ADVANCED MATERIALS TECHNOLOGIES
www.advmattechnol.de

Magnetically Functionalized Hydrogels for High-Throughput Genomic Applications

*Evan Lammertse, Siran Li, Jude Kendall, Catherine Kim, Patrick Morris, Nissim Ranade, Dan Levy, Michael Wigler, and Eric Brouzes**

Single-cell genomics has revolutionized tissue analysis by revealing the genetic program of individual cells. The key aspect of the technology is the use of barcoded beads to unambiguously tag sequences originating from a single cell. The generation of unique barcodes on beads is mainly achieved by split–pooling methods, which are labor-intensive due to repeated washing steps. Toward the automation of the split–pooling method, a simple method to magnetize hydrogel beads is developed. It is shown that these hydrogel beads provide increased yields and washing efficiencies for purification procedures. They are also fully compatible with single-cell sequencing using the BAG-seq workflow. The work opens the automation of the split–pooling technique, which will improve single-cell genomic workflows.

1. Introduction

Single-cell sequencing is commonly used to investigate the genetic heterogeneity of cell populations. Identifying cell subpopulations within tissues can lead to the discovery of new cell types[1] and the deciphering of tumor evolution.[2] Droplet microfluidics has proven to be an invaluable method for single-cell sequencing due to its ability to isolate cells, modular nature, and high throughput.[3,4] Droplet-based single-cell sequencing relies on encapsulating single cells with single barcoded beads in individual droplets.[5,6] The barcoded bead captures the nucleic acids upon cell lysis. The sequences of the captured molecules are then appended to the barcode by either elongation or reverse transcription for DNA or RNA, respectively. The physical linkage, or tagging, assures the identification of the capturing bead, enabling the aggregation of sequences originating from the same single cell.

Barcodes are typically generated by split–pooling, which sequentially and randomly adds a short sequence to the growing barcode on beads.[6,7] These beads can be solid[6] or polyacrylamide-based;[5,7] the latter can be encapsulated without the limitation imposed by Poisson's statistics.[8] Our recent BAG-seq method also employs barcoding by split–pooling.[9] We encapsulate single cells into droplets containing functionalized oligomers, cells are lysed, and droplets are subsequently polymerized into balls of acrylamide gel (BAGs). The polymerizing gel network captures the cellular nucleic acids and oligonucleotides. Notably, the gel porosity allows reagents and polymerases to access nucleic acids, which enables direct split-and-pool barcoding of individual single-cell BAGs.

Barcoding by split–pooling is thus central to single-cell genomics; however, it remains labor-intensive and would greatly benefit from automation because of its repetitive washes and buffer exchanges. The workflow is based on centrifugation, which presents inherent obstacles to its automation, such as bulkiness, complexity, and moving parts. In contrast, magnetic separation is an attractive approach. It is commonly employed in commercial automated nucleic acid purification systems, such as the BioRobot series[10] (Qiagen, Germany), KingFisher System[11] (Thermo-Fisher Scientific, USA), and MagNA Pure series[12] (Roche, Germany). Magnetization of BAGs would enable bulk separation and the automation of the split–pooling method. It would reduce labor costs, increase reproducibility, and improve sample processing throughput.

The formation of hydrogel beads benefits from droplet microfluidics, which offers precise control over gel particle size[13] and shape.[14] Particle homogeneity assures response

E. Lammertse
Department of Biomedical Engineering
Stony Brook University
Stony Brook, NY 11794, USA
S. Li, J. Kendall, C. Kim, P. Morris, N. Ranade, D. Levy, M. Wigler
Cold Spring Harbor Laboratories
Cold Spring Harbor, NY 11724, USA
E. Brouzes
Department of Biomedical Engineering
Stony Brook University
Stony Brook, NY 11794, USA
E-mail: Eric.brouzes@stonybrook.edu
E. Brouzes
Laufer Center for Physical and Quantitative Biology
Stony Brook University
Stony Brook, NY 11794, USA
E. Brouzes
Cancer Center
Stony Brook School of Medicine
Stony Brook, NY 11794, USA
E. Brouzes
Institute for Engineering Driven Medicine
Stony Brook University
Stony Brook, NY 11794, USA

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/admt.202301155

DOI: 10.1002/admt.202301155

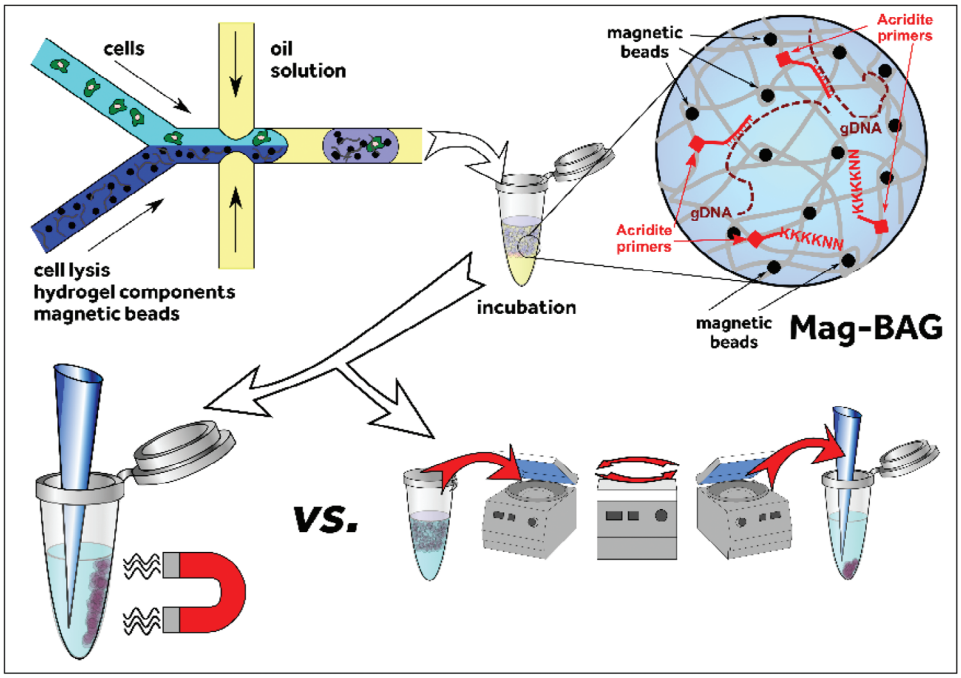

Figure 1. Mag-BAG workflow. A co-flow droplet generator is used to encapsulate the combination of a single-cell suspension with functionalized hydrogel components, lysis buffer, and magnetic beads. After incubation, the hydrogel forms, capturing genomic DNA via Acrydite primers and physically embedding 1 µm magnetic beads. The Mag-BAGs can be separated on a magnetic stand to create and maintain a flat pellet during aspiration. In contrast, centrifugation interrupts the workflow and creates pellets concentrated at the bottom of microtubes and wells.

uniformity to external stimuli such as magnetic fields. This aspect is particularly critical for efficient automation. Few examples of embedding magnetic particles in hydrogel beads using droplet microfluidics exist. They typically encapsulate small magnetic nanoparticles (5–20 nm)[13] or ferrofluids[14,15] in various gel formulations and particle geometries. While basic properties are commonly reported for these magnetic microgels, a comprehensive characterization of their bulk separation performance and a practical application are lacking.

In this work, we produced magnetic BAGs (Mag-BAGs) by physically embedding 1 µm commercially available superparamagnetic microbeads within the gel matrix. We quantified the effect of the magnetic bead and monomer concentrations, BAG size, bead and buffer types, and PCR thermocycling on the bulk separation of Mag-BAGs. We identified key parameters and predictors of efficient magnetic separation. We directly compared the Mag-BAG collection yield with centrifugation after multiple wash cycles in both microtubes and microplates. Finally, we developed and validated Mag-BAG-seq by analyzing a mixture of SKN-1 human fibroblast and SK-BR-3 human breast cancer cell lines via single-cell DNA sequencing. This demonstrates the compatibility of Mag-BAGs with the BAG-seq workflow. Mag-BAGs are compatible with commercial magnetic separation devices and suitable for single-cell encapsulation, genomic library preparation, and single-cell sequencing. This work paves the way for efficient automation of the labor-intensive split–pooling in general and single-cell sequencing using the BAG-Seq workflow in particular.

2. Results and Discussion

2.1. Mag-BAG Generation

We produce Mag-BAGs with a droplet generator that includes a co-flow for two aqueous solutions.[9] The first solution contains the single-cell suspension. The second solution contains the hydrogel component, the cell lysis buffer, and the magnetic beads. The oil phase contains the PEG-PFPE surfactant[16] dissolved at 2% weight in fluorinated oil HFE7500 and TEMED 0.4% (v/v) to catalyze the polyacrylamide polymerization. After encapsulation, we incubate Mag-BAGs overnight at 50 °C for gelation (**Figure 1**). We manipulate Mag-BAGs in aqueous buffers and process them via the split-pool method[5,6] using a series of washing and centrifugation steps.

We optimized gelation conditions by embedding magnetic beads in different acrylamide gel concentrations and varying the monomer concentration (%T). We first observed that hydrogel beads are reliably formed above 3.8%T (Figure S1, Supporting Information). Second, we investigated the impact of the monomer concentration on the separation performance. For each sample, we prepared a 500 µL 20% (v/v) suspension of Mag-BAGs in 6× SSC buffer, recorded a video of the separation, and extracted the kinetics and separation timescales (Figure S1e, Supporting Information). We established that the separation time was inversely correlated with the monomer concentration. The separation was complete at 300 s for 6.2%T, which we used in all subsequent experiments (Section S1.1, Supporting Information).

Bead variety strongly affects the separation performance of Mag-BAGs, as demonstrated by head-to-head comparison of

1 µm and 500 nm magnetic beads. 500 nm beads sediment more rapidly, which creates difficulties in maintaining a uniform concentration of beads over the Mag-BAGs population. Depending on the buffer, the smaller beads also tended to remain aggregated after magnetization. Finally, they provide a lower magnetic force per unit weight than the 1 µm beads (Figure S2, Supporting Information).

In this report, we used streptavidin-functionalized beads because they are readily available and would not interfere with nucleic acids. Experiments proved that their performance is similar to un-functionalized beads provided by the manufacturer (Figure S3a,b, Supporting Information). This result allows multifunctionalizing Mag-BAGs to capture additional molecules with capture moieties tethered to the magnetic beads. We also observed that the magnetic separation of Mag-BAGs was not affected by the high salt concentration of the 6× SSC buffer (Figure S3c,d, Supporting Information) or 40 cycles on a thermocycler (Figure S3e,f, Supporting Information). We noticed that thermocycling could induce a magnetization loss for other magnetic beads. This observation further emphasizes the need to select the appropriate bead type for specific applications. We used the 1 µm superparamagnetic beads for all other experiments because of their better overall performance. Those results also indicate that Mag-BAGs embedded with the 1 µm superparamagnetic beads used here are compatible with various bead functionalization, buffer types, and thermocycling.

2.2. Effect of Weight Bead Concentration and Mag-BAG Size on Separation Rate and Pellet Shape

BAG-Seq automation requires a fast separation rate and a pellet shape that enables efficient supernatant aspiration. Here, we investigated the effect of bead count and Mag-BAG size on these metrics. We generated a series of Mag-BAGs with 100, 65, and 45 µm diameters, each at five different bead weight concentrations doubling from 0.14 to 2.3 mg mL^{-1}.

The separation timescale at fixed Mag-BAG size decreased with increasing magnetic bead concentration. For 100 µm Mag-BAGs, the separation was complete within 60 s at 2.30 and 1.15 mg mL^{-1} (**Figure** 2a–c, green and red). Below 0.59 mg mL^{-1}, the separation timescale increased substantially with decreasing concentrations (Figure 2d).

The separation timescale at fixed magnetic bead concentration increased with decreasing Mag-BAG sizes (**Figure** 3d). At 1.15 mg mL^{-1}, the separation was only complete at 240 s for the 45 µm case compared to 60 s for the 100 µm case (in Figure 3a–c, gold). Notably, the separation timescale became less sensitive to the Mag-BAG diameter with increasing bead concentration (Figure 3d). Indeed, at 2.3 mg mL^{-1}, the timescale became independent of the Mag-BAG diameter, indicating a regime where the magnetic force largely dominated the drag force.

Indeed, magnetic and viscous drag forces scale differently with the Mag-BAG diameter. On the one hand, the magnetic force is proportional to the number of embedded magnetic beads and scales with the bead weight concentration (Section S1.6, Supporting Information). Thus, increasing the bead concentration increases the magnetic force and decreases the timescale. At fixed concentration, the bead count per Mag-BAG scales linearly with the Mag-BAG volume or the cube of its diameter. On the other hand, the Stokes drag force on a spherical Mag-BAG scales linearly with the diameter $F_d = 3\pi\mu Dv$ (4), where μ is the dynamic viscosity of the buffer, v is the separation velocity of the Mag-BAG, and D is the Mag-BAG diameter. Thus, the increase in magnetic force with increasing diameter exceeds the corresponding

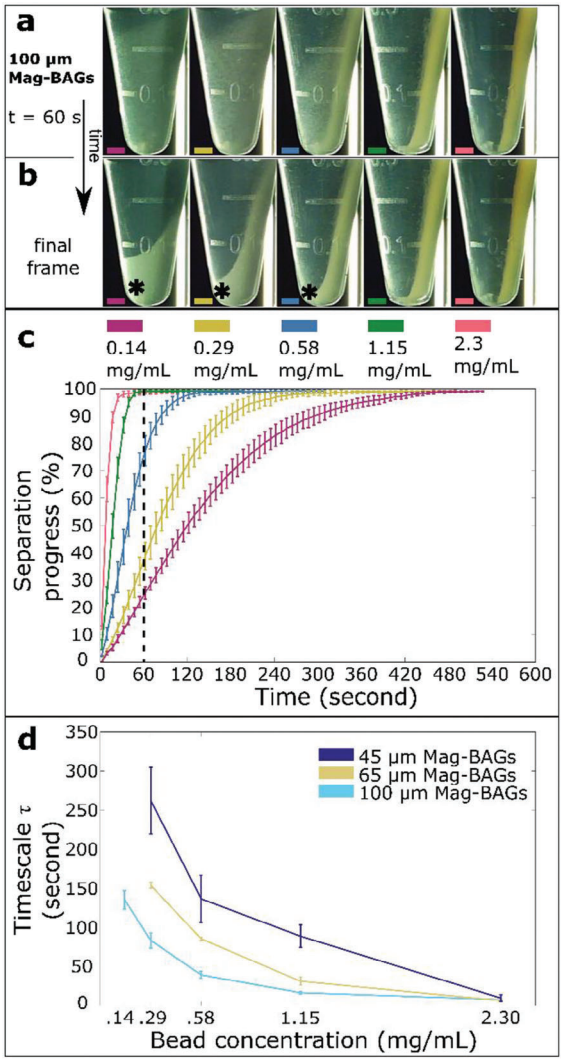

Figure 2. Higher bead concentration reduces separation time and narrows the pellet shape at a constant Mag-BAG diameter. a) Video frames showing bulk separation progress at t = 60 s of 2.3 (red), 1.15 (green), 0.58 (blue), 0.29 (gold), and 0.14 (purple) mg mL^{-1} samples. At 1.15 and 2.3 mg mL^{-1}, magnetic separation is complete within 60 s. Scalebar = 5 mm. b) The Last frame of videos analyzed in panel (b) shows the final pellet shapes. Higher bead concentrations yield flatter vertical pellets. Scalebar = 5 mm. The asterisks indicate where Mag-BAGs accumulate due to the effect of gravity. c) Bulk separation progress of 100 µm Mag-BAGs (6.2%T) at 0.14, 0.29, 0.58, 1.15, and 2.3 mg mL^{-1} bead concentration. Curves and error bars represent the mean and standard deviation of three video replicates. d) Timescale τ of the separation curves as a function of bead concentration and BAG size. τ decreases with increasing bead concentration, indicating a faster separation with increasing bead concentration. τ becomes independent of Mag-BAG size at higher bead concentrations. Curves and error bars represent the mean and standard deviation of at least three video replicates.

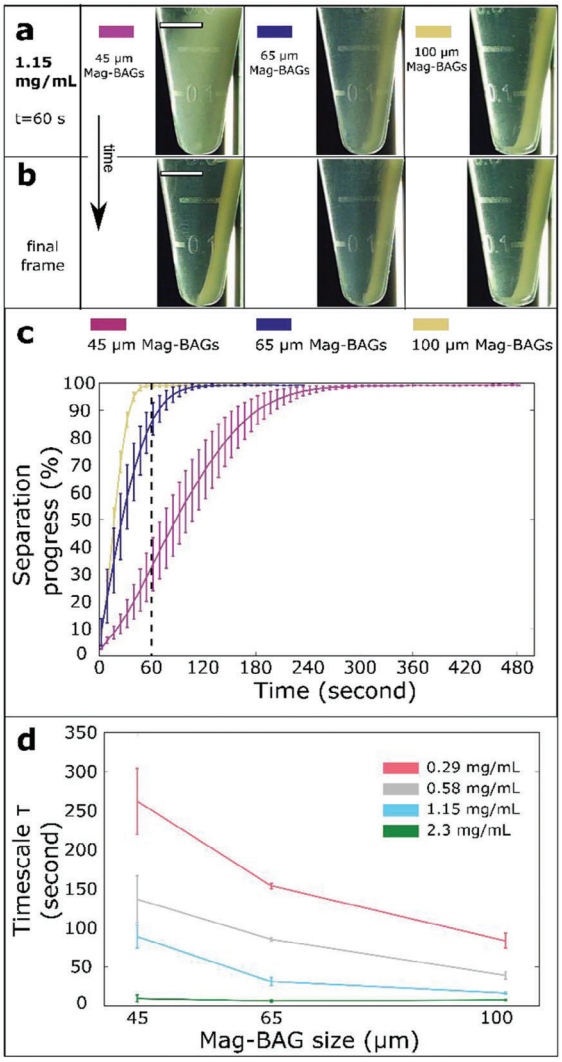

Figure 3. A larger Mag-BAG diameter reduces separation time at constant bead concentration. a) Bulk separation progress of 100, 65, and 45 µm diameter Mag-BAGs at 1.15 mg mL^{-1} bead concentration. Curves and error bars represent the mean and standard deviation of three video replicates. b) Video frames showing bulk separation progress at $t = 60$ s of 100 µm (gold), 65 µm (blue), and 45 µm (purple) Mag-BAG samples. Final separation time increases from ≈60 s for the 100 µm case to ≈150 and 300 s for the 65 and 45 µm cases, respectively. Scalebar = 5 mm. c) The Last frame of videos analyzed in panel (b) shows the final pellet shapes. Pellet shape is unaffected by BAG size at 1.15 mg mL^{-1} bead concentration. Scalebar = 5 mm. d) timescale τ as a function of bead concentration and Mag-BAG size. τ becomes independent of Mag-BAG size at higher bead concentrations (above 2.3 mg mL^{-1}). Curves and error bars represent the mean and standard deviation of at least three video replicates.

increase in the viscous drag force, resulting in a decrease in timescale.

Reducing the Mag-BAG size is necessary for specialized applications such as fluorescence-activated cell sorting, designed to sort particles in the 10–30 µm range efficiently. We could use a higher concentration of magnetic beads to maintain a suitable separation timescale for smaller Mag-BAGs. We also noted that for decreasing BAG size, a fixed bead count per BAG increased the relative effect of the magnetic force and thus made more efficient use of the beads in achieving a desired timescale. More generally, smaller Mag-BAGs offer cost reduction at scale by lowering reagent consumption per droplet and processing cost per cell. The most cost-effective design would minimize Mag-BAG size and bead concentration for the appropriate timescale.

The pellet shape is important because it affects the washing efficiency and yield, ease of use, and robustness to variation in pipette position. The ideal pellet is pulled flat against the tube wall without sedimented BAGs. This shape allows for optimal distance between the pipette tip and the pellet. It enables complete supernatant aspiration with minimal pellet disturbance and increases washing efficiency. In automated applications, this shape makes the separation more robust to position variability of the pipette tip by removing the pellet from the tip's vertical path. For 100 µm Mag-BAGs, we achieved the ideal shape for 2.3 and 1.15 mg mL^{-1} magnetic bead concentrations (Figure 2b, red and green). At lower concentrations, hydrogel particles accumulated at the bottom of the tube due to the effect of gravity (Figure 2b, asterisks in blue, gold, and purple). At a constant bead concentration of 1.15 mg mL^{-1}, the pellets for 100, 65, and 45 µm Mag-BAGs displayed an ideal shape (Figure 3b). Smaller Mag-BAGs could generate the ideal pellet shape at a lower bead count. Their smaller sizes led to a reduced sedimentation rate and a reduced drag force, allowing them to move sideways faster.

2.3. Magnetic Separation Offers Improved Yield Over Centrifugation After Repeated Wash Cycles

We compared the separation yield of Mag-BAGs (100 µm diameter, 2.3 mg mL^{-1} bead concentration) by magnetic separation and centrifugation. We measured the yield as the percent volume loss of a sedimented Mag-BAG layer in 500 µL 20% (v/v) suspensions after ten wash cycles (two split-pool steps require 15 wash cycles). We considered two use cases: i) 1.5 mL microcentrifuge tubes and ii) 96-well PCR plates. We used both 6× SSC and PCR buffers because both apply to different stages of the BAG-seq workflow.[9] Also, the higher viscosity of 6× SSC generates higher local shear forces, which could increase the loss of BAGs during buffer aspiration. We evaluated $N = 3$ replicate samples for each separation method, buffer, and container type combination.

The magnetic separation resulted in greater yield for wash cycles performed in 1.5 mL tubes, and the buffer type did not make a difference. Using a two-way ANOVA, we found a statistically significant difference in percent volume loss between magnetic separation and centrifugation ($F(1) = 20.804$, $p<0.001$) but not between 6× SSC and PCR buffers ($F(1) = 2.035$, $p = 0.192$) or for the interaction between separation and buffer types ($F(1) = 0.792$, $p = 0.399$). Critically, we aspirate 400 µL from 500 µL of a 20% v/v BAG suspension, leaving no remainder supernatant. Thus, we conclude that magnetic separation offers improved yield over centrifugation with either buffer using 1.5 mL tubes. Magnetic separation improves washing yield by allowing a greater volume to be aspirated thanks to the pellet shape produced by a side magnet configuration.

The 96-well PCR plate format provided a better yield than microtubes for both centrifugation and magnetic separation (**Figure** 4b,d). In 96-well PCR plates, magnetic separation

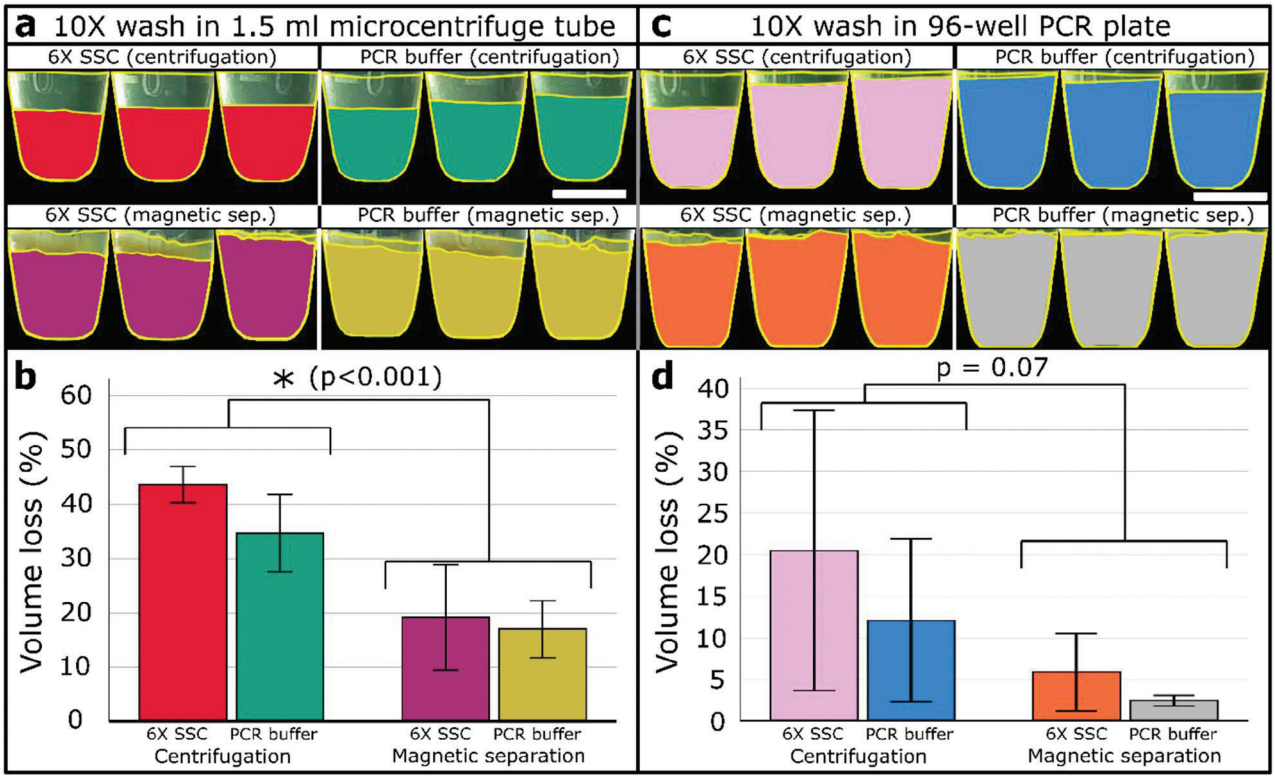

Figure 4. Magnetic separation offers improved yield over centrifugation after repeated wash cycles (100 µm diameter, 2.3 mg mL^{-1} bead concentration). a, c) ROIs indicating BAG pellet boundaries before and after ten separation/wash cycles performed in (a) 1.5 mL microcentrifuge tubes and (c) 96-well PCR plates. BAGs (no magnetic beads) were separated via centrifugation (top), and Mag-BAGs (2.3 mg mL^{-1} 1 µm magnetic beads) via magnetic separation (bottom). Washes were performed with 6× SSC (left) and PCR buffers (right). b, d) Percent volume loss for each condition of suspension buffer and separation method performed in (b) 1.5 mL microcentrifuge tubes and (d) 96-well PCR plates. The buffer type does not significantly affect the recovery efficiency in all cases. However, magnetic separation exhibits a significantly ($p < 0.001$) higher recovery efficiency than centrifugation in 1.5 mL microtubes. Recovery was also higher for magnetic separation than centrifugation in 96-well PCR plates, but the difference was not statistically significant ($p = 0.07$).

performed better than centrifugation overall. A two-way ANOVA found no statistically significant difference in percent volume loss between separation methods ($F(1) = 4.384$, $p = 0.07$), buffer type ($F(1) = 1.038$, $p = 0.338$), or their interaction ($F(1) = 0.183$, $p = 0.680$). The key benefit to magnetic separation is that centrifugation exhibited a higher variability between repeats in these conditions. We attribute this variability to the reduced precision in positioning the multichannel pipette tips relative to the BAG pellet aggregated at the bottom of the well. Aspiration from 96-well plates is more prone to accidental disturbances of the BAG pellet when using centrifugation. The advantage of the pellet shape with magnetic separation was particularly apparent when using 96-well plates. We can apply the tips of the multichannel pipette along the opposite wall of the undisturbed Mag-BAG pellets. In addition, magnetic separation maintains the pellet's aggregation force during aspiration, contributing to a higher pellet integrity.

2.4. Magnetization Does Not Affect BAG-seq Sequencing Data Quality

We would expect the mere incorporation of magnetic beads into BAGs, thus creating Mag-BAGs, and the use of magnetic versus centrifugal separation would not produce differences in the data quality arising from the BAG-seq workflow. Nevertheless, we tested the two methods for our most frequent application: detection of segmental amplifications or deletions of the genome in cancer and normal cells. We compared the data from BAGs and Mag-BAGs using a similar workflow (see Experimental Section).

In brief, both BAG and Mag-BAG protocols used in this experiment are based on the original BAG-seq workflow,[9] where Acrydite primers captured the genomic content of each single cell copolymerized into the gel ball matrix. In addition, for the Mag-BAGs protocol, we added 1 µm magnetic beads into droplets. The only other difference between these protocols is that we used magnetic separation for the Mag-BAGs and a standard centrifugation step for the BAGs during the pool-and-split steps. Two rounds of pool-and-split steps were carried out for each protocol. After the genomic DNA copies are captured and replicated, they are cut using the enzyme NLAIII to achieve a universal 3 'overhang sequence CATG. Each round introduces 96 different cell barcodes, generating 96×96 = 9216 cell barcode varieties. The first pool-and-split also introduces four bases of random sequences, the varietal tag, which, combined with the mapped sequence, produces a template-specific barcode unique to each template.

After sequencing using Illumina NovaSeq, we analyzed the sequencing data using a pipeline described in the original BAG-seq paper.[9] We extracted the cell barcode and template-specific tag information from the fastq files. Subsequently, we mapped the read pairs to the human reference genome using HISAT2.[17] Reads with the same varietal tag, mapped sequence, and cell barcode were identified as originating from the same template.

The cells were a mixture of SKN-1, a cultured human skin fibroblast, and SK-BR-3, a cultured human female breast cancer. The genome of cancerous SK-BR-3 cells has a characteristic CNV profile, while the SKN-1 cells are normal fibroblast cells with a diploid male genome. Each cell type has unique single nucleotide variants (SNVs), allowing their identification and validation of methodologies (see Experimental Section). Whereas we would ordinarily go through three rounds of pool and split for creating BAG tags, in these experiments, we only used two rounds, increasing BAG tag "collisions." However, that does not affect the comparison of methods. We used already established protocols by our group for copy number analysis.

We compared the sequence data from the two methods by the following criteria: number of templates per BAG for a given read depth; fragment length distribution of templates; extent of diffusion between BAGs of cell type-specific SNVs; correctness of subpopulation identification; and consistency of the copy number profiles obtained by the two methods. By all these criteria, the two methods are within 10–20% of each other, if not indistinguishable.

We have two libraries, one from BAGs, A, and one from Mag-BAGs, B. After provisional cell type assignment is determined by the majority of either SKN-1 or SK-BR-3 SNVs, we determine which Mag-BAG tags and BAG tags represent captured cells using the second derivative of smoothed ranked read count curves for SKN-1 and SK-BR-3 cells separately and present them in their respective cumulative distribution plots (**Figure 5a**). The resulting data set for Mag-BAGs has 3472 tags with cells totaling 281 835 031 reads; and for BAGs 2583 tags with cells totaling 341 867 689 reads. In order to facilitate comparison between the libraries, we downsampled the number of tags with cells in the Mag-BAG library to match that in the BAG library. Then we downsampled the total number of reads in the BAG library to match the total read count in the downsampled Mag-BAG data set, resulting in two libraries, A' and B', each with 2583 tags with cells totaling 207 763 305 reads.

We identified tags with a single cell type using thresholds of greater than 98% SK-BR-3 SNVs or greater than 70% SKN-1 SNVs. The other tags are considered "mixed" tags (Figure 5b). In these data sets, we have for Mag-BAGs: 1883 SK-BR-3 tags, 245 SKN-1 tags, and 455 mixed tags. For BAGs, we have 2270 SK-BR-3 tags, 102 SKN-1 tags, and 211 mixed tags. The variation in the proportion of SKN-1 cells undoubtedly represents user pipetting variability. Both BAGs and Mag-BAGs have a high collision rate, as seen in the population with "mixed" SNVs, and this is due to performing only two rounds of pool-and-split. The ratio of SNVs for threshold identifiers is skewed because the SK-BR-3 cells have a higher ploidy, there are more SK-BR-3 cells than SKN-1, and there are about fourfold more SK-BR-3 SNVs detected than SKN-1 SNVs.

First, we computed the number of templates (template tags) per BAG for the two protocols, restricted to the downsampled libraries, A' and B'. These are shown as distributions in Figure 5c. The distributions are very similar, and the median number of templates was 35 014 per BAG and 33 123 per Mag-BAG, within 10%. This result demonstrates similar capture and tagging from individual genomes, irrespective of the protocol. Second, we examined the size distribution of the templates for each protocol. They are virtually indistinguishable, as seen in Figure 5d. This result demonstrates that the biochemical processing of cellular DNA (priming, elongation, and cleavage) proceeds equivalently in the two environments. Third, we examined the diffusion rates in the two libraries (see Experimental Section), that is, the contamination of one BAG with genomic fragments from another BAG. To do this, we computed the ratio of SK-BR-3 to SKN-1 SNVs within the BAGs identified as primarily SKN-1 for each BAG or Mag-BAG and then took the mean. The mean ratios are 0.128 and 0.122, from BAGs and Mag-BAGs, respectively, within 5% of each other. This result demonstrates that the diffusion of lysed cellular nuclear DNA between the BAGs for the two protocols is similar.

Next, we clustered BAGs by their genomic copy number features and derived the copy number profiles of each cluster for each of the two protocols. To cluster by genomic features, we first used the known copy number profile of the two components, the normal diploid and the tumor cell line. We used the boundaries of the SK-BR-3 copy number profile as the boundaries of the bins used in clustering. We took 100 cells from each SKN-1 and SK-BR-3 BAG and combined them similarly for the Mag-BAGs. Reads are assigned to each bin for each BAG. We use a multinomial expectation maximization algorithm (EM) that will be described in another manuscript to cluster the populations (Figure 5e). In each case, the clustering reproduced the identity determined by SNVs without any assignment error. We derived the copy number from each cluster, compared them for the two protocols, and saw only very minor differences (Figure 5f).

We conclude that magnetic beads in Mag-BAGs do not affect the ability to distinguish between the two cell lines using BAG-seq. Mag-BAG is suitable for the general application of BAG-seq and split–pooling. These results support the purpose of our work, which is to enable us to robotize the BAG protocol. Magnetic field separation should be easier to robotize than centrifugation or filtration. Currently, the protocol is manual and labor intensive, particularly the pool-and-split steps.

Consequently, we are limited in the number of unique BAGs we can produce, and the process is prone to user error. Our objective is to reduce labor, increase reproducibility, and increase the number of cells that can be uniquely tagged from tens of thousands to millions through larger splits at the pool-and-split stage. The ultimate applications are in the analysis of tumor biopsies or the analysis of circulating blood components in healthy persons and patients with leukemia, blood dyscrasias, infections, or inflammation.

3. Conclusion

In this work, we developed magnetized balls of acrylamide gel (Mag-BAGs) by physically embedding commercially available superparamagnetic microparticles. We achieved consistent and repeatable BAG magnetization that allows bulk separation and future automation of critical workflows such as split–pooling.

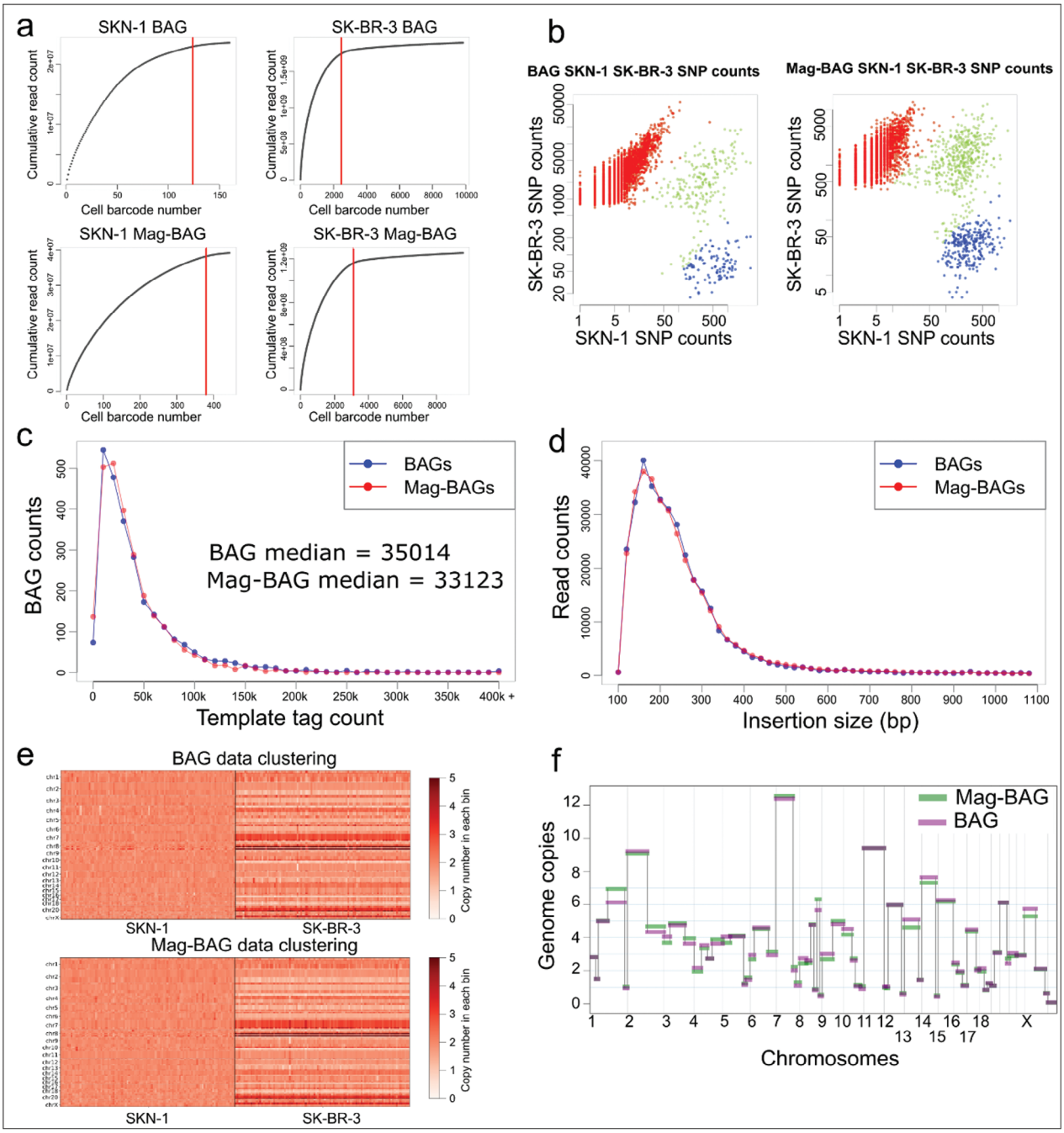

Figure 5. Comparison between BAG and Mag-BAG performance from a single-cell SKN-1/SK-BR-3 mixture experiment. a) Cumulative distribution plots indicating the cut-offs (red line) for SKN-1 and SK-BR-3 cells for both protocols. The SKN-1 and SK-BR-3 cells were separately plotted and distinguished by abundant SNPs for each BAG-barcode. b) The three populations in each condition indicate SKN-1 (blue), SK-BR-3 (red), and collisions (green). c) After downsampling to the same number of cells and total reads for both conditions, the distribution of the number of templates per BAG is shown in a histogram. d) The histogram displays the size distribution of the templates based on the mapping information for all mapped read pairs in each condition. e) The heatmaps show the copy number profiles of randomly sampled 100 SKN-1 and 100 SK-BR-3 cells for each condition. f) The aggregated copy number profiles for all the SK-BR-3 cells for each condition show similar profiles.

Magnetization, separation performance, and molecular applications strongly depend upon the physical characteristics of the magnetic particles. We identified 1 μm superparamagnetic particles that perform robust separation for typical buffer exchange and thermocycling. Larger BAGs and higher bead concentrations result in faster separation due to the differential scaling relationships of the magnetic force and viscous drag with the Mag-BAG radius. Above a critical bead concentration, the separation timescale becomes independent of the Mag-BAG size. Here, we reach the maximum utility of increased bead count per BAG. We

can thus optimize the total bead content per Mag-BAG to achieve a desired separation rate and pellet profile to reduce bead and reagent costs.

Mag-BAG washing via magnetic separation offers improved yield to centrifugation in both 1.5 mL tubes and 96-well plates, proving more robust to user and system variability. Mag-BAGs are fully compatible with single-cell sequencing workflows. We used Mag-BAGs to efficiently distinguish cells from a mixed cell population using the BAG-seq protocol. Mag-BAGs exhibited similar sequencing performance as non-magnetized BAGs.

In conclusion, we increased the functionality of hydrogel beads by embedding magnetic beads using a straightforward method. These Mag-BAGs provide increased yields and washing efficiencies for purification procedures. Our work enables automation to replace the labor-intensive methods required for split–pooling, and we expect it will have a lasting impact on single-cell genomic workflows.

4. Experimental Section

Microfluidic Designs: Microfluidic droplets containing a bead suspension and a gel monomer solution were generated using a 2-layer co-flow design (see the CAD file in Supporting Information). For the BAG magnetization experiments, three different microfluidic circuits were designed for nominal BAG diameters of 100, 65, and 45 µm. A schematic of the circuits is shown in Figure S4 (Supporting Information). Relevant channel dimensions are tabulated in Table S1 (Supporting Information). For the scDNA experiment, a commercially sourced Drop-seq device (Nanoshift, Emeryville, CA, USA) was used.[18]

Microfabrication: Microfluidic devices were fabricated in PDMS (Sylgard 184 Silicone Elastomer Kit, Dow Corning, Midland, MI) using soft lithography.[19] An established protocol to create a double-layer mold was used.[20] The microfluidic circuits using the AutoCAD software (AutoDesk, San Francisco, CA, USA) and had them printed onto mylar masks (CAD/Art services, Bandon, OR, USA) were designed. A negative photoresist (SU-8, Kayaku Advanced Materials, Westborough, MA, USA) onto 3″ silicon wafers was spincoated. A UV mask aligner (500 W UV illuminator, Newport, Irvine, CA) to pattern the designs onto the photoresist was used. The structures using PGMEA (Kayaku Advanced Materials, Inc) were developed. To facilitate PDMS release, the mold surfaces via overnight vapor deposition of a fluorinated silane compound (1H, 1H, 2H, 2H-Perfluorooctyltrichlorosilane, Gelest, Morrisville, PA) were silanized. The mold onto a custom acrylic jig to ensure consistent PDMS slab thickness was mounted. PDMS was mixed with crosslinker at a 10:1 (polymer:crosslinker) weight ratio, degassed under vacuum, and poured before curing for 1 h at 65 °C. The inlet and outlet ports with a biopsy tool (Syneo, West Palm Beach, FL, USA) were punched. PDMS slabs were bonded to 50 × 75 mm glass slides spin-coated with a 150 µm thick layer of PDMS to ensure the same surface properties of all four channel walls. Bonding was achieved after surface activation via 1 min exposure to oxygen plasma (PDC-32G, Harrick Plasma, Ithaca, NY). To ensure preferential wetting of fluorinated oil, channels are surface-treated with a solution at 5% (v/v) of (heptadecafluoro-1,1,2,2-tetrahydrodecyl)trichlorosilane (Gelest) in HFE-7500 (3 M, St. Paul, MN, USA) for 5 min right after bonding. The solution was then flushed with pure HFE-7500 oil.

Fluid Compositions: Fluid compositions were based on the scDNA BAG-seq protocol of Li et al,[9] with slight modifications for investigating BAG magnetization. Beads were added to the water component volume in the aqueous phase 2. For optimization experiments (i.e., without cells), the Linker TG primer and Proteinase K components were replaced with equal volumes of water. Otherwise, fluid compositions were taken from the scDNA BAG-seq protocol.[9]

The oil phase consisted of 0.4% (v/v) TEMED (1610801, Bio-Rad, Hercules, CA, USA) in HFE-7500 (3 M) that contained 2% (w/w) PFPE-PEG surfactant.[21] TEMED in the oil phase catalyzed the polymerization of the acrylamide gel within droplets. Aqueous phase 1 consisted of 0.05% (w/v) BSA in 1× PBS.

For scDNA applications, the aqueous phase 2 was prepared in 500 µL batches containing 90 µL Acrylamide/Bis 19:1, 40% w/v (A9926, Sigma–Aldrich, St. Louis, MO), 64.5 µL Acrylamide 40% w/v (A4058, Sigma–Aldrich, St. Louis, MO), 135.5 µL bead suspension in H_2O, 80 µL 500 µM Linker TG primer /5ACryd//iSp18/TGTGTTGG GTGTGTTTG-GKKKKKKKGKKKKKKKKNN, (Integrated DNA Technologies, Coralville, IA, USA), 25 µL EDTA 0.5 M (AM9260G, Thermo Fisher Scientific, Waltham, MA, USA), 50 µL 1 M Tris pH 7.5 (15567-027, Invitrogen), 5 µL 20% (w/v) Sarkosyl (L7414, Sigma–Aldrich) in H_2O, 10 µL Proteinase K (P4850, Sigma–Aldrich), 10 µL 0.1 M DTT (707265ML, Thermo Fisher Scientific, Waltham, MA, USA), and 30 µL 10% (w/v) ammonium persulfate (09913-100G, Sigma–Aldrich) in H_2O.

Three types of superparamagnetic beads: 1 µm streptavidin-functionalized beads (S1420S, New England Biolabs, Ipswich, MA, USA), 1 µm blank beads (i.e., non-functionalized, but otherwise identical to S1420S from New England Biolabs, provided as a special request), and 500 nm azide-functionalized beads (MGB-AZD-10-10, Luna Nanotech, Markham, ON, Canada) were investigated. Beads were washed three times in wash buffer (0.1 M PBS pH7.4) using a magnetic separation stand (Z5332, Promega) before being resuspended in H_2O.

Experimental Setup: Solutions were injected using 1 mL glass syringes (Gastight #1001, Hamilton, Reno, NV, USA) mounted on syringe pumps (NE-300, New Era Pump Systems, Farmingdale, NY, USA). Fluids were delivered to the devices via PEEK tubing (0.254 mm ID, 0.787 mm OD, Zeus, Orangeburg, SC, USA). On-chip droplet generation was monitored using an inverted brightfield microscope (Diaphot-TMD, Nikon, Tokyo, Japan) equipped with a camera (XCD-V60, Sony, Tokyo, Japan). Droplet and BAG images were captured with 10× (Plan Ph1 10/0.3 DL 160/0.17, Nikon) and 20× (Plan Ph2 20/0.4 160/1.2 ELWD, Nikon) objective lenses using a custom LabVIEW software (National Instruments, Austin, TX, USA).

BAG Generation: When generating BAGs for the magnetization and yield experiments, the aqueous phases 1 and 2 were premixed at 1:1 (v/v) into a 1 mL total volume and injected using a single syringe. Oil and aqueous phases were injected at 468.2 and 323.1 µL h^{-1}, respectively. When using 500 nm beads, these flow rates were doubled to reduce sedimentation in the syringe (see Section 2.3). The emulsion was collected into 1.5 mL microcentrifuge tubes containing 300 µL mineral oil (330779, Sigma–Aldrich) for a minimum of 1 h.

Incubation and BAG Release: After collection, samples were incubated at 50 °C overnight to ensure complete polymerization of the gel matrix. After incubation, the HFE-7500 oil layer at the bottom of the sample tube was aspirated with a 22G needle and replaced with a solution of 2% (w/w) PFPE-PEG surfactant[21] in FC-40 fluorinated oil (F9755, Sigma–Aldrich). Samples were then incubated at 95 °C for 12 min, at 55 °C for 1 h, and at room temperature for 10 min. A 22G needle to remove both the bottom layer of FC-40 oil and the top layer of mineral oil was used. The dense emulsion layer was then resuspended in 600 µL 6× SSC buffer. 150 µL of 1H,1H,2H,2H-perfluorooctanol (370533, Sigma–Aldrich) was added, and manually shook the tube for 10 s to break the droplet interfaces and release the balls of acrylamide gel (BAGs). The tube was centrifuged at 1000 rcf for 1.5 min before removing the bottom oil and top buffer layers with a 22G needle. The BAGs using 600 µL 6× SSC buffer and centrifuged at 1000 rcf for 1.5 min were again washed.

Sample Preparation: 500 µL 20% (v/v) Mag-BAG suspensions in 1.5 mL microcentrifuge tubes were prepared for recording magnetic separation videos. Equal volumes of the sedimented BAG layer (100 µL) for each sample were ensured. Before separation, the BAG suspension was agitated via pipette aspiration to break up BAG clusters and ensure a uniform dispersion.

Magnetic Separation Video Capture: Mag-BAGs in suspension were separated on a two-tube magnetic separation stand (Z5332, Promega) using the default side-magnet configuration. The separation was recorded

using a USB microscope/camera (#1061, Adafruit, NY, NY, USA). A dark backdrop was placed behind the tube, and a 3.75″ square LED backlight (LED-SP, Amscope, Irvine, CA, USA) was placed between the magnetic stand and the backdrop (Figure S5a, Supporting Information). The backlight was covered with a mylar mask with a 0.6 mm wide transparent slit to illuminate the sample with a vertical structure light (Figure S5b, Supporting Information). Videos were recorded in a dark room using the AMCap software, producing 640 × 480 pixel RBG videos in .avi format. The setup allowed for a clear contrast between the suspended Mag-BAGs and the background buffer during separation. Videos are available upon request.

Video Post-Processing and Separation Curve Generation: The procedure for generating separation progress curves from video is outlined in Figure S6 (Supporting Information). The separation progress was monitored by comparing the mean gray value of the Region of Interest (ROI) that defines the final pellet and the rest of the solution (void area). Video .avi files were imported to ImageJ/Fiji[22] as a virtual time stack and cropped to a rectangle, bounding the tube tip, the meniscus bottom, and the side edges. The red channel's 8-bit greyscale .tif stack was generated to improve the contrast between the separated Mag-BAG pellet and the "void area." "Enhance Contrast" was then applied with the "Normalize" setting and 20% pixel saturation. The average background grey value inside the void area was then subtracted from the image before applying "Enhance Contrast" again with the same settings.

ROIs bounding the pellet and void areas were established using the last slice of the time stack: 1) the polygon tool was used to manually draw an ROI that defined the entire interior perimeter of the solution (including the pellet); 2) an ROI was drawn to define the void area with the same method; 3) these two ROIs were combined with the "XOR" command, and then the "Split" command was used to generate the pellet ROI. The resulting pellet and void area ROIs formed two subdivisions of the full tube interior.

A custom ImageJ macro measured the mean grey value inside these ROIs and calculated the difference between these values (the "raw contrast" C) for each slice in the time stack. These contrast values were then normalized to the minimum and maximum values:

$$C_{norm} = \frac{C - C_{min}}{C_{max} - C_{min}} \tag{1}$$

The evolution of C_{norm} across the time stack was an exponential asymptote from 0, reflecting the homogenous distribution of Mag-BAGs in suspension, to 1, reflecting the complete magnetic separation of Mag-BAGs into a pellet along the side of the tube. $C_{norm}(t)$ is thus modeled as:

$$C_{norm}\ (t) = \ 1 - A\left(e^{\frac{-t}{\tau}}\right) \tag{2}$$

MATLAB's fit() function with "fitType" parameter "exp1" to fit A and τ to $(1-C_{norm})$ from the time-array of C_{norm} is used. Here, τ is the separation timescale; this parameter is used as a metric of separation time performance (see Results and Discussion).

Post-Thermocycling Separation Experiment: A baseline magnetic separation video of a 500 µL 20% (v/v) BAG suspension in 1× standard *Taq* reaction buffer (B9014S, New England Biolabs) was recorded. Each sample was then split into two PCR microtubes to accommodate their lower volume. Each tube was then subjected to the following thermocycling sequence: step 1) 1× of 72 °C for 300 s, step 2) 94 °C for 30 s, step 3) 40 cycles of 94 °C for 10 s, followed by 63 °C for 30 s, followed by 72 °C for 60 s. After thermocycling, the two tubes for each concentration were recombined into a 1.5 mL microcentrifuge tube to which 300 µL of 1× PCR buffer was added to restore the original 500 µL 20% (v/v) sample volume and concentration. Post-thermocycling separation videos were then recorded.

Magnetic Separation Yield Experiment (1.5 mL Microcentrifuge Tubes): To compare the magnetic separation yield performance of Mag-BAGs against centrifugation-based separation in 1.5 mL microcentrifuge tubes, 100 µm BAGs and Mag-BAGs at 2.3 mg mL^{-1} bead concentration were generated. For each sample, four sets of three replicates of 500 µL 20% v/v samples were prepared: i) bead-free BAGs in 6× SSC buffer, ii) Mag-BAGs in 6× SSC buffer, iii) bead-free BAGs in 1× PCR buffer, and iv) Mag-BAGs in 1× PCR buffer.

The separation performance is compared by measuring the percent volume loss of the BAG pellets between a pre-separation baseline and the remaining volume after ten wash cycles. The baseline volume was first recorded after concentrating Mag-BAGs with a bottom magnet and allowing BAGs to sediment. For the bead-free BAG samples, BAGs were separated by centrifuging at 1000 rcf for 1.5 min. For the Mag-BAG samples, BAGs were separated for 30 s on a magnetic separation stand (25 332, Promega) with the default side-magnet configuration.

To aspirate 400 µL of supernatant, sample tubes were placed on the separation stand and monitor them live using the video-capture setup. Using the video feed as a guide, a 1 mL micropipette was carefully inserted into the tube to the 0.3 mL mark and then slowly aspirated 400 µL of supernatant, slowly lowering the pipette tip as needed to ensure that the tip stayed submerged while also not disturbing the BAG pellet. It was then resuspended in 400 µL of fresh buffer and repeated the cycle. After ten separation/wash cycles, the post-separation volume was recorded.

Magnetic Separation Yield Experiment (96-Well PCR Plates): Using the same conditions as the separation in 1.5 mL microcentrifuge tubes, the separation was performed directly in a 96-well plate using a magnetic separator designed for PCR plates (12331D, Invitrogen) and a centrifuge.

After recording the baseline BAG pellet volumes, each sample was diluted to 1 mL (10% v/v) and split into 8 wells of a 96-well PCR plate. For each cycle, the 96-well PCR plate was placed on the magnetic separator for 2 min (for the Mag-BAGs) or centrifuged at 1000rcf for 2 min (for the bead-free BAGs). 85 µL of buffer was aspirated from the top of each well simultaneously using an eight-channel multipipette and then replaced with 85 µL of the appropriate buffer. The pipette tips were pressed against the wall of the wells opposite the magnet-side wall, slowly lowered just below the fluid surface, and fluid was gently aspirated while slowly lowering the tips to keep pace with the descending surface. After 10 cycles, the wells from each sample's row were recombined into a 1.5 mL microcentrifuge tube, centrifuged at 1000 rcf for 1.5 min, and 500 µL of supernatant buffer was aspirated to restore the starting 500 µL 20% (v/v) sample volume and concentration. Finally, the post-separation pellet volume and estimated percent volume loss were recorded.

Estimating BAG Pellet Volume Loss: To estimate the percent volume loss of the BAG pellets, the pre- and post- separation video .avi files recorded by AMCap were imported to ImageJ/Fiji for volume analysis. Using the polygon tool, ROIs around the sedimented BAG layers in the last frame of each video were manually drawn. The baseline and post-separation pellet volumes were estimated as volumes of revolution by applying Pappus' second centroid theorem:

$$V = A \times d \tag{3}$$

where, A is the area to be revolved around the axis of revolution, and d is the distance traveled by the centroid of that area over the course of a 360° revolution about the axis. A custom ImageJ macro was used to estimate the volume of revolution for each ROI. First, a blank image was created with height and width equal to that of the ROI's bounding box. A new ROI on this image with a flat top surface was then defined, using the points that defined the original ROI but replacing the vertical coordinate for each point that defined the top surface with the averaged value among all these points. The revolution axis was defined as a vertical line down the middle of the image and the ROI was split into left and right halves about this line using the "XOR" and "Split" commands. The BAG pellet volume was estimated as the average of the two volumes of revolution calculated for each half using Equation (3) above, with $d = 2\pi \times r$ where r is the distance of the half-ROI's centroid from the axis of revolution.

Single-Cell DNA (scDNA) Sequencing Experiment: The use of Mag-BAGs was assessed in a single-cell DNA (scDNA) sequencing experiment. The performance was compared to bead-free BAGs, evaluated in a previous paper,[9] using human cells from normal skin fibroblast (SKN-1) and

a breast cancer cell line (SK-BR-3). The bead-free BAGs ("no mag beads" condition) and Mag-BAGs ("mag beads" condition) were generated and processed separately but concurrently by the same operator.

Sample Preparation, Droplet Generation, and BAG Release: 1 million total cells, from a mixture of two cell sources, were suspended in a 1 mL solution containing 850 µL aqueous phase 1 (0.05% (w/v) BSA in 1× PBS) and 150 µL OptiPrep (D1556, Sigma–Aldrich) to reduce cell sedimentation in the syringe and inlet tubing. For this experiment only, the oil phase consisted of 5% (v/v) surfactant (008-Fluorosurfactant, RAN Biotechnologies, Beverly, MA, USA) in HFE-7500. For generating Mag-BAGs, a final bead concentration of 1.15 mg mL^{-1} was used; 288 µL of the bead stock volume was washed three times in 1× PBS pH 7.4, then resuspended in the H_2O volume component of aqueous phase 2. For this experiment, the cell suspension (aqueous phase 1) and aqueous phase 2 were encapsulated into microfluidic droplets using a Drop-seq[18] device (Nanoshift, Emeryville, CA, USA). Flow rates of 650 µL h^{-1} for both aqueous phases and 2,800 µL h^{-1} for the oil phase were used. Droplets were collected for ≈15 min into a 1.5 mL microcentrifuge tube containing 300 µL mineral oil. The BAGs were incubated overnight at 37 °C.

Linear Extension: The BAG tubes were incubated at 95 °C for 12 min, followed by 55 °C for 30 min and room temperature for 10 min before removing the oil layers. BAGs were successively washed with 6× SSC buffer and 1× NEBuffer 2 (B7002S, New England Biolabs) before being resuspended in a 1 mL solution containing 830 µL H_2O, 100 µL NEBuffer 2, 60 µL 10 mM dNTP (11814362001, Sigma–Aldrich), and 10 µL DNA Polymerase I (M0210M, New England Biolabs). The suspension was incubated at room temperature for 1.5 h and then at 37 °C for 30 min under rotation. The reaction was stopped using a series of STOP-X buffers, where X is the concentration of EDTA in mM; these buffers also contained 10 mM Tris pH 8.0, 0.1% Tween-20, and 0.1M KCl. After incubation, the BAG suspension was mixed with 5 mL of STOP-25 for 2 min. The BAG suspension was spun down, the supernatant removed, BAGs were resuspended in 1 mL STOP-10, passed through a 150 µm cell strainer, and then transferred to a new 1.5 mL microcentrifuge tube for each condition.

Exonuclease Treatment: After linear extension, a 3′ exonuclease treatment was used to cleave nucleotides from unused linker TG primers. BAGs were successively washed with STOP-1 and 1× Exonuclease I buffer (B0293S, New England Biolabs) before being resuspended in an 800 µL solution containing 680 µL H_2O, 80 µL 10× Exonuclease I buffer and 40 µL Exonuclease I enzyme (M0293S, New England Biolabs). This suspension was incubated at 37 °C for 1.5 h under rotation. BAGs were washed twice in STOP-25, and once in STOP-10 to stop the reaction.

4.0.0.1. CATG Overhang Cutting: DNA was cleaved using NLAIII restriction enzymes (R0125L, New England Biolabs) to generate 3′CATG overhangs. BAGs were washed successively with STOP-1 and 1× rCutSmart buffer (B6004S, New England Biolabs). BAGs were resuspended in a 1 mL solution containing 840 mL H_2O, 100 µL 10× rCutSmart buffer, and 60 µL NLAIII enzyme and incubated under rotation at 37 °C for 1.5 h. BAGs were centrifuged at 800 rcf, the supernatant was removed, and BAGs were resuspended in a 1 mL solution containing 880 µL H_2O, 100 µL 10× rCutSmart buffer, and 20 µL NLAIII enzyme, and then incubated at 37 °C for 1.5 h under rotation. After incubation, BAGs were washed twice with STOP-25 and once with STOP-10.

First Split-and-Pool: In this step, the UMI and the first BAG barcodes were added to the 3′ CATG overhangs through a ligation-and-extension reaction. First, BAGs were washed 2× in HBW buffer (10 mM Tris pH 8.0, 1 mM EDTA, and 0.1% v/v Tween-20). The BAG pellet was then resuspended in fresh HBW buffer to 350 µL total volume and mixed with 770 µL of 2× Quick Ligation Reaction Buffer (B2200S, New England Biolabs) and 110 µL 10 mM dNTP. An eight-channel pipette was then used to dispense 11 µL of this BAG suspension into each well of a 96-well PCR plate. 4 µL of well-specific 10 µM LNA (4 bp) primers were added into each well., The plate was briefly centrifuged, inverted and incubated for 5 min at room temperature, incubated for 5 min at 50 °C. The plates were moved onto a 4 °C cold plate and kept inverted for 10 min, and then rotated at 4 °C for 10 min. A solution on ice containing 220 µL H_2O, 302.5 µL 2× Quick Ligation Reaction Buffer, 82.5 µL Klenow DNA polymerase (3′ → 5′ exo-, M0212L, New England Biolabs), and 55 µL Quick Ligase (M2200S, New England Biolabs) was prepared and then 6 µL of that solution were dispensed to each well of the PCR plate (on cold rack) and mixed via pipette aspiration. The PCR plates were incubated under rotation at 4 °C for 10 min, then at 10 °C for 10 min, then at room temperature for 20 min, and finally at 37 °C for 20 min. The reaction was stopped by addition 100 µL STOP-25 to each well and incubation for 5 min at room temperature. Next, the bead-free BAGs ("no mag beads" condition) were separated via centrifugation (900 rcf for 2 min) while Mag-BAGs ("beads" condition) were separated for 10 min on a 96-well magnetic separation rack (12331D, Invitrogen). 85 µL of supernatant was removed from the top of each well using a multichannel pipette while carefully avoiding disruption of BAGs. The remaining volume in each well was then pooled into a solution of 5 mL of STOP-25. BAGs were then transferred to a 15 mL conical tube, washed 2× in STOP-10, and resuspended in 800 µL STOP-10.

Second Split-and-Pool: This step added a second BAG barcode to DNA strands via well-specific PCR primers, creating $96^2 = 9216$ unique barcodes to identify individual BAGs. BAGs were first washed 2× in HBW buffer. A universal primer/BAG suspension was prepared containing 10 µL of 100 µM TG primer (TGTGTTGGGTGTGTTT*G*G, Integrated DNA Technologies) and 880 µL of BAG suspension using H_2O. 9 µL of this suspension and 1 µL of 10 µM well-specific barcode primers were dispensed to each well of a 96well plate.- On ice, 10 µL NEBNext Ultra II Q5 Master Mix (M0544S, New England Biolabs) was added to each well. Samples were amplified using the following thermocycling sequence: 95 °C for 2 min, 98 °C for 15s, 12× of [98 °C for 15 s, 63 °C for 1 min, 65 °C for 1 min], 65 °C for 5 min, 4 °C infinite. Finally, the resulting amplified DNA was purified using AMPure XP magnetic beads (A63881, Bechman Coulter, Pasadena, CA, USA).

Final Sequencing Library Preparation: A final PCR reaction solution was prepared containing 20 µL NEBNext Ultra II Q5 Master Mix (M0544S, New England Biolabs), 1 µL of 10 µM "P5-TG" primer (AATGATACGGCGACCACCGAGATCTACAC GGAGATGTG TGTGTTGGGTGTGTTT*G*G, Illumina, San Diego, CA), and 1 µL of 10 µM N70x Nextera primer (FC-131-1096, Illumina) and added to the DNA samples after resuspension in H_2O to 40 µL. Samples were amplified via PCR using the following thermocycling sequence: 95 °C for 20s, 4× of [98 °C for 15s, 62 °C for 40s, 72 °C for 40s], 3× of [98 °C for 10s, 67 °C for 20s, 72 °C for 40s], 72 °C for 4 min, 4 °C infinite. After PCR, the product was purified using AMPure XP magnetic beads and then sequenced using the NovaSeq Illumina platform with a custom Read 1 sequencing primer (GGAGATGTG TGTGTTGGGTGTGTTTGG).

Genomics Data-Analysis Workflow—Mapping Reads to Human Reference Genome: The single-cell data analysis pipeline largely followed the methodology described in the previous paper.[9] Pair-end read mapping to the human reference genome hg19 using HISAT2 was performed. Reads with the same varietal tag (UMI), genomic mapping position, and cell barcode were identified as originating from the same template.

Genomics Data-Analysis Workflow—Identifying BAGs that Contain Cells: To determine which BAG tags represent cells rather than cellular debris or background diffusion, the read counts for each tag were sorted from highest to lowest, and the log of the read counts was plotted on the *y*-axis. This curve was smoothed using the loess function in *R* with the span parameter set to 0.1. The second-order finite central difference

$$\delta^2[f](x) = f(x-k) - 2f(x) + f(x+k) \tag{4}$$

was computed for each point and the point with the minimum value was selected as the threshold read count above which BAG tags were deemed to represent a cell. For SKN-1 and SK-BR-3 BAGs, *k* was set to 30 and 100, respectively.

Genomics Data-Analysis Workflow—Cell-Specific SNPs and Cell-Identity Classification: Using whole-genome-sequencing (WGS) libraries of SKN-1 and SK-BR-3, positions were examined with a sequencing depth greater than 30 in both libraries, from which the cell–source-specific SNPs were searched that satisfied the following rule: for one library, the alternative

allele frequency exceeded 50% of the reference allele frequency, while for the other library, the alternative allele frequency was 0. In all, 225037 SK-BR-3-specific SNPs and 245614 SKN-1-specific SNPs were identified. These SNPs specific to each cell source were applied to the single-cell libraries.

Genomics Data-Analysis Workflow—Downsampling Libraries for Yield Comparison: The Mag-BAG and BAG libraries had different numbers of cells and reads. To compare the efficiency of capture and tagging between libraries, the BAG tags in the Mag-BAG library were downsampled to the same number of tags as the BAG library. The reads in the BAG library were then downsampled to match the number of reads in the sampled tags in the Mag-BAG library. The two libraries used these downsampled data sets to compare reads and tags per BAG.

Genomics Data-Analysis Workflow—Comparing Genomic Insert Lengths Between the Two Libraries: To compare genomic insert length distributions in the two libraries, 350 000 read pairs from the bam files were sampled that were marked as primary mappings and proper pairs.

Genomics Data-Analysis Workflow—Procedure for Clustering and Copy Number Profile Comparison: For clustering based on copy number, 100 cells each from SKN-1 and SK-BR-3 from a single protocol were taken. The known copy number profile from a previously done whole genome analysis of the cell lines for binning was used. The constant copy number segment boundaries of the SK-BR-3 were used as bin boundaries to get 64 bins. Even though the sizes of the bins vary greatly, genomic locations where the copy number changes in any cell will occur only near the bin boundaries. Reads to each bin in each BAG were assigned, and a multinomial expectation maximization algorithm on unique read counts in bins. The details of the algorithm will be described in another manuscript at a later time. In each case, the clustering reproduced the identity determined by SNVs without assignment error. The copy number profiles for the SK-BR-3 clusters in both protocols were computed by normalizing the unique read counts in each bin by the average read counts in SKN-1 clusters in each protocol. Only minor differences in the copy number profiles derived in this manner for SK-BR-3 from both protocols are seen.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

The authors want to thank Dr. Phenix-Lan Quan for revising the manuscript. This work was supported by grants to M. Wigler from The Breast Cancer Research Foundation (BCRF-22-174) and the Simons Foundation, Life Sciences Founders Directed Giving-Research (519054), and an award from the National Science Foundation (NSF)-CBET (No. 1705578) to E. Brouzes.

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

automation, BAG-seq, droplet microfluidics, magnetic separation, single-cell genomics, split–pooling

Received: July 13, 2023
Revised: October 15, 2023
Published online:

[1] A.-C. Villani, R. Satija, G. Reynolds, S. Sarkizova, K. Shekhar, J. Fletcher, M. Griesbeck, A. Butler, S. Zheng, S. Lazo, L. Jardine, D. Dixon, E. Stephenson, E. Nilsson, I. Grundberg, D. Mcdonald, A. Filby, W. Li, P. L. De Jager, O. Rozenblatt-Rosen, A. A. Lane, M. Haniffa, A. Regev, N. Hacohen, *Science* **2017**, *356*, eaah4573.
[2] N. Navin, J. Kendall, J. Troge, P. Andrews, L. Rodgers, J. Mcindoo, K. Cook, A. Stepansky, D. Levy, D. Esposito, L. Muthuswamy, A. Krasnitz, W. R. Mccombie, J. Hicks, M. Wigler, *Nature* **2011**, *472*, 90.
[3] J. Clausell-Tormos, A. D. Griffiths, C. A. Merten, *Lab Chip* **2010**, *10*, 1302.
[4] E. Brouzes, M. Medkova, N. Savenelli, D. Marran, M. Twardowski, J. B. Hutchison, J. M. Rothberg, D. R. Link, N. Perrimon, M. L. Samuels, *Proc Natl Acad Sci U S A* **2009**, *106*, 14195.
[5] E. Z. Macosko, A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, S. A. Mccarroll, *Cell* **2015**, *161*, 1202.
[6] A. M. Klein, L. Mazutis, I. Akartuna, N. Tallapragada, A. Veres, V. Li, L. Peshkin, D. A. Weitz, M. W. Kirschner, *Cell* **2015**, *161*, 1187.
[7] R. Zilionis, J. Nainys, A. Veres, V. Savova, D. Zemmour, A. M. Klein, L. Mazutis, *Nat. Protoc.* **2017**, *12*, 44.
[8] A. R. Abate, C.-H. Chen, J. J. Agresti, D. A. Weitz, *Lab Chip* **2009**, *9*, 2628.
[9] S. Li, J. Kendall, S. Park, Z. Wang, J. Alexander, A. Moffitt, N. Ranade, C. Danyko, B. Gegenhuber, S. Fischer, B. D. Robinson, H. Lepor, J. Tollkuhn, J. Gillis, E. Brouzes, A. Krasnitz, D. Levy, M. Wigler, *Genome Res.* **2020**, *30*, 49.
[10] R. Kishore, W. Reef Hardy, V. J. Anderson, N. A. Sanchez, M. R. Buoncristiani, *J Forensic Sci* **2006**, *51*, 1055.
[11] S. A. Ware, N. Desai, M. Lopez, D. Leach, Y. Zhang, L. Giordano, M. Nouraie, M. Picard, B. A. Kaufman, *J. Biol. Chem.* **2020**, *295*, 15677.
[12] A. Chranioti, E. Aga, N. Margari, C. Kottaridi, A. Pappas, I. Panayiotides, P. Karakitsos, *Infect. Dis. Obstet. Gynecol.* **2011**, *2011*, 931281.
[13] J.-W. Kim, A. S. Utada, A. Fernández-Nieves, Z. Hu, D. A. Weitz, *Angew. Chem. Int. Ed.* **2007**, *46*, 1819.
[14] D. K. Hwang, D. Dendukuri, P. S. Doyle, *Lab Chip* **2008**, *8*, 1640.
[15] C.-H. Chen, A. R. Abate, D. Lee, E. M. Terentjev, D. A. Weitz, *Adv. Mater.* **2009**, *21*, 3201.
[16] C. Holtze, A. C. Rowat, J. J. Agresti, J. B. Hutchison, F. E. Angilè, C. H. J. Schmitz, S. Köster, H. Duan, K. J. Humphry, R. A. Scanga, J. S. Johnson, D. Pisignano, D. A. Weitz, *Lab Chip* **2008**, *8*, 1632.
[17] D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, *Nat. Biotechnol.* **2019**, *37*, 907.
[18] N. Habib, I. Avraham-Davidi, A. Basu, T. Burks, K. Shekhar, M. Hofree, S. R. Choudhury, F. Aguet, E. Gelfand, K. Ardlie, D. A. Weitz, O. Rozenblatt-Rosen, F. Zhang, A. Regev, *Nat. Methods* **2017**,*14*, 955.
[19] D. C. Duffy, J. C. Mcdonald, O. J. A. Schueller, G. M. Whitesides, *Anal. Chem.* **1998**, *70*, 4974.
[20] M. Sauzade, L. Li, T. Bakowski, H. H. Strey, E. Brouzes, *Biomicrofluidics* **2020**, *14*, 014116.
[21] R. Scanga, L. Chrastecka, R. Mohammad, A. Meadows, P.-L. Quan, E. Brouzes, *RSC Adv.* **2018**,*8*, 12960.
[22] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, *Nat. Methods* **2012**, *9*, 676.