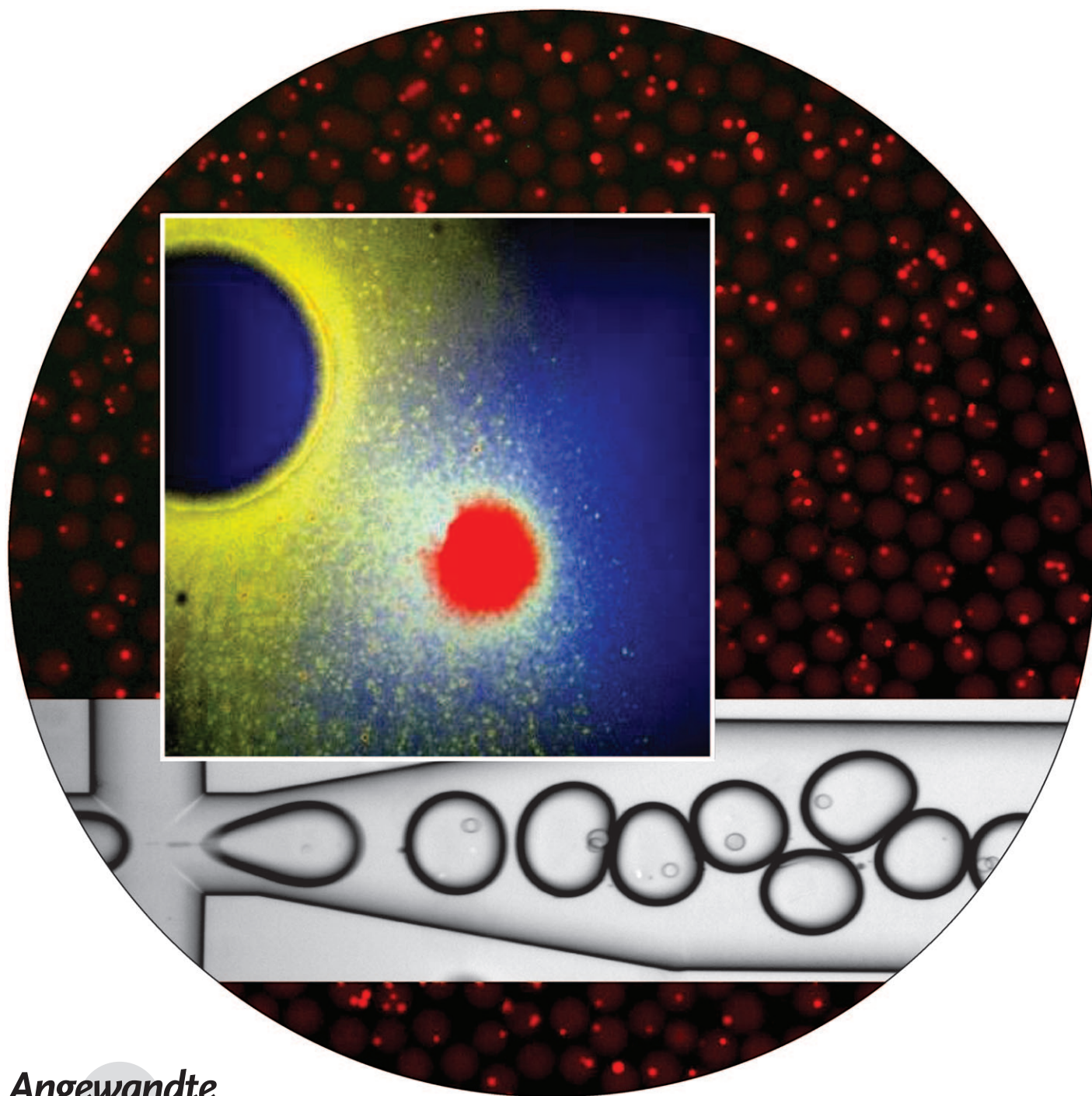


Single-Cell Multiplex Gene Detection and Sequencing with Microfluidically Generated Agarose Emulsions**

*Richard Novak, Yong Zeng, Joe Shuga, Gautham Venugopalan, Daniel A. Fletcher, Martyn T. Smith, and Richard A. Mathies**



Genetic assays, such as polymerase chain reaction (PCR), typically report on multiple cells or mixtures of genomic DNA. As a result, they cannot properly characterize the genetic heterogeneity of a cell population or detect the cooccurrence of different mutations within a single cell; these factors are key to understanding the development, progression, and treatment of cancers.^[1,2] In particular, since initial mutagenesis occurs inherently at the single-cell level, the detection and characterization of carcinogenesis will be dramatically facilitated by analytical techniques with single-cell resolution.

Cytometric sorting, limiting dilution, and micromanipulation have been previously used to perform single-cell PCR assays in 96-well PCR plates, but these approaches are not ideal for large-scale screening applications.^[3] Microfluidic technology offers fundamentally new capabilities for the manipulation of fluids, molecules, and cells that are very pertinent for the development of high-throughput single-cell analysis methods.^[4–7] Microfluidic droplet technology is particularly advantageous for single-cell/molecule analysis because it facilitates the rapid statistical compartmentalization of targets for massively parallel pico- to nanoliter-scale assays.^[8–10] In particular, microfluidic emulsion PCR (ePCR) enables high-fidelity digital single-molecule counting owing to its unique ability to ensure equal population sampling and amplification efficiencies across all reaction compartments.^[4,11,12]

To date, most reported single-cell genomic analyses have been carried out on bacterial samples.^[13–15] For mammalian cells, droplet-based genetic analyses have predominantly implemented reverse transcriptase PCR for phenotypic profiling.^[11,16] A difficulty in single-cell PCR is the persistent technical challenge of integrating a robust and scalable DNA-extraction method.^[13,14] The relative lack of suitable technologies for single-cell genomic analysis combined with the

significant genetic heterogeneity associated with cancer underscores the importance of developing new microdroplet methodologies that integrate robust single-cell genome preparation with multiplex PCR.

To address these challenges, we have developed an agarose-droplet-based platform that leverages emulsion-generator-array technology for high-throughput single-cell genetic analysis.^[4,12] Single cells were microfluidically encapsulated together with primer-functionalized beads in agarose-gel droplets for subsequent SDS lysis and proteinase K digestion to release genomic DNA. With the coencapsulated primer beads and purified genomes, we demonstrate multilocus single-cell sequencing of the control gene β -actin and the chromosomal translocation t(14;18), a mutation associated with 85–90 % of cases of follicular lymphoma.^[1,17] The coupling of our robust and high-throughput single-cell DNA-purification method with the sequencing of multiple gene targets within single cells will enable detailed studies of mutation cooccurrence and synergy during carcinogenesis.

The underlying principle of our highly parallel cell-digestion and DNA-purification method is the microfluidic encapsulation of cells in agarose droplets (Figure 1) to maintain single-genome fidelity during cell lysis and DNA purification as well as efficient multiplex emulsion PCR target amplification for subsequent analysis. Single lymphoblast cells were encapsulated along with primer-functionalized beads in 1.5 % low-melting-point agarose by using a four-channel microfluidic-emulsion-generator array (MEGA; Figure 2a).^[4] The MEGA platform is very versatile: micropump

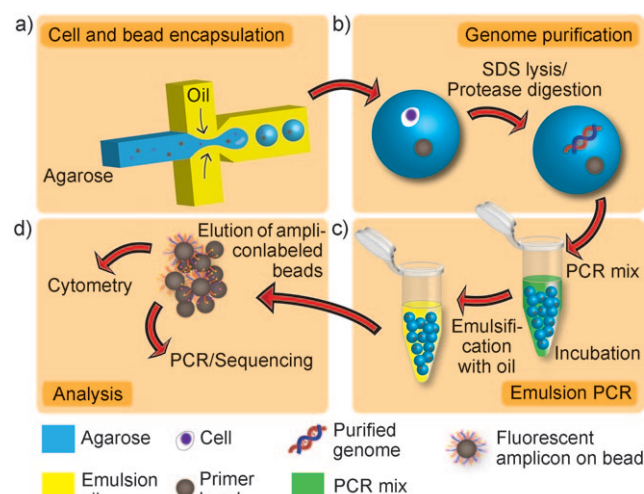


Figure 1. Workflow diagram showing the use of agarose-emulsion droplets for the genetic analysis and multilocus sequencing of single mammalian cells. a) Single cells are microfluidically encapsulated together with primer-functionalized beads in agarose-gel droplets. b) The genomes of single cells are released in the gel droplets upon SDS lysis and digestion with proteinase K according to a standard protocol. c) The agarose droplets are equilibrated in PCR buffer containing fluorescent forward primers, emulsified with oil by mechanical agitation, and thermally cycled. d) Following multiplex PCR amplification, primer beads are released by breaking the emulsion and melting the agarose. The fluorescent beads are then rapidly quantified by flow cytometry or further subjected to PCR amplification for the sequencing of target genes. SDS = sodium dodecyl sulfate.

[*] R. Novak,^[+] Dr. Y. Zeng,^[+] Dr. J. Shuga, Prof. M. T. Smith, Prof. R. A. Mathies
Center for Exposure Biology, University of California, Berkeley
307 Lewis Hall, Berkeley, CA 94720 (USA)
Fax: (+1) 510-642-3599
E-mail: ramathies@berkeley.edu

R. Novak,^[+] G. Venugopalan, Prof. D. A. Fletcher, Prof. R. A. Mathies
UC Berkeley/UC San Francisco Joint Graduate Group in Bioengineering
University of California, Berkeley (USA)

Dr. Y. Zeng,^[+] Prof. R. A. Mathies
Department of Chemistry, University of California, Berkeley (USA)
Dr. J. Shuga, Prof. M. T. Smith
School of Public Health, University of California, Berkeley (USA)

[*] These authors contributed equally.

[**] This research was supported by the trans-NIH Genes, Environment and Health Initiative, Biological Response Indicators of Environmental Systems Center Grant U54 ES016115-01 to M.T.S. and R.A.M., and by Superfund Basic Research Program NIEHS Grant P42 ES004705 to M.T.S. R.N. is supported by an NSF Graduate Research Fellowship, and J.S. is supported by the Canary Foundation and ACS Early Detection Postdoctoral Fellowship.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201006089>.

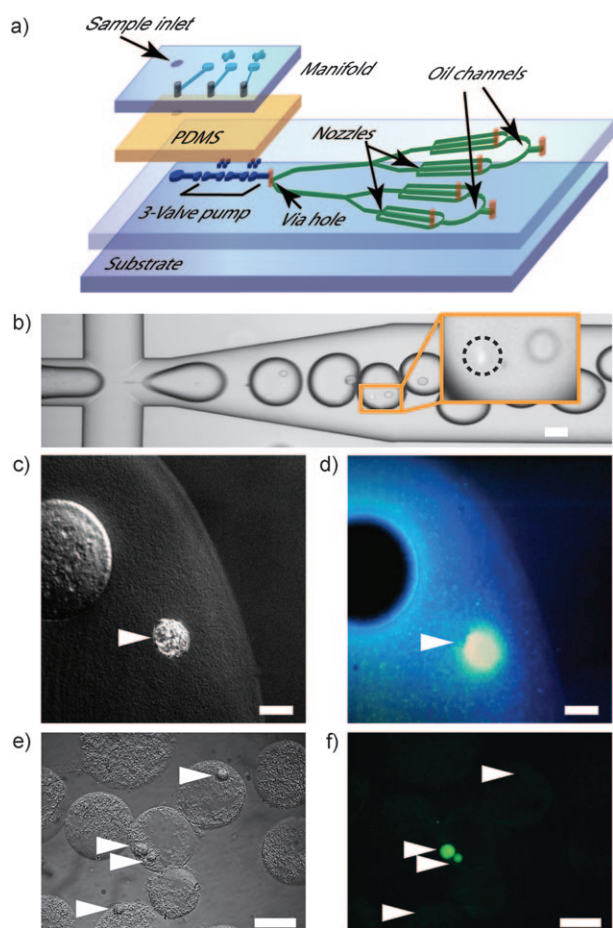


Figure 2. Microfluidic agarose encapsulation and genetic analysis of single mammalian cells. a) Four-layer glass–polydimethylsiloxane (PDMS) microfluidic-emulsion-generator array used for agarose-droplet formation. A three-valve micropump actuates an agarose–cell–bead suspension from the sample inlet toward cross-channel nozzles, where droplet generation occurs. b) Agarose-emulsion generation. Single cells are shown encapsulated along with primer-labeled beads in uniform 3 nL droplets. The inset highlights the presence of a cell (circled) and a bead in one droplet. c) Differential interference contrast (DIC) image of a cell (arrow) at the edge of an agarose droplet after protease digestion. d) Projection of the confocal micrograph in (c) showing DNA stained with acridine orange confined in the agarose after release from the cell (arrow). e) Following PCR amplification of the chromosomal translocation $t(14;18)$ and β -actin, agarose droplets containing beads (arrows) remained intact and did not merge. f) Epifluorescence micrograph of the sample in (e) showing that primer-labeled beads (arrows) become fluorescent if a $t(14;18)^+$ RL cell is present in the same droplet, but otherwise remain dark. Scale bars are 100 μm in (b), (e), and (f) and 10 μm in (c) and (d).

actuation was optimized to account for the increased viscosity of molten agarose without modification of the microfluidic design. Droplet generation at approximately 40 °C for 30 min resulted in the encapsulation of approximately 18000 cells at up to 0.3 cells per droplet (cpd) on average. Figure 2b shows the generation of uniform 3 nL agarose droplets containing primer-functionalized beads. The inset highlights an example of cell and bead coencapsulation in a single droplet.

Encapsulation in agarose droplets enables reproducible single-cell DNA extraction and isolation through the adapta-

tion of robust DNA-purification protocols. An example of an agarose droplet in which a bead and a cell were coencapsulated is shown in Figure 2c. Cell lysis and digestion of the DNA-binding histone proteins upon overnight incubation of the gel droplets in SDS lysis buffer containing proteinase K led to DNA release. The void left by the cell in the agarose was occupied by brightly fluorescent genomic DNA, which exhibited minimal diffusion into the surrounding gel (Figure 2d) as a result of the relatively small pore size of 1.5 % agarose (ca. 130 nm).^[18] An incubation temperature of 52 °C facilitated enzymatic protein digestion while preserving the integrity of the agarose droplets. By staining with propidium iodide, we were able to visualize single high-molecular-weight DNA strands protruding from the relatively small agarose droplets (see Figure S1D in the Supporting Information). This result indicated that the majority of nuclear proteins were removed by the combination of proteinase K and SDS. The agarose droplets were stable for at least one week when stored in ethanol at 4 °C, as determined by confocal imaging of DNA diffusion radii.

A key benefit of agarose encapsulation is the ability to mechanically manipulate the isolated genomic DNA without mixing the genetic content of different cells. Agarose droplets equilibrated with PCR mix were reemulsified by mechanical agitation in dispersing oil to produce uniform nanoliter-droplet “reactors” for massively parallel single-cell PCR analysis. Excess PCR mix produces microfines (emulsions less than 1 μm in diameter), which enhance emulsion stability during thermal cycling.^[19] The agarose droplets melt during the hot start phase of PCR and remain liquid throughout the amplification process, maximizing reagent and amplicon diffusion rates. We varied the initial heating rate and tested various concentrations of Triton X-100 as well as combinations of Abil em90 and Span 80 detergents in oil. A slow temperature ramp profile (0.1 °C s⁻¹) resulted in improved short-term stability, and the addition of BSA (4 mg mL⁻¹) to the PCR mix and 0.8 % Triton X-100 to the emulsion oil minimized droplet merger over the course of PCR thermal cycling (Figure 2e). Fluorescently labeled amplicons bound to the primer-functionalized beads could be seen inside the agarose droplets following PCR (Figure 2f). The 34 μm cross-linked beads were selected for their ability to amplify targets exceeding 1 kb in amounts of at least 100 amol per bead.^[12] The absence of fluorescence from beads in droplets without genomes indicated that genomic targets were not transferred between agarose droplets.

To demonstrate highly parallel genotyping with single-cell resolution, we performed a multiplex PCR assay of cancer cells harboring the $t(14;18)$ translocation at various mutant-to-wild type (RL/TK6) cell ratios. Labeling of the control gene product, β -actin, with the cyanine dye Cy5 enabled the quantification of total cell frequency, whereas the translocation $t(14;18)$ product labeled with carboxyfluorescein (FAM) spanned the *bcl-2* and *IgH* genes across their breakpoint regions and could be used to determine mutation frequency. A representative flow cytometric profile of beads following multiplex PCR amplification with 50 % RL cells at an average cell frequency of 0.3 cpd (Figure 3a) demonstrated distinct populations of negative beads, positive beads with Cy5-

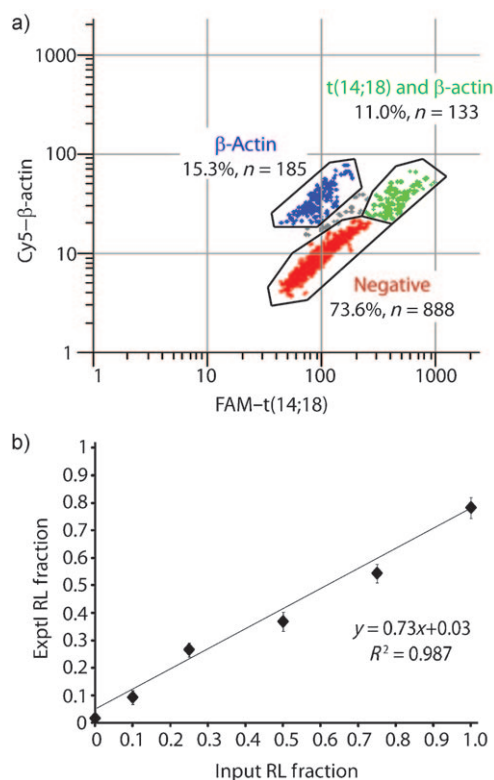


Figure 3. High-throughput digital genetic analysis of cancer cells. a) Representative flow cytometry plot and gated populations of beads from a sample containing 50% RL lymphoblast cells ($n = 1206$) at a frequency of 0.3 cells per droplet. The observed fraction of total positive events (26.3%) compares favorably with an expected frequency of 25.9% based on the Poisson distribution. b) Standard curve for the detection of $t(14;18)^+$ RL cells from mixtures of mutant RL and wild-type TK6 cells. Experimental quantification by multiplex emulsion PCR and flow cytometry exhibits a linear response between the 10 and 100% mutant-cell frequencies tested.

labeled β -actin only, and double-positive beads with FAM-labeled $t(14;18)$ and Cy5-labeled β -actin. Beads containing $t(14;18)$ only were never observed, which further indicates the conservation of single-genome integrity during cell lysis and PCR. By maintaining a constant total cell density in the single-cell regime (0.1–0.3 cpd on average) and varying the relative concentrations of mutant RL and wild-type TK6 cells, we generated a standard curve (Figure 3b) to confirm that amplification originated from single cells. In this stochastic regime, the linearity ($r = 0.993$) of the measured concentration of mutant RL cells with respect to the input in the 0–100% RL-cell-frequency range tested indicated successful genetic analysis of single cells. Importantly, in the subset of samples further tested, the ratio of total amplicon-positive bead frequency determined by flow cytometry to the microscopically observed cell-encapsulation frequency indicated high PCR efficiency ($113 \pm 24\%$).

For the single-cell sequencing of both target gene loci, the template bound to single beads diluted to the stochastic limit in a standard 96-well plate was reamplified, and the products were separated from individual wells by gel electrophoresis.

Gel analysis further confirmed amplification from single genomes, and frequencies of β -actin single-positive to double-positive events matched flow cytometry frequency results in the samples tested. Fluorescence-based sorting can be applied to amplicon-labeled beads to sequence only populations of interest and thereby increase efficiency. Size-separated products from single beads were excised from the gel for sequencing. The expected $t(14;18)$ and β -actin sequences were recovered (Figure 4; see also Figures S2 and S3 in the Supporting Information). The random nucleotide insertion sequence in $t(14;18)$ matched the unique translocation “fingerprint” determined by sequencing of RL cells in bulk.^[17]

The reamplification step enables the integration of single-cell PCR with standard molecular-biology techniques for the sequencing of multiple genetic loci in individual cells at a rate required for meaningful population analysis. Using the single-bead reamplification and size-separation approach, it is possible to sequence multiple target genes from hundreds of single cells in a single experiment and to perform meaningful statistical analysis of gene-sequence variation in a cell population. Furthermore, the number of target loci per cell is limited only by the ability to excise individual bands from a gel, provided the multiplex PCR has been validated. Adaptation of our agarose cell-encapsulation and emulsion PCR method to next-generation sequencing technologies has the potential to provide additional gains in throughput for single-cell sequencing. Although alternative single-cell-resolution approaches to genetic analysis, such as fluorescence in situ hybridization^[20,21] and in situ PCR,^[22,23] are utilized for investigating cell-mutation progression, the throughput offered by these methods is limited, and the lack of amplicon recovery prevents the detection of unknown mutations. In the case of follicular lymphoma, identification of the breakpoint sequence along the IgH and bcl-2 genes can help in the elucidation of the mechanisms of erroneous genetic recombination that cause the translocation $t(14;18)$, whereas the random insertion sequence between the two chromosomes is a unique identifier of distinct mutation events.^[17]

Encapsulation enables robust parallel cell lysis and DNA purification of cell types that are difficult to screen by using other single-cell PCR methods. The method can be adapted in future applications to encapsulate live single cells for cell culture and subsequent analysis of clonal populations.^[24] Although a similar approach was demonstrated recently for the screening and amplification of alginate-encapsulated *Escherichia coli* colonies containing plasmid libraries,^[24] our approach enables the robust genome purification of mammalian cells, requires 1% of the reagent volume for emulsion PCR, does not require droplet sorting, and maintains single-cell segregation throughout all downstream analyses as a result of the incorporation of primer-functionalized beads as amplicon substrates. Furthermore, single-cell sequencing techniques generally involve cell types that are amenable to lysis during PCR and the sorting of single cells into 96-well PCR plates.^[25,26] However, PCR amplification directly from single cells has been hindered by the lack of a DNA-purification step that would remove histones and other nuclear components that inhibit polymerase activity.^[25,27] In our approach, the incorporation of a highly parallel single-

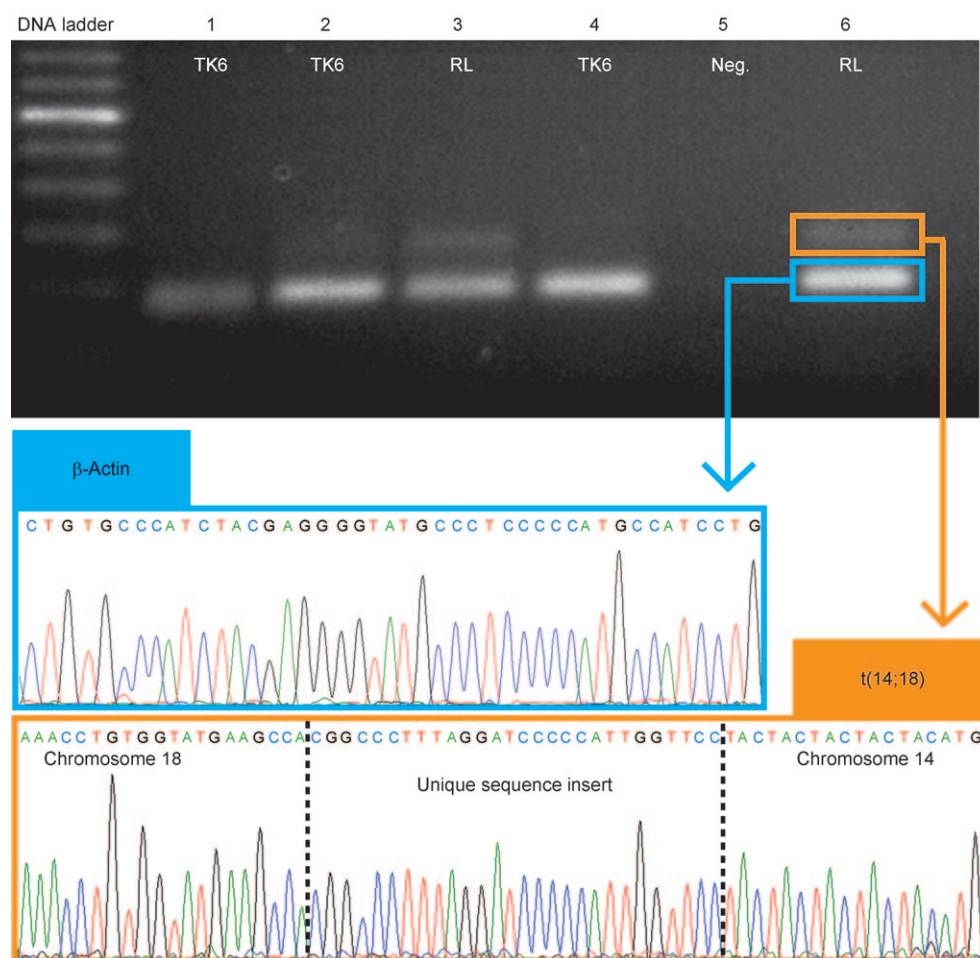


Figure 4. Multiplex single-cell sequencing. Two loci from DNA products isolated from single cells were sequenced by the reamplification of individual amplicon-labeled beads in separate PCR reactions and the separation of the two products by gel electrophoresis. This representative image of a section of the gel from a 50% RL cell sample shows amplified t(14;18) and β -actin from single beads (lanes 3 and 6), β -actin originating from TK6 cells (lanes 1, 2, and 4), the lack of β -actin and t(14;18) from negative beads (lane 5), and a 100 bp DNA ladder. A lack of t(14;18) amplification without β -actin indicates true single-cell analysis. Both products were excised and sequenced to give sequence data from two genes colocalized in a single cell. The breakpoint locations on chromosomes 14 and 18 as well as the unique insertion sequence confirmed the recovery of the expected RL lymphoblast sequence. By integrating single-cell emulsion PCR with a standard sequencing technique, it is possible to sequence multiple genes from hundreds of single cells in 2–3 days. Thus, this strategy opens up the possibility of population-level analyses of genetic heterogeneity.

cell-lysis and DNA-purification step resulted in near 100 % amplification efficiency. Finally, the ability to purify genomic DNA from single cells in a supporting matrix opens up the possibility of single-cell epigenetic analysis, such as the methylation-specific PCR of gene regulatory sequences following conversion with bisulfite.

Cancer is an evolving disease driven by genetic instability, which results in a constant clonal divergence of the tumor cell population. The accumulation of mutations in the genes coding for cellular pathways plays an important role in carcinogenesis, metastasis, and therapeutic resistance.^[28–30] Significant cellular heterogeneity may therefore exist in tumors; this heterogeneity changes dynamically at different stages of disease progression.^[21] We have developed a robust agarose-microdroplet method for detecting and sequencing

multiple genetic targets from single cells in a scalable manner. Whereas oncogene mutations, such as TP53, CDKN2A, CDKN2B, and bcl-6, have been demonstrated to correlate individually with poor prognosis in hematopoietic cancers,^[31–35] the ability to sequence multiple targets within single cells with high throughput will facilitate investigation of the synergistic effects of mutation cooccurrence and their impact on disease progression and treatment. Furthermore, the screening of large cell populations will uncover potential tumor heterogeneity at the single-nucleotide level that is otherwise obscured by the ensemble average. By overcoming the many problems associated with single-cell genetic analysis,^[36,37] including nonuniform cell lysis, DNA release and amplification, and low-throughput, our single-cell agarose encapsulation technology provides the throughput, robustness, and resolution required for probing the stochastic mechanisms of carcinogenesis, progression, and response to chemotherapy.

Received: September 29, 2010
Published online: December 3, 2010

Keywords: cancer mutations · gene sequencing · microemulsions · microfluidics · single-cell analysis

- [1] F. d'Amore et al., *Clin. Canc. Res.* **2008**, *14*, 7180–7187.
- [2] S. Riethdorf, H. Wikman, K. Pantel, *Int. J. Cancer* **2008**, *123*, 1991–2006.
- [3] B. Vogelstein, K. W. Kinzler, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9236–9241.
- [4] Y. Zeng, R. Novak, J. Shuga, M. T. Smith, R. A. Mathies, *Anal. Chem.* **2010**, *82*, 3183–3190.
- [5] H. Andersson, A. V. D. Berg, *Curr. Opin. Biotechnol.* **2004**, *15*, 44–49.
- [6] J. El-Ali, P. K. Sorger, K. F. Jensen, *Nature* **2006**, *442*, 403–411.
- [7] N. M. Toriello, E. S. Douglas, N. Thaitrong, S. C. Hsiao, M. B. Francis, C. R. Bertozzi, R. A. Mathies, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20173–20178.
- [8] A. Huebner, S. Sharma, M. Srisa-Art, F. Hollfelder, J. B. Edel, A. J. deMello, *Lab Chip* **2008**, *8*, 1244–1254.
- [9] Y. Schaerli, F. Hollfelder, *Mol. Biosyst.* **2009**, *5*, 1392.
- [10] J. Clausell-Tormos, D. Lieber, J. Baret, A. El-Harrak, O. J. Miller, L. Frenz, J. Blouwolff, K. J. Humphry, S. Köster, H. Duan, *Chem. Biol.* **2008**, *15*, 427–437.
- [11] N. Beer, E. Wheeler, L. Lee-Houghton, N. Watkins, S. Nasarabadi, N. Hebert, P. Leung, D. Arnold, C. Bailey, B. Colston, *Anal. Chem.* **2008**, *80*, 1854–1858.

- [12] P. Kumaresan, C. J. Yang, S. A. Cronier, R. G. Blazej, R. A. Mathies, *Anal. Chem.* **2008**, *80*, 3522–3529.
- [13] T. Chao, A. Ros, *J. R. Soc. Interface* **2008**, *5*, S139–S150.
- [14] R. N. Zare, S. Kim, *Annu. Rev. Biomed. Eng.* **2010**, *12*, 187–201.
- [15] E. A. Ottesen, J. W. Hong, S. R. Quake, J. R. Leadbetter, *Science* **2006**, *314*, 1464–1467.
- [16] M. M. Kiss, L. Ortoleva-Donnelly, N. R. Beer, J. Warner, C. G. Bailey, B. W. Colston, J. M. Rothberg, D. R. Link, J. H. Leamon, *Anal. Chem.* **2008**, *80*, 8975–8981.
- [17] C. M. McHale et al., *J. Natl. Cancer Inst. Monogr.* **2008**, *2008*, 74–77.
- [18] J. Xiong, J. Narayanan, X. Liu, T. K. Chong, S. B. Chen, T. Chung, *J. Phys. Chem. B* **2005**, *109*, 5638–5643.
- [19] M. Margulies et al., *Nature* **2005**, *437*, 376–380.
- [20] J. Hicks et al., *Cold Spring Harbor Symp. Quant. Biol.* **2005**, *70*, 51–63.
- [21] N. Navin et al., *Genome Res.* **2010**, *20*, 68–80.
- [22] P. M. Leppänen, J. Koponen, M. P. Turunen, T. Pakkanen, S. Ylä-Herttuala, *J. Gene Med.* **2001**, *3*, 173–178.
- [23] P. M. Mulrooney, T. I. Michalak, *J. Virol.* **2003**, *77*, 970–979.
- [24] M. Walser, R. Pellaux, A. Meyer, M. Bechtold, H. Vanderschuren, R. Reinhardt, J. Magyar, S. Panke, M. Held, *Nucleic Acids Res.* **2009**, *37*, e57.
- [25] L. Zhang, X. Cui, K. Schmitt, R. Hubert, W. Navidi, N. Arnheim, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 5847–5851.
- [26] *Current Protocols in Molecular Biology* (Eds.: F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. Seidman, J. A. Smith, K. Struhl), Wiley, Hoboken, NJ, **2001**.
- [27] D. Goldenberger, I. Perschil, M. Ritzler, M. Altwegg, *Genome Res.* **1995**, *4*, 368–370.
- [28] D. Hanahan, R. A. Weinberg, *Cell* **2000**, *100*, 57–70.
- [29] L. Vermeulen, M. Todaro, F. de Sousa Mello, M. R. Sprick, K. Kemper, M. Perez Alea, D. J. Richel, G. Stassi, J. P. Medema, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 13427–13432.
- [30] N. E. Navin, J. Hicks, *Mol. Oncol.* **2010**, *4*, 267–283.
- [31] A. I. Robles, C. C. Harris, *Cold Spring Harbor Perspect. Biol.* **2010**, DOI: 10.1101/cshperspect.a001016.
- [32] U. Vitolo et al., *Leukemia* **2002**, *16*, 268–275.
- [33] K. H. Young et al., *Blood* **2008**, *112*, 3088–3098.
- [34] I. S. Lossos, R. Levy, *Semin. Cancer Biol.* **2003**, *13*, 191–202.
- [35] D. O'Shea et al., *Blood* **2008**, *112*, 3126–3129.
- [36] R. B. Brown, J. Audet, *J. R. Soc. Interface* **2008**, *5*, S131–138.
- [37] H. Huang, Y. Bu, G. Zhou, *World J. Gastroenterol.* **2006**, *12*, 3814–3820.