Title
Microfluidic Electrophoretic Cytometry Device and Assay Development for Protein Analysis of Mammalian Cells

Author
Rosas, Elisabet

Publication Date
2018

Peer reviewed|Thesis/dissertation

Microfluidic Electrophoretic Cytometry Device and Assay Development for Protein Analysis of Mammalian Cells

By

Elisabet Rosas

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Joint Doctor of Philosophy
with University of California, San Francisco,

in

Bioengineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Amy E. Herr, Chair
Professor Clayton J. Radke
Professor Todd McDevitt

Summer 2018

Abstract

Microfluidic Electrophoretic Cytometry Device and Assay Development for Protein Analysis of Mammalian Cells

By

Elisabet Rosas

Joint Doctor of Philosophy
with University of California, San Francisco,
in Bioengineering

University of California, Berkeley

Professor Amy E. Herr, Chair

Understanding cell-to-cell phenotypic heterogeneity is crucial for elucidating the biological mechanisms of multicellular organisms. Cytometry, or the measurement of characteristics of cells, encompasses a myriad of tools and assays that measure properties from cell size and morphology to expression levels of mRNA or protein. Conventional protein measurement assays, however, generally lack the analytical sensitivity required for single-cell analysis. To fill this gap, microfluidic design utilizes length scales and time scales that enable performing perturbations and measurements at sub-cellular resolution. In this dissertation, we describe the design, development and optimization of electrophoretic cytometry tools that advance protein and nucleic acid measurement capabilities for single-cell and single-embryo analysis (from 1 to 100 s of cells).

We first examine the technical noise and reproducibility of the single-cell polyacrylamide gel electrophoresis assay, or single-cell PAGE. We then establish a novel method for quantification of surface receptor protein localization by integrating an upstream surface immunostaining step. We detect a shift in electrophoretic migration of the antibody-antigen complex, and demonstrate agreement of results between our assay and standard methods for measuring cell surface proteins (flow cytometry and immunofluorescence).

We then develop a high-specificity, multiplexed single-cell immunoblot for screening of primary, uncultured smooth muscle cells (SMCs) for a panel of maturation markers. An emerging theory for life-threatening vascular remodeling involves the existence of a subset of SMCs with immature-like phenotype. To scrutinize healthy vessels for immature-like SMCs, we first use numerical simulations to optimize microwell volume in order to minimize initial protein dilution at the *in situ* lysis step. We then perform electrophoresis for separation of the 34 to 227 kDa molecular mass range of target markers, and identify a subpopulation (< 10%) of immature-like SMCs. Our results support the recently-established mechanism that only a subset of immature-like SMCs is responsible for vascular remodeling.

Next, we focus on questions regarding the complex mechanisms governing mammalian development, such as the embryonic stage at which blastomeres first exhibit cell-fate commitment. By integrating embryo-specific sample preparation and single-embryo handling with scaled-up microfluidic immunoblotting designed for murine embryos, we measure proteins from all stages of mouse preimplantation development, from individual zygotes to single blastocyst blastomeres. Despite a lack of highly selective immunoreagents, we effectively interrogate inter-embryonic heterogeneity of embryo-specific isoforms involved in RNA-mediated gene expression. In dissociated two-cell and four-cell blastomeres, we detect significant inter-blastomeric variation in proteins implicated in cell cycle regulation, suggesting blastomere heterogeneity even in the earliest multicellular stage of preimplantation embryos. Further, 20-30 embryos recovered from a single mouse are sufficient for statistically relevant analyses, eliminating the need for sample pooling in preimplantation development studies.

Finally, we develop a novel method for dual protein-nucleic acid measurements from low starting cell numbers (from 10 s to 100 s of cells) aimed at assessing whether specific modifications in genomic DNA or alternative splicing events in preimplantation embryos translate generate different protein isoforms. To achieve the dual measurement, we fractionate cells into cytoplasmic and nuclear compartments, where cytoplasmic proteins are analyzed by electrophoretic separations and nuclei are retrieved for off-chip nuclei acid measurements. We demonstrate that the single-cell PAGE protein signal correlates strongly with protein expression prior to lysis, and measure both mRNA and DNA from retrieved nuclei. This method shows promise for determining whether the abundant splice variants and DNA modifications in preimplantation embryos translate to protein isoforms.

To My Family

Perquè ho doneu tot per mi.

Acknowledgements

Our Commencement speaker reminded us of the saying “if you see a turtle on a fence post, you know it had some help”. Just like that turtle, I would not be where I am now without the love, support and encouragement I have received from the people in my life.

My Ph.D. experience has been more than I could have ever dreamed of, and that is mainly because I chose the perfect advisor, Amy. Apart from showing us how to become curious and rigorous engineers, thank you for teaching me that it is never okay to be silent in the face of an injustice, whether it's standing up for others or your own beliefs, even when others may seem to have a louder voice. I have lost count of the times you have inspired us and shown us the power of believing in others and empowering the most disadvantaged. Thank you for your constancy, your fearlessness and for making this world a more kind and equal place.

I would like to thank my Dissertation and Qualifying Exam Committee members, Prof. Dorian Liepmann, Prof. Clay Radke, Prof. Lydia Sohn and Prof. Todd McDevitt. Thank you for lifting our spirits with your humor and your energy, and thank you for the support you have given me throughout these years. I would also like to thank my collaborators, Prof. Song Li and Prof. Lin He for working with me and for offering me your support and expertise.

I would like to thank the friends I could not have done this without. Marc: dono gràcies a Berkeley perquè m’ha regalat una de les coses més preuades; your friendship. I could simply copy and paste one of our conversations here to demonstrate how your sense of humor has gotten me through some of the hardest times of my Ph.D. Thank you for being the same person (one day we might find out which one of us is real) and bringing out the best in everyone. Monica and Ginny, thank you for giving me a family when I felt far away from mine. Elly, thank you for being my first mentor, both inside and outside of lab. Tiffany, I am so grateful that the Herr-Li collaboration placed you in my life. Thank you for our hour-long conversations (while we should have been running experiments!), your generosity and your brilliant thoughtfulness. Andrew, thank you for turning our collaboration into a reason to look forward to carrying heavy laboratory equipment across campus. Nathan, thank you for being the best roommate. Sophie and Henry thank you for shining your light in my life.

To the Herr Lab members, thank you for creating the diligent, inquisitive environment where science thrives. Hector, thank you for convincing me to talk to Amy- you changed my life! Julea, Burcu and Philippe, thank you for the fun and emotional conversations in our “corner” of the lab. Augusto, thank you for being my ‘unofficial mentor’. John, Rob, Qiong, Martha, Elaine, Sam, Anjali, Kevin, Ali, Alisha, Kristine, Andoni, Yizhe, Shaheen, Azim, thank you for your support and for making the lab such a wonderful place.

Thank you to my BioE Cohort for becoming a family from the very first day. Thank you to Kristin Olson for being so much more than an academic advisor, and to the BioE staff, Brandy Thomas, Lily Mirels, Mary West and Naima Azgui for making our work possible and always offering more than technical and administrative support.

I would also like to thank the funding sources that have made all of this possible. Gràcies a la Obra Social "la Caixa" per creure amb mi. Thank you to the California Institute for Regenerative Medicine and the Siebel Scholars Foundation for placing your trust in me.

And most importantly, I would like to thank my family. To Marcel, from Boston to Berkeley, gràcies per haver estat al meu costat durant una de les etapes més difícils de la meva vida. Thank you for being the rock that I can always turn to, no matter how difficult the situation. Gràcies per haver triat fer-me costat, we have done this together, igual que totes les coses que ens queden per viure. To my sister Núria, thank you for always admiring and supporting me. Thank you for being un exemple a seguir, for always encouraging me and for making your little sister feel like she can conquer the world. A la meva família, de la qual n'estic tant orgullosa, gràcies per fer que les distàncies semblin tant curtes. Gràcies Tieta per sempre ser-hi, i per sempre tenir a punt paraules carinyoses i de suport. Tiets i cosins, gràcies per fer-me sentir com si fós a casa tot i ser a tants kilòmetres de distància.

And finally, to the two people who have dedicated their lives to mine; my parents. No sé si mai podré agraïr-vos prou tot el que feu per a mi. Gràcies per ser els meus referents d'esforç i treball, per regalar-me la millor educació del món, i per sempre haver-me dit que podia arribar a ser i fer el que vulgués. Mami, gràcies per ser una font de carinyo il.limitada, gràcies per la teva devoció i la teva constant dedicació, i per recordar-me qui sóc quan de vegades me n'oblido. Papi, gràcies per no perdre mai la positivitat, il.lusió i motivació tant característiques tevés, gràcies per "push" me a ser la millor versió de mi mateixa- (can you believe que ja som dos Dr. Rosàs?!). Aquest doctorat és una mostra més del que una persona pot aconseguir quan se sent infinitament recolzada i profundament estimada. Vosaltres m'heu fet arribar a ser la Dra. Rosàs Canyelles. Gràcies per sempre creure amb mi and remember that everything I do is to make you proud. Us ho dec tot a la vida, us estimo.

Table of Contents

Chapter 1: Introduction

1.1 Electrokinetic Phenomena

1.1.1 The Electrical Double Layer

. The common source of electrokinetic phenomena, as originally conceived by Hermann von Helmholtz, is the electrical double layer (EDL) that forms at the interface between a charged surface and an electrolyte[1,2]. The Gouy-Chapman-Stern Model describes the EDL as a first layer of ions strongly adsorbed onto the charged surface, or the Stern Layer, followed by a second layer of diffuse ions, or the diffuse layer (Figure 1, a).

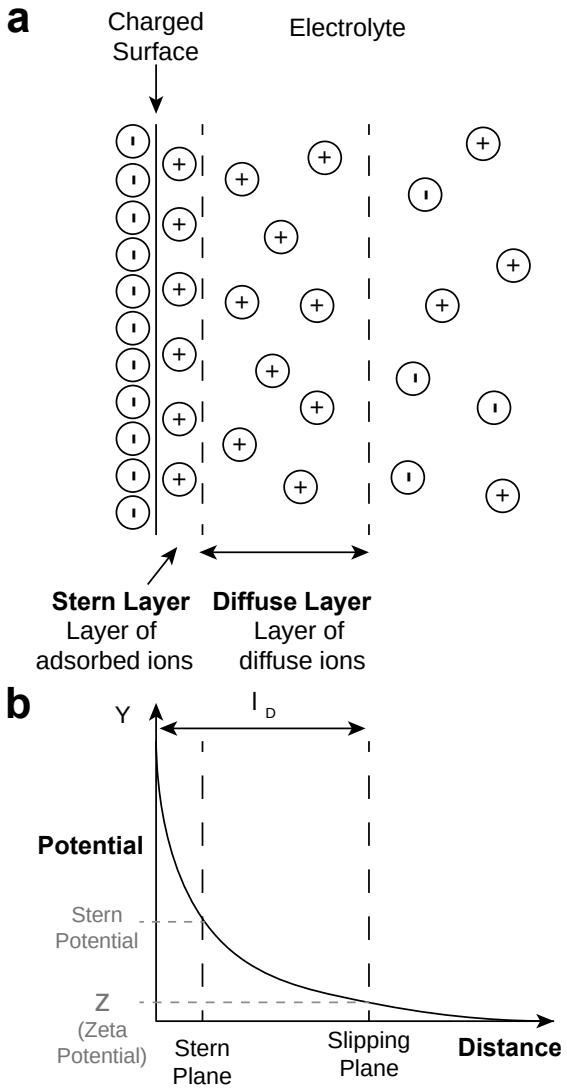


Figure 1: The electrical double layer (EDL) is formed at the interface of a charged solid surface and an electrolyte and is composed of a first layer of strongly adsorbed ions, or the Stern layer, followed by a layer of ions loosely associated with the surface, or the diffuse layer.

The formation of the EDL screens the charged surface of the solid phase from the bulk solution, causing an electrokinetic potential between the charged surface and any charge in the bulk fluid[1] (Figure 1, b). This electrokinetic potential drops as the distance from the surface increases. The potential at the Stern plane is commonly known as the Stern potential, while the potential at the boundary between the diffuse layer and the bulk fluid, or the slipping plane, the potential is known as ζ, or zeta potential (Figure 1, b). The length of the EDL is commonly known as the Debye screening length, λ_D (Figure 1, b).

Two major electrokinetic phenomena, electro-osmosis and electrophoresis, arise due to the interaction of this system with an electric field. While electro-osmosis occurs when the migration of the diffuse layer generates bulk flow in the fluid, electrophoresis refers to the migration of a solid particle in a stationary fluid. The following section is a more detailed discussion of electrophoresis, the electokinetic phenomenon used in separation science to fractionate analytes for quantification or analysis of physical or chemical properties.

1.1.2 Electrophoresis

Electrophoresis is the migration of a charged particle in an electrolyte under an applied electric field (Figure 2).

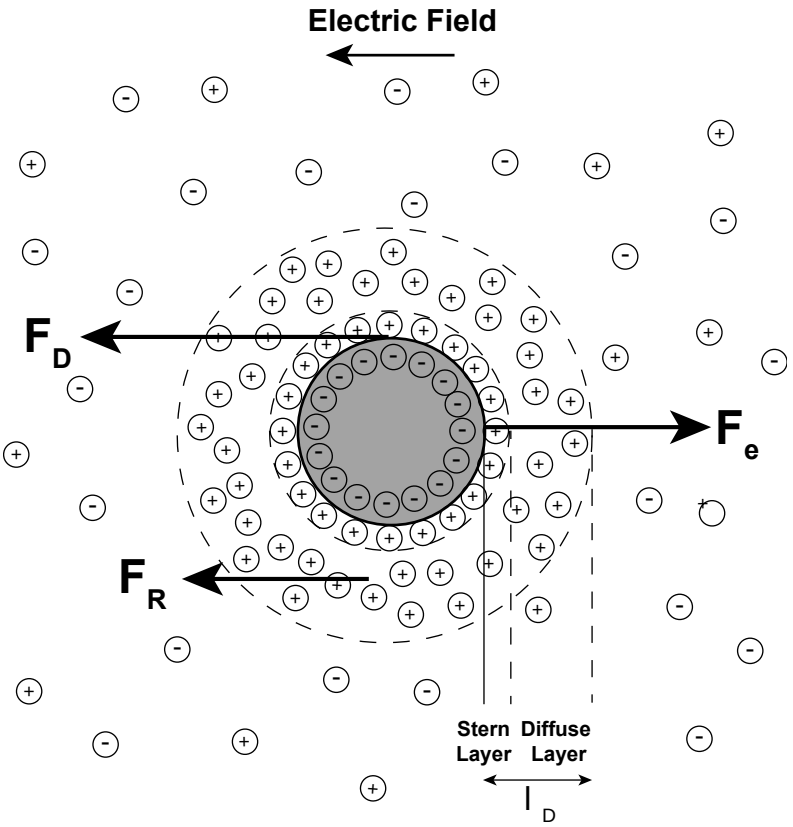


Figure 2: The forces acting on a charged particle in a fluid during electrophoresis include (i) the electrostatic force F_e, (ii) the viscous frictional force from the fluid, or drag force F_D, and (iii) an

electrophoretic retardation force, F_R, a frictional force resulting from the migration of ions in the diffuse layer. λ_D is the length of the electrical double layer, including the Stern and diffuse layers.

Under an applied electric field, the particle will experience an electrostatic force, F_e, as well as a viscous friction force, or drag force F_D, in the opposite direction due to movement in the fluid (Figure 2). An additional force results from the presence of the EDL, where movement of ions in the diffuse layer cause a frictional force called the electrophoretic retardation force, F_R (Figure 2).

The electrophoretic mobility (μ) characterizes the motion of the particle migrating in the electric field[3]. The importance of the EDL in determining electrophoretic properties can be seen through electrophoresis theory of Smoluchowski, which describes the relationship between electrophoretic mobility, μ, and the zeta potential, ζ, of a rigid spherical colloid as:

$$\mu = \frac{\epsilon_r \varepsilon_0}{\eta} \zeta \qquad \text{Equation 1}$$

where the ϵ_r, ϵ_0 and η are the relative permittivity, the permittivity of free space and the viscosity of the surrounding fluid[4]. As seen in the equation, the electrophoretic mobility is proportional to the zeta potential. Additionally, the length of the EDL is a key parameter in determining electrophoretic mobility, as the relationship only holds for a thin double layer, when $a >> 1/\kappa$, where a is the particle radius and κ is $1/\lambda_D$.

To understand how electrophoresis can be used to separate particles of varying properties, we can look at how electrophoretic mobility changes with particle size and charge. Electrophoretic mobility is defined as the velocity, *U*, normalized by the applied electric field strength, *E*:

$$\mu = \frac{U}{E} \qquad \text{Equation 2}$$

We can determine the velocity of a particle by balancing the different forces exerted on this particle; the electrostatic force, F_e, and the drag forces the particle experiences as it migrates through the fluid, F_D and F_R (Figure 2). For a small spherical particle travelling in a viscous fluid, we can use Stoke's law to describe the resulting drag force from F_D and F_R experienced by the particle:

$$F_D = 6\pi\eta a U \qquad \text{Equation 3}$$

where η is the dynamic viscosity of the fluid, a is the hydrodynamic radius of the particle, and U is the velocity. The Debye-Hückle approximation describes the electrophoretic mobility for thick double layers[3]:

$$\mu = \frac{2\epsilon\psi_0}{3\eta} \qquad \text{Equation 4}$$

where ϵ is the dielectric permittivity, and ψ_0 is the uniform surface potential of the particle. For thin double layers, where $\lambda_D << a$, Helmholtz-Smoluchowski describes the mobility of a particle as[3]:

$$\mu = \frac{\epsilon\psi_0}{\eta} \qquad \text{Equation 5}$$

If the viscosity of the fluid remains uniform, ϵ/η will remain constant. For a mixture of particles in a solution, the electrophoretic mobility will thus be a function of the surface charge density. Therefore, electrophoresis allows spatial fractionation of particles by their ψ_0.

1.1.3 Electrophoretic Protein Separations

A common method for analyzing proteins from complex mixtures, such as cell lysates, is by sodium dodecyl sulfate polyacrylamide gel electrophoresis, or SDS-PAGE. SDS is an anionic detergent composed of a polar head (sulfate group) and a 12-carbon hydrophobic tail. Above the critical micelle concentration, SDS molecules form micelles that are capable of unfolding proteins through hydrophobic interactions of micelles with the hydrophobic residues of proteins. In SDS-PAGE, SDS is used to denature proteins and form SDS-protein complexes. Complexation masks the charges on protein amino acid residues and creates what we can be considered as a uniformly-distributed negative charge along the denatured protein polymer chain[5].

Polyacrylamide gels are then employed as sieving matrices during electrophoresis. Polyacrylamide, or PA, is a hydrogel that results from copolymerizing monomers of acrylamide (CH_2=CH-$CONH_2$) and crosslinker bisacrylamide (CH_2(NH-CO-CH=CH_2)$_2$). Characterizing how the concentration of monomers (acrylamide and bisacrylamide) affects pore size of the gel and thus properties of molecules such as diffusion has been the focus of many studies[6,7]. The composition of PA gels is commonly characterized by two parameters, the total monomer concentration, or %T ([g acrylamide + g bisacrylamide] / 100 mL precursor solution x 100), and the weight percentage of crosslinker, %C (g crosslinker / [g acrylamide + g crosslinker] x100).

The electrophoretic mobility of a proteins in free solution is influenced by properties such as the primary structure and length of the protein, as well as the ionic strength, pH, viscosity, and temperature of the solution[3]. In a commonly employed approximation, proteins are treated as non-conducting spheres with uniform surface charge[3], and the derivation of electrophoretic mobility similar to that in section 1.1.2 can be considered. Similarly, the forces acting upon the molecule are the electrostatic force, the drag force, and the electrophoretic retardation force due to the migration of ions in the diffuse layer of the EDL.

When proteins electrophoretically migrate through a sieving matrix like PA, instead of free solution, an additional retardation force acts against the direction of electrophoretic migration. The basis of protein separation by electrophoresis is that proteins with varying molecular masses have different electrophoretic mobilities- the larger the hydrodynamic radius of a protein, the stronger the drag and retardation forces it experiences, and thus the slower it migrates. The Ogston model can be used to describe the migration of SDS-protein complexes through a PA gel[8]. Assuming the protein to be a rigid sphere and an average pore size comparable to the hydrodynamic radius of the protein, the Ogston model describes the electrophoretic mobility as:

$$\mu = \mu_0 e^{-bC\left(R_{gp}+r\right)^2}$$

where μ_0 is the mobility of the protein in free solution, C is the concentration of PA, R_{pg} is the hydrodynamic radius of the protein, r is the radius of the PA fiber and b is a constant[8]. Given the assumption that R_{pg} does not change with the applied electric field and that $r << R_{pg}$, we obtain the Ferguson relation:

$$\mu = \mu_0 e^{-K_R C}$$

where μ_0 is the mobility of the protein in free solution, C is the concentration of PA, and K_R is the retardation coefficient[8]. The log of the electrophoretic mobility is therefore linear with the concentration of polymer, where the K_R is the negative slope:

$$\ln \mu = \ln \mu_0 - K_R C$$

The Ferguson relation can be used to estimate the molecular mass of proteins separated by SDS-PAGE, and therefore to determine the identity of a protein[8].

1.2 Microfluidic Analysis of Biological Samples

1.2.1 Miniaturization of biological assays

Microfluidic devices manipulate fluids at 1 to 1,000 micrometer length scales[9]. The application of microfabrication techniques developed for the semiconductor industry to the biological sciences has enabled studying cellular processes at previously unattainable resolutions[9]. Precision handling of sub-microliter volumes and control of laminar flows allow matching measurements to the length and time scales of biological processes[9,10].

A common fabrication technique for microfluidic devices is photolithography, where photoresist-coated silicon wafers are exposed to UV light through a mask containing desired features. In the case of the negative epoxy-based photoresist SU-8, UV light crosslinks the photoresist, leaving behind the desired features upon UV light exposure and developing with solvent that dissolves the unpolymerized photoresist[11]. Photolithography on silicon substrates, however, requires fabricating each device in cleanroom facilities and is extremely costly when devices are not reusable. These challenges were surmounted by the introduction of polydimethylsiloxane, or PDMS, devices[12]. To cast PDMS devices, an SU-8 master is first fabricated by photolithography and then used as a mold for the polymer device. The SU-8 master can then be used to fabricate numerous inexpensive devices that also have excellent optical, mechanical, and bio-compatible properties.

Microfluidic tools take advantage of the predominance of certain forces at the microscale. For fluid flow, at these characteristic length scales, viscous forces dominate over inertial forces, leading to laminar flow regimes[8]. The well-ordered behavior of laminar flow can be easily described by flow velocity, channel dimensions, and properties of the fluid, and offers extensive advantages such as the ability to focus a suspension of disperse particles in suspension and the formation of controlled-diffusion gradients for reagent delivery[9,10,13].

Another major advantage of microsystem length scales is favorable scaling of electric fields[10]. Even with very low voltages, high electric fields can be generated over the characteristic lengths

of the devices[10]. High electric fields are necessary to counteract diffusion-based losses that rapidly dilute the contents of small sample sizes, such as proteins from single cells. In the next section, we will introduce the main aspects of the electrophoretic cytometry devices developed to measure proteins and nucleic acids from low starting numbers of mammalian cells.

1.2.2 Single-cell electrophoretic cytometry tools

Cell-to-cell heterogeneity drives biological processes, from the first differentiation event that takes place in a preimplantation embryo, to the vastly different phenotypes that characterize the cells of different tissues in an adult organism. Cytometry, or the measurement of characteristics of cells, encompasses a myriad of tools and assays that measure properties from cell size and morphology to expression levels of mRNA or protein.

Protein abundance measurements are good predictors of cell phenotype and cell state. Antibodies are commonly employed to detect and quantify expression of endogenous proteins. In assays such as immunofluorescence, antibodies are labeled with fluorophores and bound to protein targets. Fluorescence is then imaged and quantified for computation of protein expression. The use of antibodies, however, presents major specificity limitations, where antibodies often show off-target binding and low selectivity for their target[14]. Furthermore, antibodies cannot discriminate between proteoforms when isoform-specific antibodies are not available[15].

In order to add specificity to the immuno-measurement, researchers commonly perform an electrophoretic protein separation upstream of immunoassays. This assay, known as an immunoblot or western blot, spatially resolves proteins by molecular mass to discriminate signal contributions. Conventional western blotting, however, lacks the analytical sensitivity required for single-cell analysis, generally requiring tens of μg of protein, equivalent to lysate from millions of cells, per measurement[15]. To fill this gap, our research group has used microfluidic design to develop electrophoretic cytometry tools that advance protein measurement capabilities to single-cell resolution[15–20].

More specifically, this dissertation focuses on single-cell polyacrylamide gel electrophoresis, (single-cell PAGE) or single-cell immunoblotting assay. To fabricate the single-cell PAGE device, we borrow the PDMS molding step from soft lithography, and use an SU-8 silicon wafer patterned with micro-posts to cast a polyacrylamide gel with microwells resulting from the microposts. Thus, all devices can be fabricated from a single SU-8 wafer. The structural function of the PA is spatial isolation of single cells during lysis in order to minimize diffusional losses and prevent cross-contamination. During electrophoresis, the PA gel serves as the molecular sieving matrix for separation and an anti-convective medium that prevents electro-osmotic flow. Finally, at the photocapture step, the benzophenone moieties in the PA gel matrix serve as the blotting membrane onto which proteins are covalently bound.

1.3 Thesis Overview

This dissertation reports the design, development and optimization of electrophoretic cytometry tools for protein and nucleic acid measurements from low starting numbers of mammalian cells.

In Chapter 2, we examine the technical noise and reproducibility of single-cell PAGE. We study the device performance for established breast cancer cell lines, and use GFP-expressing cell line to establish a technical variability threshold under which measured variation is considered to be technical noise. We apply this threshold to study the variability in patient-derived circulating tumor cells (CTCs).

In Chapter 3, we describe a novel method for quantifying surface protein marker localization, by integrating upstream immunostaining of cell surface receptors with downstream single-cell PAGE. We examine the shifts in electrophoretic migration of the immuno-complexes with respect to free antigens and study the stability of the antibody-antigen complex under different conditions for cell lysis and electrophoresis.

In Chapter 4, we develop a workflow for handling and analysis of highly heterogeneous, primary cells (mouse aortic smooth muscle cells, SMCs). By electrophoretically separating and quantifying three differentiation protein markers from single SMCs, we profile the maturation stage of SMCs in healthy mouse vessels. In the context of vascular injury and remodeling, we identify a subpopulation of SMCs with immature-like properties, supporting the hypothesis that stem cell-like SMCs residing in healthy vessels are responsible for SMC proliferation during vascular remodeling.

In Chapter 5, we optimize the single-cell immunoblotting platform for mouse preimplantation embryos and blastomeres. We perform protein separations from all stages of preimplantation, and examine processes such as early-stage lineage bias and embryo-specific isoform expression. The workflow significantly reduces sample size requirement for protein analysis of embryos and blastomeres and grants insight into the inter- and intra-embryonic variation previously unattainable.

Chapter 6 presents a novel method for performing dual nucleic acid and protein measurements on the same mammalian cells. Designed for analysis of mRNA and protein isoforms in late-stage mouse preimplantation embryos (20-100 cells for morula and blastocyst stages), we perform simultaneous protein separations and mRNA measurements from 10-100 cells. This method shows promise for determining whether the abundant splice variants and DNA modifications in preimplantation embryos translate to protein isoforms.

1.4 References

1. Torriero, A. A. J. in 11–13 (Springer, 2015).
2. Favaro, M. *et al.* Unravelling the electrochemical double layer by direct probing of the solid/liquid interface. *Nat. Commun.* **7,** 1–8 (2016).
3. Chae, K. S. & Lenhoff, A. M. Computation of the electrophoretic mobility of proteins. *Biophys. J.* **68,** 1120–1127 (1995).
4. Kuo, Y. C. & Lin, T. W. Electrophoretic mobility, zeta potential, and fixed charge density of bovine knee chondrocytes, methyl methacrylate - Sulfopropyl methacrylate, polybutylcyanoacrylate, and solid lipid nanoparticles. *J. Phys. Chem. B* **110,** 2202–2208 (2006).
5. Bhuyan, A. K. On the mechanism of SDS-induced protein denaturation. *Biopolymers* **93,**

186–199 (2010).
6. Tong, J. & Anderson, J. L. Partitioning and diffusion of proteins and linear polymers in polyacrylamide gels. *Biophys. J.* **70,** 1505–13 (1996).
7. Williams, J. C., Mark, L. a & Eichholtz, S. Partition and permeation of dextran in polyacrylamide gel. *Biophys. J.* **75,** 493–502 (1998).
8. Chung, M., Kim, D. & Herr, A. E. Polymer sieving matrices in microanalytical electrophoresis. *Analyst* **139,** 5635–5654 (2014).
9. Duncombe, T. A., Tentori, A. M. & Herr, A. E. Microfluidics: Reframing biological enquiry. *Nat. Rev. Mol. Cell Biol.* **16,** 554–567 (2015).
10. Velve-Casquillasa, G., Berrea, M. Le, Piela, M. & Tran, P. T. Microfluidic tools for cell biological research. *Nano Today* **5,** 28–47 (2010).
11. Zhang, J., Tan, K. L., Hong, G. D., Yang, L. J. & Gong, H. Q. Polymerization optimization of SU-8 photoresist and its applications in microfluidic systems and MEMS. *J. Micromechanics Microengineering* **11,** 20–26 (2001).
12. Friend, J. & Yeo, L. Fabrication of microfluidic devices using polydimethylsiloxane. *Biomicrofluidics* **4,** 1–5 (2010).
13. Yang, M., Yang, J., Li, C. W. & Zhao, J. Generation of concentration gradient by controlled flow distribution and diffusive mixing in a microfluidic chip. *Lab Chip* **2,** 158–163 (2002).
14. Bordeaux, J. *et al.* Antibody validation. *Biotechniques* **48,** 197–209 (2010).
15. Yamauchi, K. A. & Herr, A. E. Subcellular western blotting of single cells. *Microsystems Nanoeng.* **3,** 16079 (2017).
16. Yang, M., Nelson, R. & Ros, A. Toward Analysis of Proteins in Single Cells: A Quantitative Approach Employing Isobaric Tags with MALDI Mass Spectrometry Realized with a Microfluidic Platform. *Anal. Chem.* **88,** 6672–6679 (2016).
17. Hughes, A. J. *et al.* Single-cell western blotting. *Nat. Methods* **11,** 749–55 (2014).
18. Kang, C. C., Lin, J. M. G., Xu, Z., Kumar, S. & Herr, A. E. Single-cell western blotting after whole-cell imaging to assess cancer chemotherapeutic response. *Anal. Chem.* **86,** 10429–10436 (2014).
19. Kang, C.-C. *et al.* Single cell–resolution western blotting. *Nat. Protoc.* **11,** 1508–1530 (2016).
20. Kim, J. J., Sinkala, E. & Herr, A. E. High-selectivity cytology via lab-on-a-disc western blotting of individual cells. *Lab Chip* **17,** 855–863 (2017).
21. Qiu, P. *et al.* Extracting a cellular hierarchy from high-dimensional cytometry data with SPADE. *Nat. Biotechnol.* **29,** 886–91 (2011).

Chapter 2: Measuring technical variation and reproducibility of single-cell electrophoretic cytometry

This work was conducted in collaboration with Dr. Elly Sinkala and Julea Vlassakis, and is reproduced from "Profiling protein expression in circulating tumor cells using microfluidic western blotting", Nature Communications; ***8****:14622 (2017) with permission from all authors.*

2.1 Introduction

Understanding how individual cells respond to perturbations from their environment and other cells remains a major challenge in biology[1]. Single-cell technologies have revealed important implications of heterogeneity at the single-cell level, including functional differences in genetically identical cells[2], mechanisms for intra-tumoral heterogeneity leading to drug resistance[3], and unprecedented detail of immune responses[4].

However, with the advent of single-cell technologies, questions surrounding the significance behind measured heterogeneity have begun to take center stage. Interpreting cellular heterogeneity requires accurate detection of biological variability, that can easily be masked by technical noise and technical biases[5]. Further, evaluating data reproducibility for methods with strong technical noise can become extremely challenging[6]. Although methods for accurate interpretation of heterogeneity in RNA expression have been extensively characterized[6–8], equivalent studies for single-cell protein measurements, including single-cell polyacrylamide gel electrophoresis (single-cell PAGE), are required.

In our single-cell PAGE assay, microfluidic devices consisting of a 40 μm thick polyacrylamide (PA) gel grafted to a standard microscope slide and stippled with 30 μm diameter microwells are employed to separate proteins from single-cell lysates. Cells are first sampled by gravity-based settling into the microwells. A dual-functionality lysis and electrophoresis (EP) buffer is then introduced to first lyse cells and denature proteins. Application of an electric field injects proteins into the PA and separates proteins by molecular mass. After protein separation, the migrated proteins are covalently crosslinked to the PA matrix by UV-mediated activation of benzophenone-methacrylate moieties in the PA gel. For protein detection, PA gels are incubated with primary antibodies raised against the target protein, and with fluorophore-conjugated secondary antibodies raised against the species of the primary antibody.

In this chapter, we examine technical variation introduced during sample preparation, lysis, electrophoresis and photocapture. Furthermore, we establish a technical variation threshold under which all heterogeneity can be attributed to technical noise. Finally, we investigate the device-to-device reproducibility when assaying three different breast cancer cell lines.

2.2 Materials and Methods

Chemical Reagents. Tetramethylethylenediamine (TEMED, T9281), ammonium persulfate (APS, A3678), β-mercaptoethanol (M3148), and 30%T, 2.7%C acrylamide/bis-acrylamide

(37.5:1) (A3699), bovine serum albumin (BSA, A9418) were purchased from Sigma-Aldrich. Triton X-100 (BP-151), phosphate-buffered saline (PBS, ph 7.4, 10010023) and Dulbeccos's phosphate-buffered saline (PBS, 14190144) were purchased from ThermoFisher Scientific. Premixed 10X tris-glycine EP buffer (25 mM Tris, pH 8.3; 192 mM glycine; 0.1% SDS) was purchased from BioRad. Tris buffered saline with Tween-20 (prepared from 20X TBST, sc-24953, Santa Cruz Biotechnology, Dallas, TX, USA). Deionized water (18.2 MΩ) was obtained using an Ultrapure water system from Millipore. N-[3-[(3-Benzoylphenyl)formamido]propyl] meth-acrylamide (BPMAC) was custom synthesized by Pharm-Agra Laboratories (Brevard, NC, USA).

Device Fabrication. Fabrication of the SU-8 microwell mold and PA gels was performed as described previously[9,10]. 8%T PA gels were used for the cell line and CTC experiments. Microwells were 50 μm in diameter and 60 μm in height. All PA gels were chemically polymerized with 0.08% APS and 0.08% TEMED.

Cell Culture. Cell lines representing three major breast cancer subtypes: triplenegative (BT-20), ER+ (MCF7) and HER2+ (SK-BR-3) cells were obtained from the American Type Culture Collection and authenticated using short tandem repeat analysis (Promega). All cell lines tested negative for mycoplasma. BT-20s were maintained in Eagle's minimal essential medium supplemented with 1% penicillin/streptomycin and 10% FBS. MCF7s were maintained in RPMI 1640 supplemented with 1% penicillin/streptomycin, 0.01 mg ml1 insulin (Invitrogen) and 10% FBS. SK-BR-3s were maintained in McCoy's 5A supplemented with 1% penicillin/streptomycin and 10% FBS. A GFP-expressing MCF7 cell line, used to determine technical variation, was obtained from the American Type Culture Collection and authenticated using short tandem repeat analysis (Promega), and tested negative for mycoplasma. The cell line was maintained in RPMI 1640 supplemented with 1% penicillin/ streptomycin, 0.01 mg/ml insulin (Invitrogen) and 10% FBS. All cell lines were cultured in an incubator held at 37 C under 5% CO2 and tested for mycoplasma contamination.

Patient recruitment and blood donation. Twelve patients with advanced breast cancer were recruited, with informed consent, according to a protocol approved by the Institutional Review Board (Stanford IRB 350–Panel 3–Protocol 5630) from the Department of Oncology at the Stanford School of Medicine. Blood was drawn in EDTA BD tubes, stored at room temperature and processed within 5 h after collection.

Rare-cell enrichment from blood samples. A previously reported and commercially available microfluidic tool (Vortex Biosciences, Menlo Park, CA) was used for label-free isolation of circulating cancer cells in both the cell line spiking and cancer patient blood experiments. For spiking experiments, cells were dissociated with 1.5 ml of 0.25% trypsin (15050065, Thermo Fisher) and incubated in full media at room temperature to recover from exposure to trypsin. Cells were immediately spiked into healthy blood samples with PBS and enriched through the Vortex HT chip, which uses microscale vortices to retain large cancer cells while allowing smaller blood cells to exit as effluent[11]. The device was first primed with PBS. Then, the diluted blood sample was processed through the Vortex HT chip (8 mL/min) followed by a wash step with PBS to remove contaminating red blood cells and WBCs (8 mL/min). Stopping the flow dissipates the vortices and releases the cancer cells from the microscale reservoirs for direct deposition on the top surface of the single-cell PAGE platform. The enriched volume was 300 μl and was contained

by a mesofluidic PDMS insert that sits atop the single-cell PAGE device. For patient blood experiments, the cells isolated in the vortices were directly collected into the mesofluidic PDMS insert seated on top of the single-cell PAGE PA gel for cell positioning into microwells. For both cell-line spiking and patient-derived cell experiments, a volume of blood was reserved for subsequent red blood cell lysis to perform control experiments with WBCs.

Single-cell PAGE. The single-cell PAGE assay comprises six steps. The single-cell PAGE device consists of microwells cast into a thin layer of a photoactive PA gel grafted onto a microscope glass slide. Once aliquoted into the mesofluidic insert, cell nuclei were stained with Hoechst 33342 (H1399, Thermo Fisher) to identify target cells, and a micromanipulator (Transferman, Eppendorf) and aspiration (Cell Vario, Eppendorf) manually positioned individual cells into each microwell. A combined lysis and electrophoresis buffer was poured directly onto the PA gel where the cells were lysed in-well and then subjected to PAGE (E = 40V/cm). Lysis buffer was heated in a water bath, and the temperature was recorded with a thermometer immediately before use. After the PAGE separation, proteins were immobilized in the gel via brief ultraviolet activation (Lightningcure LC5, Hamamatsu) of benzophenone methacrylamide cross-linked into the PA gel. Immobilized proteins were probed in-gel by diffusing primary and then fluorescently labeled secondary antibody probes into the PA gel layer. A fluorescence microarray scanner (Genepix 4300A, Molecular Devices) equipped with four laser lines (488, 532, 594 and 635) acquired fluorescence readout. Subsequent rounds of antibody stripping were performed for multiplexed protein analysis.

Image processing, separation performance quantification, statistical analysis. Fluorescence microscopy of GFP-MCF7 cells settled into microwells was performed with an inverted epifluorescence microscope (Andor iXon. EMCCD camera, X-cite Lumen Dynamics mercury excitation lamp, ASI motorized stage controlled in Metamorph software, Molecular Devices, 10 Olympus UPlanFLN, numerical aperture 0.45 objective, GFP filter set Chroma 49011 ET, a binning of 1 and an exposure time of 200 ms). Slides were imaged for fluorescence protein blots with a microarray scanner (Genepix 4300A, Molecular Devices). Quantification of fluorescence signal employed in-house scripts written in MATLAB (R2017a, Mathworks). Parameters such as peak location and peak width were obtained by fitting Gaussian curves to protein bands using MATLAB's Curve Fitting Toolbox. Gamma distribution fitting, goodness of fit tests, agglomerative hierarchical clustering, and clustergram heatmap rendering were performed with the Matlab Statistics and Machine Learning Toolbox. Statistical tests (d'Agostino & Pearson normality tests, Mann Whitney U tests) were performed using GraphPad Prism 7.0b. Fiji was used to false-color fluorescence micrographs and overlay channels to create composite images.

Flow Cytometry. Flow cytometry analysis was performed on MCF7 cells to ascertain the effects of enzymatic detachment on the EpCAM antigen. Briefly, MCF7 cells were detached from tissue culture plates either by trypsin-EDTA (0.25%, Gibco 25200072) or by EDTA alone (Ultrapure 0.5M EDTA Gibco 15575020 diluted in PBS to 5 mM). Half the cells were labelled with anti-EpCAM AlexaFluor488 (mouse, monoclonal antibody, 53-8326-42, eBioscience) and half were labelled with mouse IgG AlexaFluor488 (A21202) as an isotype control. For labelling, $2x10^6$ cells were resuspended in 100 ml of 3% BSA (Sigma A2058) in PBS containing antibody at a concentration of 1 μg/ml and incubated over ice for 30 min. Cells were washed three times in PBS, then resuspended in resuspension buffer (1% BSA, 5mM EDTA in PBS) to prevent aggregation.

Cells were analysed on a Guava flow cytometer (Millipore). A total of 10,000 events were collected per sample, four samples per experimental group (N = 4) and data were compiled and analysed using FlowJo software.

2.3 Results and Discussion

2.3.1 Studying the effects of sample preparation on surface protein measurements

We first examined if sample preparation introduces variation in protein expression measurements. More specifically, we sought to understand if enzymatic cleavage of cell adhesion molecules during the cell harvesting step alters the surface receptor protein targets of interest, such as the epithelial cellular adhesion molecule, or EpCAM, a marker that regulates growth and tissue repair, as well as pathologic cell proliferation, migration, and invasion in many types of cancer[12,13].

Most single-cell analyses begin with harvesting cells from tissue culture vessels and dissociating cell-to-cell adhesions. Adherent cells grown *in vitro* are commonly detached from vessel surfaces and dissociated into single cells by a combination of enzymatic cleavage of adhesion proteins (e.g. cadherins) and chelators that remove divalent cations (Mg^{2+} and Ca^{2+}) that cell surface adhesion receptors (i.e. integrins) use to maintain cell-to-substrate and cell-to-cell adhesion[14]. In the case of single-cell PAGE, cells are detached by incubation with the chelator ethylenediaminetetraacetic acid, or EDTA, and the proteolytic enzyme trypsin, a serine protease that cleaves proteins at the arginine and lysine residues[15].

Given that the proteolytic activity of trypsin can degrade surface proteins and glycoproteins when used for cell detachment and dissociation[16,17], we examined if detaching cells with trypsin affects the levels of EpCAM on the surface of cells. To do so, we detached MCF7 cells from their vessels with either EDTA or trypsin, and immune-stained live cells with AlexaFluor488-anti-EpCAM antibodies. An aliquot of unstained, EDTA-detached cells serves as a negative control. We analyzed fluorescence of our negative control (MCF7 cells not stained with AlexaFluor488-anti-EpCAM) and experimental groups (cells stained with AlexaFluor488-anti-EpCAM after detachment with either EDTA or trypsin) using a flow cytometer. We found that in both cells detached with EDTA and trypsin, more than 98% of cells were positive for EpCAM ($99 \pm 0.06\%$ for trypsin/EDTA and $98 \pm 0.05\%$ for EDTA, N = 4 replicates for both groups) (Figure 1, A), suggesting that the EpCAM epitope is not cleaved by trypsin and, thus, enzymatic dissociation does not affect our ability to detect the EpCAM antigen.

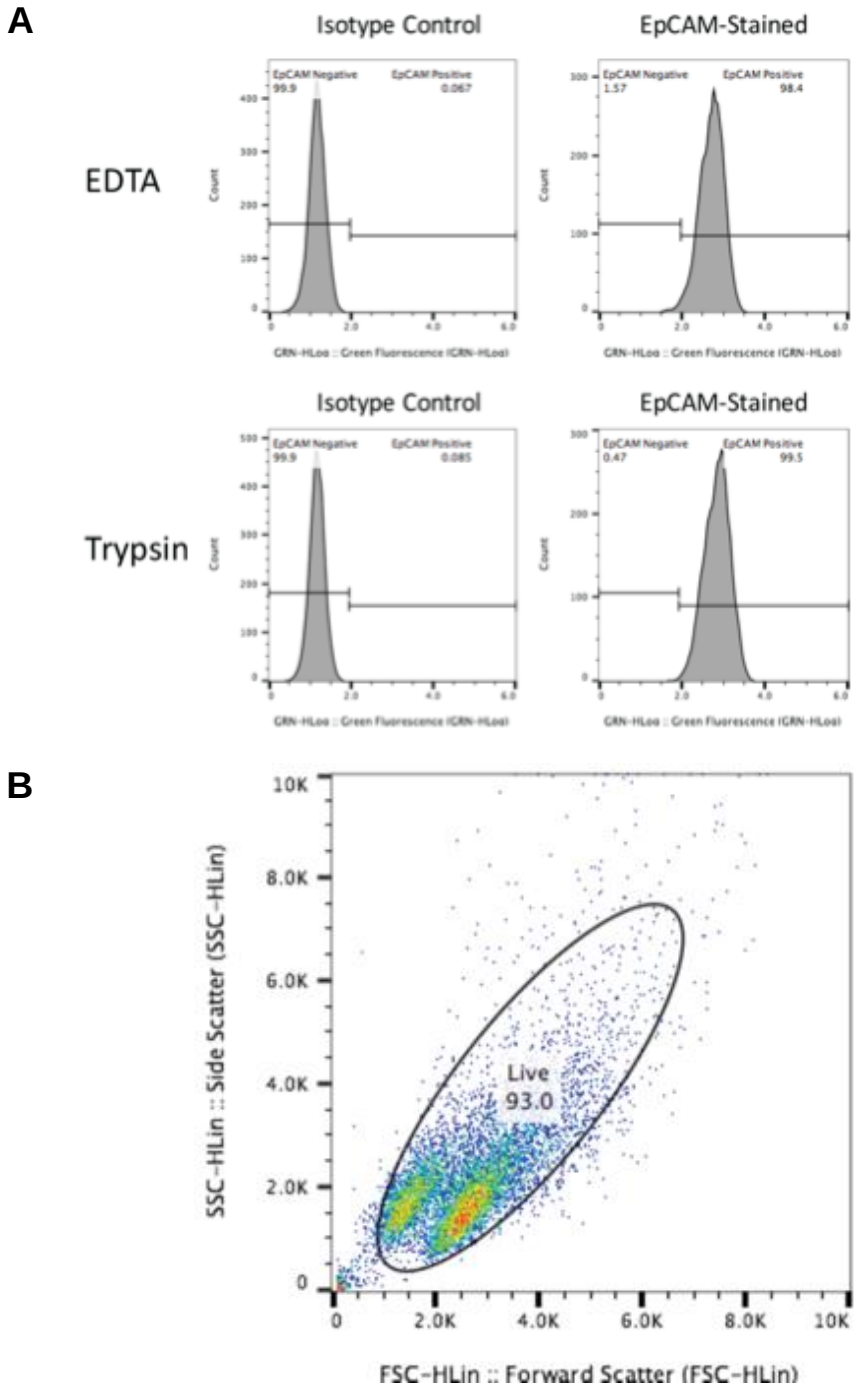


Figure 1: Sensitivity of EpCAM antigen to enzymatic detachment. (A) Fluorescence intensity histograms of MCF7 cells labeled with isotype control (AlexaFluor488-mouse IgG) or anti-EpCAM-Alexa Fluor 488 antibody. Cells were detached either by 5mM EDTA (top) or trypsin-EDTA (bottom). In both cases, >98% of labeled cells were positive for EpCAM, using the isotype controls for EpCAM-negative gating (99±0.06% for trypsin/EDTA and 98±0.05% for EDTA, N = 4 replicates for both groups). (B) Forward scatter vs. side scatter plot of the Trypsin, EpCAM-stained sample, showing the gating strategy for live MCF7 cells.

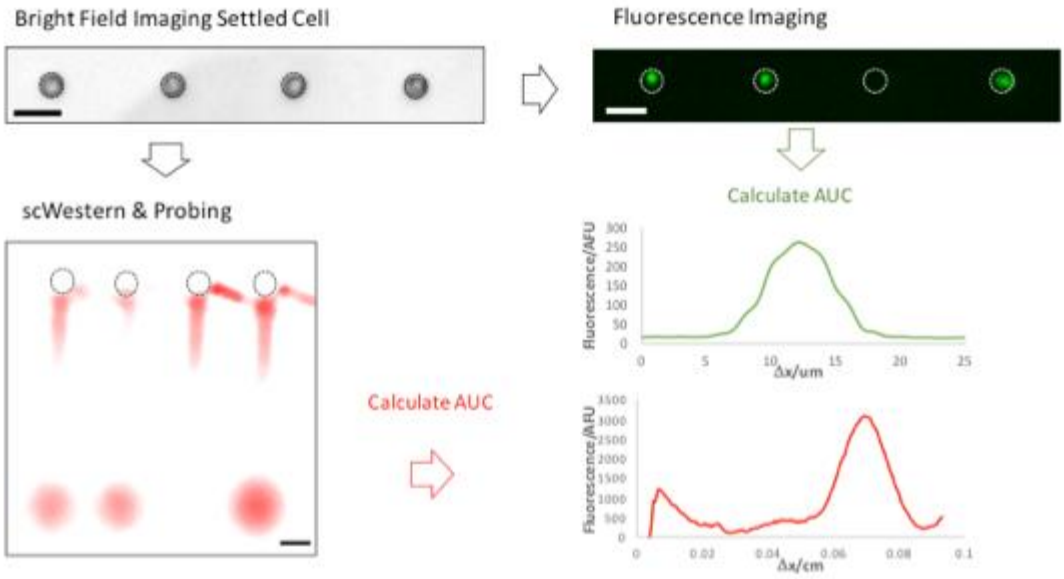
2.3.2 Establishing a technical variation threshold for single-cell PAGE measurements

Next, we sought to establish a technical variation threshold for any variability between microwells of the same device introduced during the lysis, electrophoresis, photocapture and immunoprobing steps. To do so, we used GFP-expressing MCF7 cells to compared variation in GFP expression levels obtained by (i) fluorescence imaging of whole cells with (ii) immunoprobed GFP signal from the same cells. We first harvested a suspension of GFP-expressing MCF7 cells at 1 million cells/ml (in PBS) and pipetted cells onto the single-cell PAGE device to allow cells to settle into microwells. After washing off excess cells with PBS, we performed whole-cell imaging by epifluorescence microscopy to record GFP fluorescence from the MCF7-GFP cells seated in the microwells (Figure 2, A). After imaging, cells were assayed by lysis with 55^{o}C RIPA-like lysis buffer (0.5% SDS, 0.25% Na-DOC, 0.1% Triton X-100 in 0.5 Tris-glycine), followed by PAGE (20 s at 40 Vcm^{-1}), and photo-immunoblotting (45 s). Devices were washed with TBST (2 x 1 hr washes).

We then immunoprobed devices for GFP (1:10 dilutions of anti-GFP antibody in TBST with 5% BSA, 2 hr). After washing in TBST (30 min x 2), gels were incubated with secondary antibody (1:10 anti-goat AlexaFluor 555-conjugated secondary antibody in TBST, 1 hr) and washed in TBST (30 min x 2). We rinsed the PA gels with deionized water and dried them under a nitrogen stream before imaging the immunoprobed signal with the microarray scanner. To quantify the fluorescence signal from (i) settled GFP-MCF7 cells and (ii) immunoprobed GFP signal, a fluorescence intensity profile was generated in the microwell or PA gel area abutting the microwell. After fitting a Gaussian distribution to the intensity profile, the AUC was calculated.

Finally, in order to establish the technical variation threshold, cells with similar in-microwell GFP signal (< 5% variation) were binned and considered a homogeneous GFP-expressing sample, with a 1.27–3.37% difference in area-under-the-curve signal (AUC) from the lowest and highest GFP AUC of each bin observed (Figure 2, B). We then calculated the coefficient of variation (CV = S.D./mean %) in final immunoprobed GFP signal for all bins (Figure 2, B) and defined a technical variation cutoff as 3 S.D. above the average CV% of immuprobed signal (32.4%, for a 99.7% confidence interval).

A



B

Cell	GFP AUC	Probed AUC	% Difference GFP	CV Probed (%)
1	166	95046	2.99	10.0
2	170	77774		
3	171	88965		
4	275	120137	1.27	5.5
5	275	132753		
6	278	132067		
7	367	145921	3.37	20.4
8	372	164309		
9	380	214914		
10	408	168585	3.14	15.9
11	414	153812		
12	421	208114		
13	1139	241967	1.63	3.3
14	1152	246009		
15	1157	230755		

CV Average	CV StDev	CV Threshold=avrg+3*stdev
11.0	7.1	**32.4**

Figure 2: (A) Brightfield and fluorescence images of MCF7-GFP cells in microwells, representative fluorescence micrograph and intensity profiles of GFP in the well and anti-GFP antibody probe signal, and table containing GFP and antibody probe signal values used to estimate intra-assay technical variation threshold. GFP-expressing cells are imaged by fluorescence microscopy in single-cell PAGE wells prior to cell lysis and fluorescence of the cell is quantified. The single-cell PAGE is performed (as described in the main text) and GFP is detected with anti-GFP antibodies. Antibody probe signal (area-under-the curve, AUC) and coefficient of variation (CV) is quantified for cells that had <5% variation in intact GFP cell fluorescence. (B) Table showing the fluorescence intensity prior to lysis (AUC), the probed fluorescence intensity (AUC), the difference in fluorescence intensity for each bin in %, and finally the CV for probed fluorescence between cells in the same bin. The technical variation threshold is calculated as the mean CV

(11.0) plus three standard deviations (7.1, for a 99.7% confidence interval) yielding a CV threshold of 32.4%.

We then used the technical variation cutoff to determine whether the variation detected through single-cell PAGE of circulating tumor cells (CTCs) corresponds to true biological variation. We isolated CTCs from the blood of breast cancer patients, assayed them by single-cell PAGE and immunoprobed for a panel of 8 protein markers, including loading controls (GAPDH and β-tubulin), a cell-type control (CD45) and protein markers known to be upregulated in tumor cells (ER, ERK, elF4E, EpCAM and panCK) (Figure 3, A). Notably, the CVs in protein expression for the panel of 8 protein markers are above the technical variation threshold for all patients (Figure 3, B), indicating our ability to measure biological variation in patient-derived CTCs.

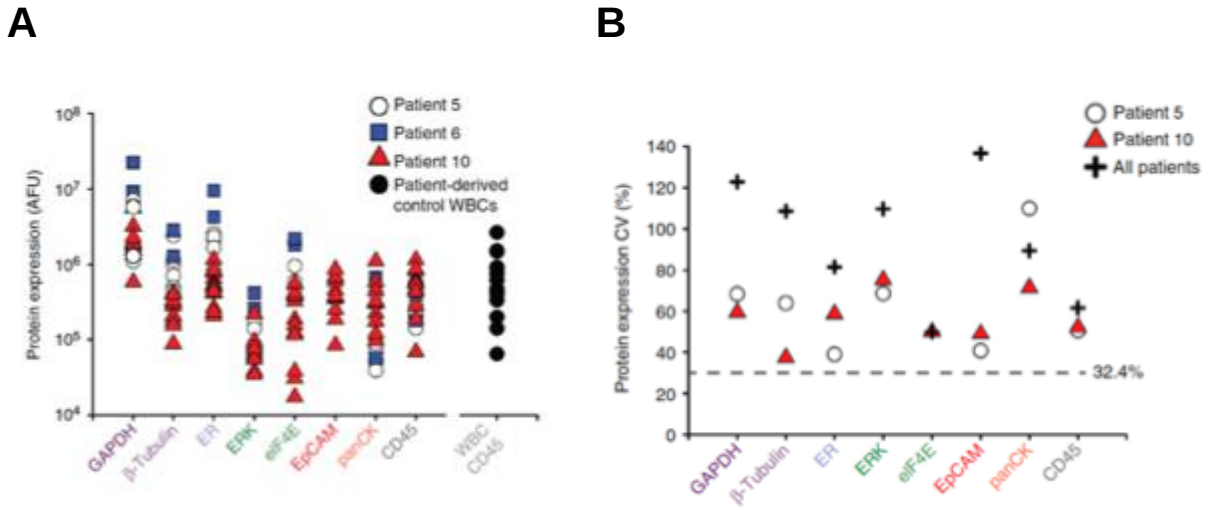


Figure 3: Immunoblotting profiles for eight proteins in each individual CTC derived from three ER positive metastatic breast cancer patients. (A) Expression for each protein marker and each patient-derived CTC, with comparison with CD45 levels from single-cell PAGE analyses of pure WBC controls. (B) CVs for protein expression (AUC) from the patient-derived CTCs. Dashed line indicates the threshold in protein expression variation established using GFP-expressing MCF7 cells.

2.3.3 Examining device-to-device reproducibility

Next, we examined the device-to-device reproducibility and performance variation of our single-cell PAGE measurements. To enable comparisons of data sets acquired on difference devices, the technical variation across different runs cannot exceed the biological variation. To examine whether the run-to-run technical variation introduced during settling, lysis, electrophoresis and probing masks biological variation, we tested whether the expression level distributions of a commonly employed loading control, GAPDH, differs when assaying cells on different devices.

For these experiments, we used the three breast cancer cell lines BT-20, SK-BR-3 and MCF-7. We harvested cells and prepared two aliquots of each cell type to perform two technical replicates on separate devices. Each cell suspension was pipetted over a different device, and cells were gravity-settled into microwells, with excess cells washed off using PBS. After completing the single-cell PAGE protocol, expression levels of the loading control GAPDH were quantified (Figure 4).

Statistical equivalence of the GAPDH expression distributions between the technical replicates was tested using the Mann–Whitney U-test (Mann–Whitney U-test p values: 0.1257, 0.7578 and 0.7815 for BT-20 (N = 59 and 65 cells), SK-BR-3 (N = 34 and 30 cells) and MCF7 (N = 42 and 40), respectively) (Figure 4). Thus, the GAPDH expression distributions across the technical replicates can be considered to be equivalent, demonstrating inter-assay reproducibility.

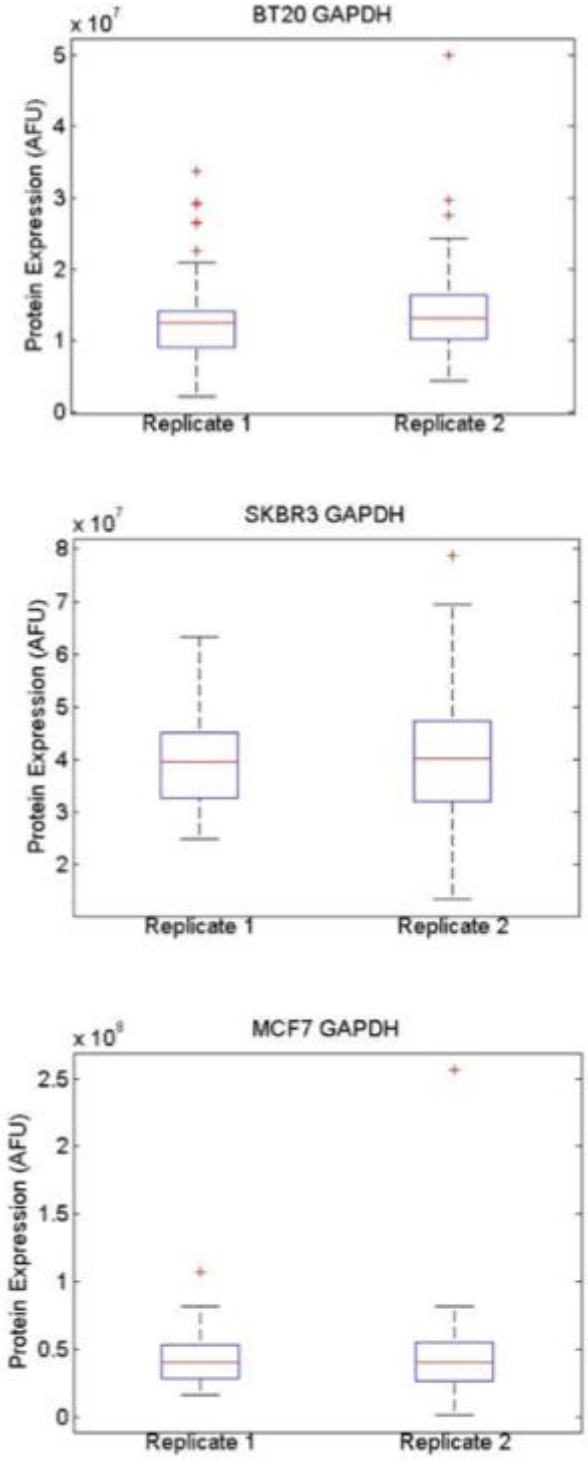


Figure 4: Inter-assay technical variation of protein expression in single-cell PAGE does not mask biological variation. Comparison of GAPDH protein expression in single-cell PAGE technical replicates using the indicated cell lines (single-cell PAGE performed as described in the main text). Mann-Whitney U-test p-values were 0.1257, 0.7578 and 0.7815 for BT20 (N = 59 and 65), SKBR3 (N = 34 and 30) and MCF7 (N = 42 and 40) respectively, confirming the null hypothesis that the distributions of the technical replicates are equal. The red horizontal line represents the mean, and the upper and lower whiskers represent the 75th and 25th percentile respectively.

2.4 Conclusions

The single-cell PAGE platform offers insight into the protein expression profiles of thousands of single cells. However, distinguishing true biological variability from technical noise is essential for performing single-cell measurements. Here we examined the effects of technical noise and biases introduced during the single-cell PAGE workflow on the protein expression profiles measured in cultured cells. We first established that the enzymatic cleavage of adhesion molecules during sample preparation does not alter the detection of surface markers of interest (i.e. EpCAM). We then designed an assay for computing a technical variation cutoff, where samples assayed were homogeneous purified protein solutions partitioned into microwells of the single-cell PAGE device. The coefficient of variation of circulating tumor cells derived from patients were all above the technical variation cutoff, indicating that the single-cell PAGE device is able to detect true biological variation even with small sample sizes (< 10 cells). When examining device-to-device reproducibility, we found that analysis of technical replicates on two separate devices was equivalent across three different cell lines studied (MCF7, SK-BR-3 and BT-20). As single-cell technologies continue to emerge, rigorous validation of the detection of genuine heterogeneity in cell populations will continue to be imperative for advancing our understanding how single cells drive complex biological processes.

2.5 References

1. Altschuler, S. J. & Wu, L. F. Cellular heterogeneity : when do differences make a difference? *Cell* **141,** 559–563 (2010).
2. De Souza, N. Single-cell methods. *Nat. Methods* **9,** 35 (2010).
3. Janiszewska, M. *et al.* In situ single-cell analysis identifies heterogeneity for PIK3CA mutation and HER2 amplification in HER2-positive breast cancer. *Nat. Genet.* **47,** 1212–1219 (2015).
4. Herderschee, J., Fenwick, C., Pantaleo, G., Roger, T. & Calandra, T. Emerging single-cell technologies in immunology. *J. Leukoc. Biol.* **98,** 23–32 (2015).
5. Yuan, G. C. *et al.* Challenges and emerging directions in single-cell analysis. *Genome Biol.* **18,** 1–8 (2017).
6. Bhargava, V., Head, S. R., Ordoukhanian, P., Mercola, M. & Subramaniam, S. Technical Variations in Low-Input RNA-seq Methodologies. *Sci. Rep.* **4,** 1–10 (2015).
7. Rizvi, A. H. *et al.* Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. *Nat. Biotechnol.* **35,** 551–560 (2017).
8. Lee, M.-C. W. *et al.* Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E4726-35 (2014).

9. Kang, C.-C. *et al.* Single cell–resolution western blotting. *Nat. Protoc.* **11,** 1508–1530 (2016).
10. Sinkala, E. *et al.* Profiling protein expression in circulating tumour cells using microfluidic western blotting. *Nat. Commun.* **8,** (2017).
11. Dhar, M. *et al.* High efficiency vortex trapping of circulating tumor cells. *Biomicrofluidics* **9,** 1–12 (2015).
12. Munz, M., Baeuerle, P. A. & Gires, O. The emerging role of EpCAM in cancer and stem cell signaling. *Cancer Res.* **69,** 5627–5629 (2009).
13. Baeuerle, P. A. & Gires, O. EpCAM (CD326) finding its role in cancer. *Br. J. Cancer* **96,** 417–423 (2007).
14. Zhang, K. & Chen, J. The regulation of integrin function by divalent cations. *Cell Adh. Migr.* **6,** 20–29 (2012).
15. Olsen, J. V., Ong, S.-E. & Mann, M. Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues. *Mol. Cell. Proteomics* **3,** 608–614 (2004).
16. Heng, B. C., Cowan, C. M. & Basu, S. Comparison of enzymatic and non-enzymatic means of dissociating adherent monolayers of mesenchymal stem cells. *Biol. Proced. Online* **11,** 161–169 (2009).
17. Huang, H. *et al.* Trypsin-induced proteome alteration during cell subculture in mammalian cells. 1–10 (2010).

Chapter 3: Integration of live-cell immunostaining with single-cell polyacrylamide gel electrophoresis.

This work was conducted in collaboration with Elly Sinkala.

Cells are equipped with cell surface receptors that allow interpretation and translation of extracellular signals. Measuring the abundance of cell receptors on the cell surface allows for cell-type identification and provides valuable insight on cell phenotype. However, it cannot sufficiently describe the specific downstream signaling pathways being activated or detect activity of receptors that have been internalized or that constitutively active in other subcellular compartments. Methods such as immunofluorescence and flow cytometry that could measure surface-bound and intracellular targets require fixation and suffer from specificity challenges due to the low specificity of antibody probes. In this chapter, we describe a method for integrating single-cell polyacrylamide gel electrophoresis (PAGE) with live MCF7s cells stained for surface-bound epithelial cellular adhesion molecule (EpCAM) receptors.

3.1 Introduction

Cell surface receptors are responsible for responding to local or distal soluble factors, binding ligands on the surface of other cells to mediate cell-to-cell interactions and sensing and transducing physical cues from their microenvironment[1,2]. Measuring the expression of surface receptors on the surface of a cell has thus become a useful strategy for identifying cell types and examining cell phenotypes[3–6]. However, measuring the expression of surface receptors on the surface of a cell alone is not enough to fully describe cellular state[7]. First, the *localization* of surface receptors more accurately depicts the phenotype of a cell than the total expression. For instance, measuring receptors not bound to the cell surface becomes crucial in cases where constitutively active surface receptors can signal from intercellular compartments[8] or with the expression of constitutively active isoforms that lack the extracellular domain or are not bound to the membrane[9]. Second, when establishing surface receptor-mediated signaling, measuring the abundance and activation of proteins in the downstream signaling pathways is as important as measuring surface receptors. This becomes extremely important in cases where a given receptor activates multiple signaling pathways, so measuring abundance of the surface receptor when only on the surface cannot reveal the specific proteins and genes involved[10,11]. Thus, in order to fully characterize cellular phenotype and state, we require tools to measure (i) the abundance of surface markers on the surface of cells in conjunction with (ii) internalized surface receptors and (iii) intracellular proteins in the downstream signaling pathways.

Gold standard tools that measure surface receptors along with intracellular targets, i.e. flow cytometry and immunofluorescence, employ antibodies probes for specificity[12–14]. However, antibody probes present confounding limitations including cross-reactivity with species other than target protein and inability to detect protein isoforms when lacking isoform-specific antibodies[15,16]. Furthermore, in both cases, cells must be fixed and permeabilized when measuring intracellular targets. This can obscure discerning between surface-localized and internalized receptors, as well as introduce fixation artifacts such as epitope masking and changes in morphology and protein localization due to formation of diffusional gradients as fixation occurs[17–19].

A common method for overcoming the lack of specificity of immunoassays is to incorporate a protein separation (e.g., western blotting) before the immunoassay. Separating proteins by molecular weight resolves target protein signal from off-target binding events, as well as achieving detection of protein isoforms even when lacking an isoform-specific antibody. However, a standard western blotting requires ~10^3 cells for analysis and cannot produce the single-cell resolution achievable with flow cytometry and IF. The recently introduced single-cell polyacrylamide electrophoresis, or single-cell PAGE[20–22], enables performing single-cell measurements by employing microfluidic transport and photo-activatable protein capture chemistry to minimize diffusional losses during the electrophoresis (EP) and protein immobilization, or immunoblotting. Furthermore, covalent immobilization of separated proteins enables stripping of bound antibodies and re-probing for new targets, enables multiplexing of over 10 targets per single cell. Nonetheless, because this method performs whole-cell lysis before the EP step, it still cannot discern surface versus internal proteins.

Hence, in this chapter, we develop a workflow for single-cell PAGE of previously immuno-stained live cells. We focus on the epithelial cellular adhesion molecule or EpCAM, a surface receptor that is widely targeted for cancer therapy and regenerative medicine and for which certain mutations can cause the receptor to be absent from the cell membrane[23]. When staining cells with anti-EpCAM, we corroborated injection of the antibody-antigen immunocomplex into the polyacrylamide gel and measured a significant decrease in electrophoretic mobility of the EpCAM immunocomplex with respect to the free EpCAM antigen. We found that the migration of the antibody-antigen complex does not interfere with the migration of other proteins. When investigating the binding kinetics of the antibody-antigen complex under varying lysis and EP conditions, we found that the concentration of detergents in the lysis buffer is the primary factor in regulating immunocomplex stability. The ability to prepend surface receptor immunostaining with single-cell PAGE extends our understanding of how the under- or over-expression of surface receptor proteins controls the complex regulatory systems in single cells.

3.2 Materials and Methods

Antibodies. The primary antibodies in the surface staining characterization experiments include EpCAM-FITC (mouse, mAb, SAB4700424, Sigma), IgG-FITC (mouse, mAB, SA1-12320, Pierce), EpCAM-AlexaFluor 488 (mouse, mAb, 53-8326-42, Ebioscience), primary protein antibodies to GAPDH (goat pAb; SAB2500450, Sigma), β-Tubulin (rabbit pAb; ab6046, Abcam). Secondary antibodies to goat IgG prelabeled with Alexa Fluor 488 and 555 (A11055 and A21432) were purchased from Invitrogen.

Chemicals. 30%T, 2.7%C acrylamide/bis-acrylamide (37.5:1) (A3699), ammonium persulfate (APS, A3678), and tetramethylethylenediamine (TEMED, T9281) were purchased from Sigma-Aldrich. Triton X-100 (BP-151) was purchased from ThermoFisher Scientific. Premixed 10× Tris/glycine/SDS electrophoresis buffer (25 mM Tris, pH 8.3; 192 mM glycine; 0.1% SDS) was purchased from BioRad. Deionized water (18.2 MΩ) was obtained using an Ultrapure water system from Millipore. *N*-[3-[(3-Benzoylphenyl)formamido]propyl] methacrylamide (BPMAC) was custom synthesized by PharmAgra Laboratories[20,21].

SU8 and polyacrylamide (PA) gel fabrication. SU8 fabrication to generate the master and PA gel fabrication were performed as described previously[21]. The cell line experiments used a 7%T PA gel, and the microwell diameter and depth of 30 μm and 40 μm respectively. All PA gels on the single-cell PAGE slides were chemically polymerized with 0.08% APS and 0.08% TEMED.

Cell lines and surface staining. MCF7 cells were obtained from the American Type Culture Collection (ATCC) and authenticated (Promega). The MCF7 cell line was maintained in RPMI 1640 supplemented with 1% penicillin/streptomycin, 0.01mg/mL insulin (Invitrogen) and 10% FBS (Gibco). Cells were kept in a 37ºC incubator at 5% CO2. For surface staining, cells were harvested with 0.25% Trypsin EDTA and resuspended in 4ºC phosphate buffer solution (PBS) + 3% bovine serum albumin (BSA) at a concentration of ~10^7 cells/mL. In a 1.5 mL Eppendorf tube, 4ºC PBS+3% BSA, 5 μL of cell suspension, and the staining solution (anti-EpCAM Alexa 488 or IgG control) were added to a total volume of 500 μL. A control tube included only the PBS+3% BSA and cell suspension. Cells were stained for 30 min in the dark on ice. Tubes were centrifuged at 1000 RCF, and the supernatant was carefully removed. For the washing steps, the staining solution was aspirated and discarded from the single-cell PAGE and 400 μL of fresh 4ºC PBS was added to the cell pellet (2x). Prior to single-cell PAGE, cells were resuspended to a concentration of ~10^6 cells/mL. For the titration experiments, we tested a range of antibody concentration (0.1, 1.0, 3.0 and 5.0 μg/mL) and 3.0 μg/mL was used for the remaining experiments.

Single-cell PAGE protocol. Cells were pipetted over the PA gel and allowed to settled by gravity into the microwells patterned in the PA gel. Lysis buffer heated in a water bath to 50ºC was poured over the PA gel in order to lyse the cells in the microwells. An electric field (E = 40V/cm).was applied to inject and separate proteins in the PA gel abutting the microwell. After separation, proteins were immobilized to the gel matrix via UV activation (Lightningcure, LC5 Hamamatsu) of benzophenone methacrylamide cross-linked into the PA gel. Immobilized proteins were probed in-gel by diffusing fluorescently labeled antibody probes into the PA. A fluorescence microarray scanner (Genepix 4300A, Molecular Devices) equipped with 4-laser lines (488, 532, 594, 635) acquired fluorescence readout. Subsequent rounds of antibody stripping were performed for multiplexed protein analysis as detailed previously[21].

Single-cell PAGE and flow cytometry comparison experiments. MCF7s were stained with 0.1, 1.0, 3.0 and 5.0 μg/mL of the antibody solution, as previously described, and placed into four separate Eppendorf tubes. A negative control was prepared with unstained cells. For each tube, a 200 μL of the cell suspension was placed into a 96-well plate for flow analysis (Guava EasyCyte 6HT). The remaining cells were processed with by single-cell PAGE. The negative control with unstained cells was used for thresholding, and a total of 5000 events were counted per antibody concentration.

Immunoblot signal quantification and statistical analysis. The data sets generated during and analyzed during the current study are available from the corresponding author on request. Quantification of fluorescence signal from immunoblots used in-house scripts written in MATLAB (R2014b) as previously described[27]. Briefly, Gaussian curves were fit to fluorescence intensity profiles in MATLAB (R2014b, Curve Fitting Toolbox) and area-under-the-curve (AUC) was performed to quantify immunoblot signal. Statistical tests were performed with GraphPad Prism 7.0b.

3.3 Results and Discussion

3.3.1 Antibody-antigen immunocomplex readily migrates through the PA gel.

Given the importance of correlating surface receptor localization to activation of intracellular signaling, we sought to understand if surface EpCAM immunocomplexes are detectable using single-cell immunoblotting. Single-cell immunoblotting can analyze 100s to 1000s of individual cells in ~4 hours. The device consists of a 40 μm thick PA gel affixed to a standard microscope slide. The thin gel layer is stippled with 30 μm diameter microwells (Fig. 1, A). To investigate migration of immunocomplexes, cells are first immuno-stained with fluorophore-labeled antibodies to a surface receptor target (Fig. 1, B). Stained cells are then sedimented into the microwells designed to maximize single-cell-per-microwell occupancy (Fig. 1, C). A dual-functionality cell lysis and electrophoresis buffer is used to lyse cells, solubilize proteins, and support electrophoresis. After cell lysis, an electric field is applied to (i) electrophoretically inject proteins into the PA gel and (ii) separate proteins by single-cell PAGE. After protein separation, the migrated proteins are covalently bound to the PA gel by UV-mediated activation of benzophenone-methacrylamide monomers crosslinked into the PA gel. For protein detection, PA gels are immunoprobed with primary and fluorescently tagged secondary antibodies.

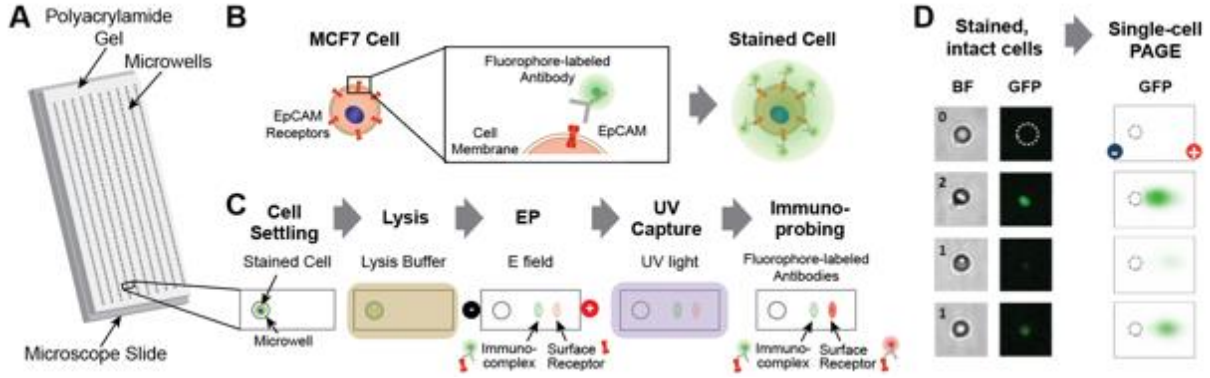


Figure 1. Concurrent detection of surface EpCAM immunocomplex and intracellular EpCAM using mobility shift single-cell electrophoresis. (A) Brightfield image of the single-cell PAGE device, showing a thin PA gel layer grafted on a microscope slide and stippled with an array of microwells. (B) Schematic of EpCAM receptors on surface of MCF7 cells stained with fluorophore-labeled anti-EpCAM antibodies. (C) Integration of live-cell immunofluorescence with the single-cell PAGE assay. Unfixed cells stained with fluorescently conjugated antibodies are settled into microwells for subsequent lysis, protein PAGE, and photo-blotting of separated proteins. Unstained targets can be immunoprobed with additional fluorophore-conjugated antibodies. (D) Bright field (BF) and fluorescence (GFP) micrographs of MCF7 cells stained with anti-EpCAM* and seated in microwells. False-colored fluorescence micrographs show surface EpCAM immunocomplex fluorescence blots after migration into the PA gel.

3.3.2 Antibody-antigen immunocomplexes do not interfere with the migration of other protein targets.

We next assessed the dual capacity of the surface receptor immunostaining complexes to (i) electromigrate into the molecular sieving gel (immunocomplex is ~290 kDa) and (ii) remain associated, even after single-cell lysis and electrophoresis. To support electrophoresis of a wide molecular mass range, we selected a moderate pore-size (7-8%T) PA gel and 3-4x longer separation distance than previously employed (1.5 – 2 mm)(Hughes et al)[17]. We simultaneously assayed unstained MCF7 cells, a population of stained MCF7 cells, and cells stained with an isotype-matched antibody control also labeled with AlexaFluor488. To control for device-to-device variation in electrophoretic migration, we interrogated cells for the loading control GAPDH. We did not see any significant difference the migration of GAPDH across the three experimental groups (Kruskall-Wallis statistic = 3.251, p value = 0.1968, 3 groups tested, N = 37 total) (Figure 2).

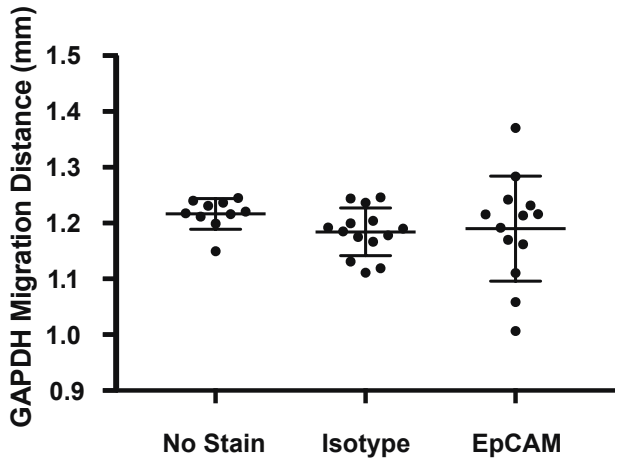


Figure 2. The migration distance of GAPDH for MCF7 cells stained with unstained MCF7 cells ("No Stain"), anti-EpCAM-1 stained MCF7 cells ("Stained"), and MCF7 cells stained with an isotype-matched antibody ("Isotype") shows no significant difference between the three groups tested (Kruskal-Wallis test, Kruskal-Wallis statistic = 3.251, p value = 0.1968, 3 groups with total $N = 37$).

We first examined the isotype-matched control and the unstained MCF7 experimental group for immunocomplex bands. As expected, neither group showed any signal for the AlexaFluor488 immunocomplex (Figure 3, a). Upon immunoprobing for EpCAM post single-cell PAGE, both experimental groups showed signal for free EpCAM (Figure 3, A). Next, we examined the single-cell PAGE blots for previously stained MCF7 cells. In this case, we detected immunocomplex bands after single-cell electrophoresis, with the intact complex peak detected 0.209 ± 0.02336 mm away from the center of the microwell on the separation axis (mean ± SD, $N = 37$, Figure 3a, green band). When probing the blots with fresh anti-EpCAM antibody, we detected a lower molecular mass protein band (Figure 3, A, blue band). To examine whether the two bands indeed correspond to different species, where the lower molecular mass is free EpCAM and the high molecular mass present before immunoprobing is the immunocomplex, we compared the migration distance of the peaks. Given that the bands indeed show a significant difference in migration distance (Mann Whitney U test, p value < 0.0001, $N = 37$) (Figure 3, B), we attribute the high molecular mass band to the anti-EpCAM-EpCAM complex; we hypothesize that the lower molecular mass peak corresponds to the intercellular fraction of EpCAM, which was not available for antibody binding

at the surface of the cells. Hence, integration of live-cell immuno-staining with subsequent single-cell PAGE provides a method for quantifying the localization of receptors to the surface of cells.

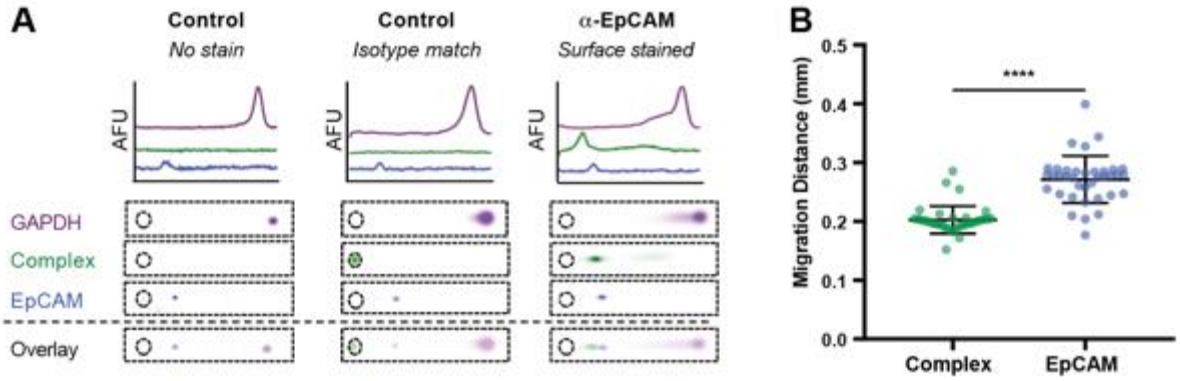


Figure 3. The scWB assay detects an electrophoretic mobility shift of the EpCAM immunocomplex. (a) Stained, unstained and isotype matched stained MCF7s were assayed by single-cell PAGE (Lysis buffer: 0.5% SDS, 0.1% Triton X-100, 0.25%, Na-DOC, 50°C, lysis for 20 s and EP for 25 s). Fluorescence intensity profiles for false-colored fluorescence micrographs of GAPDH, antibody-EpCAM immunocomplex and EpCAM are shown for the three conditions. (b) Violin plots of migration distance of the immunocomplex and free EpCAM demonstrates a reduction in electrophoretic mobility of the EpCAM immunocomplex with respect to the free EpCAM (Mann Whitney U test, p value < 0.0001, $N = 37$).

3.3.3 Benchmarking single-cell PAGE to flow cytometry and IF.

To benchmark the detection of immunocomplexes by single-cell PAGE with conventional single-cell protein technologies, we examined whether the fluorescence intensity distributions of anti-EpCAM stained cells obtained by single-cell PAGE were consistent with the gold standard assays of flow cytometry and immunofluorescence (IF).

In a first set of experiments, we examined whether flow cytometry and single-cell PAGE render similar distribution of fluorescence intensity over a range of initial anti-EpCAM-AlexaFluor 488 concentrations. To do so, we harvested MCF7 cells and stained them with varying concentrations of anti-EpCAM-AlexaFluor 488 (0.1, 1.0, 3.0 and 5.0 μg/mL). The cell suspensions stained at each concentration were then split into 2 vials for analysis by either single-cell PAGE or flow cytometry.

Fluorescence intensity distributions for single-cell PAGE and flow cytometry showed similar trends, where stained cells show high overlap when stained at 1, 3 and 5 μg/uL (for flow cytometry: 85.8% for 1 mg/mL vs 3 mg/mL, 81.2% for 1 mg/mL vs. 5 mg/mL and 87.8% for 3 mg/mL vs. 5 mg/mL; for single-cell PAGE: 61.7% for 1 mg/mL vs. 3 mg/mL, 67.7 % for 1 mg/mL vs. 5 mg/mL and 67.4 mg/mL for 3 mg/mL vs. 5 mg/mL) (Figure 4, a & b). Cells stained with 0.1 μg/mL, however, showed low overlap with high concentrations for both single-cell PAGE and flow cytometry (overlap of 0.1 μg/mL for flow cytometry: 6.62% for 1 μg/mL, 5.76% for 3 μg/mL and 7.44% for 5 μg/mL; overlap of 0.1 μg/mL for single-cell PAGE: 26.4% with 1 μg/mL, 29.9% with 3 μg/mL and 32.4 with 5 μg/mL). These results suggest that surface receptors were not saturated with anti-EpCAM antibody at 0.1 μg/mL, which can be determined through both flow cytometry and single-PAGE. In the case of flow cytometry, the ability to run a negative control allowed for

validation that 0.1 μg/mL shows higher overlap with the negative control (69.6% for negative control vs. 0.1 mg/mL) (Figure 4, a & b) than for the higher concentrations supports that at surface receptors were not saturated at 0.1 μg/mL.

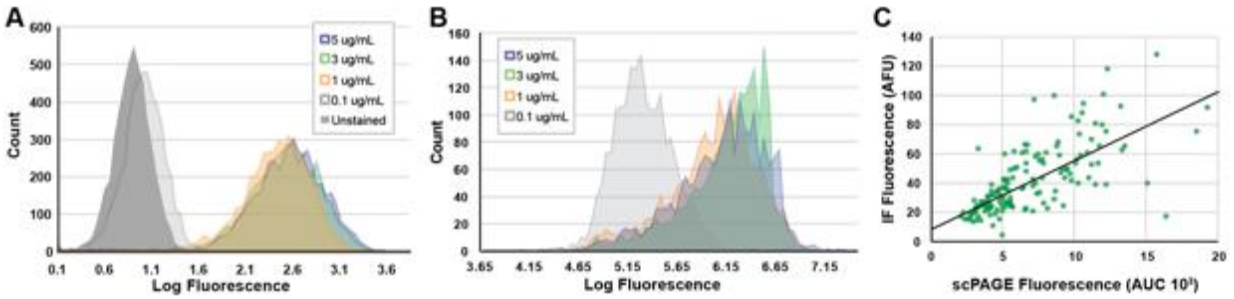


Figure 4. Benchmarking of single-cell PAGE with gold standard single-cell surface receptor measurements. Histograms of fluorescence intensity (Log Fluorescence) for MCF-7 cells stained with AlexaFluor488-labeled anti-EpCAM measured by (A) flow cytometry and (B) single-cell PAGE result in high distribution overlaps for antibody concentrations of 1.0. 3.0 and 5.0 μg/mL, but not 0.1 μg/mL. (C) Bivariate plot of fluorescence for FITC-anti-EpCAM stained MCF7 cells measured by immunofluorescence (IF) prior to lysis and single-cell PAGE after lysis and electrophoresis, showing a significant positive, linear correlation between fluorescence intensity measured by IF and single-cell PAGE (Pearson correlation, $\rho = 0.694$, p value < 0.00001, n = 148 microwells containing single cells).

We next investigated if the EpCAM immunocomplex signal obtained from IF correlated with the immunocomplex signal obtained from single-cell PAGE. MCF7 cells were stained with FITC-anti-EpCAM at 3 mg/mL in order to saturate EpCAM receptors. Cells were then settled into microwells of a single-cell PAGE device and imaged by fluorescence microscopy for bound FITC-anti-EpCAM. Single-cell PAGE was then performed and photocaptured anti-EpCAM immunoblots were quantified. We found a significant positive linear correlation between the IF-based measurement of bound anti-EpCAM and immunoblotted anti-EpCAM (Pearson correlation, $\rho = 0.694$, p value < 0.00001, n = 148 microwells containing single cells, Figure 4, b), suggesting that single-cell PAGE accurately estimates the anti-EpCAM antibodies bound to EpCAM receptors even after lysis, electrophoresis and photocapture.

3.3.4 Immunocomplex stability is affected by lysis and electrophoresis conditions

We next examined whether the degree of preservation of the EpCAM immunocomplex is a function of the lysis and EP parameters. We modified several parameters including lysis buffer composition, lysis time (t_{lysis}), and EP time (t_{EP}) applied to stained MCF7 cells and compared (i) EpCAM immunocomplex signal (AUC) and (ii) protein peak migration distance between different conditions.

We first investigated the effects of lysis buffer composition EpCAM immunocomplex disruption. The hybrid lysis-electrophoresis buffer includes the non-ionic detergent Triton X-100, that lyses cells by disrupting the phospholipid bilayer of the cellular membrane, and ionic detergents (sodium dodecyl sulfate (SDS) and sodium deoxycholate (Na-DOC)) that denature cellular proteins[24]. Varying detergent concentration can affect the stability of immunocomplexes by disrupting

antibody-antigen binding. To test the role of lysis buffer composition on immunocomplex stability, we formulated three lysis buffers, which we termed mild (0.5% SDS/0.1% triton X-100/0.25% Na-DOC), moderate (1.0% SDS/0.1% triton X-100/0.25% Na-DOC), and harsh (1.0% SDS/1.0% triton X-100/0.5% Na-DOC) and assayed MCF-7 cells stained with anti-EPCAM-AlexaFluor488. We first scrutinized the effects of buffer composition on immunocomplex electrophoretic migration. We observed a significant increase in the migration distance of the EpCAM immunocomplex protein peak from mild to moderate (130 ± 24.46 μm and 146.5 ± 8.06 μm, Mann Whitney U test p-value < 0.0001 for n = 387 and 670, respectively) and from moderate to harsh (146.5 ± 8.06 μm and 215 ± 19.56 μm, Mann Whitney U test p-value < 0.0001 for n = 670 and 59, respectively). We attribute the increase in migration distance to the increasing conductivity due to increasing concentration of detergents.

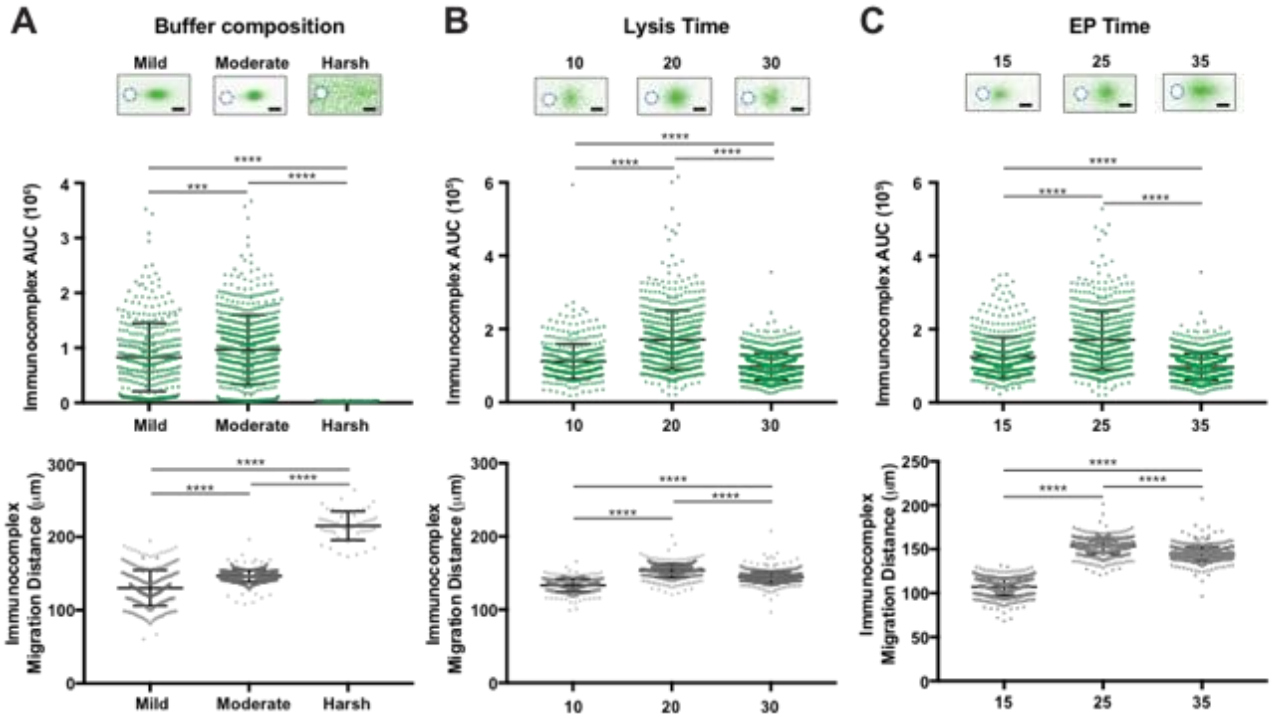


Figure 5. Cell lysis conditions determine compatibility between surface antibody immunofluorescence and single-cell western blotting. The effects of (A) buffer composition, (B) lysis time and (C) electrophoresis (EP) time on immunocomplex AUC and migration distance are displayed in the corresponding false-colored fluorescence micrographs (top) and violin plots (bottom) for each condition. Mann Whitney U tests, where '***' and '****' represent p values of <0.001 and < 0.0001. (A) N = 387, 670 and 59 cells for mild, moderate and harsh, respectively. (B) N = 364, 567 and 854 cells for t_{lysis} = 10, 20 and 30 s, respectively. (C) N = 539, 567 and 854 cells for t_{EP} = 15, 25 and 35 s, respectively.

We then inspected immunocomplex fluorescence signal (AUC) as a function of lysis buffer composition. We found that increasing the detergent concentration from the mild to moderate lysis buffer increased the immunocomplex signal (AUC, Mann Whitney U test, p-value < 0.001 for n = 387 and 670, respectively) (Figure 5, a), indicating higher degree of solubilization of proteins and thus increased injection of immunocomplex into the PA gel. On the other hand, increasing lysis buffer strength from moderate to the harsh resulted in a significant reduction in immunocomplex signal (Mann Whitney U test, p-value < 0.0001, for n = 59 and 670, respectively). Reduction in

EpCAM immunocomplex signal suggests that the concentration of detergents in the harsh lysis buffer disrupts antigen-antibody interactions.

Next, to scrutinize the effects of lysis duration on immunocomplex migration and preservation, we considered the trade-off between (i) allowing enough time for effective solubilization of cellular proteins, and (ii) minimizing the protein losses associated with the lysis step, that occur due to diffusion of lysate from the microwell to the buffer solution above the microwell. To investigate this trade-off, we varied the lysis time (t_{lysis}) for MCF7 cells (t_{lysis} = 10, 20, and 30 s) with the mild lysis buffer while maintaining a constant EP time of 25 s. We found a significant increase in EpCAM immunocomplex signal (AUC) from 10 s to 20 s, followed by a significant reduction in EpCAM immunocomplex signal from 20 s to 30 s (111753 ± 47532, 170859 ± 80099 and 97434 ± 36264, mean ± s.d. for 10, 20 and 30 s, respectively; Mann Whitney U test p-values were all < 0.0001) (Figure 5, b). We attribute these differences in immunocomplex to the balance between adequate solubilization of proteins, and thus effective injection of protein into the gel, and higher diffusion losses at longer lysis times. The reduction of intensity at 30s is attributed to increased protein loss as compared to 20 s. Hence, a lysis time of 20 s is appropriate for injecting the immunocomplex into the gel while minimizing protein loss by diffusion.

Finally, we investigated if preserving EpCAM immunocomplexes is also a function of EP duration. MCF7 cells were lysed for 20 s with mild lysis buffer and assayed at three EP times (t_{EP} = 15, 25, and 35 s). We observed a significant increase in EpCAM immunocomplex signal for 25 s and 35 s over 15 s (Mann Whitney U test, p values < 0.0001 for 15 s vs 25 s and 15 s vs 35 s, with N = 539, 567 and 854 cells for 15, 25 and 35 s respectively) (Figure 5, C). Interestingly, we found that increasing the tEP from 25 to 35 s caused a decrease in immunocomplex signal (Mann Whitney U test, p value < 0.0001, for N = 567 and 854, respectively) (Figure 5, C). As seen previously for the t_{lysis} = 10 s group, we attribute the lower immunocomplex signal at t_{EP} = 15 s to minimal injection of the complex into the gel. As the duration of EP is increased, the complex readily migrates into the gel and thus is retained in the PA matrix at the time of photocapture. We further hypothesize that the increasing temperature during longer EP times (i.e., 35 s) due to Joule heating also increases diffusional losses, leading to a lower immunocomplex signal than for the 25 s time point.

3.4 Conclusions

Surface staining of intact cells plays a critical role in identifying specific cell subpopulations. We introduce a single-cell mobility shift assay that reports surface receptor levels in unfixed mammalian cells. We surface stain cells with fluorescently labeled antibody after isolating individual cells in microwells. After isolation, we perform chemical cell lysis and single-cell protein PAGE, measuring a mobility shift between surface EpCAM immunocomplex and intracellular EpCAM. The mobility difference stems from the fact that EpCAM that is localized to the cell surface is accessible to anti-EpCAM* surface stain while EpCAM localized to the intracellular compartments is not and, thus, does not form an immunocomplex with the anti-EpCAM* stain. We detect the unbound, intracellular EpCAM by immunoprobing of single-cell PAGE.

During single-cell PAGE, we observe no notable interference between the large immunocomplex and intracellular proteins (i.e., no detectable impact on GAPDH migration distance for surface-

stained and unstained cells). Gold standard flow cytometry reports a similar anti-EpCAM* fluorescence distribution to that reported by single-cell PAGE for a range of anti-EpCAM concentrations (0.1 – 5 μg/mL). The results support the assertion that single-cell PAGE accurately measures surface-bound receptors through formation and mobility shift of immunocomplexes.

We were curious about the stability of the surface EpCAM immunocomplexes under cell lysis and electrophoresis conditions, and scrutinized the stability for a range of cell lysis buffers, lysis times, and electrophoresis times. Our analysis suggestions that surface EpCAM immunocomplex stability is most sensitive to the detergent concentration of lysis buffer, for the system and conditions described here. The integration of surface receptor staining of live cells with single-cell PAGE provides a single mobility shift readout representative of protein target localization and cell phenotype.

3.5 References

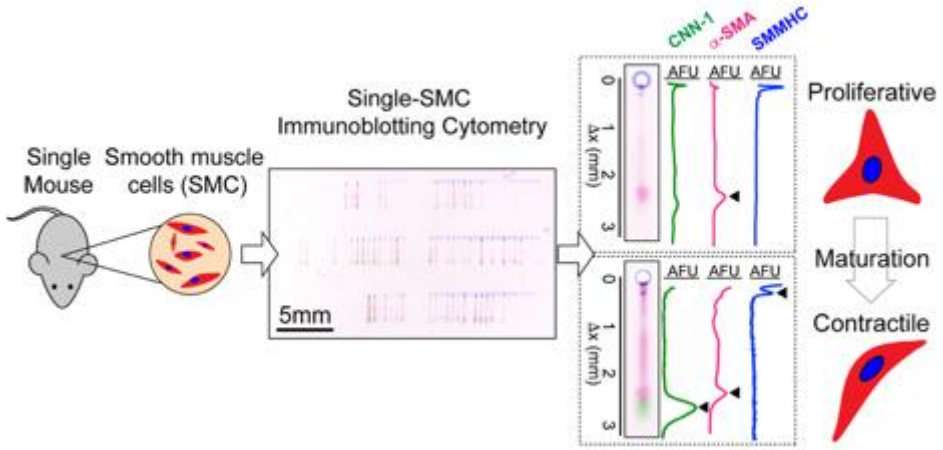
1. Uings, I. J. & Farrow, S. N. Cell receptors and cell signalling. *J Clin Pathol Mol Pathol* **53,** 295–9 (2000).
2. Schwartz, M. A. & DeSimone, D. W. Cell adhesion receptors in mechanotransduction. *Curr. Opin. Cell Biol.* **20,** 551–556 (2008).
3. Brockhoff, G., Hofstaedter, F. & Knuechel, R. Flow Cytometric Detection and Quantitation of the Epidermal Growth-Factor Receptor in Comparison To Scatchard Analysis in Human Bladder-Carcinoma Cell-Lines. *Cytometry* **17,** 75–83 (1994).
4. Fenderson, B. A., Miguel, M. P. De, Pyle, A. D. & Donovan, P. J. Antibodies to Stage-Specific Embryonic Antigens. *Methods Mol. Biol.* **325,** 207–224
5. Zhang, S. *et al.* Identification and Characterization of Ovarian Cancer-Initiating Cells from Primary Human Tumors. *Cancer Res.* **68,** 4311–4320 (2008).
6. Cheang, M. C. U. *et al.* Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Natl. Cancer Inst.* **101,** 736–750 (2009).
7. Lustberg, M. B. *et al.* Heterogeneous atypical cell populations are present in blood of metastatic breast cancer patients. *Breast Cancer Res.* **16,** (2014).
8. Prasad, B. M. Methods to detect cell surface expression and constitutive activity of GPR6. *Methods Enzym.* **484,** 179–195 (2010).
9. Ward, T. M. *et al.* Truncated p110 ERBB2 induces mammary epithelial cell migration, invasion and orthotopic xenograft formation, and is associated with loss of phosphorylated STAT5. *Oncogene* **32,** 2463–74 (2013).
10. Mitsuhashi, M. & Payan, D. G. Multiple signaling pathways of histamine H2 receptors. *J. Biol. Chem.* **264,** 18356–18362 (1989).
11. Wang, D. *et al.* Protein C receptor stimulates multiple signaling pathways in breast cancer cells. *J. Biol. Chem.* **293,** 1413–1424 (2018).
12. Perfetto, S. P., Chattopadhyay, P. K. & Roederer, M. Seventeen-colour flow cytometry: Unravelling the immune system. *Nat. Rev. Immunol.* **4,** 648–655 (2004).
13. Bendall, S. C., Nolan, G. P., Roederer, M. & Chattopadhyay, P. K. NIH Public Access. *Trends Immunol* **33,** 323–332 (2012).
14. Cram, L. S. Flow cytometry, an overview L. *Methods Cell Sci.* **24,** 1–9 (2002).
15. Bradbury, A. & Plückthun, A. Reproducibility: Standardize antibodies used in research. *Nature* **518,** 27–29 (2015).

16. Berglund, L. *et al.* A Genecentric Human Protein Atlas for Expression Profiles Based on Antibodies. *Mol. Cell. Proteomics* **7,** 2019–2027 (2008).
17. Schnell, U., Dijk, F., Sjollema, K. A. & Giepmans, B. N. G. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods* **9,** 152–158 (2012).
18. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *Elife* **5,** 1–24 (2016).
19. Chatterjee, S. Artefacts in histopathology. *J Oral Maxillofac Pathol* **18,** S111–S116 (2014).
20. Kang, C. C., Lin, J. M. G., Xu, Z., Kumar, S. & Herr, A. E. Single-cell western blotting after whole-cell imaging to assess cancer chemotherapeutic response. *Anal. Chem.* **86,** 10429–10436 (2014).
21. Hughes, A. J. *et al.* Single-cell western blotting. *Nat. Methods* **11,** 749–55 (2014).
22. Kang, C.-C. *et al.* Single cell–resolution western blotting. *Nat. Protoc.* **11,** 1508–1530 (2016).
23. Schnell, U., Cirulli, V. & Giepmans, B. N. G. EpCAM: Structure and function in health and disease. *Biochim. Biophys. Acta* **1828,** 1989–2001 (2013).
24. Qualtiere, L. F., Anderson, A. G. & Meyers, P. Effects of Ionic and Nonionic Detergents on Antigen-Antibody Reactions. *J. Immunol.* **119,** 1645–1651 (1997).

Chapter 4: Single-Smooth Muscle Cell Immunoblotting

*This work was conducted in collaboration with Tiffany Dai, Danny Wen-Chin Huang and Professor Song Li, and is reproduced from "Mouse-to-mouse variation in maturation heterogeneity of smooth muscle cells", Lab on a Chip; **18**:1875 (2018) with permission from all authors.*

Smooth muscle cell (SMC) heterogeneity plays an important role in vascular remodeling, a life-threatening hallmark of many vascular diseases. However, characterization of SMCs at the single-cell level is stymied by drawbacks of contemporary single-cell protein measurements, including antibody probe cross-reactivity, chemical fixation artifacts, limited isoform-specific probes, low multiplexing and difficulty sampling cells with irregular morphologies. This chapter focuses on the development of a high specificity, multiplexed single-cell immunoblot for unfixed, uncultured primary cells, in order to scrutinize healthy vessels for subpopulations of SMCs with proliferative-like phenotypes (Schematic 1).



Schematic 1: Schematic of SMC maturation profiling by single-cell electrophoretic cytometry.

4.1 Introduction

Understanding cell-to-cell phenotypic heterogeneity is crucial for elucidating the biological mechanisms of multicellular organisms.[1] Currently, the most prevalent single-cell studies involve transcriptomic analysis of cultured cells, which are readily available and can be expanded without limitation.[2–4] However, culturing cells often leads to the loss of their *in vivo* molecular phenotype.[2,5] Furthermore, although single-cell nucleic acid tools have led to tremendous advances, mRNA levels do not necessarily correlate with protein expression.[6] Immunofluorescence (IF) is the *de facto* standard for detecting and measuring protein expression of unmodified endogenous proteins in primary single cells.[7–9] Although invaluable, quantitative IF presents major drawbacks. Variable non-specific background signal results from ubiquitous antibody cross-reactivity[10,11], and accessing intracellular markers with antibody probes requires fixation of cells, which introduces critical artifacts, including epitope masking, changes in cell morphology and disruption of molecular binding events due to generation of diffusional gradients as fixation occurs.[12,13] In addition, image analysis algorithms used to segment tissue micrographs into individual cells yield variable results when cell morphologies are complex[14], when borders

between cells are low-contrast,[15] or when samples contain crowded cells, such as the closely associated SMCs in the blood vessel wall.[9,16,17] Other widely used techniques that release cells from tissues, such as fluorescence activated cell sorting (FACS), also suffer from low antibody selectivity and require large sample sizes (thousands to millions of cells), often requiring pooling of samples from multiple animals. Although sample pooling can help reach these high sample requirements, it can also lead to: (a) biological averaging, where the assumption that protein expression in the pooled sample is equivalent to the mean of individual samples does not hold for all genes, (b) variance reduction, where sample pooling can hide relevant biological variance, and (c) dilution effects, where proteins showing high expression in individual samples can be diluted or lost when pooling with other samples.[18–20] Consequently, a critical gap exists in high-selectivity, single-cell resolution protein analysis tools suitable for profiling of sparingly available primary cells from a single donor.

In the field of vascular biology, single-cell IF studies of mouse aortas revealed an important milestone in our understanding of vascular remodeling. Vascular remodeling is the disruption of the discrete layers of blood vessels due to the abnormal proliferation of vascular smooth muscle cells (SMCs), and is a hallmark of numerous vascular diseases such as hypertension, diabetic macroangiopathy and atherosclerosis.[21–24] The established paradigm depicted SMCs as a homogeneous population exhibiting high contractility and low proliferation, that in response to injury, *de-differentiate* into a proliferative phenotype.[5] However, this generally accepted theory was recently challenged by IF studies that discovered a subpopulation of SMCs with proliferative capabilities.[25] In these studies, culturing primary SMCs revealed that only a subset of cells is capable of proliferating, and that this subpopulation of proliferative SMCs has a distinct phenotype in their native tissue; high expression of early stage differentiation marker α smooth muscle actin (α-SMA), and low or negative expression of middle and late-stage differentiation markers calponin 1 (CNN-1) and smooth muscle myosin heavy chain (SMMHC), respectively. Lineage tracing studies of blood vessels in transgenic mice further confirmed that proliferation of the smooth muscle layer is not a unified and escalating process among SMCs, as previously speculated, but results from the proliferation of just one or two single SMCs.[26,27] Given this paradigm shift, new questions arise over whether these proliferative SMCs exist as a subpopulation in the healthy vessel wall, as well as whether this subset of SMCs can be characterized through protein expression of key markers. Answering these questions requires that primary SMCs be derived and directly analyzed from tissue for single-cell protein expression profiles. However, no single-cell technique capable of analyzing the heterogeneity of SMCs in native blood vessels addresses the challenges of (a) low antibody selectivity or (b) low sample availability.

To enhance selectivity beyond IF, researchers commonly perform an electrophoretic protein separation, that separates proteins by molecular mass, upstream of immunoassays. This assay, known as an immunoblot or western blot, spatially resolves proteins to discriminate signal contributions by molecular mass. Although western blots grant superior selectivity compared to IF, until recently the assay has lacked the analytical sensitivity needed for single-cell resolution.[9] Microfluidic technologies have advanced single-cell measurement capabilities by working with micrometer scales suitable for handling minute sample volumes, comparable to the size of single cells.[28] We recently introduced microfluidic single-cell lysate electrophoresis (EP) and immunoblotting for highly specific protein measurements.[29–32] However, existing systems have been primarily designed for spherical cell morphologies, and use gravity-based settling [9,29–32]

and/or use cell-isolation in microwells for cell selection by size to ensure one-cell-per-microwell occupancies.[32,33] Analysis of primary cells is hindered by low sample availability and non-spherical cellular morphologies.

To directly analyze SMC heterogeneity in native blood vessels at the single-cell level, we designed an assay for high-specificity protein profiling of primary cells. We applied our assay to freshly isolated murine aortic SMCs in search of distinct subpopulations, to support or refute the existence of cells that are preconditioned to proliferate in response to injury. We assayed SMCs for markers α-SMA, CNN-1 and SMMHC, characterizing early, middle, and late differentiation stages. We demonstrate high-selectivity measurements of protein expression in sparse cell populations (starting cell populations of 100s of SMCs) to report single-SMC profiling of individual mouse aortas, eliminating the need for sample pooling or cell expansion through *in vitro* culture. The straightforward integration of the microfluidic device with ubiquitous micromanipulation systems shows promise as a widespread bioanalytical method in biological laboratories.

4.2 Materials and Methods

Chemical reagents. Tetramethylethylenediamine (TEMED, T9281), ammonium persulfate (APS, A3678), β-mercaptoethanol (M3148), and 30%T, 2.7%C acrylamide/bis-acrylamide (37.5:1) (A3699), bovine serum albumin (BSA, A9418) were purchased from Sigma-Aldrich. Triton X-100 (BP-151), phosphate-buffered saline (PBS, ph 7.4, 10010023) and Dulbeccos's phosphate-buffered saline (PBS, 14190144) were purchased from ThermoFisher Scientific. Premixed 10X tris-glycine EP buffer (25 mM Tris, pH 8.3; 192 mM glycine; 0.1% SDS) was purchased from BioRad. Tris buffered saline with Tween-20 (prepared from 20X TBST, sc-24953, Santa Cruz Biotechnology, Dallas, TX, USA). Deionized water (18.2 MΩ) was obtained using an Ultrapure water system from Millipore. N-[3-[(3-Benzoylphenyl)formamido]propyl] meth- acrylamide (BPMAC) was custom synthesized by Pharm-Agra Laboratories (Brevard, NC, USA).

Device Fabrication. Devices were fabricated using SU-8 wafers as previously reported.[30] Microwells were 100 μm in diameter and 80 μm deep. Microwell spacing was 5 mm along the separation axis and 1 mm traverse to the separation axis. Each device was fabricated with an array of 120 microwells, with lateral spacing between microwells of 1 mm, and vertical spacing of 5 mm. The 5-mm gaps served as the separation lane length. Polyacrylamide precursor solution (7%T acrylamide/bis-acrylamide, 3 mM BPMAC) was chemically polymerized with 0.08% APS and 0.08% TEMED.

Enzymatic digestion of mouse aorta to dissociate SMCs. All experimental procedures with mice were approved by the ACUC committee at UC Berkeley and carried out according to institutional guidelines. All efforts were made to minimize the suffering and number of animals used. Before experiments, SMMHC-CreER^{T2}/LoxP-tdTomato adult mice were given daily intraperitoneal injections of 2 μg of tamoxifen in 100 μL of corn oil for 5 days and used for analysis a week afterward. Mice were euthanized via CO_2 exposure with a CO_2 flow rate of 2 L/min for 5 min. Mice were then inspected for cessation of movement and respiration for at least 1 min. A secondary form of euthanasia, cervical dislocation, was performed.

After harvesting, the aorta was submerged in a dish of cold sterile PBS and stripped of the outer tunica adventitia layer consisting of fibroblasts and connective tissue. The inner endothelial cell layer was denudated by passing the rounded tip of a surgical wire through the vessel back and forth 3 times. The remaining tunica media layer comprised of smooth muscle cells (SMCs) was cut into short segments and placed in a microcentrifuge tube filled with cold sterile PBS. To isolate SMCs from the aorta, the vessel segments were first incubated in a solution of 1 mg/mL collagenase in DPBS containing calcium and magnesium for 10 min. The solution was then discarded, as preliminary digestion of the aorta with collagenase served to minimize contamination of SMCs by other cell types. Next, the aorta was incubated in a solution of 1 mg/mL collagenase and 0.125 mg/mL elastase in DPBS containing calcium and magnesium. During digestion, the aorta was agitated on an orbital shaker set to 50 rpm at 37 ^{o}C for 50 min. The solution was pipetted up and down to further disperse the digested extracellular matrix and spun down at 1500 rpm for 5 min. The cell pellet was then resuspended in PBS and passed through a cell strainer with 40 μm pores to remove ECM debris and cell clusters**.**

Cell settling & single-cell lysate separations. The final cell suspension contained ~1,000 tdTomato-positive SMCs, 50-100 of which were sampled into each device, depending on whether SMCs from one or two mice were sampled onto the same chip. To place single tdTomato-expressing SMC into microwells, we used a micromanipulator (Transferman®, Eppendorf, Germany) under an inverted epifluorescence microscope (Zeiss AxioObserver, Carl Zeiss MicroImaging GmbH, Germany). *In situ* cell lysis was performed by pouring lysis buffer (25 mM Tris-glycine buffer at pH 8.3 with 1% SDS, 0.5% sodium deoxycholate, and 1% Triton X-100) warmed to 55^{o}C. An electric field (E = 40 V / cm) was applied to separate the proteins (PowerPac^{TM} Basic Power Supply, Bio-Rad Laboratories). Protein bands were immobilized by UV activation of the benzophenone moieties (Lightningcure LC5, Hamamatsu, Japan). Slides were probed with antibodies and scanned with a fluorescence microarray scanner (Genepix 4300A, Molecular Devices, CA, USA).

Antibodies. Antibodies employed for analysis of SMCs include: rabbit anti-□-Tubulin (ab6046), mouse anti-SMMHC (ab683), mouse anti-α-SMA (ab7817), rabbit anti-CNN-1 (ab46794) from Abcam, Cambridge, MA. Donkey secondary antibodies AlexaFluor 647-conjugated anti-mouse (A31571), AlexaFluor 594-conjugated anti-mouse (A21203) and AlexaFluor 488-conjugated anti-rabbit (A21206) were purchased from ThermoFisher Scientific CA, USA. All antibodies were used at 1:10 dilution in 2% BSA in TBST.

Image processing, separation performance quantification, statistical analysis and clustering. Quantification of fluorescence signal in protein blots employed in-house scripts written in MATLAB (R2017a, Mathworks). Parameters such as peak location and peak width were obtained by fitting Gaussian curves to protein bands using MATLAB's Curve Fitting Toolbox. Gamma distribution fitting, goodness of fit tests, agglomerative hierarchical clustering, and clustergram heatmap rendering were performed with the Matlab Statistics and Machine Learning Toolbox. Statistical tests (d'Agostino & Pearson normality tests, Mann Whitney U tests) were performed using GraphPad Prism 7.0b. Fiji was used to false-color fluorescence micrographs and overlay channels to create composite images. Any differences in final cell number reported per mouse arise from the sampling step (i.e., the number of SMCs initially sampled onto the device) and post-assay

analysis of immunoblots, where some separation lanes contained particulates that prevented accurate Gaussian fitting and area-under-the-curve-analysis of the fluorescence intensity profile.

4.3 Designing a single-SMC protein separation tool for analysis of murine aortic SMCs

Differentiation of SMCs into a mature, contractile state can be traced by three maturation markers: α-SMA, CNN-1, and SMMHC (Fig. 1, A). To overcome artifacts introduced by poor antibody selectivity when measuring endogenous proteins, we designed a workflow for assaying primary SMCs that prepends a molecular mass-based separation to the immunoaffinity measurements.

The SMC analysis workflow (Fig. 1, B) begins with isolation of SMCs from the aorta of one mouse. Enzymatically dissociated SMCs are then sampled into individual microwells patterned onto a thin (80 μm) polyacrylamide (PA) gel layer using a micromanipulator (Fig. 1, B). Care was taken to ensure unbiased sampling of SMCs into microwells. Namely, cells were sampled if they appeared as a completely dissociated individual cell, independently of their size (Fig. 1, C). Furthermore, to prevent contamination of microwells with other vascular cell types, we used transgenic mice labeled with tdTomato under the control of the SMMHC promoter (specific for SMCs) and fluorescence microscopy to select only tdTomato-positive cells. If the use of non-transgenic mice is required, the multiplexing capability of the immunoblot allows for probing of cell-specific markers to rule out contamination with other cell types.

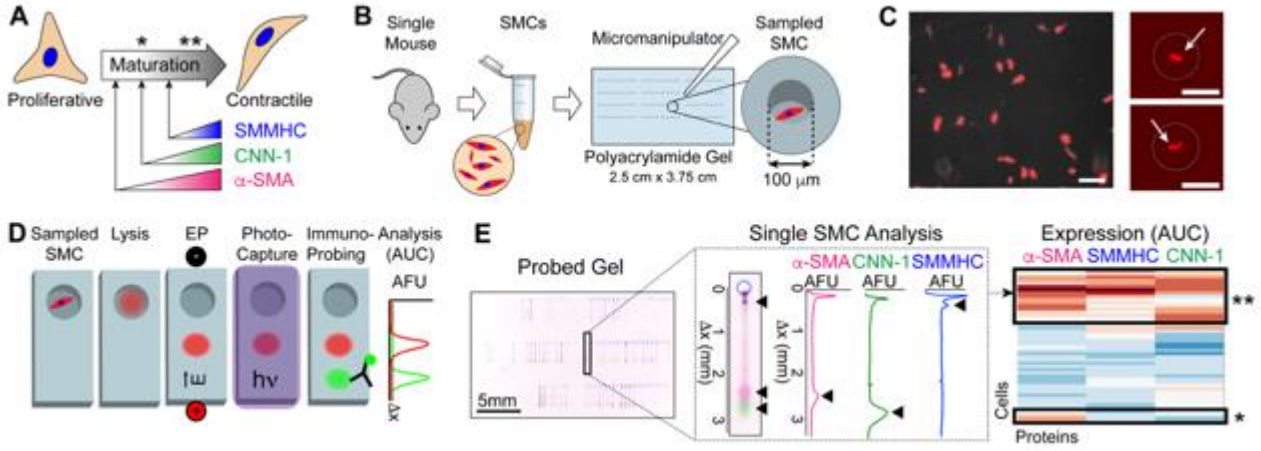


Figure 1 Immunoblotting cytometry profiles aortic SMCs freshly dissociated from a single mouse aorta. (A) Schematic of SMC differentiation. Onset of expression of three proteins mark the differentiation of SMCs to contractile phenotype (α-SMA, CNN-1 and SMMHC). (B) Schematic of cell isolation and sampling. The aorta of a single mouse is enzymatically digested. tdTomato-positive aortic SMCs are then placed with a micromanipulator into microwells (100 μm diameter) of a thin polyacrylamide (PA) gel (80 μm height) grafted onto a glass microscope slide. (C) Fluorescence micrographs of freshly isolated tdTomato-positive SMCs in cell suspension (left) and sampled into microwells (right). Arrows indicate SMCs, microwells are traced with dashed line. Scale bars are 100 μm. (D) Immunoblotting cytometry stages. SMCs are lysed and solubilized proteins are electrophoresed across the PA gel layer by application of an electric field. UV light immobilizes migrated proteins into the PA gel by activation of photoactive benzophenone methacrylamide moeities in the gel matrix. Gel is incubated with solutions of fluorescently-

labeled antibodies and area-under-the-curve (AUC) analysis is performed on protein band fluorescence intensity profiles to quantify protein expression. (E) On left, false-colored micrograph of PA gel immunoprobed for all three differentiation markers, from which a single separation lane is shown (center) along with corresponding fluorescence intensity profiles. Black arrows mark position of protein bands and peaks. On right, clustergram of α-SMA, CNN-1 and SMMHC expression for all single SMCs analyzed from an individual mouse. Asterisks mark subpopulations identified and mapped back to maturation stage in panel (A).

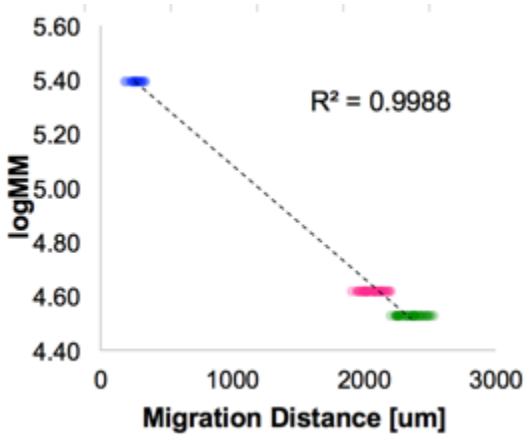


Figure 2: Biaxial plot demonstrating the log-linear relation between protein molecular mass and migration distance ($R^2 > 0.99$). Protein markers are SMMHC (blue, 227 kDa), α-SMA (magenta, 42 kDa) and CNN-1 (green, 34 kDa) (N = 47 single SMCs from one mouse).

With a buffer-exchange step, settled SMCs are treated with ionic and non-ionic detergents to achieve *in situ* cell lysis and protein solubilization (Fig. 1D). Application of an electric field causes injection of proteins through the PA gel interface at the microwell wall and migration of proteins through the sieving gel matrix. Each microwell has an associated 5 mm long *separation lane*, or area of gel in which proteins are separated by molecular mass. Exposure to UV light activates benzophenone methacrylamide moieties incorporated in the hydrogel matrix, that covalently capture the separated proteins. Covalent protein immobilization facilitates the buffer exchanges necessary for immunoprobing, while minimizing dilution and protein losses. Area-under-the-curve (AUC) analysis of proteins bands is performed to quantify expression of SMC maturation markers and allows identification of SMC subpopulations based on variable expression of α-SMA, CNN-1 and SMMHC (Fig. 1E). We verified a log-linear relation between protein molecular mass and migration distance ($R^2 = 0.99$), consistent with SDS-PAGE separations (Figure 2), and achieved separation of proteins with molecular mass range of 34 to 227 kDa. To assess cross-contamination between immunoblots on neighboring microwells, immunoblotting of the contents of empty microwells are employed as negative controls. For all devices, we observed no detectable immunoblot signal for the microwells with zero cells per microwell occupancy (Figure 3). To facilitate investigating inter-mouse and intra-mouse heterogeneity in SMC maturation stage, the workflow is designed to assay single-SMCs from starting populations of just 100s of SMCs from a single mouse aorta.

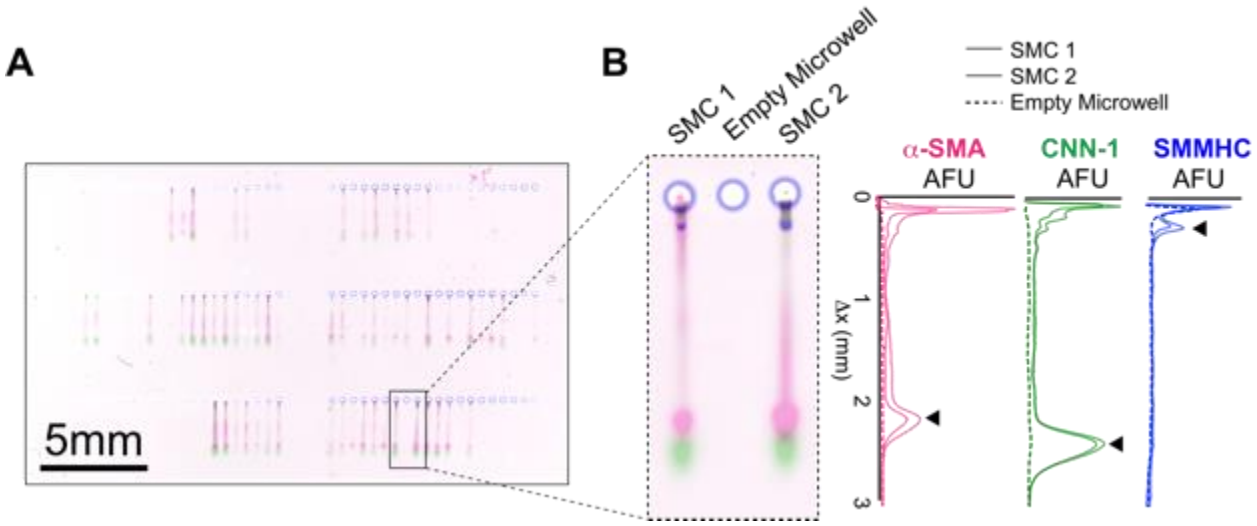


Figure 3: Empty microwells serve as negative controls. (A) False-colored fluorescence micrograph of single-cell immunoblot probed for α-SMA, CNN-1 and SMMHC. (B) False-colored micrograph of three adjacent single-cell immunoblots and corresponding fluorescence intensity profiles. Microwells loaded with SMCs (left and right) show immunoblot signal for all three markers (solid lines), while the negative control (empty microwell) shows no immunoblot signal for either marker (dashed lines). Black arrows mark position of protein peaks.

4.3.1 Microfluidic Sample Preparation and Immunoblot Development for Single-SMC Lysate

To enable profiling of SMC differentiation stage, we designed our single-cell immunoblot to address a pair of contradictory constraints, namely, the need to minimize dilution of protein lysate while: (a) providing microwell geometries that are large enough for isolation of large, anisotropic SMCs, and (b) providing sufficient protein solubilization and EP durations to size large molecular mass maturation marker proteins.

In satisfying our first constraint on microwell geometry and lysate dilution, we observed that primary SMCs conserved their spindle-shaped morphology and displayed a range of cell lengths (80 to 120 μm) (Figure 4, A). To isolate individual cells, cylindrical microwells are well-suited for capturing spherical cells of similar diameters, such as isolation of cells from suspensions of dissociated cultured cells. Size exclusion from the microwells ensures occupancies of one-cell-per-microwell [30]. However, settling spindle-shaped SMCs into cylindrical microwells requires large microwell diameters to accommodate the long axis of the SMCs, that leads to settling of multiple SMCs into the same microwell. We discarded the use of non-cylindrical microwells, such as a projection of an oblong feature, as these would reduce the probability of sampling due to limited ability to orient cells during sedimentation. To ensure one-cell-per-microwell occupancies, we actively seated individual SMCs into microwells using a micromanipulator, following established protocols for microtransfer of single cells.[33,34] We observed that micromanipulation allowed settling single SMCs into a minimum microwell diameter of 100 μm, where cell deformability and microwell height (80 μm) aided in settling of SMCs with lengths exceeding 100 μm.

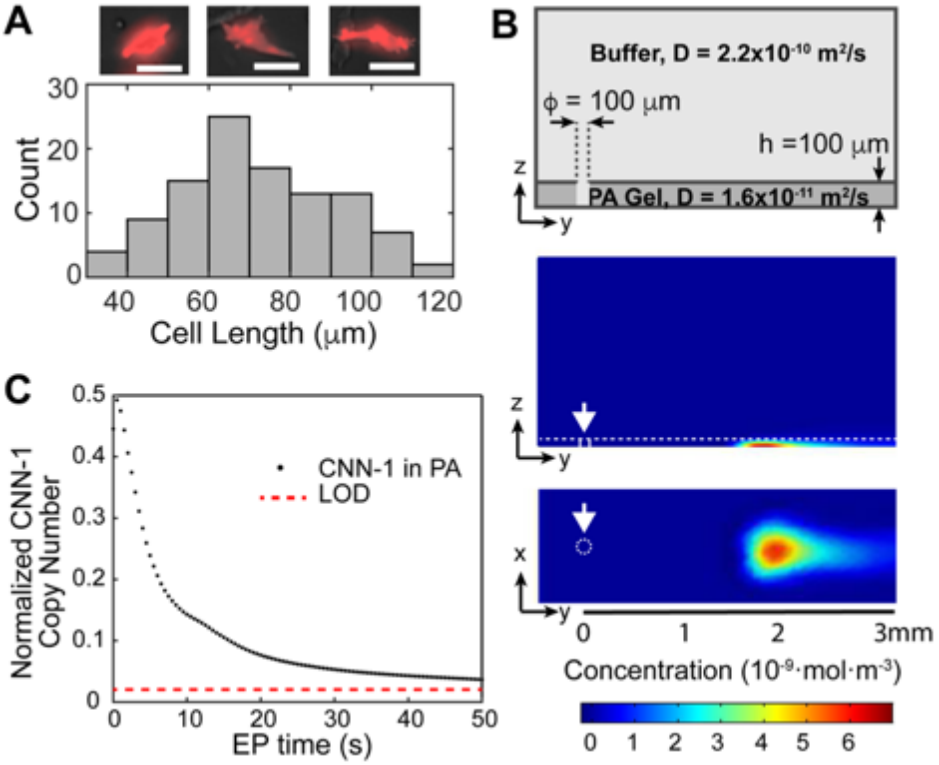


Figure 4: Numerical simulation of protein losses during lysis and EP determine final concentrations of target markers remain above the LOD. (A) Histogram of SMC length (N = 105). Fluorescence micrographs show representative tdTomato-tagged SMCs with lengths of 60 μm, 70 μm and 100 μm. Scale bars are 50 μm. (B) Numerical simulation of CNN-1 protein diffusion during lysis and EP migration. Schematic of microwell and separation lane (top). Heatmaps of CNN-1 concentration after 40 seconds of lysis and 50 seconds of EP, computed with Comsol® Multiphysics (bottom). White arrows mark location of microwell. (C) Scatter plot of retained CNN-1 proteins in the PA gel (normalized to initial protein copy number) during EP. LOD of 27,000 molecules is represented by red dotted line.

During cell lysis, solubilized lysate diffuses both in-plane (into microwell volume and surrounding PA gel layer) and out-of-plane (buffer layer above PA gel layer). Although thermodynamic partitioning helps confine cell lysate to the microwell ($0.001 < K < 0.1$ for 7 %T, or total monomer concentration in g / 100 mL) [35], a ~60x dilution occurs as the ~10 pL SMC lysate fills the 628 pL microwell volume (100 μm diameter x 80 μm tall). Consequently, we selected cell lysis conditions to rapidly lyse SMCs and quickly solubilize even large proteins in SMC lysate. To assess protein solubilization, we monitored electromigration of α-SMA into the PA gel comprising the microwell walls, with electromigration as a proxy for solubilization. For lysis buffer composition, we fixed lysis and EP conditions (lysis duration 25 s, electromigration duration 25 s, E = 40 V/cm, 55ºC) and determined performance with the commonly employed RIPA buffer (25 mM Tris-glycine buffer at pH 8.3, 0.5 % SDS, 0.1 % Triton X-100) and a more concentrated buffer previously optimized for lysis-hardy cells (25 mM Tris-glycine buffer at pH 8.3 with 1% SDS, 0.5% sodium deoxycholate, and 1% Triton X-100[33]). While we observed no noticeable electromigration of α–SMA from the microwell and into the PA with RIPA buffer, under the same conditions we observed full electromigration of the α-SMA from the microwell and into the PA gel layer for the

harsher buffer (Figure 5). Given the composition of the latter, we anticipate notable Joule heating during EP of each single-SMC lysate, which exacerbates both in-plane and out-of-plane diffusional losses of lysate proteins.

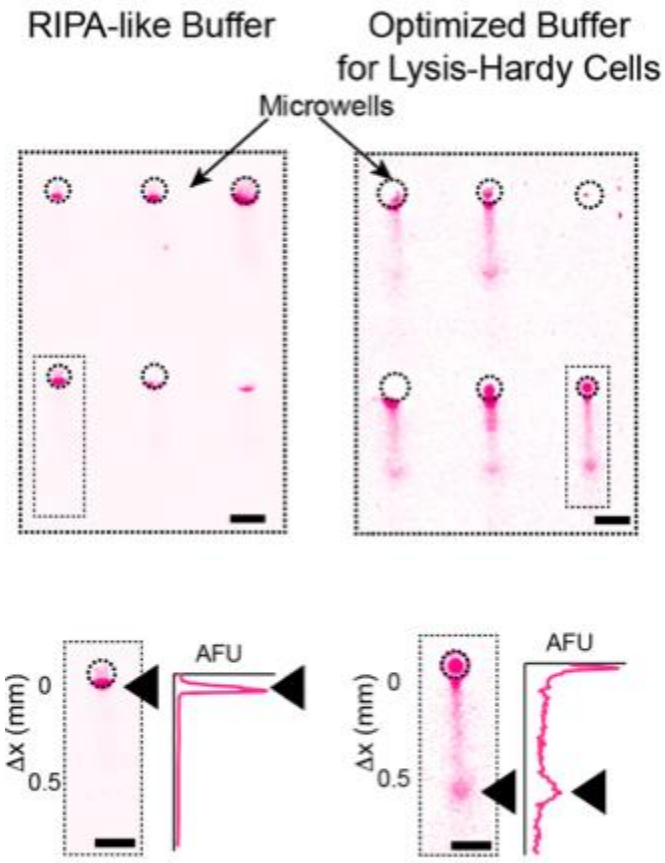


Figure 5: Experimental determination of SMC lysis. False-colored micrographs of probed α-SMA after lysis with RIPA-like lysis buffer (left) and lysis buffer optimized for lysis-hardy cells (right). On bottom, fluorescence intensity profiles are shown for representative samples of each. RIPA-like lysis buffer demonstrates high signal at edge of microwell (left micrograph, black arrow), indicating poor solubilization of α-SMA. The buffer optimized for lysis-hardy cells, however, showed a band in separation lane (right, black arrow) and signal at edge of microwell. For both buffers, running conditions were: lysis duration 25 s, electromigration duration 25 s and E = 40 V/cm, 55ºC.

To satisfy our second constraint on protein marker dilution, protein solubilization, EP duration, and the limit of detection (LOD) of the single-cell immunoblot, we numerically simulated the transport of CNN-1 during in-microwell cell lysis, electromigration into the EP sieving gel, and subsequent EP analysis (Figure 6, B). The LOD was previously determined by partitioning known concentrations of proteins into PA gels and using fluorescently labeled antibodies to probe for captured proteins [29]. For this analysis, we considered CNN-1, as this species has the lowest molecular mass (34 kDa) of the maturation marker set and thus, the highest diffusivity. We set the lysis time to 40 s (experimentally determined as time required for full lysis of SMCs using the harsher buffer warmed to 55 ºC) and the EP time to 50 s (time required at an E = 40 V / cm for injection of MHC, the protein target with highest molecular mass, 227 kDa and thus, slowest

electrophoretic mobility). Based on the simulation results, we estimate a ~ 50% loss of protein during lysis in the open microwell geometry. Nevertheless, after the 40 s cell lysis period and 50 s of elapsed EP separation time, the CNN-1 protein copy number remains above the LOD of the open fluidic immunoblot assay (Figure 4, C).

4.3.2 Testing correlation between protein signal and off-target signal

We experimentally verified our simulation results by assaying single SMCs under the simulated conditions (100 μm-diameter microwells, 40 s lysis, 50 s EP) and successfully separating and detecting CNN-1, α-SMA, and SMMHC. The triad of protein markers – α-SMA, CNN-1 and SMMHC – were assayed in SMC populations from individual aortas at an average of 65 ± 19 SMCs per device (mean ± S.D. for N = 3 separate devices). Upon performing single-SMC protein separations, we observed a spurious band associated with the α-SMA antibody (Figure 6, A), consistent with several reports in which the use of the same antibody showed non-specific bands at higher molecular masses [36]. Interestingly, we observed that this off-target signal appeared only in a subset of cells, and with varying intensity (Figure 6, B). To determine if the off-target signal was correlated with the α-SMA protein band signal, we quantified both bands for a set of single SMCs (Figure 6, C, N = 47). When plotting the protein band AUC against the off-target signal AUC, we observed a low positive correlation (R^2 = 0.65, Figure 6, C). This low correlation suggests that an assay that uses immuno-affinity alone, such as IF, could lead to inaccurate measurements of protein expression. For instance, as shown in Figure 6, B, while Cell 1 and Cell 3 have similar α-SMA expression (Cell 3 to Cell 1 ratio = 1.09), if we add the non-specific band signal contribution is to the α-SMA signal, Cell 3 appears to have a 2.3-fold increase in α-SMA signal.

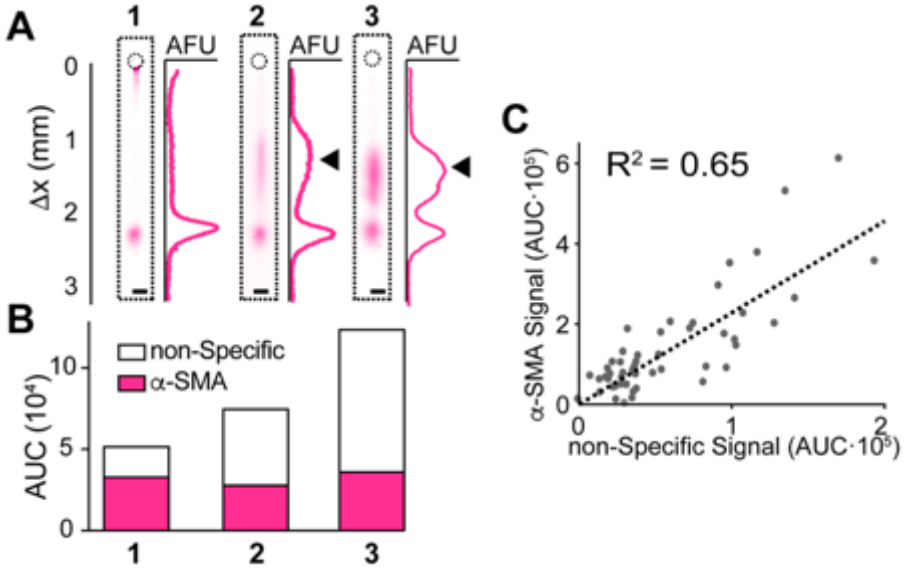


Figure 6 Immunoblotting cytometry detects SMC-to-SMC variation in off-target antibody signal. (A) False-colored micrographs and corresponding intensity profiles for three analyzed SMCs showing varying non-specific band intensity (black arrows). (B) Stacked bar graph of quantified α-SMA bands (magenta) and non-specific bands (white) for three SMCs shown in (A). (C) Bivariate scatter plot illustrating the poor linear correlation (R^2 = 0.65) between the AUC of the α-SMA and the non-specific bands (N = 47 single SMCs for one independent experiment with an individual mouse).

These results emphasize the need to perform thorough antibody validation experiments while taking into account how different sample preparation methods can mask epitopes or modify binding affinities. Namely, caution must be taken when validating antibodies by one method and then applying them to an assay with different sample preparation. For instance, validation of an antibody by western blotting, where proteins are usually denatured and treated with reducing agents, does not guarantee successful application in IF, where proteins not only maintain their native structure, but may also be subject to artifacts due to cell fixation, like epitope masking due to methylene bridge formation when using paraformaldehyde.[11,37]

4.3.3 Mouse-to-Mouse Variation in SMC Maturation

Low sample requirements of 100 s of cells allow us to assay SMCs from individual mice and examine mouse-to-mouse variation. We first validated run-to-run reproducibility to ensure comparison of data collected on different devices. To perform this analysis, we compared the protein band peak width of SMCs assayed on two separate devices. Because the final peak width of a protein band depends on (a) injection dispersion, which is affected by the degree of protein solubilization, and (b) diffusive band broadening that occurs during EP, which is a function of gel pore size, we used this separation performance metric to validate run-to-run lysis efficiency and gel-to-gel reproducibility. We thus assayed SMCs from one mouse on two different devices, and compared the peak width distributions of CNN-1, the marker with highest diffusivity. The peak width was calculated as 4σ, where σ is the standard deviation evaluated by fitting a Gaussian curve to the fluorescence intensity profile. The two devices showed peak widths of 627 ± 66.7 μm and 591 ± 66.8 μm, which were not significantly different, demonstrating reproducibility between gels and similar levels of solubilization between runs (mean ± S.D., N = 22 and 11 SMCs, Mann Whitney U Test, *P* value threshold 0.05, *P* value = 0.2381).

Next, to investigate whether differences in protein distribution exist between mice, we assayed SMCs from two different mice for maturation markers CNN-1, α-SMA, and SMMHC. We first validated that protein expression distributions for all protein markers fit a gamma distribution, consistent with single cell protein expression profiles reported for mammalian cells and which results from stochasticity in transcription and translation.[38,39] For Mouse 1, the expression distributions of all markers were well described by gamma distributions, as determined by fitting a gamma distribution to each data set and performing a Chi Squared Goodness of Fit Test between the data and the fitted Gamma distribution (*P* values above the threshold of 0.1005, 0.4247 and 0.2158 for α-SMA, CNN-1 and SMMHC, respectively). For Mouse 2, while CNN-1 and SMMHC expression distributions were well described by a Gamma distribution (*P* values of 0.2070 and 0.1866, respectively), the distribution of α-SMA was poorly described by a Gamma distribution (*P* value = 0.0169). For both mice, the lower *P* values for α-SMA suggest that α-SMA is a stronger classifier of subpopulations of SMCs, as compared to other protein markers studied here. Given that α-SMA is a marker of earlier differentiation, these results suggest the existence of immature-like SMCs in the healthy aortic wall and support the theory of subpopulations of SMCs being responsible for proliferation.

Next, we calculated the coefficient of variation (CV = S.D. / average x 100) for each protein marker to validate the detection of biological variation, and not technical noise. CVs for α-SMA and CNN-

1 were similar for both mice: 70.9 % and 70.4 % for α-SMA and 60.9 % and 60.2 % for CNN-1, for Mice 1 and 2, respectively. Mouse 1 showed a higher CV for SMMHC than Mouse 2 (80.5 % and 66.6%, respectively). All CV values were above the technical variation threshold CV 32.4 % determined previously.[33] Briefly, this threshold was calculated by running single-cell separations with GFP-expressing cells, binning cells with < 5 % difference in initial GFP fluorescence and calculating the CV of the probed signal. CV values are all above this threshold, indicating that the variation detected is true biological variation and not technical noise.

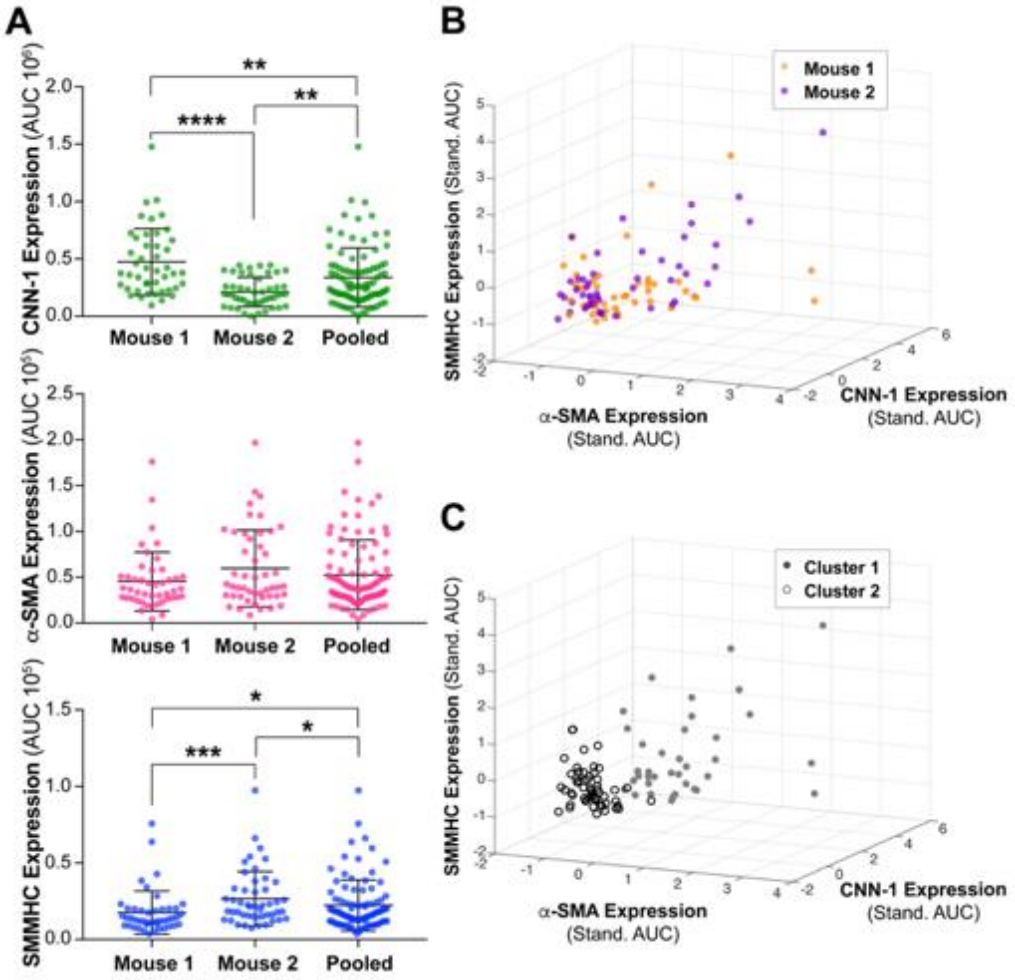


Figure 7: Mouse-specific expression of maturation stage markers reveal mouse-to-mouse heterogeneity, which cannot be recovered when samples are pooled. (A) Violin plots displaying the protein expression (AUC) of CNN-1, α-SMA, and SMMHC of SMCs isolated from two mice, shown individually and pooled. Horizontal black lines represent the mean and mean ± SD. '*', '**', '***' and '****' indicate Mann-Whitney U Tests that resulted in *P* values of < 0.05, < 0.01, < 0.001 and < 0.0001, respectively. An independent experiment was conducted for each individual mouse. Mice 1 and 2 had *N* = 45 and 48 single SMCs, respectively. (B) Three-dimensional scatter plot showing standardized expression of CNN-1, α-SMA and SMMHC by single SMCs for Mice 1 and 2. (C) K means clustering (cosine similarity, k = 2) was performed on pooled, standardized CNN-1, α-SMA and SMMHC expression data. Resulting clusters, shown by open "o" or gray filled "•" circles on the three-dimensional scatter plot of CNN-1, α-SMA and SMMHC expression, demonstrate a low classification accuracy (Rand Index of 0.495).

We investigated mouse-to-mouse heterogeneity by comparing expression distributions for Mouse 1 and Mouse 2. α-SMA expression did not show a significant difference between mice (Mann Whitney U Test, *P* value = 0.114). CNN-1 and SMMHC, however, showed different expression distributions in Mouse 1 compared with Mouse 2 (*P* values < 0.0001 and 0.0004 for CNN-1 and MHC, respectively). To determine whether the pooled distribution is representative of SMCs within the individual mice, we compared each protein expression distribution to the pooled distribution. α-SMA was the only marker for which distributions were not significantly different (*P* values for Mann Whitney U test between pooled distribution and Mice 1 and 2 were 0.1144 and 0.3720, respectively). For CNN-1 and SMMHC, however, the pooled distribution was significantly different from the individual protein expression distributions (Mann Whitney U test, P values for pooled distribution vs. Mouse 1 were < 0.0001 for CNN-1 and 0.0004 for SMMHC, and for pooled distribution vs. Mouse 2 were 0.0040 for CNN-1 and 0.0466 for SMMHC) (Fig. 6A). These results highlight the importance of performing analyses at the single-mouse resolution, as for two out of three markers studied here, CNN-1 and SMMHC, pooling samples results in distributions that are not representative of SMCs within the aorta of the two individual mice.

Finally, we tested whether inter-mouse variation could be recovered from a pooled sample. We standardized expression data for all three markers for Mouse 1 and Mouse 2 (Figure 7, B) and performed clustering on pooled data (K means clustering, cosine similarity, with k = 2 clusters). Resulting clusters are shown in Figure 7, C. For Mouse 1, 40.0% of SMCs were classified into Cluster 1 and 60.0% into Cluster 2. Similarly, 39.6% of SMCs from Mouse 2 were clustered into Cluster 1 and 60.4% into Cluster 2. To evaluate clustering accuracy, we calculated an external validation index, the Rand Index (RI). The values of RI range between 0 and 1, where values approaching 1 indicate a high level of agreement between clustering and the natural classes. For our clustering results, we obtained an RI of 0.495, indicating low accuracy when clustering the data pooled from two mice. These results demonstrate that mouse origin cannot be easily recovered once cells have been pooled. Single-mouse studies therefore render valuable animal-to-animal phenotypic variation that cannot be recovered if cells from multiple mice are initially pooled.

4.3.4 Identifying Subpopulations of SMCs

The quantitative and specific nature of immunoblotting – as compared to IF – provides insight into subtle differences among single SMCs. To investigate if SMCs showed distinct subpopulations, we performed agglomerative hierarchical clustering on SMCs from an individual mouse aorta (Figure 8, A). In order to exclude any effects of cell size, we used cosine similarity as the distance metric. The majority of SMCs (92 %) were clustered into two major populations demonstrating similar levels of expression across all markers, consistent with the expected expression profile of a mature SMC. One of these populations shows generally high expression for all markers (Figure 8, A, top gray rectangle), whereas the other shows generally low expression across the three maturation markers (Figure 8, A, middle gray rectangle). A representative micrograph and corresponding fluorescence intensity profile for these two populations are displayed in Figure 8, B, marked with '*' and '**' for the high and low expression populations, respectively. In the second cluster, a minority of cells were segmented into a subgroup of SMCs exhibiting higher expression of α-SMA than SMMHC and CNN-1 (Figure 8, A, red rectangle). We observed a similar trend for SMCs from a different individual mouse aorta, where 6% of the population

showed high α-SMA and low expression of CNN-1 and SMMHC, consistent with immature-like phenotype (Figure 9). These results are consistent with immunofluorescence studies that demonstrated the existence of a subpopulation of proliferative SMCs, comprising less than 10% of the total SMC population.[25] The representative micrograph and protein separation of this subpopulation (Figure 8, B, '***'), demonstrate the elevated expression of α-SMA with respect to CNN-1 and SMMHC. The identification of subpopulations with marked phenotypic differences indicates that an inherently heterogeneous population of SMCs exists within the normal blood vessel wall, even without injury. Furthermore, the detection of a subpopulation (<10%) with immature-like phenotype, namely, high expression of early-stage maturation marker α-SMA but low expression of late-stage maturation markers CNN-1 and SMMHC, corroborate findings that suggest that only subset of proliferative SMCs may be responsible for vascular remodeling [25–27].

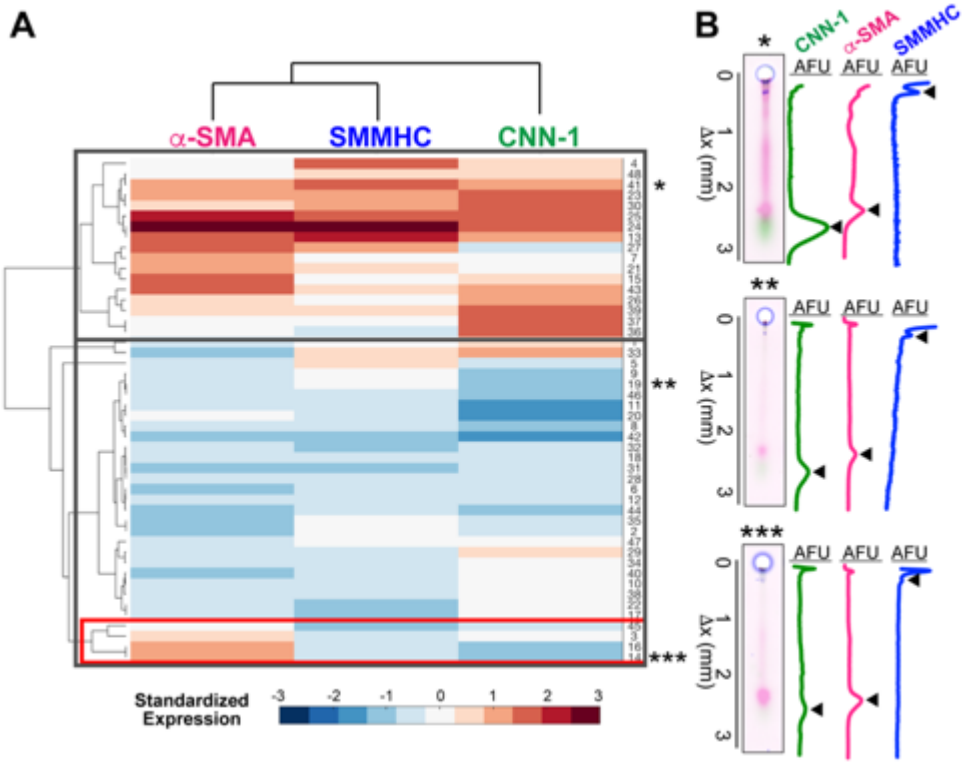


Figure 8: Hierarchical clustering of individual SMCs by expression of α-SMA, CNN-1 and SMMHC identifies SMC subpopulations. (a) Heat map of SMCs clustered by expression of α-SMA, CNN-1 and SMMHC. Cells are plotted on the y-axis, while protein markers are plotted on the x-axis (N = 48 single SMCs for one independent experiment). The majority (92 %) of SMCs are clustered into two major populations, one displaying generally high (top gray rectangle) and the other generally low (middle gray rectangle) expression of all three markers. The latter cluster includes a subpopulation of SMCs with high α-SMA and low CNN-1 and SMMHC expression (red rectangle). (b) False-colored micrographs and corresponding fluorescence intensity profiles for representative samples from the subpopulations marked in panel (a). Black arrows mark the position of protein peaks. Scale bars are 100 μm.

Although three protein markers were sufficient for profiling SMC maturation stage in this study, other biological questions may require profiling a larger panel of protein targets. To validate our multiplexing capability beyond 3 targets, we calculated the theoretical *peak capacity*, or maximum number of resolved peaks that fit in a separation lane. Considering the average peak width 798 ±

133 μm of the marker with highest diffusivity, CNN-1 (average ± S.D., N = 92) and total length of the separation lane, 5 mm, we obtained a peak capacity of 7.3. With 7 fully resolved theoretical peaks and four spectral channels of our current imaging system, we anticipate a multiplexing capability of > 30 markers. In addition, we can strip the gel of bound antibodies (by incubation with 2% SDS, 0.8% □-mercaptoethanol and 62.5 mM Tris base at 55°C) and re-probe for new targets. Multiplexing beyond 30 markers extends the applicability of our assay to address complex biological questions that may require studying signaling pathways or screening large panels of protein markers.

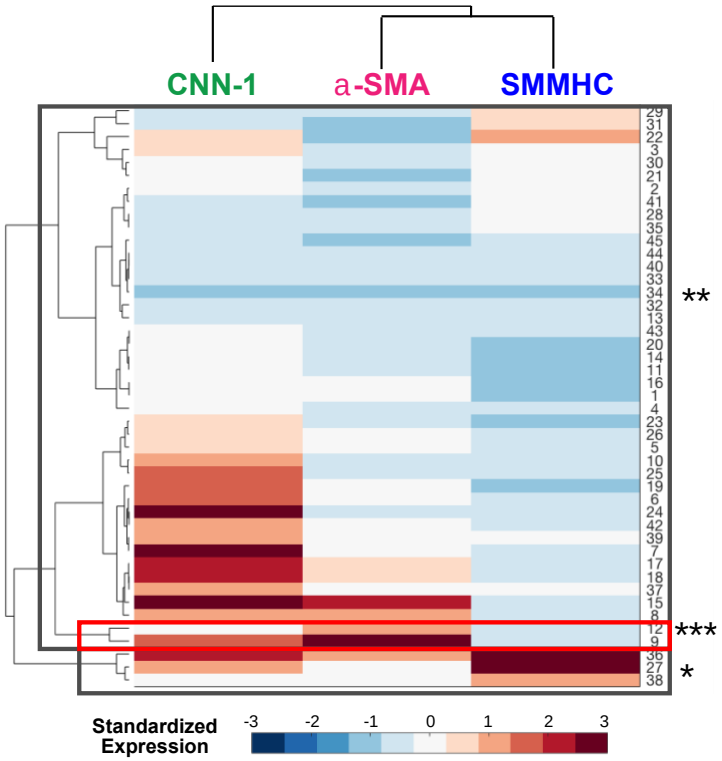


Figure 9: Hierarchical clustering of individual SMCs by expression of α-SMA, CNN-1 and SMMHC identifies subpopulations of SMCs. Heat map of hierarchical clustering of SMCs by expression of α-SMA, CNN-1 and SMMHC, where cells are plotted on the y-axis and protein markers are plotted on the x-axis (cosine similarity, N = 45 single SMCs). The majority (96 %) of SMCs are clustered into two major populations, one displaying generally high (*) and the other generally low (**) expression of all three markers. Inside the latter group, a subpopulation (6%) of SMCs show high α-SMA and low CNN-1 and SMMHC expression (red rectangle, ***), corresponding with the phenotype of immature-like SMCs.

4.4 Maturation profiling of human vascular SMCs isolated from injured vs. healthy vessels

We next explored if our findings in mouse aortic SMC translated to SMCs in human vessels. We thus profiled the maturation stage of vascular SMCs isolated from a healthy vessel and an injured vessel belonging to the same human donor, by measuring proteins related to early-stage and late-stage maturation. In the case of human cell lines, SOX-10 and CNN-1 serve as early and late-stage SMC differentiation markers. SOX-10 is expressed in a subpopulation of SMCs believed to show

stem cell-like *in vivo* phenotypes as well as to show proliferative capabilities when isolated and expanded when cultured *in vitro*[25].

4.4.1 Adapting the immunoblotting device for morphologically heterogeneous human SMCs

Our first observation was that although human SMCs presented a rounded morphology, unlike the spindle-shaped mouse SMCs, they displayed heterogeneity in cell size (~ 20 to 100 μm in length, Figure 9, A). This unexpected heterogeneity caused (i) sampling bias towards cells of < 30 mm in diameter (corresponding to the diameter of the microwells) (Figure 9, A, black arrow shows cell too large to enter microwell), and (ii) settling of multiple < 30 mm diameter cells into the same microwells (Figure 10, A, white arrow).

To overcome the sampling bias and occupancies larger than one-cell-per-well, we modified the diameter of the microwells patterned on the PA gel. Instead of patterning microwells of homogeneous diameter, we fabricated microwells with a gradient-diameter (20, 30, 50, 75 and 100 mm) (Figure 10, B). We achieved settling of all cells, including those over 30 μm in diameter, and confined smaller cells in microwells with diameters closer to their own (Figure 10, B).

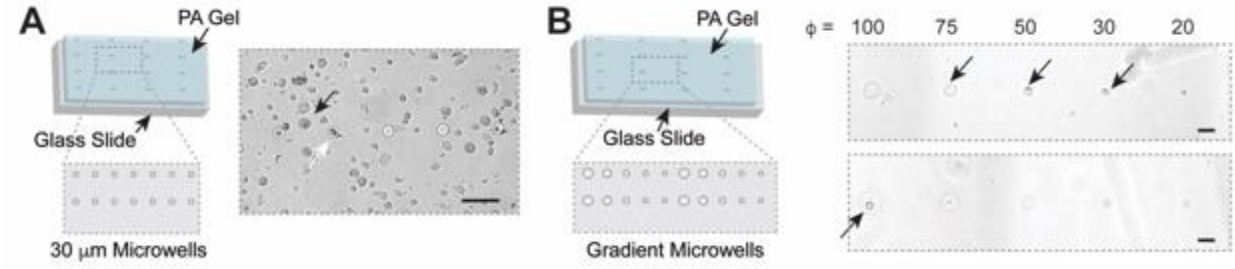


Figure 10: Adapting the microfluidic immunoblotting microwells to accommodate a gradient in cell sizes. (A) On the left, schematic of immunoblotting device and the 30 μm microwells patterned on the PA gel. To the right, bright field micrograph of a SMCs (human cell line isolated from an injured vessel) settled onto the immunoblotting device. Black arrow points to a cell with a diameter larger than the microwell diameter, while the white arrow points to a microwell containing two cells of diameters < 30 um. (B) On the left, schematic of the modified device, where PA gel is patterned with microwells of a gradient in diameter (100, 75, 50, 30 and 20 μm). Bright field micrographs of sections of the gradient microwell PA gel will SMCs settled into microwells. Black arrows mark microwells containing single cells. Scale bars are 100 μm.

4.4.2 Maturation-stage profiling of diseased vs. healthy human SMCs

After successful unbiased sampling of both human SMC cell lines, we optimized lysis and EP conditions to resolve SOX-10 and CNN-1 from one other as well as from non-specific bands. By lysing cells with RIPA-like lysis buffer (25 mM Tris-glycine buffer at pH 8.3, 0.5 % SDS, 0.1 % Triton X-100) warmed to 55ºC for 20 s, followed by EP at E = 40 V/cm for 30 s and photocapture for 45 s at 100 % power, we successfully resolved SOX-10 and CNN-1 (Figure 9, A). In the SOX-10 immunoblots, we observed several non-specific bands (Figure 11, A), suggesting that

quantifying total immunoprobed signal without first resolving proteins by molecular mass yields inaccurate results.

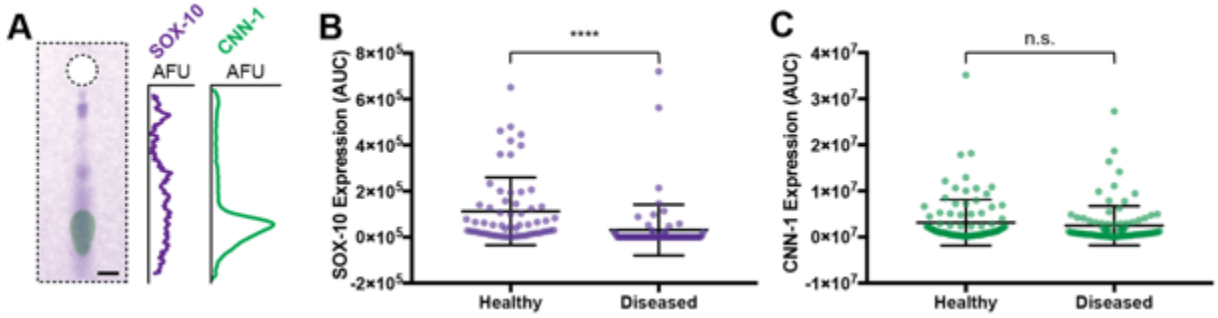


Figure 11: Single-SMC microfluidic immunoblotting allows for maturation-stage profiling of healthy vs. diseased human SMCs. (A) False-colored overlaid micrograph of SOX-10 and CNN-1 immunoblot from a single human SMC. Corresponding fluorescence intensity profiles for SOX-1 and CNN-1 are shown to the right of micrograph. (B) (Mann Whitney U tests, p value 0.0893 for CNN-1 with N = 99 and 96 for healthy and diseased and p value < 0.0001 for SOX-10 with N = 61 and 73 SMCs for healthy and diseased, respectively).

We performed AUC analysis on the intensity profiles of the single-SMC immunoblots and found that while the distributions for CNN-1 expression did not significantly differ (Mann Whitney U tests, p value 0.0893 for CNN-1 with N = 99 and 96 for healthy and diseased, respectively), SMCs isolated from the injured vessel showed a significant decrease in SOX-10 expression compared to SMCs isolated from a healthy vessel (Mann Whitney U test, p value < 0.0001 for SOX-10 with N = 61 and 73 SMCs for healthy and diseased, respectively)(Figure 11, B). We hypothesize that the lower SOX-10 expression in SMCs from the injured vessel suggests that cells presenting immature-like phenotype differentiate in response to injury. This behavior supports the model that immature-like cells already residing in the healthy vessel are responsible for proliferation in response to injury.

4.5 Conclusions

Studying cell-to-cell phenotypic variation of unfixed, primary cells is a crucial capability, as culturing cells leads to losses in molecular phenotype. Here we introduce a high-selectivity, multiplexed immunoblotting cytometry assay for interrogating single primary cells. We applied our assay to demonstrate maturation stage profiling of aortic SMCs freshly isolated from individual mice. After ensuring unbiased sampling of SMCs (80 – 120 μm in length), we performed single-SMC electrophoretic protein separations, that resolve protein signal from off-target antibody binding, and immunoblotted for differentiation markers α-SMA, CNN-1 and SMMHC (targets ranging from 34 kDa to 227 kDa). Using our assay, we demonstrate the detection of a population of SMCs with an immature phenotype (high α-SMA and low SMMHC and CNN-1), which supports the theory that healthy blood vessels harbor a subset of SMCs responsible for the abnormal proliferation of SMCs seen in numerous vascular diseases. When comparing the maturation stage of SMCs isolated from healthy and injured vessels from a human donor, we found

that SMCs from the injured vessel expressed lower SOX-10 than those from a healthy vessel, suggesting that indeed (i) the healthy human vessel contained cells with immature-like phenotypes and (ii) SMCs from the injured vessel may have begun to respond to injury by differentiating into mature-like SMCs.

Beyond advancing knowledge in vascular biology, our key goal to minimize primary sample requirements reveals important insight into variation in cell phenotype not only within the same donor, but also between different donors. Although human cell-based *in vitro* models are gaining importance, animal models remain the gold standard for both drug testing and biological discovery.[40,41] However, isolating specific tissues or cell groups is difficult and often results in small yields. High sample requirements of protein assays (western blotting or FACS) often necessitate pooling cells from multiple mice, which not only obscures essential differences in individual responses, but also significantly increase the number of animals required. 16 billion dollars are spent annually on animal models,[42] with low successful rates of translation- less than 8% in clinical cancer trials, for instance.[43] Low reproducibility may also be connected with the inability of bulk assays to detect patterns or responses unique to individual organisms. Single-mouse resolution assays, such as the one described in this paper, provides insight into mouse-to-mouse phenotypic variability, reduces the economic costs associated with animal models, and minimizes animal lives sacrificed.

4.6 References

1. Oikawa, A. & Saito, K. Metabolite analyses of single cells. *Plant J.* **70,** 30–38 (2012).
2. Pan, C., Kumar, C., Bohl, S., Klingmueller, U. & Mann, M. Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. *Mol. Cell. Proteomics* **8,** 443–450 (2009).
3. Hakulinen, J. K. *et al.* Structural study on the architecture of the bacterial ATP synthase Fo motor. *Proc. Natl. Acad. Sci. U. S. A.* **109,** E2050-6 (2012).
4. Lee, M.-C. W. *et al.* Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E4726-35 (2014).
5. Slomp, J. *et al.* Differentiation, dedifferentiation, and apoptosis of smooth muscle cells during the development of the human ductus arteriosus. *Arterioscler. Thromb. Vasc. Biol.* **17,** 1003–9 (1997).
6. Newman, J. R. S. *et al.* Single-cell proteomic analysis of S. cerevisiae reveals the architecture of biological noise. *Nature* **441,** 840–6 (2006).
7. Gosselin, E. J., Cate, C. C., Pettengill, O. S. & Sorenson, G. D. Immunocytochemistry: Its evolution and criteria for its application in the study of epon-embedded cells and tissue. *Am. J. Anat.* **175,** 135–160 (1986).
8. Burry, R. W. Controls for Immunocytochemistry: An Update. *J. Histochem. Cytochem.* **59,** 6–12 (2011).
9. Yamauchi, K. A. & Herr, A. E. Subcellular western blotting of single cells. *Microsystems Nanoeng.* **3,** 16079 (2017).
10. Schnell, U., Dijk, F., Sjollema, K. A. & Giepmans, B. N. G. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods* **9,** 152–158 (2012).

11. Bordeaux, J. *et al.* Antibody validation. *Biotechniques* **48,** 197–209 (2010).
12. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *Elife* **5,** 1–24 (2016).
13. Chatterjee, S. Artefacts in histopathology. *J Oral Maxillofac Pathol* **18,** S111–S116 (2014).
14. Dorini, L. B., Minetto, R. & Leite, N. J. White blood cell segmentation using morphological operators and scale-space analysis. *XX Brazilian Symp. Comput. Graph. Image Process.* 294–304 (2007).
15. Chen, S., Zhao, M., Wu, G., Yao, C. & Zhang, J. Recent advances in morphological cell image analysis. *Comput. Math. Methods Med.* **2012,** (2012).
16. Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7,** R100 (2006).
17. Kamentsky, L. *et al.* Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* **27,** 1179–80 (2011).
18. Diz, A. P., Truebano, M. & Skibinski, D. O. F. The consequences of sample pooling in proteomics: An empirical study. *Electrophoresis* **30,** 2967–2975 (2009).
19. Zhang, S. D. & Gant, T. W. Effect of pooling samples on the efficiency of comparative studies using microarrays. *Bioinformatics* **21,** 4378–4383 (2005).
20. Rajkumar, A. P. *et al.* Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics* **16,** 548 (2015).
21. Hadrava, V. *et al.* Vascular smooth muscle cell proliferation and its therapeutic modulation in hypertension. *Am Hear. J* **122,** 1198–203 (1991).
22. Chistiakov, D. A., Orekhov, A. N. & Bobryshev, Y. V. Vascular smooth muscle cell in atherosclerosis. *Acta Physiol.* **214,** 33–50 (2015).
23. Caplice, N. M. *et al.* Smooth muscle cells in human coronary atherosclerosis can originate from cells administered at marrow transplantation. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 4754–9 (2003).
24. Haas, S. *et al.* Real-time monitoring of relaxation and contractility of smooth muscle cells on a novel biohybrid chip. *Lab Chip* **10,** 2965 (2010).
25. Tang, Z. *et al.* Differentiation of multipotent vascular stem cells contributes to vascular diseases. *Nat. Commun.* **3,** 1–13 (2012).
26. Feil, S. *et al.* Transdifferentiation of vascular smooth muscle cells to macrophage-like cells during atherogenesis. *Circ. Res.* **115,** 662–667 (2014).
27. Chappell, J. *et al.* Extensive Proliferation of a Subset of Differentiated, yet Plastic, Medial Vascular Smooth Muscle Cells Contributes to Neointimal Formation in Mouse Injury and Atherosclerosis Models. *Circ. Res.* **119,** 1313–1323 (2016).
28. Yang, M., Nelson, R. & Ros, A. Toward Analysis of Proteins in Single Cells: A Quantitative Approach Employing Isobaric Tags with MALDI Mass Spectrometry Realized with a Microfluidic Platform. *Anal. Chem.* **88,** 6672–6679 (2016).
29. Hughes, A. J. *et al.* Single-cell western blotting. *Nat. Methods* **11,** 749–55 (2014).
30. Kang, C.-C. *et al.* Single cell–resolution western blotting. *Nat. Protoc.* **11,** 1508–1530 (2016).
31. Kang, C. C., Lin, J. M. G., Xu, Z., Kumar, S. & Herr, A. E. Single-cell western blotting after whole-cell imaging to assess cancer chemotherapeutic response. *Anal. Chem.* **86,** 10429–10436 (2014).
32. Kim, J. J., Sinkala, E. & Herr, A. E. High-selectivity cytology via lab-on-a-disc western

blotting of individual cells. *Lab Chip* **17,** 855–863 (2017).
33. Sinkala, E. *et al.* Profiling protein expression in circulating tumour cells using microfluidic western blotting. *Nat. Commun.* **8,** (2017).
34. Hannemann, J. *et al.* Quantitative high-resolution genomic analysis of single cancer cells. *PLoS One* **6,** (2011).
35. Tong, J. & Anderson, J. L. Partitioning and diffusion of proteins and linear polymers in polyacrylamide gels. *Biophys. J.* **70,** 1505–13 (1996).
36. Shin, H. Western blot abreview for anti-alpha smooth muscle Actin antibody [1A4] (ab7817). *Abcam product website* http://www.abcam.com/alpha-smooth-muscle-Actin-ant (2015).
37. Uhlen, M. *et al.* A proposal for validation of antibodies. *Nat. Methods* **13,** 823–827 (2016).
38. Cohen, A. A. *et al.* Protein dynamics in individual human cells: Experiment and theory. *PLoS One* **4,** (2009).
39. Furusawa, C., Suzuki, T., Kashiwagi, A., Yomo, T. & Kaneko, K. Ubiquity of Log-normal Distributions in Intra-cellular Reaction Dynamic. *Biophysics (Oxf).* **1,** 25–31 (2005).
40. Esch, E. W., Bahinski, A. & Huh, D. Organs-on-chips at the frontiers of drug discovery. *Nat. Rev. Drug Discov.* **14,** 248–260 (2015).
41. Ghaemmaghami, A. M., Hancock, M. J., Harrington, H., Kaji, H. & Khademhosseini, A. Biomimetic tissues on a chip for drug.pdf. *Drug Discov Today* **17,** 173–181 (2013).
42. Peggs, K. An insufferable business: Ethics, nonhuman animals and biomedical experiments. *Animals* **5,** 624–642 (2015).
43. Mak, I. W., Evaniew, N. & Ghert, M. Lost in translation: animal models and clinical trials in cancer treatment. *Am. J. Transl. Res.* **6,** 114–8 (2014).

Chapter 5: Single-embryo and single-blastomere immunoblotting reports protein expression heterogeneity in early-stage preimplantation embryos

This work was conducted in collaboration with Dr. Andrew J. Modzelewski and Prof. Lin He

Understanding how a zygote develops from a single cell into a multicellular organism has benefitted from single-cell tools, including RNA sequencing (RNA-Seq) and immunofluorescence (IF). However, scrutinizing inter- and intra-embryonic phenotypic variation is hindered by two fundamental limitations: the loose correlation between transcription and translation and the cross-reactivity of immunoreagents. To address these challenges, in this chapter describe a high-specificity microfluidic immunoblot optimized to quantify protein expression from all stages of mouse preimplantation development. Despite limited availability of isoform-specific immunoreagents, the immunoblot resolves inter-embryonic heterogeneity of embryo-specific isoforms (i.e., DICER-1). We observed significantly higher DICER-1 isoform expression in oocytes when compared to two-cell embryos, and further find that protein expression levels follow the same trend as mRNA for both the full-length and truncated DICER-1 isoforms. At the morula stage, we assayed both whole and disaggregated embryos for loading controls (β-tubulin, GAPDH) and markers that regulate cell fate decisions (CDX-2, SOX-2). In disaggregated morula, we found that cell volume showed positive, linear correlation with expression of β-tubulin and SOX-2. In dissociated two-cell and four-cell blastomeres, we detect significant inter-blastomeric variation in GADD45a expression, corroborating suspected cellular heterogeneity even in the earliest multicellular stage of preimplantation embryos. As RNA-Seq and other transcript-centric approaches continue to further probe preimplantation development, the demand for companion protein-based techniques rises. The reported microfluidic immunoblot serves as an essential tool for understanding mammalian development by providing high-specificity and direct measurements of protein targets at single-embryo and single-blastomere resolution.

5.1 Introduction

The initiating events and specific proteins involved in the first cell fate commitment within pre-implantation blastomeres still remain open questions in developmental biology. Although functional studies and embryonic plasticity suggest that blastomeres remain equivalent until the compacted morula[1–3], growing evidence of inter-blastomeric differences in early-stage embryos point to heterogeneous configurations at even the earliest multicellular stages[4–14]. Measurement tools with single-embryo and single-blastomere resolution, including RNA-Seq, have greatly advanced our knowledge. Nevertheless, companion protein expression and state measurements within single embryos are required to test and validate the transcript-based predictions. Direct assessment of protein expression is required.

While immunofluorescence (IF) reports protein abundance and localization in embryos, IF is stymied by: (i) ubiquitous immunoreagent cross-reactivity that renders it unsuitable for detection of small protein variations or multiplexing beyond ~5 targets[15], (ii) proteoform 'blind spots' arising from limited isoform-specific immunoreagents that reduce the detectable repertoire of targets[16],

and (iii) confounding but required chemical fixation prior to IF measurement of endogenous intracellular proteins (i.e., epitope masking, cell morphology modifications, and perturbation of protein localization by diffusional gradients formed as fixation occurs)[17,18]. Flow cytometry and mass cytometry suffer from similar specificity and fixation concerns as in IF[19]. Mass spectrometry does not yet afford the sensitivity to analyze single-mouse embryos or blastomeres (tens of ng of protein in oocytes to < 1 ng of protein in blastomeres at the blastocyst stage)[19,20]. Although protein analysis tools have been recently introduced to measure protein targets in single cells[21–24], fundamental differences between cultured cell lines and mammalian embryos have hindered the study of early mammalian development. Key differences include cell size and composition, membrane structure and sample availability of ~20 embryos per mouse, equivalent to less than 1 μg of protein[25–28]. To complement the repertoire of existing measurements and resolve the intriguing questions surrounding mammalian development, such as when the first cell fate decisions are made, precision protein tools with higher selectivity are needed.

Here we report microfluidic immunoblotting for direct analysis of proteoforms across all stages of mouse preimplantation, from whole embryos to single blastomeres. With a dynamic detection range spanning two-orders of magnitude (oocytes at 10^{-15} mole to blastocyst blastomeres at 10^{-17} mole), we scrutinize widely used housekeeping proteins (β-tubulin, α-actinin, and GAPDH), embryo-specific isoforms (DICER-1) and regulators of cell fate decisions (SOX-2 and GADD45a). In whole morula, we assess expression of three targets (GAPDH, CDX-2 and SOX-2) having small 1-2 kDa differences in molecular mass. Applied to the study of intra-embryonic heterogeneity, we observe statistically significant differences in GADD45a expression at the four-cell and even two-cell stages. Microfluidic immunoblotting provides the resolution necessary to quantify both inter- and intra-embryonic heterogeneity.

5.2 Materials and Methods

Animals and Ethics Statement. As a matter of caution and compliance, all appropriate authorizations have been acquired from institutional and/or federal regulatory bodies prior to performing this protocol. All mouse use, including but not limited to housing, breeding, production, sample collection for genotyping, and euthanasia, is in accordance with the Animal Welfare Act, the AVMA Guidelines on Euthanasia and are in compliance with the ILAR *Guide for Care and Use of Laboratory Animals*, and the UC Berkeley Institutional Animal Care and Use Committee (IACUC) guidelines and policies. Our animal care and use protocol has been reviewed and approved by our IACUC for this project.

Chemical Reagents. Tetramethylethylenediamine (TEMED, T9281), ammonium persulfate (APS, A3678), β-mercaptoethanol (M3148), 30%T/2.7%C acrylamide/bis-acrylamide (37.5:1) (A3699), bovine serum albumin (BSA, A9418), Tyrode's solution (T1788), trypsin 10X (59427C), Accutase® (A6964) and 3-(trimethoxylsilyl)propyl methacrylate (440159) were purchased from Sigma-Aldrich. Triton X-100 (BP-151) and phosphate-buffered saline (PBS, pH 7.4, 10010023) were purchased from Thermo Fisher Scientific. Premixed 10X tris-glycine EP buffer (25 mM Tris, pH 8.3; 192 mM glycine; 0.1% SDS) was purchased from Bio-Rad. Tris buffered saline with Tween-20 (TBST) was prepared from 20X TBST (sc-24953, Santa Cruz Biotechnology, Dallas, TX). Deionized water (18.2 MΩ) was obtained using an Ultrapure water system from Millipore. Alexa555-labeled bovine serum albumin (A34786) was purchased from Invitrogen. N-[3-[(3-

Benzoylphenyl)formamido]propyl] methacrylamide (BPMAC) was custom synthesized by Pharm-Agra Laboratories (Brevard, NC). Gel Slick™ (50640) was purchased from Lonza.

Device Fabrication. Devices were fabricated using SU-8 wafers as previously reported. Microwell height and diameter were varied to accommodate different embryo stages (microwell diameters ranged from 20 μm for dissociated blastocyst blastomeres to 150 μm for oocytes, with height maintaining an aspect ratio of 1.3). Polyacrylamide precursor solution including acrylamide/bis-acrylamide (7-12%T) and 3 mM BPMAC was degassed with sonication for 9 min. 0.08% APS and 0.08% TEMED were added to precursor solution and solution was pipetted between the SU-8 wafer (rendered hydrophobic with Gel Slick™ solution) and a glass microscope slide functionalized with 3-(trimethoxylsilyl)propyl methacrylate (to ensure covalent grafting of PA gel to glass surface). After chemical polymerization (20 min), devices (glass with grafted PA gel layer) were lifted from wafer, rinsed with deionized water and stored dry until use.

Mouse Embryo Isolation and Culture. Three-to-five-week-old female C57BL/6J mice (000664, Jackson Laboratory) were superovulated by intraperitoneal (IP) injection of 5IU of Pregnant Mare Serum Gonadotropin (PMSG, Calbiochem, 367222, Millipore) and 46-48 hours later, IP injection of 5IU Human Chorion Gonadotropin (hCG, Calbiochem, 230734, Millipore). Superovulated females were housed at a 1:1 ratio with 3-8 month old C57BL/6J males to generate 1-cell zygotes at 0.5 days post coitum. Using forceps under a stereomicroscope (Nikon SMZ-U or equivalent), the ampulla of oviduct was nicked, releasing fertilized zygotes and oocytes associated with surrounding cumulus cells into 50 μL M2 + BSA media (M2 media (MR-015-D, Millipore) supplemented with 4 mg/mL bovine serum albumin (BSA, A3311, Sigma)). Using a handheld pipette set to 50 μL, zygotes were dissociated from cumulus cells after the cumulus oocyte complexes were transferred into a 200 μL droplet of Hyaluronidase/M2 solution (MR-051-F, Millipore), incubated for up to 2 min, and passed through five washes in the M2+BSA media to remove cumulus cells.

From this point on, embryos were manipulated using a mouth-controlled assembly consisting of a capillary pulled from glass capillary tubes (P0674, Sigma) over an open flame attached to a 15-inch aspirator tube (A5177, Sigma). Embryos were passed through five washes of M2+BSA to remove cumulus cells. Embryos were then transferred to KSOM + BSA media (KCl-enriched simplex optimization medium with amino acid supplement, ZEKS-050, Zenith Biotech) that was equilibrated to final culturing conditions at least 3–4 hr prior to incubation. Embryos were cultured in 20 μl droplets of KSOM + BSA overlaid with mineral oil (ES-005-C, Millipore) in 35 × 10 mm culture dishes (627160, CellStar Greiner Bio-One) in a water-jacketed CO_2 incubator (5% CO_2, 37 °C and 95% humidity).

Single-Embryo Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

All single embryo cDNA was prepared using a modified version of the Single Cell-to-Ct qRT-PCR kit (4458236, Life Technologies). Whole embryos were isolated at the desired developmental stage. Using a mouth pipette, embryos were then passed through three washes of PBS. With a hand-held pipette set to 1 μL, a single embryo was collected and transferred to one tube of an 8 well PCR strip. Presence of embryo was visually confirmed in each tube prior to cDNA synthesis. To account for the larger sample input, twice the protocol recommended volume of Lysis/DNAse

(20 μL) was added to each embryo and allowed to incubate at room temperature (RT) for 15 minutes. Then, 2 μL of Stop Solution was added and incubated for 2 min. At this point, half of the reaction was stored in -80°C conditions as a technical replicate and the remaining sample (11 μL) continued through the original Single Cell-to-Ct protocol. All qRT-PCR reactions were performed using SSO Universal SYBR Green SuperMix, as per manufacturer instructions (1725275, BioRad). Primer sequences used were Rfx1 (5'AGT GAG GCT CCA CCA CTG GCC G, 5'TGG GCA GCC GCT TCT C), Dicer-1 (5'GGA TGC GAT GTG CTA TCT GGA, 5'GCA CTG CTC CGT GTG CAA) and Dicer° (5'CTC TTT CCT TTG AAT GTA CAG CTA C, 5'CAG TAA GCA GCG CCC CTC). All qRt-PCR analyses were performed on the StepOnePlus Real Time PCR system (437660, Thermo).

Single-Embryo and Single-Blastomere Microfluidic Immunoblotting Once the desired developmental stage was reached, embryos were transferred to a ~50 μL droplet of acid Tyrode's Solution (T1788, Sigma) and incubated at 37°C for up to 2 min to remove the *zona pellucida*. If dissociation into blastomeres was required, embryos were first incubated with a 1:1 solution of Accutase® and 10X trypsin (15090046, Thermo) at 37°C (time varied with development stage, ranging from 5 min for two-cell embryos to up to 5 hr for blastocysts). Embryos were then mechanically disrupted by passing embryo through a capillary repeatedly until dissociation. Single embryos or blastomeres were washed with PBS and deposited into microwells of the PA gel. Microwells were imaged by brightfield microscopy to collect data on size, morphology and ensure occupancies of one embryo or blastomere per microwell. Lysis conditions, including buffer composition, temperature and treatment time, were optimized for each developmental stage (Table 1). Electrophoresis was performed at 40 V/cm for varying times (from 20 to 60 s, depending on developmental stage and protein targets). Immobilization of proteins by photocapture was carried out by illumination with UV light source (100% power, 45 s, Lightningcure LC5, Hamamatsu). Gels were washed in 1X TBST for at least 1 hr before probing with antibodies. Primary antibodies were incubated at 1:10 dilution (40 μL/gel, in 2% BSA in 1X TBST), while fluorophore-conjugated secondary antibodies were incubated at 1:20 dilution (40 μL/gel, in 2% BSA in TBST). In order to strip bound antibodies and reprobe for new targets, gels were treated with 2% SDS, 0.8% β-mercaptoethanol and 62.5 mM Tris base at 55°C for three hours, washed in TBST (1 hr) twice and then reprobed.

Table 1: Conditions for assaying embryos and blastomeres of murine preimplantation stages. All assays were run at an E field strength of 40 V/cm.

Sample Type (zona-free)	Lysis Buffer	Lysis T & time	EP time	PA Gel %T
Oocyte, two-cell & four-cell (whole or disaggregated)	*	55°C, 60-75 s	75 s	7, 10
Whole morula	*	55°C, 60 s	50 s	7
Disaggregated morula blastomeres	*	55°C, 60 s	45 s	7
Disaggregated blastocyst blastomeres	**	55°C, 35 s	20 s	10

* 25 mM Tris-glycine buffer at pH 8.3 with 1% SDS, 0.5% sodium deoxycholate, and 1% Triton X-100. ** 25 mM Tris-glycine buffer at pH 8.3 with 0.5% SDS, 0.25% sodium deoxycholate, and 0.1% Triton X-100.

Antibodies. Antibodies employed for analysis of embryos include: rabbit anti-β-Tubulin (Abcam, ab6046, polyclonal, LOT: GR31927-5), mouse anti-Dicer-1 (Santa Cruz, sc-136981, A-2, LOT: I1817), mouse anti-CDX-2 (Abcam, ab157524, CDX2-88, LOT: GR300552-6), rabbit anti-SOX-2 (EMD Millipore, AB5603, polyclonal, LOT: NG1863962), goat anti-GAPDH (Sigma, SAB2500450, polyclonal, LOT: 6377C2), rabbit anti-GADD45a (Invitrogen, MA5-17014, D.81.E, LOT: R12274975). Donkey secondary antibodies AlexaFluor 647-conjugated anti-mouse (A31571), AlexaFluor 594-conjugated anti-mouse (A21203) and AlexaFluor 488-conjugated anti-rabbit (A21206) were purchased from ThermoFisher Scientific CA, USA.

Image Processing, Signal Quantification and Statistical Analysis. The datasets generated during and analyzed during the current study are available from the corresponding author on request. Statistical tests were performed using GraphPad Prism 7.0b. Quantification of fluorescence signal of protein immunoblots employed in-house scripts written in MATLAB (R2017a, Mathworks)[33]. Briefly, Gaussian curves were fit to protein band fluorescence intensity profiles using MATLAB's Curve Fitting Toolbox. Gaussian fit parameters of protein peak location and standard deviation (σ) were used to compute area-under-the-curve (AUC) by integrating the fluorescence intensity profile for the peak width defined as 4σ. Fiji (ImageJ) was used to false-color fluorescence micrographs and overlay channels to create composite images. ImageJ was used to compute cell volume[50]. Briefly, cell boundaries were traced using the freehand selection tool. For area traces with circularity of > 0.9, we assumed spherical morphology, computed cell diameter from traced area and cell volume was calculated from the computed cell diameter.

5.3 Results and Discussion

5.3.1 Workflow for microfluidic immunoblotting of single embryos and blastomeres

We first sought to measure protein expression in cells ranging from single oocytes (~80 μm in diameter) to single blastomeres from disaggregated blastocysts (<20 μm in diameter at 3.5-4.0 days post coitus, dpc) (Figure 1, a). Sample preparation of harvested murine embryos includes (i) removal of the *zona pellucida* by incubating with acidic Tyrode's solution and, if studying disaggregated blastomeres, (ii) dissociation of embryos into individual blastomeres by incubation with trypsin and Accutase®. The microfluidic immunoblot, comprised of a 100-150 μm-thick polyacrylamide (PA) gel layered on a glass microscope slide, is stippled with an array of microwells (100-160 μm in diameter), where each microwell is designed to isolate an individual embryo or blastomere, depending on the assay. Using a standard mouth-controlled capillary tube assembly[29], single embryos or blastomeres are seated into individual microwells (Figure 1, b). Once isolated in each microwell, cell samples are chemically lysed (60-70 s). An electric field (E = 40 V/cm; 35-75 s) is applied to drive protein polyacrylamide gel electrophoresis (PAGE) in a 3 mm-long separation lane abutting each lysate-containing microwell. Protein blotting (immobilization) to the PA gel is achieved by UV-mediated activation of benzophenone methacrylamide moieties crosslinked into the gel matrix. Size-resolved immobilized protein bands are probed using primary antibodies and fluorophore-conjugated secondary antibodies, resulting in single-embryo or single-blastomere immunoblots. By probing for the protein loading control β-tubulin in lysate from single oocytes down to individual blastomeres from a disaggregated

blastocyst (Figure 1, c), we determined a dynamic detection range spanning femtomoles (10^{-15}) to tens of attomoles (10^{-17}), assuming a starting protein target concentration in the μM range[30].

We next scrutinized single morula (3.0 dpc) for SOX-2 and CDX-2, two transcription factors that regulate pluripotency and differentiation[31,32], and the loading control GAPDH. SOX-2, CDX-2 and GAPDH have molecular masses of 37, 38 and 39 kDa, respectively (Figure 1, d). By employing a combination of (i) primary antibodies raised in different animals (goat-anti-GAPDH, rabbit-anti-SOX-2, and mouse-anti-CDX-2), (ii) secondary antibodies conjugated to different fluorophores (donkey-anti-goat, rabbit and mouse conjugated to AlexaFluor 555, 488 and 594, respectively), and (iii) reprobing after gel stripping using a reducing stripping buffer, the microfluidic immunoblot resolved all three targets from intact morula. The observed log-linear relationship between molecular mass and migration distance (Figure 1, d) enables distinguishing target protein bands from non-specific antibody signal and demonstrates that single-morula PAGE resolves protein targets with molecular mass as close as 1-2 kDa.

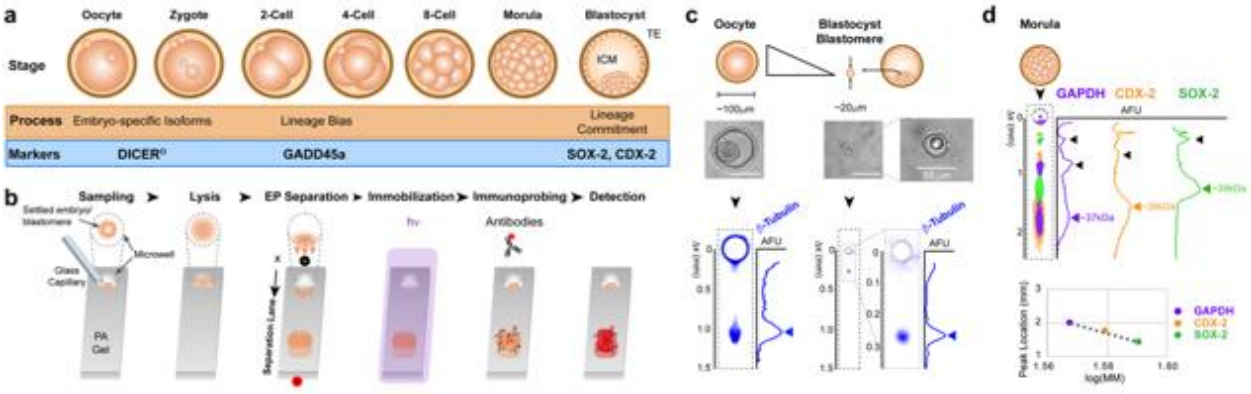


Figure 1: Microfluidic immunoblotting of single embryo and single blastomeres measures protein markers in all stages of preimplantation development. (a) During preimplantation development, a fertilized oocyte develops into a blastocyst. Protein markers investigated in this study are related to embryo-specific isoform expression, early-stage lineage biases and cell fate specification. (b) The microfluidic immunoblotting workflow begins with sampling a single embryo or single blastomere into a microwell patterned on a polyacrylamide (PA) gel. Samples are lysed and electrophoresed into the PA, achieving separation of proteins by molecular mass. Proteins are immobilized to gel matrix by UV-activated benzophenone chemistry and probed with fluorophore-conjugated antibodies. (c) Loading control β-tubulin was measured from single oocytes down to single blastocyst blastomeres. Brightfield micrographs of a settled oocyte and blastomere are shown above false-colored micrographs of resulting β-tubulin immunoblots and corresponding fluorescence intensity profiles. Arrows mark position of protein bands and scale bars are 100 μm, unless specified. (d) Single morula assayed for multiple targets that differ by 1-2 kDa (GAPDH, CDX-2 and SOX-2) show a strong log-linear relationship between migration distance and molecular mass (R^2 = 0.9842).

5.3.2 Validating detection of inter- and intra-embryonic biological variation

Given our ability to immunoblot dissociated blastomeres, we next examined (i) if embryo disaggregation artificially alters the protein abundance of the whole embryo and (ii) if we can reconstruct the expression profile of the whole embryo, even when constituent blastomeres are assayed individually.

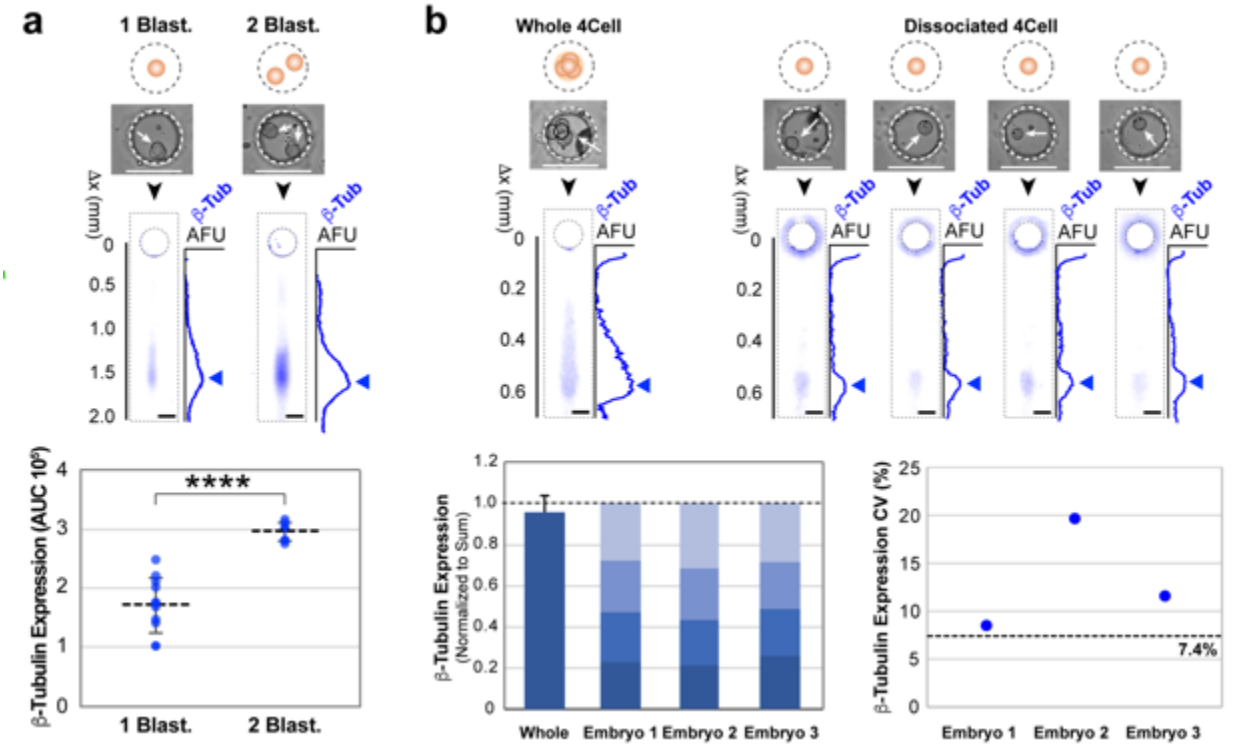


Figure 2: Microfluidic immunoblotting detects intra-embryonic biological variation in β-tubulin expression. (a) β-tubulin titration experiment. One or two blastomeres of dissociated four-cell embryos are sampled into microwells and assayed for β-tubulin. Brightfield images show blastomeres settled into microwells prior to lysis. Under these, false-color fluorescence micrographs and corresponding β-tubulin intensity profiles of resulting immunoblots. Arrows mark the position of protein bands. Dot plot of β-tubulin signal for immunoblots of one and two blastomeres demonstrate an increase in detection of β-tubulin for two blastomeres over one blastomere (horizontal bars indicate mean ± S.D., Mann Whitney U test, p value < 0.001, N = 11 and 7 for microwells with one and two blastomeres, respectively). (b) Reconstruction of whole embryo from disaggregated blastomeres. Bright field micrographs of whole and disaggregated four-cell embryos settled into microwells (top), with corresponding false-colored fluorescence micrographs of β-tubulin immunoblots. Intensity profiles are shown to the right of immunoblots, with blue arrow marking the position of the protein bands. Stacked bar graphs show individual blastomere contributions to total β-tubulin expression of four-cell embryos. Whole embryos assayed alongside dissociated blastomeres show similar levels of total β-tubulin expression (error bar indicates S.D., Wilcoxon matched-pairs signed rank test, p value = 0.5, *N* = 3 independent experiments), indicating sum of individually assayed blastomeres is equivalent to a whole embryo. Dot plot of β-tubulin expression coefficient of variation (CV %) for blastomeres of three disaggregated four-cell embryos. All CV values are above the technical $CV_{threshold}$ of 7.4% (Figure 3).

We first inquired if loading a pre-determined increase of protein in the microfluidic immunoblot yields a concomitant increase in protein signal. We thus performed titrations where we controlled loaded protein by using individual blastomeres from dissociated four-cell embryos (2.0 dpc) as discrete and easily manipulable loads of protein. We loaded either one or two blastomeres into microwells and assayed the microwell lysate for β-tubulin (Figure 2, a). We observed a proportional increase in β-tubulin expression (area-under-the-curve signal or AUC) from microwells loaded with two blastomeres as compared to microwells loaded with one blastomere

(Mann Whitney U Test, p value = 6.28×10^{-5}, with N = 7 and 11 microwells, respectively). This observation corroborates the supposition that two blastomeres contain two-fold more protein than a single blastomere (Figure 2, a).

Finally, we assessed whether the source of the observed inter-blastomeric variation in β-tubulin AUC was attributable to biological variation or confounding technical variation. First, we established a technical variation threshold by quantifying immunoblots of microwells uniformly loaded with purified protein. Briefly, we partitioned a solution of purified bovine serum albumin (BSA, 1 μM in PBS) into the microwells by incubating PA gels in BSA solution for 30 min. We then performed the immunoblotting assay and quantified BSA protein band AUC. We calculated the coefficient of variation in BSA AUC (CV_{AUC} % = AUC standard deviation (S.D.) / mean AUC x 100) and computed a technical variation threshold defined as > 3 x S.D. of the mean CV_{AUC} [33] ($CV_{threshold}$ = mean CV_{AUC} + 3 S.D. = 7.4%, where mean CV_{AUC} = 4.69% and S.D. = 0.92%, Fig. S1). For all dissociated four-cell embryos studied, the inter-blastomeric β-tubulin expression CV exceeded the technical variation threshold (CVs = 8.3%, 19.6% and 11.3% for embryos, Figure 2, b). Consequently, we attribute the inter-blastomeric variation to biological variation and not to technical variation.

Figure 3: Determination of the technical variation threshold of the microfluidic immunoblot. (a) Schematic of purified protein immunoblotting assay. The polyacrylamide (PA) gel of the microfluidic immunoblotting device is incubated with a solution of fluorescently-labeled bovine serum albumin (BSA) for 30 min for BSA to partition into microwells. Assay is then run as described in main text. (b) False-colored fluorescence micrograph of resulting BSA immunoblots and corresponding fluorescence intensity profiles used to perform area-under-the-curve (AUC) quantification. The coefficient of variation (CV%) was calculated as S.D./mean x 100 for N = 9 replicates. The technical variation threshold was computed as the mean CV (4.7 %) plus three standard deviations for a 99% confidence interval (S.D. = 0.9%) for a CV threshold of 7.4%.

The technical CV threshold calculated here is lower than the threshold computed in Chapter 2, section 2.3.2, 32.4%. We believe that several sources of variability account for this difference in computed CV thresholds. First, the protein loads used in Chapter 2 were single GFP-MCF-7 cells, compared with a solution of purified protein partitioned into the microwells. The variation in protein loading between microwells is thus less variable in the case of the purified protein. Second, the fluorescence mechanism of the two loaded proteins is not equivalent- while whole-cell GFP before lysis depends on the folding of GFP, the loaded BSA is conjugated to a fluorophore. Because accurate folding is required for fluorescence, any GFP that is not correctly folded would not be accounted for in the initial GFP measurement. Third, the mechanism for protein loss over the top of the microwell is different for these two assays. For the GFP-MCF7 cells, cells must first lyse (~15 s) before released proteins can diffuse away. Thus, proteins are protected from the

convective stream lines during the first ~10 s of lysis, when the pouring of lysis buffer causes convective flow over the top of the microwell. In the case of purified protein, however, proteins can diffuse out of the microwell as soon as the lysis buffer is introduced, suggesting that the purified protein assay is more sensitive to variable losses due to non-uniform convective flow during the lysis step.

5.3.3 Investigating β-tubulin for normalization by cell volume in morula blastomeres

RNA-Seq studies suggest that cells regulate transcription to maintain mRNA concentration in response to chaftablenges in cell volume[34,35]. Thus, cellular mRNA concentration is a more accurate indicator of cell phenotype than is cellular mRNA abundance[34,35]. Normalizing by a loading control that is strongly correlated with cell size is therefore crucial for elucidating phenotypic differences between cells. However, RNA-based studies show that commonly employed housekeeping genes (e.g., GAPDH and β-tubulin) are not stably and homogeneously expressed across different samples, experimental conditions or treatments[36]. The issue is further exaggerated as the sample size diminishes, reducing the averaging effect of a larger pooled samples. Whether this variability, and thus unreliability, of loading controls for single-cell studies prevails at the protein level remains to be explored.

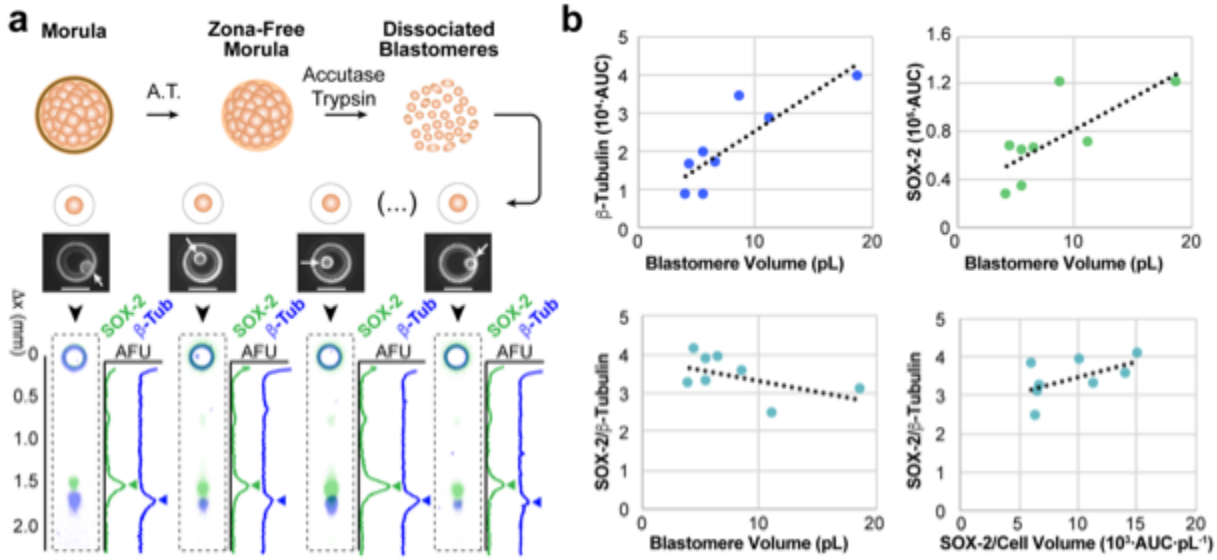


Figure 4: Single-blastomere immunoblotting identifies correlations between cell volume and marker expression in dissociated morula blastomeres. (a) Immunoblotting dissociated morula blastomeres for β-tubulin and SOX-2. Schematic (top) for dissociation of whole morula into individual blastomeres, which are seated into microwells of an immunoblotting device as shown in bright field images. False-colored fluorescence micrographs show β-tubulin and SOX-2 protein bands, with intensity profiles shown adjacent to micrographs. Arrows mark the position of protein bands. (b) Bivariate plot of blastomere cell volume and β-tubulin expression shows significant positive linear correlation (Pearson correlation, N = 8, ρ = 0.582, p value = 0.0064). While cell volume also shows significant positive linear correlation with SOX-2 expression, the same is not true when normalizing SOX-2 expression by β-tubulin expression (Pearson correlation, N = 8, ρ = 0.7381 and -0.5232 with p values = 0.0366 and 0.1835, respectively). Bivariate plot

of SOX-2 expression normalized by β-tubulin and by cell volume shows a positive, but non-significant, association (Pearson correlation, N = 8, ρ = 0.5522, p value = 0.1559).

Hence, we next tested whether the widely used loading control β-tubulin is an accurate indicator of cell volume in preimplantation embryos. We assayed dissociated morula blastomeres for β-tubulin and compared β-tubulin expression (AUC) to cell volume (computed from brightfield images of cells seated in microwells) (Figure 4, a). We noted a significant positive, linear correlation between cell volume and β-tubulin expression (Pearson correlation, ρ = 0.8582, N = 8, p value = 0.0064) (Figure 4, b). To determine if this correlation indicates that β-tubulin can be used as a proxy for cell volume, we examined whether normalizing SOX-2 expression by β-tubulin expression is equivalent to normalizing by computed cell volume. We first observed that while SOX-2 expression showed significant positive correlation with cell volume (Pearson correlation, ρ = 0.7660, N = 8, p value = 0.0267) (Figure 4, b), SOX-2 expression normalized by β-tubulin was not correlated with cell volume (negative association, Pearson correlation ρ = -0.5866, N = 7, p value = 0.1668) (Figure 4, b). Furthermore, SOX-2 normalized by cell volume was not significantly correlated with SOX-2 normalized by β-tubulin (Pearson correlation, ρ = 0.5522, p value = 0.1558) (Figure 4, b). Hence, careful validation of loading controls as accurate predictors of cell volume is indispensable, even if expression of loading controls is well correlated with cell volume. Indexing endpoint immunoblot results with micrographs of the originating and intact cell sample allows us to normalize target protein expression by cell volume – a normalization that is impossible in endpoint immunoblots of lysate from pooled cell populations.

5.3.4 Microfluidic immunoblotting detects truncated DICER-1 isoform expression in oocytes and two-cell embryos

Alternative splicing is frequent during early embryonic development in mouse and human[37–39]. However, efforts to investigate whether the corresponding alternate protein isoforms are ultimately and stably generated require pooling tens of thousands of collected embryos from each stage, losing intra-blastomeric information in the process[20]. Thus, single embryo and blastomere approaches capable of resolving proteoforms resulting from alternative splicing are needed.

To this end, we examined one of the earliest known examples of a protein isoform that exists in mouse development. DICER-1 is essential for small RNA-mediated gene expression regulation. By processing small RNAs into their mature form, DICER-1 generates the sequence-specific guides required by effector complexes to target cognate mRNAs and repress their translation [40]. Bulk analyses of mouse oocytes found high expression of an N-terminally truncated isoform, denoted DICER^{O} (Figure 5, a). DICER^{O} demonstrates higher catalytic activity than its full-length form (DICER-1) and is believed to drive the high activity of endogenous small interfering RNAs (endo-siRNAs) in mouse oocytes, but not in somatic cells[40]. The Dicer^{O} transcript persists until the fertilized zygote stage, but it remains unclear whether the DICER^{O} protein isoform is exclusive to oogenesis or is maternally inherited by the preimplantation embryo.

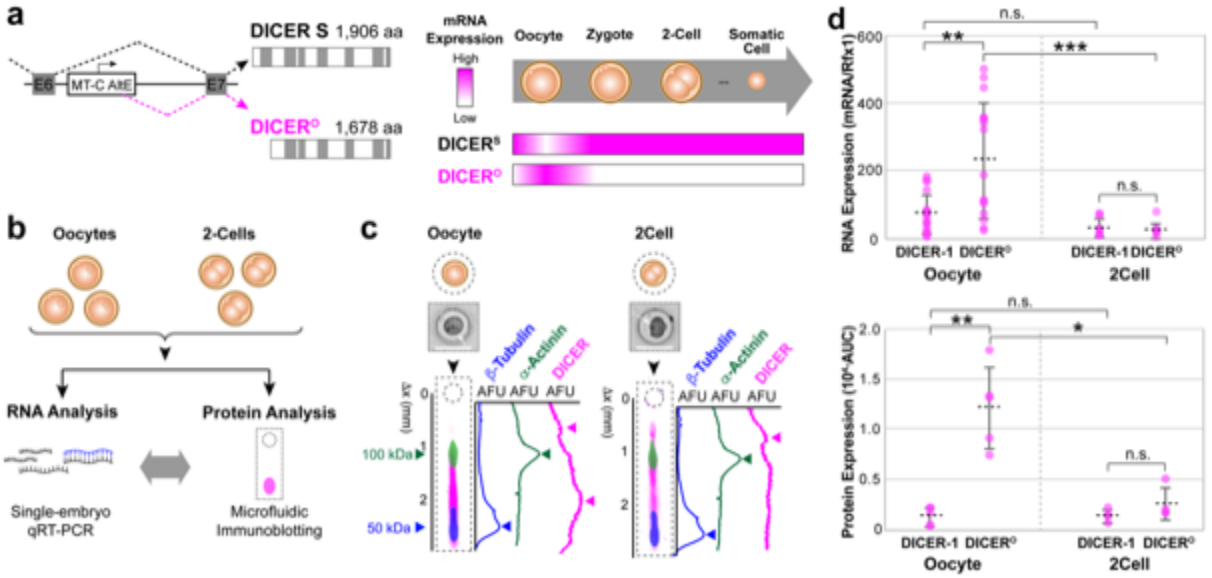


Figure 5: Higher DICER-1 isoform expression in oocytes than in two-cell embryos correlates with mRNA levels. (a) DICER^{O}, a truncated isoform of DICER-1, appears only at the oocyte stage and is a product of alternative promoter usage. (b) Schematic of oocytes and two-cell embryos analyzed either by microfluidic immunoblotting or by companion qRT-PCR analysis. (c) Bright field micrographs of a settled oocyte and two-cell embryo. Corresponding overlaid false-colored fluorescence micrographs and intensity profiles show protein bands for loading controls (α-actinin and β-tubulin) and DICER-1, where oocyte immunoblot demonstrates presence of a full-length DICER-1 (top arrow) and a lower molecular weight isoform (bottom arrow). Scale bars are 100 μm. (d) Dot plots of DICER isoform mRNA levels normalized by endogenous control Rfx1 (top) and protein expression (bottom) for single oocytes and single two-cell embryos. Expression of the truncated isoform is higher than the full-length DICER-1 for both mRNA and protein in oocytes (mRNA/Rfx1$_{DICER\text{-}1}$ vs. mRNA/Rfx1$_{DICER^{o}}$: Mann-Whitney U test, p value = 0.0052 for N = 18; for AUC$_{DICER\text{-}1}$ vs. AUC$_{DICER^{o}}$: Mann-Whitney U test, p value = 0.0079 for N = 5), but not in two-cell embryos (Mann-Whitney U tests for mRNA/Rfx1$_{DICER\text{-}1}$ vs. mRNA/Rfx1$_{DICER^{o}}$: p value = 0.9551 for N = 7 for DICER-1 and N = 8 for DICER^{O}; for AUC$_{DICER\text{-}1}$ vs. AUC$_{DICER^{o}}$: p value = 0.20 for N = 4). Oocytes show higher mRNA and protein expression than two-cells for the truncated isoform (Mann Whitney U tests for mRNA/Rfx1$_{DICER^{O}}$: p value = 0.0004 for N = 18 oocytes and 8 two-cells; AUC$_{DICER^{O}}$: p value = 0.0159 for N = 5 oocytes and 4 two-cell embryos), but not the full-length DICER-1 (Mann Whitney U tests for mRNA/Rfx1$_{DICER\text{-}1}$:, p value = 0.084 for N = 18 oocytes and 7 two-cell embryos; AUC$_{DICER\text{-}1}$: p value = 0.9048 for N = 5 oocytes, 4 two-cell embryos) (horizontal bars indicate mean ± S.D.).

To explore whether DICER^{O} is specific to the oocyte stage, we assayed oocytes and two-cell embryos for isoforms of DICER-1. We collected oocytes and two-cell embryos and divided each sample for analysis of either protein by microfluidic immunoblotting or mRNA analysis by single-embryo quantitative reverse transcription polymerase chain reaction (qRT-PCR) (Figure 5, b). Despite of a lack of an isoform-specific antibody, the electrophoretic separation of proteins resolved multiple DICER-1 isoforms by molecular mass. We observed that both oocytes and two-cell embryos expressed both the full-length and the truncated DICER-1 (Figure 5, c). For oocytes, we observed significantly higher expression of the truncated isoform over the full-length DICER-1 for both mRNA (normalized by endogenous control Rfx1) and protein (AUC) (for mRNA/Rfx1$_{DICER\text{-}1}$ vs. mRNA/Rfx1$_{DICER^{o}}$: Mann-Whitney U test, p value = 0.0052 for N = 18; for AUC$_{DICER\text{-}1}$ vs. AUC$_{DICER^{o}}$: Mann-Whitney U test, p value = 0.0079 for N = 5) (Figure 5, d). On the other hand, we found no significant difference in expression of truncated and full-length

isoforms of DICER-1 in two-cell embryos (for mRNA/Rfx1$_{DICER-1}$ vs. mRNA/Rfx1$_{DICER^o}$: Mann-Whitney U test, p value = 0.9551 for N = 7 for DICER-1 and N = 8 for DICER^O; for AUC$_{DICER-1}$ vs. AUC$_{DICER^o}$: Mann-Whitney U test, p value = 0.20 for N = 4) (Figure 5, d). When comparing expression levels between embryonic stages, we observed that for both mRNA and protein the expression of full-length DICER-1 was not significantly different between oocytes and two-cell embryos (Mann Whitney U tests, mRNA/Rfx1$_{DICER-1}$: p value = 0.084 for N = 18 oocytes and 7 two-cell embryos; AUC$_{DICER-1}$: p value = 0.9048 for N= 5 oocytes and 4 two-cell embryos) (Figure 5, d). For the truncated isoform, however, we observed a significant decrease in both mRNA levels and protein levels from the oocyte to the two-cell stage (Mann Whitney U tests, mRNA/Rfx1$_{DICER^O}$: p value = 0.0004 for N = 18 oocytes and 8 two-cells; AUC$_{DICER^O}$: p value = 0.0159 for N = 5 oocytes and 4 two-cell embryos) (Figure 5, d). Hence, protein PAGE from single-embryo lysates grants the selectivity required for measuring protein isoforms, even when specific antibodies are the only reagent available.

5.3.5 Scrutiniting single blastomeres for GADD45a expression heterogeneity in two- and four-cell embryos

We finally sought to inspect early-stage embryos for lineage biases by measuring protein expression from disaggregated two-cell and four-cell embryos. The exact stage and circumstances by which blastomeres acquire certain fates remains unknown. On the one hand, it is thought that embryonic plasticity supports blastomere symmetry up to the 8-cell embryo, where embryos can compensate for the loss of one blastomere as early as the two-cell stage[41]. On the other hand, studies point to early asymmetry, where sister blastomeres show consistent bimodal expression of genes related to differentiation, suggesting that the involved factors may not be inherited equally by all blastomeres[6].

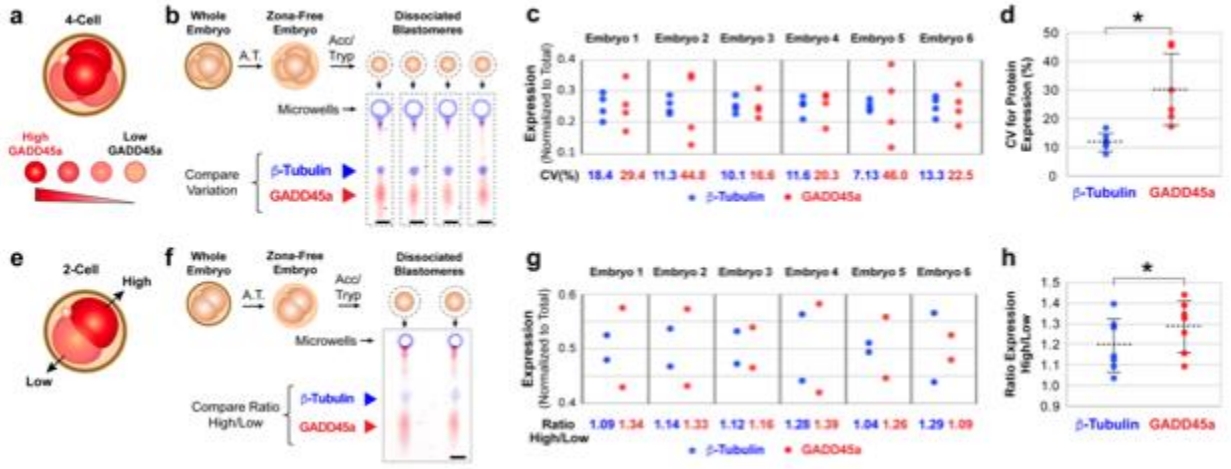


Figure 6 Microfluidic immunoblotting measures intra-embryonic heterogeneity in GADD45a expression in four-cell and two-cell embryos. (a) Four-cell embryos have been suggested to show early lineage bias by heterogeneous expression of GADD45a. (b) Design for testing this hypothesis includes removing *zona pellucida* from four-cell embryos and dissociation into individual blastomeres for subsequent immunoblotting for GADD45a and loading control β-tubulin. False-colored fluorescence micrographs show bands for both protein targets for one dissociated four-cell embryo. (c) Dot plot of expression of β-tubulin (blue) and GADD45a (red) by individual blastomeres, normalized to total expression, with

corresponding inter-blastomeric coefficient of variation (CV) for β-tubulin and GADD45a expression. (d) Dot plot of CVs shows that inter-blastomeric variation in GADD45a is significantly different from inter-blastomeric variation in β-tubulin (Wilcoxon signed-rank test, p value = 0.0312, $N = 6$). We investigated whether this heterogeneity can be traced back to the two-cell stage (e). (f) Schematic of two-cell embryo sample preparation by removal of *zona pellucida* and dissociation into individual blastomeres. False-colored fluorescence micrographs show β-tubulin and GADD45a immunoblots for two-cell sister blastomeres. (g) Dot plots of β-tubulin and GADD45a expression by sister blastomeres, normalized to sum of expression of sister blastomeres. (h) Dot plot of ratio between the AUC of high-expressing and low-expressing blastomeres for β-tubulin and GADD45a (paired t test p value = 0.0251, N = 11 dissociated two-cell embryos) (horizontal bars in (d) and (h) indicate mean ± S.D.).

As such, to quantitatively examine intra-embryonic heterogeneity in cell fate related markers, we assayed early-stage blastomeres for GADD45a, a protein involved in DNA damage repair that has been reported to show bimodal transcription at the two-cell and four-cell stages[4] (Figure 6, a). We compared the intra-embryonic heterogeneity of GADD45a expression to that of β-tubulin, to control for stochasticity of protein partitioning at cell division [42] (Figure 6, b). We observed that the intra-embryonic variation in GADD45a expression is significantly higher than the variation in β-tubulin expression ($CV_{GADD45a} = 31.5 \pm 13.5\%$; $CV_{\beta\text{-tub}} = 13.4 \pm 5.4\%$; mean ± S.D., Wilcoxon signed-rank test, p value = 0.0312, $N = 6$, where all CVs > $CV_{threshold}$ of 7.4%, indicating that the measured heterogeneity is attributable to biological, not technical, variation) (Figure 6, c & d). These findings suggest that blastomeres of four-cell embryos show heterogeneous expression of GADD45a, in agreement with the mRNA and IF-based findings of Biase *et al.*[4].

We next investigated if heterogeneity in GADD45a expression arises in the earlier two-cell embryo. Unlike in the four-cell stage, bimodality in GADD45a protein expression at the two-cell stage remains unexplored. We thus assayed dissociated two-cell embryos to understand the intra-embryonic distribution of GADD45a. To test whether one blastomere consistently showed higher GADD45a AUC than the other, we immunoblotted dissociated two-cell blastomeres for GADD45a and β-tubulin (Figure 6, e and f) and computed the ratio of expression between the blastomere with high expression and low expression for both markers (Figure 6, g). We found that the ratio of high-to-low GADD45a expression (1.26 ± 0.109) is significantly larger than the high-to-low ratio for β-tubulin (1.16 ± 0.11; mean ± S.D., paired t test p value = 0.0251, N = 11 two-cell embryos, distributions for GADD45a and β-tubulin passed the D'Agostino & Pearson normality test with p values = 0.6842 and 0.2497, respectively) (Figure 6, h). Hence, we posit that two-cell blastomeres display significant heterogeneity in GADD45a protein expression, providing protein-based evidence for heterogeneity in GADD45a expression even at the earliest multicellular preimplantation stage. Functional studies will allow us to determine whether differences in GADD5a protein expression levels are indicative of differential cellular phenotypes.

5.4 Conclusions

Open questions surround the timing and mechanism by which the first cell fate decisions are made during mammalian pre-implantation development. Do developing blastomeres remain homogeneous and functionally equivalent until the compacted morula stage [1,2]? Or, alternately, do these developing blastomeres exhibit symmetry-breaking heterogeneous configurations, perhaps as early as the two-cell stage [7,9–11]? In mice, zygotic genome activation (ZGA) occurs shortly after

fertilization and is not fully realized until the two-cell stage, at which point nascent mRNA populate the embryonic transcriptome [43]. While recent single-cell RNA-Seq experiments reveal sister blastomere transcriptome heterogeneity as early as the two-cell embryo in both human and mice[4,11], functional studies suggest these differences may not matter until the four-cell stage[3,12].

Here, microfluidic design provides an avenue for a cellular-resolution form of protein immunoblotting applicable to mammalian development as early as the oocyte stage of a murine model. To advance linking mRNA levels to protein expression when tackling questions on cell fate determination, we scrutinized two-cell and four-cell embryos for signs of heterogeneity in protein expression of GADD45a, a gene shown to be bimodally transcribed in early embryonic stages. The microfluidic immunoblot detects higher heterogeneity in GADD45a expression than loading control β-tubulin between the blastomeres in early embryos, providing the first protein-based validation of recent single cell RNA-Seq predictions. Companion functional competency measurements guided by the ever-increasing resolution of single-cell transcript approaches will help in determining the proteins and isoforms involved in key cell fate decisions events.

Directly scrutinizing and independently validating single-cell transcriptional data describing early embryonic development requires direct measurement of proteins with single-cell resolution. The asynchrony observed between mRNA and protein expression may simply reflect the uncoupled relationship between transcription and translation in the early embryo (~15 hr), synchronizing more tightly in later cleavage stages [44]. Transcript abundance is not an accurate determinant of protein abundance[45,46], and whether the required activating post-translational modifications (PTMs) are present has been difficult to ascertain[47]. Beyond PTMs and protein-mediated signaling, isoforms of proteins such as DICER1 are involved in regulation of small RNA-mediated gene expression. Microfluidic immunoblotting resolved DICER-1 isoforms in single oocytes and two-cell embryos, with the truncated DICER-1 protein isoform as the dominant isoform in the oocyte stage. Significantly lower abundance of the truncated DICER-1 mRNA and protein in two-cell embryos compared to oocytes suggests that these may be inherited from the oocyte stage. These findings support previous studies suggesting that the truncated DICER-1 isoform, which shows higher activity than the full-length DICER-1, is oocyte-specific. Previously unattainable single-embryo lysate protein separations enable resolving and quantifying protein isoforms, even when isoform-specific antibodies are not available.

As detailed here, the ~20 embryos harvested from a single mouse donor are sufficient not just for one immunoblot, but for multiple single-embryo and single-blastomere immunoblots. The precision in sample handling and in enhanced sensitivity notably reduces the conventional PAGE sample requirements of several hundreds or thousands of embryos[20,40]. The implications are multi-fold. First, as single-embryo immunoblots inherently and dramatically lower sample requirements, the burden of animal sacrifice is likewise reduced. Second, statistical interpretation of single-embryo and single-blastomere immunoblot results is feasible, revealing intra-embryonic heterogeneity, as well as significant differences between embryos of the same fertilization event and between donors.

Lastly, the possibility arises of companion immunoblot and mRNA assays on embryos of the same donor, thus enhancing the biological accuracy of correlations between mRNA levels and protein expression at different stages of the preimplantation embryo. Such insight into the expression dynamics would clarify how modulation in transcription dictates cellular phenotype [48]. Moreover,

with the advent of new gene editing technologies, (e.g., CRISPR, genomic screening methods including targeted, exome or whole genome sequencing) screening for off-target activity has become critical [49]. Protein assays that can complement genomic screening, such as the one described in this study, will be crucial for screening embryos for protein-level effects of both on-target and off-target mutations, even when the latter occur in non-coding regions.

5.5 References

1. Motosugi, N., Bauer, T., Polanski, Z., Solter, D. & Hiiragi, T. Polarity of the mouse embryo is established at blastocyst and is not prepatterned. *Genes Dev.* **19,** 1081–1092 (2005).
2. Alarcon, V. B. & Marikawa, Y. Unbiased Contribution of the First Two Blastomeres to Mouse Blastocyst Development. *Mol. Reprod. Dev.* **72,** 354–361 (2005).
3. Fujimori, T., Kurotaki, Y., Miyazaki, J. & Nabeshima, Y. Analysis of cell lineage in two- and four-cell mouse embryos. *Development* **130,** 5113–5122 (2003).
4. Biase, F. H., Cao, X. & Zhong, S. Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing. *Genome Res.* **24,** 1787–1796 (2014).
5. Xue, Z. *et al.* Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* **500,** 593–597 (2013).
6. Casser, E. *et al.* Totipotency segregates between the sister blastomeres of two-cell stage mouse embryos. *Sci. Rep.* **7,** 1–15 (2017).
7. Torres-Padilla, M. E., Parfitt, D. E., Kouzarides, T. & Zernicka-Goetz, M. Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* **445,** 214–218 (2007).
8. Goolam, M. *et al.* Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell* **165,** 61–74 (2016).
9. White, M. D. *et al.* Long-Lived Binding of Sox2 to DNA Predicts Cell Fate in the Four-Cell Mouse Embryo. *Cell* **165,** 75–87 (2016).
10. Plachta, N., Bollenbach, T., Pease, S., Fraser, S. E. & Pantazis, P. Oct4 kinetics predict cell lineage patterning in the early mammalian embryo. *Nat. Cell Biol.* **13,** 117–123 (2011).
11. Shi, J. *et al.* Dynamic transcriptional symmetry-breaking in pre-implantation mammalian embryo development revealed by single-cell RNA-seq. *Development* **142,** 3468–3477 (2015).
12. Bischoff, M., Parfitt, D.-E. & Zernicka-Goetz, M. Formation of the embryonic-abembryonic axis of the mouse blastocyst: relationships between orientation of early cleavage divisions and pattern of symmetric/asymmetric divisions. *Development* **135,** 953–962 (2008).
13. Piotrowska-Nitsche, K. & Zernicka-Goetz, M. Spatial arrangement of individual 4-cell stage blastomeres and the order in which they are generated correlate with blastocyst pattern in the mouse embryo. *Mech. Dev.* **122,** 487–500 (2005).
14. Zheng, Z., Li, H., Zhang, Q., Yang, L. & Qi, H. Unequal distribution of 16S mtrRNA at the 2-cell stage regulates cell lineage allocations in mouse embryos. *Reproduction* **151,** 351–367 (2016).
15. Bordeaux, J. *et al.* Antibody validation. *Biotechniques* **48,** 197–209 (2010).
16. Trenchevska, O., Nelson, R. W. & Nedelkov, D. Mass spectrometric immunoassays for

discovery, screening and quantification of clinically relevant proteoforms. *Bioanalysis* **8,** 1623–1633 (2016).
17. Schnell, U., Dijk, F., Sjollema, K. A. & Giepmans, B. N. G. Immunolabeling artifacts and the need for live-cell imaging. *Nat. Methods* **9,** 152–158 (2012).
18. Teves, S. S. *et al.* A dynamic mode of mitotic bookmarking by transcription factors. *Elife* **5,** 1–24 (2016).
19. Zhu, Y. *et al.* Nanodroplet processing platform for deep and quantitative proteome profiling of 10 – 100 mammalian cells. *Nat. Commun.* DOI: 10.1038/s41467-018-03367-w (2018).
20. Gao, Y. *et al.* Protein Expression Landscape of Mouse Embryos during Pre-implantation Development. *Cell Rep.* **21,** 3957–3969 (2017).
21. Hughes, A. J. *et al.* Single-cell western blotting. *Nat. Methods* **11,** 749–55 (2014).
22. Kang, C.-C. *et al.* Single cell–resolution western blotting. *Nat. Protoc.* **11,** 1508–1530 (2016).
23. Kang, C. C., Lin, J. M. G., Xu, Z., Kumar, S. & Herr, A. E. Single-cell western blotting after whole-cell imaging to assess cancer chemotherapeutic response. *Anal. Chem.* **86,** 10429–10436 (2014).
24. Kim, J. J., Sinkala, E. & Herr, A. E. High-selectivity cytology via lab-on-a-disc western blotting of individual cells. *Lab Chip* **17,** 855–863 (2017).
25. Tsichlaki, E. & Fitzharris, G. Nucleus downscaling in mouse embryos is regulated by cooperative developmental and geometric programs. *Sci. Rep.* **6,** 1–7 (2016).
26. Epifano, O., Liang, L. F., Familari, M., Moos, M. C. & Dean, J. Coordinate expression of the three zona pellucida genes during mouse oogenesis. *Development* **121,** 1947–1956 (1995).
27. Martín-Coello, J., González, R., Crespo, C., Gomendio, M. & Roldan, E. R. S. Superovulation and in vitro oocyte maturation in three species of mice (Mus musculus, Mus spretus and Mus spicilegus). *Theriogenology* **70,** 1004–1013 (2008).
28. Marangos, P. in *Oogenesis: Methods and Protocols* (ed. Nezis, I. P.) 209–215 (Springer New York, 2016).
29. Chen, S., Lee, B., Lee, A. Y. F., Modzelewski, A. J. & He, L. Highly efficient mouse genome editing by CRISPR ribonucleoprotein electroporation of zygotes. *J. Biol. Chem.* **291,** 14457–14467 (2016).
30. Mozziconacci, J., Sandblad, L., Wachsmuth, M., Brunner, D. & Karsenti, E. Tubulin dimers oligomerize before their incorporation into microtubules. *PLoS One* **3,** 1–8 (2008).
31. Strumpf, D. *et al.* Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. *Development* **132,** 2093–2102 (2005).
32. Zhang, S. Sox2, a key factor in the regulation of pluripotency and neural differentiation. *World J. Stem Cells* **6,** 305 (2014).
33. Sinkala, E. *et al.* Profiling protein expression in circulating tumour cells using microfluidic western blotting. *Nat. Commun.* **8,** (2017).
34. Padovan-Merhar, O. *et al.* Single Mammalian Cells Compensate for Differences in Cellular Volume and DNA Copy Number through Independent Global Transcriptional Mechanisms. *Mol. Cell* **58,** 339–352 (2015).
35. Kempea, H., Schwabeb, A., Crémazya, F., Verschurea, P. J. & Bruggemanb, F. J. The volumes and transcript counts of single cells reveal concentration homeostasis and capture biological noise. *Mol. Biol. Cell* **26,** 797–804 (2015).

36. Jeong, J.-K. *et al.* Evaluation of reference genes in mouse preimplantation embryos for gene expression studies using real-time quantitative RT-PCR (RT-qPCR). *BMC Res. Notes* **7,** 675 (2014).
37. Revil, T., Gaffney, D., Dias, C., Majewski, J. & Jerome-Majewska, L. A. Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics* **11,** 399 (2010).
38. Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat. Genet.* **40,** 1413–1415 (2008).
39. Wang, E. T. *et al.* Alternative isoform regulation in human tissue transcriptomes. *Nature* **456,** 470–476 (2008).
40. Flemr, M. *et al.* A retrotransposon-driven dicer isoform directs endogenous small interfering rna production in mouse oocytes. *Cell* **155,** 807–816 (2013).
41. Morris, S. A., Guo, Y. & Zernicka-Goetz, M. Developmental Plasticity Is Bound by Pluripotency and the Fgf and Wnt Signaling Pathways. *Cell Rep.* **2,** 756–765 (2012).
42. Huh, D. & Paulsson, J. Random partitioning of molecules at cell division. *Proc. Natl. Acad. Sci.* **108,** 15004–15009 (2011).
43. Schultz, R. M. Regulation of zygotic gene activation in the mouse. *BioEssays* **15,** 531–538 (1993).
44. Nothias, J. Y., Miranda, M. & DePamphilis, M. L. Uncoupling of transcription and translation during zygotic gene activation in the mouse. *EMBO J.* **15,** 5715–25 (1996).
45. Darmanis, S. *et al.* Simultaneous Multiplexed Measurement of RNA and Proteins in Single Cells. *Cell Rep.* **14,** 380–389 (2016).
46. Schwanhausser, B. *et al.* Corrigendum: Global quantification of mammalian gene expression control. *Nature* **495,** 126–127 (2013).
47. Snider, N. T. & Omary, M. B. Post-translational modifications of intermediate filament proteins: Mechanisms and functions. *Nat. Rev. Mol. Cell Biol.* **15,** 163–177 (2014).
48. Macaulay, I. C., Ponting, C. P. & Voet, T. Single-Cell Multiomics: Multiple Measurements from Single Cells. *Trends Genet.* **33,** 155–168 (2017).
49. Zischewski, J., Fischer, R. & Bortesi, L. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol. Adv.* **35,** 95–104 (2017).
50. Johnson, D. E., Ostrowski, P., Jaumouillé, V. & Grinstein, S. The position of lysosomes within the cell determines their luminal pH. *J. Cell Biol.* **212,** 677–692 (2016).

Chapter 6: Dual nucleic acid and protein isoform measurements on low starting cell numbers

This work was conducted in collaboration with Dr. Andrew Modzelewski.

6.1 Introduction

Embryo-specific nucleic acid modifications, including retrotransposon-derived genomic modifications and alternative splicing of mRNA, is crucial for the development of mammalian embryos[1]. The importance of alternative splicing as a developmental regulatory mechanism, however, has been established by monitoring mRNA isoform levels and not protein levels. Resolving if all genomic modifications and mRNA isoforms translate to protein variations remains an intriguing question that requires simultaneously measuring (i) nucleic acids (DNA variations and mRNA isoforms) and (ii) protein isoforms in early-stage embryos.

Recently introduced technologies allow interrogation of the genome, epigenome, transcriptome, metabolome and protein at single-cell resolution[2–6], with some tools capable of performing measurements of proteins and DNA and/or RNA from single cells[2]. However, many of these tools perform the protein measurements on different set of single cells than the nucleic acid measurements[2]. Furthermore, the specificity of the protein measurement of these methods relies on antibodies alone, that are subject to nonspecific cross-reactivity and cannot detect isoforms without isoform-specific antibodies. As a result, identifying different proteoforms arising from modifications to the DNA or mRNA therefore remains extremely challenging.

In this chapter, we present a method for performing dual nucleic acid and protein isoform measurements on low starting cell numbers (1-100), a range that includes the cell numbers in the morula (16-30) and blastocyst (30-100) stages[7]. Our technique integrates fractionation polyacrylamide gel electrophoresis (fPAGE) of 1-100 cells with off-chip analysis of nuclei. We designed polyacrylamide gel-based device, or GelBond™-PA gel device, consisting of a 100-150 um polyacrylamide (PA) gel covalently grafted to a polyester film (GelBond™). After settling cells into microwells patterned on the PA gel, we perform fractionation single-cell polyacrylamide gel electrophoresis (fractionation PAGE). The plastic GelBond™ substrate allows us to then laser-excise areas of the gel, or gel rafts, containing the microwells with the fractionated nuclei. The nuclei-containing gel rafts are then collected to perform the genomic measurement, as well as transcriptomic measurements, given the conservation of general expression differences of genes between nuclei and whole cells[8,9]. In this way, we are able to measure expression of protein isoforms from the cytoplasmic fraction of 1-100 cells while achieving analysis of either DNA or mRNA retained in the nuclei.

6.2 Materials and Methods

Chemical Reagents. Tetramethylethylenediamine (TEMED, T9281), ammonium persulfate (APS, A3678), β-mercaptoethanol (M3148), 30%T/2.7%C acrylamide/bis-acrylamide (37.5:1) (A3699), bovine serum albumin (BSA, A9418), Tyrode's solution (T1788), trypsin 10X (59427C),

digitonin (D141), sucrose (S0389-500G), magnesium chloride (M8266) and HEPES (90909C) were purchased from Sigma-Aldrich. Triton X-100 (BP-151), phosphate-buffered saline (PBS, pH 7.4, 10010023), SYBR Gold (S11494), agarose (BP-1356-500) and TRIzol™ (15596026) were purchased from Thermo Fisher Scientific. Premixed 10X tris-glycine EP buffer (25 mM Tris, pH 8.3; 192 mM glycine; 0.1% SDS) was purchased from Bio-Rad. Tris buffered saline with Tween-20 (TBST) was prepared from 20X TBST (sc-24953, Santa Cruz Biotechnology, Dallas, TX). Deionized water (18.2 MΩ) was obtained using an Ultrapure water system from Millipore. N-[3-[(3-Benzoylphenyl)formamido]propyl] methacrylamide (BPMAC) was custom synthesized by Pharm-Agra Laboratories (Brevard, NC). GelSlick™ (50640) and Lonza™ GelBond™ PAG Film for Acrylamide Gels (BMA54746) was purchased from Lonza. Taq PCR kit (E5000S), proteinase K (P8107S) were purchased from New England Biosciences.

Buffer Compositions. Fractionation lysis buffer: 0.125 mg/mL digitonin, 1% v/v Triton X-100 and 0.5 X Tris-glycine. Nuclei wash buffer: 320 mM sucrose, 5 mM MgCl2, 10 mM HEPES.

Cell Culture. U251 human glioblastoma cells were obtained from the UC Berkeley Tissue Culture Facility via the American Type Culture Collection and stably transduced with TurboGFP via lentiviral infection (multiplicity of infection=10). Cells were cultured in high-glucose Dulbecco's modified eagle medium (DMEM) (11965, Life Technologies) supplemented with 1× MEM nonessential amino acids (11140050, Life Technologies), 100 U mL^{-1} penicillin-streptomycin (15140-122, Life Technologies), 1 mM sodium pyruvate (11360-070), and 10% fetal bovine serum (JR Scientific, Woodland) in an incubator at 37 °C with humidified 5% CO_2 air.

Device Fabrication. SU-8 wafers, fabricated by photolithography as previously reported[10], were used as molds to cast PA gel devices. SU-8 posts on wafers, which later translate into microwells in the PA gel, were 200 μm in diameter and 200 μm in height. Briefly, PA precursor solution including acrylamide/bis-acrylamide (10%T) and 3 mM BPMAC was degassed with sonication for 9 min. 0.08% APS and 0.08% TEMED were added to precursor solution and solution was pipetted between the SU-8 wafer (rendered hydrophobic with Gel Slick™ solution) and a GelBondTM Film cut to the size of a standard glass microscope slide (25 mm x 75 mm). After chemical polymerization (20 min) the GelBond™-PA gel devices (thin PA gel layer covalently grafted onto the GelBond™ surface) were lifted from wafer, rinsed with deionized water and stored in hydrated (DI water) at 4°C until use.

Fractionation PAGE of 5-100 U251-TurboGFP cells. TurboGFP-expressing U251 cells were harvested from tissue culture plates by incubation in trypsin/EDTA (15090046, Thermo) at 37°C for 5 min. Trypsin was inactivated by addition of FBS, and cells were pelleted by centrifugation at 100 x g. After removal of supernatant, cells were resuspended in PBS at $1x10^6$ cells / mL. 1 mL of this cell suspension was pipetted over the PA gel device and cells were allowed to settle into microwells for 10 min. Excess cells not settled into microwells were then washed off the PA gel surface with PBS and microwells were imaged by bright field and fluorescence microscopy to collect data on number of cells per microwell and TurboGFP expression. The device was placed into an EP chamber, fractionation lysis buffer (RT, 12 mL) was poured over the gel followed by incubation for 1 min. Electrophoresis was performed at 40 V/cm for 2-3 min (depending on the assay). Immobilization of proteins by photocapture was carried out by illumination with a UV light

source (100% power, 45 s, Lightningcure LC5, Hamamatsu). Gels were quickly placed in ice-cold nuclei wash buffer and buffer was exchanged three times before proceeding to laser excision.

Laser excision of PA gel device into gel rafts. The PA gel was kept hydrated at all times with nuclei wash buffer and kept over ice between excision events. A CO_2 laser (HL40-5G-110, Full Spectrum Laser) was used to excise gel rafts from GelBond™-PA gel devices. The device was placed with the PA gel face down onto a clear acrylic sheet (McMaster-Carr) engraved with a 5 x 5 mm grid (Figure 1, a). Using a bright-field microscope, microwells were aligned to be horizontally centered above a grid square and approximately 1 mm away from the top edge of each square (Figure 1, b). The laser was aligned over the left corner of a grid square and programmed to cut at 10% power, 20 speed and 1 pass (Figure 1, b).

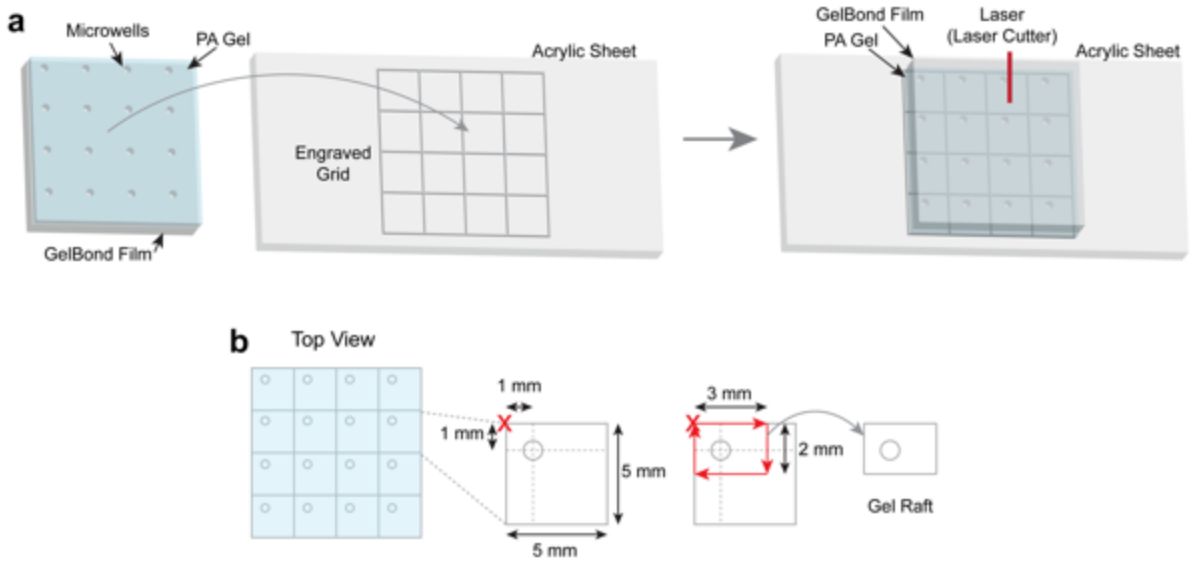


Figure 1: Set up for laser excision of PA gel device into gel rafts. (a) Schematic of PA gel device (left) and acrylic sheet onto which a 5 mm x 5 mm grid has been engraved. PA gel device is placed with PA gel facing down onto the grid, so laser can cut through the GelBondTM film first, then the PA gel. (b) Top view of the PA gel and grid assembly. PA gel is aligned over the grid so that microwells are approximately 1 mm away from the top left edge of their corresponding grid square. Laser is aligned over the top left edge and programmed to cut a 3 mm x 2 mm area, creating the gel raft.

Immunoprobing and microarray imaging of photocaptured protein and immunoblots. After laser excision, devices were washed in deionized water, and a 25 mm x 75 mm coverslip was placed over the hydrated PA gel. Devices were imaged in a microarray scanner (Genepix 4300A, Molecular Devices) for photocaptured TurboGFP protein with the cover slip facing down (Figure 2). Devices were then washed in 1X TBST for at least 1 hr before probing with antibodies. Primary antibodies were incubated at 1:10 dilution (80 µL/gel, in 2% BSA in 1X TBST), while fluorophore-conjugated secondary antibodies were incubated at 1:20 dilution (80 µL/gel, in 2% BSA in TBST). Devices were scanned again for fluorescence immunblot signal. In order to strip bound antibodies and reprobe for new targets, gels were treated with 2% SDS, 0.8% □-mercaptoethanol and 62.5 mM Tris base at 55ºC for three hours, washed in TBST (1 hr) twice and then reprobed.

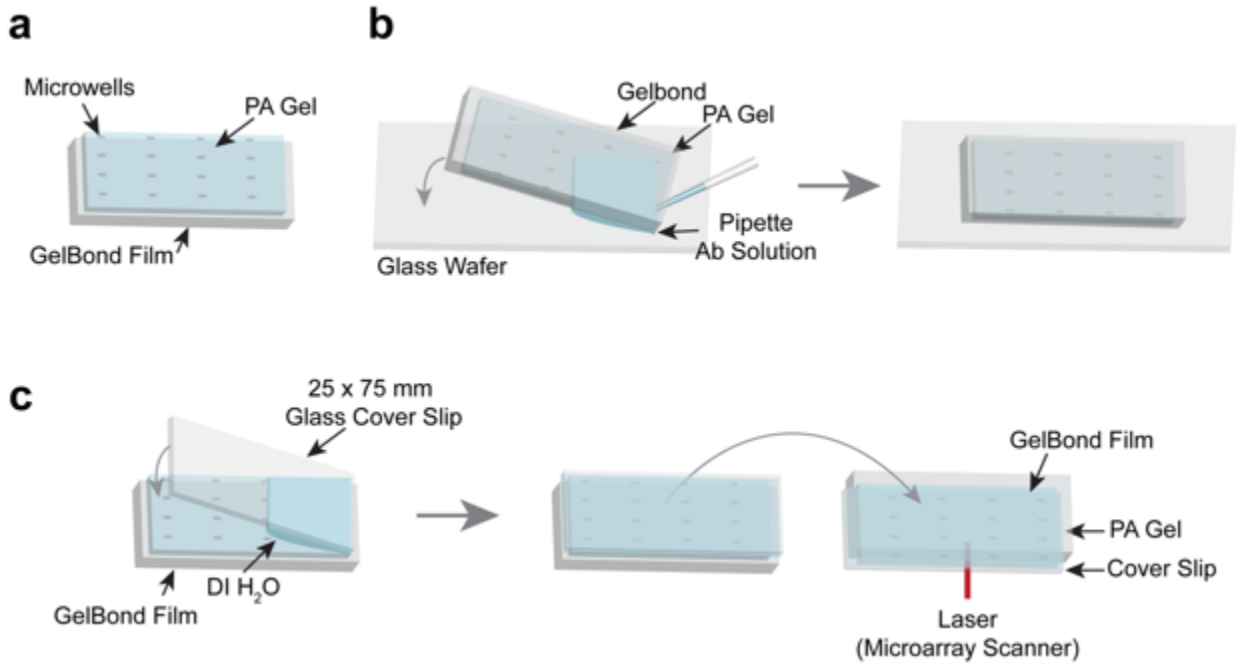


Figure 2: Gel incubation set ups. (a) Schematic of the PA gel device structure. (b) Antibody solution incubation. PA gel device is placed facing down over a glass wafer. Antibody solution is pipetted between the PA gel and the glass wafer, allowing the solution to wick across the surface of the gel. (c) Schematic for hydrating gel in preparation for microarray scanner read-out.

Antibodies. Antibodies employed for analysis of embryos include: rabbit anti-□-tubulin (Abcam, ab6046, polyclonal), rabbit anti-TurboGFP (PA5-22688), Alexa Fluor™ 647-conjugated donkey anti-rabbit secondary (A31572, Thermo Fisher) and Alexa Fluor™ 555-conjugated donkey anti-rabbit secondary (A31572, Thermo Fisher).

Single-Gel Raft PCR. After laser excision, gel rafts were placed into a 0.5 mL PCR tube containing 2.5 μL Molecular Grade water, 1 μL SDS (17 μM to final concentration of 3.4 μM) and 1.5 μL proteinase K. Tubes were incubated at 45oC for 15 min followed by proteinase K inactivation by incubation at 95oC for 20 min. Next, the following were added to each tube: 2.5 μL TurboGFP primers (at 500 nM, purchased from Integrated DNA Technologies, sequences: (5'TGA TGG GCT ACG GCT TCT A, 5'GTG TTG CTG TGA TCC TCC TC), 1 μL dNTPs (at 200 μM, Taq PCR Kit), 0.25 μL Taq polymerase (Taq PCR Kit), 5 μL of Standard Taq Reaction Buffer 10X (Taq PCR Kit) and water up to 50 μL. Template DNA (~200 ng/μL) extracted TurboGFP-U251 lysate was added to positive control tubes. Negative controls did not contain DNA or gel rafts. For amplification of the TurboGFP gene, the following cycle steps were programmed using a thermal cycler (PTC-100TM, MJ Research Inc): the first stage at 95oC for 10 min, the second stage (annealing at 51oC for 30 s, extension at 72oC for 30 s, denaturation at 95oC for 30 s) for 45 amplification cycles, and a final stage at 72oC for 10 min. PCR products were analyzed on a 1% agarose gel by electrophoresis. SYBR Gold was used at 1X to stain agarose gels and a ChemiDocTM XRS+Gel Imaging System (Bio-Rad) was used to image the DNA bands. Gels were analyzed by densitometry using ImageJ software11.

Single-Gel Raft Quantitative Reverse Transcription Polymerase Chain Reaction
Once excised, each gel raft was transferred to one centrifuge tube, immediately followed by the addition of 100 uL of Trizol Reagent (15596026, Thermo). Sample were stored in -80^{o}C until RNA preparation. RNA was isolated following the manufacturer's protocol but with the following alterations: Overall volumes were reduced to a starting input of 100 uL, following the transfer of the supernantant, an additional 100uL of water was added to the sample in order to collect residual RNA not collected in the first transfer, 1 uL of glycogen (R0561, Thermo) was added in order to aid in recovery of total RNA, RNA pellet was resuspended in 10 uL of water, all of which was used for cDNA synthesis as per manufacturer's protocol (iScript, 1725037, BioRad). All qRT-PCR reactions were performed using SSO Universal SYBR Green SuperMix, as per manufacturer instructions (1725275, BioRad). Primer sequences used were GAPDH (5' AGCCACATCGCTCAGACAC, 5' GCCCAATACGACCAAATCC)[12] and TurboGFP (5'TGA TGG GCT ACG GCT TCT A, 5'GTG TTG CTG TGA TCC TCC TC). All qRt-PCR analyses were performed on the StepOnePlus Real Time PCR system (437660, Thermo).

6.3 Results and Discussion

6.3.1 Design of assay for measuring protein and nucleic acids from the same 1-100 mammalian cells

To perform dual protein-nucleic acid measurements on the same 1-100 mammalian cells, we designed an assay that integrates (i) electrophoretic separation of cytoplasmic proteins and (ii) extraction of nucleic acids from the nuclei. Our assay begins with settling cells into the microwells patterned onto the GelBond^{TM}-PA gel device, a 200 μm thick polyacrylamide (PA) gel covalently bound to the treated surface of a flexible polyester film (Gelbond^{TM} PAG Film) (Figure 3, a, left image). After cells settle into microwells by gravity, a buffer exchange step from PBS to a fractionation lysis buffer[13] achieves *in situ* lysis of the cytoplasmic fraction of cells (Figure 3, a). An electric field is then applied to (i) inject solubilized proteins through the microwell wall and into the PA gel layer and (ii) achieve polyacrylamide gel electrophoresis (PAGE), which separates proteins by molecular mass along the separation lane, or region abutting the microwell. Proteins are then photocaptured by UV-light activation of benzophenone moieties incorporated in the PA gel matrix. After cytoplasmic protein PAGE, the nuclei remain intact in the microwells. Nuclei are extracted from the device by CO_2 laser-excision of a 2 mm x 3 mm area of the GelBond^{TM}-PA device into *gel rafts* containing the intact nuclei. These gel rafts are then placed into reaction vessels to perform extraction and off-chip analysis of either DNA or mRNA. The remaining GelBond^{TM}-PA gel device is then probed for proteins with fluorescently-labeled antibodies, yielding protein immunoblots from the original settled cells.

To optimize our workflow, we utilized TurboGFP-transduced U251 human glioblastoma cells, where the expression of fluorescent TurboGFP was a useful protein model for visualizing protein lysis, injection, PAGE and photocapture. We first settled TurboGFP-U251 cells stained with nuclear Hoechst dye settled into a microwell of a GelBond^{TM}-PA gel device (Figure 3, b, top). After fractionation lysis, PAGE and photocapture, we observed a TurboGFP band in the separation lane along with absence of TurboGFP fluorescence in the microwell, suggesting complete lysis and injection of the cytoplasmic proteins into the PA gel was achieved (Figure 3, b, bottom).

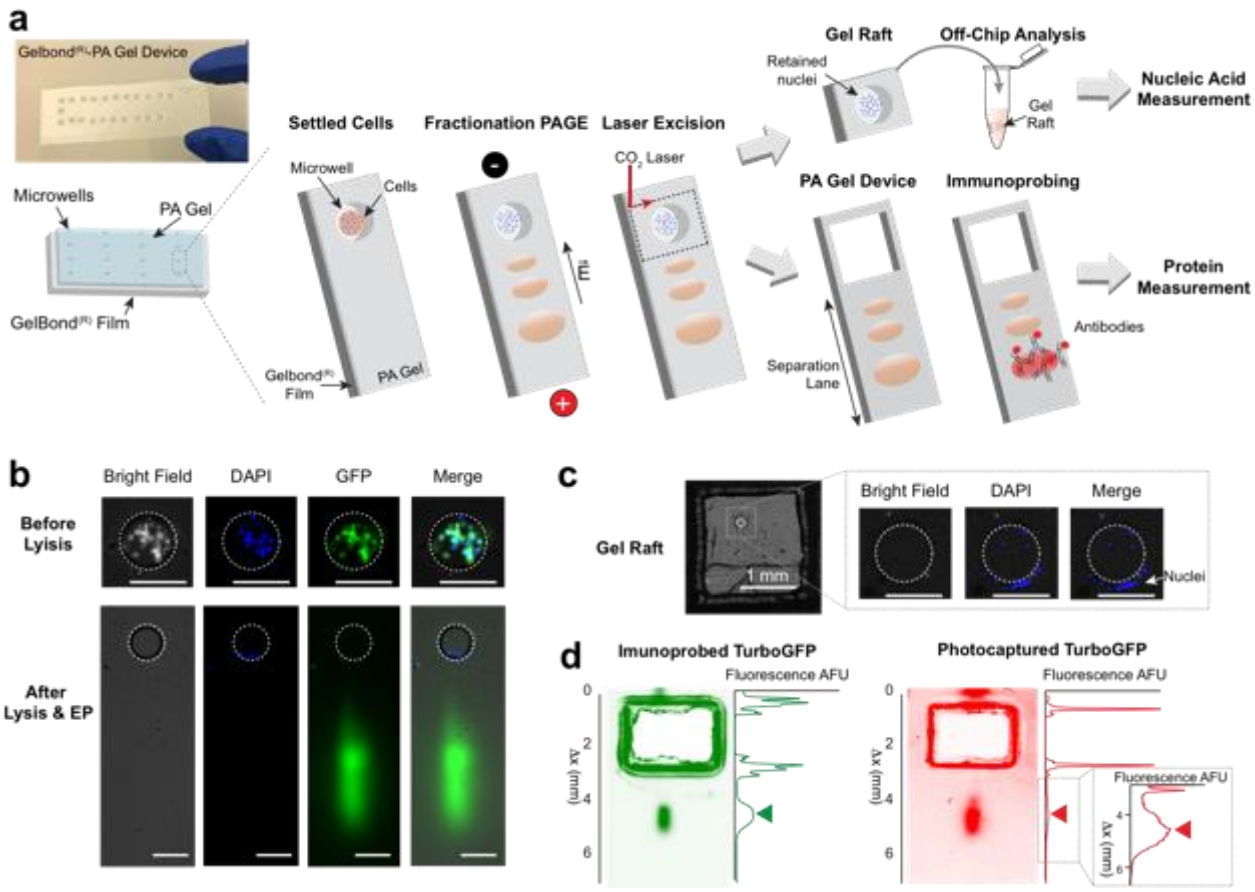


Figure 3: Fractionation PAGE is coupled with laser excision of microwells into gel rafts for off-chip analysis of nucleic acids. (a) The GelBondTM-PA gel device is comprised of a thin polyacrylamide (PA) gel covalently grafted onto a GelBondTM polymer film substrate and stippled with microwells. 1-100 cells are settled into microwells of the GelBondTM-PA device and lysed with fractionation buffer. An electric field is then applied to inject the solubilized cytoplasmic proteins into the PA gel and separate them by molecular mass, which are then immobilized to the gel by UV-light activation of benzophenone methacrylamide moieties in the PA matrix. A region of the GelBondTM-PA gel device that contains retained nuclei is excised to create a gel raft, that can then be analyzed off-chip for DNA or mRNA. The remaining GelBondTM-PA gel device can be probed for proteins with fluorescently-labeled antibodies, yielding protein immunoblots from the original settled cells. (b) Fractionation PAGE retains nulcei in microwells. Top row displays bright field, DAPI and GFP micrographs of TurboGFP-U251 cells settled into a microwell of a GelBondTM-PA gel device, prior to the cell lysis step. On bottom row, bright field, DAPI and GFP micrographs of microwell and PA gel abutting the microwell (separation lane) after fractionation PAGE, when cytoplasmic proteins have been electrophoresed into the PA gel while nuclei are retained in the microwell. (c) Excision of GelBondTM-PA gel into gel rafts. Excision of microwells from GelBondTM-PA gel creates gel rafts (left) that allow extraction of nuclei for off-chip analysis of nucleic acids. Bright field micrograph shows one gel raft. The retention of the nuclei can be verified by the fluorescence imaging of the Hoechst-stained nuclei, as displayed in the bright field, DAPI fluorescence and merged micrographs of a gel raft microwell. (d) Immunoblots of photocaptured and immunoprobed TurboGFP. On the left, false-colored micrograph of photocaptured TurboGFP protein after EP separation. Corresponding intensity profile is shown on adjacent to micrograph. On the right, false-colored micrograph of the immunoprobed TurboGFP immunblot, with corresponding intensity profile. Arrow heads mark the position of the protein peak.

Next, after placing the GelBond™-PA gel device in nuclei wash buffer in order to maintain the integrity of the nuclei, we excised gel rafts containing the microwells (Figure 3, c, left). Fluorescence imaging of the Hoechst-stained nuclei confirmed the presence of the nuclei in the microwells (Figure 3, c, right). The remaining GelBond™-PA gel device was imaged for native TurboGFP signal and then incubated with primary antibodies against TurboGFP followed by Alexa Fluor™ 555-conjugated secondary antibodies and imaged for resulting TurboGFP immunoblots (Figure 3, d).

6.3.2 Laser excision into gel rafts does not compromise integrity of nuclei

To validate the integrity of nuclei and viability of nucleic acid analysis after retrieval, we performed amplification of the TurboGFP gene by polymerase chain reaction (PCR) from gel rafts containing a single nucleus. We designed microwells to isolate single TurboGFP-expressing U251 cells (32 μm in diameter, 40 μm in height). After fractionation and PAGE of the cytoplasmic fraction, single nuclei retained in the microwells were excised into gel rafts. We inspected the gel rafts under an epifluorescence microscope for the Hoechst-stained nuclei to verify retention of the nucleus. Gel rafts were then placed into separate reaction vessels (centrifuge tubes) for PCR amplification of the TurboGFP gene. PCR products were analysed on an agarose gel, revealing successful amplification of the TurboGFP from the gel raft (Figure 4, b & c). A 78-fold increase compared to the negative control of the gel raft band, and a 33-fold decrease with respect to the positive control indicates amplification of TurboGFP gene from the gel raft sample, validating the viability of DNA extraction from nuclei in gel rafts.

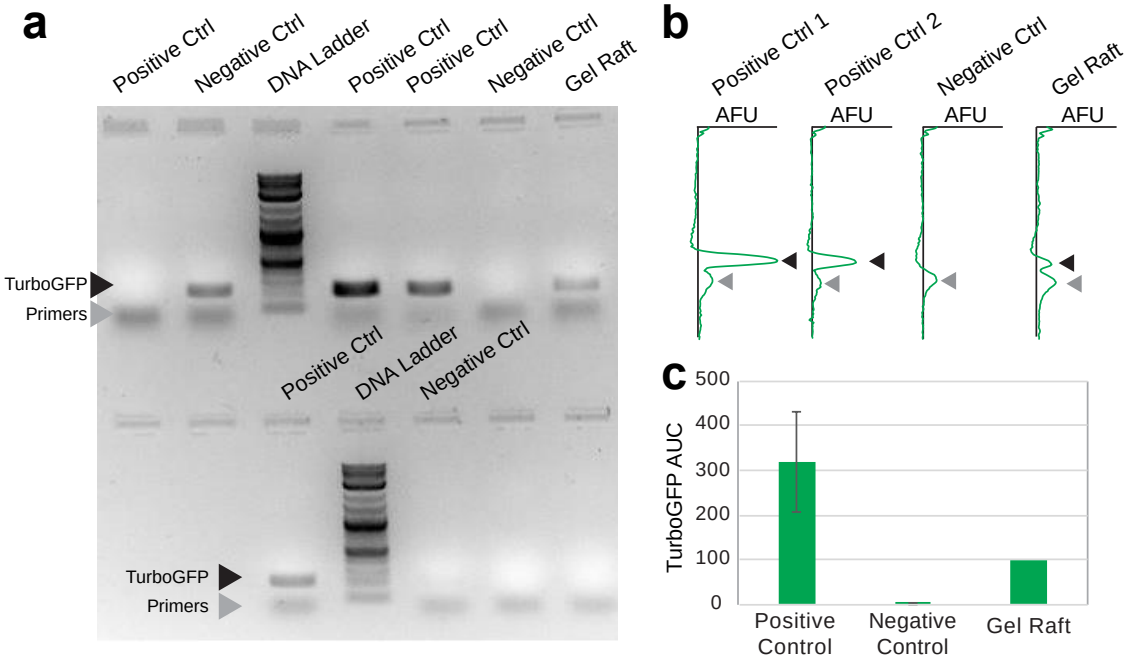


Figure 4: PCR amplification of DNA from a gel raft containing a single nucleus. (a) Agarose gel electrophoresis of DNA amplified for the TurboGFP gene by PCR from samples including positive controls (10 ug DNA extracted from TurboGFP-U251 lysate), negative controls (no DNA) and gel raft containing one TurboGFP-U251 nucleus. (b) Intensity profiles for agarose gel lanes corresponding to positive controls, negative control and the gel raft that contained a single TurboGFP-U251 nucleus. The gel raft band shows a 78-fold increase with regards to the negative control and 33-fold decrease with respect to the positive control.

6.3.3 Photocaptured and immunoblotted protein fluorescence signal correlates with protein expression prior to lysis

We next evaluated whether protein measured after lysis, EP separation, and immunoblotting accurately measures protein abundance prior to lysis. We used the TurboGFP protein in TurboGFP-expressing U251 cells as our protein target in order to use fluorescence as a convenient measure of protein abundance. We first imaged TurboGFP-U251 cells settled into microwells prior to lysis and computed fluorescence intensity (AFU). We then ran fractionation PAGE and scanned the PA gel device for photocaptured, native TurboGFP fluorescence. We calculated total fluorescence by performing area-under-the-curve (AUC) analysis. Finally, we immunoprobed the PA gel devices with primary antibodies against TurboGFP (rabbit-anti-TurboGFP), followed by fluorophore-conjugated secondary antibodies (Alexa Fluor ™ 555 donkey-anti-rabbit), and computed total immunoprobed signal (AUC).

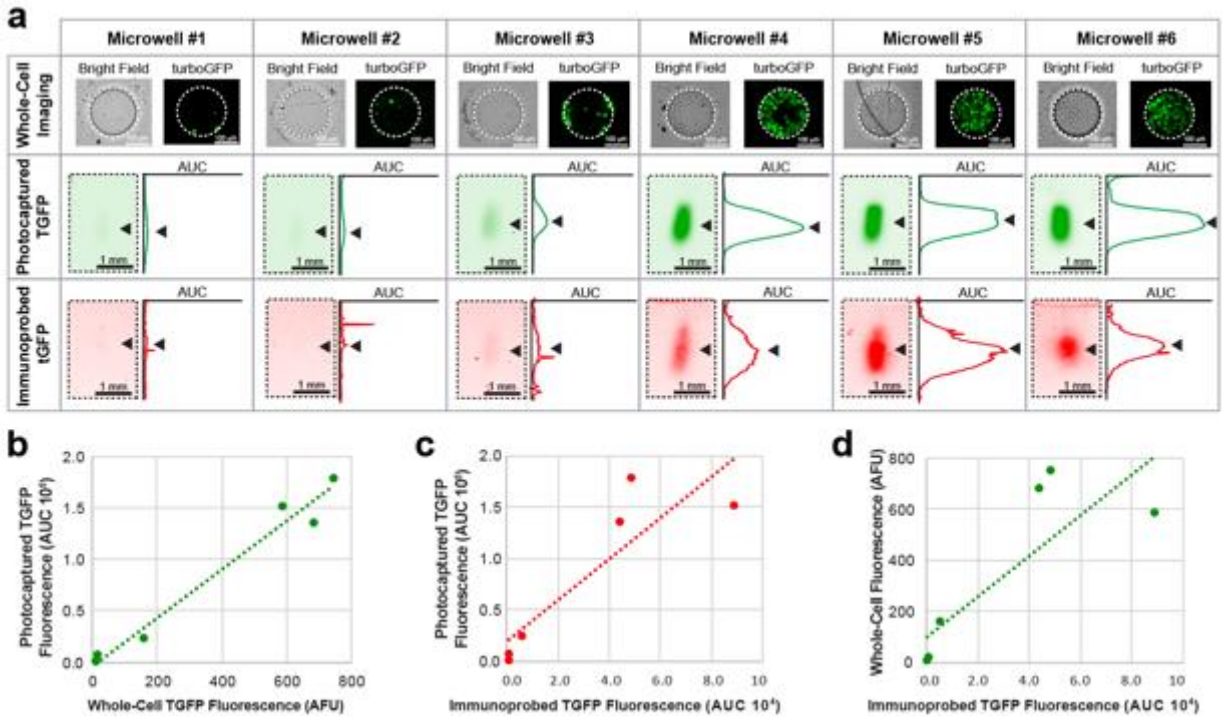


Figure 5: Photocaptured and immunoblotted signal correlates with in-well fluorescence signal prior to lysis in turbo-GFP expressing cells. (a) Bright field and false-colored fluorescence micrographs of TurboGFP-expressing U251 cells settled into microwells, with corresponding false-colored fluorescence micrographs of TurboGFP immunoblots. Fluorescence intensity profiles are shown to the right of immunoblots, with black arrows marking the position of protein peaks. (b) Bivariate plot of whole-cell TurboGFP fluorescence prior to lysis (AFU) and photocaptured TurboGFP fluorescence (AUC) demonstrating a significant, positive linear correlation (Pearson correlation, $\rho = 0.9857$, $R^2 = 0.9715$, p value = 0.00031, N = 6 microwells). (c) Bivariate plot of photocaptured and immunoprobed TurboGFP fluorescence (AUC) displaying a strong linear correlation (Pearson correlation, $\rho = 0.0.8787$, $R^2 = 0.7722$, p value = 0.02118, N = 6 microwells). (d) Bivariate plot of whole-cell fluorescence prior to lysis (AFU) and immunoprobed TurboGFP fluorescence (AUC), also showing a positive linear correlation (Pearson correlation, $\rho = 0.8254$, $R^2 = 0.68134$, p value = 0.04307, N = 6 microwells).

When comparing (i) whole-cell TurboGFP prior to lysis, (ii) native signal from the photocaptured TurboGFP, and (iii) immunoprobed signal from fluorophore-conjugated antibodies against TurboGFP (Figure 5, a), we found that whole-cell TurboGFP fluorescence demonstrated a significant positive, linear correlation with signal from both photocaptured TurboGFP and immunoblotted fluorescence signal (Pearson correlation, N = 6 microwells, where $\rho = 0.9857$, $R^2 = 0.9715$, p value = 0.00031 for whole-cell TurboGFP fluorescence versus photocaptured TurboGFP signal, and $\rho = 0.0.8787$, $R^2 = 0.0.7722$, p value = 0.02118 for whole-cell TurboGFP fluorescence versus immunoprobed TurboGFP signal) (Figure 5, b&c). Likewise, photocaptured and immunoblotted signal showed a significant positive, linear correlation (Pearson correlation, $\rho = 0.8254$, $R^2 = 0.68134$, p value = 0.04307, N = 6 microwells) (Figure 5, d). These results indicate that measuring signal from resulting immunoblots accurately estimates protein abundance in cells prior to lysis, EP, and photocapture.

6.3.4 TurboGFP mRNA levels show higher correlation with immunoprobed, and not whole-cell fluorescence or photocaptured, TurboGFP signal

We finally examined whether we could recover and measure mRNA from excised microwells. After fractionation PAGE and excision of GelbondTM-PA gel into gel rafts, the gel rafts were placed into separate reaction vessels containing TRIzol to extract mRNA from the intact nuclei. Isolated RNA was analyzed for TurboGFP and reference gene GAPDH using quantitative real-time PCR analysis (qRT-PCR). Results show that whole-cell fluorescence and photocaptured protein signal do not show a significant correlation with mRNA levels (Pearson correlation, N = 5 microwells, $\rho = 0.514$ and 0.559, p values = 0.327 and 0.376 for whole-cell TurboGFP fluorescence and photocaptured TurboGFP, respectively) (Figure 6, a&b). Conversely, TurboGFP mRNA levels show a significant positive association with immunoprobed TurboGFP signal (Pearson correlation, N = 5 microwells, $\rho = 0.882$, p value 0.0476) (Figure 6, c).

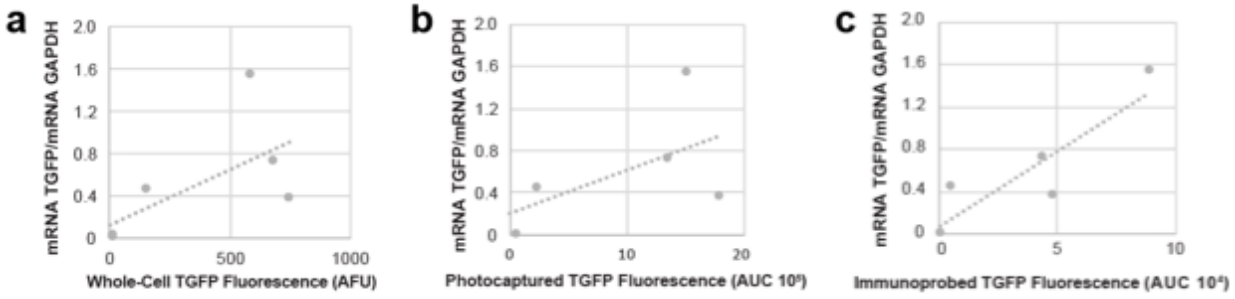


Figure 6: mRNA TurboGFP levels correlate with immunoprobed TurboGFP, but not whole-cell TurboGFP fluorescence and photocaptured TurboGFP. (a) Bivariate plot of whole-cell TurboGFP fluorescence prior to lysis (AFU) and TurboGFP mRNA normalized by GAPDH demonstrating a non-significant, positive linear correlation (Pearson correlation, ρ =0.514, p value = 0. 0.376, N = 5 microwells). (c) Bivariate plot of photocaptured and TurboGFP mRNA normalized by GAPDH demonstrating displaying a non-significant positive correlation (Pearson correlation, $\rho = 0.559$, p value = 0.327, N = 5 microwells). (d) Bivariate plot of immunoprobed TurboGFP fluorescence (AUC) and TurboGFP mRNA normalized by GAPDH demonstrating a significant positive linear correlation (Pearson correlation, $\rho = 0.0.882$, p value = 0. 04761, N = 5 microwells).

Interestingly, immunoprobed TurboGFP is the only protein signal significantly correlated with TurboGFP mRNA. These results suggest that native TurboGFP fluorescence may not be an accurate proxy for protein abundance because it requires correct folding of the TurboGFP protein. Immunoprobing, on the other hand, is performed in denaturing conditions in order to ensure epitopes of all proteins are available to antibodies. These results suggest the potential for measuring unlabeled endogenous proteins, where the starting concentration prior to lysis cannot be determined through fluorescence.

6.4 Conclusions

Assessing whether specific modifications in genomic DNA and frequent alternative splicing drive important mechanisms in preimplantation development requires measuring both nucleic acids and protein isoforms. Here we designed an assay for simultaneous measurement of protein isoforms and nucleic acids from low starting numbers of mammalian cells. We demonstrated that signal from immunoprobed protein correlates strongly with protein expression prior to lysis in TurboGFP-expressing cells. We also measured both mRNA and DNA from retrieved nuclei, with positive amplification of TurboGFP gene and mRNA, demonstrating our ability to recover, isolate and amplify nucleic acids from gel rafts. The cell number range over which we performed these measurements (from 1 to approximately 100 cells) includes the cell numbers in the latest stages of the preimplantation embryo, the morula (16-30) and blastocyst (30-100). The future application of this tool to morula and blastocysts will help reveal the mechanisms by which embryo-specific nucleic acid modifications to both genomic DNA and mRNA orchestrate the growth and development of mammalian embryos.

6.5 References

1. Revil, T., Gaffney, D., Dias, C., Majewski, J. & Jerome-Majewska, L. A. Alternative splicing is frequent during early embryonic development in mouse. *BMC Genomics* **11,** 399 (2010).
2. Macaulay, I. C., Ponting, C. P. & Voet, T. Single-Cell Multiomics: Multiple Measurements from Single Cells. *Trends Genet.* **33,** 155–168 (2017).
3. Gawad, C., Koh, W. & Quake, S. R. Single-cell genome sequencing: Current state of the science. *Nat. Rev. Genet.* **17,** 175–188 (2016).
4. Schwartzman, O. & Tanay, A. Single-cell epigenomics: Techniques and emerging applications. *Nat. Rev. Genet.* **16,** 716–726 (2015).
5. Wu, M. & Singh, A. K. Single-Cell Protein Analysis. *Curr. Opin. Biotechnol.* **23,** 83–88 (2012).
6. Stegle, O., Teichmann, S. A. & Marioni, J. C. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* **16,** 133–145 (2015).
7. Wang, H. & Dey, S. K. Roadmap to embryo implantation: Clues from mouse models. *Nat. Rev. Genet.* **7,** 185–199 (2006).
8. Tang, F. *et al.* RNA-Seq analysis to capture the transcriptome landscape of a single cell. *Nat. Protoc.* **5,** 516–535 (2010).
9. Grindberg, R. V. *et al.* RNA-sequencing From Single Nuclei. *Proc. Natl. Acad. Sci.* **110,** 19802–19807 (2013).
10. Kang, C.-C. *et al.* Single cell–resolution western blotting. *Nat. Protoc.* **11,** 1508–1530

(2016).
11. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ : 25 years of image analysis HISTORICAL commentary NIH Image to ImageJ : 25 years of image analysis. *Nat. Methods* **9,** 671–675 (2012).
12. Zheng, Z., Li, H., Zhang, Q., Yang, L. & Qi, H. Unequal distribution of 16S mtrRNA at the 2-cell stage regulates cell lineage allocations in mouse embryos. *Reproduction* **151,** 351–367 (2016).
13. Yamauchi, K. A. & Herr, A. E. Subcellular western blotting of single cells. *Microsystems Nanoeng.* **3,** 16079 (2017).

Appendix A: Matlab script for mapping smooth muscle cell maturation

The following script was written in order to map the differentiation of smooth muscle cells (SMCs) from immature-like phenotype (high a-SMA, low CNN-1 and SM-MHC) to their characteristic contractile, mature phenotype (high expression of all three maturation markers). Assuming that at any given time, SMCs exist as a heterogeneous population with regards to maturation stage, we follow a similar approach to the SPADE algorithm (Spanning-tree Progression Analysis of Density-normalized Events[21]) developed for differentiation of stem cells, to represent multiparametric data on a simple 2D plot that maps the progression of cell differentiation.

```
function [c T M]=hierClust3(A)

%Input: A=[aSMA CNN1 MHC]

%1. Agglomerative Hierarchical Clustering
%   1. Find distance b/n objects using 'pdist'
%   2. Group into binary tree using 'linkage'
%   3. Determine where to cut tree using 'cluster'
Y=pdist(A);
Y=squareform(Y);
Z=linkage(Y);
%graph dendrogram for visualization
figure ('Name', 'Dendrogram')
%dendrogram(Z);%will simplify graph, not include all leaf nodes
dendrogram(Z,0);%graphs ALL observations, from leaf nodes
c=cophenet(Z,Y);%check consistency between linkage distances and pdist distances
if c<0.7%if low, recalculate using different methods
    Y=pdist(A,'cityblock');
    Z=linkage(Y,'average');
end
c=cophenet(Z,Y);
I=inconsistent(Z); %Check inconsistency coefficients of links

%create clusters using 'maxclust': defining max number of clusters
%   clustering using natural divisions did not work very well: T=cluster(Z,'cutoff',1.2);
T=cluster(Z,'maxclust',6);
%graph clusters
figure('Name', 'Agglomerative Clustering')
hold on
for i=1:max(T);
    scatter3(A(T==i,1),A(T==i,2),A(T==i,3), 'MarkerEdgeColor', [1-0.14*i 0 0.14*i])
```

```
end
%calculate centroids
cent=zeros(max(T),3);
for j=1:max(T);
    for k=1:3;
        cent(j,k)=mean(A(T==j,k));
    end
end
%plot centroids
i=1;
for i=1:max(T);
    scatter3(cent(i,1),cent(i,2),cent(i,3),'o', 'MarkerEdgeColor', [0 0 0])
end
xlabel aSMA
ylabel CNN1
zlabel MHC


%3. Minimum Spanning Tree
%build the 'graph' structure containing the centroids as nodes
n=max(T);

%calculate vectors s and t (pairs)
n_vect=zeros(n,1);
i=1;
for i=1:n
    n_vect(i)=i;
end
len_pairs=0.5*n^2-0.5*n;
s=zeros(len_pairs,1);
t=zeros(len_pairs,1);
i=1;
while s(i)<n
    s(i)=n_vect(i);
    t(i)=n_vect(i+1);
    j=1;
    while t(j)<n
        j=j+1;
        s(j)=s(i);
        t(j)=t(j-1)+1;
    end
    i=j;
end

if n==2
    s =[1]; %[first column of possible pairs]
    t =[2]; %[second column of possible pairs]
elseif n==4
    s =[1;1;1;2;2;3]; %[first column of possible pairs]
    t =[2;3;4;3;4;4]; %[second column of possible pairs]
elseif n==5
    s =[1;1;1;1;2;2;2;3;3;4]; %[first column of possible pairs]
    t =[2;3;4;5;3;4;5;4;5;5]; %[second column of possible pairs]
elseif n==6
    s =[1;1;1;1;1;2;2;2;2;3;3;3;4;4;5]; %[first column of possible pairs]
    t =[2;3;4;5;6;3;4;5;6;4;5;6;5;6;6]; %[second column of possible pairs]
elseif n==7
    s =[1;1;1;1;1;1;2;2;2;2;2;3;3;3;3;4;4;4;5;5;6]; %[first column of possible pairs]
    t =[2;3;4;5;6;7;3;4;5;6;7;4;5;6;7;5;6;7;6;7;7]; %[second column of possible pairs]
elseif n==8
   s =[1;1;1;1;1;1;1;2;2;2;2;2;2;3;3;3;3;3;4;4;4;5;5;6]; %[first column of possible pairs]
   t =[2;3;4;5;6;7;8;3;4;5;6;7;8;4;5;6;7;8;5;6;7;6;7;7]; %[second column of possible pairs]
end


%calculate the weights
weights=zeros(length(s),1);%[distances between pairs]
for i=1:length(s);
    weights(i)=sqrt((cent(s(i),1)-cent(t(i),1))^2+(cent(s(i),2)-cent(t(i),2))^2+(cent(s(i),3)-
cent(t(i),3))^2);
end
G = graph(s,t,weights);
figure ('Name', 'Minimum Spanning Tree')
p = plot(G,'EdgeLabel',G.Edges.Weight);
[M,pred] = minspantree(G);
highlight(p,M)
```

```
hold on
%graph MST on the centroids graph
%   look for the centroid pairs (edges) corresponding to each branch & graphs them
%   M3D contains pairs of linked centroids (edges) in 3D space
M3D=zeros(2,3,height(M.Edges));%need to use 'height' for a table! can't use length
i=1;
j=1;
for i=1:height(M.Edges);
    M3D(1,:,i)=cent(M.Edges.EndNodes(i,1),:);
    M3D(2,:,i)=cent(M.Edges.EndNodes(i,2),:);
end
%   graph
figure ('Name', 'MST on Centroids')
hold on
i=1;
for i=1:height(M.Edges);
    scatter3(M3D(:,1,i),M3D(:,2,i),M3D(:,3,i));
    line(M3D(:,1,i),M3D(:,2,i),M3D(:,3,i));
end
xlabel aSMA
ylabel CNN1
zlabel MHC

%4. Principle Component Analysis
%   Change Visualization in 2D space, with heatmaps for each marker
%   PCA
[COEFF, SCORES, latent, Tsquared] = pca(A);
    %COEFF (pxp),SCORES (nxp)= representation of A in the PC space (rows: observ, cols: PCs)
    %latent (n)= eigenvalues of the covariance matrix of A, Tsquared (p)= measure of multivariate
distance of each observ to the center of data set
varContrib=cumsum(latent)./sum(latent); %shows how much variance is accounted for by each PC

%   bring cluster centroids to PC space
i=1;%PC saves centroid links (edges) in PC space
PC=zeros(2,2,height(M.Edges)); %transform M3D into PC space, where Column1=PC1, Column2=PC2;
for i=1:height(M.Edges);
    PC(1,1,i)=COEFF(1,1)*M3D(1,1,i)+COEFF(2,1)*M3D(1,2,i)+COEFF(3,1)*M3D(1,3,i);
    PC(1,2,i)=COEFF(1,2)*M3D(1,1,i)+COEFF(2,2)*M3D(1,2,i)+COEFF(3,2)*M3D(1,3,i);
    PC(2,1,i)=COEFF(1,1)*M3D(2,1,i)+COEFF(2,1)*M3D(2,2,i)+COEFF(3,1)*M3D(2,3,i);
    PC(2,2,i)=COEFF(1,2)*M3D(2,1,i)+COEFF(2,2)*M3D(2,2,i)+COEFF(3,2)*M3D(2,3,i);
end
%  plot centroids and links (edges)
i=1;
figure ('Name', 'MST in PC Space')
hold on
for i=1:height(M.Edges)
    plot(PC(:,1,i),PC(:,2,i));
    plot(PC(:,1,i),PC(:,2,i), 'X', 'MarkerSize', 12);
end
%superimpose clusters
%   color clusters with heatmap of markers
%   meanExp: matrix containing average expression of markers
meanExp=zeros(max(T),3);
for i=1:max(T);
    meanExp(i,1)=mean(A(T==i,1));%aSMA
    meanExp(i,2)=mean(A(T==i,2));%CNN1
    meanExp(i,3)=mean(A(T==i,3));%MHC
end

%MATCH COLORS TO EACH CLUSTER!!! Heatmap style
%   normalize marker expression from 0 to 1 (for clusters) to use rgb format
normS =(meanExp(:,1)-min(meanExp(:,1)))/(max(meanExp(:,1))-min(meanExp(:,1)));%aSMA
normC=(meanExp(:,2)-min(meanExp(:,2)))/(max(meanExp(:,2))-min(meanExp(:,2)));%CNN1
normM=(meanExp(:,3)-min(meanExp(:,3)))/(max(meanExp(:,3))-min(meanExp(:,3)));%aSMA

%   transform cent (matrix containing centroids in 3D space) to PC space
i=1;
centPC=zeros(length(cent),2);
for i=1:length(cent)
    centPC(i,1)=COEFF(1,1)*cent(i,1)+COEFF(2,1)*cent(i,2)+COEFF(3,1)*cent(i,3);
    centPC(i,2)=COEFF(1,2)*cent(i,1)+COEFF(2,2)*cent(i,2)+COEFF(3,2)*cent(i,3);
end

%   graph the aSMA heatmap cluster graph
figure ('Name', 'aSMA Expression');
i=1;
```

```
hold on
for i=1:height(M.Edges)
    scatter(SCORES(T==i,1),SCORES(T==i,2), 40,[1-normS(i) 0 normS(i)],'filled');%plot cluster
    plot(PC(:,1,i),PC(:,2,i),'k');%plot edges in black
    scatter(centPC(i,1),centPC(i,2), 60, [1-normS(i) 0 normS(i)], 'd');%plot centroids
end
xlabel PC1
ylabel PC2

%   graph the CNN-1 heatmap cluster graph
figure ('Name', 'CNN-1 Expression');
i=1;
hold on
for i=1:height(M.Edges)
    scatter(SCORES(T==i,1),SCORES(T==i,2), 40,[1-normC(i) 0 normC(i)],'filled');%plot cluster
    plot(PC(:,1,i),PC(:,2,i),'k');%plot edges in black
    scatter(centPC(i,1),centPC(i,2), 60, [1-normC(i) 0 normC(i)], 'd');%plot centroids
end
xlabel PC1
ylabel PC2

%   graph the MHC heatmap cluster graph
figure ('Name', 'MHC Expression');
i=1;
hold on
for i=1:height(M.Edges)
    scatter(SCORES(T==i,1),SCORES(T==i,2), 40,[1-normM(i) 0.5 normM(i)],'filled');%plot cluster
    plot(PC(:,1,i),PC(:,2,i),'k');%plot edges in black
    scatter(centPC(i,1),centPC(i,2), 60, [1-normM(i) 0.5 normM(i)], 'd', 'LineWidth',1.5);%plot
centroids
end
xlabel PC1
ylabel PC2

figure
scatter(SCORES(:,1),SCORES(:,2))

end
```