* and *M1AP* were

236   identified (Table 1).

237   Each candidate gRNA was transfected to the HEK 293T cells using the CRISPR activation

238   system accompanied by p*MYC*-promoter-Dendra2. Because Dendra2 was positive only in the

239   cells transfected with *M1AP* gRNA (Fig. 3A-D), *M1AP* was considered as the most promising

240   gene that would increase *MYC* transcriptional activity. Then, double-stranded DNA fragments

241   coding the full length of *M1AP* induced into Doxycycline-inducible lentiviral vector (pTRIPZ-

242   *M1AP-GFP*, Fig. 4A). The pTRIPZ-*M1AP-GFP* was introduced into HEK 293T cells. Seventy-

243   two hours after the induction *M1AP* with DOX (1 μg/ml), high *MYC* expression was observed in

244   DOX-induced cells by RT qPCR (*t*-test, $p < 0.05$, Fig. 4B).   Furthermore, in the western blots,

245   an increase of *MYC* expression was also observed in DOX-induced cells (Fig. 4C).

246   Additionally, a dual-luciferase assay conducted with a reporter containing the *MYC* promoter

247   region showed a significant increase of promoter activity in *M1AP*-overexpressing cells at both

248   48 and 72 hours (Fig. 4D and 3S). On the other hand, the *MYC* mRNA, which was isolated at the

249   same time points, showed mildly higher in *M1AP*-overexpressing cells at 48 hours but not at 72

250   hours (Fig. 4E and 3S). Taken together, although *M1AP* was confirmed to activate the *MYC*

251   promoter strongly, the *MYC* mRNA itself might be regulated also by other factors.

252

253   ***Correlation Between MYC and M1AP in the Database***

254   According to the GEPIA web server (*Tang Z et al., 2017*) (http://gepia2.cancer-pku.cn/),

255   high *M1AP* expression is observed in cervical carcinoma, acute myeloid leukemia (AML),

256   ovarian carcinoma, thyroid carcinoma, uterine endometrial carcinoma, and uterine

257   carcinosarcoma when compared to normal tissue (Fig. 4S A).   Among these carcinomas, a

258   correlation analysis using cBioPortal revealed a mild positive correlation between *MYC* and

259   *M1AP* only in AML (Fig. 5A, Fig. 4S B-F). Therefore, we focused on hematopoietic cells, and

260   *M1AP* expression was analyzed using the BloodSpot database. As a result, *M1AP* was

261   significantly higher in AML tissue compared to normal bone marrow tissue (Fig. 5B).

262   Interestingly, *M1AP* was highly expressed in AML, especially with t (8;21) (Fig. 5C).

263

264   **Discussion**
265   In this study, we created a new reporter vector that incorporates the *MYC* promoter region

266   across the proximal -3.1 kb from the TSS and photoconversion fluorescent protein to evaluate

267   the transcriptional activity of endogenous *MYC*. Ideally, a large region of the *MYC* promoter on

268   Chr.8q24 would be applied to better understand the full *MYC* promoter mechanism. However,

269   even short region on the *MYC* promoter region are still not fully understood. In designing this

270   new vector, 3.1 kb of the promoter region was applied because of the limited length of effective

271   DNA insertion into the vector. Future improved experiments on larger region will help to clarify

272   the whole mechanism of *MYC* regulation via the promoter region.

273   By applying this tool to the CRISPR library, we screened molecules that increase the

274   expression of endogenous *MYC*. According to previous reports, by introducing the CRISPR

275   library into a malignant tumor cell line and carrying out selection with a drug, molecules that

276   cause susceptibility or resistance to that drug can be identified (*Konermann et al., 2015; Gilbert*

277   *et al., 2014; Kampmann, 2018; Kurata et al., 2018; Shalem et al., 2014; Wang et al., 2014;*

278   *Koike-Yusa et al., 2014; Doench et al., 2016; Jain et al., 2016*). However, this is the first specific

279   report of a molecule that activates transcription upstream of *MYC*.

280   *MYC* is an indispensable factor for cell growth and proliferation in normal tissues, but at the

281   same time, it is one of the most well-known proto-oncogenes, with high amplification in various

282   carcinomas. Moreover, *MYC* is widely involved in tumor development and growth (*Dang, 2012*).

283   Elucidation of factors that regulate the transcriptional activity of *MYC* is crucial for establishing

284   new therapeutic strategies for tumors. Although *MYC* expression is strictly regulated by mRNA

285   and MYC protein, in 1998, *He et al.* reported that inactivation of *APC* located downstream of

286   Wnt signaling pathway results in excessive accumulation of β-catenin and promotes endogenous

287   *MYC* expression by acting on TCF present in the promoter region of *MYC*.

288   Regarding the research of the regulatory mechanism of *MYC*, *Vo et al.* (2018) transplanted

289   CRISPR/dCas9-VP160 and gRNAs that activated *MYC* transcription in the mouse brain,

290   resulting in the development of medulloblastoma. They concluded that this model is useful for

291   future drug research by suppressing *MYC* expression.

292   Because this screening was adapted to about 70,290 kinds of gRNAs, initially, multiple

293   molecules were expected to be identified; however, the screening eventually led to the

294   identification of only one molecule, *M1AP*. The reason for this is that, on performing single-cell

295   collection using flow cytometry, to avoid collecting false-positive cells, the cells with the highest

296   fluorescence intensity were collected; as a result, the number of collected cells was extremely

297   small. Furthermore, since the proliferation of the cells after the collection was poor, it is expected

298   that collected cells included cells that were highly brightened due to cell senescence. Moreover,

299   oncogene-induced cell senescence can be one explanation for why the survival rate of collected

300   single cells was very low. Oncogene-induced senescence, which is now a well-known

301   phenomenon, causes cell growth arrest where the oncogene is excessive. This was first reported

302   in *RAS* (*Serrano et al., 1997*), and the same mechanism has been reported for *MYC*, where *MYC*

303   oncogene overexpression results in inactivation of *MYC*, followed by the arrest of cell growth

304   and proliferation (*Wu et al., 2007*). Excessive *MYC* expression due to the CRISPR activation

305   system may have induced cell senescence and caused a low survival rate of sorted cells. We

306   would like to use a next-generation sequencing system and adjust the multiplicity of infection in

307   future screenings.

308   By newly established experimental line in this study, *M1AP*, which was found as a transcription

309   factor for *MYC*, was proved to function strongly as a *MYC* activator by promoter assay. On the

310   other hand, *MYC* expression was mildly elevated in the *M1AP*-overexpressing cells collected at

311   the same time as the luciferase assay. These results suggest that the *MYC* mRNA is complexly

312   controlled by multiple factors and multiple pathways.

313   In conclusion, the present study confirmed that this experimental method applying the CRISPR

314   library technology functioned effectively for the identification of the molecule that activates

315   endogenous *MYC*. This method will help elucidate the regulatory mechanism of *MYC* expression,

316   as well as facilitating further drug research against malignant tumors. However, more

317   investigation is required to establish a more accurate and precise screening technique.

318

319

320

321

322

323

324 **Acknowledgments**

325   The authors would like to thank tor the member of stem cell laboratory Tokyo Medical and

326 Dental University, for cell sorting and her technical assistance and advice.

327

328 **References**

329 Bagger FO, Sasivarevic D, Sohi SH, Laursen LG, Pundhir S, Sonderby CK, Winther O, Rapin N,

330   Porse BT. 2016. BloodSpot: a database of gene expression profiles and transcriptional

331   programs for healthy and malignant haematopoiesis. *Nucleic acids research*, 44:D917-24.

332 Cancer Genome Atlas Research Network, Weinstein JN, Collisson EA, Mills GB, Shaw KR,

333   Ozenberger BA, Ellrott K, Shmulevich I, Sander C, Stuart JM. 2013. The Cancer Genome

334   Atlas Pan-Cancer analysis project. *Nature genetics*, 45:1113-1120.

335 Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer

336   ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. 2012. The cBio

337   cancer genomics portal: an open platform for exploring multidimensional cancer genomics

338   data. *Cancer discovery*, 2:401-404.

339 Dang CV. 2012. MYC on the path to cancer. *Cell*, 149:22-35.

340 Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z,

341   Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE. 2016. Optimized sgRNA design to

342   maximize activity and minimize off-target effects of CRISPR-Cas9. *Nature biotechnology*,

343   34:184-191.

344 Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, Sun Y, Jacobsen A, Sinha R,

345   Larsson E, Cerami E, Sander C, Schultz N. 2013. Integrative analysis of complex cancer

346   genomics and clinical profiles using the cBioPortal. *Science signaling*, 6:pl1.

347   Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH, Guimaraes C,

348   Panning B, Ploegh HL, Bassik MC, Qi LS, Kampmann M, Weissman JS. 2014. Genome-Scale

349   CRISPR-Mediated Control of Gene Repression and Activation. *Cell*, 159:647-661.

350   Jain IH, Zazzeron L, Goli R, Alexa K, Schatzman-Bone S, Dhillon H, Goldberger O, Peng J,

351   Shalem O, Sanjana NE, Zhang F, Goessling W, Zapol WM, Mootha VK. 2016. Hypoxia as a

352   therapy for mitochondrial disease. *Science (New York, N.Y.)*, 352:54-61.

353   Kampmann M. 2018. CRISPRi and CRISPRa Screens in Mammalian Cells for Precision Biology

354   and Medicine. *ACS chemical biology*, 13:406-416.

355   Koike-Yusa H, Li Y, Tan EP, Velasco-Herrera Mdel C, Yusa K. 2014. Genome-wide recessive

356   genetic screening in mammalian cells with a lentiviral CRISPR-guide RNA library. *Nature

357   biotechnology*, 32:267-273.

358   Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib

359   N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F. 2015. Genome-scale transcriptional

360   activation by an engineered CRISPR-Cas9 complex. *Nature*, 517:583-588.

361   Kong YW, Cannell IG, de Moor CH, Hill K, Garside PG, Hamilton TL, Meijer HA, Dobbyn HC,

362   Stoneley M, Spriggs KA, Willis AE, Bushell M. 2008. The mechanism of micro-RNA-

363   mediated translation repression is determined by the promoter of the target gene. *Proceedings

364   of the National Academy of Sciences of the United States of America*, 105:8866-8871.

365   Kurata M, Rathe SK, Bailey NJ, Aumann NK, Jones JM, Veldhuijzen GW, Moriarity BS,

366   Largaespada DA. 2016. Using genome-wide CRISPR library screening with library resistant

367   DCK to find new sources of Ara-C drug resistance in AML. *Scientific reports*, 6:36199.

368   Kurata M, Wolf NK, Lahr WS, Weg MT, Kluesner MG, Lee S, Hui K, Shiraiwa M, Webber BR,

369   Moriarity BS. 2018. Highly multiplexed genome engineering using CRISPR/Cas9 gRNA

370   arrays. *PloS one*, 13:e0198714.

371   Kurata M, Yamamoto K, Moriarity BS, Kitagawa M, Largaespada DA. 2018. CRISPR/Cas9

372   library screening for drug target discovery. *Journal of human genetics*, 63:179-186.

373   Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y, Zhang Z, Lin X, He X. 2002. Control of

374   beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell*, 108:837-847.

375   Novak A, Dedhar S. 1999. Signaling through beta-catenin and Lef/Tcf. *Cellular and molecular*

376   *life sciences : CMLS*, 56:523-537.

377   Sampson VB, Rong NH, Han J, Yang Q, Aris V, Soteropoulos P, Petrelli NJ, Dunn SP, Krueger

378   LJ. 2007. MicroRNA let-7a down-regulates MYC and reverts MYC-induced growth in Burkitt

379   lymphoma cells. *Cancer research*, 67:9762-9770.

380   Serrano M, Lin AW, McCurrach ME, Beach D, Lowe SW. 1997. Oncogenic ras provokes

381   premature cell senescence associated with accumulation of p53 and p16INK4a. *Cell*, 88:593-

382   602.

383   Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelson T, Heckl D, Ebert BL, Root

384   DE, Doench JG, Zhang F. 2014. Genome-scale CRISPR-Cas9 knockout screening in human

385   cells. *Science (New York, N.Y.)*, 343:84-87.

386   Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. 2017. GEPIA: a web server for cancer and normal

387   gene expression profiling and interactive analyses. *Nucleic acids research*, 45:W98-W102.

388   Tseng YY, Moriarity BS, Gong W, Akiyama R, Tiwari A, Kawakami H, Ronning P, Reuland B,

389   Guenther K, Beadnell TC, Essig J, Otto GM, O'Sullivan MG, Largaespada DA, Schwertfeger

390   KL, Marahrens Y, Kawakami Y, Bagchi A. 2014. PVT1 dependence in cancer with MYC

391   copy-number increase. *Nature*, 512:82-86.

392   Vo BT, Kwon JA, Li C, Finkelstein D, Xu B, Orr BA, Sherr CJ, Roussel MF. 2018. Mouse

393   medulloblastoma driven by CRISPR activation of cellular Myc. *Scientific reports*, 8:8733-018.

394   Wang T, Wei JJ, Sabatini DM, Lander ES. 2014. Genetic screens in human cells using the

395   CRISPR-Cas9 system. *Science (New York, N.Y.)*, 343:80-84.

396   Wu CH, van Riggelen J, Yetil A, Fan AC, Bachireddy P, Felsher DW. 2007. Cellular senescence

397   is an important mechanism of tumor regression upon c-Myc inactivation. *Proceedings of the*

398   *National Academy of Sciences of the United States of America*, 104:13028-13033.

399

400   **Figure Legends**

401

402   **Figure 1. A scheme for *MYC* transcriptional activator screening with the CRISPR**
403   **activation library.**

404

405   **Figure 2. Dendra2-positive cells using fluorescence microscope.**
406   (A)-(L) correspond Clone 1-12. p*MYC*-promoter-Dendra2 was transfected into 12 clones that
407   were successfully cloned after CRISPR activation screening. Clones 3 and 5 showed Dendra2
408   positivity.

409

410   **Figure 3. Validation experiments for candidate genes as *MYC* activators.**
411   (A) *PLEKHO2* (B) *MICU3* (C) *MBTPS1* (D) *M1AP*
412   With gRNAs of the four nominated genes, p*MYC*-promoter-Dendra2 was transfected into HEK
413   293T cells by applying the CRISPR activation system. Only *M1AP*-introduced cells were
414   Dendra2 positive.

415

416   **Figure 4. *MYC* mRNA expression analyzed by real-time quantitative polymerase chain**

417   **reaction.**

418   (A) Vector map of Dox-inducible *M1AP* expression vector. (B) High *MYC* expression was

419   observed in DOX-induced cells (*t*-test, $p < 0.05$). (C) Western blotting analysis showed

420   increasing of MYC protein in DOX- induced cells. (D) Luciferase activity at 48 hours was

421   significantly higher in *M1AP*-overexpressing cells (*t*-test, $p < 0.05$). (E) *MYC* mRNA expression

422   at 48 hours was mildly higher in *M1AP*-overexpressing cells (*t*-test, $p < 0.05$).

423

424   **Figure 5. Correlation between *MYC* and *M1AP* in the database.**

425   (A) In cBioPortal using TCGA data, a correlation analysis between *MYC* and *M1AP* in AML

426   revealed a mild positive correlation.

427   (B) Using the Bloodspot database *M1AP* mRNA expression in acute myeloid leukemia (AML,

428   N=989) was significantly higher than in normal tissue (N=397). (Mann–Whitney U test, $p <$

429   0.05)

430   (C) The BloodSpot database was used to retrieve gene expression data. *M1AP* expression was

431   higher in each subtype of AML, especially with t (8;21). (Kruskal–Wallis test, *: $p < 0.01$)

432

433   **Table 1. List of Candidate Genes for *MYC* Activation**

434

435   **Supplementary Figures**

436

437   **Fig. 1S. Construct of *MYC* promoter-reporter system.**

438

439   **Fig. 2S. Photoswitchable attribute of Dendra2.**

440

441   **Fig. 3S. Luciferase activity in *M1AP*-overexpressing cells and *MYC* mRNA expression at 72**

442   **hours.** *MYC* promoters are strongly activated at 72 hours after induction of *M1AP* expression,

443   but the mRNA of *MYC* was not significantly upregulated at 72 hours.

444

445   **Fig. 4S. *M1AP* expression and correlation with *MYC* in databases.**

446   (A) Comparison of *M1AP* expression in normal and tumor tissue in GEPIA.

447   (B)-(F) Correlations between *MYC* and *M1AP* in cBioPortal.

448   (B) Cervical carcinoma. (C) Ovarian carcinoma. (D) Thyroid carcinoma. (E) Uterine endometrial

449   carcinoma. (F) Uterine carcinosarcoma.

450

List of Candidate Genes for *MYC* Activation

1   **Table 1**

Sequence of gRNA	Gene name
GGGCGGGAAATGGGTGGGGA	*PLEKHO2*
TTGTGACACTACTCCAGCCT	*MICU1*
AGGATCCCCGAAAAGGAGCA	*MBTPS1*
TCTCAGGGTATCTAGGGACT	*M1AP*

Figure 1

A scheme for *MYC* transcriptional activator screening with the CRISPR activation library.

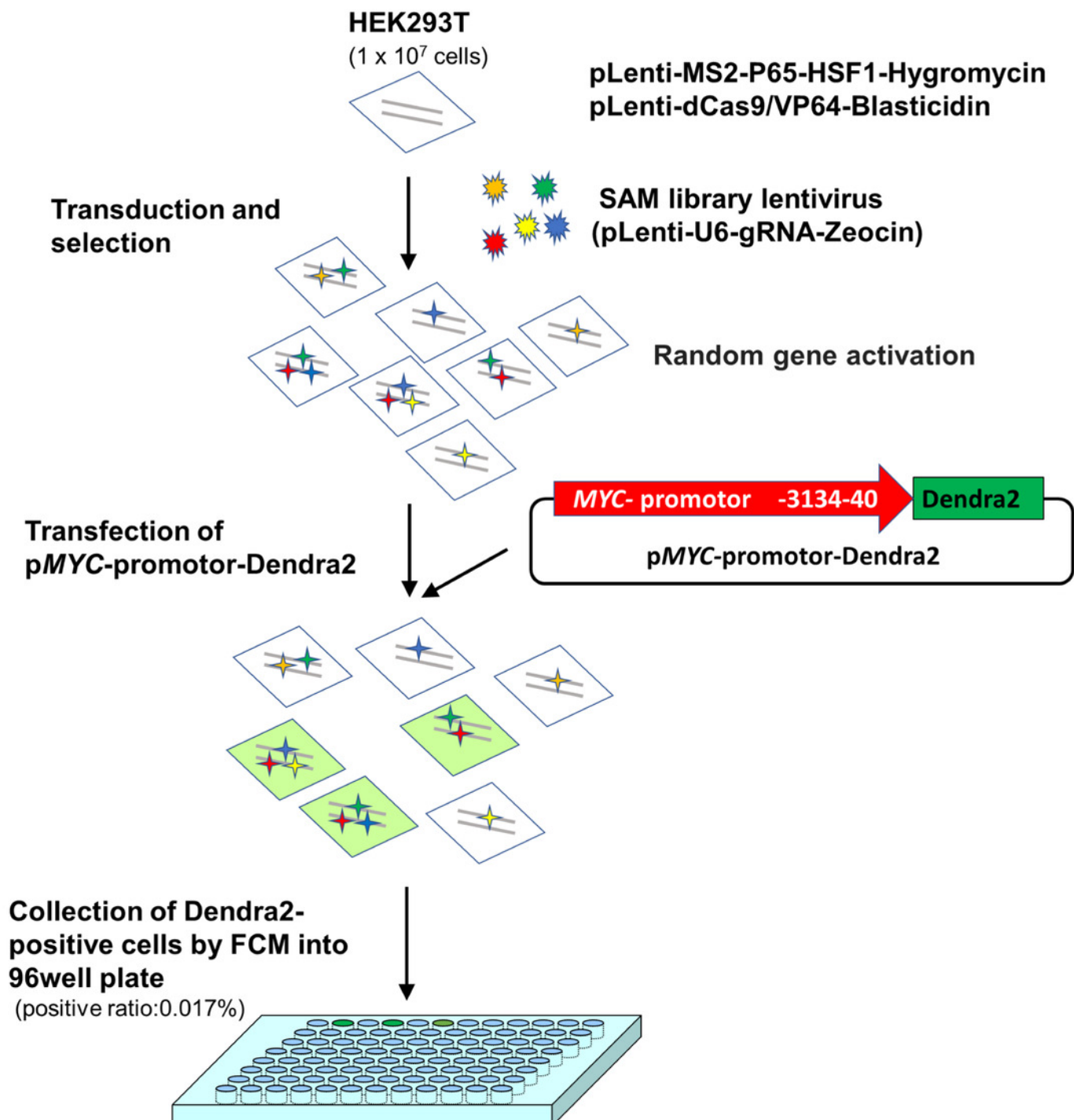



Figure 2

Dendra2-positive cells using fluorescence microscope.

(A)-(L) correspond Clone 1-12. p MYC- promoter-Dendra2 was transfected into 12 clones that were successfully cloned after CRISPR activation screening. Clones 3 and 5 showed Dendra2 positivity.

A

B

C

D

E

F

G

H

I

J

K

L

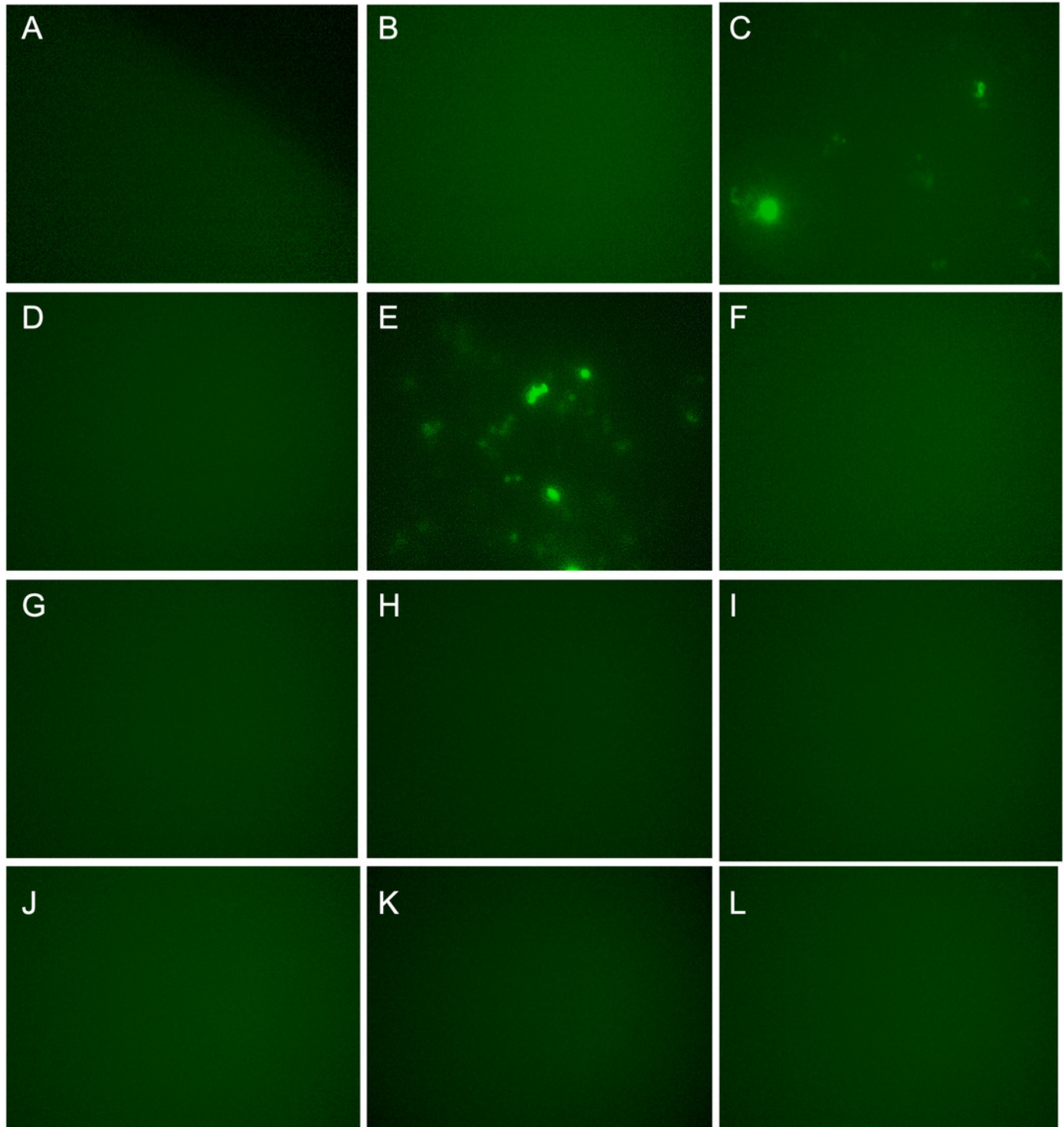



Figure 3

Validation experiments for candidate genes as *MYC* activators

(A) *PLEKHO2* (B) *MICU3* (C) *MBTPS1* (D) *M1AP* With gRNAs of the four nominated genes, p MYC- promoter-Dendra2 was transfected into HEK 293T cells by applying the CRISPR activation system. Only *M1AP*-introduced cells were Dendra2 positive.

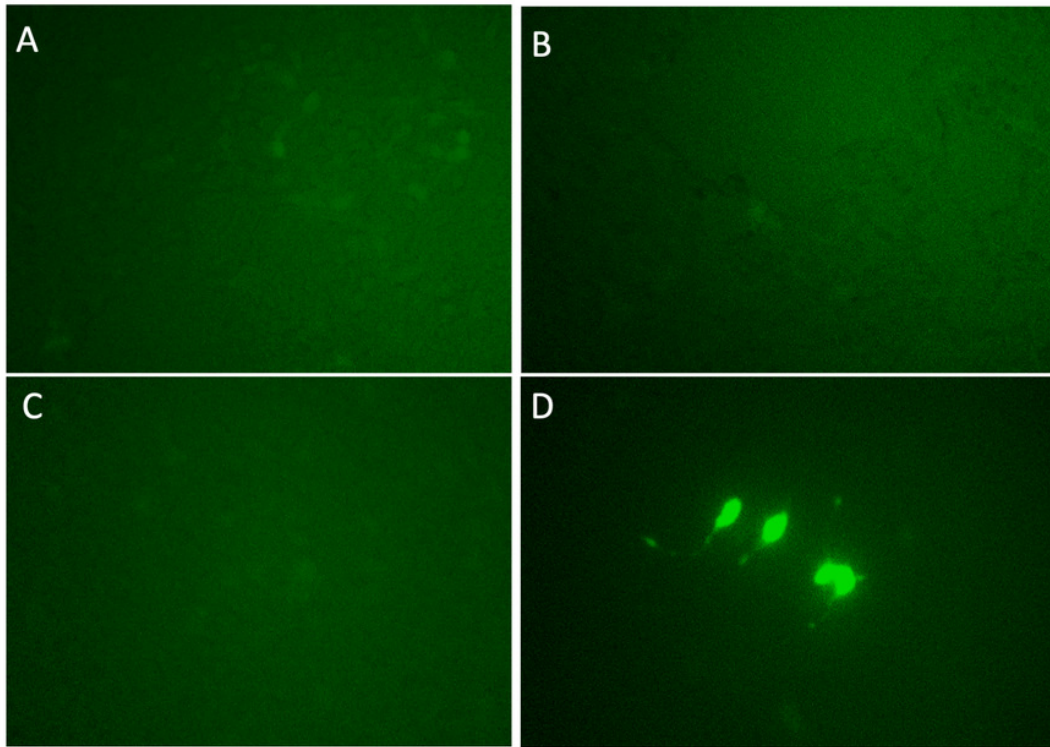
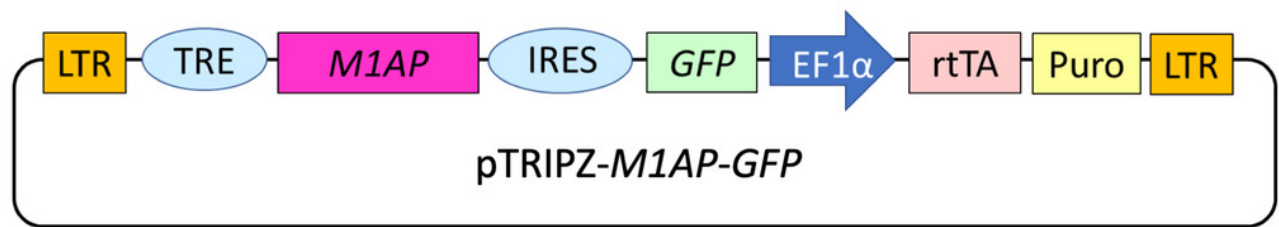
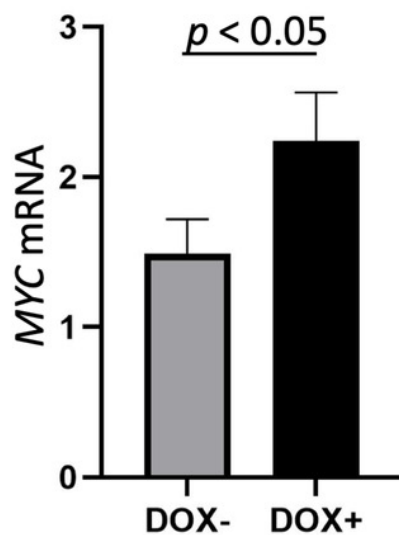
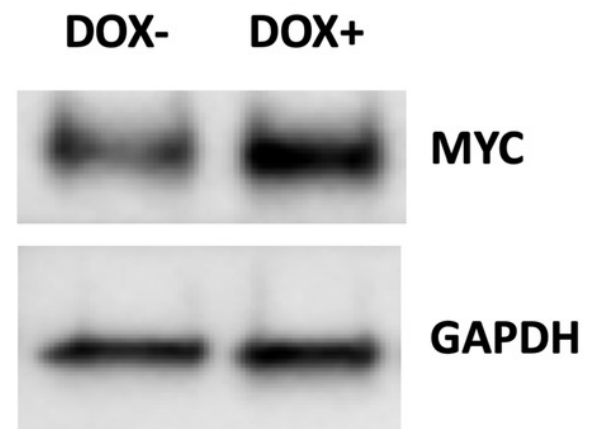
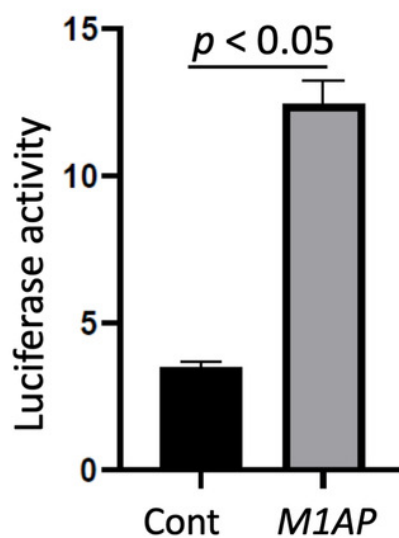



Figure 4

MYC mRNA expression analyzed by real-time quantitative polymerase chain reaction.

(A) Vector map of Dox-inducible *M1AP* expression vector. (B) High *MYC* expression was observed in DOX-induced cells (*t*-test, $p < 0.05$). (C) Western blotting analysis showed increasing of MYC protein in DOX- induced cells. (D) Luciferase activity at 48 hours was significantly higher in *M1AP*-overexpressing cells (*t*-test, $p < 0.05$). (E) *MYC* mRNA expression at 48 hours was mildly higher in *M1AP*-overexpressing cells (*t*-test, $p < 0.05$).

A

B   **C**

D   **E**

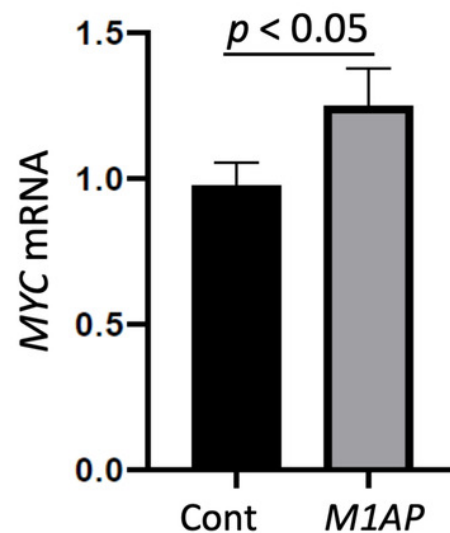



Figure 5

Correlation between *MYC* and *M1AP* in the database.

(A) In cBioPortal using TCGA data, a correlation analysis between *MYC* and *M1AP* in AML revealed a mild positive correlation. (B) Using the Bloodspot database *M1AP* mRNA expression in acute myeloid leukemia (AML, N=989) was significantly higher than in normal tissue (N=397). (Mann–Whitney U test, $p < 0.05$) (C) The BloodSpot database was used to retrieve gene expression data. *M1AP* expression was higher in each subtype of AML, especially with t (8;21). (Kruskal–Wallis test, *: $p < 0.01$)

A

Spearman: 0.24
p = 1.521 e-3

B

$p < 0.05$

Normal
(N=989)

AML
(N=397)

C

Normal
(N=989)

Complex
(N=87)

Inv (16)
(N=77)

t (15;17)
(N=87)

t (8;21)
(N=98)

t (11q23)/MLL
(N=58)