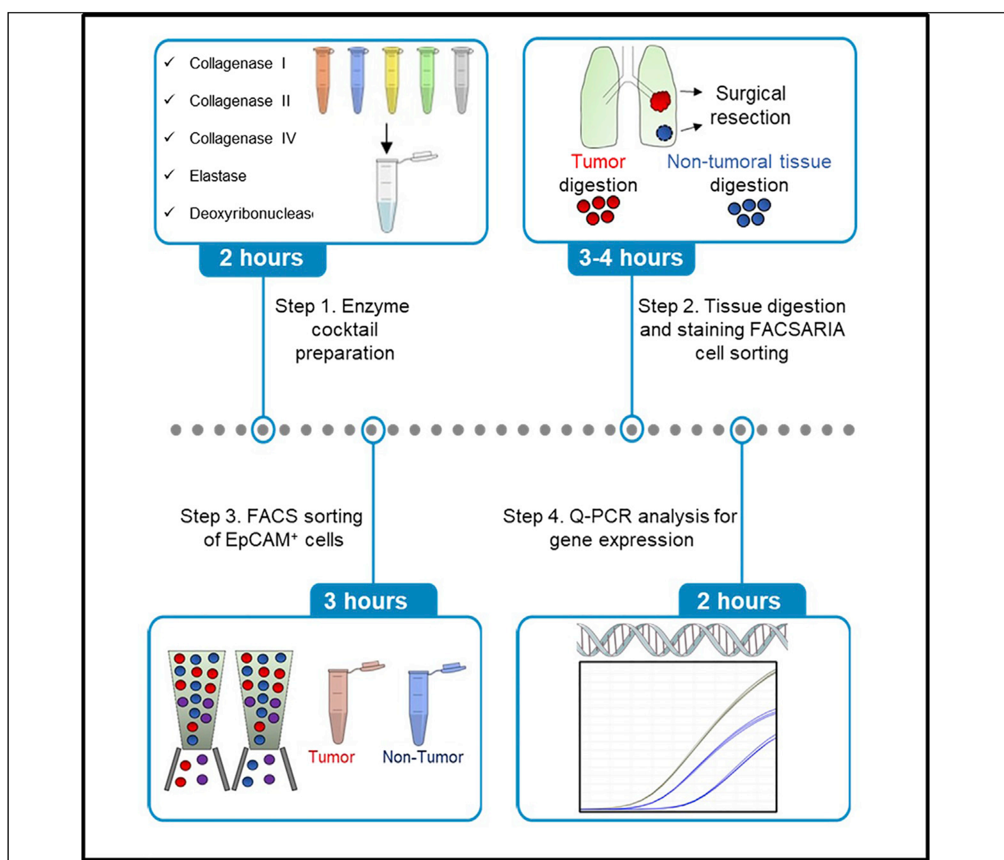
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for fluorescence-activated cell sorting of human EpCAM$^+$ lung cancer cells for gene expression analysis of Rac guanine-nucleotide exchange factors



Here, we describe a protocol for fluorescence-activated cell sorting (FACS) of human EpCAM$^+$ cells from fresh surgically resected specimens. We then use Q-PCR to identify specific molecular targets associated with the metastatic phenotype. This combined approach enables a qualitative and quantitative gene expression analysis of lung cancer samples. We describe how to use the protocol for Rac GEFs, but it can be applied broadly to other molecular targets.

Neil T. Sullivan, Mariana Cooke

marcooke@pennmedicine.upenn.edu

Highlights

Tumoral and adjacent non-tumoral human lung tissue enzymatic digestion protocol

Fluorescence-activated cell sorting of EpCAM$^+$ cells from surgically resected specimens

Q-PCR analysis of specific molecular targets associated with the metastatic phenotype

Rac-GEF expression analysis in isolated human lung cancer EpCaM$^+$ cells

Sullivan & Cooke, STAR Protocols 3, 101367
June 17, 2022 © 2022 The Author(s).
https://doi.org/10.1016/j.xpro.2022.101367

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for fluorescence-activated cell sorting of human EpCAM$^+$ lung cancer cells for gene expression analysis of Rac guanine-nucleotide exchange factors

Neil T. Sullivan[1,4] and Mariana Cooke[2,3,5,*]

[1]Division of Thoracic Surgery, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, PA 19104, USA

[2]Department of Systems Pharmacology and Translational Therapeutics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104, USA

[3]Department of Medicine, Einstein Medical Center Philadelphia, Philadelphia, PA 19141, USA

[4]Technical contact

[5]Lead contact

*Correspondence: neiltsullivanpenn@gmail.com (N.T.S.), marcooke@pennmedicine.upenn.edu (M.C.)
https://doi.org/10.1016/j.xpro.2022.101367

SUMMARY

Here, we describe a protocol for fluorescence-activated cell sorting (FACS) of human EpCAM$^+$ cells from fresh surgically resected specimens. We then use Q-PCR to identify specific molecular targets associated with the metastatic phenotype. This combined approach enables a qualitative and quantitative gene expression analysis of lung cancer samples. We describe how to use the protocol for Rac GEFs, but it can be applied broadly to other molecular targets.
For complete details on the use and execution of this protocol, please refer to Cooke et al. (2021) and Quatromoni et al. (2015).

BEFORE YOU BEGIN

Lung adenocarcinoma is the most prevalent histologic subtype of non-small cell lung cancer (NSCLC) and a major contributor to cancer-related deaths worldwide. Most common oncogenic alterations in lung adenocarcinoma occur in KRAS, receptor tyrosine kinases (RTKs, e.g., EGFR) and tumor suppressor genes (e.g., p53), which ultimately lead to aberrant activation of oncogenic and metastatic signaling cascades. Among the key signaling nodes downstream of RTKs, we identified Rac guanine-nucleotide exchange factors (Rac-GEFs) as paramount players of NSCLC cell motility and invasion. Specific Rac-GEFs were found to be markedly up-regulated in NSCLC, and in some cases their deregulated expression associates with poor patient prognosis (Cooke et al., 2021).

Surgical approach is the mainstream procedure and the standard of care for medically operable patients with early-stage NSCLC. Minimally invasive techniques such as video-assisted thoracoscopic surgery (VATS) and robot-assisted thoracoscopic surgery (RATS) have developed within the past two decades; however, most lung resections continue to be accomplished via a thoracotomy (Ettinger et al., 2021; Yeo et al., 2020; Upham and Onaitis, 2018; Baig et al., 2021). NSCLC tumor cells express epithelial cell adhesion molecules (EpCAM), an established epithelial cell marker used in pathological examinations (Baeuerle and Gires, 2007; Osta et al., 2004). Here, we describe a method for isolation of EpCAM$^+$ cancer cells from fresh surgically resected specimens from lung cancer patients followed by the subsequent identification of specific molecular targets. In our recent study (Cooke et al., 2021), we successfully determined the expression of Rac-GEFs in the purified EpCaM cells using Q-PCR. The current protocol embodies an efficient and valuable tool towards quali-quantitative gene expression analysis in isolated EpCAM$^+$ tumor cells and subsequent molecular analysis.

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

General preparation

The protocol described here was achieved by a close collaborative agreement within a thoracic surgeon and the voluntary participation of patients undergoing scheduled surgical resection of NSCLC tumors at the University of Pennsylvania. Prior to tissue procurement, all enrolled patients signed an Institutional Review Board (IRB) informed consent document approved by the University of Pennsylvania, by which patients gave their specific consent to obtain a portion of tumor and non-tumoral tissue specimens to be used for research purposes. All patients selected for entry into the study met the following criteria: (a) histologically confirmed stage I-II adenocarcinoma (early stage), (b) had no prior chemotherapy or radiation therapy, (c) no other solid malignancies, and (d) were negative for blood borne pathogens (such as HIV, HBV and HCV). Despite all patients being screened and tested negative for blood borne pathogen, given the nature of working with human samples, all specimens were handled under BSL2 setting with appropriate PPE and materials were properly discarded in accordance with university regulations.

Enzyme cocktail preparation for lung-tissue sample digestion

Timing: 2 h

Note: The enzyme cocktail was optimized for enhanced cellular yield from human lung tissue samples while maintaining high cellular viability and cell surface marker expression. Although other enzyme cocktails and digesting buffers are commercially available, a previous study demonstrated that their usage reached lower cellular yield, higher cell death and to varying degrees, affected composition of crucial cell surface receptors (Quatromoni et al., 2015). Even though this protocol specifically describes the methodology for lung human tissue digestion followed by the isolation of tumoral EpCAM$^+$ cells, it can be easily adapted to an expanded spectrum of human solid malignancies (e.g., mesothelioma, head and neck tumors, etc) with the subsequent isolation of further tumor infiltrating cellular components (i.e., T-cells, neutrophils, macrophages, etc). All the following steps should be performed under sterile conditions in a biosafety cabinet.

Final concentration for all enzyme cocktail components:

L-15 Leibowitz media.

Collagenase I (170 mg/L).

Collagenase II (56 mg/L).

Collagenase IV (170 mg/L).

Elastase (25 mg/L).

Deoxyribonuclease I (25 mg/L).

1. Dissolve the lyophilized collagenases I, II, and IV individually in 40 mL of L-15 culture media in a 50 mL conical followed by vigorous vortexing for 1–2 min. In the same way, dissolve deoxyribonuclease I and elastase.
2. Once dissolved, transfer the enzymes into the opened 300 mL L-15 culture media bottle (the 200 mL from this 500 mL bottle was removed to reconstitute the enzymes) and add the deoxyribonuclease I and elastase to obtain a ''stock'' concentrated enzymes mixture. Incubate the mixture for 1 h at 20C. The total final volume of this ''stock'' concentrated enzyme mixture is 500 mL.
3. Transfer 100 mL of the stock concentrated enzyme mix into five separate 500 mL L-15 media flasks and add 5 mL of penicillin/streptomycin to each flask. The stock concentrated enzyme mix is now

distributed among 3 L of L-15 to obtain final enzyme mix. Filter with a 0.22 μM pore-size membrane and store individual 25 mL aliquots at -80C. Enzyme cocktails can be used up to one year.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies (1:1,000 dilution)		
Alexa Fluor 488 anti-human CD326 (EpCAM)	BioLegend (clone 9C4)	Cat# 324210
Brilliant Violet 785™ anti-human CD45	BioLegend (clone 2D1)	Cat# 304048
Biological samples		
Human lung tissue and tumor tissue	University of Pennsylvania	N/A
Chemicals, peptides, and recombinant proteins		
Collagenase I (170 mg/L)	Worthington Biochemical	Cat# LS004196
Collagenase II (56 mg/L)	Worthington Biochemical	Cat# LS004174
Collagenase IV (170 mg/L)	Worthington Biochemical	Cat# LS004188
Deoxyribonuclease I (DNAse) (25 mg/L)	Worthington Biochemical	Cat# LS002139
Elastase (25 mg/ L)	Worthington Biochemical	Cat# LS002292
L-15 Leibowitz media (6 bottles)	VWR, Hyclone	Cat# SH30525.01
DMEM: F12 media	VWR, Hyclone	Cat# SH30023.01
100× Penicillin/streptomycin (100 mL)	Life Technologies	Cat# 15140-122
HyClone™ Fetal Bovine Serum (United States), Embryonic Stem (ES) Cell Screened, HyClone products (Cytiva)	VWR	Cat# 82013-578
BSA 30% in DPBS	Sigma-Aldrich	Cat# A9576
FACS stain buffer (BSA)	BD Biosciences	Cat# 554657
eBioscience Fixable Viability Dye eFluor™ 450	Thermo Fisher Scientific	Cat# 65-0863-14
RBC lysis buffer, 10×	Santa Cruz Biotechnology	Cat# 296258
Parafilm paper	Fisher Scientific, Bemis™	Cat# PM999
Trypan blue	VWR	Cat# C838R90
L-Glutamine	Corning	Cat# 25-005-CL
UltraPure 0.5 M EDTA, pH8.0	Thermo Fisher Scientific / Invitrogen	Cat# 15575020
1× DPBS without Ca2+ Mg2+	Gibco	Cat# 14190-136
Critical commercial assays		
RNeasy Mini Kit (250)	QIAGEN	Cat# 74106
QIAshredder (250)	QIAGEN	Cat# 79656
RNase-free DNase Set (50)	QIAGEN	Cat# 79254
RNase free water	QIAGEN	Cat# 1018017
Random Hexamers	Invitrogen	Cat# N8080127
TaqMan Reverse Transcription	Applied Biosystems	Cat# N8080234
0.2 mL Thin walled 8 tube & flat cap strips	Thermo Fisher Scientific	Cat# AB-1182
SimplyAmp Thermal Cycler	Thermal Cycler-Applied Biosystems	Cat# A24811
Probes	Thermo Fisher Scientific (TaqMan® Assays)	N/A
TaqMan® Fast Advanced Master Mix (2×)	Thermo Fisher Scientific	Cat# 4444556
Nuclease free water	QIAGEN	Cat# 129114
MicroAmp® EnduraPlate™ Optical 96-Well Clear Reaction Plate with Barcode	Applied Biosystems	Cat# 4483354
MicroAmp™ Optical adhesive film	Applied Biosystems	Cat# 4311971
QuantStudio 3	Applied Biosystems	Cat# A28137
Dead cell removal kit	Miltenyi Biotec	Cat# 130-090-101
Software and algorithms		
BD FACSDiva™ software	BD Biosciences	https://www.bdbiosciences.com/en-us/products/software/instrument-software/bd-facsdiva-software#Overview
Flowjo	BD Biosciences	https://www.bdbiosciences.com/en-us/products/software/flowjo-v10-software
Other		
QuadroMACS Separator and Starting kits	Miltenyi Biotec	Cat# 130-090-976

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
Surgical scissors	Fine Science Tools	Cat# 14568-12
Forceps	Fine Science Tools	Cat# 11627-12
Whatman Benchkote Plus absorbent protector sheets, 50 × 60 × m	Tisch Scientific	Cat# 2301-6150
Shaking incubator 37C	Corning	Cat# 6750
Sterile FACS 5 mL tube	Falcon	Cat# 352063
Filters 70 μm	Fisher Scientific	Cat# 22-363-548
Pre separation Filter (30 μm)	Miltenyi Biotec	Cat# 130-041-407
Centrifuge 7 inches diameter rotor		N/A
Stericup Quick release Millipore Express Plus 0.22 μM Filter Flask	Millipore	Cat# S2GPU05RE
37C water bath		N/A
100 mm TC-treated culture petri dishes	Corning	Cat# 30293
UltraComp eBeads™ Plus Compensation Beads	Invitrogen	Cat# 01-3333-42
Accudrop beads	BD Biosciences	Cat# 661612
FACSAria Fusion sorter	BD Biosciences	N/A

MATERIALS AND EQUIPMENT

Complete cell culture media		
Reagent	Final concentration	Volume
DMEM: F12 media	N/A	440 mL
Heat Inactivated FBS	10%	50 mL
L-Glutamine	1%	5 mL
Penn/Strep	1%	5 mL
Total	N/A	500 mL
Media should be stored at 4C.		

MACS isolation buffer		
Reagent	Final concentration	Volume
1× DPBS without Ca^{2+}, Mg^{2+}	N/A	981 mL
BSA	0.5%	5 mL
0.5 M EDTA	2 mM	4 mL
Penn/Strep	1%	10 mL
Total	N/A	1,000 mL

Note: Buffer should be stored at 4C (make new monthly).

Wash buffer		
Reagent	Final concentration	Amount
RPMI	N/A	970 mL
Penn/Strep	1%	10 mL
FBS	2%	20 mL
Total	N/A	1 L

Note: Buffer should be stored at 4C (make new monthly).

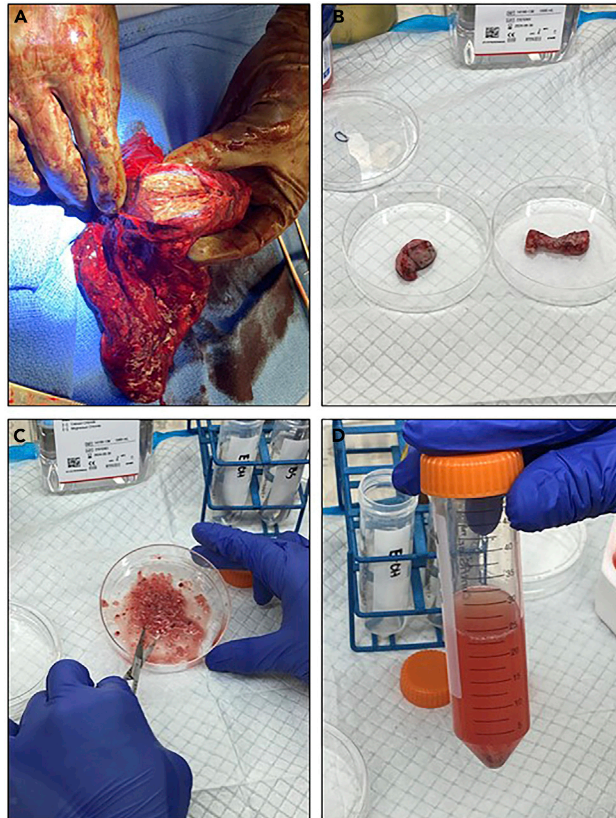


Figure 1. Workflow for digestion of tumoral and non-tumoral lung tissues for EpCAM$^+$ cells isolation

(A) Tumor tissue and non-tumor tissue specimens procured from the operating room.

(B) Pieces of tumor and non-tumor tissues placed in a petri dish before mincing.

(C) Minced tissue to demonstrate the homogenous mixture needed.

(D) Minced tissue resuspension with 1× enzyme cocktail mixture.

STEP-BY-STEP METHOD DETAILS

Tumor and adjacent non-tumoral lung tissue digestion

Timing: 3.5–4 h

Note: This section outlines the critical steps needed for efficient tissue digestion. Following these steps will generate high-quality single cell digest, which is essential for isolation of EpCAM$^+$ cells. Fresh non-tumoral and tumor tissue specimens should be obtained from the operating room in a sterile specimen container with sterile saline (Figure 1A). Processing of tissue should commence within the first 15 min after tissue procurement.

Note: To obtain high cell yield and viability, tissue should be processed within 15 min of procuring it. However, if this is not possible, tissue should be fully submerged in 4C wash media. In doing so, cell yield and viability will likely be compromised. We don't recommend letting tissue sit at 4C longer than 2 h.

1. Thaw a tube of stock enzyme cocktail described in the previous section in a 37C water bath for approximately 10 min. Dilute the enzyme cocktail mixture in 25 mL of L-15 culture media to obtain a final volume of 50 mL. This dilute enzyme is our working stock with final concentrations indicated in section above. The volume of enzyme cocktail needed is proportional to the tissue mass, with an approximate ratio of 25 mL per 0.5 grams of tissue sample.

2. Place tissue specimens in separate petri dishes for weighting and processing (Figure 1B). Add 10 mL of 1× DPBS to each dish. Using forceps, gently press tissue samples and remove remaining air/blood (non-tumoral tissue sample is usually highly vascularized). For tumor tissue, all areas of visible tissue necrosis must be trimmed away and discarded.
3. Transfer the tissue specimens to a new petri dish for mincing. Use micro-dissecting scissors to cut up tissues into progressively smaller pieces, aiming a size range of 1–2 mm^3.
4. Continue until the mincing process gives a homogenous mix with no remaining large tissue pieces (Figure 1C). This step should be performed quickly to avoid drying of the tissue samples.

   Note: If large tissue begins to dry out (pink smearing on petri dish), add a small volume (500 μL-2 mL) of enzyme cocktail mixture to keep the tissue samples hydrated.

5. Add 25–30 mL of enzyme cocktail mixture to each petri dish.

   Note: If the tissue does not smoothly pass into the opening of a 25 mL serological pipette, it means that it has not been minced thoroughly. If proper mincing is not reach, proper digestion will be impacted, and this will significantly diminish cell yield and viability.

6. Transfer the tissue and enzyme cocktail mixture into a new 50 mL tube (the final volume of tissue it should not exceed the 7.5 mL mark on the conical tube; if it exceeds this mark, add more enzyme cocktail mixture, and split the sample evenly into two 50 mL tubes). (Figure 1D).

   Note: Not surpassing the 7.5 mL mark on the conical tube with total tissue is an important aspect for tissue digest. The final volume of enzyme in the tube is 25–30 mL, and the tissue within that 25–30 mL should not settle above the 7.5 mL demarcation on the tube, because this allows for optimal tissue to enzyme ratio. If the tissue exceeds this mark, there will be too much tissue and not enough enzyme for proper digestion.

7. Incubate the 50 mL tubes in a shaker incubator (1 × *g*) for 45 min at 37C.
8. Remove tissue samples from incubator, using a 10 mL serological pipette to disaggregate the solid component by pipetting up and down against the inside edge of the tube. Repeat 15–20 times, depending on tissue hardness.
9. Place the tubes back in the shaker incubator for 1 h at 37C.
10. Repeat step 8.
11. Filter tissue sample using a 70 μm pore-size filter. Gentle pressure is then applied with the black rubber bottom of a 5 mL syringe (Fisher Scientific) to any remaining, partially digested tissue on the cell strainer. 25 mL of DPBS is used to wash the filter one last time. Typically, <5% of the tissue (consisting of mainly acellular-connective tissue) remains on the strainer at the end of this step. Centrifuge at 252 × *g* for 5 min.
12. Prepare RBC lysis mixture: dilute 5 mL of 10× RBC lysis with 45 mL sterile water at 20C.
13. Discard the supernatant and resuspend pellets in 25 mL 1× RBC lysis mixture and then incubate for 9 min at 20C.
14. Add 10 mL of 1× DPBS and centrifuge at 252 × *g* for 5 min.
15. Discard supernatant and resuspend in 5–10 mL of L-15 culture media.
16. Count cells and determine the percentage of live cells using trypan blue.

   Note: After digestion, cell viability should be ~80–100%. If total cell viability is below 80%, use a "dead cell removal kit". Follow manufacture protocol as stated using the *QuadroMACS Separator and Starting kits* (https://www.miltenyibiotec.com/upload/assets/IM0001218.PDF). For this enzyme cocktail, it is expected that the processed tissue will yield ~6×10^6 cell per gram of tissue. This cocktail is optimal to have minimal effect on the cell surface receptors of the major immune cell populations and preserve their functionality. See Quatromoni et al. (2015) for more information.

17. Place samples on ice or proceed immediately with cell surface staining.

Flow surface staining for EpCAM$^+$ cells

Timing: 1 h

Note: All steps should be performed under sterile conditions in a laminar flow hood; samples should be kept on ice. Protect from light.

18. Transfer 1–10 million single cell suspension into a 5 mL FACS tube and add 3 mL 1× DPBS, followed by centrifugation at 252 × *g* for 5 min.
19. Aspirate and discard supernatant.
20. Prepare a 1/1,000 dilution of Live Dead Aqua Fixable Stain or Fixable Viability Dye eFluo with 1× DPBS. Resuspend sample suspension in 100–200 μL of 1× viability stain and incubate for 5 min.
21. Add 3 mL 1× DPBS and centrifuge at 252 × *g* for 5 min.
22. Add 50 μL of staining buffer and for every 1 million cells add 1 μL of EpCAM antibody (Alexa Fluor 488) and 1 μL of CD45 antibody (BV785). Incubate 30 min on ice.
23. Wash 3 times with 3 mL of 1× DPBS, with centrifugation at 252 × *g* for 5 min between washes. Aspirate supernatant and resuspend sample in 500 μL of 1× MACS buffer.
24. Filter sample with a pre-separation filter (30 μm pore size membrane) before running on cytometer to prevent clogs.

EpCAM$^+$ cells sorting by BD FACSAria™ fusion flow cytometer

Set up the FACSARIA fusion sorter

Timing: 3 h

Note: This is an operator-dependent procedure. Core facilities with trained personnel might be required. FACSARIA Fusion sorter runs on FACSDiva software.

25. Turn on cytometer, computer, compressor, and AMS system, and set temperature at 4C. Start the BD *FACSDiva™ software*.
26. Settings: cytometer tab, select the nozzle size and correct configuration. Select and install the 100 μm nozzle and holder for 2-way sorting. Insert sterile, 5 mL FACS collection tube containing 500 μL of MACS buffer (so viable EpCAM$^+$ cells are not sorted into dry FACS tubes).
27. Align the stream/droplets with the waste drawer. Next, set the stream breakoff point and ensure there is proper droplet formation. For 100 μm nozzle, set sorting parameters as followed (values are starting points):
    a. Amplitude: 12.
    b. Frequency: 32.
    c. Drop 1: 150.
    d. Gap: ∼12.
28. Next perform Accudrop to calculate proper drop delay. Create a master mix tube of beads by adding 1 drop of Accudrop beads to 500 μL 1× DPBS/staining buffer.
29. Load the FACS tube of Accudrop beads and turn on the optical filter in the 100 μm side stream, aiming for an event rate of 1,000–1,500 events/sec. Keep the waste drawer in position. Run the automatic accudrop delay program.
30. Set up a new FACS sorting experiment for EpCAM$^+$ cells isolation.
31. Under cytometer settings, select the colors needed for the sort. The violet and blue laser will be needed to detect Live Dead Aqua Dye, EpCAM (Alexa Fluor 488/FITC) and CD45 (BV785) fluorochrome. The violet laser will excite at 405 nm; live dead aqua will emit 525/50 and BV785 emit at 780/60. The blue laser will excite at 488 nm and Alexa Fluor 488/ FITC will emit at 530/30.
32. Set up compensatory controls and calculate compensation matrix.

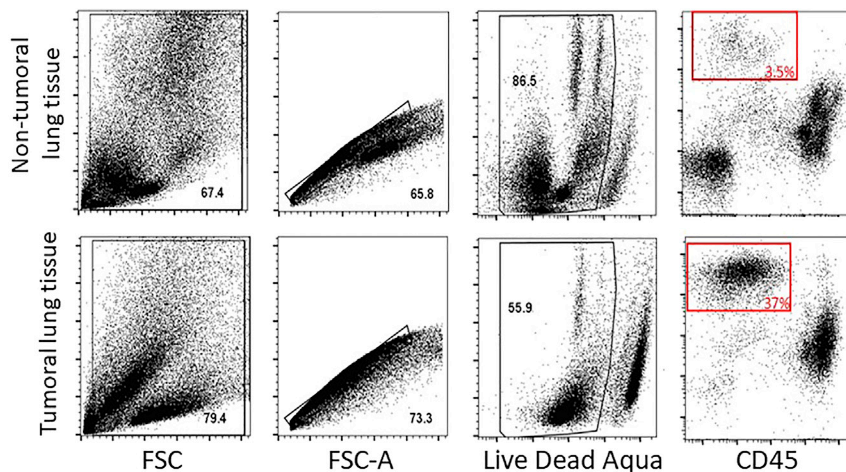


Figure 2. EpCAM+ FACS sorting gating strategy on the BD FACSAria cell sorter

To FACS sort EpCAM+ cells from non-tumoral and tumor tissue, all events, minus debris, were gated on Forward scatter (FSC) and Side scatter (SSC) plots. Following this, cells were discriminated and gated on only live cells. Lastly, a EpCAM *vs.* CD45 plot was made and only Live, EpCAM+CD45- cells were sorted (indicated in the *red* box).

33. Create plots for analysis and sorting (Figure 2).

   Note: There should be 4 plots: a) forward scatter by side scatter (FSC vs SSC), b) forward scatter area by height (FSC-H vs. FSC-A), c) side scatter by live dead aqua (SSC vs. live dead aqua), and d) EpCAM (FITC) vs. CD45 (BV785).

34. Collect at least 10,000 EpCAM+ cells to ensure enough yield for obtaining high quality RNA. (By using the FACSAria Fusion sorter, we are able to isolate EpCAM+ cells, with purity of ∼99%).
35. Once sorting is completed, wash FACS tube with 5 mL 1× DPBS.
36. Place cell in a sterile Eppendorf tube and wash cell pellet twice with 1× DPBS. You can store samples at -80 C as a dry pellet for 16–18 h or proceed immediately to RNA extraction.

RNA extraction (based on QIAgen RNA extraction kit)

Timing: 3 h

Note: All steps should be carried out at 20C.

Prepare in advance following solutions:

70% Ethanol in RNase-free water (3.5 mL ethanol + 1.5 mL water).

RLT buffer + 1% β-mercaptoethanol (10 mL RLT + 100 μL β-mercaptoethanol).

60 μL DNase aliquots should be stored at 4C. At the time of the experiment, mix an aliquot with 420 μL RDD buffer.

37. Add 350 μL RLT buffer to sample. Sample can be stored at -80C if needed, as opposed to dry pellet.
38. Transfer sample to a QIAshredder column (purple). Centrifuge 2 min at 12,000 × *g*.
39. Remove the column.
40. Add 350 μL 70% ethanol to the sample.
41. Transfer all volume to the QIAgen column (pink). Centrifuge 1 min at 12,000 × *g*.

42. Discard the supernatant and keep the column.
43. Add 350 μL RW1 buffer to the pink column. Centrifuge 15 s at 12,000 × *g*.
44. Discard the supernatant and keep the pink column for subsequent steps.
45. Mix 10 μL DNase per column. Let it sit on the column for 15 min at 20C.
46. Add 350 μL RW1 buffer directly to the pink column. Centrifuge for 15 s at 12,000 × *g*.
47. Discard the supernatant and keep the column.
48. Add 500 μL RPE directly to the column. Centrifuge for 1 min at 12,000 × *g*.
49. Repeat step #47 and discard the supernatant.
50. Now place the column into a new collection tube. Centrifuge 2 min at 12,000 × *g*.
51. Discard supernatant and add 50 μL RNase-free water. Incubate 2 min at 20C.
52. Centrifugate 2 min at 12,000 × *g*.
53. This is now your high-quality RNA collect from FACS sorted $EpCAM^{+}$ cells.
54. For quality check of RNA integrity, quantify RNA by spectrometry measuring 1 μL/sample in duplicates, and assess purity by A260/280 nm ratio (it should be >1.8). Calculate μL/μg RNA per sample.

Q-PCR protocol

Reagent	Volume/well (μL)
TaqMan® Fast Advanced Master Mix (2×)	10
Probe	1
cDNA template	9
Final volume	20

cDNA synthesis

Timing: 1 h

Note: All reagents should be kept on ice.

55. cDNA synthesis was performed following standard protocol for Taqman Reverse Transcription reagents kit.
56. Prepare Mix1 for cDNA conversion as indicated below:
    a. RNA: 1 μg (complete with RNase free water to obtain final volume of 6.6 μL).
    b. Random hexamers 1 μL.
    c. Final volume: 7.7 μL.
    d. Add Mix #1 into the PCR thermocycle and initiate denaturalization phase. For random hexamers, the thermocycler settings are as followed; 25C for 10 min, 37C for 30 min, 95C for 5 min, and a 4C hold.
57. Leave cDNA in thermocycler at 4C while Mix #2 is prepared. Mix #2 components are indicated below.
    a. TaqMan buffer 10×: 2 μL.
    b. dNTPs: 4 μL.
    c. $MgCl_2$: 4.4 μL.
    d. RNase inhibitor: 0.4 μL.
    e. RT: 0.5 μL.
    f. Final volume: 11.3 μL.
58. Add 11.3 μL of Mix #2 to each tube containing Mix #1, reaching a final volume of 20 μL. The cDNA and master mix for q-PCR has been made. The only component omitted from this mix is the specific target probes (Rac-GEFs and control probes). Proceed to PCR cDNA amplification steps for Rac-GEFs in the next section.

Q-PCR for Rac-GEFs in 96-well plates

Timing: 2 h

Note: While this protocol was used primarily for determination of Rac-GEF expression, it can be easily adapted for measuring other genes of interest in Epcam$^+$ cells.

59. Assure proper compatibility between the 96-well plate and the Q-PCR system.
60. Thaw the TaqMan® Fast Advanced Master Mix (2×) reagent and the desired probes for specific Rac-GEFs (or other genes of interest) on ice.
61. Prepare a mix with the following components in the indicated proportion. Add 10% overage for pipetting loss.
62. Gently vortex (avoid vigorous mixing).
63. Spin down the mix tube to eliminate bubbles/air in a bench microcentrifuge.
64. Add 19 μL the mix to each well.
65. Add 1 μL of cDNA template (1–100 ng) to each well.
66. Seal the plate with the Optimal adhesive film (avoid bubbles/air trapping by applying compression).
67. Spin down the 96-well plate in a macro-centrifuge at maximum velocity.
68. Set up experimental conditions in the Real-Time PCR system as follows in PCR cycling conditions table.
69. Run the reaction.
70. Analyze data (see below).

PCR cycling conditions

Steps	Temperature	Time	Cycle
Initial denaturation*	95C	2 min	1
Denaturation	95C	3 s	40
Annealing**	60C	30 s	
Extension	72C	10 min	
Final extension	72C	10 min	1
Hold	4C	Infinite	

Settings were determined as "Fast" mode based on our laboratory equipment QuantStudio™ 6 and 7 Flex Real-Time PCR System.
*UNG (uracil-containing DNA) activation,
***Select anneal temperature based on T_m of the primers.*

EXPECTED OUTCOMES

The protocol described here provides a reliable method for EpCAM$^+$ cells isolation based on both tumor tissue and adjacent normal tissue digestion, followed by FACS analysis. The purified EpCAM$^+$ cells could be easily subjected to several molecular biology approaches for gene expression analysis including RNA Seq, single cell expression gene assay by quantitative RT-PCR, and multiple gene screening using 96-well plates, either commercial or customized. On average, for stage I and II lung adenocarcinoma patients, the recovered tissue mass weights from 0.1 to 4 grams, with a yield ranging from <1 million to >100 million cells for a single cell suspension. For more information, see the previous published study by Quatromoni et al. (2015). To obtain high quality RNA for examining multiple targets, at least 10,000 Epcam+ cells are needed to be isolated.

QUANTIFICATION AND STATISTICAL ANALYSIS

Q-PCR analysis

Analytical approach differs within conventional *vs.* Q-PCR. For conventional PCR, amplicon is detected by end-point analysis, while in Q-PCR amplicon accumulation is measured and detected as

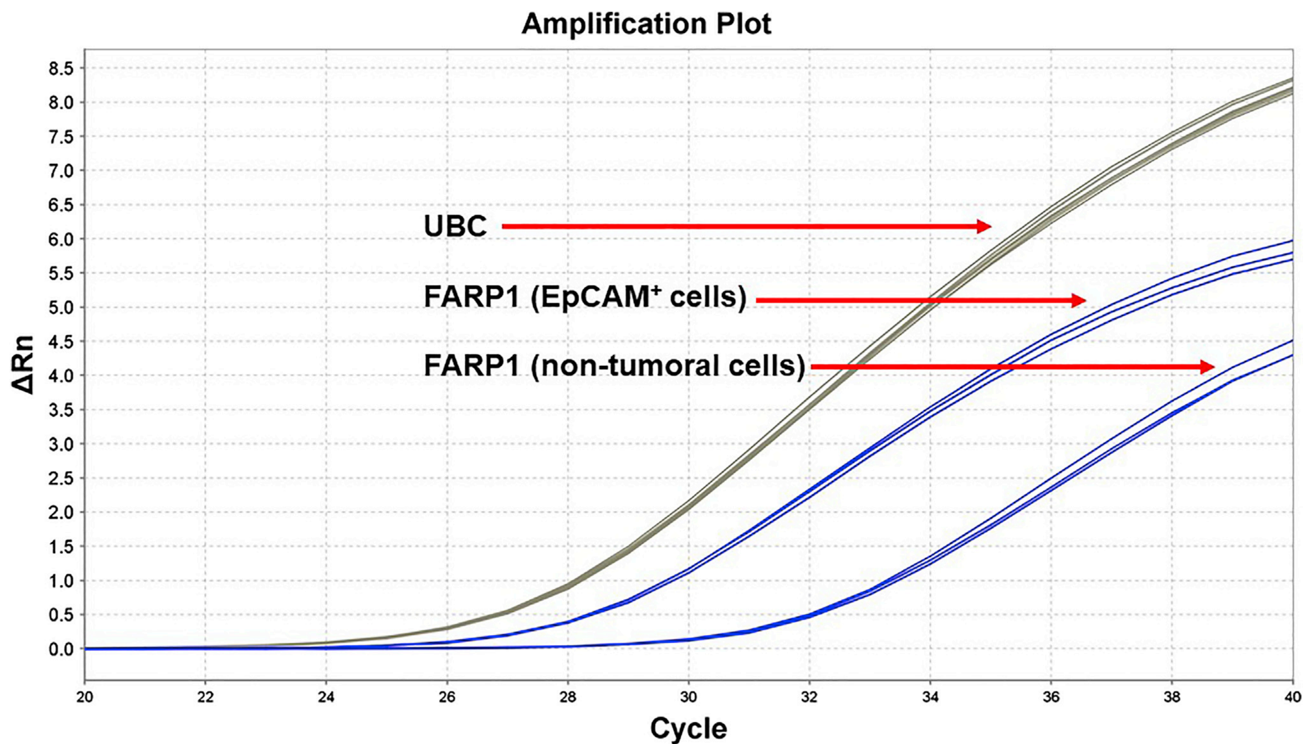


Figure 3. Plot displays a graph showing a representative Q-PCR amplification curve in a linear scale
Each curve represents different genes (reference gene in *gray*, and FARP1 target gene in *blue*). Each gene was analyzed in either triplicate.

the reaction progresses. Plot depicted in Figure 3 displays a graph showing a typical Q-PCR amplification curve. The proportional relationship between the fluorescence of the amplification reaction and the amplified product is shown on the *y* axis, while PCR cycle numbers are shown in the *x* axis. Fluorescence in the initial cycles stay at background readings until enough amplification of the product accumulates and generates a detectable signal. The cycle number at which fluorescence of the accumulated product reaches a detectable signal is known as the ''threshold cycle'' (C_T). C_T values are inversely proportional to the amount of amplified product.

Q-PCR analysis can be accomplished through two different methods (absolute and relative) based on experimental purposes. Briefly, for absolute quantification (i.e., gene copy number) a standard curve is made to interpolate the quantity of the amplification product. Instead, relative quantification is used to compare levels of gene expression, particularly addressing the *fold-change* for gene expression across different experimental conditions. To achieve relative quantification analysis, samples must be normalized against either the unit mass or a reference gene. We have conducted our studies using the later method. The reference gene should be a well-known gene whose expression is stable, i.e., it does not change expression upon any experimental condition *(''housekeeping gene'')*. The most common method for relative expression analysis is the Livak or $\Delta\Delta C_T$ method. To proceed with this analysis, both target and housekeeping genes should have around 100% efficiency of amplification and less than 5% of variation between each other. Essentially, there are three steps to follow: a) normalization of C_T of target genes/C_T reference gene; b) normalization of ΔC_T of treated sample/ΔC_T of control sample; and c) calculation of the fold-change of expression. All these are exemplified in Q-PCR analysis table.

Note: When EpCAM$^+$ cell isolation was initially attempted by means of a standard magnetic beads technique, the Q-PCR analysis for Rac-GEFs gene expression gave misleading results.

For example, we found increased expression of PREX1 (a Rac-GEF classically overexpressed in neutrophils) in 5 different patient's tumor samples using the magnetic bead approach, whereas no express was found using the FACS purification approach described above. Therefore, this was an indication that our magnetic bead purified samples were contaminated with neutrophils, and potentially with other $CD45^+$ cells, which made the sample not suitable for the desired gene expression assays.

Q-PCR analysis

	Gene	
Tissue sample	C_T target (FARP1)	C_T reference (UBC)
Normal	30.35	29.61
Tumor	27.52	29.69
ΔC_T Normal	C_T Target (FARP1) - C_T Reference (UBC)	
ΔC_T Normal	30.35–29.61 = 0.74	
ΔC_T Tumor	C_T Target (FARP1) - C_T Reference (UBC)	
ΔC_T Tumor	27.52–29.69 = −2.17	
ΔC_T Tumor - ΔC_T Control	−2.17–0.74 = −2.91	
Normalized expression ratio= $2^{\wedge-\Delta\Delta CT}$		
$2^{-(-2.91)}$ = 7.52		
Tumor cells express FARP1 **7.52-fold** higher level than control cells		

LIMITATIONS

A recognized limitation is that the $EpCAM^+$ cells yield may largely depend on the different intrinsic tumor properties within the subject population, such as disease stage, central tumor necrosis and histologic type of the tumor (i.e., higher tumor stages, larger central necrotic tissues, and certain histological types such as squamous cell carcinoma, would render lower yield of $EpCAM^+$ cells). Furthermore, our approach cannot be subjected to significant technical changes, such as replacement of the FACS methodology for any other standard purification techniques like the "magnetic beads" approach, since this will drastically decrease the purity of tumor cells yielded and introduce significant biases for gene expression purposes. For instance, the sample will suffer deleterious contamination with blood cells ($CD45^+$ cells), which could jeopardize the experimental stringency of the molecular analysis. Although technically challenging, the FACS approach will assure maximum purity of the $EpCAM^+$ cells population, thus allowing a reliable and straightforward use of this preparation for any systematic and desired molecular gene expression approach. Another limitation relates to the size of the tumor and adjacent tissue used for experimental purposes, which is variable from patient to patient.

TROUBLESHOOTING

Problem 1

Tissue cannot be processed within 15 min of procuring tissue (step 1).

As stated in the tumor and adjacent non-tumoral lung tissue digestion section, tissue should be processed immediately, within 15 min of procuring it. If tissue isn't processed immediately, the viability and cell yield will suffer.

Potential solution

While it is not ideal, tissue can be stored short term, for up to 2 h in sterile wash buffer before processing if completely submerged and stored at 4C.

Problem 2

Low viability of $EpCAM^+$ cells (step 16).

We have stated in the tumor and adjacent non-tumoral lung tissue digestion section that there is patient-to-patient variation and tumor-to-tumor variation with respect to cell death. If there is a high level of cell death with respect to the single cell suspension, the user should use the dead cell removal kit. This will also remove dead $EpCAM^+$ cells. However, this will unbiasedly remove all dead cells, not just $EpCAM^+$ cells. To ensure we are only examining Rac-GEFs from viable, $EpCAM^+$ cells, we used the FACSAria Fusion sorted to remove all dead $EpCAM^+$ cells and only analyze viable ones.

Potential solution

Do not deviate from the protocol outlined above. $EpCAM^+$ cells should be isolated by FACS sorting and not magnetic beads or other means. Only FACS sorting with give purity of ~99% and all of them viable, as opposed to beads, which is not as pure and will give a mixture of live and dead cells.

Problem 3

Higher ratio of $CD45^+$ cells to $EpCAM^+$ cells (step 33).

As shown in Figure 2, the percentage of CD45+ cells is typically greater than $EpCAM^+$ cells. However, this protocol allows the user to identifying tumor specific changes in transcriptional profiles by comparing with non-cancerous adjacent $EpCAM^+$ cells. In some cases, $EpCAM^+$ cells are very low and it would be difficult to FACS sort enough pure $EpCAM^+$ cells to obtain high quality RNA for RNA gene expression analysis, especially with non-cancerous adjacent tissue.

Potential solution

To ensure you have enough sample to achieve a minimum of 10,000 $EpCAM^+$ cells, before using the FACSAria sorter, do a quick phenotype check of the single cell suspension as shown in Figure 2. Determine the frequency of $EpCAM^+$ cells in both distant and tumor digest and multiple this by total number of cells in single cell suspension. If the final number doesn't seem possible, do not proceed further with sorting. This will save valuable time and resources.

Problem 4

Clogging of 70 μm filters (step 11).

Tumors have a very heterogeneous nature, where some tumor single cell suspension have a high level of debris or are very mucus. Two examples can clog the 70 μm filter.

Potential solution

Multiple filters will have to be used during the filtering step, followed by using the plunger to apply gentle pressure to help pass through as much digested tissue as possible. In addition, using a pre-separation filter following RBC lysis will ensure a better starting material for flow cell surface staining for $EpCAM^+$ cells.

Problem 5

Elevated C_T within gene expression analysis (step 70).

Elevated C_T within gene expression analysis might be attributed to an inadequate storage of cellular pellet sample prior to molecular analysis approach, low yielded mRNA, poor cDNA quality and/or inadequate A260/280 nm ratio.

Potential solution

First, the cellular pellet should be stored at -80C for no longer than 24–48 h. Isolation of RNA immediately after FACS sorting is preferred. For mRNA isolation, we recommend that this is done within the first 24–48 h after the pellet is obtained. We recommend not to proceed with the gene expression analysis if A260/280 nm ratio of mRNA is below 1.8. Revise every step of the mRNA extraction

protocol and check if any step was oversighted or missed. Also, if possible, use more than one housekeeping gene (UBC, B2M, GAPDH) in the gene expression analysis. In many cases, it may be convenient to average more than one housekeeping gene if variability in their expression is observed.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Mariana Cooke (marcooke@pennmedicine.upenn.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

We acknowledge support by T32 CA009140 from the NIH to N.T.S. We also thank the Children's Hospital of Philadelphia Flow Cytometry core for FACSARIA sorting maintenance, daily quality control checks, and training.

AUTHOR CONTRIBUTIONS

Conceptualization, Manuscript Writing and Editing, Data Analysis, and Figure Preparation, N.T.S. and M.C.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Baeuerle, P.A., and Gires, O. (2007). EpCAM (CD326) finding its role in cancer. Br. J. Cancer *96*, 417–423. https://doi.org/10.1038/sj.bjc.6603494.

Baig, M.Z., Razi, S.S., Stroever, S., Weber, J.F., Connery, C.P., and Bhora, F.Y. (2021). Anatomic resection has superior long-term survival compared with wedge resection for second primary lung cancer after prior lobectomy. Eur. J. Cardiothorac. Surg. *59*, 1014–1020. https://doi.org/10.1093/ejcts/ezaa443.

Cooke, M., Kreider-Letterman, G., Baker, M.J., Zhang, S., Sullivan, N.T., Eruslanov, E., Abba, M.C., Goicoechea, S.M., Garcia-Mata, R., and Kazanietz, M.G. (2021). FARP1, ARHGEF39, and TIAM2 are essential receptor tyrosine kinase effectors for Rac1-dependent cell motility in human lung adenocarcinoma. Cell Rep. *37*, 109905.

Ettinger, D.S., Wood, D.E., Aisner, D.L., Akerley, W., Bauman, J.R., Bharat, A., Bruno, D.S., Chang, J.Y., Chirieac, L.R., D'Amico, T.A., et al. (2021). NCCN guidelines insights: non-small cell lung cancer, version 2.2021. J. Natl. Compr. Canc. Netw. *19*, 254–266. https://doi.org/10.6004/jnccn.2021.0013.

Osta, W.A., Chen, Y., Mikhitarian, K., Mitas, M., Salem, M., Hannun, Y.A., Cole, D.J., and Gillanders, W.E. (2004). EpCAM is overexpressed in breast cancer and is a potential target for breast cancer gene therapy. Cancer Res. *64*, 5818–5824. https://doi.org/10.1158/0008-5472.CAN-04-0754.

Quatromoni, J.G., Singhal, S., Bhojnagarwala, P., Hancock, W.W., Albelda, S.M., and Eruslanov, E. (2015). An optimized disaggregation method for human lung tumors that preserves the phenotype and function of the immune cells. J. Leukoc. Biol. *97*, 201–209. https://doi.org/10.1189/jlb.5TA0814-373.

Upham, T.C., and Onaitis, M.W. (2018). Video-assisted thoracoscopic surgery versus robot-assisted thoracoscopic surgery versus thoracotomy for early-stage lung cancer. J. Thorac. Cardiovasc. Surg. *156*, 365–368. https://doi.org/10.1016/j.jtcvs.2018.02.064.

Yeo, J.H., Shariati, N.M., Pelz, G.B., Dozier, J.M., and Rizk, N.P. (2020). Lean lobectomy: streamlining video-assisted lobectomy to increase the value of lung cancer care. JCO Oncol. Pract. *16*, e823–e828. https://doi.org/10.1200/JOP.19.00590.