CANCER

Delineating cancer evolution with single-cell sequencing

Nicholas E. Navin[1,2,3*]

Single-cell sequencing methods are revolutionizing cancer research and medicine by providing powerful tools to investigate intratumor heterogeneity and rare subpopulations.

Tumors evolve from single cells. During this evolutionary process, cancer cells diverge and form distinct lineages, resulting in intratumor heterogeneity. This clonal diversity is likely to play an important role during cancer progression, when the population of tumor cells encounters selective pressures in the tumor microenvironment, including the immune system, hypoxia, nutrient deprivation, geographical barriers, pH changes, and chemotherapy. However, there is currently a large gap in knowledge regarding our understanding of clonal diversity and its role in important biological processes that occur in cancer, such as invasion, clonal evolution, metastasis, and evolving resistance to therapy. Our poor understanding of clonal diversity stems from the fact that we cannot accurately measure it. Standard genomic techniques such as microarrays and next-generation sequencing methods require micrograms of DNA or RNA as input material and thus report an average signal from a complex population of cells. These methods average together mutations from multiple subpopulations and mask the presence of rare tumor clones that may play an important role in tumor progression. To address this problem, we pioneered the development of a single-cell sequencing (SCS) method to study cancer progression (*1*), which helped to establish the field of single-cell genomics. This field has shown tremendous growth over the past four years, due to the myriad of applications in cancer research and biomedicine.

INTRATUMOR HETEROGENEITY

My interest in intratumor heterogeneity took root when I was a graduate student at the Cold Spring Harbor Laboratory, working under the mentorship of Michael Wigler and James Hicks. During this time, we discovered the existence of normal copy number variants (CNVs) in the human population (*2*), and I became enthralled with the question of whether genomic diversity existed within tumor cell populations. However, at this time, the tools needed to study this problem did not exist. Thus, I spent my graduate career developing a method that combined macrodissection of spatial regions in tumors and flow-sorting by DNA ploidy to isolate tumor subpopulations and study aneuploidy evolution (*3*). Although these data on tumor subpopulations were informative, they still reflected a complex admixture of tumor cell genomes. To overcome this problem, I developed a computational algorithm (PROBER) to design tiling-oligonucleotide FISH probes (*4*) and applied it to study the clonal substructure of breast tumors at single-cell resolution. This approach could resolve intratumor heterogeneity at single-cell resolution, but it was limited to reporting copy number aberrations at targeted regions in the genome. Therefore, it was not possible to accurately reconstruct tumor evolution, which required a large number of genomic markers. It soon became clear that genome-wide single-cell sequencing data were needed, despite the formidable technical challenge it presented at the time.

SINGLE-CELL GENOMICS

As a postdoctoral student, I overcame this challenge by developing a single-cell genome sequencing method (*1*). This study was important for two reasons: (i) it demonstrated the technical feasibility of sequencing the genome of a single cell, and (ii) it showed that single-cell genomic data could be used to reconstruct tumor evolution. The method we developed was called single-nucleus sequencing (SNS), and it combined flow-sorting, whole-genome amplification, and next-generation sequencing to generate genome-wide datasets from single cancer cells. We applied SNS to study copy number changes in breast tumors, revealing a punctuated model of chromosome evolution, in which hundreds of rearrangements occurred in short bursts, followed by stable clonal expansions to form the tumor mass. These data challenged the paradigm of gradual evolution (*5*, *6*), which posited that copy number changes accumulated gradually and sequentially over time, leading to more malignant stages of cancer.

After this work, I established an independent research laboratory at the MD Anderson Cancer Center. Our group focuses on applying SCS methods to study clonal evolution and diversity in the context of invasion, metastasis, and resistance to chemotherapy in breast cancer. However, the initial difficulty we faced was that these processes were driven by point mutations and indels, which required high-coverage single-cell data. The technical hurdle was that the original SNS method was limited to generating about 10% physical coverage of a single cell's genome, which was sufficient for measuring large-scale (54 kb) copy number changes but insufficient for resolving mutations at base-pair resolution. To overcome this challenge, we developed a method called nucleus sequencing (Nuc-Seq), which combines single-cell flow-sorting, multiple-displacement-amplification, tagmentation, and next-generation sequencing to achieve high-coverage (>90%) data from single mammalian cells (*7*, *8*).

We applied Nuc-Seq to study several important questions in cancer biology. First, we investigated the mutator phenotype hypothesis, which posits that cancers are driven by increased mutation rates (*9*). The question we posed is whether most human cancers show an increased number of mutations after each cell division, or alternatively whether tumors accumulate large frequencies of mutations through increased cell proliferation at a low or normal error rate (about 1 error per cell division). By sequencing the genomes of single breast tumor cells, we showed that triple-negative breast cancer patients had increased mutation rates (mean 13X) compared to normal breast cells, but estrogen receptor positive patients did not. Furthermore, the single-cell data revealed that breast tumors harbor hundreds of ultra-rare mutations (URMs) that occur at low frequencies (<1%) in the tumor mass. In addition, these data clearly showed that no single tumor cells are genetically identical

[1]Department of Genetics, University of Texas, MD Anderson Cancer Center, Houston, TX 77030, USA. [2]Department of Bioinformatics and Computational Biology, University of Texas, MD Anderson Cancer Center, Houston, TX 77030, USA. [3]Graduate Program in Genes and Development, Graduate School of Biomedical Sciences, University of Texas Health Science Center at Houston, Houston, TX 77030, USA.
*E-mail: nnavin@mdanderson.org

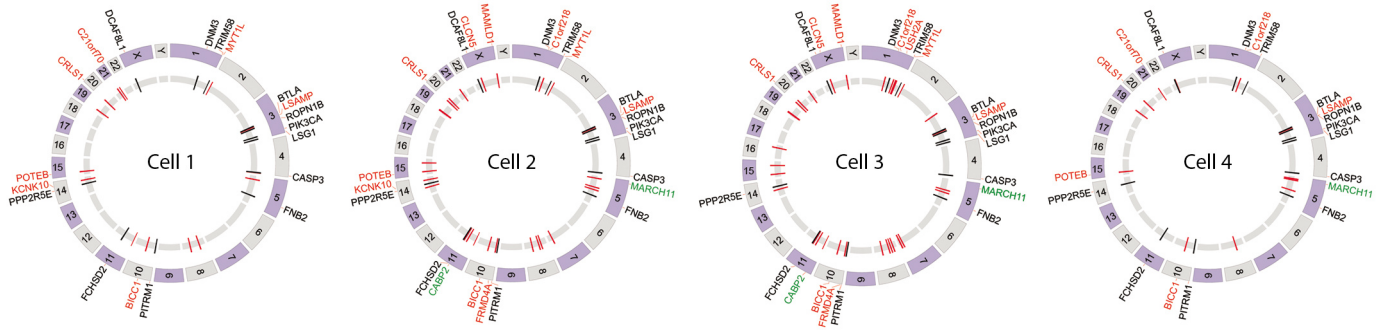

Fig. 1. No tumor cells are identical. Whole-genome single-cell sequencing of four tumor cells from an estrogen receptor–positive breast cancer patient shows that no two tumor cells are genetically identical. Clonal mutations are shown in black, subclonal mutations in intergenic regions are colored in red, and subclonal mutations in exonic regions are shown in green.

(Fig. 1). By directly comparing copy number changes and point mutations, we discovered that two distinct molecular clocks were operating during tumor evolution. Copy number aberrations occurred early in punctuated bursts, followed by stable clonal expansions, whereas point mutations were acquired gradually over time, resulting in extensive genomic diversity.

Our current research is moving beyond primary tumors, to understand the role of clonal diversity in invasion, metastasis, and chemoresistance evolution in breast cancer. In early-stage breast cancers, such as ductal carcinoma in situ (DCIS), we are using SCS methods to identify rare clones that escape the ducts and invade the surrounding tissues. We are also applying SCS methods to study the key intermediates of metastasis, circulating tumor cells (CTCs), with the goal of understanding their genetic relationship to the primary and metastatic tumors. To study the evolution of chemoresistance, we are using SCS methods to determine if rare chemoresistant subclones are preexisting in the tumor mass, or alternatively, are induced in response to the therapeutic agent.

CONCLUSIONS

SCS methods will have numerous translational applications in the clinic, particularly in the area of early detection, prognostics, noninvasive monitoring, and guiding targeted therapy (*10*). For early detection, SCS methods provide a powerful approach to detect malignant tumor cells in scarce clinical samples, such as blood, sputum, urine, fecal matter, or ductal fluids that normally cannot be profiled by genomic methods. An important prognostic application will be to measure a diversity index of each patient's tumor, which is expected to predict whether a patient will show poor response to therapy or poor survival. Perhaps the most immediate clinical application is noninvasive monitoring, in which single-cell sequencing of CTCs can provide a "liquid biopsy" of the primary and metastatic tumors to monitor therapy response. SCS can also improve therapeutic targeting and personalized medicine by delineating the clonal substructure of a tumor and identifying molecular targets that are present in the majority of tumor cells or, alternatively, in the most malignant subpopulations.

In summary, our development and application of single-cell sequencing technologies has helped to establish an emerging field of biology (single-cell genomics), which has had a major impact on cancer research and medicine. Our studies have illuminated the role of genomic diversity in primary tumors and led to the discovery of punctuated copy number evolution, mutator phenotypes, and rare subclonal mutations in breast cancer patients. By translating these tools into the clinic in the coming years, we anticipate that our work will have a large impact on reducing morbidity and improving the quality of life for cancer patients.

REFERENCES AND NOTES

1. 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, Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
2. J. Sebat, B. Lakshmi, J. Troge, J. Alexander, J. Young, P. Lundin, S. Månér, H. Massa, M. Walker, M. Chi, N. Navin, R. Lucito, J. Healy, J. Hicks, K. Ye, A. Reiner, T. C. Gilliam, B. Trask, N. Patterson, A. Zetterberg, M. Wigler, Large-scale copy number polymorphism in the human genome. *Science* **305**, 525–528 (2004).
3. N. Navin, A. Krasnitz, L. Rodgers, K. Cook, J. Meth, J. Kendall, M. Riggs, Y. Eberling, J. Troge, V. Grubor, D. Levy, P. Lundin, S. Månér, A. Zetterberg, J. Hicks, M. Wigler, Inferring tumor progression from genomic heterogeneity. *Genome Res.* **20**, 68–80 (2010).
4. N. Navin, V. Grubor, J. Hicks, E. Leibu, E. Thomas, J. Troge, M. Riggs, P. Lundin, S. Månér, J. Sebat, A. Zetterberg, M. Wigler, PROBER: Oligonucleotide FISH probe design software. *Bioinformatics* **22**, 2437–2438 (2006).
5. E. R. Fearon, B. Vogelstein, A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
6. M. Höglund, D. Gisselsson, G. B. Hansen, T. Säll, F. Mitelman, Multivariate analysis of chromosomal imbalances in breast cancer delineates cytogenetic pathways and reveals complex relationships among imbalances. *Cancer Res.* **62**, 2675–2680 (2002).
7. Y. Wang, J. Waters, M. L. Leung, A. Unruh, W. Roh, X. Shi, K. Chen, P. Scheet, S. Vattathil, H. Liang, A. Multani, H. Zhang, R. Zhao, F. Michor, F. Meric-Bernstam, N. E. Navin, Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
8. M. L. Leung, Y. Wang, J. Waters, N. E. Navin, SNES: Single nucleus exome sequencing. *Genome Biol.* **16**, 55 (2015).
9. L. A. Loeb, Human cancers express mutator phenotypes: Origin, consequences and targeting. *Nat. Rev. Cancer* **11**, 450–457 (2011).
10. N. E. Navin, Cancer genomics: One cell at a time. *Genome Biol.* **15**, 452 (2014).

Acknowledgments: I thank Martin Wachtel, Rose Wachtel, and AAAS for awarding me the 2015 AAAS Martin and Rose Wachtel Cancer Research Award. **Funding:** N.E.N. is a Nadia's Gift Foundation Damon Runyon-Rachleff Innovator (DRR-25-13) and is supported by the Lefkofsky Family Foundation. This work is supported by grants to N.E.N. from the National Cancer Institute (1RO1CA169244-01) and the MD Anderson Cancer Center (Moonshot Knowledge Gap and the Center for Genetics & Genomics).

10.1126/scitranslmed.aac8319

Citation: N. E. Navin, Delineating cancer evolution with single-cell sequencing. *Sci. Transl. Med.* **7**, 296fs29 (2015).

Science Translational Medicine

Delineating cancer evolution with single-cell sequencing

Nicholas E. Navin

Sci Transl Med **7**, 296fs29296fs29.
DOI: 10.1126/scitranslmed.aac8319