8292–8301 *Nucleic Acids Research, 2016, Vol. 44, No. 17* *Published online 16 August 2016*
doi: 10.1093/nar/gkw700

Selective single cell isolation for genomics using microraft arrays

Joshua D. Welch[1,2,†], Lindsay A. Williams[3,4,†], Matthew DiSalvo[5,†], Alicia T. Brandt[6], Raoud Marayati[3], Christopher E. Sims[7], Nancy L. Allbritton[3,5,7,8], Jan F. Prins[1,2], Jen Jen Yeh[3,8,9,*] and Corbin D. Jones[1,6,9,*]

[1]Curriculum in Bioinformatics and Computational Biology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA, [2]Department of Computer Science, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA, [3]iBGS—Integrative Program for Biological & Genome Sciences,3356 Genome Sciences Bldg, CB #7100 Chapel Hill, NC 27599-7100, USA, [4]Department of Epidemiology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA, [5]Joint Biomedical Engineering Program, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA, [6]Department of Biology, Campus Box 3280, Coker Hall, UNC-Chapel Hill, Chapel Hill, NC 27599-3280, USA, [7]Department of Chemistry, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA, [8]Department of Pharmacology, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA and [9]Department of Surgery, The University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3280, USA

Received November 6, 2015; Revised July 28, 2016; Accepted July 29, 2016

ABSTRACT

Genomic methods are used increasingly to interrogate the individual cells that compose specific tissues. However, current methods for single cell isolation struggle to phenotypically differentiate specific cells in a heterogeneous population and rely primarily on the use of fluorescent markers. Many cellular phenotypes of interest are too complex to be measured by this approach, making it difficult to connect genotype and phenotype at the level of individual cells. Here we demonstrate that microraft arrays, which are arrays containing thousands of individual cell culture sites, can be used to select single cells based on a variety of phenotypes, such as cell surface markers, cell proliferation and drug response. We then show that a common genomic procedure, RNA-seq, can be readily adapted to the single cells isolated from these rafts. We show that data generated using microrafts and our modified RNA-seq protocol compared favorably with the Fluidigm C1. We then used microraft arrays to select pancreatic cancer cells that proliferate in spite of cytotoxic drug treatment. Our single cell RNA-seq data identified several expected and novel gene expression changes associated with early drug resistance.

INTRODUCTION

A fundamental problem in modern biology is identifying genetic and genomic characteristics that determine the functional or phenotypic properties of individual tissues and cells in a multicellular organism. New genomics techniques, such as RNA-seq, ATAC-seq and Hi-C, have revealed hidden details about how the genome is organized and how that organization shapes gene expression to produce phenotypes. These high-throughput techniques are indispensable tools, but they are most commonly performed on bulk tissue samples containing millions of cells. Such bulk analyses inherently blur the properties of individual cells within a tissue (e.g. (1)). An aggregate view may hide strong heterogeneity among cells within tissue, mask the effects of small, phenotypically distinct subpopulations of cells and drive a false impression of similarity across tissues.

Targeting and genomic characterization of individual cells within a tissue resolves this problem and facilitates connecting genotype and phenotype at the level of individual cells. Recently, several microfluidic methods have been developed to enable isolation of dozens to tens of thousands of cells at once (2–5). The Fluidigm C1, for example, is a widely used microfluidic single cell sorting system that performs cell lysis, RNA isolation, and cDNA creation for 96 cells at once on a single chip (6). The C1 offers automated single cell isolation, but is unable to select specific cell types from a heterogeneous population, requiring the user to load a pre-selected set of cells. Pre-selection based on fluorescent

*To whom correspondence should be addressed. Tel: +1 919 962 4443; Fax: +1 919 962 1625; Email: cdjones@email.unc.edu
Correspondence may also be addressed to Jen Jen Yeh. Tel: +1 919 962 4443; Fax: +1 919 962 1625; Email: jen_jen_yeh@med.unc.edu
†These authors contributed equally to the paper as first authors.

markers can be performed using flow cytometry or similar approaches, but once cells enter the C1 chip, the user cannot determine which 96 cells will be captured from their starting pool.

In addition, even if a heterogeneous population of cells is pre-sorted based on fluorescence, many cellular phenotypes of interest are too complex to be captured by fluorescent markers. These approaches cannot capture many important cellular characteristics that can be measured only as 'complex' phenotypes. Complex phenotypes can involve a temporal component, such as proliferation, cell mobility, extracellular matrix invasion and drug resistance that cannot be characterized by fluorescent markers. This inability to select cells based on temporally or spatially varying phenotypes limits the ability of existing single cell capture technologies to fully define specific individual cell types and increases the risk that heterologous cells will be treated as a single population.

We have developed a novel protocol for single cell isolation and genomic analysis to address these limitations and enable the linking of genotype to phenotype at the individual cell level. Our method allows for selection of individual cells from a heterogeneous population based on complex phenotypes including cell surface markers, cell proliferation and drug response. This enables genomic characterization at the single cell level by allowing the measurement of cellular phenotypes before cell isolation. We illustrate this approach by performing single cell RNA-seq on individual cells that were selected for specific phenotypes from a heterogeneous population of cells. We focused on RNA-seq as it is particularly susceptible to the problems of bulk tissue analysis, it is currently one of the most commonly used single cell approaches and it is most readily comparable to the C1 technology (1,6).

MATERIALS AND METHODS

Cell line and culture conditions

CFPAC-1 pancreatic cancer cells were purchased from American Type Culture Collection (Manassas, VA, USA) and were used for all experiments. They were cultured in RPMI plus 10% fetal bovine solution and 1× penicillin/streptomycin. Prior to use for the sequencing only experiments, CFPAC-1 cells were infected with mCherry lentivirus and flow cytometry sorted to enrich for the cells that highly express mCherry.

C1 single cell isolation and sample preparation for sequencing

C1 selection of single mCherry CFPAC-1 cells was performed according to the manufacturer's recommended protocol using a starting cell suspension of 10 000 cells/ml. Following single cell isolation, the C1 chip was visualized under a confocal microscope using the Texas Red (546) filter to identify nests on the chip containing no cells, single cells and two or more cells. Following visualization, cDNA was created on the C1 chip using the manufacturer protocol and the ClonTech SMARTer Ultra Low Input kit. Next, RT-PCR was performed on a selection of five single cells, five double cells and five empty wells from the C1 chip using GAPDH (Hs02758991_g1) and EpCAM (hs00901888_g1) to ensure proper samples were selected for subsequent analyses. Once confirmed, the RNA-sequencing libraries were created using the manufacturer's protocol with modifications specific to the C1 for lower input with the Nextera XT library prep kit.

C1 single cell isolation of healthy mouse blood spiked with mCherry CFPAC-1 cells

About 2 ml of healthy mouse blood was spiked with 2000 mCherry CFPAC-1 cells. Ficoll separation and mouse lineage and CD45 depletion was performed (Miltenyi Mouse Linage Cell Depletion Kit and Mouse CD45 Microbeads) using the MACS Separation Columns and MACS magnet. The depleted fraction was loaded onto the C1 per the manufacturer's protocol. C1 chip was imaged as described above and Specific Target Amplification was used with the following TaqMan Probes (Life Technologies): (all CD probes are mouse specific) Cd19 (Mm0515420_m1), Cd11b (Mm00434455_m1), Cd2 (Mm00488928_m1), Cd41 (Mm00439741_m1), Cd123 (Mm00434273_m1), Cd45 (Mm01293577_m1), Cd235a (Mm00494848_m1), Cd62p (Mm0129531_m1), GAPDH (Hs02758991_g1), ACTIN B (Hs01060665_g1), EpCAM (human) (Hs00901888_g1). RT-PCR was performed on a subset of samples using 40 cycles.

Microfabrication of microraft arrays

Microraft arrays containing 44 100 microraft elements that were each 100 μm square, 20 μm deep and 15 μm apart were fabricated as reported previously (7,8). Briefly, standard UV photolithography was employed to create array templates in negative photoresist on glass. Polydimethylsiloxane (PDMS) replica molding from these templates produced PDMS microwell arrays, which were dip-coated in a solution of magnetic polystyrene. Due to discontinuous de-wetting, a bead of polystyrene remained within each microwell and formed a microraft after baking off the solvent. The completed arrays were adhered to a polycarbonate cassette, oxygen plasma treated for 5 min, sterilized with ethanol, rinsed three times with 1× sterile phosphate buffered saline (PBS), and then coated with a thin layer of glucose to enhance surface wettability.

Microraft array single cell isolation similar to C1 isolation

Cell isolation was carried out similarly to that described in prior publications (7,8) (see also, http://www.cellmicrosystems.com/product-applications/). To isolate mCherry CFPAC-1 cells from the microraft array to create a batch of cells as similar to the C1 cells as possible, ∼10 000 cells were added to the array and incubated in 1 ml RMPI containing 10% Fetal Bovine Serum (FBS), 1× penicillin/streptomycin and 0.5 mg/ml geneticin for 24 h to allow cells to adhere. The array was inspected using an epifluorescence microscope (Nikon Instruments) outfitted with a needle-based release device mounted on the 4× objective (7,8). As each microraft containing single, double or no cells was randomly identified, it was released from

the array using the needle device, collected using a magnetic wand, rinsed once with sterile 1× PBS and then deposited into a polymerase chain reaction (PCR) tube containing 4 μl of sterile 1× PBS as the collection media, in which the cells were lysed without stripping them from the microraft culture surface. The entire release and collection process required 15–20 s to perform per sample, and sample tubes were immediately placed on dry ice and stored at −80°C. The five single, five double, five empty microrafts, two samples of ∼1000 mCherry CFPAC-1 cells, 200 pg of the SMARTer kit positive control and a water sample were processed for isolation and amplification of the mRNA using the ClonTech SMARTer Ultra Low Input Kit for Sequencing-V3 kit per the manufacturer's recommendations. The resulting cDNA was analyzed with RT-PCR using GAPDH (Hs02758991_g1) and EpCAM (hs00901888_g1) probes to determine concentration alongside a standard curve of CFPAC-1 cDNA. Libraries were created using the C1 recommended Nextera XT kit with the C1 modifications for low concentration reactions.

Microraft array single cell isolation with RNase out and ERCC spike-ins

A second set of cells was isolated from the microraft arrays in the same manner as described above, with the exception that RNase Out (Invitrogen) was added to the 1× PBS collection buffer at the recommended concentration of 2 units/μl. External RNA Controls Consortium (ERCC) controls (9) were added to each sample at a dilution of 1:10 000 prior to RNA extraction. This dilution was selected based on RT-PCR using probes for ERCC 0130 (Ac03459943_a1) and ERCC 00170 (Ac03460062_a1) at various dilutions. The five selected single, five double, five empty microrafts, two samples of ∼1000 mCherry CFPAC-1 cells, 200 pg of the SMARTer kit positive control and a water sample were processed for mRNA isolation and amplification using the ClonTech SMARTer Ultra Low Input Kit for Sequencing-V3 kit. RT-PCR was performed to determine concentration of the cDNA products as described above. Libraries were then created using the standard Nextera XT library prep kit per ClonTech's recommendations.

Gemcitabine treatment and phenotypic selection for single cells on microarrays

CFPAC-1 cells were stained with 10 μM CellTrace Far Red (DDAO-SE) (Invitrogen C34553) for 30 min at 37°C prior to seeding. Four microraft arrays were seeded with ∼10 000 cells each and incubated overnight in standard 10% FBS and 1× penicillin/streptomycin RPMI media containing no additional treatment, 0.9% saline (as vehicle), 2 nM gemcitabine and 5 nM gemcitabine. The arrays were imaged daily on an Olympus IX81 microscope using custom image acquisition software written using MATLAB and Micro-Manager (Vale Lab, UCSF) (10). This imaging step took place within a humidified, 37°C, 5% CO_2 stage incubator on the microscope. Each day's imaging run required approximately 30 min to acquire images of all 44 100 microrafts on the array (15 minute focus scan + 5 min brightfield + 5 min fluorescence channel 1 + 5 min fluorescence channel 2). After 72 h of culture, the cells were stained for EpCAM by incubating them in a 1:50 dilution of PE conjugated anti-CD326 (Life Technologies A15782) for 30 min at room temperature. In total, 31 proliferative and 31 non-proliferative colonies were identified manually by their specific location on the array using the arrays' brightfield, CellTrace and anti-CD326 time-lapse image data. Proliferative colonies were defined as colonies that grew to three or more cells starting from a single cell, while non-proliferative colonies were defined as those that remained as single cells for the entire 72 h culture period (see Supplementary Figure S6). Specific microrafts of interest were isolated into 1× PBS plus RNase Out (2 units/μl) as described above. The entire collection procedure required ∼30 minutes and was performed at room temperature and atmosphere. RT-PCR was performed on all of the microraft samples prior to RNA sequencing in order to determine adequacy of RNA content for sequencing. RNA was extracted using the SMARTer kit and ERCC spike-ins were used as previously described.

RNA-seq data analysis

RNA-seq reads were aligned to hg19 using MapSplice (11) v. 2.1.4 and to the ERCC reference transcripts (9) using bowtie2 (12) v. 2.2.5. Genomic read distribution and gene body coverage plots were generated using RSeQC (13) v. 2.4 and UCSC hg19 known gene annotations. RSEM (14) v. 1.2.8 was run with UCSC hg19 known gene annotations to calculate gene expression levels and with the flag—estimate-rspd. Only samples with at least 1500 genes detected (detection defined as at least 1 fragment per kilobase (FPKM)) were used in subsequent analyses. Gene set enrichment analysis of the C1 and microraft cells was performed using GSEA (15) v. 2.2.0 run with default parameters and the Gene Ontology All gene set database (c5.all.v5.0.symbols). A one-tailed, homoscedastic *t*-test at $P = 0.05$ was used to determine differential expression; only genes with a mean expression level of at least 10 FPKMs in both proliferating and non-proliferating cells were tested. Enriched gene ontology terms were identified using DAVID (16) in background list mode. Pearson correlations were computed using the cor function in R. Heatmaps were generated using the HeatMapImage module (v. 6) of the GenePattern (17) suite. We used the quasi-likelihood model in edgeR (glmQLFTest, 18) to perform differential gene expression analysis of bulk versus single cell data.

RNA sequencing

All three sets of samples for each isolation (0, 1 and 2-cell nests or microrafts) were pooled and run on a single lane of a MiSeq for 50 pb paired end sequencing. Five MiSeq runs were completed for samples described above.

The flowcell and a reagent cartridge were loaded according to Illumina specification on a MiSeq. Flowcells produced a median of 34 million reads each, with a median of 98% of reads passing filter and median Q-score of 37. Only 2.1% of reads had ambiguous barcodes. Given that some samples were empty, the balance between barcodes was close to expectation (5.26% each) with a median coefficient of variation of 0.44. Raw data has been deposited in GEO (GSE85183).

RESULTS

Overview of approach

We developed a novel method for single cell selection, isolation and sequencing based on microraft arrays (Figure 1, Supplementary Figure S1 and Supplementary File 5) (7,8). These arrays consist of thousands of individual square magnetic polystyrene cell culture sites that are each contained within elastomeric microwells. Cells are plated on the microraft array as a dilute suspension and settle onto individual microrafts. Once these cells have adhered, individual microrafts chosen by the investigator can be dislodged from the underlying array using a motorized microneedle that pierces the array from below. Dislodged microraft cell carriers can then be magnetically manipulated into separate sample containers where the attached cells are lysed without being stripped from the microraft carrier.

Microraft arrays enable investigators to selectively isolate single cells of interest, even from small numbers of cells, based on a rich variety of cellular phenotypes, such as cellular morphology, proliferation and cell surface marker expression levels. Individual cells on microrafts are isolated gently without the need to strip the cells from their substrate or perform microfluidic sorting. Furthermore, microraft arrays present an ideal environment for *in vitro* culture, enabling examination of cell morphology and growth properties over extended periods of time or in response to treatment/conditions. Cell isolation performed using microraft arrays yields highly viable cells that can be assayed or maintained in culture for other downstream applications.

Comparison of microrafts and the Fluidigm C1

We compared RNA sequencing on microraft arrays and the Fluidigm C1 by isolating and sequencing pancreatic cancer cells using both systems. We recognize that the two systems utilized different techniques to isolate and sequence single cells and that there may be systematic technical variation between the results from each system. It is also possible that the process of cell isolation changes gene expression levels. We investigated these effects by sequencing samples containing 0, 1 and 2 cells from the same CFPAC-1 cell line using both the C1 and microrafts (Figure 1A and Supplementary Figure S2). Overall, both approaches produced data of high quality, but we observed several notable differences. We compared the read coverage (Figure 2A), 3′ bias (Figure 2B), genomic read distribution (Figure 2C) and number of genes detected (Figure 2D) in the resulting RNA-seq data. The proportion of reads aligning to the genome, 3′ bias and genomic read distribution for the C1 and microraft data are highly similar. We noticed that the number of genes detected shows more variation for single microraft cells than single C1 cells, and the median number of genes detected using microrafts is smaller (Figure 2D). Part of this difference reflects a difference in the multiplexing used for sequencing the two sets of samples as we were able to multiplex more microraft samples per run than the C1. To show the relationship between number of genes detected and sequencing depth, we plotted genes detected both as a function of number of aligned reads, as well as genes detected per 100 000 reads (Supplementary Figure S3). We also speculate that this difference in genes detected could be due to the greater efficiency of reagent delivery in the microfluidic system, which allows the tiny amount of starting RNA to be concentrated in a microscale reaction volume. This difference was *much smaller* when comparing samples containing two cells (Figure 2D), likely because of the larger amount of starting RNA generated by two cells.

Comparing the C1 empty samples with the microraft empty samples shows that the C1 empty samples contain a much higher proportion of reads mapping to the human genome (Figure 2E). These mapped reads from empty C1 samples are strongly enriched for coding regions of the genome, indicating that these reads come from RNA, not DNA contamination (Figure 2E). The C1 empty samples show a much higher proportion of reads mapping to the 3′ UTR (Figure 2E), suggesting that the reads may originate from broken or degraded transcripts. These reads may have originated from cells that lysed after being deposited in the C1, resulting in free-floating RNA in the cell suspension that subsequently was carried within the microfluidics. We also performed qPCR on cells isolated on the C1 from a starting cell suspension mixture of mouse blood cells and mCherry CFPAC-1 cells, which were identified on the C1 chip under the Texas Red filter. We found mouse RNA in both empty nests and nests containing human cells (see Supplementary Table S1). Such cross-contamination was also observed in a similar recent evaluation of the Fluidigm C1 (2). In contrast, we observed few reads in the empty microraft samples (Figure 2E), likely because we rinsed and deposited the microrafts into separate reaction wells before RNA extraction, removing any contaminating RNA.

It is difficult to assess how much contamination between wells on the C1 affects single cell RNA-seq results for two reasons. First, in the examples of contamination that we show in Supplementary Table S1, the CT values for the contaminating RNA present in the empty wells are comparable to the CT values from wells containing single cells, indicating that the amount of contaminating RNA is significant. Also, although computational filtering can be used to remove contaminating reads from another species, as in the example of mouse reads contaminating human RNA-seq data, there is no way of computationally disambiguating RNA from two cells with the same genome. Thus, the real problem is potential mixing of RNA from different cells of the same species, blurring the identities of the individual cells under consideration.

To assess the technical variability of the RNA-seq data generated from microrafts, we added ERCC spike-in transcripts to a subset of the cells we sequenced (total of 55 cells with spike-ins). The spike-in transcripts were added manually to each individual cell after isolation; in contrast, spike-ins are loaded onto the C1 along with other reagents and then dispersed to all cells on the chip through the microfluidic system. The sequencing data that we generated from the Fluidigm C1 did not contain spike-in transcripts, so we used previously published Fluidigm C1 data for comparison (19,20). Using these data, we found similar levels of technical variation in measured expression and also similar correlation between the true and measured expression levels of the spike-in transcripts (Supplementary Figure S4). Thus the manual pipetting of spike-ins does not introduce

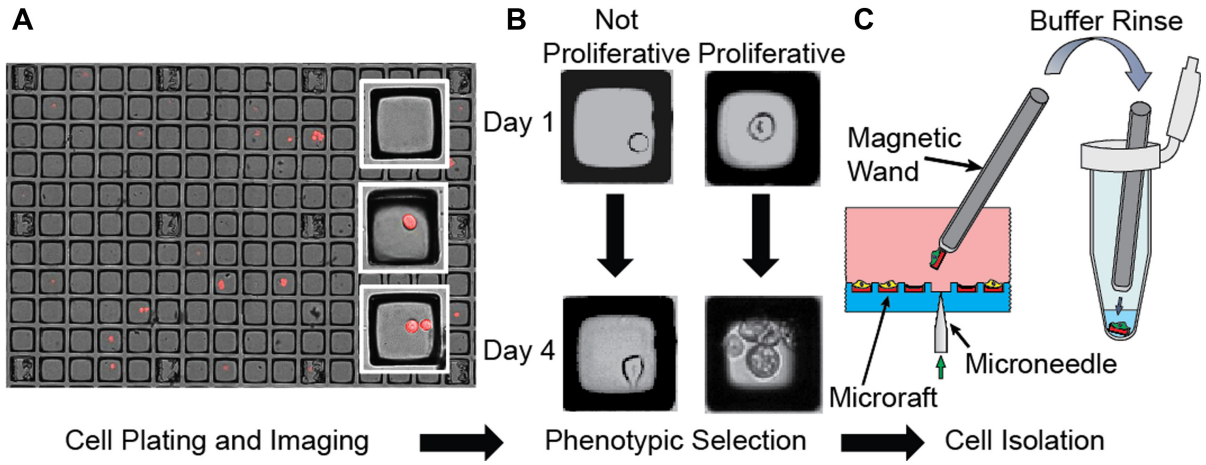

Figure 1. Selective cell isolation using microraft arrays. (**A**) Composite fluorescence and bright-field microscopy of mCherry cells on microraft arrays. (**B**) Selection of cells based on a phenotype such as proliferation. (**C**) A microneedle dislodges microraft cell-carriers, which are manipulated through multiple buffer washes and into lysis buffer using a magnetic wand.

significant levels of additional technical variation. Because we could manually add spike-ins to each cell after isolation, we used two separate spike-in mixes with different concentrations of the individual spike-in transcripts. This experiment is not possible on the Fluidigm C1, because spike-ins for all cells are loaded in a single batch. We used this information to show that the spike-in transcripts with different expression levels between the two mixes show the expected differential expression across the set of cells that we studied (Supplementary Figure S4).

We computed correlations between the gene expression profiles of cells and found that pairs of samples sequenced using the same platform showed strong concordance, indicating that both the microrafts and the C1 produce reproducible results across replicates (Figure 2F and Supplementary Figure S5). However, the correlations between pairs of samples from different platforms showed lower agreement (Figure 2F and Supplementary Figure S5). In computing these correlations, we restricted the set of genes to those with an average expression level of at least 1 FPKM across both C1 and microraft samples to ensure that the effect was not simply due to differences in number of genes detected.

To further investigate the differences between C1 and microraft samples, we performed gene set enrichment analysis (GSEA, 15) and found several gene sets with significantly higher transcript abundance (FDR: false discovery rate $<$ 10%) in the microraft samples (Supplementary File 1). No gene sets show significant upregulation in the C1 samples after FDR correction (10% FDR), but several have significant nominal P-values ($P < 0.05$), which is suggestive of differences rather than conclusive and may reflect the fact there is less power for the C1 comparison because there are fewer C1 samples than microraft samples. An interesting theme emerges from the gene sets that show differences between the two methods. Among the top gene sets upregulated in the raft cells are electron transport, mitochondrial protein, ribosome, nucleoside kinase, chromosome condensation (Figure 2G) and mitosis. All of these gene sets are required for active cell growth and division. Among the top gene sets showing up-regulation in the C1 cells, in contrast, are cell junction (Figure 2H), cell adhesion and cell migration.

We confirmed the reliability of the microraft platform for single cell RNA-seq with two additional analyses that control for artifacts unique to the microraft system. First, to control for possible variations in the microraft manufacturing process, we compared RNA-seq results from cells grown on separate microraft arrays under identical conditions. Pairwise correlations among the cells showed strong concordance (Figure 3 and Supplementary Figure S6), indicating that systematic differences between microraft arrays are not a significant source of experimental variation. We also examined RNA-seq data from cells grown on the microrafts and RNA-seq from bulk grown cells whose RNA was pooled and split into single cell volumes before sequencing. Because the growing conditions are slightly different and the RNA from bulk grown cells was pooled before sequencing, we expect the gene expression profiles to be quite similar but not identical. Most genes appeared highly similar (Pearson $\rho = 0.71$, Spearman $\rho = 0.70$), with differences possibly caused by the slightly different growing conditions for the bulk and single cells. Differential expression analysis did, however, reveal 588 differentially expressed genes (DEGs) (2% of expressed genes), with only a handful of GO terms in common among genes upregulated in bulk versus microraft cells (Supplementary File 2). The only noticeable commonality among GO terms was translation, with terms including 'ribosome', 'ribosomal protein' and 'translational elongation' appearing on the list (Supplementary File 2). These data show that for the CFPAC-1 cells characterized here in bulk and as single cells are largely similar, but that single cell resolution can identify subtle differences between patterns inferred from bulk data versus single cells.

Selective cell isolation based on proliferation after gemcitabine treatment

To show how microrafts may be used to study and select different cellular phenotypes within a population of

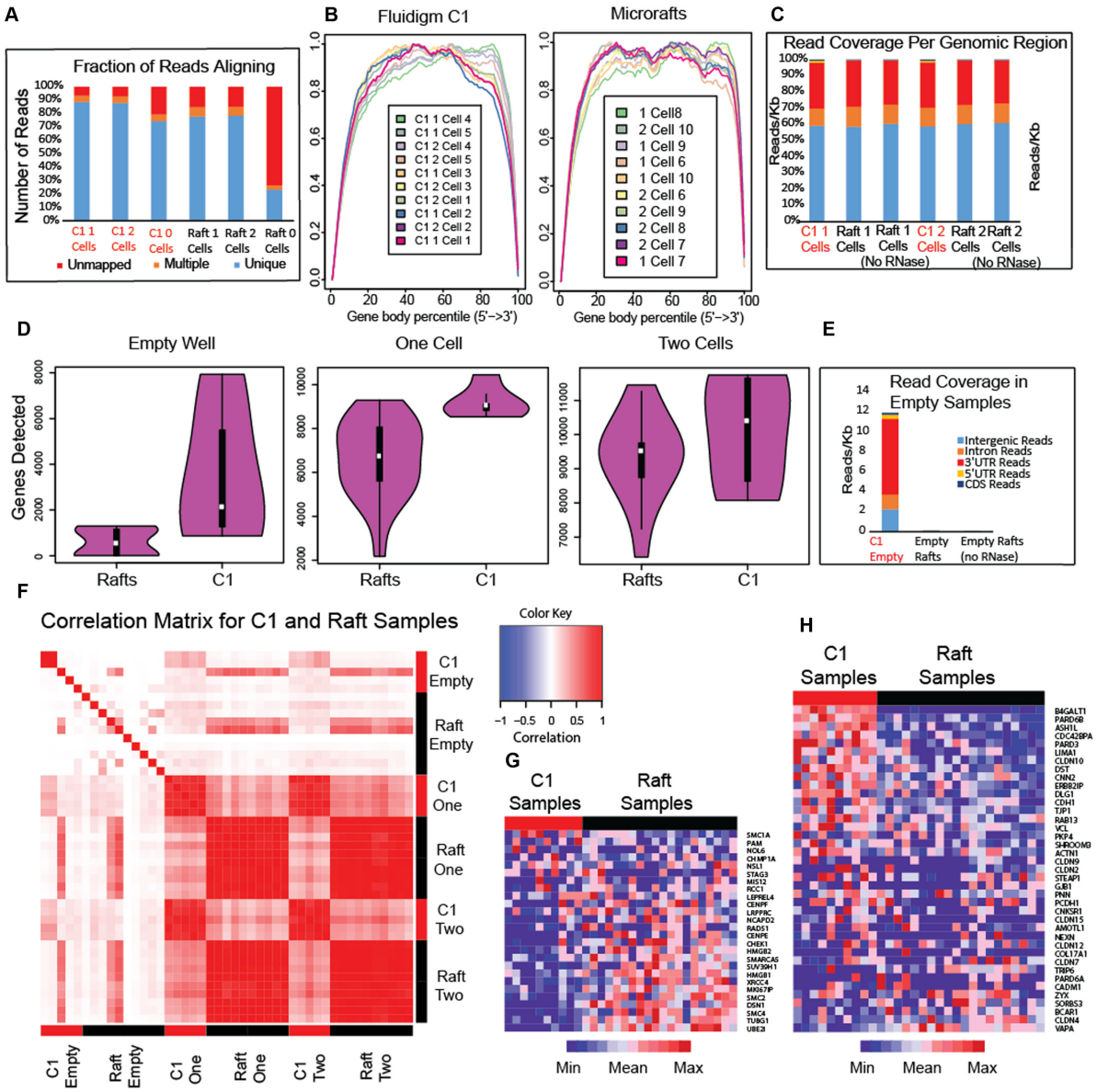

Figure 2. Comparing RNA-seq from cells isolated with microrafts and the Fluidigm C1. (**A**) Fraction of reads aligning to the human genome. (**B**) Gene body coverage plots measuring 3′ bias in read coverage. (**C**) Distribution of read coverage by genomic region. (**D**) Number of genes detected. (**E**) Read coverage for empty samples. (**F**) Heatmap of Pearson correlations among C1 and microraft cells. (**G**) Heatmap of chromosome condensation gene set, which shows strong upregulation in raft cells. (**H**) Heatmap of cell junction gene set, which shows slight upregulation in C1 cells.

cells, we performed a pilot study to evaluate the differences in proliferative capacity of single cancer cells in response to chemotherapy treatment. Proliferation is an important property of tumor cells that cannot be used as a basis for sorting using the C1 system. CFPAC-1 cells were plated onto a microraft array, treated with 2 or 5 nM gemcitabine and allowed to grow for 4 days. Control cells were plated and subjected to a placebo treatment and no treatment, respectively. We also stained the cells in this experiment for EpCAM, which illustrates the compatibility of microrafts and cell staining (Supplementary Figure S7). Tracking the cells on the array during the growth period showed that the placebo treatment had essentially no effect on cell growth, but the 5 nM gemcitabine treatment strongly restricted the growth of many cells compared to the control arrays (Figure 4). The cells treated with 2 nM gemcitabine showed only a moderate restriction of cell growth, so we used only the data from cells treated with 5 nM gemcitabine in subsequent analyses. From each microraft array, we selected and harvested (1) cells that divided and (2) cells that did not divide (Figure 5A and C). Note that microrafts with cells that divided contained a small clonal population of 2–4 cells rather than a single cell. These multi-celled samples provided more starting material, making RNA extraction and sequencing likely to succeed. On the other hand, only expression changes that are shared across the two or more cells on the selected microraft will be detectable, but this is

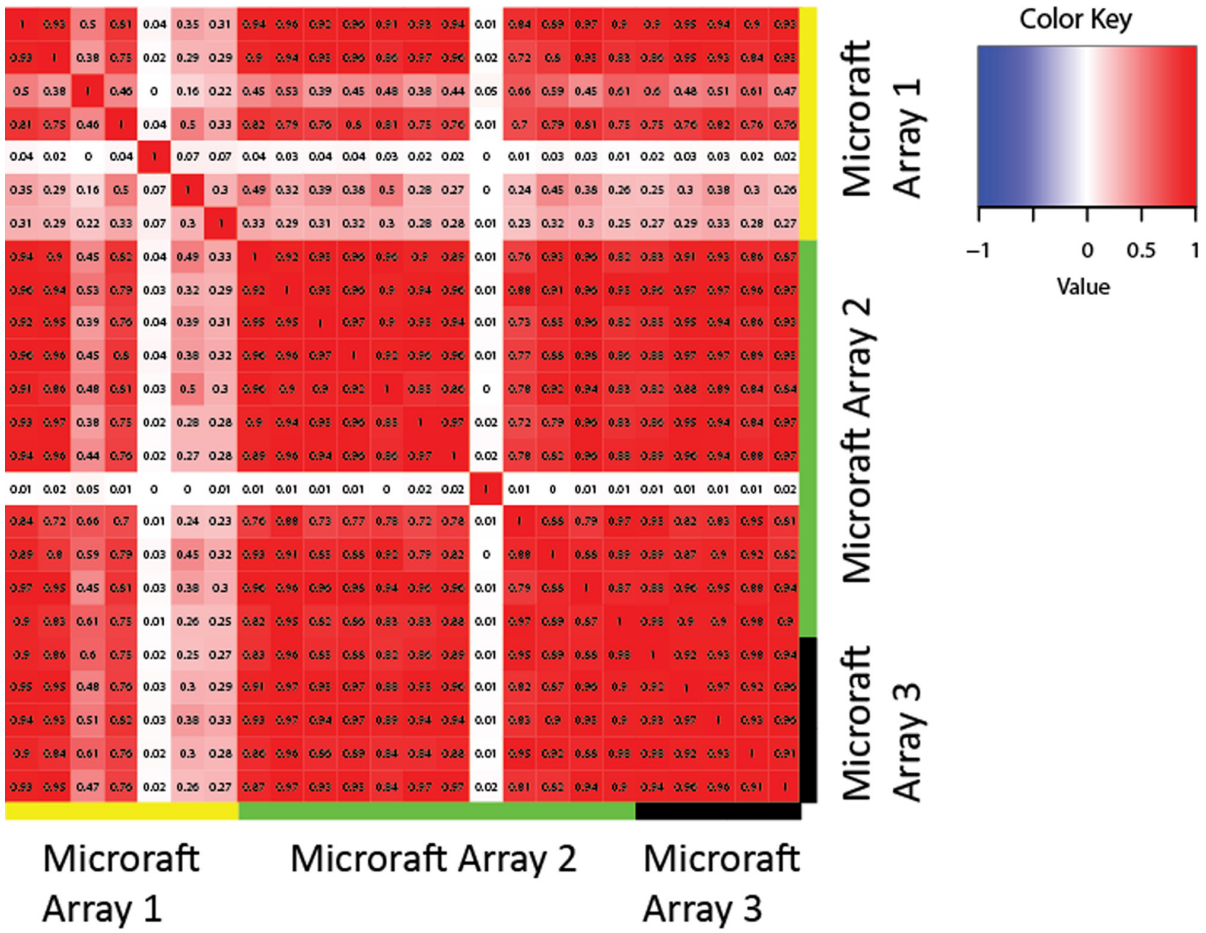

Figure 3. Reproducibility of microraft results. Heatmap shows the correlations among expression profiles of cells grown under identical conditions on three separate microraft arrays. The correlation among cells grown under the same conditions on either the same or different microraft array suggests strong reproducibility, although the correlation with some cells on array 1 is lower. With the exception of two cells, all of the cells on array 1 have <3000 genes detected, which partially reflects the multiplexing of array 1. In contrast, the smallest number of genes detected in any of the other cells shown is 5000. The correlations between cells on arrays 2 and 3 are all very high, although we observe some biological variation, as expected, among individual cells.

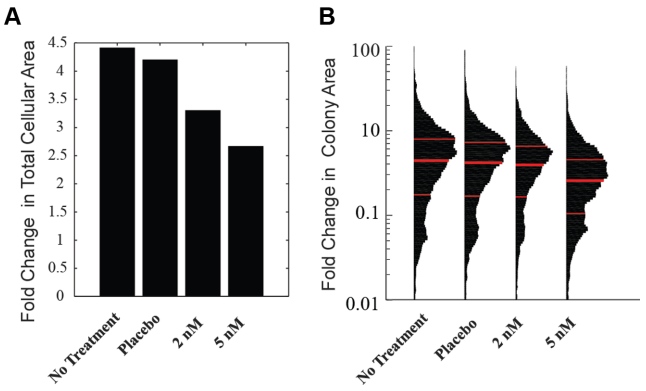

Figure 4. Quantification of CFPAC-1 colony growth. (**A**) Bulk cell measurements. (**B**) Single-colony measurements. Upper, middle and lower red lines indicate the 75th, 50th and 25th percentiles, respectively. The 50th percentile fold change was 4.46, 4.21, 3.95 and 2.58 for cells receiving no treatment, placebo, 2 and 5 nM drug, respectively.

not unreasonable given that each such microraft contains a clonal colony formed from a single cell.

As expected, DEGs involved in cell growth and division were seen in cells that divided compared to non-dividing cells on the control 'no treatment' and placebo treatment arrays (Figure 5B and Supplementary File 3). We searched for DEGs in dividing versus non-dividing cells from the drug treatment array and then removed from these candidate DEGs any genes that were also differentially expressed in the dividing versus non-dividing cells on the control arrays (Figure 5D and Supplementary File 4). We further refined the list by removing DEGs that were from the same gene family as genes differentially expressed in the control experiment (we removed 47 genes, leaving 42), because differences in these genes may be artifacts of read alignments.

Strikingly, roughly 30% of the genes that show specific response to gemcitabine have been previously implicated in tumor progression and chemotherapy resistance (Supplementary File 3). For example, Immediate Early Response Gene X-1 (*IEX-1*) has been implicated in pancreatic cancer cell resistance to gemcitabine treatment (21) and used as a prognostic biomarker in pancreatic cancer (22). Previous studies have linked changes in expression of *IEX-1* to pancreatic cancer patient survival in different settings (23); we

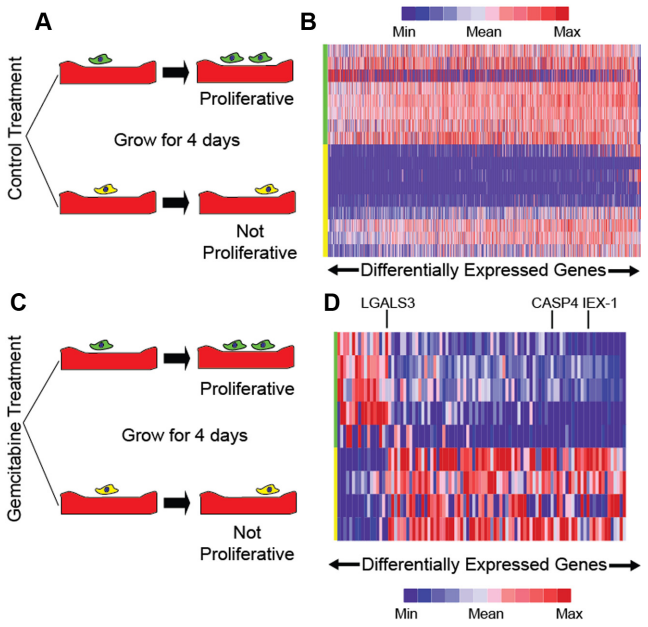

Figure 5. Linking gene expression changes to differences in proliferation after drug treatment. (**A**) Cells on no treatment and placebo treatment microraft arrays are selected based on proliferation. (**B**) Heatmap of genes that change in dividing versus non-dividing cells from the no treatment and placebo treatment microraft arrays. (**C**) Cells on gemcitabine treatment microraft arrays are selected based on proliferation. (**D**) Heatmap of genes differentially expressed in cells that divide after gemcitabine treatment but not control cells.

observed lower expression of *IEX-1* in cells dividing after gemcitabine treatment. Similarly, we observed downregulation of *CASP4* in cells that proliferate after drug treatment. This result is consistent with prior reports of *CASP4* knockdown in pancreatic cells leading to decreased apoptosis after bortezomib treatment (24). We also noted an example of up-regulation after drug treatment; *LGALS3* was previously shown to induce cell proliferation and invasion in pancreatic cancer cells by activating Ras signaling (25). Finally, several of our candidate genes overlapped with those seen in a somewhat similar bulk RNA-seq study looking at gemcitabine resistance (26).

Continued proliferation during gemcitabine treatment is also associated with DEGs functioning in secretion of proteins to the endoplasmic reticulum (ER) and, to a lesser extent, mRNA surveillance and splicing (Supplementary File 3). Changes in ER stress response are known to allow tumors to survive chemotherapy treatment in some cases (27). Gemcitabine is a nucleoside analog that results in increased DNA damage, and response to DNA damage is known to be intertwined with ER stress response (27). We hypothesize that changes in the expression of this set of genes modifies ER stress response, which allows drug-treated cells to handle gemcitabine-induced damage.

Finally, we looked at the *within-group* gene expression differences among the cells from each treatment group, a comparison that highlights a key benefit of single cell resolution. Given the small sample size and relatively low sequencing depth, our power to make conclusive statements is limited. Nevertheless, we observed several interesting patterns. A set of genes showed much greater variation in the gemcitabine-treated cells than the control cells (Supplementary Figure S8). Among such highly variable genes in the drug treated cells, we noticed DNA repair associated genes (*NSMCE1*, *RAD23A*) and/or oncogenes (*RNF103-CHMP3*). This indicates that gemcitabine treatment changed the expression levels of these genes, but possibly only in a subset of the cells, hinting that in this experiment there are multiple mechanisms for continuing to proliferate in the presence of gemcitabine.

DISCUSSION

Microraft arrays provide two key advantages over existing single cell isolation approaches for genomics: (i) selective isolation without the need for a pre-selection step and (ii) the ability to sort cells based on complex temporal or spatial phenotypes, as well as by cellular markers. Here, we reported the application of this cellular isolation system to connect a proliferative pancreatic cancer cell phenotype in the presence of chemotherapeutic to these cells' altered transcriptome. We also showed that microrafts produce data of comparable quality to the C1 but with less contamination between cells during the isolation process.

Microrafts promise to be broadly useful. To date more than 30 cell lines and a number of primary cell types have been successfully cultured on the arrays, which can be coated with any desired extracellular matrix to promote cell health and growth (7,28–31). Many dynamic phenotypes, such as cell mobility, invasiveness, morphology or biomarker expression, are of immense interest in basic research fields such as cell and developmental biology. Microrafts are also useful for isolating rare cells, and in principle can be used to isolate cells from a suspension containing only a few dozen cells of interest. While the current report describes a manual protocol for cell isolation and analysis, the method has the potential to perform high-throughput cell isolation, because a single microraft array can be scaled to contain thousands to millions of microrafts and is amenable to incorporation in an automated platform. Indeed, work is ongoing to develop automated high-throughput instrumentation that will carry out the entire procedure of array scanning, target-cell identification, microraft release, collection and transfer under temperature, humidity and CO_2 controlled conditions at much faster rates than is currently possible. By virtue of indexing on the array, cells in culture can also be used to provide time-series measurements of phenotypic phenomena as was done in the current work or to provide samples for multiple downstream analyses.

The selective ability of microraft arrays can also be exploited for several other genomic approaches beyond RNA-seq. Individual cells or small selected clonal populations from microrafts can be used for Hi-C (32), which measures the 3D organization of the genome; ATAC-seq (33), which measures nucleosome positioning; bisulfite sequencing for DNA methylation (34); or ChIP-seq, for histone modifications (35). Similarly, the selectivity of microraft arrays can be applied to metagenomic and microbiome studies to distinguish, select and sequence members of a complex community.

There are also a number of potential applications for microrafts in clinical research. For example, tumors often exhibit strong intratumoral gene expression and phenotypic heterogeneity. Microrafts could be used to make the link between transcriptomic and functional heterogeneity within tumors. Some functional differences among tumor cells, such as the ability of cells to leave the solid tumor, have important implications for patient outcomes. Microrafts provide an invaluable tool for studying such phenomena by connecting gene expression with phenotypic measurements from individual cells.

DISCLOSURE DECLARATION

N.L.A. and C.E.S. are inventors of microraft arrays (US Patent #9,068155) and have financial interests in Cell Microsystems, Inc.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful for discussions with Erin Osborne, Sophia Tintori and Dan McKay regarding the nuances of single cell genomics. Thanks to Scott Magness and the Advanced Analytics Core of the UNC Center for GI Biology and Disease for assistance with the Fluidigm C1.

FUNDING

NCBC [2013-MRG-1110, in part to C.D.J.]; UCRF (in part to C.D.J.); National Science Foundation (NSF) [ABI/EF0850237 to J.F.P.]; National Institutes of Health (NIH) [R43EB019752 to N.L.A.]; NSF Graduate Research Fellowship [DGE-1144081 to J.D.W.]; NIH BD2K Fellowship [T32 CA201159 to J.D.W.]. Funding for open access charge: NCBC [2013-MRG-1110].

Conflict of interest statement. N.L.A. and C.E.S. are inventors of microraft arrays (US Patent #9,068155) and have financial interests in Cell Microsystems, Inc.

REFERENCES

1. Saliba,A.-E., Westermann,A.J., Gorski,S.A. and Vogel,J. (2014) Single-cell RNA-seq: advances and future challenges. *Nucleic Acids Res.*, **42**, 8845–8860.
2. Macosko,E.Z., Basu,A., Satija,R., Nemesh,J., Shekhar,K., Goldman,M., Tirosh,I., Bialas,A.R., Kamitaki,N., Martersteck,E.M. *et al.* (2015) Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*, **161**, 1202–1214.
3. Klein,A.M., Mazutis,L., Akartuna,I., Tallapragada,N., Veres,A., Li,V., Peshkin,L., Weitz,D.A. and Kirschner,M.W. (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, **161**, 1187–1201.
4. Bose,S., Wan,Z., Carr,A., Rizvi,A.H., Vieira,G., Pe'er,D. and Sims,P.A. (2015) Scalable microfluidics for single cell RNA printing and sequencing. *Genome Biol.*, **16**, 120.
5. Kimmerling,R.J., Lee Szeto,G., Li,J.W., Genshaft,A.S., Kazer,S.W., Payer,K.R., de Riba Borrajo,J., Blainey,P.C., Irvine,D.J., Shalek,A.K. *et al.* (2016) A microfluidic platform enabling single-cell RNA-seq of multigenerational lineages. *Nat. Commun.*, **7**, 10220.
6. Kolodziejczyk,A.A., Kim,J.K., Svensson,V., Marioni,J.C. and Teichmann,S.A. (2015) The technology and biology of single-cell RNA sequencing. *Mol. Cell*, **58**, 610–620.
7. Gach,P.C., Wang,Y., Phillips,C., Sims,C.E. and Allbritton,N.L. (2011) Isolation and manipulation of living adherent cells by micromolded magnetic rafts. *Biomicrofluidics*, **5**, 32002–3200212.
8. Wang,Y., Phillips,C., Xu,W., Pai,J.-H., Dhopeshwarker,R., Sims,C.E. and Allbritton,N.L. (2010) Micromolded arrays for separation of adherent cells. *Lab. Chip.*, **10**, 2917–2924.
9. Jiang,L., Schlesinger,F., Davis,C.A., Zhang,Y., Li,R., Salit,M., Gingeras,T.R. and Oliver,B. (2011) Synthetic spike-in standards for RNA-seq experiments. *Genome Res.*, **21**, 1543–1551.
10. Edelstein,A., Amodaj,N., Hoover,K., Vale,R. and Stuurman,N. (2010) Computer control of microscopes using manager. *Curr. Protoc. Mol. Biol.*, doi:10.1002/0471142727.mb1420s92.
11. Wang,K., Singh,D., Zeng,Z., Coleman,S.J., Huang,Y., Savich,G.L., He,X., Mieczkowski,P., Grimm,S.A., Perou,C.M. *et al.* (2010) MapSplice: accurate mapping of RNA-seq reads for splice junction discovery. *Nucleic Acids Res.*, **38**, e178.
12. Langmead,B. and Salzberg,S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat. Methods*, **9**, 357–359.
13. Wang,L., Wang,S. and Li,W. (2012) RSeQC: quality control of RNA-seq experiments. *Bioinformatics*, **28**, 2184–2185.
14. Li,B. and Dewey,C.N. (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*, **12**, 323.
15. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 15545–15550.
16. Huang,D.W., Sherman,B.T. and Lempicki,R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
17. Reich,M., Liefeld,T., Gould,J., Lerner,J., Tamayo,P. and Mesirov,J.P. (2006) GenePattern 2.0. *Nat. Genet.*, **38**, 500–501.
18. Robinson,M.D., McCarthy,D.J. and Smyth,G.K. (2010) edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
19. Treutlein,B., Brownfield,D.G., Wu,A.R., Neff,N.F., Mantalas,G.L., Espinoza,F.H., Desai,T.J., Krasnow,M.A. and Quake,S.R. (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. *Nature*, **509**, 371–375.
20. Buettner,F., Natarajan,K.N., Casale,F.P., Proserpio,V., Scialdone,A., Theis,F.J., Teichmann,S.A., Marioni,J.C. and Stegle,O. (2015) Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat. Biotechnol.*, **33**, 155–160.
21. Hamidi,T., Algül,H., Cano,C.E., Sandi,M.J., Molejon,M.I., Riemann,M., Calvo,E.L., Lomberk,G., Dagorn,J.-C., Weih,F. *et al.* (2012) Nuclear protein 1 promotes pancreatic cancer development and protects cells from stress by inhibiting apoptosis. *J. Clin. Invest.*, **122**, 2092–2103.
22. Sasada,T., Azuma,K., Hirai,T., Hashida,H., Kanai,M., Yanagawa,T. and Takabayashi,A. (2008) Prognostic significance of the immediate early response gene X-1 (IEX-1) expression in pancreatic cancer. *Ann. Surg. Oncol.*, **15**, 609–617.
23. Wu,M.X., Ustyugova,I. V, Han,L. and Akilov,O.E. (2013) Immediate early response gene X-1, a potential prognostic biomarker in cancers. *Expert Opin. Ther. Targets*, **17**, 593–606.
24. Nawrocki,S.T., Carew,J.S., Pino,M.S., Highshaw,R.A., Dunner,K., Huang,P., Abbruzzese,J.L. and McConkey,D.J. (2005) Bortezomib sensitizes pancreatic cancer cells to endoplasmic reticulum stress-mediated apoptosis. *Cancer Res.*, **65**, 11658–11666.
25. Song,S., Ji,B., Ramachandran,V., Wang,H., Hafley,M., Logsdon,C. and Bresalier,R.S. (2012) Overexpressed galectin-3 in pancreatic cancer induces cell proliferation and invasion by binding Ras and activating Ras signaling. *PLoS One*, **7**, e42699.
26. Shen,Y., Pan,Y., Xu,L., Chen,L., Liu,L., Chen,H., Chen,Z. and Meng,Z. (2015) Identifying microRNA-mRNA regulatory network in gemcitabine-resistant cells derived from human pancreatic cancer cells. *Tumour Biol.*, **36**, 4525–4534.
27. Dicks,N., Gutierrez,K., Michalak,M., Bordignon,V. and Agellon,L.B. (2015) Endoplasmic reticulum stress, genome damage, and cancer. *Front. Oncol.*, **5**, 11.
28. Gracz,A.D., Williamson,I.A., Roche,K.C., Johnston,M.J., Wang,F., Wang,Y., Attayek,P.J., Balowski,J., Liu,X.F., Laurenza,R.J. *et al.*

(2015) A high-throughput platform for stem cell niche co-cultures and downstream gene expression analysis. *Nat. Cell Biol.*, **17**, 340–349.
29. Attayek,P.J., Hunsucker,S.A., Wang,Y., Sims,C.E., Armistead,P.M. and Allbritton,N.L. (2015) Array-based platform to select, release, and capture epstein-barr virus-infected cells based on intercellular adhesion. *Anal. Chem.*, **87**, 12281–12289.
30. Koh,J., Hogue,J.A., Wang,Y., DiSalvo,M., Allbritton,N.L., Shi,Y., Olson,J.A. and Sosa,J.A. (2016) Single-cell functional analysis of parathyroid adenomas reveals distinct classes of calcium sensing behaviour in primary hyperparathyroidism. *J. Cell. Mol. Med.*, **20**, 351–359.
31. Gordon,K.R., Wang,Y., Allbritton,N.L. and Taylor,A.M. (2015) Magnetic alignment of microelements containing cultured neuronal networks for high-throughput screening. *J. Biomol. Screen.*, **20**, 1091–1100.
32. Nagano,T., Lubling,Y., Stevens,T.J., Schoenfelder,S., Yaffe,E., Dean,W., Laue,E.D., Tanay,A. and Fraser,P. (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature*, **502**, 59–64.
33. Buenrostro,J.D., Wu,B., Litzenburger,U.M., Ruff,D., Gonzales,M.L., Snyder,M.P., Chang,H.Y. and Greenleaf,W.J. (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. *Nature*, **523**, 486–490.
34. Smallwood,S.A., Lee,H.J., Angermueller,C., Krueger,F., Saadeh,H., Peat,J., Andrews,S.R., Stegle,O., Reik,W. and Kelsey,G. (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. *Nat. Methods*, **11**, 817–820.
35. Rotem,A., Ram,O., Shoresh,N., Sperling,R.A., Goren,A., Weitz,D.A. and Bernstein,B.E. (2015) Single-cell ChIP-seq reveals cell subpopulations defined by chromatin state. *Nat. Biotechnol.*, **33**, 1165–1172.