ESI mutagenesis: A one-step method for introducing point mutations into bacterial artificial chromosome transgenes.

Arnaud Rondelet[1]*, Andrei Pozniakovsky[2]*, Marit Leuschner[2], Ina Poser[2], Andrea Ssykor[2], Julian Berlitz[1], Nadine Schmidt[1], Anthony A Hyman[2], Alexander W Bird[1]

[1]Max Planck Institute of Molecular Physiology, Dortmund, Germany
[2]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

*authors contributed equally

Abstract

Bacterial artificial chromosome (BAC)-based transgenes have emerged as a powerful tool for controlled and conditional interrogation of protein function in higher eukaryotes. While homologous recombination-based recombineering methods have streamlined the efficient integration of protein tags onto BAC transgenes, generating precise point mutations has remained less efficient and time-consuming. Here we present a simplified method for inserting point mutations into BAC transgenes requiring a single recombineering step followed by antibiotic selection. This technique, which we call ESI (**E**xogenous/**S**ynthetic **I**ntronization) mutagenesis, relies on co-integration of a mutation of interest along with a selectable marker gene, the latter of which is harboured in an artificial intron adjacent to the mutation site. Cell lines generated from ESI-mutated BACs express the transgenes equivalently to the endogenous gene, and all cells efficiently splice out the synthetic intron. Thus, ESI-mutagenesis provides a robust

and effective single-step method with high precision and high efficiency for mutating BAC transgenes.

Introduction

The ability to precisely query functional hypotheses of protein function in cells requires the capacity to express rationally mutated proteins from genes under their native physiological regulation. Traditionally, transgenes expressed in higher eukaryotes are derived from cDNAs, and thus lack native cis-regulatory elements or alternative splicing isoforms, often resulting in overexpression and deregulation artefacts. This may hinder proper phenotypic functional characterization of mutated genes, as well as the determination of the precise localization and interaction partners of the protein products. While the development of sequence-specific nucleases has enabled mutation of specific genes at their endogenous loci(Gaj *et al.*, 2013; Ran *et al.*, 2013), mutations in essential genes that are lethal will prevent growth and recovery of viable cells. Additionally, deleterious mutations are prone to accumulate suppressive changes in chromosome integrity or gene expression during the procedure of selecting and expanding cells for analysis, particularly in the genomically instable cancer cell lines frequently employed.

The use of bacterial artificial chromosomes (BACs) as transgenes largely overcomes these limitations. BACs are bacterial vectors containing fragments of an eukaryotic genome that are large enough to encode an entire gene in its genomic context, including *cis*-regulatory elements. Transgenes expressed from BACs thus deliver near-physiological expression in eukaryotic cells(Kittler *et al.*,

23 2005; Bird and Hyman, 2008; Poser *et al.*, 2008; Bird *et al.*, 2011). BAC transgenes
24 modified to be RNAi-resistant allow for conditional exposure of recessive
25 mutations through selective depletion of endogenous protein(Kittler *et al.*, 2005;
26 Bird and Hyman, 2008; Ding *et al.*, 2009; Bird *et al.*, 2011; Scolz *et al.*, 2012; Zheng
27 *et al.*, 2014).
28
29 Because of their size, BACs do not lend themselves to restriction/ligation-based
30 modification techniques. Instead, homologous recombination-based
31 recombineering techniques are used to engineer BACs in *Escherichia coli* (Murphy,
32 1998; Zhang *et al.*, 1998; Yu *et al.*, 2000). Recombineering utilizes phage proteins
33 (Redα, β and γ from phage λ; or RecE and RecT from the Rac prophage) to promote
34 homologous recombination, facilitating a wide variety of DNA modifications. The
35 high efficiency of the single-step recombineering procedure to insert a protein tag
36 attached to a selectable marker gene on a DNA vector has allowed its adaptation to
37 genome-scale high-throughput pipelines(Sarov *et al.*, 2006; Poser *et al.*, 2008;
38 Sarov *et al.*, 2012; Hasse *et al.*, 2016; Sarov *et al.*, 2016). The subsequent generation
39 of transgenic cell lines based on these libraries has further been used to precisely
40 determine the cellular localisation and the quantitative interactome of more than
41 a thousand proteins (Hubner *et al.*, 2010; Hutchins *et al.*, 2010; Hein *et al.*, 2015).
42 On the contrary, current techniques to introduce point mutations in BACs still
43 require either a counterselection-based two-step procedure(Bird *et al.*, 2011;
44 Wang *et al.*, 2013) or a lower efficiency one-step procedure requiring extensive
45 PCR-screening (Lyozin *et al.*, 2014). Here we present a simple and efficient one-
46 step procedure to introduce point mutations in BAC transgenes, harnessing

47 introns to carry selectable markers, which reduces the time and cost of generating
48 mutagenized constructs.

49 Modifying intronic sequences is an attractive approach to modifying or
50 gaining gene functionality with minimal perturbation of the host gene protein
51 product, and has met use in a variety of applications such as conditional knock-out
52 and mutant constructs, or miRNA expression(Gu *et al.*, 1994; Kaulich *et al.*, 2015;
53 Wassef *et al.*, 2017). Synthetic introns have been designed based on the four core
54 splicing signals necessary for spliceosome processing: two splice sites (SS) located
55 at the 5' and 3' intron boundaries (5'SS and 3'SS, respectively), a branchpoint
56 sequence (BP) located approximately 25 nucleotides upstream of the 3'SS, and a
57 polypirimidine tract (PPT) directly upstream of the 3'SS (Figure 1A) (Lin *et al.*,
58 2003; Wang and Burge, 2008; Mercer *et al.*, 2015). Most of the sequence
59 determinants for splicing are thus located within the intron. In most eukaryotes,
60 the consensus sequences of the core splicing elements located within exons are
61 limited to a (C/A) A G sequence directly upstream of the intron and a G
62 downstream of the intron (Figure 1 A) (Zhang, 1998). Synthetic introns designed
63 from these minimal core splicing signals are indeed efficiently spliced out of an
64 exon, and have been used in multiple applications, such as making conditional
65 knockouts, or expressing shRNAs or miRNAs(Lin *et al.*, 2003; Greber and
66 Fussenegger, 2007; Lin and Ying, 2013; Seyhan, 2016; Guzzardo *et al.*, 2017).

67 Here we present a one-step BAC recombineering procedure in which point
68 mutations are introduced into a BAC along with a synthetic intron. An antibiotic-
69 resistance gene located within the intron allows the selection of positive clones
70 after the single recombineering step in bacteria, and is seamlessly spliced out of
71 mRNA sequences of the harbouring transgene in eukaryotic cells. This procedure

72 significantly increases the efficiency and decreases the screening time of BAC
73 mutagenesis.
74
75 **RESULTS**
76 **ESI-mutagenesis: a one-step BAC recombineering strategy for introduction**
77 **of point mutations.**
78 The ESI-mutagenesis strategy is outlined in Figure 1. First, a synthetic
79 intron "cassette" is amplified by PCR for use in the recombineering reaction (Fig
80 1B). This synthetic intron cassette contains the core intron-based splicing signal
81 sequences, as well as an antibiotic resistance gene. We have created multiple
82 cassettes with varying bacterial and eukaryotic selectable markers, presented in
83 Figure S1. In order to become a functional intron, the synthetic intron cassette
84 must be recombined into a site in the exon where it is directly bordered by an
85 upstream (C/A) A G sequence and a downstream G (Figure 1A, and C). While this
86 (C/A) A G | G sequence would occur every 128 nucleotides assuming random
87 distribution, similar sequences can be mutated to this sequence at amino acid
88 wobble positions during cassette amplification (see Figure S2 for an example),
89 significantly increasing the frequency of useable integration sites. Furthermore,
90 this consensus is not a strict requirement, and a variety of sequences are found in
91 cells and predicted to function(Zhang, 1998; Nguyen *et al.*, 2018; Ohno *et al.*, 2018).
92 Therefore, it is unlikely that a potential synthetic intron integration site will not be
93 found within the useable vicinity of a mutation site of interest. In addition, if the
94 desired mutation is close enough to an existing intron, the antibiotic resistance
95 cassette can be directly targeted to the endogenous intron by generating only one
96 new splice site, with the opposite homology arms targeting intronic sequences (see

97 Figure 1C, right side). This is preferable when possible, because not only does it
98 maintain more closely the native intron/exon structure, but it also avoids the
99 creation of unusually small exons (less than ~50bp), which may be prone to exon
100 skipping (Dominski and Kole, 1991).
101
102 An example of how to design ESI-mutagenesis is presented in supplementary
103 Figure S2. Once a target site for synthetic intron insertion near the mutation site is
104 identified, primers are designed to amplify the synthetic intron cassette and
105 containing 5'-extended 50 bp arms homologous to sequences directly surrounding
106 the insertion site. One or both homology arms contain the mutation(s) to be
107 integrated. The cassette is amplified by PCR and the resulting product is inserted
108 into the target BAC via recombineering (Figure 1 C). Correct recombination events
109 are selected using the antibiotic resistance encoded within the synthetic intron. A
110 protocol for the ESI mutagenesis recombineering procedure as well as generating
111 mammalian transgenic cell lines is provided in Figure S3.
112
113 Upon transfection of the modified BAC into eukaryotic cells of interest, cells that
114 have integrated the BAC in their genome may also be selected using the antibiotic
115 resistance encoded within the synthetic intron. If desired, the antibiotic resistance
116 gene may be removed by application of Cre recombinase via loxP sites located
117 within the cassette (Figure S1). Once the BAC transgene is transcribed the
118 synthetic intron is spliced out leading to the formation of a messenger RNA
119 carrying the desired mutation.
120
121 **ESI mutated BAC transgenes yield proteins with the expected size and**

122 **localization**

123 To confirm that BAC transgenes containing synthetic introns yielded
124 proteins of the correct size and localisation when expressed in human cells, we
125 started with a panel of 10 GFP-tagged BAC transgenes, and modified each using our
126 ESI-mutagenesis strategy to introduce mutations to render them resistant to
127 specific siRNA targeting sequences (Table S1). We then generated Hela cell lines
128 stably expressing either the "parental" GFP-tagged genes, or the "ESI-mutated" (i.e.
129 RNAi-resistant) GFP-tagged genes (Figure 2A). Because these mutations should
130 not change the corresponding protein sequence, this allowed us to directly
131 compare the expression and localization of the mutated transgenes to that of the
132 parental, while offering an assay for the presence and function of the introduced
133 mutation (RNAi resistance). Depending on the location of the mutation with
134 respect to an existing intron, we either targeted the synthetic intron along with the
135 desired mutation into an exon (4/10 BACs) (Figure 1 C left panel, Figure 2B), or
136 made use of an existing exon-intron border to introduce the desired mutation and
137 the antibiotic resistance (6/10 BACs) (Figure 1 C right panel, Figure 2B). All stably
138 integrated "ESI-mutated" transgenes led to the expression of proteins of sizes
139 identical to the one expressed from the corresponding parental transgenes (Figure
140 2 C). Moreover, we could not observe any differences in the cellular localisation of
141 proteins expressed from "ESI-mutated" transgenes as compared to those
142 expressed from parental transgenes (Figure 2 D). We next performed siRNA
143 experiments on six of the parental / "ESI-mutated" BAC lines pairs to confirm
144 functionality of the mutations. When the parental BAC-GFP lines were treated with
145 the siRNA against the tagged transgene, we observed a reduction in the
146 corresponding mRNA levels comparable to that observed in Hela wild type cells

147 (Figure 3 A). In contrast, ESI-mutated BAC lines did not show any reduction in the
148 corresponding mRNA levels upon siRNA treatment. We then analysed RNAi-
149 resistance on the protein level for two of the ESI-mutated BAC lines for which we
150 had the corresponding specific antibody. Both CEP135 and AURKB ESI-mutated
151 BAC clonal lines showed expression of their respective GFP-tagged transgenes at a
152 level similar to the corresponding endogenous proteins (Figure 3 B), and
153 treatment of both ESI-mutated lines with the corresponding siRNA resulted in
154 depletion of the endogenous protein while the transgene could still be detected
155 (Figure 3 B).
156
157 To further demonstrate that ESI-mutagenesis allows the introduction of functional
158 mutations in a BAC transgene, we used the procedure to mutate two EB1-
159 interacting motifs (SxIP) of the microtubule-associated protein GTSE1 to abolish
160 its interaction with the plus end tracking protein EB1(Honnappa *et al.*, 2009; Scolz
161 *et al.*, 2012). In this example, we started with a GFP-tagged GTSE1 BAC, integrated
162 the synthetic intron between the two sites to be mutated, and included the two
163 mutations (SxNN) within the 5' and 3' homology arms (Figure S2). U2OS cells were
164 generated stably expressing the BAC transgene and analysed for GTSE1 expression
165 and interaction with EB1. The ESI-mutated BAC yielded a protein of the same size
166 as GTSE1 endogenously tagged with GFP, as expected(Figure 3 C, left panel)
167 (Bendre *et al.*, 2016). However, the mutated protein was no longer efficiently
168 pulled-down by GST-EB1 when compared to the endogenous GTSE1 or GTSE1-GFP
169 (Figure 3 C), verifying that the mutation was present.
170
171

172 **A functional ESI-mutated BAC transgene is expressed in all cells of a clonal**
173 **line**

174 We showed above that ESI-mutated BACs yield proteins of the expected size
175 and the correct cellular localisation, and that mutation may efficiently confer RNA-
176 resistance or loss of interaction partner when analysed at the cell-population level.
177 We next further assessed whether an ESI-mutated transgene was functionally
178 expressed, i.e. able to sustain the cellular functions of its endogenous counterpart,
179 uniformly in all the cells of a population. This is critical for functional and
180 phenotypic analysis of mutants on the single cell level. To address this question,
181 we ESI-mutated the gene encoding Aurora A kinase (AURKA) on a GFP-tagged BAC
182 transgene to render it resistant to RNAi, transfected it into U2OS cells, and selected
183 a cell clone (U2OS AURKA-GFP siRES (ESI)) that expressed Aurora A-GFP at the
184 same level as the endogenous Aurora A (Figure 4 A, sixth lane). Treatment of this
185 line with siRNA against AURKA led to the depletion of the endogenous Aurora A,
186 while the GFP-tagged Aurora A remained expressed (Figure 4 A, fifth and sixth
187 lanes, lowest and uppermost band respectively) and correctly localised to the
188 centrosomes and the spindle (Figure 4 B and C). U2OS cells depleted of Aurora A
189 accumulate in prometaphase and show smaller bipolar spindles (Marumoto et al
190 JBC 2003, Bird and Hyman 2008) (Figure 4 D, E, and F). In contrast, cells containing
191 the RNAi-resistant Aurora A-GFP ESI maintained a normal distribution of mitotic
192 phases, similar to control-treated cells (Figure 4 D). Furthermore, we could not
193 observe any differences in the length of the metaphase spindle between U2OS
194 treated with control-siRNA and the U2OS AURKA-GFP siRES (ESI) clone treated
195 with the siRNA against AURKA (Figure 4 E and F). This result indicates that all cells
196 depleted for the endogenous Aurora A express a functional ESI mutated Aurora A-

197 GFP ESI transgene. Taken together, these data show that all cells of the U2OS
198 AURKA-GFP siRES (ESI) clone uniformly express a functional Aurora A-GFP
199 protein.
200
201 **DISCUSSION**
202 We have shown that by co-integrating a selectable marker located within an intron,
203 site-directed mutations can be quickly and efficiently introduced into BAC
204 transgenes using a one-step recombineering procedure. This recombineering
205 method is thus analogous to that for integrating protein tags: a single PCR reaction
206 to generate a cassette with homology arms followed by recombineering and
207 selection. The high efficiency (>90%) of the latter has facilitated high-throughput
208 genome-scale application of protein tagging(Poser *et al.*, 2008; Sarov *et al.*, 2012;
209 Hein *et al.*, 2015; Hasse *et al.*, 2016; Sarov *et al.*, 2016). While we have focused on
210 applications in human cell lines, this method would also be applicable to BAC or
211 fosmids commonly used as transgenes in other model organisms.
212
213 Critically, BAC transgenes ESI-mutated to be RNAi resistant yield proteins of the
214 correct size and localization that can functionally rescue the phenotypes observed
215 upon depletion of the endogenous counterpart. Our investigation of transgenic
216 lines at the single-cell level suggest that the synthetic intron is correctly spliced out
217 in all the cells of a population. Furthermore, although we formally describe the ESI-
218 mutagenesis technique here for the first time, we have previously successfully
219 employed it to generate mutations in BAC transgenes, for example, to functionally
220 analyze a microcephaly-associated mutation in the centriolar protein CPAP at the
221 single cell level in HeLa cells and neural progenitor cells(Zheng *et al.*, 2014; Gabriel

222 *et al.*, 2016).

223

224 We cannot exclude that a small portion of synthetic introns within individual cells
225 are misspliced, yet without detectable consequences in our functional assays.
226 Indeed, although the core essential sequences for defining introns and splice sites
227 are located primarily within the intron and a few nucleotides into the exons, the
228 precise regulation of efficient splicing is a more complex process and impacted by
229 additional sequence information, including that within exons (Chasin, 2007).
230 Nevertheless, mis-spliced mRNAs are probably minimal and degraded by
231 nonsense-mediated decay.

232

233 One of our original concerns in this study was that the new synthetic intron or the
234 gene within may impact expression of the transgene to an unuseful extent. While
235 introns have been well characterized to generally increase gene expression(Nott
236 *et al.*, 2003; Shaul, 2017), nested intronic genes found natively in eukaryotic
237 genomes generally impart a negative effect on the "host" gene expression(Yu *et al.*,
238 2005; Kumar, 2009). Thus it is possible that expression of a transgene could be
239 impacted by synthetic intron-mutagenesis. While we did not observe gross
240 changes in gene expression in the cell lines generated here, this will likely be highly
241 context-dependent (gene, location of intron, organism, tissue, etc) and we have
242 observed a case where a nested antibiotic resistance gene in a synthetic intron
243 directly integrated into the genome via CAS9 has decreased host gene expression
244 (unpublished results). We have thus also designed a version of the synthetic intron
245 cassette for BAC recombineering in which the antibiotic resistance only carries a
246 bacterial promoter, eliminating the necessity to recombine out with Cre. In the end,

247 because BACs integrate randomly into the genome by standard transfection
248 procedures, clonal variations in expression (+/- 2-fold) in transfected cell lines,
249 likely dependent on position effect and copy number, are anyway observed. Such
250 variations could compensate for synthetic intron effects, and it is thus good
251 practice, prior to phenotypic functional analysis, to select for clonal cell lines
252 expressing BAC transgenes at endogenous levels.
253
254 The general strategy outlined here of harbouring a selectable marker gene in an
255 artificial intron adjacent a mutation could in principle be used to increase the
256 efficiency of isolating point mutations via Cas9-induced homologous
257 recombination directly into the genome. This would particularly be useful in hard-
258 to-transfect cell lines where extensive screening may be necessary without a
259 selection step. In our preliminary studies applying this ESI-mutagenesis strategy
260 to direct Cas9-mediated human genome modification, we are indeed able to
261 recover modified cell lines at high frequency. We also found, however, more readily
262 observable impacts on gene expression from the synthetic intron when located in
263 the native gene, which could be alleviated by removing the antibiotic resistance
264 gene after Cre application. While further development of the synthetic intron
265 cassette design/strategy towards direct genome modification holds promise for
266 high efficiency selection of Cas9-mediated site-directed mutation, sensitive
267 applications (i.e. anything to be used in a clinical setting) would not be advisable
268 due to potential changes induced by the synthetic intron.
269
270 The use of BAC transgenes remains a powerful genetic tool for exploration of
271 protein function. The ESI mutagenesis method described here allows for

272 straightforward and rapid site-directed mutagenesis of BAC transgenes to further
273 improve their ease of use.
274
275 **Methods and Materials**
276 **Bacterial Artificial Chromosomes and recombineering**
277 All BACs used and generated, including modified sequences, are presented in
278 Table S1. LAP- and NLAP-tagged "parental" BACs were generated by
279 recombineering as described in Poser et al 2008. "ESI-mutated" BACs were
280 generated from "parental" BACs following the procedure outlined in Figure S3.
281
282 **Antibodies**
283 Mouse anti-GFP (11 814 460 001, Roche), rabbit anti-GTSE1 (custom generated;
284 described in Scolz *et al.*, 2012), and rabbit anti-Aurora kinase B (ab2254, Abcam)
285 were used for Western blots. Mouse anti-α-tubulin (DM1α, Sigma-Aldrich), rabbit
286 anti-Cep135 (custom generated, described in Bird and Hyman, 2008), goat anti-
287 GFP (MPI Dresden; described in Poser *et al.*, 2008), and goat anti-Aurora A (sc-
288 27883, Santa Cruz) were used for both Western Blot and immunofluorescence.
289 Donkey anti-goat Alexa488 (Jackson Immunoresearch; 705 545 147), donkey
290 anti-mouse Alexa594 (Bethyl; A90-337D4), and donkey anti-rabbit Alexa650
291 (Bethyl; A120-208D5) were used in immunofluorescence. Donkey anti-goat HRP
292 (Santa Cruz; SC-2020), sheep anti-mouse HRP (Amersham; NXA931-1ml), and
293 donkey anti-rabbit HRP (Amersham; NXA934-1ml) were used for Western Blots.
294
295 **Cell Lines and Cell Culture**

296 U2OS, HeLa cells, and derivatives were grown at 5% CO2 and 37 °C in DMEM (PAN
297 Biotech) supplemented with 10% filtered Fetal Bovine Serum (FBS, Gibco), 2 mM
298 L-glutamine (PAN Biotech), 0.1 mg/mL streptomycin and 100 U/ml penicillin
299 (PAN Biotech, Pen/Strep mix). BACs were transfected into cells using the
300 Effectene kit (Qiagen) following the manufacturer's instructions, and stable cell
301 lines were selected on antibiotics.
302
303 **RNAi**
304 siRNA sequences used are presented in Table S1. siRNAs were purchased from
305 Thermofisher/Ambion. Unless specified otherwise, control siRNA was Silencer
306 negative control No. 2 (Thermofisher). A reverse transfection approach was used
307 to deplete AurkA and CEP135, using 50 nM and 40 nM siRNA, respectively. The
308 transfection procedure for a 24-well plates with a final volume of 500 µL is
309 described. If required, the procedure was scaled up. The siRNA and 2.5 µL
310 Oligofectamine (Invitrogen) were prepared separately in a final volume of 50 and
311 15 µL OptiMEM (Invitrogen), respectively. After 5 min incubation at RT, the siRNA
312 and transfection reagent were mixed and further incubated 20 min at RT.
313 Meanwhile, HeLa cells, U2OS cells or their derivatives were seeded at the desired
314 confluency into a 24 well plates containing or not a coverslip. The transfection mix
315 was finally added onto the cells and the volume adjusted to 500 µL with pre-
316 warmed medium. The medium was changed 7 to 8 hours post-transfection. For
317 CEP135 depletion, cells were harvested for Western Blot 48 hours post-
318 transfection. For AurkA depletion, 12 hours post-transfection cells were
319 synchronized with 3.3 µM Nocodazole for 18 hours, and harvested for Western
320 Blot. Alternatively, cells seeded on coverslips were fixed and processed for

321 microscopy 30 hours post-transfection. A forward transfection approach was
322 used to deplete AurkB using 50nM siRNA. Day before transfection, cells were
323 seeded at ~40 % confluency in a 3.5 cm dish. The siRNA and 5 µL of Oligofectamine
324 (Invitrogen) were prepared separately into 200 µL OPTIMEM. After 5 min
325 incubation at RT siRNA and Oligofectamine solutions were mixed and further
326 incubated 20 min at RT. Meanwhile, cells were washed twice with PBS and 1.6 mL
327 OPTIMEM were added on top of them. The transfection mix was then added onto
328 the cells. Cells were harvested for Western Blot 48 hours post-transfection. In
329 experiments measuring mRNA levels in HeLa cells, a control scrambled RNA
330 (customed, Ambion, sense: 5'-UUCUCCGAACGUGUCACGUtt-3') was used, and
331 siRNA was transfected using Interferin (Polyplus Transfection) and 100nM siRNA
332 per well of a 6-well plate in forward or reverse transfection (140,000 cells).
333 Duration of the depletion is indicated for each gene in Table S1.
334
335 **Western Blot**
336 Cells were lysed on ice for 10 to 15 min in RIPA buffer or cell lysis buffer (50 mM
337 Na2HPO4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM EGTA, 1.5 mM
338 MgCl2, 50 mM HEPES pH 7.2, 1 mM DTE) both supplemented with 2X protease
339 inhibitor mix (SERVA). Debris were eliminated by centrifugation at 4 °C at max-
340 speed in a benchtop centrifuge. Samples were prepared in Laemmli buffer and run
341 on SDS-PAGE before transfer onto a nitrocellulose membrane. 5% milk (Carl Roth)
342 in PBS Tween20 0.1% (SERVA electrophoresis) was used to block the membrane
343 before incubation with the primary and secondary antibodies diluted in the same
344 blocking solution. Membranes were incubated with ECL reagent (GE Healthcare)
345 before development onto Hyperfilm (Amersham) or imaging on the ChemiDoc™

346 MP imaging system (BioRad). Digital data were obtained by scanning Hyperfilms
347 or using the ImageLab software (BioRad). Fiji was used to adjust levels, generate
348 8-bit tiffs, and crop images.
349
350 **qPCR analysis**
351 cDNA was isolated from 250,000 cells per condition using the Cells-to-cDNA II Kit
352 (Ambion AM1722) following the manufacturer's two-step RT-PCR protocol.
353 qPCR was performed with SYBR Green (ThermoFisher) on a Mx3000 qPCR
354 Machine (Stratagene). Primers are indicated in Table S1.
355
356 **Pull-downs**
357 Cells from a confluent 15 cm dish were lysed for 15 min on ice in 2mL cell lysis
358 buffer (50 mM Na2HPO4, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1 mM
359 EGTA, 1.5 mM MgCl2, 50 mM HEPES pH 7.2, 1 mM DTE) supplemented with 2X
360 protease inhibitor mix (SERVA) and 1mM DTE (15min on ice). Debris were
361 removed by centrifugation at max-speed at 4 °C in a benchtop centrifuge. Total
362 protein concentration was measured by Bradford. Equivalent amounts of cell
363 lysate were incubated 20 hours at 4 °C on GST or GST-EB1 beads. Beads were
364 washed thrice with 1 mL cell lysis buffer supplemented with 2X Protease inhibitor.
365 Proteins were eluted with 45 µL Laemmli 2x.
366
367 **Immunofluorescence**

368 For most experiments, cells were seeded onto coverslips and fixed in ice-cold
369 methanol for 12 min at -20 °C. Cells were blocked at RT with 5% BSA (Sigma)

370 dissolved into PBS. For images of Aurora B-GFP and CHMP4B-GFP cell lines, cells
371 were fixed in 3% PFA supplemented with 5 mM EGTA, 1mM $MgCl_2$ and 2%
372 sucrose, they were washed, then permeabilized using PBS/0.1% Triton X-100 and
373 eventually blocked with PBS/0.2% Fish Skin Gelatin (Sigma, Germany). Primary
374 antibodies were diluted into 5% BSA in PBS and incubated onto coverslips for 1 h
375 at 37 °C in a wet chamber. The same procedure was applied for secondary antibodies.
376 Coverslips were mounted using Prolong Gold antifade mounting with DAPI (Molecular
377 Probes and Thermo Fisher Scientific). Cells were washed thrice with PBS between
378 every step.

379 **Microscopy**

380 Immunofluoresence images for the AurkA experiments were acquired using a
381 spinning-disk confocal system (Marianas, Intelligent Imaging Innovation, 3i) built
382 on an Axio Observer Z1 microscope (Zeiss) with an ORCA-Flash 4.0 camera
383 (Hamamatsu Photonics). Images were taken with a 63× 1.4 NA Plan-Apochromat
384 oil objective (ZEISS). The spindle size was measured using the CEP135 signal in
385 the Slidebook software (Marianas, Intelligent Imaging Innovation, 3i, V5.5 or
386 later).
387 Experiments assessing the localization of the GFP transgenes, were all performed
388 on live cells at the exception of the AurkB and CHMP4B transgenes. Live cells were
389 imaged in CO2-independent visualization media (Gibco). Images were acquired
390 using a DeltaVision imaging system (GE Healthcare) with an sCMOS camera (PCO
391 Edge 5.5) and a 60x 1.42NA Plan Apo N UIS2 objective (Olympus).
392

393 **BIBLIOGRAPHY**

394 Bendre, S., Rondelet, A., Hall, C., Schmidt, N., Lin, Y.-C., Brouhard, G. J., and Bird, A.
395 W. (2016). GTSE1 tunes microtubule stability for chromosome alignment and
396 segregation by inhibiting the microtubule depolymerase MCAK. The Journal of
397 Cell Biology *215*, 631–647.

398 Bird, A. W., and Hyman, A. A. (2008). Building a spindle of the correct length in
399 human cells requires the interaction between TPX2 and Aurora A. The Journal of
400 Cell Biology *182*, 289–300.

401 Bird, A. W., Erler, A., Fu, J., Hériché, J.-K., Maresca, M., Zhang, Y., Hyman, A. A., and
402 Stewart, A. F. (2011). High-efficiency counterselection recombineering for site-
403 directed mutagenesis in bacterial artificial chromosomes. Nat Meth *9*, 103–109.

404 Chasin, L. A. (2007). Searching for splicing motifs. Adv. Exp. Med. Biol. *623*, 85–
405 106.

406 Ding, L. *et al.* (2009). A genome-scale RNAi screen for Oct4 modulators defines a
407 role of the Paf1 complex for embryonic stem cell identity. Cell Stem Cell *4*, 403–
408 415.

409 Dominski, Z., and Kole, R. (1991). Selection of splice sites in pre-mRNAs with
410 short internal exons. Molecular and Cellular Biology *11*, 6075–6083.

411 Gabriel, E. *et al.* (2016). CPAP promotes timely cilium disassembly to maintain
412 neural progenitor pool. Embo J *35*, 803–819.

413 Gaj, T., Gersbach, C. A., and Barbas, C. F., III (2013). ZFN, TALEN, and CRISPR/Cas-
414 based methods for genome engineering. Trends in Biotechnology, 1–9.

415 Greber, D., and Fussenegger, M. (2007). Multi-gene engineering: Simultaneous
416 expression and knockdown of six genes off a single platform. Biotechnol. Bioeng.
417 *96*, 821–834.

418 Gu, H., Marth, J. D., Orban, P. C., Mossmann, H., and Rajewsky, K. (1994). Deletion
419 of a DNA polymerase beta gene segment in T cells using cell type-specific gene
420 targeting. Science *265*, 103–106.

421 Guzzardo, P. M., Rashkova, C., Santos, dos, R. L., Tehrani, R., Collin, P., and
422 Bürckstümmer, T. (2017). A small cassette enables conditional gene inactivation
423 by CRISPR/Cas9. Nature Publishing Group *7*, 16770.

424 Hasse, S., Hyman, A. A., and Sarov, M. (2016). TransgeneOmics â€“ A transgenic
425 platform for protein localization based function exploration. Methods *96*, 69–74.

426 Hein, M. Y. *et al.* (2015). A Human Interactome in Three Quantitative Dimensions
427 Organized by Stoichiometries and Abundances. Cell *163*, 712–723.

428 Honnappa, S. *et al.* (2009). An EB1-Binding Motif Acts as a Microtubule Tip
429 Localization Signal. Cell *138*, 366–376.

430 Hubner, N. C., Bird, A. W., Cox, J., Splettstoesser, B., Bandilla, P., Poser, I., Hyman,
431 A., and Mann, M. (2010). Quantitative proteomics combined with BAC
432 TransgeneOmics reveals in vivo protein interactions. The Journal of Cell Biology
433 *189*, 739–754.

434 Hutchins, J. R. A. *et al.* (2010). Systematic Analysis of Human Protein Complexes
435 Identifies Chromosome Segregation Proteins. Science *328*, 593–599.

436 Kaulich, M., Lee, Y. J., Lonn, P., Springer, A. D., Meade, B. R., and Dowdy, S. F.
437 (2015). Efficient CRISPR-rAAV engineering of endogenous genes to study protein
438 function by allele-specific RNAi. Nucleic Acids Research.

439 Kittler, R., Pelletier, L., Ma, C., Poser, I., Fischer, S., Hyman, A. A., and Buchholz, F.
440 (2005). RNA interference rescue by bacterial artificial chromosome transgenesis
441 in mammalian tissue culture cells. Proceedings of the National Academy of
442 Sciences *102*, 2396–2401.

443 Kumar, A. (2009). An overview of nested genes in eukaryotic genomes.
444 Eukaryotic Cell *8*, 1321–1329.

445 Lin, S.-L., and Ying, S.-Y. (2013). Gene silencing in vitro and in vivo using intronic
446 microRNAs. Methods Mol Biol *936*, 209–229.

447 Lin, S.-L., Chang, D., Wu, D.-Y., and Ying, S.-Y. (2003). A novel RNA splicing-
448 mediated gene silencing mechanism potential for genome evolution. Biochem
449 Biophys Res Commun *310*, 754–760.

450 Lyozin, G. T., Bressloff, P. C., Kumar, A., Kosaka, Y., Demarest, B. L., Yost, H. J.,
451 Kuehn, M. R., and Brunelli, L. (2014). Isolation of rare recombinants without
452 using selectable markers for one-step seamless BAC mutagenesis. Nat Meth *11*,
453 966–970.

454 Mercer, T. R., Clark, M. B., Andersen, S. B., Brunck, M. E., Haerty, W., Crawford, J.,
455 Taft, R. J., Nielsen, L. K., Dinger, M. E., and Mattick, J. S. (2015). Genome-wide
456 discovery of human splicing branchpoints. Genome Research *25*, 290–303.

457 Murphy, K. C. (1998). Use of bacteriophage lambda recombination functions to
458 promote gene replacement in Escherichia coli. Journal of Bacteriology *180*,
459 2063–2071.

460 Nguyen, H., Das, U., Wang, B., and Xie, J. (2018). The matrices and constraints of
461 GT/AG splice sites of more than 1000 species/lineages. Gene *660*, 92–101.

462 Nott, A., Meislin, S. H., and Moore, M. J. (2003). A quantitative analysis of intron
463 effects on mammalian gene expression. Rna *9*, 607–617.

464 Ohno, K., Takeda, J.-I., and Masuda, A. (2018). Rules and tools to predict the
465 splicing effects of exonic and intronic mutations. Wiley Interdiscip Rev RNA *9*.

466 Poser, I. *et al.* (2008). BAC TransgeneOmics: a high-throughput method for
467 exploration of protein function in mammals. Nat Meth *5*, 409–415.

468 Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013).
469 Genome engineering using the CRISPR-Cas9 system. Nat Protoc *8*, 2281–2308.

470 Sarov, M. *et al.* (2012). A genome-scale resource for in vivo tag-based protein
471 function exploration in C. elegans. Cell *150*, 855–866.

472 Sarov, M. *et al.* (2016). A genome-wide resource for the analysis of protein
473 localisation in Drosophila. eLife *5*, e12068.

474 Sarov, M., Schneider, S., Pozniakovski, A., Roguev, A., Ernst, S., Zhang, Y., Hyman,
475 A. A., and Stewart, A. F. (2006). A recombineering pipeline for functional
476 genomics applied to Caenorhabditis elegans. Nat Meth *3*, 839–844.

477 Scolz, M. *et al.* (2012). GTSE1 is a microtubule plus-end tracking protein that
478 regulates EB1-dependent cell migration. PLoS ONE *7*, e51259.

479 Seyhan, A. A. (2016). A multiplexed miRNA and transgene expression platform
480 for simultaneous repression and expression of protein coding sequences. Mol
481 Biosyst *12*, 295–312.

482 Shaul, O. (2017). How introns enhance gene expression. Int. J. Biochem. Cell Biol.
483 *91*, 145–155.

484 Wang, H., Bian, X., Xia, L., Ding, X., Muller, R., Zhang, Y., Fu, J., and Stewart, A. F.
485 (2013). Improved seamless mutagenesis by recombineering using ccdB for
486 counterselection. Nucleic Acids Research.

487 Wang, Z., and Burge, C. B. (2008). Splicing regulation: from a parts list of
488 regulatory elements to an integrated splicing code. Rna *14*, 802–813.

489 Wassef, M., Luscan, A., Battistella, A., Le Corre, S., Li, H., Wallace, M. R., Vidaud, M.,
490 and Margueron, R. (2017). Versatile and precise gene-targeting strategies for
491 functional studies in mammalian cell lines. Methods, 1–10.

492 Yu, D., Ellis, H. M., Lee, E. C., Jenkins, N. A., Copeland, N. G., and Court, D. L. (2000).
493 An efficient recombination system for chromosome engineering in Escherichia
494 coli. Proceedings of the National Academy of Sciences *97*, 5978–5983.

495 Yu, P., Ma, D., and Xu, M. (2005). Nested genes in the human genome. Genomics
496 *86*, 414–422.

497 Zhang, M. Q. (1998). Statistical features of human exons and their flanking
498 regions. Human Molecular Genetics *7*, 919–932.

499 Zhang, Y., Buchholz, F., Muyrers, J. P., and Stewart, A. F. (1998). A new logic for
500 DNA engineering using recombination in Escherichia coli. Nat Genet *20*, 123–
501 128.

502 Zheng, X. *et al.* (2014). Conserved TCP domain of Sas-4/CPAP is essential for
503 pericentriolar material tethering during centrosome biogenesis. Proc. Natl. Acad.
504 Sci. U.S.a. *111*, E354–E363.

505

506

507

508 **FIGURE LEGENDS**

509 Figure 1: ESI-mutagenesis: a one step recombineering procedure to introduce
510 point mutations in BAC

511 **(A)** Scheme showing the organisation of an eukaryotic gene with exons (in blue)
512 and their flanking introns. Introns are spliced out from pre-messenger RNAs to
513 produce mature messenger RNAs (mRNAs) containing only exons. The core
514 splicing signals constitute the minimal information required for the splicing of an
515 intron, and include a 5' and 3' splicing signal sequences (5'SS and 3'SS,
516 respectively), a branch point (B.P.), and a poly-pyrimidine tract (P.P.T). The
517 position of these signals relative to the exon / intron border is indicated with the
518 corresponding nucleotides shown in pink (in introns) or light orange (in exons).
519 ESI mutagenesis is a one step recombineering procedure that relies on the
520 introduction into a BAC of a synthetic intron coding a selectable marker along with
521 the mutation of interest, thereby allowing for the easy selection of correct
522 recombinants. After transcription of the transgene in eukaryotic cells, the
523 synthetic intron is spliced out to produce a mutated mRNA. **(B)** The synthetic
524 intron constitutes of the intronic core splicing signal (in pink), an antibiotic
525 resistance cassette (Atb.^{R}) under the control of a dual eukaryotic/prokaryotic
526 promoter (in green), and two loxP sites flanking the antibiotic resistance cassette.
527 In order to serve as a template for BAC recombineering, the synthetic intron is
528 amplified by PCR with primers containing ~50 bp long homology arms to the
529 targeted sequence (in blue), the mutation(s) to introduce (in orange) and the

530 exonic part of the 5'SS and 3'SS (in light orange). **(C)** Scheme showing the
531 introduction of a point mutation into a BAC by ESI-mutagenesis. Depending on the
532 position of the sequence to mutate, the point mutation can be inserted along with
533 the synthetic intron into an exon (Left panel), or, if the mutation is located in the
534 proximity of a pre-existing intron, the part of the synthetic intron encoding the
535 antibiotic resistance can be targeted into this intron (Right panel). The synthetic
536 intron is spliced out during RNA maturation and only the desired mutation is
537 present in the mRNA.
538
539 Figure 2: ESI-mutated BAC transgenes yield proteins of the right size and
540 localisation
541 **(A)** GFP tagged BAC transgenes (Parental BAC, P) were mutated by ESI
542 mutagenesis to be RNAi resistant (ESI mutated BAC, ESI) and transfected into
543 HeLa cells. **(B)** Depending on the position of the sequence to mutate, the synthetic
544 intron was either targeted into an exon (left panel, light blue label in B, C and D)
545 or into the neighbouring intron by making use of pre-existing 5'SS, and / or 3'SS
546 (right panel, light brown label in B, C and D). Names of the mutated transgene are
547 indicated in the panel corresponding to their mutation strategy. The ESI mutated
548 AURKA transgene is analysed in Figure 4. The GTSE1 transgene was ESI mutated
549 at its SxIP motifs and is analysed in Figure 3. **(C)** GFP-tagged BAC transgenes ESI
550 mutated to be RNAi resistant yield proteins of the expected size. Immunoblotting
551 on cell-lysate of pools of HeLa cells transfected with either parental BACs (P) or
552 ESI-mutated BACs (ESI). GFP antibody was used as a probe. Transgenes names are
553 indicated at the top. **(D)** GFP-tagged transgenes ESI mutated to be RNAi resistant
554 show the same cellular localization as their parental GFP-tagged transgenes. Still

555 images of live cell imaging on HeLa stably transfected with the parental or the ESI-
556 mutated BAC. CHMP4B and AurkB -GFP transgenes were detected by
557 immunofluorescence with anti-GFP antibody. Because of differences in expression
558 level within cell pool, each picture was acquired and scaled independently of the
559 others. Scale Bar 10 µm.
560
561 Figure 3: ESI-mutated BAC transgenes show the expected phenotype
562 **(A)** BAC transgenes ESI mutated to be RNAi resistant show RNAi resistance at the
563 mRNA level. HeLa wild type, parental BAC lines, and ESI-mutated BAC lines were
564 treated with control-siRNA or transgene-specific siRNAs. Levels of mRNA
565 corresponding to the transgenes were determined by qPCR and normalised to
566 GAPDH mRNA levels. For each cell lines, the % of mRNA remaining in the
567 transgene-siRNA treated cells as compared to the control-siRNA treated cells are
568 presented (N=2 exp.). A panel of 3 transgenes mutated by targeting the synthetic
569 intron into an exon (light blue label) and 3 transgenes mutated by targeting the
570 synthetic intron into a pre-existing intron (light brown label) were analysed. **(B)**
571 BAC transgenes ESI mutated to be RNAi resistant show RNAi resistance at the
572 protein level. HeLa and clonal HeLa lines expressing ESI-mutated CEP135 or
573 AurkB GFP-tagged BAC transgenes were treated with control, CEP135 or AurkB -
574 siRNAs. Levels of AurkB, CEP135 and Tubulin were monitored by immunoblot
575 using specific antibodies. **(C)** Mutation of GTES1 SxIP motifs into SxNN by ESI
576 mutagenesis disrupts the interaction of a GTSE1-GFP BAC transgene with EB1. Cell
577 lysates from U2OS, U2OS expressing an ESI-mutated GTSE1-GFP SxNN BAC
578 transgene, and U2OS expressing an endogenously GFP-tagged GTSE1 (GTSE1-GFP
579 Cas9), were used in pull-downs with GST or GST-EB1 as baits. Pull-downs inputs

580 and outputs were probed by immunoblot using GTSE1 antibody. GST and GST-EB1
581 fusion were visualized by Coomassie Blue.
582
583 Figure 4: An AURKA-GFP BAC transgene ESI-mutated to carry RNAi resistance
584 rescues endogenous AURKA depletion**. (A)** An AURKA-GFP transgene ESI-
585 mutated to be RNAi resistant is expressed at endogenous level and is RNAi
586 resistant. Immunoblotting on U2OS and a U2OS clone expressing the ESI-mutated
587 AURKA-GFP BAC transgene (U2OS AURKA-GFP siRES (ESI)), both treated with
588 control- or AURKA- siRNA and blocked in mitosis. Antibodies against AURKA and
589 Tubulin were used as probe. **(B,C)** The ESI-mutated AURKA-GFP transgene shows
590 the correct cellular localization. Immunofluorescence on U2OS and U2OS AURKA-
591 GFP siRES (ESI) treated with control- and/or AURKA- siRNA. AURKA antibody (B)
592 or AURKA and Tubulin antibodies (C) were used as a probe. The ESI mutated
593 AURKA-GFP transgene rescues the mitotic arrest **(D)** and spindle collapse **(E,F)**
594 observed after endogenous AURKA depletion. The percentage of cells in each
595 mitotic phase is presented (D) for U2OS and U2OS AURKA-GFP siRES (ESI) clone
596 treated with control- or AURKA- siRNA (N= 3 exp.). U2OS and U2OS AURKA-GFP
597 siRES (ESI) clone were treated with control- or AURKA- siRNA and stained by
598 immunofluorescence using DAPI and an antibody against CEP135 (E). The pole-
599 to-pole distance (μm) in metaphase cells is presented (F). In AURKA-depleted
600 U2OS, spindle length was measured in prometaphase cells with nearly aligned
601 chromosomes. The number of mitotic cells used in each condition (n) and
602 statistically significant differences to U2OS treated with control-siRNA are
603 indicated (Kruskal Wallis, followed by paired Wilcoxon test). All scale bar 5 μm.
604

605 Figure S1: Available synthetic intron cassettes.

606 The synthetic intron cassettes generated in this study are depicted (A to D). The
607 resistance genes and the promoters (bacterial and eukaryotic) controlling their
608 expression are shown. When present, the loxP sites allow the removal of the
609 antibiotic resistances by the CRE recombinase. The position of the core intron-
610 based splicing signal sequences (intronic splicing donor dite, branch point (B.P.),
611 polypirimidine tract (PPT), intronic splicing acceptor site) is shown. The plasmid
612 from which each cassette can be amplified is indicated. Numbers indicate the
613 nucleotide positions within the cassette.

614

615 Figure S2: ESI mutagenesis of the GTSE1 SxIP motifs

616 **(A)** Scheme representing the sequence of exon 9 of GTSE1 containing the two SxIP
617 motifs (in orange) and a potential insertion site for the synthetic intron (in pink).
618 Distance (in bp) of the potential insertion site for the synthetic intron to the
619 border with introns 8-9 and 9-10 are presented. **(B)** Primers used to amplify the
620 synthetic intron are presented. They are composed of (i) homology arms to GTSE1,
621 (ii) the SxIP mutations (in orange) and (iii) a synonymous mutation (in orange)
622 allowing the potential insertion site (in pink) to fit the CAGG consensus. **(C)** The
623 amplified synthetic intron is recombined onto the GTSE1 BAC. The synthetic
624 intron is spliced out from the pre-mRNA to yield a mRNA coding GTSE1 mutated
625 at both SxIP motifs.

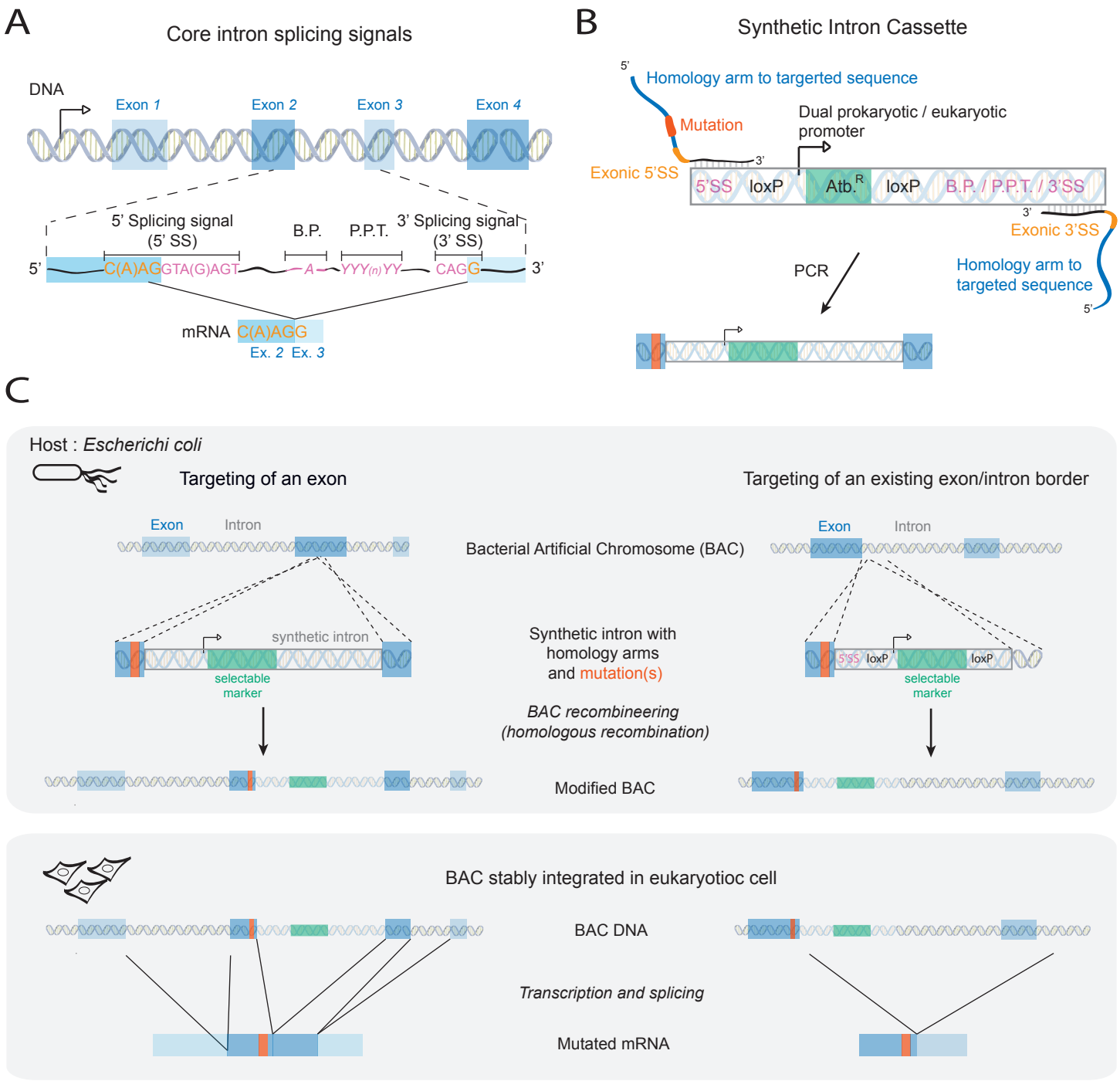
626

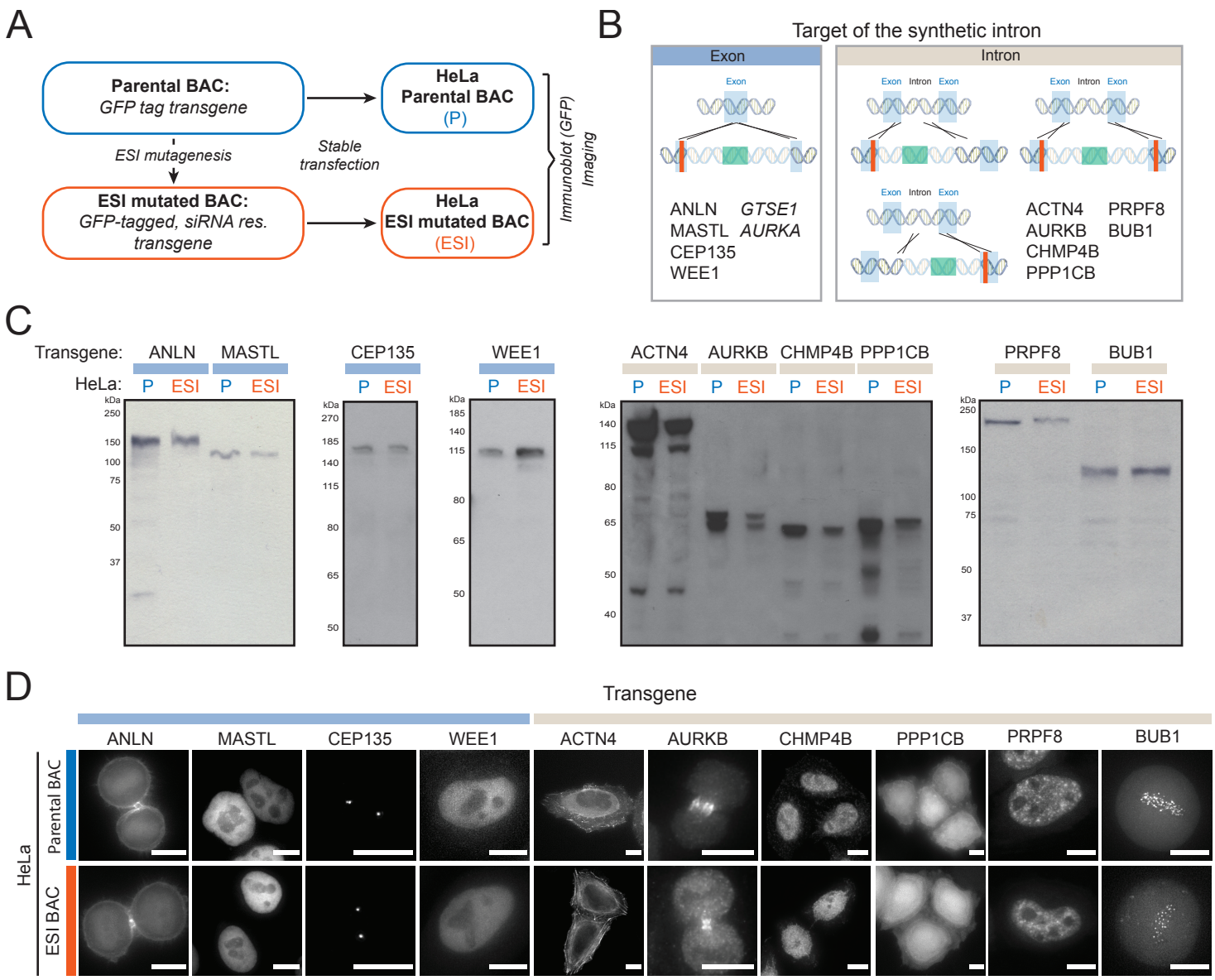
627 Figure S3: The ESI mutagenesis procedure at a glance.

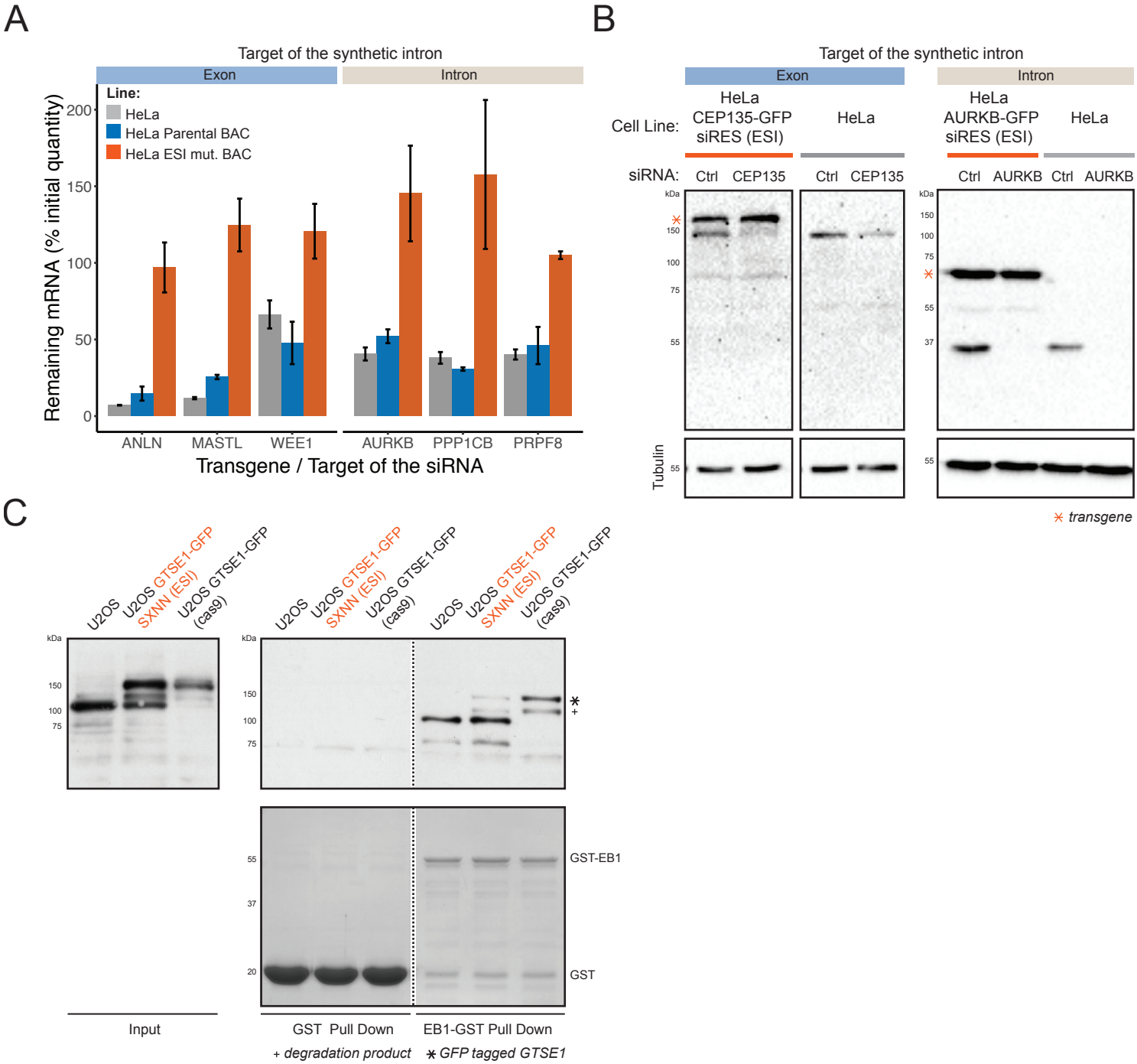
628 (A) Scheme presenting the BAC recombineering protocol. Important parameters
629 to pay attention to are indicated. (B) Scheme showing the procedure to obtain the

630 ESI mutagenesis BAC recombineering template. Three PCRs (see Figure 1B and S1
631 for primer design) are performed to amplify the synthetic intron and pooled for
632 subsequent PCR purification. The resulting product is used as recombineering
633 template. (C) The ESI mutated BAC is transfected into eukaryotic cells. Stably
634 transfected cells are selected using the antibiotic resistance coded by the synthetic
635 intron. If the BAC carries a fluorescent protein tagged transgene, the antibiotic
636 resistance associated to the tag should be use for selection.
637
638

Figure 1


A
Core intron splicing signals
DNA
Exon 1
Exon 2
Exon 3
Exon 4
5' Splicing signal
(5' SS)
B.P.
P.P.T.
3' Splicing signal
(3' SS)
5'
C(A)AGGTA(G)AGT
A
YYY(n)YY
CAGG
3'
mRNA C(A)AGG
Ex. 2 Ex. 3
B
Synthetic Intron Cassette
5'
Homology arm to targerted sequence
Mutation
Dual prokaryotic / eukaryotic
promoter
Exonic 5'SS
3'
5'SS
loxP
Atb.R
loxP
B.P. / P.P.T. / 3'SS
3'
Exonic 3'SS
PCR
Homology arm to
targeted sequence
5'
C
Host : Escherichi coli
Targeting of an exon
Targeting of an existing exon/intron border
Exon
Intron
Bacterial Artificial Chromosome (BAC)
Exon
Intron
synthetic intron
selectable
marker
Synthetic intron with
homology arms
and mutation(s)
5'SS
loxP
loxP
selectable
marker
BAC recombineering
(homologous recombination)
Modified BAC
BAC stably integrated in eukaryotioc cell
BAC DNA
Transcription and splicing
Mutated mRNA


A
Parental BAC:
GFP tag transgene
HeLa
Parental BAC
(P)
ESI mutagenesis
Stable
transfection
ESI mutated BAC:
GFP-tagged, siRNA res.
transgene
HeLa
ESI mutated BAC
(ESI)
Immunoblot (GFP)
Imaging
B
Target of the synthetic intron
Exon
Intron
Exon
Exon Intron Exon
ANLN
MASTL
CEP135
WEE1
GTSE1
AURKA
ACTN4
AURKB
CHMP4B
PPP1CB
PRPF8
BUB1
C
Transgene:
HeLa:
P
ESI
kDa
D
Transgene
HeLa
Parental BAC
ESI BAC


A
Target of the synthetic intron
Exon
Intron
Line:
HeLa
HeLa Parental BAC
HeLa ESI mut. BAC
Remaining mRNA (% initial quantity)
200
150
100
50
0
ANLN
MASTL
WEE1
AURKB
PPP1CB
PRPF8
Transgene / Target of the siRNA
B
Target of the synthetic intron
Exon
Intron
Cell Line:
HeLa CEP135-GFP siRES (ESI)
HeLa
HeLa AURKB-GFP siRES (ESI)
HeLa
siRNA:
Ctrl
CEP135
Ctrl
CEP135
Ctrl
AURKB
Ctrl
AURKB
kDa
150
100
75
55
37
Tubulin
55
transgene
C
U2OS
U2OS GTSE1-GFP SXNN (ESI)
U2OS GTSE1-GFP (cas9)
kDa
150
100
75
55
37
20
GST-EB1
GST
Input
GST Pull Down
EB1-GST Pull Down
+ degradation product
GFP tagged GTSE1

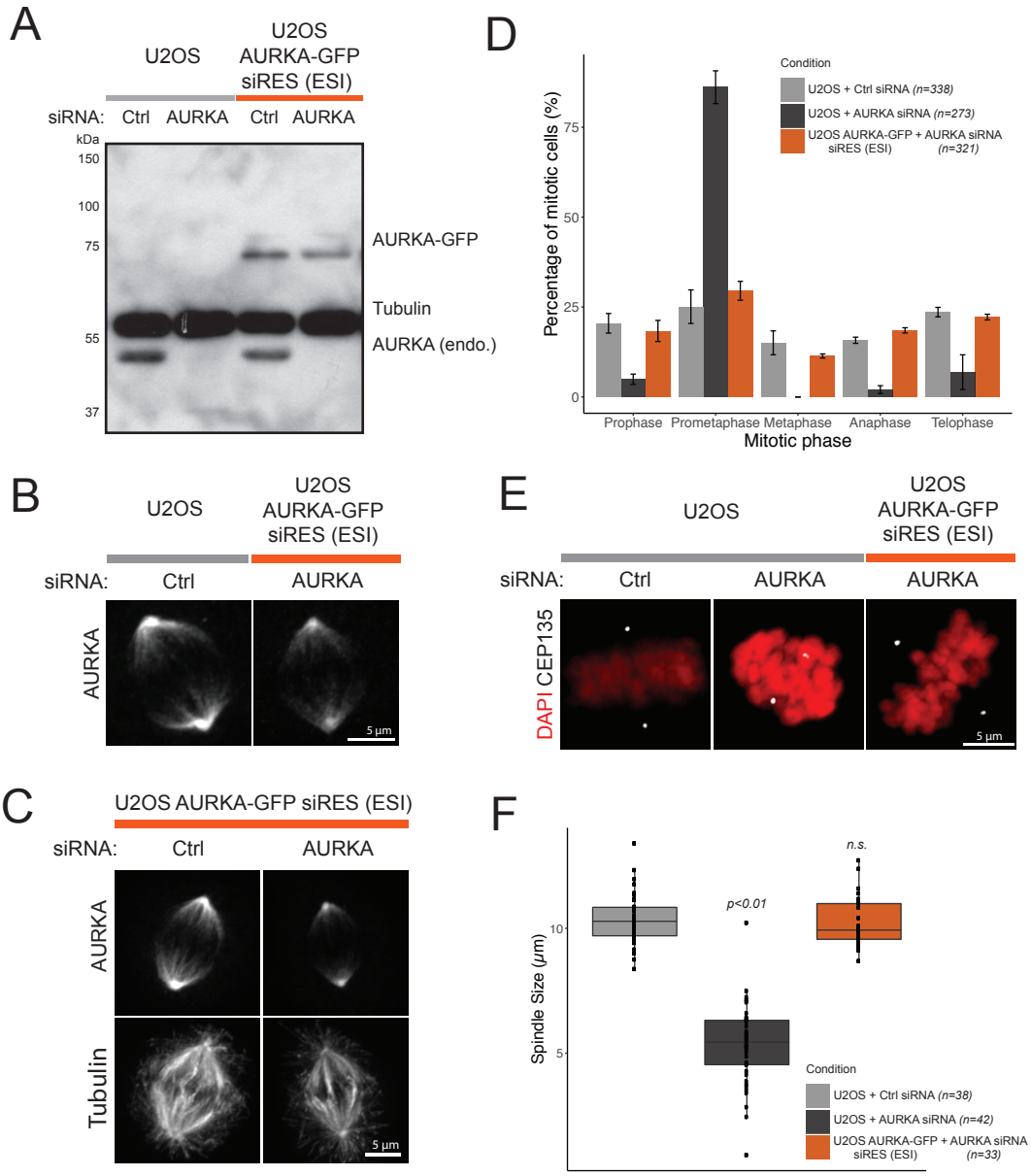
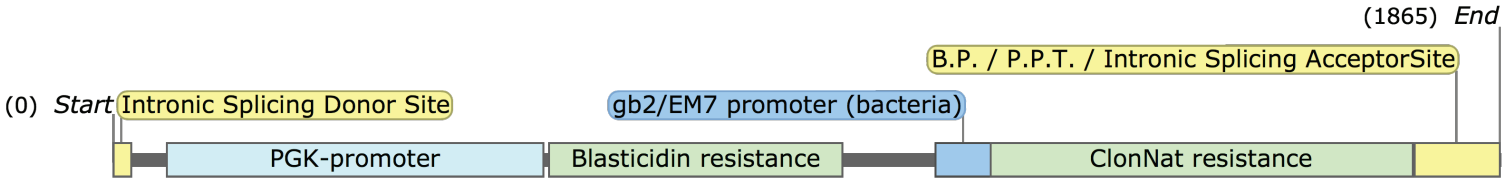

A
U2OS
U2OS AURKA-GFP siRES (ESI)
siRNA: Ctrl AURKA Ctrl AURKA
kDa
150
100
75
55
37
AURKA-GFP
Tubulin
AURKA (endo.)
B
U2OS
U2OS AURKA-GFP siRES (ESI)
siRNA: Ctrl AURKA
AURKA
5 µm
C
U2OS AURKA-GFP siRES (ESI)
siRNA: Ctrl AURKA
AURKA
Tubulin
5 µm
D
Condition
U2OS + Ctrl siRNA (n=338)
U2OS + AURKA siRNA (n=273)
U2OS AURKA-GFP + AURKA siRNA siRES (ESI) (n=321)
Percentage of mitotic cells (%)
0
25
50
75
Prophase
Prometaphase
Metaphase
Anaphase
Telophase
Mitotic phase
E
U2OS
U2OS AURKA-GFP siRES (ESI)
siRNA: Ctrl AURKA AURKA
DAPI CEP135
5 µm
F
Spindle Size (µm)
5
10
p<0.01
n.s.
Condition
U2OS + Ctrl siRNA (n=38)
U2OS + AURKA siRNA (n=42)
U2OS AURKA-GFP + AURKA siRNA siRES (ESI) (n=33)

Figure S1

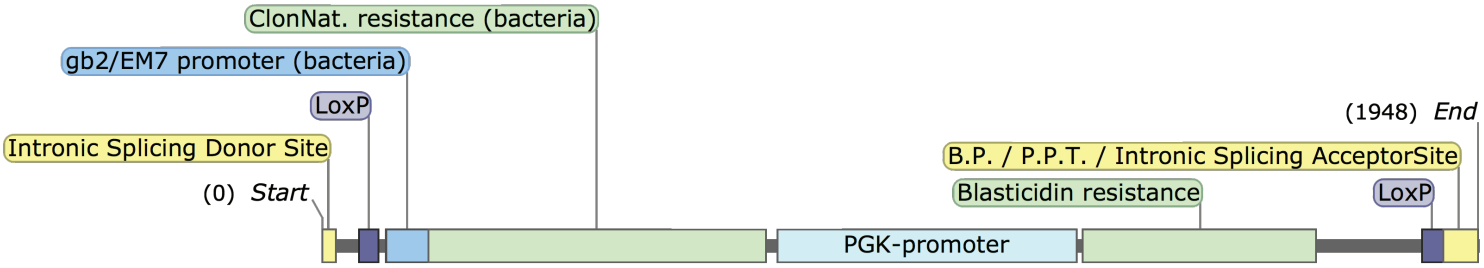
A

Synthetic intron *(plasmid TH0496)* **with :**
Blasticidin resistance under eukaryotic promoter
ClonNat resistance under bacterial promoter



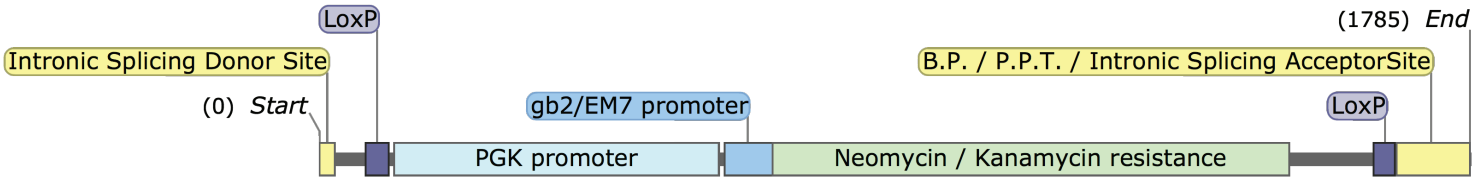
B

Synthetic intron *(plasmid TH0823)* **with :**
Blasticidin resistance under eukaryotic promoter
ClonNat resistance under bacterial promoter
loxP sites for removal of antibiotic resistances by the CRE recombianse



C

Synthetic intron *(plasmid TH0480)* **with :**
Neocyn / Kanamycin resistance under dual eukaryotic / bacterial promoter
loxP sites for removal of antibiotic resistances by the CRE recombianse



D

Synthetic intron *(plasmid TH0821)* **with :**
Zeocin resistance under dual eukaryotic / bacterial promoter
loxP sites for removal of antibiotic resistances by the CRE recombianse

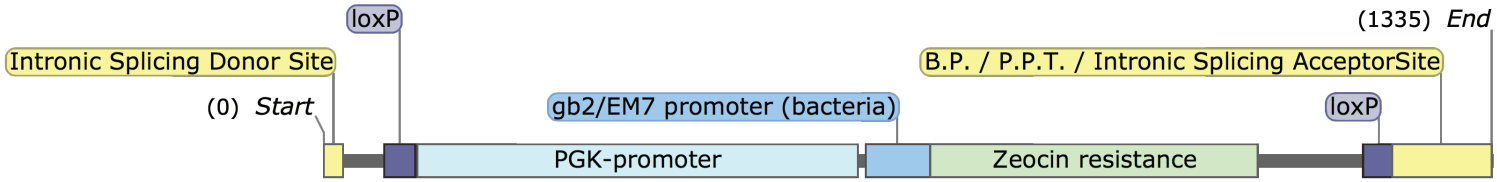


Figure S2

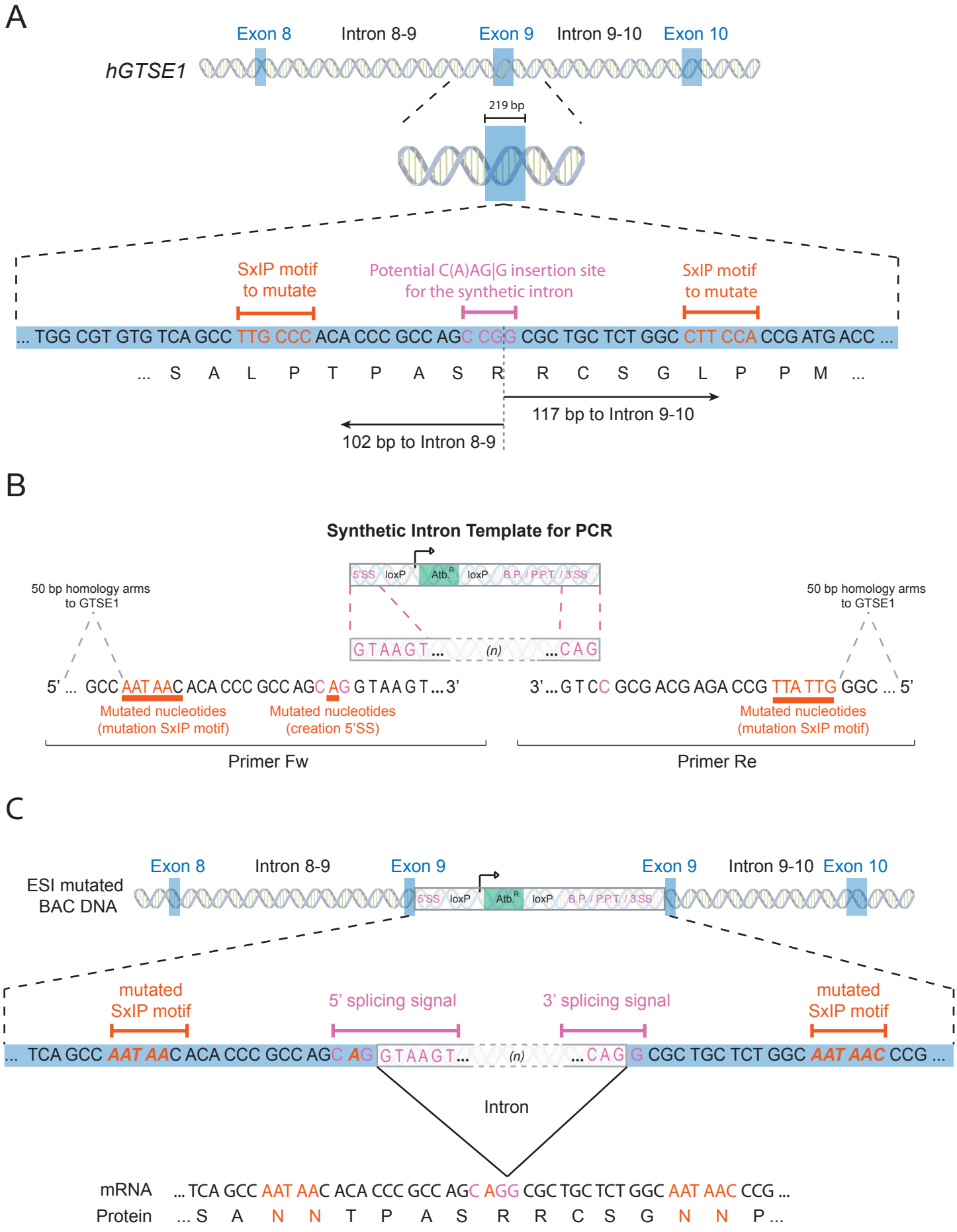
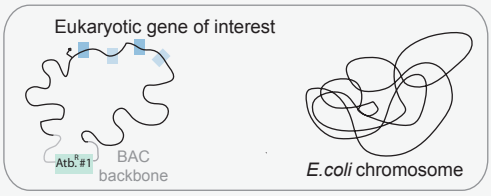

A
Exon 8
Intron 8-9
Exon 9
Intron 9-10
Exon 10
hGTSE1
219 bp
SxIP motif to mutate
Potential C(A)AG|G insertion site for the synthetic intron
SxIP motif to mutate
... TGG CGT GTG TCA GCC TTG CCC ACA CCC GCC AGC CGG CGC TGC TCT GGC CTT CCA CCG ATG ACC ...
... S A L P T P A S R R C S G L P P M ...
117 bp to Intron 9-10
102 bp to Intron 8-9
B
Synthetic Intron Template for PCR
5'SS
loxP
Atb.R
loxP
B.P. / P.P.T. / 3'SS
GTAAGT...
(n)
...CAG
50 bp homology arms to GTSE1
50 bp homology arms to GTSE1
5' ... GCC AAT AAC ACA CCC GCC AGC AG GTAAGT...3'
3'...GTCC GCG ACG AGA CCG TTA TTG GGC ... 5'
Mutated nucleotides (mutation SxIP motif)
Mutated nucleotides (creation 5'SS)
Mutated nucleotides (mutation SxIP motif)
Primer Fw
Primer Re
C
ESI mutated BAC DNA
Exon 8
Intron 8-9
Exon 9
Exon 9
Intron 9-10
Exon 10
mutated SxIP motif
5' splicing signal
3' splicing signal
mutated SxIP motif
... TCA GCC AAT AAC ACA CCC GCC AGC AG GTAAGT... (n) ...CAG G CGC TGC TCT GGC AAT AAC CCG ...
Intron
mRNA ... TCA GCC AAT AAC ACA CCC GCC AGC AGG CGC TGC TCT GGC AAT AAC CCG ...
Protein ... S A N N T P A S R R C S G N N P ...

Figure S3

A

E.coli strains available from BAC collections

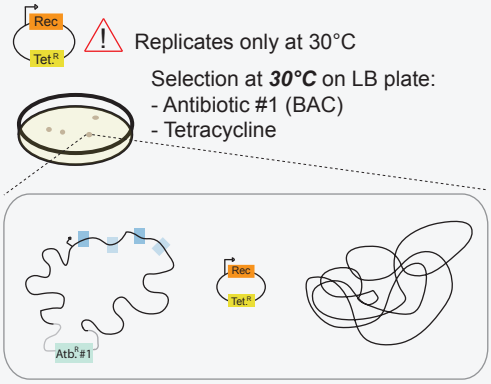


NB: the procedure can be applied on BAC carrying a tagged transgene.
In this case Antibiotic #1 corresponds to the antibiotic resistance associated to the tag.

Day 0 to day 2

Electroporation of plasmid coding Rec α,β,γ

Day 0: Pre-culture of BAC clone (LB, Atb. #1)
Day 1: 1) Liquid culture from pre-culture (LB, Atb. #1)
2) Electroporation of Rec α,β,γ plasmid

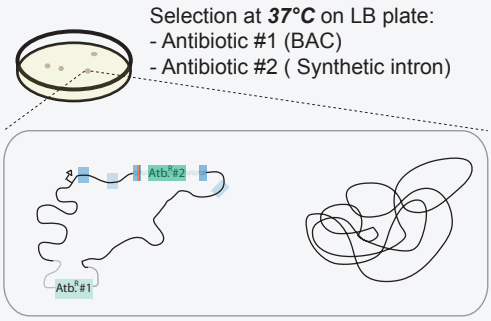


Day 2 to day 4

Recombineering of the synthetic intron

Day 2: Pre-culture transformed clone (LB, Atb. #1, Tet)
Day 3: 1) Liquid culture from pre-culture (LB, Atb. #1, Tet.)
2) Induction of Rec α,β,γ for 70 min. (at *OD 0.3-0.4*)
3) Electroporation of synthetic intron **(see B)**
4) Incubation 50 min at 37°C

Pre-culture and culture at 30°C. Keep bacteria on ice while preparing them for induction and electroporation.

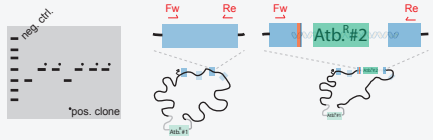


Day 4

Colony PCR on Antibiotic #2 resistant clones

1) Use colonies for PCR and restreak on (LB, Atb. #2) plates
2) Sequencing of positive PCR product

Synthetic intron integration results in a shift in the PCR product size shift



Day 5 and 6

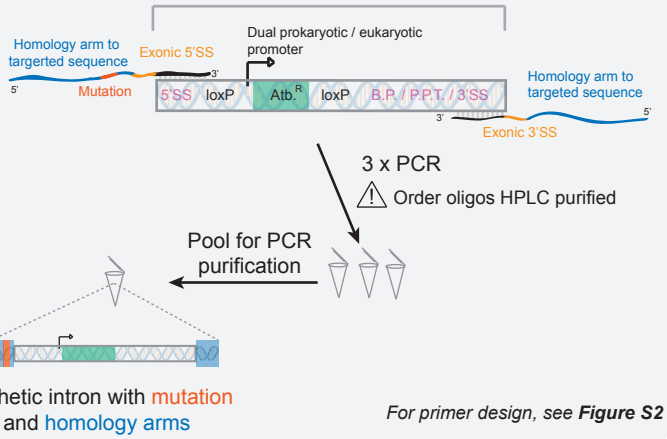
BAC purification

Day 5: Pre-culture of PCR positive clones (LB, Atb. #2)
Day 6: 1) Large scale culture from pre-culture (LB, Atb. #2)
2) BAC purification using a BAC prep. Kit

Do not freeze BAC DNA (keep at 4°C)

NB : The integritry of BACs that have integrated the synthetic intron can be further checked by comparing their restriction profile with the one of the template BAC

B



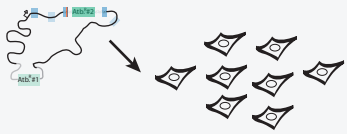
*For primer design, see **Figure S2***

C

Day 5 to day 7

BAC transfection into host cell line

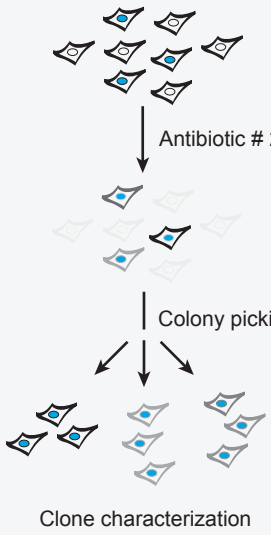
Day 5: Seed cells for transfection (60-70% confluent on day 6)
Day 6: Transfect BAC DNA into cells
Day 7: Change media to fresh media



from Day 8

Selection of transfected clones expressing the ESI mutated transgene

Day 8: Change media to selection media (Antibiotic #2)
Day 21-28: Regularly change media to fresh selection media
Pick and expand resistant clones



If the BAC harbours a tagged transgene it is advised to select clones using the antibiotic corresponding to the resistance attached to the Tag.