protocols.io

🌐 Intein-assisted Bisection Mapping (IBM)

Trevor YH Ho[1], Baojun Wang[1]

[1]University of Edinburgh

Nov 18, 2020

1 | *Works for me*   dx.doi.org/10.17504/protocols.io.bpqdmms6

Trevor Ho
University of Edinburgh

ABSTRACT

Split inteins are powerful tools for seamless ligation of synthetic split proteins. Yet, their use remains limited because the already intricate split site identification problem is often complicated by the requirement of extein junction sequences. To address this, we augmented a mini-Mu transposon-based screening approach and devised the intein-assisted bisection mapping (IBM) method. IBM robustly revealed clusters of split sites on four proteins, converting them into AND or NAND logic gates. In this protocol we provide details in executing the IBM method.

GUIDELINES

It would be best to read the manuscript "Intein-assisted bisection mapping systematically splits proteins for Boolean logic and inducibility engineering" prior to engaging this protocol.

An overview of protocol is given in Introduction. Please peruse before proceeding to the rest of the protocol. Please also refer to our manuscript for the literature on which this method was built upon.

MATERIALS TEXT

Most materials used in this protocol can be obtained from common commercial vendors, except the MuA transposase.

⊠ MuA Transposase **Thermo**

While the **Fisher Catalog #F750** is available, our MuA transposase was purified and supplied by Domus Biotechnologies (Turku, Finland).

With the transposase comes with:

- MuA dilution buffer
- 2x MIX

The following solution needs to be prepared by need to be prepared by users beforehand
- 2.5 M NaCl + 0.25 M MgCl$_2$ 1µl
- 1.25% Triton X-100

Wherever necessary, users might need to prepare their own electrocompetent cells.

Other reagents and materials used in this protocol:

Kits and consumables

⊠ Gene Pulser/MicroPulser Electroporation Cuvettes 0.1 cm gap **Bio-rad Laboratories Catalog #1652089** [In 2 steps]

⊠ NEB 10-beta Electrocompetent E.coli - 6x0.1 ml **New England Biolabs Catalog #C3020K** [Step 10]

⊠ NEB 10-beta/Stable Outgrowth Medium - 4x25 ml **New England Biolabs Catalog #B9035S** [Step 10.1]

⊠ QIAprep Spin Miniprep Kit Print **Qiagen Catalog #27104** [Step 2]

⊠ ZymoPURE™ II Plasmid Midiprep Kit **Zymo Research Catalog #D4201** [Step 2]

⊠ Monarch PCR and DNA Cleanup Kit - 250 preps **New England Biolabs Catalog #T1030L** [Step 2]

⊠ Aluminium Seal **StarLab Catalog #E2796-9792** [Step 39]

Inducers:

⊠ L-( )-Arabinose **Sigma Aldrich Catalog #A3256** [Step 33]

⊠ 24-Diacetylphloroglucinol **Cambridge Bioscience Catalog #16345** [Step 33]

Enzymes:

⊠ BsaI-HFv2 **New England Biolabs Catalog # R3733L** [In 3 steps]

⊠ FastAP Thermosensitive Alkaline Phosphatase (1 U/µL) **Thermo Fisher Catalog #EF0651** [Step 3.2]

⊠ BbsI-HF - 1,500 units **New England Biolabs Catalog #R3539L** [In 2 steps]

⊠ BglII - 10,000 units **New England Biolabs Catalog #R0144L** [Step 4.2]

⊠ SapI - 1,250 units **New England Biolabs Catalog #R0569L** [In 2 steps]

⊠ T4 DNA Ligase **New England Biolabs Catalog #M0202** In 2 steps

⊠ Taq DNA Polymerase with ThermoPol® Buffer **New England Biolabs Catalog #M0267L** Step 39

Equipment (necessary):

MicroPulser Electroporator	
Electroporator	
Bio-Rad laboratories	1652100

Equipment (optional):

Eppendorf™ Concentrator Plus with Pump and GB Plug	
Concentrator	
Eppendorf	5305000568

OT-2	
Liquid handler	
Opentrons	OT-2

SAFETY WARNINGS

No particular hazard in executing this protocol.
Observe common biosafety precautions.

DISCLAIMER:

DISCLAIMER – FOR INFORMATIONAL PURPOSES ONLY; USE AT YOUR OWN RISK

The protocol content here is for informational purposes only and does not constitute legal, medical, clinical, or safety advice, or otherwise; content added to protocols.io is not peer reviewed and may not have undergone a formal approval of any kind. Information presented in this protocol should not substitute for independent professional judgment, advice, diagnosis, or treatment. Any action you take or refrain from taking using or relying upon the information presented here is strictly at your own risk. You agree that neither the Company nor any of the authors, contributors, administrators, or anyone else associated with protocols.io, can be held responsible for your use of the information contained in or linked to this protocol or any of our Sites/Apps and Services.
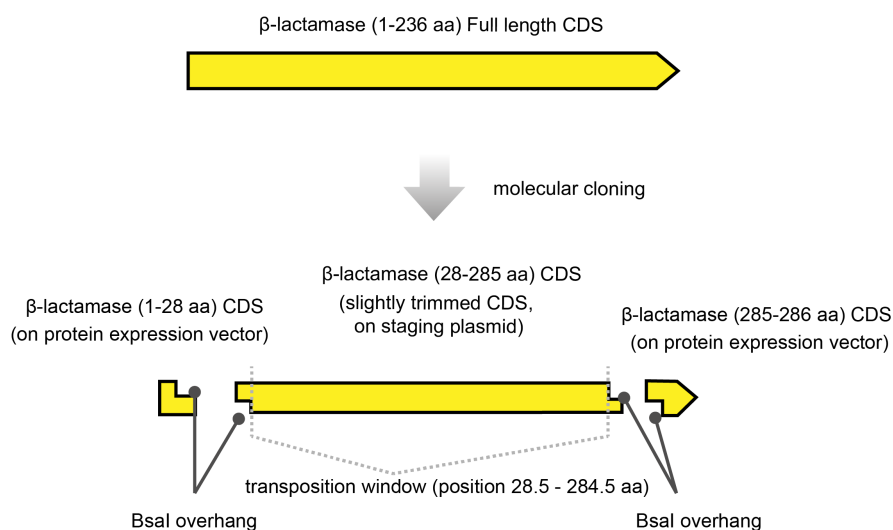
ABSTRACT

Split inteins are powerful tools for seamless ligation of synthetic split proteins. Yet, their use remains limited because the already intricate split site identification problem is often complicated by the requirement of extein junction sequences. To address this, we augmented a mini-Mu transposon-based screening approach and devised

the intein-assisted bisection mapping (IBM) method. IBM robustly revealed clusters of split sites on four proteins, converting them into AND or NAND logic gates. In this protocol we provide details in executing the IBM method.

BEFORE STARTING

Before you start, you will need to perform some molecular cloning on your gene of interest. You need to separate most of the middle part of the coding DNA sequence (termed slightly trimmed CDS in this protocol) and subclone it onto a staging vector.

β-lactamase (1-236 aa) Full length CDS

molecular cloning

β-lactamase (1-28 aa) CDS
(on protein expression vector)

β-lactamase (28-285 aa) CDS
(slightly trimmed CDS,
on staging plasmid)

β-lactamase (285-286 aa) CDS
(on protein expression vector)

transposition window (position 28.5 - 284.5 aa)

BsaI overhang

BsaI overhang

Plasmid construction prior to IBM execution.

This ensures that all transposition events happen within the CDS of your choice. Please also note how the transposition window was defined. In this protocol, most of the steps will be illustrated on an example of conducting intein-assisted bisection mapping on the gene encoding TEM-1 β-lactamase, where the protein will be split by the gp41-intein such that both the N-lobe (our way of calling the N-terminal split protein fragment) and the C-lobe need to be expressed to gain resistance to ampicillin.

Introduction

1
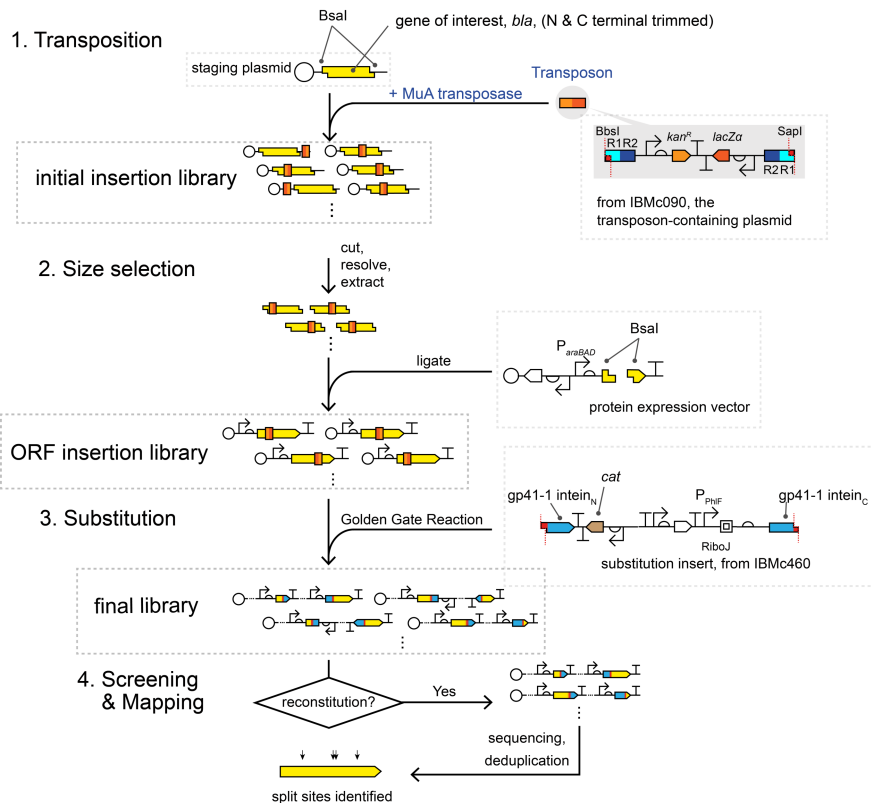
Summary workflow of IBM. This overview describes the split of β-lactamase by the gp41-1 intein.

IBM starts with a transposition reaction: a staging vector carries a 5' and 3' trimmed gene of interest (GOI) with the internal BsaI, BbsI and SapI sites removed. The staging plasmid is mixed in vitro with the MuA transposase and the mini-Mu transposon, which generates the initial insertion library. After in vivo amplification (transformation into *E. coli* and overnight growth), the insertion library is cut by BsaI and resolved on a DNA agarose gel. The band with the correct size corresponding to GOI with the insertion is purified and ligated into a linearized, protein expression vector with BsaI-generated overhangs. This produces the Open Reading Frame (ORF) insertion library. The ORF insertion library is again amplified in vivo and then an aliquot is mixed with a DNA fragment carrying the split gp41-1 intein intein in a Golden Gate reaction, using the restriction enzymes BbsI and SapI. The DNA fragment contains a different selection marker (*cat*) to facilitate selection of the library with the transposon replaced. In the end, the final library is screened for individual strains that showed proper reconstitution of protein function, i.e. resistance against ampicillin upon expression of both the N- and C-lobes (AND logic gate behavior). Those strains were then subjected to Sanger sequencing the C-terminal joint to map the split sites. Identical split sites were then deduplicated and consolidated with activity data to generate the intein-bisection map.

This protocol describes the main procedures of transposition, size selection, substitution, and mapping. Molecular cloning to generate the starting DNA constructs are not covered. However, this protocol will start with describing the necessary treatment on those constructs once they are available.

The most important objective in this protocol is to obtain a final library with large enough "library complexity, or diversity", in other words, ensuring that during construction and the final phase, the libraries contain enough random events so one can be confident that most possible splits are covered, and that a function split site for intein can be screened out.

Kits

2 We found good quality of DNA from each step is essential for the success of the protocol.
We use the following kits:

⊠ QIAprep Spin Miniprep Kit

1. Print **Qiagen Catalog #27104**   for miniprep

2. ⊠ ZymoPURE™ II Plasmid Midiprep Kit **Zymo**
   Research Catalog #D4201   for midiprep

3. ⊠ Monarch PCR and DNA Cleanup Kit - 250 preps **New England**
   Biolabs Catalog #T1030L   for gel and all
   other DNA purifications and concentrating steps.

| Preparation of DNA materials | 2h |

3 Prepare linearized protein expression vector

3.1 Grow and harvest 2 * 🥛**10 mL bacterial culture** of the protein expression vector.
Perform a mini-prep the next day.
Pool the extracted DNA into one tube.

3.2 Digest 🥛**5 µg plasmid of protein expression vector** from miniprep using   2h

⊠ BsaI-HFv2 **New England**

Biolabs Catalog # R3733L   .

Simultaneously treat with

⊠ FastAP Thermosensitive Alkaline Phosphatase (1 U/µL) **Thermo**

Fisher Catalog #EF0651

.

Number of reactions	1	5
10X Cutsmart Buffer	2	10
BsaI-HFv2 (20 U/µL)	1	5
FastAP (1 U/µL)	1	5
DNA (concentration: 89 ng/µL)	11.2	56
water	4.8	24

Example digestion recipe for linearizing the protein expression vector. All volumes in tables are in µL.

Run reaction at 🌡 **37 °C** for at least ⏲**02:00:00** .

3.3 Resolved digested DNA on 0.5 - 1% agarose gel, depending on the size of the vector. Load
🥛**1 µg digested DNA** per lane.
Excise the band corresponding to the linearized vector.

3.4 Purify the DNA.
Load at most 2 gel slices (1 slice per lane on gel) per column.
Elute DNA from each column using 🥛**10 µl nuclease free water** .
Pool eluted DNA together and quantify the concentration.

4 Prepare the linearized mini-Mu transposon

4.1 Grow and harvest 🥛**50 mL bacterial culture** for the plasmid carrying the MuA-transposon
(plasmid ID = IBMc090).
Perform a midi-prep the next day.

4.2 Digest 🧪**20 µg mini-Mu containing plasmid** from miniprep using

⊠ BglII - 10,000 units **New England Biolabs Catalog #R0144L**   and

⊠ BsaI-HFv2 **New England Biolabs Catalog # R3733L**   .

	1X	20
10X Cutsmart Buffer	2	40
DNA (1000 ng/µL)	1	20
BglII (10 U/µL)	1	20
BsaI-HFv2 (20 U/µL)	1	20
water	15	300

Example digestion recipe for linearizing the mini-Mu transposon. All volumes in tables are in µL.

Run reaction at 🌡 **37 °C** , for at least ⏱ **02:00:00** , or 🌙 **Overnight** if time is permitted.

4.3 Resolved digested DNA on 1% agarose gel. Load 🧪**1 µg digested DNA** per lane.

Run the gel at slow speed, e.g. 75 V for 60 - 90 mins helps to increase band sharpness.

> 📈
>
> LT release example.png
>
> Expected result of digested IBMc090 resolved on agarose gel.
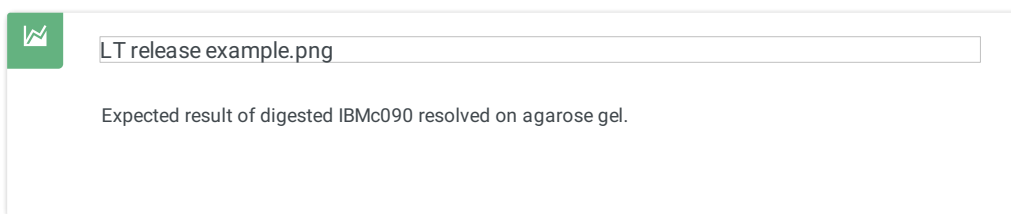
Excise the band with size of 1667 bp, which corresponds to the linear transposon.

4.4 Purify the DNA.

Load at most 2 gel slices (1 slice per lane on gel) per column.

Elute DNA from each column using 🧪**10 µl nuclease free water**

Pool eluted DNA together and quantify the concentration.

> 📈
>
> It is important to obtain > 80 ng/µL of purified linear transposon for downstream processes. If your purification efficiency is too low, try using a
>
> > Eppendorf™ Concentrator Plus with Pump and GB Plug
> > Concentrator
> > **Eppendorf**   5305000568
>
> to concentrate the DNA.

5 Prepare the staging plasmid

5.1 Grow and harvest ☐**10 mL bacterial culture** for the staging plasmid.
Perform a mini-prep the next day.

6 Prepare the linear substitution insert DNA

6.1 Grow and harvest ☐**50 mL bacterial culture** for the plasmid carrying the substitution insert.
Perform a midi-prep using the next day.

6.2 Digest ☐**10 µg substitution insert containing plasmid** from midiprep using   2h

⊠ BbsI-HF - 1,500 units **New England**
Biolabs Catalog #R3539L   and
⊠ SapI - 1,250 units **New England**
Biolabs Catalog #R0569L

NEB buffers and enzymes	1	10
10X CS Buffer	2	20
DNA (819 ng/µL)	1.2	12
BbsI-HF (20U/µL)	1	10
SapI (10U/µL)	1	10
ddH2O	14.8	148

Example recipe to release substitution insert. All volumes in tables are in µL.

Note: DO NOT try to be thrift with SapI. It will result in incomplete digestion.

Run reaction at 🌡 **37 °C** , for at least ⏲ **02:00:00** , or 🌙 **Overnight** if time is permitted.

6.3 Resolved digested DNA on 1% agarose gel. Load ☐**1 µg digested DNA** per lane.

Expected result of digested IBMc460 resolved on agarose gel.

Excise the band with size of 2591 bp, which corresponds to the substitution insert.

6.4 Purify the DNA.

Load at most 2 gel slices (1 slice per lane on gel) per column.

Elute DNA from each column using 🔲**10 µl nuclease free water**

Pool eluted DNA together and quantify the concentration.

Transposition using an engineered mini-Mu transposon   6h 50m

7 ⚠️

Prepare MuA transposase

The MuA transposase should be stocked at ❄ **-80 °C** at high concentrations to maximize shelf-life.

Dilute the concentrated MuA transposase to 220 ng/µL right before setting up the transposition reaction.

Using leftover MuA transposase from previous experiments, even if kept at ❄ **-20 °C** , could lead to a much reduced transposition efficiency.

8 Set up 6 * 🔲**25 µl transposition reaction** .

6h 10m

Number of reactions	1	6
2X Mix	12.5	75
2.5 M NaCl + 0.25 M MgCl2	1	6
1.25% Triton X-100	1	6
MuA Transposase (220 ng/µl)	3	18
Transposon (69.6 ng/µl)	4.3	25.8
Vector (158 ng/µl)	1.3	7.8
H2O	1.9	11.4

Example recipe for transposition reaction. All volumes in tables are in µL.

Aliquot all mixtures into 2 * 🔲**75 µl reactions** in 200 µl PCR tubes.

Run reactions on a thermocycler for

1. ❄ **30 °C** for ⏱ **06:00:00**

2. ❄ **80 °C** for ⏱ **00:10:00**

9 Pooled the reaction mixture together and purify the DNA library.

Use a single column and elute all DNA in 🔲**10 µl nuclease free water** .

It is not necessary to quantify the DNA concentration.

Electroporation of DNA library into *E. coli*

10 Electroporate the DNA library into

⊠ NEB 10-beta Electrocompetent E.coli - 6x0.1 ml **New England Biolabs Catalog #C3020K** .

10.1 Aliquot 4 * 🔲**1 mL** of

30m

⊠ NEB 10-beta/Stable Outgrowth Medium - 4x25 ml **New England Biolabs Catalog #B9035S**

(come with the electrocompetent cells) into microcentrifuge tubes or 30 mL universal tubes / test tubes

and warm up the medium in tubes at 🌡 **37 °C** for at least 🕐**00:30:00** .

10.2 Thaw and then mix 🧊**100 µl electrocompetent cells** with the eluted DNA library in a pre-chilled microcentrifuge tube. Place the tube on ice for 🕐**00:05:00**

5m

10.3 Aliquot 🧊**27.5 µl DNA-competent cell mixture** into a

⊠ Gene Pulser/MicroPulser Electroporation Cuvettes 0.1 cm gap **Bio-rad Laboratories Catalog #1652089**

5m

Ensure that the mixture is loaded in the gap. Place the cuvette on ice for 🕐**00:05:00**

10.4 Take out the warmed up Stable Outgrowth Medium.
Wipe down any traces of water on the cuvettes. Carry out electroporation using

> MicroPulser Electroporator
> Electroporator
> **Bio-Rad laboratories** 1652100

Use protocol EC1. Between each cuvette, jot down the time constants for electroporation event. The value should be between 4 to 5.5 ms. Even if the value is out-of-range, keep going.
Immediate after electroporation, add 🧊**972.5 µl Stable Outgrowth Medium** into the electroporation cuvette. Gently suspend cells by mixing up and down for 1-2 times. Transfer all suspended cells in medium into the microcentrifuge or universal tube where the medium was warmed. For samples of the same library, pool the suspended cells together.
It may be helpful to do this aseptically, but this is not strictly required. Speed is more important.

10.5 After finishing the electroporation steps for all samples, incubate the tubes at
🧫**160 rpm, 37°C, 01:00:00** for recovery.

10.6 Retrieve the recovered cells.
For each library, there should be ~ 4 mL of recovered cells. Take out 🧊**20 µl recovered cells** and dilute it into 🧊**180 µl 0.85% saline** . Serially dilute it again 10-fold for 5 times. Label the tubes from A (lowest dilution) to E (highest dilution).
Spread 🧊**100 µl diluted cells** onto 90 mm LB agar + KAN + CHL plates labelled A to E. This is to quantify the "library complexity".

10.7 For the rest of the recovered cells, spread onto 4 * 145 mm LB agar + KAN + CHL plates, each using ~ 🧊**1 mL recovered cells** . Allow the plates to dry before inverting them upside down.

Incubate all plates overnight at 🌡 **37 °C**

Retrieve the plates and determine library complexity.

11

For each library, tally the colony counts from plates A - E and create a table like below:

Plate	# of colonies	dilution factor	Initial number of transformants
A	TMTC	400	-
B	TMTC	4000	-
C	TMTC	40000	-
D	48	400000	19200000
E	4	4000000	16000000

Example record table for calculating library complexity. TMTC = too many to count

For n bp there is n-1 possibilities of insertion, and the transposon can insert in both forward and reverse orientations.

Total number of colonies (aka library complexity) was $1.92 * 10^7$.

Size of staging plasmid is 2859 bp.

Therefore, library fold coverage was $1.92 * 10^7 / ((2859 - 1) * 2) = 3359$ fold

If the library fold coverage is > 100. Proceed. Otherwise, discard the library.

Cell scraping and midiprep

12

Scrap down all cells from the 4 * 145 mm plates using a L-shape spreader and 0.85% saline. For each plate, add **20 mL 0.85% saline** to each plate, scrape down cells and transfer cells in saline to a 50 mL centrifuge tube. Repeat this once.

13

Centrifuge all tubes at **5000 rpm, Room temperature , 00:10:00** , remove most of the supernatant until the volume is < **10 mL** , then resuspend the cell pellets completely. Pool all resuspended cells into one of the four tubes. Repeat centrifugation with a balance.

14

⚠️

Remove supernatant again until total volume is about **5 mL** . At this the suspension is very dense and probably a bit viscous. Take out **1 mL resuspended cells** and save it as a glycerol stock.

The glycerol stock can be re-propagated and will save huge amount of time if you need to reuse the library.

15

Repeat the centrifuge again and this time, discard all supernatant. Perform a midiprep to extract the initial insertion library.

Size selection on initial insertion library to generate ORF insertion library   2h

6h

16

Digest **10 µg extracted, initial insertion library** using

⊠ BsaI-HFv2 **New England Biolabs Catalog # R3733L**

Number of reactions	1	10
10X CS Buffer	2	20
BsaI-HFv2 (20U/µL)	1	10
DNA (769 ng/µL)	1.3	13
water	15.7	157

Example digestion recipe for size selection. All volumes in tables are in µL.

Run reaction at 🌡 **37 °C** , for at least ⏱**06:00:00** , or ⏱**Overnight** if time is permitted.

17 ⚠️

Resolved digested DNA on 0.5 - 1% agarose gel. The gel percentage will depend on the expected sizes.

Load 🧪**1 µg digested DNA** per lane.

Critical: Run the gel at slow speed, e.g. 75 V for 90 mins. Resolution between bands is extremely important here, because the band corresponding to the trimmed CDS inserted with the transposon could be faint and close to other bands.

Excise the band which corresponds to the trimmed CDS inserted with the linear transposon.

Example in calculating the theoretical size of the trimmed CDS inserted with the transposon
- The staging plasmid for β-lactamse is 2859 bp.
- When digested by BsaI, it yields 771 bp (trimmed CDS) and 2088 bp (backbone).
- Therefore, after transposition:, each receives + 1667 + 5 bp
- Expected sizes of digested library are therefore 771, 2088, *2443*, 3760 bp
- The *2443 bp* band is the band containing the trimmed CDS inserted with transposon
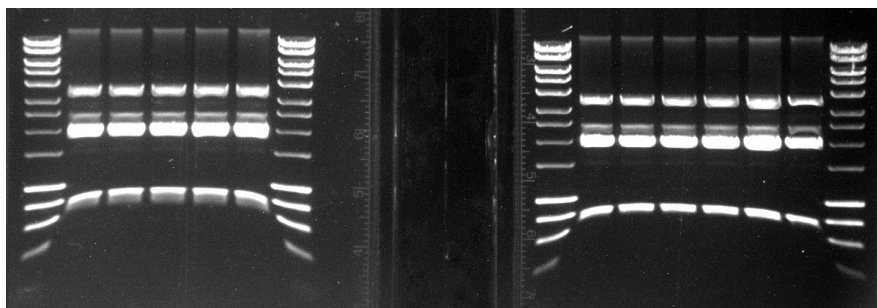
Example size selection step during IBM on β-lactamse, pre-excision. Notice how faint the desired band could be.
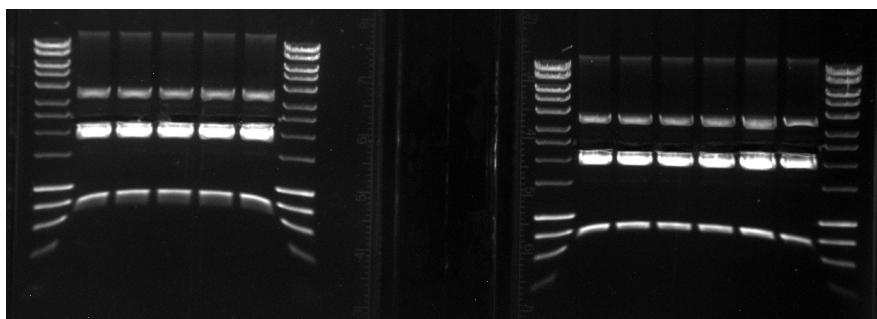
Example size selection step during IBM on β-lactamse, post-excision. It might not be always possible to exclude all other fragments while recover everything single bit of the desired band, but it is ok to proceed.

18 Purify the DNA.

Load at most 2 gel slices (1 slice per lane on gel) per column.

Elute DNA from each column using 🧊 **10 µl nuclease free water** .

Pool eluted DNA together and quantify the concentration.

19 Set up ligation of purified DNA to protein expression vector using the ⊠ T4 DNA Ligase **New England**

Biolabs Catalog #M0202

1h 30m

Use 1:2 insert:vector ratio to increase the chance that each insert will receive a vector.

At this stage, the ligation volume is typically 🧊 **100 µl** , but this can vary.

Number of reactions	1
10X T4 DNA Ligase Buffer	10
Vector (119 ng/µL)	9
Insert (6 ng/µL)	46
T4 DNA Ligase	10
water	25

Example ligation recipe to ligate size selected DNA into protein expression vector. All volumes in tables are in µL.

Run the ligation at 🌡 **Room temperature** for 🕐 **Overnight**

20 Purify the ligated DNA library.

Use a single column and elute all DNA in 🧊 **10 µl nuclease free water** .

It is not necessary to quantify the DNA concentration.

Electroporation and library complexity quantification for ORF insertion library

21 From this point onward, it is generally acceptable to move away from expensive commercial electrocompetent cells to in-house prepared electrocompetent cells. The procedures are therefore nearly identical to the steps from ↻ **go to step #10** to ↻ **go to step #11** with minor changes. To make the protocol easier to follow, the steps are reproduced below with the changes incorporated.

22 Electroporate the ligated, ORF insertion DNA library into in-house electrocompetent cells.

22.1 Aliquot 2 * 🧊 **1 mL SOC medium** into microcentrifuge tubes or 30 mL universal tubes / test tubes and warm up the medium in tubes at 🌡 **37 °C** for at least 🕐 **00:30:00** .

22.2 Thaw and then mix 🧊 **100 µl electrocompetent cells** with the eluted DNA library in a pre-chilled microcentrifuge tube. Place the tube on ice for 🕐 **00:05:00**

22.3 Aliquot 🔲**55 µl DNA-competent cell mixture** into a

⊠ Gene Pulser/MicroPulser Electroporation Cuvettes 0.1 cm gap **Bio-rad Laboratories Catalog #1652089**

Ensure that the mixture is loaded in the gap. Place the cuvette on ice for ⏲ **00:05:00**

22.4 Take out the warmed up SOC medium.
Wipe down any traces of water on the cuvettes. Carry out electroporation using

MicroPulser Electroporator	
Electroporator	
Bio-Rad laboratories	1652100

Use protocol EC1. Between each cuvette, jot down the time constants for electroporation event. The value should be between 4 to 5.9 ms. Even if the value is out-of-range, keep going.
Immediate after electroporation, add 🔲**945 µl SOC** into the electroporation cuvette. Gently suspend cells by mixing up and down for 1-2 times. Transfer all suspended cells in medium into the microcentrifuge or universal tube where the medium was warmed. For samples of the same library, pool the suspended cells together.
It may be helpful to do this aseptically, but this is not strictly required. Speed is more important.

22.5 After finishing the electroporation steps for all samples, incubate the tubes at
🔬**160 rpm, 37°C, 01:00:00** for recovery.

22.6 Retrieve the recovered cells.
For each library, there should be ~ 2 mL of recovered cells. Take out 🔲**10 µl recovered cells** and dilute it into 🔲**90 µl of 0.85% saline** . Serially dilute it again 10-fold for 5 times. Label the tubes from A (lowest dilution) to E (highest dilution).
Spread 🔲**50 µl diluted cells** onto 90 mm LB agar + KAN + TET plates labelled A to E. This is to quantify the "library complexity".

Note that the antibiotics used change here to reflect to selection of different genetic components. TET is for selecting the protein expression vector, it may change if you use a different protein expression vector. KAN is for selecting the transposon.

22.7 For the rest of the recovered cells, spread onto 2 * 145 mm LB agar + KAN + TET plates, each using ~ 1 mL of the cells. Allow the plates to dry before inverting them upside down.

Incubate all plates overnight at 🌡 **37 °C**

23 Retrieve the plates and determine library complexity.

For each library, tally the colony counts from plates A - E and create a table like below:

Plate	# of colonies	dilution factor	Initial number of transformants
A	TMTC	400	-
B	108	4000	432000
C	13	40000	520000
D	1	400000	400000
E	0	4000000	0

Example record table for calculating library complexity for the ORF insertion library.

Total number of colonies (aka library complexity) was $4.32 * 10^5$.

Size of trimmed CDS is 767 bp.

Therefore, library fold coverage was $4.32 * 10^5 / ((767-1) * 2) = 282$ fold

NOTE: when calculating library fold coverage for the ORF insertion library, the formula becomes:
coverage = number of colonies / ((size of transposition window) * 2)
the size of transposition = size of the trimmed CDS -1

If the library fold coverage is > 100. Proceed. Otherwise, discard the library.

24 The rest of the steps including cell scrapping and midiprep are the same.

Substitution of transposon with split intein	4h

25 Take DNA from midipreped ORF insertion library, mix with the substitution insert in a 1:5 for library : insertion ratio and add to Golden Gate Assembly reaction.   4h
The substitution insert is provided in excess to increase the chance that all transposon within the library is replaced.

Set up 6 * 🗋**20 µl of Golden Gate Assembly reaction** in total.

⊠ BbsI-HF - 1,500 units **New England**
Enzymes used are **Biolabs Catalog #R3539L** ,

⊠ SapI - 1,250 units **New England**
Biolabs Catalog #R0569L , and

⊠ T4 DNA Ligase **New England**
Biolabs Catalog #M0202

Note: the maximum amount of ORF insertion library DNA per 20 µL of Golden Gate reaction mix is 60 ng.

Number of reactions	1	6
10X T4 Ligase Buffer	2	12
BbsI-HF	1	6
SapI	1	6
T4 DNA Ligase	1	6
ORF insertion library (452 ng/µL)	0.133	0.798
substitution insert (42 ng/µL)	2.6	15.6
water	12.267	73.602

Example recipe for a Golden Gate substitution reaction. All volumes in tables are in µL.

Aliquot the master mix into 2 * ▭**60 µl of reactions** .

Place them on thermocyclers and run the following program:

Step		Temperature (°C)	Time
1	Digestion	37	3 mins
2	Ligation	16	4 mins
Repeat steps 1 - 2 for 25 cycles			
3	Final Digestion	37	30 mins
4	Heat Inactivation	65	20 mins
5	Storage	16	infinite

Thermocycler program for Golden Gate substitution reaction.

This protocol typically takes ⏱ **04:00:00** .

26   Pool all the Golden Gate Assembly reaction mixtures together and purify the final library.

Use a single column and elute all DNA in ▭**10 µl nuclease free water** .

It is not necessary to quantify the DNA concentration.

Electroporation and library complexity quantification for the final library

27   Repeat ↻ **go to step #21** to ↻ **go to step #22.5** .

Note: if the target protein needs a reporter plasmid to manifest its function, then it is necessary to make bespoke electrocompetent cells that carry the reporter plasmid.

28   After the recovery step, there should be ~ 2 mL of recovered cells per library. Take ▭**1 mL recovered cells** and save as a glycerol stock.

29   Continue with ↻ **go to step #22.6** and ↻ **go to step #23** to estimate library complexity for the final library.

Screening for AND / NAND gate behavior

30   There is no fixed way to screen the final library. There could be different ways in which steps of enrichment or selection are conducted. Yet, in any case, the objective is to first isolate single colonies from the library that clearly show restoration of function when both the N- and C-lobes are expressed. The following steps provide a general description of what happens next.

31   Once single colonies that show restoration of function are obtained, they are picked into 96-well plates with medium supplemented with appropriate antibiotics. Typically, > 500 colonies are picked.

If your lab uses TipOne tips, the 3D printable Colony Picking Assistance might come in handy.

32   Grow the plates at 🌡**1000 rpm, 37°C, 16:00:00** (overnight) to give a saturated culture.

33   Dilute ▭**2 µl overnight culture** into 96-well plates with ▭**198 µl medium** supplemented with appropriate antibiotics and with 4 different combinations of inducers.

⊗ L-( )-Arabinose **Sigma**

Assuming the inducer **Aldrich Catalog #A3256**   drives expression of the N-lobe and

the inducer ⊠ 24-Diacetylphloroglucinol **Cambridge Bioscience Catalog #16345** (DAPG) drives expression of the C-lobe, the 4 combinations would be

1. water + DMF (solvent for DAPG)
2. arabinose + DMF
3. water + DAPG
4. arabinose + DAPG

Grow inoculated plates at 🌡 **1000 rpm, 37°C** . The end time of this screening assay depends on the experiment context.

34 Assay the plates and identify clones that restore the original protein function when both inducers are present. Cherry-pick and consolidate the individual wells into a new 96-well plate. Usually 93 - 195 colonies are good enough for subsequent steps.

Using a robotic liquid handler like OT-2 is strongly recommended, otherwise you risk transferring the wrong wells because this task is prone to mental exhaustion and hence mistakes.

OT-2
Liquid handler

Opentrons   **OT-2**

35 Grow the consolidated plate(s) at 🌡 **1000 rpm, 37°C, 16:00:00** . This plate is now the stock. Duplicate this plate and make a 96-well 🌡 **-80 °C** glycerol stock plate by adding 🧪 **100 µl bacterial culture** to 🧪 **50 µl 50% glycerol** . Freeze immediately in 🌡 **-80 °C freezer**

36 Repeat ↻ **go to step #32** to ↻ **go to step #33** on the isolated clones for a rigorous experiment.
Clones might behave different from the first time they are screened out. Keep only the clones that give consistent behavior.

| Whole-cell PCR and sequencing to identify split sites |

37 Once the clones are confirmed to show consistent AND or NAND behavior, proceed to whole-cell PCR.

You only needs to amplify and sequence the region where the intein-extein CDS region join together. Both the N- or C-junction could be sequenced. In most cases the analysis is easier if the $intein_C$-$extein_C$ junction is sequenced. This option is assumed for the steps below.

38 Design and order primers
The forward primer should anneal ~ 50 - 100 bp before the 3' end of the $intein_C$ CDS and the reverse primer should anneal ~ 50 - 100 bp after the 3' end of the permitted transposition window.
While designing primers, estimate the maximum size possible for the PCR product.

39 Carry out whole-cell PCR using
⊠ Taq DNA Polymerase with ThermoPol® Buffer **New England Biolabs Catalog #M0267L**

This is equivalent to colony PCR just that bacterial cells were obtained from liquid cultures instead of solid medium.

Assuming you have 93 candidate clones to sequence, prepare 2 tubes of the master mixes below, each in a 1.5 mL microcentrifuge tube:

Number of reactions	1	52
10X ThermoPol Buffer	2.5	130
1mM dNTP	5	260
FW primer	0.5	26
RV primer	0.5	26
Template bacteria	0.5	-
Taq Pol	0.125	6.5
water	15.875	825.5

Example recipe for whole-cell PCR. All volumes in tables are in µL.

Aliquot 🟦**24 µl PCR master mix** into each well of a 96-well PCR plate, then add 🟦**1 µl bacterial culture** to the wells. There is no need to mix by pipetting up and down, which introduces bubbles. Seal the plate with an

⊠ Aluminium

Seal **StarLab Catalog #E2796-9792** .

Run the following protocol

Step		Temperature (°C)	Time
1	Initial Denaturation	95	2 mins
2	Denaturation	95	15 secs
3	Annealing	51	15 secs
4	Extension	68	1 mins
Repeat steps 2 - 4 for 35 cycles			
5	Final Extension	68	5 mins
6	Storage	12	infinite

Example whole-cell PCR thermocycling program. Assumes 1 kb of maximum PCR product size and that Taq's extension speed is 1 kb/min

40 ✪

Resolve 🟦**2.5 µl PCR product** on agarose gel to affirm PCR is successful.

41 Purify the PCR product.

Elute DNA from each sample in 🟦**20 µl nuclease free water** .

Dilute the eluted DNA to appropriate concentrations (10 ng/µL for Source BioSciences) and send to sequencing.

Note: this step can be done manually but it is best to be facilitated using a robotic liquid handler.

> OT-2
> Liquid handler
> **Opentrons**　　　**OT-2**

42 Customize the Python scripts from our [Github repository](#) and analyze the sequencing data to identify split sites for each candidate clones. Merge split site information and protein activity data to produce an intein-bisection map.

The intein-bisection map of β-lactamase.
Left panel, optical densities of the controls that provide the references as to how cells with or without intact β-lactamase (BLA) would grow when challenged by ampicillin. Right panel, bisection map of β-lactamase split by the split gp41-1 intein. Each vertical group of spots represents an identified split site on the x axis, aligned to the β-lactamase secondary structure (from PDB: 1ZG4) below. A total of 55 candidate strains showing AND gate behavior were characterized and sequenced to generate this map. y locations and error bars are mean and std of optical densities from independent experiments performed on three different days. Vertical dashed lines bound the permitted transposition window.