Protocol

*In vivo* gain-of-function cDNA library screening for colonization genes in a mouse model of pulmonary metastasis

We provide a protocol for gain-of-function (GOF) cDNA screen of genes that foster cancer cell colonization of secondary tissues, the last and most lethal step of the metastasis cascade. We present techniques for cDNA viral library preparation and delivery leading up to the recovery of colonization-promoting sequences in a proof-of-concept DU145-based mouse model of pulmonary metastasis. Adapted to other cDNA libraries and cancer models, this approach would prove widely useful in enumerating intrinsic genetic determinants underlying metastatic colonization.

Publisher’s note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.
Protocol

In vivo gain-of-function cDNA library screening for colonization genes in a mouse model of pulmonary metastasis

Zhenbo Tu$^{1,4}$ and Antoine E. Karnoub$^{1,2,3,4,5,*}$

$^1$Department of Pathology and Cancer Center, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA
$^2$Harvard Stem Cell Institute, Cambridge, MA 02138, USA
$^3$Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA
$^4$Technical contact
$^5$Lead contact
*Correspondence: ztu@bidmc.harvard.edu (Z.T.), akarnoub@bidmc.harvard.edu (A.E.K.)
https://doi.org/10.1016/j.xpro.2022.101413

SUMMARY

We provide a protocol for gain-of-function (GOF) cDNA screen of genes that foster cancer cell colonization of secondary tissues, the last and most lethal step of the metastasis cascade. We present techniques for cDNA viral library preparation and delivery leading up to the recovery of colonization-promoting sequences in a proof-of-concept DU145-based mouse model of pulmonary metastasis. Adapted to other cDNA libraries and cancer models, this approach would prove widely useful in enumerating intrinsic genetic determinants underlying metastatic colonization.

For complete details on the use and execution of this protocol, please refer to Tu et al. (2021).

BEFORE YOU BEGIN

Institutional permissions

In vitro and animal experiments were performed in accordance with and following approval by Harvard Committee on Microbiological Safety (COMS) and Institutional Animal Care and Use Committee (IACUC) of Beth Israel Deaconess Medical Center. Investigators following this protocol need to first seek approval from their own institutional regulatory agencies for the use of microbiological tools and animals in experimental research.

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Bacterial and virus strains** | | |
| DH10B Competent Cells | Thermo Fisher Scientific | Cat# EC0113 |
| **Chemicals, peptides, and recombinant proteins** | | |
| Dulbecco’s modified Eagle’s medium | Corning | Cat# 10-017-CV |
| Fetal bovine serum | Gibco | Cat# 26400044 |
| Phosphate-Buffered Saline (PBS) 10x | Multiscell | Cat# 311-012-CL |
| Glycerol | Thermo Fisher Scientific | Cat# 158920025 |
| Agar | Fisher Bioreagents | Cat# BP1423-500 |
| Agarose | KSE Scientific | Cat# BMK-A1705 |
| LB broth | Fisher Bioreagents | Cat# BP1426-2 |

(Continued on next page)
| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Ampicillin          | Corning| Cat# 69-52-3 |
| Polybrene           | Santa Cruz | Cat# sc-134220 |

Critical commercial assays

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Fugene HD           | Roche  | Cat# 04709705001 |
| Lipofectamine 2000  | Invitrogen | Cat# 11668019 |
| Qiagen DNase Blood & Tissue Kits | QIAGEN | Cat# 69906 |
| Qiagen Maxi Prep    | QIAGEN | Cat# 12162 |
| DNA ladder 1 kb+    | Invitrogen | Cat# 10787026 |
| PicoMaxx High Fidelity PCR System | Agilent | Cat# 600424 |
| Advantage™-GC 2 PCR Kit | Clontech | Cat# 639119 |

Experimental models: Cell lines

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| HEK293T (Passage <10) | ATCC  | Cat# ACS-4500 |
| DU145 (Passage <10) | ATCC  | Cat# HTB-81 |

Experimental models: Organisms/strains

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| NOD/SCID (mouse) (8–10 weeks males) | Charles River | Strain code 394 |

Oligonucleotides

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| pLIB-3’ primer      | Integrated DNA Technologies; Tu et al. (2021) | 5'-ACC TAC AGG TGG CTT TCA TTC CC-3' |
| pLIB-5’ primer      | Integrated DNA Technologies; Tu et al. (2021) | 5'-AGC CCT CAC TCC TCT TTC AG-3' |

Recombinant DNA

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| Human prostate retroviral library | Clontech | Cat# 634204 |
| pCMV-VSV-G           | Stewart et al. (2003) | Addgene Cat# 8454 |
| pUMVC                | Stewart et al. (2003) | Addgene Cat# 8449 |
| pCMV-dR8.2 dvpr      | Stewart et al. (2003) | Addgene Cat# 8455 |
| pRRL3-GFP            | Whitehead Institute | N/A |

Software and algorithms

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| GraphPad Prism 8    | GraphPad Software | https://www.graphpad.com/scientific-software/prism/ |

Other

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| C1000 Thermal Cycler | Bio-Rad | C1000 Touch |
| SteREO stereomicroscope | Zeiss | Discovery. V8 |

**MATERIALS AND EQUIPMENT**

**LB/Amp agar plate**

| REAGENT | Final concentration | Amount |
|---------|---------------------|--------|
| LB broth | n/a | 25 g |
| Agar | n/a | 15 g |
| ddH₂O | n/a | up to 1 L |
| Ampicillin | 100 μg/mL | 1 mL |
| Total | | 1 L |

Autoclave (121°C, 30 min) LB broth mixture and allow to cool until 50°C before adding ampicillin. Pour mixture into 15-cm dishes (15–20 mL/plate) and allow solidification at room temperature (25°C). Store LB/Amp plates at 4°C (maximum storage time < 1 month).

**LB/Amp medium**

| REAGENT | Final concentration | Amount |
|---------|---------------------|--------|
| LB broth | n/a | 25 g |
| ddH₂O | n/a | up to 1 L |
| Ampicillin | 100 μg/mL | 1 mL |
| Total | | 1 L |

Autoclave (121°C, 30 min) LB mixture and allow to cool until 50°C before adding ampicillin. Store medium at room temperature (25°C) (maximum storage time < 1 week).
**STEP-BY-STEP METHOD DETAILS**

**Setting up the plasmid library**

© Timing: 3 days

This section describes determination of the titer and amplification of the plasmid library.

1. Determining plasmid library titer.
   a. Prepare LB/Amp agar plates. Measure 25 g broth powder and 15 g bacteriological agar and swirl to mix in 1 L ultrapure water. Autoclave on liquid cycle (121 degree, 30 min) then cool to around 50°C in a temperature-controlled water bath. Add Ampicillin (final concentration 100 μg/mL) then pour the mixture into 18–20 15 cm plates (15–20 mL per plate) and allow so solidify at room temperature (25°C).
   b. Dilute 1 μL of the bacterial plasmid library into 1 mL of LB broth and mix by gentle vortexing as Dilution A.
   c. Dilute 1 μL of Dilution A into 1 mL of LB broth and mix by gentle vortexing as Dilution B.
   d. Remove 50 μL aliquots from Dilution B and spread onto each prewarmed LB/Amp agar plate.
   e. Leave the plate at room temperature (25°C) for 20 min and then incubate it inverted at 37°C overnight.
   f. Count the total number of colonies and calculate the titer by following this formula:

\[(\text{colony#} \text{ from Dilution B}/\text{Plating volume}) \times 10^3 \times 10^3 \times 10^3 = \text{colonies forming unit (cfu) per mL} \]

**Note:** The titer should be at least 10-times higher than the number of independent clones present in the library. In this case, since our library size was \(\approx 8.0 \times 10^6\), our cfu was >\(10^8\).

2. Amplification of plasmid library.
   a. Prepare enough of the cDNA library for transfection. Typically, this represents about 1 μL of 1,000-fold-diluted cDNA library culture (in E.coli DH10B) diluted in 150 μL of LB broth before spreading. It takes about 20 plates (150-mm) to amplify the whole cDNA library.
   b. Incubate plates at 37°C overnight.
   c. Add 5 mL of LB medium to each plate and scrape off colonies into 2 liters of LB/Amp broth. Repeat to ensure as much as possible was recovered from plates.
   d. Incubate cultures in a shaking incubator at 37°C to obtain a saturated culture (OD600 > 2.0).
   e. Proceed with plasmid isolation using QIAGEN Maxi Prep (Catalog number 12162) (https://www.qiagen.com/us/resources/download.aspx?id=0bd0c5fb-c271-43e7-af43-32d539374fa9&lang=en) to obtain highly purified plasmid library.

3. Verification of cDNA representation.
   a. Use the 5' and 3' LIB primers (sequences in the key resources table) to amplify cDNA inserts by Advantage™-GC 2 PCR Kit (https://www.takarabio.com/resourcedocument/x32404) in C1000 Thermal Cycler.

---

**Cell culture medium**

| Reagent                   | Final concentration | Amount |
|---------------------------|---------------------|--------|
| Dulbecco’s modified Eagle’s medium | n/a                 | 500 mL |
| Fetal bovine serum        | 10%                 | 50 mL  |
| Total                     |                     | 500 mL |

Store media at 4°C (maximum storage time < 2–3 weeks). Warm the medium to 37°C before using.
b. Conduct gel electrophoresis on the reactions to see a smear pattern consistent with the library size intervals. For the pLIB library we used here, that pattern should be between 0.5 kb to 5 kb size (Figure 1A).

Note: The library used here was commercially procured, but cDNA libraries can be prepared in-house from any RNA pool. Flanking linkers can be added to the 5’ and 3’ insert ends using sequences compatible with 5’ and 3’ LIB primers or any other linker sequences determined fitting by the investigator.

△ CRITICAL: If the need arises, the amplified bacterial library can be stored frozen in 33% glycerol at −80°C for later use.

Preparation of viral library

△ Timing: 3 days

This section describes packaging plasmid library into viral particles.

4. Culture of packaging cells.
   a. Plate HEK293T cells at a density of 6.0 × 10⁶ cells per 100-mm plate 12 h before transfection.

△ CRITICAL: Replace old medium with fresh medium before transfection (1–2 h).

Note: The number of packaging cell plates should be calculated and scaled based on library size. For example, the pLib-human prostate retroviral library clone size we used here was ~10 × 10⁶, so 10-times more packaging cells were needed, i.e., 100 × 10⁶ cells.

5. Transfection of packaging cells.
   a. Transfect 10 μg of plasmid DNA along with 1 μg pCMV-VSV and 10 μg pUMVC using 63 μL Lipofectamine 2000 per 100-mm plate. Adapt viral envelope and packaging plasmid based on the nature of the plasmid (e.g., retroviral or lentiviral plasmid) carrying library inserts.

---

### PCR reaction mix

| Reagent                          | Amount  |
|---------------------------------|---------|
| GC Melt (5 M)                   | 5 μL    |
| Primer mix (10 μM each)         | 2 μL    |
| Advantage-GC 2 Pol. Mix (50×)   | 1 μL    |
| dNTP mix (10 mM each)           | 1 μL    |
| DNA template                    | ~100 ng/μL |
| PCR-grade H2O                   | up to 50 μL |

### PCR cycling conditions

| Steps                  | Temperature | Time | Cycles |
|------------------------|-------------|------|--------|
| Initial Denaturation   | 94°C        | 3 min| 1      |
| Denaturation           | 94°C        | 30 s | 35     |
| Annealing/Extension    | 68°C        | 1.5 min| 1       |
| Final extension        | 68°C        | 3 min| 1      |
Note: Gently swirl plates to make sure the transfection solution is distributed uniformly throughout the culture surface.

6. Virus recovery.
   a. Aspirate old medium at 10–12 h post-transfection, then wash the cells twice with 1× PBS and add 5–6 mL of fresh medium.
   b. Aspirate medium 32–38 h after transfection and pass through 0.45 μm filters.
   c. Media containing undiluted viral particles can be immediately used, stored at 4°C for use within 24 h, or frozen at −80°C for later use.

Cancer cell infection

© Timing: 2 days

This section describes viral library transduction of cancer cells.

7. Preparation of GFP-DU145 cells.
   a. Infect DU145 cells with pRRL3 lentivirus to generate GFP-DU145 cells.
   b. Plate GFP-DU145 cells at a density of 2.0 × 10^6 cells per 100-mm plate 12 h before infection.
   c. Generate 1:10 dilution of filtered library virus (preferably fresh) and administer to the GFP-DU145 cells in the presence of 5 μg/mL polybrene for 18 h.
   d. To ensure the library is well represented, conduct genomic PCR of infected cells after 48–72 h using 5' and 3' LIB primers. A smear between 0.5-5 kb should be obtained, indicating adequate transcript representation (Figure 1B). Of note, comparative library sequencing could also be conducted to ensure maintenance of library complexity.

△ CRITICAL: First, dilute the virus-containing media at least two-fold with fresh media. Second, it is very important to maintain a virion number that is substantially lower than the number of cells to be infected in order to reduce the risk of more than one cDNA insertion per cell. In our case here, and as an example, we used a multiplicity of infection (MOI) of 0.2, which we determined using pLAPSN, an expression plasmid containing the sequence of bacterial alkaline phosphatase. Assessment of alkaline phosphatase activity in the infected population allowed for a calculation of percentage of
cells transduced, which, when combined with the dilution factor used and the number of cells plated provided us with a surrogate measure of viral titer and virion numbers to utilize.

**Cellular library delivery in vivo and recovery of candidate colonization genes**

© Timing: 3–4 weeks

This section describes generation of metastases and verification of candidate colonization genes.

8. Generation of lung metastases.
   a. Mix $1.5 \times 10^5$ of the infected GFP-DU145 cells in $100 \mu$L PBS supplemented with 1% FBS and intravenously inject via tail vein into recipient NOD/SCID mice (8–10 weeks old).
   b. After 18 days, sacrifice these mice and excise and wash lungs in ice cold PBS. Lay down extracted lungs flat on a clean TC plate and examine the metastatic colonies using fluorescence-equipped dissecting SteREO stereomicroscope (Zeiss, Axio). Surgically cut pieces of lungs harboring distinct visible (~1 mm in diameter) GFP-positive colonies, minimizing the carryover of peripheral lung tissues (Figure 2). (troubleshooting 1).

9. Colonization gene recovery.
   a. Isolate genomic DNA from GFP-positive colonies using Qiagen DNeasy (https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj1xebG86_3AhUkhlEHeJkAFMQFnoECAUQAX&url=https%3A%2F%2Fwww.qiagen.com%2Fbr%2Fresources%2Fdownload.aspx%3Fid%3D63e22fd7-6eed-4bcb-8097-7ec77bcd4de6%26lang%3Den&usg=AOvVaw345qmlTImF31pltxHmQKeF).
   b. To identify pLIB-library-derived integrated sequences, amplify DNA using manufacturer-supplied 5’ and 3’ linker primers following PicoMaxx High Fidelity PCR system guidelines. (troubleshooting 2).

**PCR reaction mix**

| Reagent                        | Amount |
|--------------------------------|--------|
| 10X PicoMaxx reaction buffer   | 5 μL   |
| 100 mM dNTP mix (25 mM of each dNTP) | 0.4 μL |
| Primer #1 (100 ng/μL)         | 1 μL   |

(Continued on next page)
c. Resolve amplification reactions using 1% agarose electrophoresis (Figure 3).

d. Cut out resolved sequences and send out for DNA sequencing, then BLAST for gene identification. (troubleshooting 3).

EXPECTED OUTCOMES

Inserted sequences isolated from genomic DNA derived from macro-colonies are expected to represent cDNA sequences that are postulated, when re-expressed on their own, to be instigators of cancer cell metastatic growth in mouse lungs (e.g., Aldolase A; Figure 3). Also, it is expected for visible >1 mm macro-colonies to be clonal and derived from one cell harboring integrated cDNA sequence(s).

LIMITATIONS

It is our experience with the pLIB library used here that many of the sequences recovered by PCR from mouse lungs represented in fact pieces of cDNAs rather than full-length gene sequences. However, this did not preclude the identification of real modifiers of the colonization phenotype, as we describe in our paper of Tu et al. (2021). Since the pLIB platform was commercially generated, we could not optimize to purge such partial sequences from the library. This aspect limited the number of full-length sequences recovered from our otherwise proof-of-concept screen. Of note, cDNA libraries can be generated in-house from any tissue, and flanking linker sequences can be added as well to facilitate sequence recovery using PCR.

TROUBLESHOOTING

Problem 1

Insufficient recovery of genomic DNA quantities in step 9a.

Potential solution

For rather smaller nodules, recovering sufficient genomic DNA presented a significant challenge. To circumvent this step, we recurred to harvesting the nodules, digesting them with collagenase (Type I) under tumbling overnight at 37°C and plating the cells for expansion in tissue culture. Within 7–14 days, recovered cancer cells (which were GFP-labeled, hence their growth was followed using fluorescence microscopy) were then harvested for genomic DNA isolation and PCR. This approach has consistently led to better outcomes for subsequent reactions, although the fact that cancer cell populations drift in culture and may vary in their heterogeneity is a prominent drawback/concern with this expansion.

Problem 2

Failed amplification reactions from macro-colonies in step 9b.

---

| Reagent                        | Amount   |
|--------------------------------|----------|
| Primer #2 (100 ng/μL)          | 1 μL     |
| 50X dNTP mix (10 mM each)      | 1 μL     |
| PicoMaxx high fidelity PCR system | 1 μL     |
| PCR-grade H2O                  | up to 50 μL |

**PCR cycling conditions**

| Segment | Temperature °C | Time   | Cycles |
|---------|----------------|--------|--------|
| 1       | 95              | 2 min  | 1      |
| 2       | 95              | 40 s   | 35     |
|         | Primer Tm-5°C   | 30 s   |        |
| 3       | 72              | 1 min/kb | 1     |
|         | 72              | 10 min |        |
Potential solution
We have experienced a higher than expected ratio of failed PCR reactions compared to the successful ones in which bands were easily identifiable. This could be due to impurities that are inadvertent when surgically removing metastatic nodules (in this case, from the lungs), suboptimal PCR conditions for genomic DNA integration sites of the library sequences in the specific nodule isolated, or insufficient quantities of genomic DNA recovered. We found that using aseptic techniques in lung excision, minimizing surgical handling on ice, extensive washing of the lung pieces with cold 1× PBS or 1× HBSS, and immediate processing of tissues for genomic DNA isolation can help with minimizing impurities and in avoiding genomic DNA degradation. Regarding genomic PCR, we have experimented with several genomic DNA extraction and PCR kits and have rested on PicoMaxx High Fidelity PCR system as a successful system to amplify picogram quantities of recovered DNA. We also found that inclusion of 2%–5% DMSO in the reactions improved yields.

Problem 3
Amplification of multiple sequences from a single macro-colony in step 9c.

Potential solution
We have experienced the appearance of several bands when amplifying genomic DNA from isolated colonies. These occurrences could derive from (1) the fact that the macro-colony was derived in effect from multiple clones that coexisted/cooperated to generate the macro-metastasis, (2) the insertion of multiple library cDNA sequences into the genome of cancer cells that generated the macro-colony, and (3) non-specific priming during the amplification step. While we cannot control the first possibility, and the MOI of 0.2 minimized (but did not absolutely preclude) the occurrence that more than one cDNA was inserted per infected cell, we were able however, to improve PCR by modifying the cycling/temperature parameters and by increasing DMSO concentrations in the amplification reactions.

Problem 4
Failed sequencing reactions in step 9d.

Potential solution
Many of our submitted sequencing reactions failed because of poly-T’s and slippage. Although there is nothing that we could do experimentally to avoid such outcomes, we found that submitting >100–200 ng of PCR product decreased the chances of this happening. We have also found that replacing the 3’ library primer with Oligo-dT primer in the PCR assays lessened such occurrences.
Problem 5
Use of retroviral library that is only amplified in dividing cells.

Potential solution
We used here a commercial retroviral library for proof-of-principle purposes. We recognize, however, that retroviruses carrying certain cDNAs may be lost in subsequent steps if they infect non-dividing cells, and are hence unable to integrate into genomic DNA, thus leading to loss of representation. Although we used here highly proliferative HEK293T and DU145 cells, this remains a possibility considering the heterogeneity of these populations and the dynamics of quiescence and division among pre-existing clones. To circumvent this problem and transduction of all cells, use of lentiviral vectors for library cloning would be recommended.

RESOURCE AVAILABILITY
Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Dr. Antoine E. Karnoub (akarnoub@bidmc.harvard.edu).

Materials availability
This study did not generate new unique reagents.

Data and code availability
This study did not generate/analyze [datasets/code].

ACKNOWLEDGMENTS
We acknowledge institutional support from Beth Israel Deaconess Medical Center (to A.E.K.), Department of Defense BCRP grants BC160702 and BC142139 (to A.E.K.), and NCI grants CA207322 and CA249223 (to A.E.K.).

AUTHOR CONTRIBUTIONS
Z.T. and A.E.K. wrote the manuscript. Both authors contributed to the manuscript and approved it for publication.

DECLARATION OF INTERESTS
The authors declare no competing interests.

REFERENCES
Stewart, S.A., Dykxhoorn, D.M., Palliser, D., Mizuno, H., Yu, E.Y., An, D.S., Sabatini, D.M., Chen, J.S., Hahn, W.C., Sharp, P.A., et al. (2003). Lentivirus-delivered stable gene silencing by RNAi in primary cells. RNA 9, 493–501. https://doi.org/10.1261/rna.2192803.

Tu, Z., Hou, S., Zheng, Y., Abudul, M., Onder, T., Intlekofer, A.M., and Karnoub, A.E. (2021). In vivo library screening identifies the metabolic enzyme aldolase A as a promoter of metastatic lung colonization. Science 24, 102425. https://doi.org/10.1016/j.sci.2021.102425.