VIEWPOINT

SCIENTIFIC DISCOVERY AND THE FUTURE OF MEDICINE

Single Molecules Meet Genomics
Pinpointing Precision Medicine

Xiaoliang Sunney Xie, PhD
Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts; and Biodynamic Optical Imaging Center, Peking University, Beijing, China.

Corresponding Author: Xiaoliang Sunney Xie, PhD, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford St, Cambridge, MA 02138 (xie@chemistry.harvard.edu).

Since the 1990s, advances in single-molecule imaging and manipulation have led to the emergence of single-molecule biology (ie, the probing and understanding of biological behaviors on a single-molecule basis). Due to contributions from many laboratories around the world, this has changed the way many biological problems are addressed and generated much new knowledge.

One early example of single-molecule biology is the study of turnovers of the enzyme cholesterol oxidase.[1] The enzyme contains a flavin moiety that is naturally fluorescent in its oxidized form, but not in its reduced form. Each on/off cycle of fluorescence emission corresponds to an enzymatic cycle, enabling the observation of enzymatic reactions in real time. On a single-molecule basis, a chemical reaction occurs in a stochastic way (ie, waiting time for a chemical reaction to occur is random) rather than deterministic as is the case of a large number of molecules. The ability to observe a single-molecule chemical reaction in real time is important because many biomolecules, such as DNA, exist as single molecules in live cells.

Single-molecule enzymology also has practical implications. Single-molecule sequencers can directly read a DNA sequence by monitoring 1 DNA polymerase molecule incorporating individual fluorescently labeled nucleotides into a single-stranded DNA template.[2] The single-molecule sequencers offer the unique property of long-sequencing lengths, although they are not (yet) competitive in cost, accuracy, and throughput compared with bulk sequencers. It was the emergence of the next-generation bulk sequencers (since 2007) rather than the single-molecule sequencers that created the new era of personalized medicine.

On the fundamental side, however, single-molecule biology has yielded understanding of the workings of many macromolecular systems, not only in vitro but also in live cells. Because the copy number of a gene in a cell can be 1 or 2, gene expression takes place stochastically. Consequently, single-molecule biology is germane to single-cell biology. Extensive live cell studies have been performed to monitor the stochastic production of messenger RNA and protein one at a time, allowing the central dogma of molecular biology to be quantitatively described at the single-molecule level.[3]

The fact that DNA exists as single molecules (chromosomes) in a human cell also means that changes in the genome occur stochastically. Accordingly, every germ cell of an individual is different because of recombination, and cancer cells in a primary tumor are highly heterogeneous because of drastic genomic variations. Most common genomic changes in a cell include single-nucleotide variations, which are single base-pair mutations, and copy-number variations. One single-nucleotide variation in the 6-billion base-pair human genome can cause a hereditary disease. The copy number of a particular gene in a human germ cell is 1, and in a somatic cell is 2 (1 from each parent). The copy-number variations usually arise from chromosome segregation errors in germ cells and from genomic rearrangements, including insertions, deletions, inversions, and translocations in cancer cells. Because single-nucleotide variations and copy-number variations are not synchronized in different cells, every cell has a different genome. Hence single-cell sequencing is necessary, but had not been feasible until the recent advances in single-cell genomics.[4]

To sequence the genome of a single human cell requires its whole-genome amplification. The newly developed whole-genome amplification method of multiple annealing– and looping-based amplification cycles (MALBAC)[5] offers a more accurate copy-number variation and single-nucleotide variation detection than the widely used multiple displacement amplification. In a human cell, MALBAC allows digital counting of the copy number of a gene, and can detect 1 single-nucleotide variation with no false-positive results but certain false-negative results.[5]

Examples of recent applications of MALBAC involve in vitro fertilization and cancer diagnosis, both of which have important implications for precision medicine. Assisting in vitro fertilization, MALBAC has been applied to preimplantation genomic screening and preimplantation genetic diagnosis to avoid chromosome abnormalities and the inheritance of monogenic diseases. Chromosome abnormality (ie, copy-number variation at the chromosome scale) is a major cause of miscarriages as well as genetic disorders such as Down syndrome. Likewise, approximately 7000 monogenic hereditary diseases can seriously affect human health and are associated with heavy burdens both for the affected families and the health care system.

Although polymerase chain reaction, fluorescent in situ hybridization, and DNA microarray have been previously used for preimplantation genetic diagnosis and preimplantation genomic screening, MALBAC offers the advantage of higher accuracy and simultaneous avoidance of chromosome abnormality and point mutations.[6] In 2014, we demonstrated the first successful "MALBAC babies," who were free of the monogenic disease carried by either parent.[7] This brings the exciting news to couples with hereditary disorders that

they will now be able to stop passing on potentially devastating mutations to their offspring with a high success rate.

In this case, the extremely low rate of false-positive and false-negative results of the MALBAC procedure has met the compelling desire of parents to avoid the transmission of serious heterozygous mutations in their children. On the other hand, the option for parents to reduce the propensity for certain diseases in their children (such as breast cancer), by selecting a less probable allele (with specific *BRCA* gene sequence) with a similar procedure, although technically feasible, should call for conscious discussions about ethical concerns.

Cancer is a genomic disease.[8] MALBAC also has been used to sequence circulating tumor cells, which enter peripheral blood and seed metastases that account for 90% of cancer-related deaths. In analyzing single circulating tumor cells from patients with lung cancer, it has been discovered that, unlike the highly heterogeneous point mutations, the copy-number variation patterns of the circulating tumor cells are similar within a patient and within different patients of the same cancer type, but are distinctly different among different types of cancers. Furthermore, the reproducible copy-number variation pattern of circulating tumor cells is the same as that for the individual's metastatic tumor.[9] It became apparent that gains and losses in copy number of certain chromosome regions are selected for metastases. The finding that the copy-number variation patterns of circulating tumor cells are cancer dependent offers the potential for noninvasive cancer diagnosis based on the copy-number variation patterns. Further development of single-cell, whole-genome amplification methods would provide not only better understanding of tumor heterogeneity[10] and cancer metastases but also hope for personalized treatment.

Summary

DNA exists as single molecules in an individual cell. Consequently, gene expression and genomic variations occur stochastically, necessitating single-cell and single-molecule measurements. Single-cell genomics is where single molecules meet genomics. The ability to count the copy numbers of a gene and detect a point mutation in a single cell is now not only possible, but is critically important as well. Such single-molecule methods have allowed for probing, understanding, and bettering life at the single-molecule level and provided a tangible example of precision medicine.

ARTICLE INFORMATION

Published Online: May 14, 2015. doi:10.1001/jama.2015.5533.

Conflict of Interest Disclosures: The author has completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Dr Xie is a co-founder of Yikon Genomics, a single-cell genomic start-up company.

Funding/Support: Dr Xie's research on single-molecule enzymology, gene expression, and cancer genomics has been funded by National Institutes of Health Director's Pioneer Award. MALBAC development was funded by the National Institutes of Health National Human Genome Research Institute. The in vitro fertilization work was funded by Peking University.

Role of the Funder/Sponsor: The funders/sponsors had no role in preparation, review, or approval of the manuscript.

Additional Contributions: I am indebted to current and past group members and collaborators for their contributions, particularly Jie Qiao, MD, and Fuchou Tang, PhD, for their work on in vitro fertilization and Jie Wang, MD, and Fan Bai, PhD, for their work on circulating tumor cells.

Additional Information: Dr Xie is the recipient of the 2015 Albany Medical Center Prize in Medicine and Biomedical Research.

REFERENCES

1. Lu HP, Xun L, Xie XS. Single-molecule enzymatic dynamics. *Science*. 1998;282(5395):1877-1882.

2. Eid J, Fehr A, Gray J, et al. Real-time DNA sequencing from single polymerase molecules. *Science*. 2009;323(5910):133-138.

3. Li GW, Xie XS. Central dogma at the single-molecule level in living cells. *Nature*. 2011;475(7356):308-315.

4. Method of the year 2013. *Nat Methods*. 2014;11(1):1-112.

5. Zong C, Lu S, Chapman AR, Xie XS. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science*. 2012;338(6114):1622-1626.

6. Hou Y, Fan W, Yan L, et al. Genome analyses of single human oocytes. *Cell*. 2013;155(7):1492-1506.

7. Heger M. Peking University reports first success in trial of single-cell sequencing for PGD. http://bernstein.harvard.edu/pages/news/genomeweb_PKU_reports_PGD.pdf. Accessibility verified May 5, 2015.

8. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558.

9. Ni X, Zhuo M, Su Z, et al. Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Proc Natl Acad Sci U S A*. 2013;110(52):21083-21088.

10. Navin N, Kendall J, Troge J, et al. Tumour evolution inferred by single-cell sequencing. *Nature*. 2011;472(7341):90-94.