Advances in Experimental Medicine and Biology 1068

Jianqin Gu · Xiangdong Wang *Editors*

Single Cell Biomedicine

Advances in Experimental Medicine and Biology

Volume 1068

Editorial Board
IRUN R. COHEN, *The Weizmann Institute of Science, Rehovot, Israel*
ABEL LAJTHA, *N.S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA*
JOHN D. LAMBRIS, *University of Pennsylvania, Philadelphia, PA, USA*
RODOLFO PAOLETTI, *University of Milan, Milan, Italy*
NIMA REZAEI, *Children's Medical Center, Tehran University of Medical Sciences, Tehran, Iran*

More information about this series at http://www.springer.com/series/5584

Jianqin Gu • Xiangdong Wang
Editors

Single Cell Biomedicine

Editors
Jianqin Gu
Henan Provincial People's Hospital
Zhengzhou, Henan, China

Xiangdong Wang
Shanghai Institute of Clinical
Bioinformatics
Zhongshan Hospital, Fudan University
Shanghai, China

ISSN 0065-2598 ISSN 2214-8019 (electronic)
Advances in Experimental Medicine and Biology
ISBN 978-981-13-0501-6 ISBN 978-981-13-0502-3 (eBook)
https://doi.org/10.1007/978-981-13-0502-3

Library of Congress Control Number: 2018946643

© Springer Nature Singapore Pte Ltd. 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Contents

1 **Can the Single Cell Make Biomedicine Different?** 1
Yuming Wang, Li Li, Xiangdong Wang, and Jianqian Gu

2 **Automated Single-Cell Analysis and Isolation System: A Paradigm Shift in Cell Screening Methods for Bio-medicines.** 7
Kenji Tatematsu and Shun'ichi Kuroda

3 **Single-Cell Non-coding RNA in Embryonic Development** 19
Qiang Fu, Chuan-Jiang Liu, Zhen-Sheng Zhai, Xu Zhang, Tao Qin, and Hong-Wei Zhang

4 **High Throughput Single Cell RNA Sequencing, Bioinformatics Analysis and Applications** 33
Xiaoyun Huang, Shiping Liu, Liang Wu, Miaomiao Jiang, and Yong Hou

5 **Circulating Tumor Cells: The Importance of Single Cell Analysis** 45
D. Heymann and M. Téllez-Gabriel

6 **Super-Resolution Fluorescence Microscopy for Single Cell Imaging** 59
Han Feng, Xiaobo Wang, Zhiwei Xu, Xiaoju Zhang, and Yongju Gao

7 **Single Cell Proteomics for Molecular Targets in Lung Cancer: High-Dimensional Data Acquisition and Analysis** . . . 73
Zheng Wang and Xiaoju Zhang

8 **Therapeutic Antibody Discovery in Infectious Diseases Using Single-Cell Analysis** 89
Alexandria Voigt, Touyana Semenova, Janet Yamamoto, Veronique Etienne, and Cuong Q. Nguyen

9 **Single Cell Genetics and Epigenetics in Early Embryo: From Oocyte to Blastocyst** 103
Yubao Wei, Helong Zhang, Qian Wang, and Cuilian Zhang

10 The Potential Roles and Advantages of Single Cell Sequencing in the Diagnosis and Treatment of Hematological Malignancies 119
Mingyue Shi, Xiaoyan Dong, Lei Huo, Xiaobin Wei, Fang Wang, and Kai Sun

11 Application of Single Cell Sequencing in Cancer 135
Lan Yu, Hua Zhao, Li Meng, and Cuilian Zhang

12 Emergence of Bias During the Synthesis and Amplification of cDNA for scRNA-seq 149
Qiankun Luo and Hui Zhang

13 Detection and Application of RNA Editing in Cancer 159
Mengjia Qian, Claudio Spada, and Xiangdong Wang

14 Is Pooled CRISPR-Screening the Dawn of a New Era for Functional Genomics 171
Jufang Yao and Hui-Li Dai

15 Roles of Single Cell Systems Biomedicine in Lung Diseases 177
Yiming Zeng, Xiaoyang Chen, and Xiangdong Wang

16 The Significance of Single-Cell Biomedicine in Stem Cells 187
Weishan Zhuge, Furong Yan, Zhitu Zhu, and Xiangdong Wang

Check for updates

1 Can the Single Cell Make Biomedicine Different?

Yuming Wang, Li Li, Xiangdong Wang, and Jianqian Gu

Abstract

The single-cell as the basic unit of biological organs and tissues has recently been considered an important window to furthermore understand molecular mechanisms of organ function and biology. The current issue with a special focus on single cell biomedicine is the first effort to collect the evidence of disease-associated single cell research, define the significance of single cell biomedicine in the pathogenesis of diseases, value the correlation of single cell gene sequencing with disease-specific biomarkers, and monitor the dynamics of RNA processes and responses to microenvironmental changes and drug resistances.

Keywords

Single cell · Sequencing · Imaging · Bioinformatics · Systems biology

1.1 Introduction

The single-cell as the basic unit of biological organs and tissues has recently been considered an important window to furthermore understand molecular mechanisms of organ function and biology. With the development of single cell biotechnology, the single cell biomedicine becomes more and more important area to understand the heterogeneity among cells, identify disease-specific biomarkers, and explore molecular regulations and signals. The single cell systems biology is emphasized as an approach to understand single-cell mechanical phenotypes, single-cell biology, heterogeneity and organization of genome function [1]. Multi-dimensional, multi-layer, multi-crossing and stereoscopic single-cell biology definitely will benefit the discovery and development of disease-specific biomarkers, translation of single-cell systems biology into clinical phenotype, and understanding of single-cell gene sequencing and function in patient response to therapies. As a part of single cell biomedicine, single cell RNA sequencing (scRNA-seq) is used as a critical and initial tools to define the alterations of transcriptomes, development of intratumor and intercellular heterogeneity, and genotoxicity in response to drugs [2]. scRNA-seq can detect somatic mutations and epigenetic alterations in evolution, post-transcriptional RNA modifications, and RNA editing. It is also important to illuminate the effects of single-cell RNA

Author contributed equally with all other contributors. Yuming Wang and Li Li

Y. Wang · L. Li · X. Wang (✉) · J. Gu (✉)
Henan Province People's Hospital, Henan University, Zhengzhou, China
e-mail: Xiangdong.wang@clintransmed.org

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_1

isoform diversity on gene, protein expression, and regulation. The current issue with a special focus on single cell biomedicine is the first effort to collect the evidence of disease-associated single cell research, define the significance of single cell biomedicine in the pathogenesis of diseases, value the correlation of single cell gene sequencing with disease-specific biomarkers, and monitor the dynamics of RNA processes and responses to microenvironmental changes and drug resistances.

1.2 Rapid Development of Single Cell Measurements

The accuracy of targeted single cell isolation, purification, and measurement is the critical step to ensure the trust of single cell biomedicine. Tatematsu and Kuroda as the leading scientists developed an automated robot that facilitates non-invasive isolation of a single cell with the most favorable properties from arrays containing >10^5 cells, and described the system as a "single-cell robot" to compare with a conventional fluorescence-activated cell sorter [3]. Such system can carry out a high-throughput screening for single cell isolation with targeted labelling and perform the comprehensive analysis the biological function of receptor-associated signaling between single cells. In addition, they also clearly described the advantages between single colony-based and single cell-based breeding methods, and between "single cell robot" with conventional single-cell analysis, automated single-cell analysis, or other cell screening methods enabled by automated single-cell analysis. It is more critical to monitor alterations of single-cell dynamic phenotypes during evolution, microenvironmental changes, disease progression, and therapy, with the development of single-cell technologies in the deep understanding and value of the constituents within dynamic phenotypes [4]. The detection of single-cell dynamic phenotypes will require more precise readouts from "single cell robot", not only on the findings but also the meaning of tumor heterogeneity and evolution to carcinogenicity, metastasis, and responses to targeted therapies.

Huang et al. systemically addressed the methodology of high throughput single cell RNA sequencing, bioinformatics analysis and applications, by combining their own innovative experience in the current volume [5]. Hou's group as one of the pioneer scientists on the development of high throughput scRNA-seq published their early work to measure the single-cell exome sequencing on a clear cell renal cell carcinoma and its adjacent kidney tissue to better understand the intratumoral genetics underlying mutations of cancer in 2012 [6]. It was proposed that such quantitative population genetic analysis as new ways could identify the clonal subpopulations, mutation spectrums, or detailed intratumoral genetic landscape at a single-cell level. They furthermore described the difference of large-scale scRNA-seq library preparations in the accuracy and throughput of scRNA-seq, and the importance of computational analysis of scRNA-seq data. When analyzing scRNA-seq data, we should clearly define what the quality control, criteria of subpopulations identified, or differential expression and transcript isoforms across conditions are, how expression estimation and normalization, pseudotemporal ordering, or interrogation of spatial information are settled out, as well as why network inference, differential Splicing, or allelic expression patterns should have the special attentions.

More approaches are used to determine dynamic phenotypes of single cells by the cross-disciplinary nature of these techniques, e.g. quantitative live cell imaging, time series analysis, computational modeling, and statistical testing on multi-dimensional data sets. Ruderman headlined the computational models as a predictor reflecting the quantitative phenotypes of cells, new theories as a system screening the key response variables of phenotypes, or multidisciplinary dynamic phenotype research teams [7]. Feng et al. overviewed and discussed recent applications of super-resolution techniques in single cell imaging for multi-dimensional, multicolor, live-cell imaging [8]. The quality and potentials of multimodal imaging are compared

among stimulated emission depletion, structured illumination, single-molecule localization, and other super-resolution microscopies. It will significantly improve our vision and understanding of single cells if cell imaging can be integrated with molecular biology, signaling, regulation, and bio-computing algorithms.

1.3 Single Cell Biology in Development and Evolution

Single-cell transcription kinetics and variability play an important role in cell development and evolution through gene regulation. Of those, noncoding RNA (ncRNA) is suggested to regulate cell mechanic changes and volume flexibility. Fu et al. illuminated the emerging single-cell RNA sequencing technique and he expression of ncRNAs during embryo development [9]. The expression of ncRNAs within single cells measured with single-cell RNA-seq techniques can vary with stages of embryonic development. ncRNA, especially lncRNAs and miRNAs, can regulate and prevent embryonic cell development from the disorder. Although partial functions of single-cell lncRNAs and miRNAs was explored, the most of single-cell circRNAs, piRNAs, or snoRNAs functions remain unclear in embryonic development. Wei et al. described the contributions of single cell genetics and epigenetics in early embryo from basic research to reproductive medical application and the knowledge of programming/reprogramming and the epigenetics dynamics in the cell lineage differentiation [10]. This is a special vision from the reproductive medicine to evaluate the meaning of embryo or polar body scRNA-seq to genetic diagnosis and prediction. The single cell techniques and bioinformatics analyses for early embryo were listed and compared with other tissue cells. Single-cell biomedicine in the development will provide the details of each cell origination and sources as well as molecular mechanisms by the landscape shaped itself. Single cell DNA methylation will demonstrate the mechanisms of cell lineage differentiation, gene expression heterogeneity in the pluripotent state of mouse embryonic stem cells, or the start of a lineage transition or a transient phase of altered sensitivity to lineage-specific signals.

1.4 Heterogeneity of Single Circulating Tumor Cells

The cancer heterogeneity can be described by single cell sequencing and comprehensive molecular characterizations of cancer cells, including hereditary and somatic gene changes and mutations. The specificity, characterization, and roles of cancer cell heterogeneity can decide the sensitivity and resistance of cells to therapies and be considered as the critical factor to develop target-driven therapies and strategies applied in clinical trials based on a proposed precise self-validation system [11]. Cancer heterogeneity can act as a potential cause of drug resistance to targeted therapy, contribute to tumor evolution and adaptation, and influence the efficacy of personalized-medicine strategies. The influence of tumor heterogeneity on drug efficacy and resistance should be monitored by disease- and biology-specific biomarkers [12]. The intelligent single-cell robot of human cells were proposed to integrate together systems information of molecules, genes, proteins, organelles, membranes, architectures, signals, and functions to assist clinicians in the decision-making, molecular understanding, risk analyzing, and prognosis predicting [13]. Heymann and Téllez-Gabriel pointed out the characterization of heterogeneity among circulating tumor cells (CTCs) at the single cell level could be an important approach to explore the causes and progression of disease and the accurate selection of molecular biomarkers [14]. This is an initiative of disease-orientated figure to enrich, isolate, purify, and measure the single CTCs at different levels, including RNA, DNA, protein and epigenetic events. In addition to the value of cancer indication, the single CSCs will provide more clinical and biological importance to identify the heterogeneity, origin, subtypes, and malignancy of the cancer. The single circulating cell will be the major source and play the

critical role in identification and validation of disease-specific biomarkers metastases, drug resistance, prognosis, phenotypes, metabolism, or proliferation. With advances in single-cell sequencing technologies, the complete genome of the single CTC can be defined and compared with corresponding primary and metastatic tumor single cells to monitor genomic variations in metastatic and recurrent tumors, infer tumor evolution during treatment, and examine mechanisms of the epithelial-mesenchymal transition [15]. The sequencing of single CTC genomes and transcriptomes is even more complex and difficult, e.g. eliminating backgrounds of white blood cells, isolating and collecting cells without damaging or losing DNA and RNA, obtaining unbiased and even whole-genome and transcriptome amplification material, and analyzing sequencing data.

1.5 Single Cell Values in Cancer

The cancer is a major area where the single cell technologies were applied mostly to define the heterogeneity of intra- or inter-tumor cells, rare cancer cell types, gene mutation and characters, evolution and developmental lineage relationships, or sensitivity to therapies. Lung cancer is one of the most severe cancers with the highest incidence and mortality, with a complex mechanisms and available targeted therapies. A large number of lung cancer-associated biomarkers have been developed to monitor the severity, duration, subtypes, and transit from chronic lung diseases to cancer [16–19]. Wang and Zhang brought out single cell proteomics as a front point of single cell biomedicine with a clear focus on lung cancer and summarized potential technologies to measure single cell protein profiles [20], including flow cytometry, mass cytometry, microfluidics and chip technologies, chemical cytometry, single-cell western blotting, or quantity and functions of proteins. The single-cell proteomics are mainly applied for the identification and screening of diagnostic biomarkers and therapeutic targets for prevention, early detection, prognosis, and response to therapy, as well as for the understanding of mechanisms. While, the single cell sequencing is often used to identify gene mutations and intercellular heterogeneity. It would be important to define the correlation and biological consequences between the gene mutation and protein expression at the single cell. As Yu et al. summarized in the current book [21], the single-cell sequencing has been widely applied in cancer research, e.g. breast cancer, ovarian cancer, lung cancer, hematopoietic tumors, renal cell cancer, glioblastoma, circulating tumor cells, or cancer stem cells. In addition to the gene mutation, the single cell sequencing is expected to provide more indications or potential evidence on which clinicians can consider or select the individualized or targeting therapies. The single cell sequencing can benefit to identify the new sub-populations of cancer cells, the variation between cancer cells and cancer stem cells, and the development of drug resistance.

Hematological malignancies are one of challenging cancers with poor prognosis and non-specific therapies due to the downregulation of target antigens and the immunosuppressive environment against the host immune response [22]. A number of potential immunotherapies, e.g. T cells, NK cells, or monoclonal antibodies, or inducing and/or recovering T cell activation, provide the exciting future for the patients, while the large number of blood cancer cell heterogeneity as an important factor in response to treatment may influence or decline the efficacy of therapies. Chu et al. recently emphasized there is a great heterogeneity among subclones and their extensions, especially in hematological malignancies and called special attention to define the aggregate populations, intra-clonal and inter-clonal heterogeneity, and its frequency, using single cell sequencing [23]. It seems that single cell systems biology may generate more unique and important information on cancer cell subtypes, heterogeneities, or epigenetics to assist clinicians in the diagnosis and therapeutic design for diseases and in the prognosis of patients with individualized therapies. Shi et al. furthermore overviewed potential roles of single cell sequencing in the diagnosis and treatment of hematologic malignancies and tried to headline the advantages

of the single cell biology from the clinical point [24]. This particular article collected the scientific evidence from studies and summarized genomic, transcriptomic, and epigenomic findings in single cells of acute leukemia, multiple myeloma, or chronic myeloid leukemia. The most valuable point of single cell biology, e.g. circulating tumor cell sequencing, is to monitor minimal residual disease of hematologic malignancies and define functional heterogeneity and clonal evolution in such life-threatening hematological diseases.

Clustered regularly interspaced short palindromic repeats (CRISPR)/Cas (CRISPR associated) system has been applied in many aspects to understand the molecular mechanisms of a gene, signal pathway, and regulatory function, as "the solution for gene editing's Gordian knot" [25]. For example, CRISPR was proposed to play the important role in the understanding of drug genotoxicity and resistance, during which how gene changes, mutations, and heterogeneity may control and dominate the cell signaling, regulation, and sensitivity to drugs [26]. During the interaction between cells and drugs, the perturbation and phenotype of a cell can be changed and monitored using the single-cell CRISPR screening. In addition, CRISPR can be one of the most important genome editing-assisted gene knock-in technologies, to repair genetic changes and cure inherit diseases [27]. In the current volume, Qian and Wang demonstrated that the RNA editing as a RNA structure research tool also plays important role in cancer research, especially in the understanding of biological function of RNA species, structures, and expression [28].

Nowadays, transcriptomics studies mainly focus on three aspects, the RNA species (mRNA and non-coding RNA), the RNA structure (start sites, splicing patterns and post-transcriptional process) and the expression levels of RNA. Among them, the RNA structure research tool, RNA editing, remains the least popular one which we still have more to explore on the role of it in cancer research [28]. RNA editing enzymes such as ADARs and APOBECs are all promising targets in cancer therapeutic strategy. Here we listed several examples of RNA editing studies in some cancers. However, their pathways are differentially regulated in cancers which should be further clearly studied. The best tool to study RNA editing is NGS. Here, we also discussed the challenges and the possible ways to overcome them. We are sure to believe that RNA editing performed by NGS has the ability in studying transcriptomes, even at single cell level. It will be sure to help a lot in cancer diagnosis and treatment in the near future.

1.6 Other Biological Significance

Single-cell-based biotechnologies can be also used in multiple aspects. Voigt A et al. initially proposed to develop protein-based therapies, e.g. antigen-specific monoantibody, through single cell system [29]. The specificity of protein-based therapy can be screened and validated in the single cell system. Single cell analysis can define the heterogeneity-associated efficacy among cells. The measurement of such heterogeneity is fully dependent upon the quantitative accuracy of scRNA-seq, including the protocol, RNA reverse transcription, or cDNA pre-amplification [30]. In addition, scRNA-seq is suggested as a powerful tool to measure the heterogeneity and germline of stem cells [31]. Furthermore, scRNA-seq was applied in pulmonary epithelial cells isolated and harvested from the lung of animals or patients suffered from diseases [32].

In conclusion, the current issue with a special focus on single cell biomedicine is the first effort to collect the evidence of disease-associated single cell research, define the significance of single cell biomedicine in the pathogenesis of diseases, value the correlation of single cell gene sequencing with disease-specific biomarkers, and monitor the dynamics of RNA processes and responses to microenvironmental changes and drug resistances.

References

1. Niu F, Wang DC, Lu J, Wu W, Wang X (2016) Potentials of single-cell biology in identification and validation of disease biomarkers. J Cell Mol Med 20(9):1789–1795. https://doi.org/10.1111/jcmm.12868
2. Wang W, Gao D, Wang X (2017) Can single-cell RNA sequencing crack the mystery of cells? Cell Biol Toxicol https://doi.org/10.1007/s10565-017-9404-y
3. Tatematsu K, Kuroda S (2018) Automated single-cell analysis and isolation system: a paradigm shift in cell screening methods for bio-medicines. Adv Exp Med Biol 1068
4. Wang W, Zhu BJ, Wang X (2017) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427
5. Huang X, Liu S, Wu L, Jiang M, Hou Y (2018) High throughput single cell RNA sequencing, bioinformatics analysis and applications. Adv Exp Med Biol 1068
6. Xu X, Hou Y, Yin X, Bao L, Tang A, Song L et al (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148(5):886–895. https://doi.org/10.1016/j.cell.2012.02.025
7. Ruderman D (2017) The emergence of dynamic phenotyping. Cell Biol Toxicol 33(6):507–509. https://doi.org/10.1007/s10565-017-9413-x
8. Feng H, Wang X, Xu Z, Zhang X, Gao Y (2018) Super-resolution fluorescence microscopy for single cell imaging. Adv Exp Med Biol 1068
9. Fu Q, Liu CJ, Zhai ZS, Zhang X, Qin T, Zhang HW (2018) Single-cell non-coding RNA in embryonic development. Adv Exp Med Biol 1068
10. Wei Y, Zhang H, Wang Q, Zhang C (2017) Chapter 9: single cell genetics and epigenetics in early embryo: from oocyte to blastocyst. Adv Exp Med Biol 1068
11. Wang DC, Wang W, Zhu B, Wang X (2017) Lung cancer heterogeneity and new strategies for drug therapy. Annu Rev Pharmacol Toxicol https://doi.org/10.1146/annurev-pharmtox-010716-104523
12. Wu D, Wang DC, Cheng Y, Qian M, Zhang M, Shen Q, Wang X (2017) Roles of tumor heterogeneity in the development of drug resistance: a call for precision therapy. Semin Cancer Biol 42:13–19
13. Wang DC, Wang X (2017) Tomorrow's genome medicine in lung cancer. Semin Cancer Biol 42:39–43
14. Heymann D, Téllez-Gabriel M (2017) Circulating tumor cells: the importance of single cell analysis. Adv Exp Med Biol 1068
15. Zhu Z, Qiu S, Shao K, Hou Y (2017) Progress and challenges of sequencing and analyzing circulating tumor cells. Cell Biol Toxicol https://doi.org/10.1007/s10565-017-9418-5
16. Wang X (2016) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361. https://doi.org/10.1007/s10565-016-9350-0
17. Xu M, Wang X (2017) Critical roles of mucin-1 in sensitivity of lung cancer cells to tumor necrosis factor-alpha and dexamethasone. Cell Biol Toxicol 33(4):361–371. https://doi.org/10.1007/s10565-017-9393-x
18. Bao L, Zhang Y, Wang J, Wang H, Dong N, Su X, Xu M, Wang X (2016) Variations of chromosome 2 gene expressions among patients with lung cancer or non-cancer. Cell Biol Toxicol 32(5):419–435. https://doi.org/10.1007/s10565-016-9343-z
19. Bao L, Diao H, Dong N, Su X, Wang B, Mo Q et al (2016) Histone deacetylase inhibitor induces cell apoptosis and cycle arrest in lung cancer cells via mitochondrial injury and p53 up-acetylation. Cell Biol Toxicol 32(6):469–482
20. Wang Z, Zhang X (2018) Single cell proteomics for molecular targets in lung Cancer: high-dimensional data acquisition and analysis. Adv Exp Med Biol 1068
21. Yu L, Zhao H, Meng L, Zhang C (2018) Application of single cell sequencing in cancer. Adv Exp Med Biol 1068
22. Lin C, Chen S, Li Y (2017) T cell modulation in immunotherapy for hematological malignancies. Cell Biol Toxicol 33(4):323–327. https://doi.org/10.1007/s10565-017-9397-6
23. Chu MP, Kriangkum J, Venner CP, Sandhu I, Hewitt J, Belch AR, Pilarski LM (2017) Addressing heterogeneity of individual blood cancers: the need for single cell analysis. Cell Biol Toxicol 33(2):83–97. https://doi.org/10.1007/s10565-016-9367-4
24. Shi M, Dong X, Wei X, Wang F, Huo L, Sun K (2018) The potential roles and advantages of single cell sequencing in the diagnosis and treatment of hematologic malignancies. Adv Exp Med Biol 1068
25. Fang H, Wang W (2016) Could CRISPR be the solution for gene editing's Gordian knot? Cell Biol Toxicol 32(6):465–467
26. Wang W, Wang X (2017) Single-cell CRISPR screening in drug resistance. Cell Biol Toxicol 33(3):207–210. https://doi.org/10.1007/s10565-017-9396-7
27. Sakuma T, Yamamoto T (2017) Magic wands of CRISPR-lots of choices for gene knock-in. Cell Biol Toxicol 33(6):501–505. https://doi.org/10.1007/s10565-017-9409-6
28. Qian M, Wang X (2018) Detection and application of RNA editing in cancer. Adv Exp Med Biol 1068
29. Voigt A, Semenova T, Yamamoto J, Tienne V, Nguyen CQ (2018) Therapeutic antibody discovery in infectious diseases using single-cell analysis. Adv Exp Med Biol 1068
30. Zhuge W, Wang X (2018) The significance of single-cell analysis in stem cells. Adv Exp Med Biol 1068
31. Wang X, Zeng Y (2018) Single cell sequencing in respiratory diseases. Adv Exp Med Biol 1068
32. Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J, et al (2016) Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. JCI Insight 1(20):e90558

2 Automated Single-Cell Analysis and Isolation System: A Paradigm Shift in Cell Screening Methods for Bio-medicines

Kenji Tatematsu and Shun'ichi Kuroda

Abstract

We have developed an automated robot that facilitates non-invasive isolation of a single cell with the most favorable properties from arrays containing >10^5 cells, thus allowing the establishment of new cell screening methods for bio-medicines. In this chapter, an outline of the proposed automated single-cell analysis and isolation system (hereafter called 'single-cell robot') is reviewed by comparison with a conventional fluorescence-activated cell sorter (FACS). The single-cell robot could perform high-throughput screening for both mammalian cells secreting the highest amount of bio-medicines (*e.g.* Chinese hamster ovary (CHO) cells or hybridomas), and stem cells with the highest pluripotency (*e.g.*, embryonic stem (ES) cells), from huge number of cell libraries based on the recently proposed concept of "single cell-based breeding". The rational screening method for the *de novo* agonist design could also be performed using yeast cells expressing functional mammalian cytokine receptors (*e.g.*, epidermal growth factor receptor (EGFR), somatostatin G protein-coupled receptor (SSTR5), and interleukin 5 receptor (IL5R)). Furthermore, the single-cell robot could comprehensively analyze the reaction between olfactory sensory neurons and specific odorants, which will shed light on how odorants are recognized by olfactory receptors. Taken together, these unique features of the proposed single-cell robot will contribute to the high-throughput development of forthcoming bio-medicines.

Keywords

Time-lapse single-cell array cytometry · High-throughput screening · Single cell-based breeding · CHO cells · Hybridomas · ES cells · Yeast cells · Olfactory sensory neurons

2.1 Introduction

Non-invasive single-cell isolation, separating individual target cells from vast cell libraries, is very important for cell analysis in the field of bio-science field, establishing high-producing cells in the bio-engineering and bio-medical fields, and performing cell diagnostics in the bio-medical field. Conventionally, the fluorescence-activated cell sorter (FACS) has been widely used for this purpose; however, FACS cannot be used for the samples containing very few target cells (content percentage <0.1%), rare samples that will be

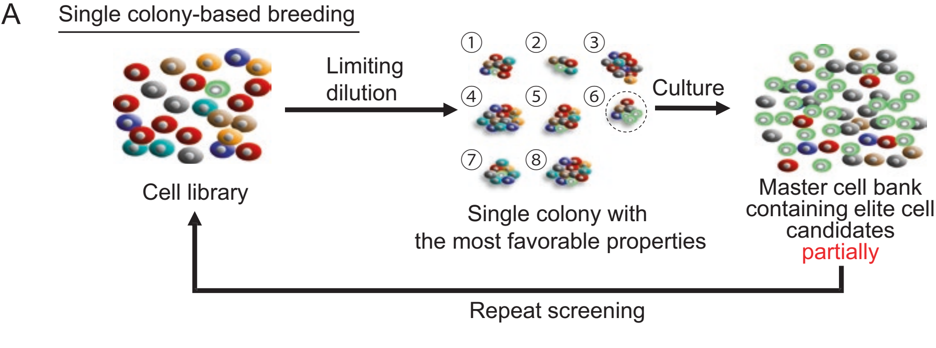
K. Tatematsu · S. Kuroda (✉)
Department of Biomolecular Science and Reaction, The Institute of Scientific and Industrial Research (ISIR), Osaka University, Ibaraki, Osaka, Japan
e-mail: kenji44@sanken.osaka-u.ac.jp; skuroda@sanken.osaka-u.ac.jp

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068, https://doi.org/10.1007/978-981-13-0502-3_2

reused, or samples that are highly susceptible to various stresses. Furthermore, FACS cannot sort cells based on time-dependent changes in cellular properties. This chapter presents an outline of an automated single-cell analysis and isolation system (hereafter called “single-cell robot”) that the authors have recently developed and commercialized in Japan [19], and introduces new cell screening methods for developing bio-medicines based on the “single cell-based breeding” concept [20].

2.2 Single Colony-Based Breeding *Versus* Single Cell-Based Breeding

In order to select microorganisms possessing the most favorable properties from a cell library, contemporary methods require single-colony isolation followed by cultivation for further evaluation steps. Similarly, mammalian cells possessing the most favorable properties are currently selected from cell libraries using the limiting dilution method followed by colony formation and cultivation. In both cases, as a single colony is formed from single cell monoclonally, it has long been believed that the phenotype of each cell in the resulting colony is identical and unchanged. However, due to the epigenetic effect, a ‘stochastic fluctuation’ in the gene expression of each cell has been indicated to occur during colony formation, resulting in ‘cellular heterogeneity’ among the cells in these colonies [2] (Fig. 2.1a). This differentiation in expression has forced scientists to perform laborious steps to permanently maintain the conventional “single colony-based breeding” in cell-based research projects and industries. It is known, however, that “elite cells” exhibiting the most favorable properties with less cellular heterogeneity can be found in cell libraries. These elite cells have been isolated as a result of long-term research and development efforts, and have proven to be able to sustain the mass production of bio-materials and bio-medicines in cell-based industries. Therefore, this work aimed to establish a rapid and high-throughput selection method for isolating elite cells that no longer requires



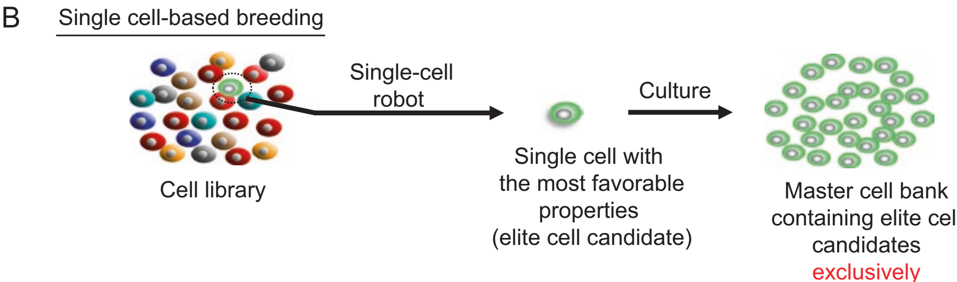


Fig. 2.1 Single colony-based breeding method (conventional, A) and single cell-based breeding method (new, B)

either time-consuming colony formation or the laborious handling of a large number of colonies (shifting the paradigm from "single colony-based breeding" to "single cell-based breeding") [20]. For example, it is probable that a master cell bank established by single colony-based breeding will be contaminated by non-elite cells, whereas a colony established using single cell-based breeding could be obtained through short-term steps and will contain exclusively elite cells or elite cell candidates (Fig. 2.1b).

2.3 Conventional Single-Cell Analysis and Isolation System

FACS has been recognized as the *de facto* standard machine in the process of isolating single target cells from cell libraries consisting of huge numbers of cells. Using FACS, droplets containing single cells are aligned in laminar flow. The properties of each cell (*e.g.*, shape, size, and cell-surface markers) are analyzed via scatter-patterns of light and fluorescence, and each droplet containing a target cell is independently charged and then sorted in a charge-dependent manner. Although FACS possesses high-throughput processing capability ($>10^4$ cells/s) the sorted cells are often damaged by chemical (sheath solution) and physical stresses (high voltage, high hydraulic pressure, ultrasound), resulting in a low survival rate of the sorted cells. Furthermore, FACS is not able to identify rare target cells (content percentage <0.1%) because of its low spectrophotometric resolution, and is unable to analyze entire samples, because pre-run selections of samples are required for gate adjustment and samples in the dead space of the flow path cannot be used. To address the cell damage issues of FACS, a micro total analysis system (microTAS; microfluidics technology) has recently been applied to single-cell analysis and isolation systems (*e.g.*, On-chip Sort, On-Chip Biotechnologies Co., Ltd., Japan). This system facilitates single-cell isolation in a non-invasive manner; however, it is limited by a low processing capability (approximately 4000 cells/s).

Table 2.1 FACS *versus* single-cell robot

Specifications	FACS	Single-cell robot
Cell sample:		
Min. content of target cells	>0.1%	~0.0003%
Optimum cell concentration	10^6 ~ 10^7 cells/ml	~3×10^5 cells/ml
Max. number of cells	∞	~3.4×10^5 cells (φ10-μm chamber)
Cell suspension buffer	Sheath solution/ buffer	Cell culture medium
Isolation of cell aggregates	No	Yes
Instrument:		
Analyzing speed	~7×10^4 cells/sec	~2.5×10^4 cells/20 min
Sorting speed	~3×10^4 cells/sec	96 cells/30 min
Disposability of cell-contacted parts	No	Yes
Pre-run for gate adjustment	Necessary	Not necessary
Reusability of used cells	No	Yes
Available colors	>8 colors	~3 colors
Time-laps analysis	No	Yes
Observation of cell morphology	No	Yes

Moreover, both FACS and microTAS are incapable of sorting cells based on time-dependent changes of cellular properties (*e.g.*, transient changes of second messengers (Ca^{2+}, cAMP), fluorescence, or morphology). New developments in single-cell analysis and isolation systems are expected to include time-lapse cytometric functions to expand their applicability into the cellular engineering and bio-medical fields (Table 2.1).

2.4 Automated Single-Cell Analysis and Isolation System

As compared with conventional single-cell analysis and isolation systems, the system proposed in this work is expected to possess the following

five performance features: target cells should be recovered without chemical or physical stress, rare target cells (content percentage <0.1%) can be identified and recovered non-invasively, entire samples can be subjected to analysis and target cell isolation, samples should be reusable for future analyses, and finally the system should measure time-dependent changes in cellular properties. In 2003, it was considered that most of these issues could be solved by changing conventional systems from flow cytometry to array cytometry. In 2013, the development of "an automated single-cell analysis and isolation system" (hereafter called single cell robot) has finally proved successful [19]. The resulting robot consists of four units (the single-cell robot (main unit), perfusion pump, light source, and personal computer) (Fig. 2.2a). The main unit is composed of two parts: the analysis part contains an inverted fluorescence microscope, CCD camera, optical unit, temperature control unit, and microchamber array; and the isolation part contains a glass capillary-equipped micropump on a micromanipulator (for Z-axis movement) and reservoir plate on a motorized stage (for XY-axis movement). Seven types of glass capillaries are available for selection. For example, the inner diameter of a rank 0 capillary is 8–14 μm (fitting yeast cells), whereas that of a rank 3 capillary is 25–29 μm (most cells) and that of a rank 6 capillary is 39–45 μm (cell aggregates). Microchamber arrays can also be selected with three well diameters (10, 20, and 30 μm), with respective well numbers of 338,560 (for yeast cells and blood-derived cells), 196,000, and 84,640 (for most cells) (Fig. 2.2c). Microchamber arrays are made of low fluorescence polystyrene, and the surface of this material can be modified with several functional groups (for an immunochamber, described below) and/or plasma (hydrophilic treatment).

If approximately 8.5×10^4 hybridoma cells in a culture medium are introduced to a 30 μm microchamber array via brief centrifugation (conditions: $50 \times g$, 1 min, room temperature), approximately 70% of wells in the microchamber

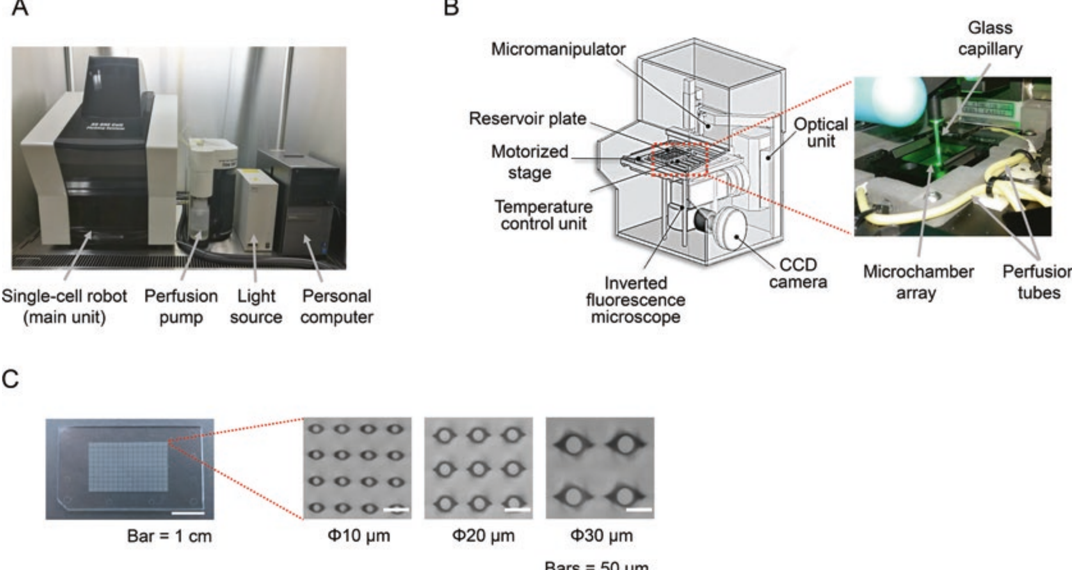


Fig. 2.2 Automated single-cell analysis and isolation system. (**a**) Overview of the single-cell robot, composed of 4 units. (**b**) Transparent view of the single-cell robot. (**c**) Microchamber arrays (3 types: well diameters 10, 20, and 30 μm)

array could be occupied by hybridomas. When twice this number of hybridomas are used under the same conditions, >90% of wells can be occupied [19]. In the analysis mode, both fluorescence and transmission images of all cells on microchamber array (in culture medium) are automatically captured with a resolution of approximately 1 µm by an inverted fluorescence microscope (three colors are available for excitation). Gate adjustment was performed for collecting target cells based on the histogram of fluorescent intensity of each cell. Conversely, in the isolation mode the positions of target cells in the microchamber array are automatically recognized by the robot, and these cells are retrieved from the microchamber array with the selected glass capillary and then released into the culture medium on the reservoir plate. Both fluorescence and transmission images of all target cells in the microchamber array are automatically saved before and after the isolation step. The robot utilizes an array cytometry-based system, whereas FACS utilizes a flow cytometry-based system that creates a larger dead volume. Thus, small volume of cell libraries can be analyzed by the single-cell robot without unnecessary waste for pre-runs and dead volume (Table 2.1). Furthermore, the array cytometry-based system allows for the identification of rare target cells in a sample (theoretically down to a content percentage of ~0.0003%, or 1 cell out of approximately 3.4×10^5 cells on 10-µm microchamber array) and the identification of target cells showing time-dependent changes in their cellular properties. As for the viability of sorted cells, mouse embryonic stem (ES) cells isolated by the robot were shown to grow from single cells with a 96% survival rate (23 live cells out of 24 isolated cells), whereas the same cells isolated using FACS had a 25% survival rate (24 live cells out of 96 isolated cells) (unpublished data). These results suggest that the proposed single-cell robot is less invasive to target cells than FACS. The single-cell robot may prove able to sufficiently solve the five aforementioned issues that exist in conventional single-cell analysis and isolation systems.

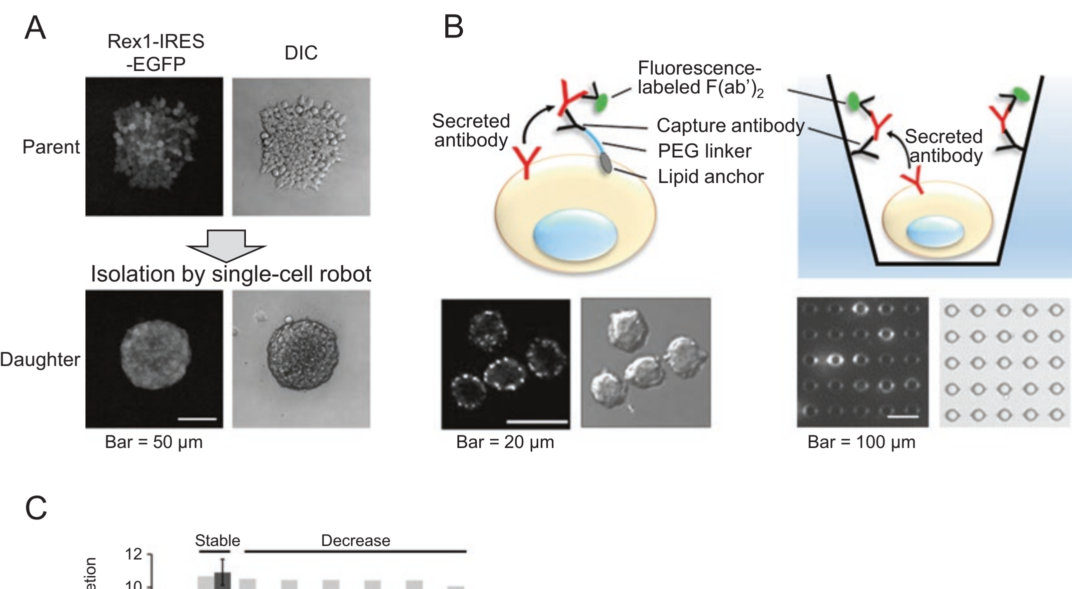
2.5 Cell Screening Methods Enabled by Automated Single-Cell Analysis and Isolation System

2.5.1 ES Cells

Stem cells including ES cells, induced pluripotent stem (iPS) cells, and mesenchymal stem cells (MSCs) have played a pivotal role in regenerative therapy. It is important for stem cells to exhibit high pluripotency over an extended period; however, they often lose their pluripotency in the manner of stochastic fluctuation [6]. This work therefore established mouse ES cells containing the Rex1 gene (a marker gene for pluripotency), internal ribosome entry site (IRES) sequence, and enhanced green fluorescence protein (EGFP) gene. From a sample of approximately 1.0×10^5 cells of the mouse ES cells, the single-cell robot was able to isolate 23 cells exhibiting the highest levels of EGFP-derived fluorescence within 1 h (Fig. 2.3a). After a 3 weeks culturing period, five single cells had produced colonies in which nearly all cells were found to evenly express higher levels of EGFP [19]. This result indicates that the single cell-based breeding technique enabled by the single-cell robot is effective for isolating elite-cell candidates from stem cell libraries.

2.5.2 Antibody-Producing Cells

Recently, the biomedical field has seen a marked increase in interest in antibody medicines. For this purpose, it is important to isolate the cells secreting the highest number of antibodies from a cell library (*e.g.*, hybridomas, Chinese hamster ovary (CHO) cells) and to identify the cells with the most efficient and stable antibody-secreting abilities. Using conventional methods, cell libraries undergo limited dilution, after which all single cells are allowed to form colonies for 2–3 weeks. In this experiment, all colonies were cultured on a small scale for 1 week before being subjected to enzyme-linked immunosorbent



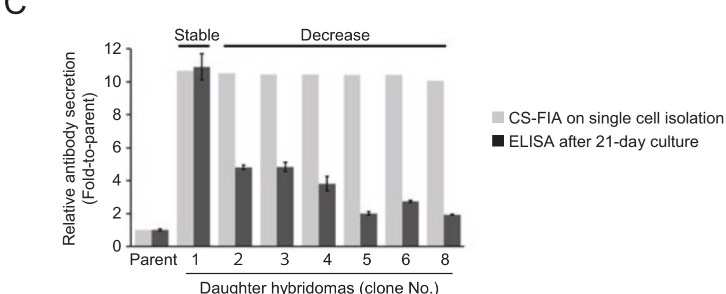


Fig. 2.3 Single cell-based breeding of ES cells and hybridomas. (**a**) Colony of mouse ES cells containing Rex1-IRES-EGFP gene (parent cell), expressing EGFP-derived fluorescence in a scattering manner. Colony formation of single mouse ES cell showing the highest expression of Rex1-IRES-EGFP gene (daughter cell), evenly expressing EGFP-derived fluorescence. Left panels, fluorescence images; right panels, DIC images. (**b**) Cell-surface fluorescence-linked immunosorbent assay (CS-FIA) (upper left panels). Lower panels, hybridomas in CS-FIA (left panel, fluorescence image; right panel, DIC). Immunochamber array (upper right panels). Lower panels, hybridomas in immunochamber array (left panel, fluorescence image; right panel, DIC). (C) Fluctuations in antibody secretion rate of hybridomas after 21 day propagation from single cells. From parent hybridoma, seven single hybridomas showing the highest secretion rates (daughter hybridomas) were isolated by the single-cell robot using CS-FIA (gray bars). After a 21-day culture, the media were analyzed using conventional ELISA to detect secreted antibodies (black bars). The daughter hybridoma No. 1 was found to have kept a higher secretion rate for at least the 21 day propagation from single cell isolation, while the other daughter hybridomas did not. Error bars, $p < 0.05$, $n = 6$

assay (ELISA) to evaluating the antibody-secreting ability of each colony. This single colony-based breeding is not a high-throughput cell screening system, because the colony formation and culture steps are excessively time-consuming (usually 1 month) and the simultaneous culture of all colonies is laborious (usually less than 1×10^4 colonies/round show the desired results). This work therefore tried to adapt the single-cell robot to single cell-based breeding by performing real-time measurements of the antibodies secreted from single cells. After evenly modifying the chosen cell surfaces with lipid-polyethylene glycol (PEG)-conjugated anti-IgG Fc antibody (capture antibody), nascent antibodies secreted from the single cell in question were promptly captured on the cell surface (within 15 min). These secreted antibodies could

then be detected with fluorescence-labeled anti-IgG F(ab′)$_2$ F(ab′)$_2$ (detection antibody) in using the sandwich FIA protocol (Fig. 2.3b, left panel). This method, designated as a cell surface-fluorescence immunosorbent assay (CS-FIA), was able to detect nascent antibodies secreted from single cell at the femtogram level (*e.g.*, 6.25 fg/cell for hybridomas [10], 0.66 fg/cell for CHO cells (Tatematsu et al., submitted)). From approximately 5.0×10^4 parental hybridomas, the combination of the single-cell robot and CS-FIA was able to isolate seven daughter hybridomas secreting >10-fold higher amounts of antibodies than the parental hybridomas within 2 h. One daughter hybridoma was found to keep this high antibody-secreting characteristic for >3 weeks, making it a potential elite-cell candidate among hybridomas (Fig. 2.3c; [19]). The single cell-based breeding enabled by the single-cell robot facilitates rapid identification of such elite-cell candidates from huge numbers of bio-medicine-secreting cells without the most prohibitively time-consuming steps (*i.e.*, colony formation, cell culture, and conventional ELISA).

Since modifying the cell-surface modification with a capture antibody may reduce the viability of isolated single cells, this work also more recently developed an immunochamber array, in which an amino group was introduced to the surface for conjugation with capture antibody (Fig. 2.3b, right panel; Tatematsu et al. submitted). Nascent antibodies secreted from single cells were promptly captured around the well before being quantified by the detection antibody in a sandwich FIA. Another research group has also developed a microchamber array possessing similar function, the immunospot array assay on a chip (ISAAC) [8].

The immune system contains B lymphocytes, naturally occurring antibody-producing cells. To identify B cells that react with specific antigens, it is useful to note that intracellular Ca^{2+} concentration is transiently increased by antigen stimulation [1]. Several research groups have therefore treated mouse and human B cells with Ca^{2+} indicator (*e.g.*, Fluo-4, Fura-2), introduced them into a microchamber array, stimulated them with antigens, analyzed the time-dependent fluorescence change, and then isolated positive single cells manually. Recent variants of these research projects have succeeded in isolating single B cells producing antigen-specific antibodies [13, 17, 18]. If such projects are combined with a single-cell robot equipped with time-lapse single-cell array cytometry (see below; [16]) in the future, these endogenously antibody-producing cells could be isolated in a high-throughput manner, significantly facilitating the generation of human monoclonal antibodies for antibody medications.

2.5.3 Yeast Cells Expressing Functional Mammalian Cytokine Receptors

Agonists and antagonists of cytokine receptors, recognized as important targets for drug discovery, have been screened from huge numbers of chemical libraries. Conventionally, each chemical compound has been examined using mammalian cells expressing the receptor of interest. However, preparing these chemical libraries and performing such global screenings using the chemical libraries and mammalian cells is costly, time-consuming, and laborious. Although the emergence of automated screening robots has reduced the labor required for these processes, it has not changed the methodology of drug screening. Furthermore, it is difficult to quantitatively measure the activation of receptors of interest in mammalian cells because endogenous signal cascades often interfere with the receptor-mediated signal cascade (due to unpredictable signaling crosstalks or similar phenomena). This work therefore tried to reconstitute the signal cascades of mammalian cytokine receptors in yeast cells, as yeast signaling molecules have negligible effects on mammalian receptors. This was successful in the functional expression of receptor Tyr kinase (epidermal growth factor receptor (EGFR); [21]), non-receptor Tyr kinase (interleukin-5 receptor (IL5R); [22]), and G protein-coupled receptors (somatostatin receptor (SSTR5); [7]) in yeast cells (Fig. 2.4). Next, in order to adapt the

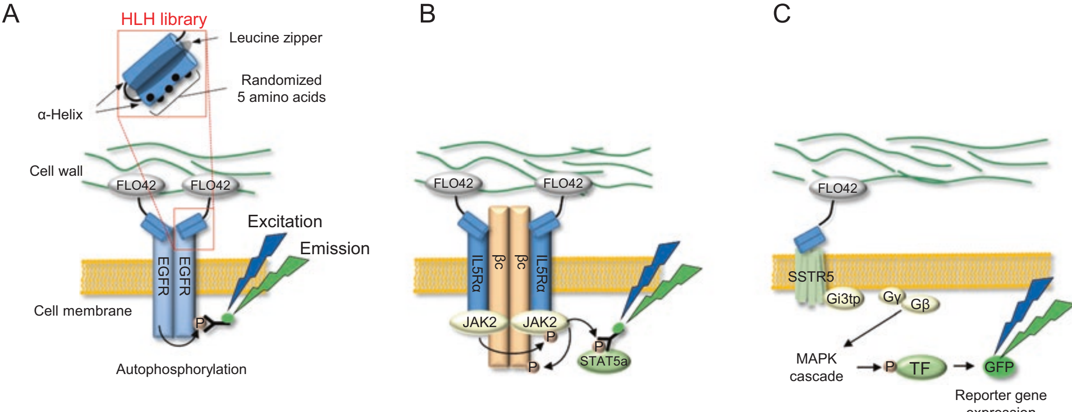


Fig. 2.4 *De novo* agonist screening with yeast cells expressing functional mammalian cytokine receptors. (**a**) EGFR. A cell wall-anchored form of HLH peptide library (HLH-FLO42) consisting of two α-helixes (blue) linked with a hepta-Gly loop, in which the second α-helix contains five randomized amino acids on its solvent-accessible surface. When HLH acts as an EGFR agonist, EGFR starts clustering and thereby auto-phosphorylating cytoplasmic Tyr residues form, which can be detected by fluorescence-labeled anti-phospho-Tyr antibodies. (**b**) IL5R. When HLH acts as an IL5R agonist, IL5R starts clustering, activating JAK2 kinase, phosphorylating transcription factor STAT5a, and thereby translocating STAT5a to nucleus, which can be detected by fluorescence-labeled anti-phospho-STAT5a antibodies. (**c**) SSTR5. When HLH acts as a SSTR5 agonist, SSTR5 starts activating the MAPK cascade via heterotrimeric GTP-binding protein, and thereby induces EGFP expression by MAPK-dependent Fig. 2.1 promoter

mammalian receptor-expressing yeast cells to the single-cell robot, a randomized peptide library (as a substitute for chemical library) was designed to display in the yeast cell walls via fusion with FLO42 peptide (a cell wall-anchoring moiety of FLO1; [15]). This facilitated the activation of proximal receptors in an autocrine manner as well as the subsequent identification of individual yeast cells displaying agonistic peptide ligands. Linear peptide libraries generally display flexible structures, which could show a weak, induced fit interaction with receptors and thereby unexpectedly generate pseudo-positive peptides. To address this issue, a helix-loop-helix (HLH) library was utilized, consisting of two α-helixes supported by leucine zipper interaction in which the second α-helix contained five randomized amino acids on its solvent-accessible surface (Fig. 2.4a; [4]). Since the HLH library shows rigid structure, it is expected to lower the possibility of emerging pseudo-positive peptides as well as allow for easy identification of the pharmacophores in positive HLH-based peptides for the *in silico* drug design of chemical compounds [3]. Cyclic peptides showing limited flexibility have recently attracted attention in the development of middle-molecule medicines (molecular mass range, 500–5000 Da) [23]. The affinity to receptors of HLH-based peptides (about 4 kDa) could be additionally increased by enhancing the structural rigidity via cyclization [5, 9, 12].

About 2.0×10^5 yeast cells coexpressing EGFR and the HLH library were fixed with paraformaldehyde, spheroplasted with Zymolyase, treated with fluorescence-labeled anti-phospho-Tyr antibodies, and then entered as a sample in the single-cell robot (Fig. 2.4a; [21]). Eight yeast cells exhibiting the highest fluorescent intensity were isolated, and the genes encoding HLH-based peptide were amplified via single-cell PCR. These HLH-based peptides expressed in *Escherichia coli* were purified and then added to the EGFR-overexpressing A431 cells. Finally, six out of eight HLH-based peptides were found to act as EGFR agonists with *de novo* structures. Besides expressing EGFR, yeast cells coexpress-

ing IL5α chain, β chain, JAK2 Tyr kinase, and STAT5a transcription factor were able to respond to exogenous IL5 stimulation, start the clustering of IL5α and β chains, activate JAK2 kinase, and then phosphorylate STAT5a. The nuclear translocation of STAT5a was detected by fluorescence-labeled anti-phospho-STAT5a antibodies (Fig. 2.4b; [22]). Meanwhile, yeast cells coexpressing SSTR5, a chimeric G protein (Gi3tp, a hybrid of human Gαi3 and yeast Gpa1), and EGFP (placed downstream of Fig. 2.1 promoter) were able to respond to exogenous somatostatin (SSTR5 ligand), initialize the activation of the MAPK cascade using Gi3tp protein, and then induce EGFP expression through MAPK-dependent Fig. 2.1 promoter (Fig. 2.4c; [7]). In combination, these results revealed that yeast cells could express a wide range of mammalian receptors and allow for agonist-dependent activation of these receptors. Relative to conventional receptor agonist screening procedures, the combination of the single-cell robot and yeast cells coexpressing mammalian receptors and HLH library shows promise for dramatically reducing the amount of time required (total 10 days from yeast library), lowering costs by eliminating the use of huge chemical libraries, and consequently contributing to the development of middle-molecule medicines with *de novo* structures.

2.5.4 Olfactory Sensory Neurons

Mammals are able to identify specific odorants in >10^5 species of odorants, although the number of mammalian odorant receptors (ORs) is limited (*e.g.*, ~400 ORs for humans and ~1100 ORs for mouse). At this time, there is no definitive answer for how such a small number of ORs is able to discriminate between this much greater numbers of odorants. One present working hypothesis is that a single OR could be activated to varying degrees by multiple different odorants and that this pattern of activated ORs is recognized by the central nerve system (for additional details, see a concept of OR repertory; [11]). To understand how each OR contributes to the separate identification of various odorants, it is essential to comprehensively analyze which ORs can be activated by specific odorants. However, the process of using current technologies to isolate all the olfactory sensory neurons (OSNs) responding to a specific odorant simultaneously in a single cell-based manner has proven too difficult. Since mammalian cells cannot express a fully functional form of mammalian ORs due to the inefficiency of membrane translocation and incorrect folding seen in ectopically synthesized ORs [14], it is impossible to prepare all human ORs in their functional forms simultaneously for the purpose of evaluating their activation by a specific odorant. Therefore, this work isolated primary mouse olfactory epithelial cells containing OSNs, treated them with Ca^{2+} indicator Fluo-4, introduced them into a microchamber array equipped with a perfusion apparatus, and then stimulated them with the odorant of interest (Fig. 2.5). When an OR was activated by an odorant, adenylate cyclase was activated by the OR through the $G\alpha_{olf}/\beta/\gamma$ complex, the subsequent production of cAMP activated cyclic nucleotide-gated ion channel (CNG), and consequently Fluo-4-derived

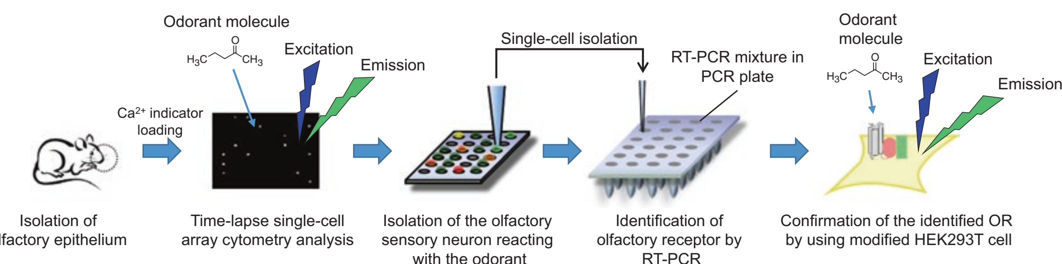


Fig. 2.5 Flow chart of high-throughput functional screening of mouse OSNs responding to specific odorants

fluorescence was increased by the influx of extracellular Ca^{2+}. Using a single-cell robot equipped with time-lapse single-cell array cytometry, it was possible to isolate all primary mouse OSNs responding to the specific odorants (*e.g.*, 2-pentanone, pyridine, 2-butanone) simultaneously by viewing the transient increases (average 30 s) of Fluo-4-derived fluorescence. The isolated ORs could then be reactivated by the respective odorant in modified HEK293T cells [16], strongly suggesting that the single-cell robot equipped with time-lapse single-cell array cytometry was able to both decipher the entirety of the mammalian OR repertory and deorphanize mammalian orphan receptors.

2.6 Conclusion

The proposed automated single-cell analysis and isolation system (the single-cell robot) has many advantages in comparison to FACS, although the high-throughput processing capability is still low (Table 2.1). In this chapter, many single cell-based applications of the single-cell robot were introduced which have been difficult or impossible up until this point using FACS. It is believed that the single-cell robot has nearly infinite possibilities for future applications, and the authors anticipate the creativity of researchers around the world in developing new single-cell applications.

References

1. Baba Y, Kurosaki T (2011) Impact of Ca2+ signaling on B cell function. Trends Immunol 32:589–594
2. Ben-David U, Benvenisty N (2011) The tumorigenicity of human embryonic and induced pluripotent stem cells. Nat Rev Cancer 11:268–277
3. El-Haggar R, Kamikawa K, Machi K, Ye Z, Ishino Y, Tsumuraya T, Fujii I (2010) Molecular design of small organic molecules based on structural information for a conformationally constrained peptide that binds to G-CSF receptor. Bioorg Med Chem Lett 20:1169–1172
4. Fujii I, Takaoka Y, Suzuki K, Tanaka T (2001) A conformationally purified α-helical peptide library. Tetrahedron Lett 42:3323–3325
5. Fujiwara D, Kitada H, Oguri M, Nishihara T, Michigami M, Shiraishi K, Yuba E, Nakase I, Im H, Cho S, Joung JY, Kodama S, Kono K, Ham S, Fujii I (2016) A cyclized helix-loop-helix peptide as a molecular scaffold for the design of inhibitors of intracellular protein-protein interactions by epitope and arginine grafting. Angew Chem Int Ed Engl 55:10612–10615
6. Gardner JS, Chiu AL, Maki NE, Harris JF (1985) A quantitative stability analysis of human monoclonal antibody production by heteromyeloma hybridomas, using an immunofluorescent technique. J Immunol Methods 85:335–346
7. Ishii J, Moriguchi M, Hara KY, Shibasaki S, Fukuda H, Kondo A (2012) Improved identification of agonist-mediated Gα(i)-specific human G-protein-coupled receptor signaling in yeast cells by flow cytometry. Anal Biochem 426:129–133
8. Jin A, Ozawa T, Tajiri K, Obata T, Kishi H, Muraguchi A (2011) Rapid isolation of antigen-specific antibody-secreting cells using a chip-based immunospot array. Nat Protoc 6:668–676
9. Kawabata K, Nagai H, Konishi N, Fujiwara D, Sasaki R, Ichikawa T, Fujii I (2014) Peptide-based immunoadsorbents: molecular grafting of IgG-Fc-binding epitopes of protein A onto a de novo-designed helix-loop-helix peptide. Bioorg Med Chem 22:1845–1849
10. Kida A, Iijima M, Niimi T, Maturana AD, Yoshimoto N, Kuroda S (2013) Cell surface-fluorescence immunosorbent assay for real-time detection of hybridomas with efficient antibody secretion at the single-cell level. Anal Chem 85: 1753–1759
11. Malnic B, Hirono J, Sato T, Buck LB (1999) Combinatorial receptor codes for odors. Cell 96:713–723
12. Nishihara T, Kitada H, Fujiwara D, Fujii I (2016) Macrocyclization and labeling of helix-loop-helix peptide with intramolecular bis-thioether linkage. Biopolymers 106:415–421
13. Ozawa T, Kinoshita K, Kadowaki S, Tajiri K, Kondo S, Honda R, Ikemoto M, Piao L, Morisato A, Fukurotani K, Kishi H, Muraguchi A (2009) MAC-CCD system: a novel lymphocyte microwell-array chip system equipped with CCD scanner to generate human monoclonal antibodies against influenza virus. Lab Chip 9:158–163
14. Saito H, Kubota M, Roberts RW, Chi Q, Matsunami H (2004) RTP family members induce functional expression of mammalian odorant receptors. Cell 119:679–691
15. Sato N, Matsumoto T, Ueda M, Tanaka A, Fukuda H, Kondo A (2002) Long anchor using Flo1 protein enhances reactivity of cell surface-displayed glucoamylase to polymer substrates. Appl Microbiol Biotechnol 60:469–474

16. Suzuki M, Yoshimoto N, Shimono K, Kuroda S (2016) Deciphering the receptor repertoire encoding specific odorants by time-lapse single-cell array cytometry. Sci Rep 6:19934
17. Tokimitsu Y, Kishi H, Kondo S, Honda R, Tajiri K, Motoki K, Ozawa T, Kadowaki S, Obata T, Fujiki S, Tateno C, Takaishi H, Chayama K, Yoshizato K, Tamiya E, Sugiyama T, Muraguchi A (2007) Single lymphocyte analysis with a microwell array chip. Cytometry A 71:1003–1010
18. Yamamura S, Kishi H, Tokimitsu Y, Kondo S, Honda R, Rao SR, Omori M, Tamiya E, Muraguchi A (2005) Single-cell microarray for analyzing cellular response. Anal Chem 77:8050–8056
19. Yoshimoto N, Kida A, Jie X, Kurokawa M, Iijima M, Niimi T, Maturana AD, Nikaido I, Ueda HR, Tatematsu K, Tanizawa K, Kondo A, Fujii I, Kuroda S (2013) An automated system for high-throughput single cell-based breeding. Sci Rep 3:1191
20. Yoshimoto N, Kuroda S (2014) Single-cell-based breeding: rational strategy for the establishment of cell lines from a single cell with the most favorable properties. J Biosci Bioeng 117:394–400
21. Yoshimoto N, Tatematsu K, Iijima M, Niimi T, Maturana AD, Fujii I, Kondo A, Tanizawa K, Kuroda S (2014) High-throughput de novo screening of receptor agonists with an automated single-cell analysis and isolation system. Sci Rep 4:4242
22. Yoshimoto N, Ikeda Y, Tatematsu K, Iijima M, Nakai T, Okajima T, Tanizawa K, Kuroda S (2016) Cytokine-dependent activation of JAK-STAT pathway in Saccharomyces cerevisiae. Biotechnol Bioeng 113:1796–1804
23. Zorzi A, Deyle K, Heinis C (2017) Cyclic peptide therapeutics: past, present and future. Curr Opin Chem Biol 38:24–29

Kenji Tatematsu, PhD, is an Assistant Professor of Biomolecular Sciences and Reaction, the Institute of Scientific and Industrial Research, Osaka University. His main research is focused on single cell-based bioengineering, odorant receptor array system, and design of ubiquitin ligase for degradation of specific intracellular proteins. He published more than 30 scientific papers in well-recognized international journals, three chapters of books, and seven patents.

Shun'ichi Kuroda, PhD, is a Professor of Biomolecular Sciences and Reaction, Vice Director, the Institute of Scientific and Industrial Research, Osaka University. His main research is focused on single cell-based bioengineering, odorant receptor array system, oriented immobilization of biosensing molecules, and virus-mimicking drug delivery systems. He published more than 170 scientific papers in well-recognized international journals, 54 reviews, 54 chapters of books, and 57 patents. He is also member of the Editorial board of four international scientific journals.

3 Single-Cell Non-coding RNA in Embryonic Development

Qiang Fu, Chuan-Jiang Liu, Zhen-Sheng Zhai, Xu Zhang, Tao Qin, and Hong-Wei Zhang

Abstract

Non-coding RNAs (ncRNAs) have significant regulatory functions on the regulation of gene expression of various life activities after transcription, even though they do not encode proteins. During the development of embryos, ncRNAs, such as long non-coding RNAs (lncRNAs), microRNAs (miRNAs), circular RNAs (circRNAs), small nucleolar RNAs (snoRNAs), and Piwi-interacting RNAs (piRNAs), have been widely proven as key regulators. The emerging single-cell RNA sequencing technique is powerful for profiling "cell-to-cell" variability at the genomic level. It has been applied to detect the expression of ncRNAs during embryo development. In this chapter, we pay close attention to single-cell ncRNA expression and summarize their roles in embryo development.

Keywords

Single-Cell · Non-coding RNA · Embryo

Q. Fu · C.-J. Liu · Z.-S. Zhai · X. Zhang
T. Qin (✉) · H.-W. Zhang (✉)
Department of Surgery, Henan Provincial People's Hospital, Zhengzhou, Henan Province, China

3.1 Introduction

The process of embryo development involves the proliferation and differentiation of embryonic stem cells (ESCs). This dynamic and complicated process is coupled with an orchestrated reorganization of the epigenome that shapes the chromatin environment to prepare for the subsequent developmental stages [1]. ESCs, which are inner cells isolated from the blastocyst at d4 or d5, possess self-renewal and pluripotency properties [2].

The mammalian genome is classified into two clusters by function, "coding" and "non-coding" regions. There are many genes involved in the regulation of embryo development including ncRNAs. Most of the non-coding genome, which is transcribed into RNAs but does not code for proteins, was originally thought to be junk DNA. Only a small number of genes, which encode proteins, and some structural non-coding RNAs have been considered as essential for cell functions (Fig. 3.1). Small structural RNAs that include small nuclear RNAs, snoRNAs, rRNAs, and tRNAs can regulate gene splicing, ribosome biogenesis and translation [3–6]. Additional research in recent years has demonstrated that ncRNAs, such as lncRNAs, miRNAs, circRNAs, snoRNAs, and piRNAs, are involved in the regulation of gene expression, especially embryo development and cell regulation [7–11].

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_3

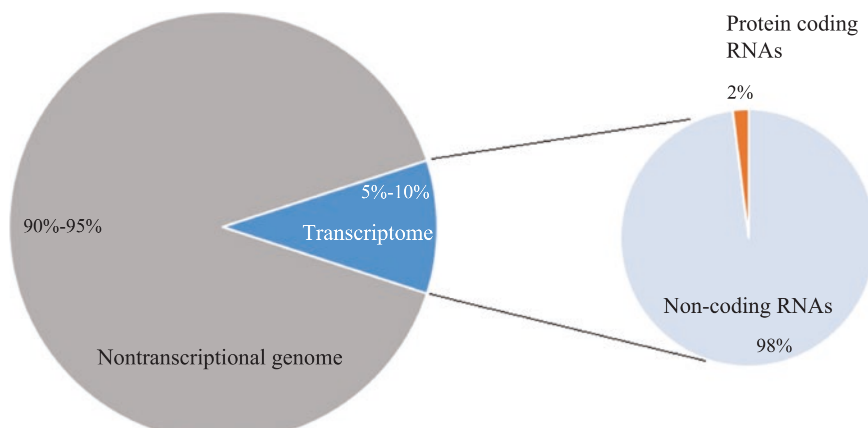


Fig. 3.1 Approximately 5%–10% of the human genome sequence is stably transcribed. NcRNAs constitute a majority (98%) of the transcriptome with protein-coding RNAs making up the rest (2%)

An emerging technique, single-cell RNA sequencing, is powerful for profiling "cell-to-cell" variability at the genomic level and has been broadly applied for genetic analysis. Analysis of single cells utilizing various tools might impact the current conceptual understanding of diverse biological processes. Studying cell genomes in individual cells provides a rare opportunity to dissect the interactions between the inner cell process and external stimuli. For example, cell fate can depend on these circumstances [12]. In this article, we focus our attention on the expression of ncRNAs in single-cell ESCs and preimplantation embryos and then summarize their role in embryo development.

3.2 Single-Cell lncRNAs in Embryonic Development

LncRNAs, a class of RNAs longer than 200 nucleotides, are mainly distributed in the nucleus, though they also are present in the cytoplasm. LncRNAs were originally considered genomic transcription "noise" and a by-product of RNA polymerase II transcription without biological functions. However, further studies identified more and more lncRNAs. LncRNAs display numerous features, types and multi-modes of action [13]. They can be classified into intergenic lncRNAs and intragenic lncRNAs (lincRNAs) according to their genome localization. LincRNAs can also be further subdivided into the following four categories depending on the lncRNA position related to an associated protein-coding gene: "sense lncRNAs", "antisense lncRNAs", "bidirectional lncRNAs" and "intronic lncRNAs" [14].

LncRNAs play an important role in biological processes including epigenetic regulation, dose compensation, cell cycle and cell differentiation, proliferation, apoptosis through gene imprinting, chromatin remodelling, transcriptional activation, transcriptional interference, nuclear splicing regulation, mRNA degradation and translational regulation. More and more studies have shown that lncRNAs are associated with various diseases including cancer, metabolic diseases, neurodegenerative diseases and psychiatric disorders, cardiovascular disease and hypertension, immune dysfunction and autoimmune diseases, especially in embryo development [15, 16]. Various functional lncRNAs are important during embryonic development. Some of them potentially decide one cell's fate and differentiation direction during embryogenesis to finally form different organs or special tissues that contain various cells expressing specific and stable genes [17].

X-inactive specific transcript (Xist), the first identified lncRNA associated with embryo development, is important for the induction of mammalian X-chromosome cis-inactivation.

In individual human embryo cells, Xist expression begins at the 4-cell stage on d3 and continues asynchronously; it is essential and pivotal for X-genome inactivation and mammalian dosage compensation [18, 19]. Abnormal expression of some lincRNAs can result in a number of human development disorders, such as DBE-T, SNORD115 and SNORD116 for "facioscapulohumeral muscular dystrophy", "Prader-Willi Syndrome" [20, 21].

In 2013, a study showed that there were 2733 novel lncRNAs among 8701 lncRNAs using single-cell sequencing analysis of 124 individual cells from human ESCs and human preimplantation embryos at different passages, and many of these lincRNAs were expressed in specific developmental stages [22]. Julien et al. completed a meta-analysis of RNAs in this article and another study at the single-cell level, though they did not analyse lncRNA profiles specifically [23]. The results showed lncRNAs expression from the beginning of human embryo development; lncRNAs became the predominant category of transcripts after embryonic gene activation at the 4-cell to 8-cell stage. Some lncRNAs, which were detected at crucial and specific steps of embryonic development, could represent candidate markers of embryonic gene activation and embryo competency. These results provide strong evidence to support the notion that lncRNAs are integral to the dynamic changes in transcript expression that occurs during human early embryo development and participate in the dramatic morphological changes during development [7]. A previous report performed a transcriptome assembly and identified 4438 putative lincRNAs in mice before embryo implantation depending on the published single-cell RNA sequencing. Moreover, a referable catalogue, which included 5808 lincRNAs in preimplantation embryonic development, was established and systematically analysed the lincRNAs. The catalogue of mouse pre-implantation embryonic development lincRNAs is now a valuable database for deep functional analysis and is helpful for elucidating regulatory mechanisms before embryo implantation [24].

"Weighted gene co-expression network analysis" is a powerful technique and effective method that is used to broadly analyse and discover the relationships between genes, networks and phenotypes [25]. By using expression profile analysis and "weighted gene co-expression network analysis," Qiu et al. completed a comprehensive and comparison analysis of lncRNAs during embryonic development using existing single-cell RNA sequencing data from humans and mice. The results showed that lncRNAs were expressed a "developmental stage-specific manner" pattern during the human early embryo stage that was different from the "temporal-specific expression pattern" during mouse embryo development. "Weighted gene co-expression network analysis" revealed that lncRNAs were tightly connected with mitochondrial functions and numerous significant processes, including oocyte maturation and zygotic genome activation during early-stage embryonic development. Additionally, lncRNA networks involved in zygotic genome activation were conserved in both human and mouse embryos, whereas no strong correlation was observed in other stages between human and mouse embryos [26]. In general, lncRNAs play an absolutely necessary role in embryonic development.

3.3 Single-Cell miRNAs in Embryonic Development

MiRNAs, which are approximately 22–25 nucleotides long, are endogenously small ncRNAs that play important roles in mammals [27]. Lin-4 was the first discovered miRNA; it regulates Caenorhabditis elegans development by repressing lin-14 expression [28]. To date, more than 15000 miRNAs have been studied in over 140 species [29].

Several kinds of RNA enzymes are involved in the step-by-step generation of mature miRNAs. First, "RNA polymerase II" transcribes the pri-miRNA. Then, "nuclear RNase III Drosha" and "DGCR8/Pasha" process the pri-miRNA into the pre-miRNA, which is approximately 60–70 nucleotides in length and possess a "stem loop

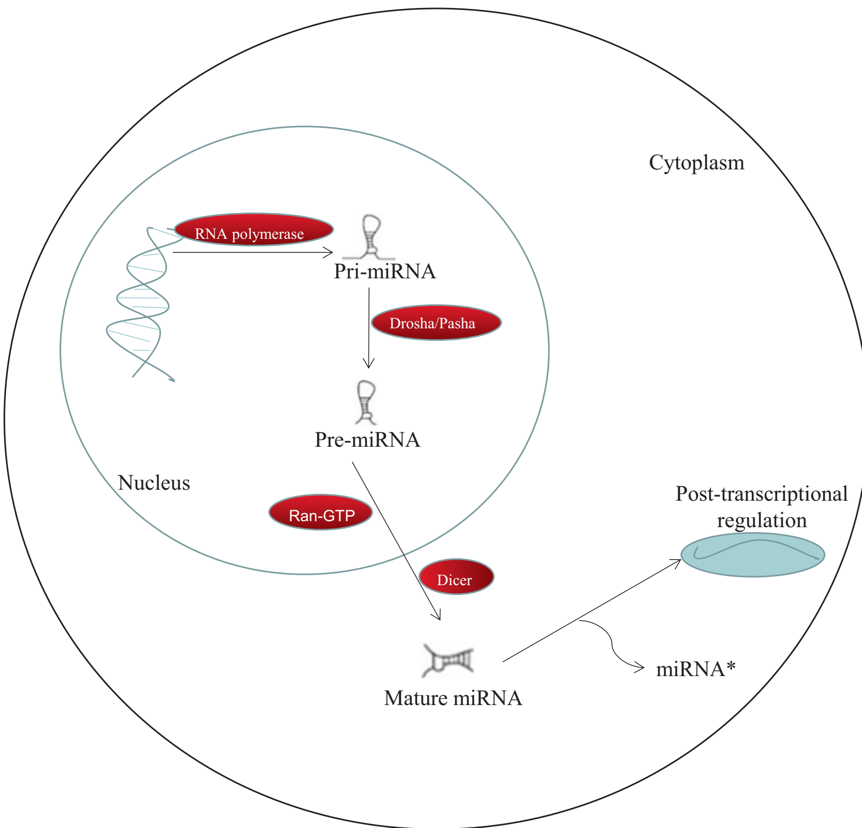


Fig. 3.2 The miRNA biogenesis pathway. RNA polymerase II, nuclear RNase III Drosha and DGCR8/Pasha, process the pre-miRNA. Ran-GTP transports the pre-miRNA to the cytoplasm where it is further processed by a second RNase III, Dicer, which generates a mature small miRNA. After mature miRNAs release their complementary miRNA*, then post-transcriptionally regulate function

structure" [30–32]. "Ran-GTP" transports the pre-miRNA to the cytoplasm where it is further processed by a second RNase III, Dicer, which generates a mature small miRNA by cutting off the terminal loop. After mature miRNAs release their complementary miRNA*, then post-transcriptionally regulate function (Fig. 3.2). For example, miR-135a and miR-22, which we discovered in rat, could bind to ErbB3 and Ptk2 mRNA, respectively, and trigger either translation inhibition or mRNA degradation (Fig. 3.3) [5, 33, 34].

Most studies have demonstrated that miRNAs can cause target mRNA degradation if they match them perfectly; in contrast, miRNAs repress target mRNAs when imperfect target duplexes are formed [35, 36]. In recent years, the functions of miRNAs have been to be complex and that miRNAs can up-regulate target genes; however, this requires further studies [37]. MiRNAs are involved in a variety of cell types under physiological and pathological conditions and are essential in various cellular processes such as "development, proliferation, apoptosis, metabolism and morphogenesis" [38, 39].

MiRNAs have been demonstrated as significant regulators of ESCs. ESCs with genetically deleted Dicer−/− or Dgcr−/−, both of which are essential for miRNA maturation, show abnormal differentiation [40, 41]. DGCR8-deficient ESCs express a few differentiation markers, but some of the pluripotency markers can be still detected, which confirms the important function of miRNAs in ESC differentiation [42]. Studies on

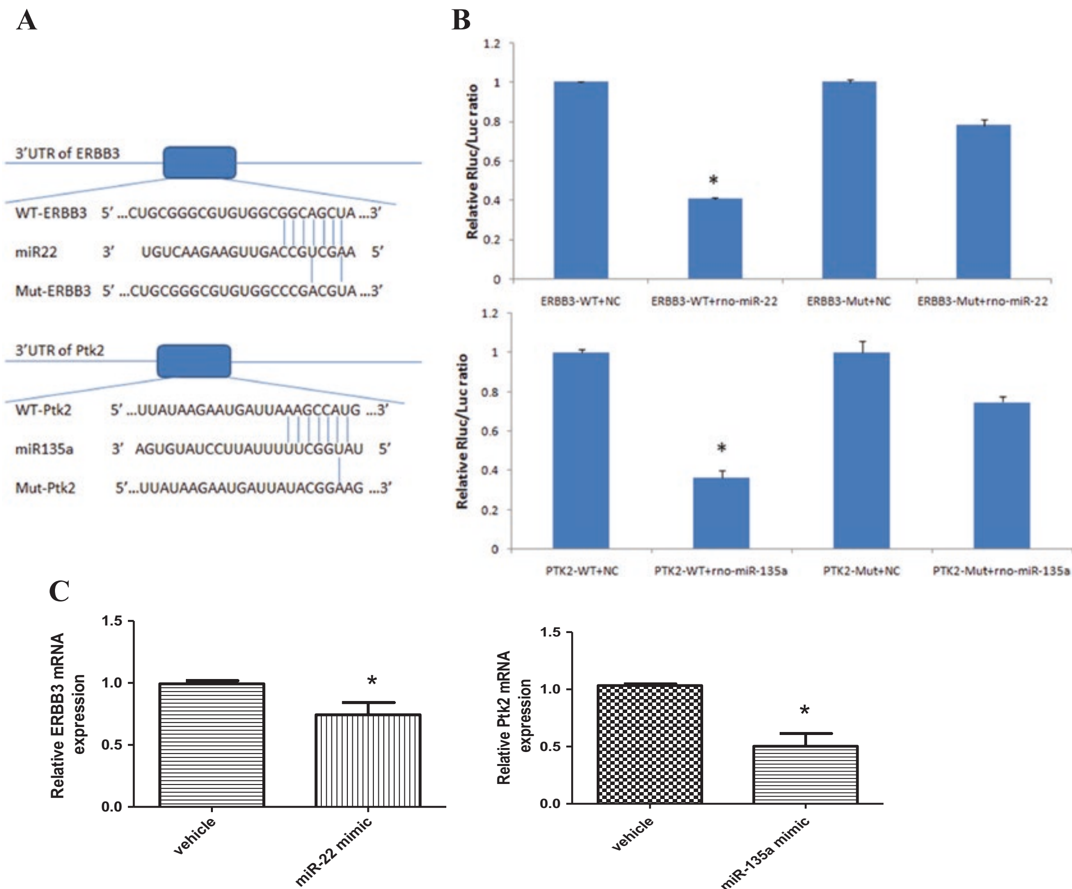


Fig. 3.3 miR-22 directly targets ERBB3 and miR-135a directly targets Ptk2 by binding to the 3'UTR. (**a**)The predicted miR-22 (miR-135a) binding sites within the 3'UTR of ERBB3 (Ptk2) and mutant version generated by site mutagenesis are shown. (**b**) Luciferase activity was determined 48 h after transfection. The ratio of normalized sensor to control luciferase activity is shown. Data are shown as the mean ± S.D. and were obtained from three independent experiments performed in triplicate. *Significant difference from control miR-transfected cells ($P < 0.05$). (**c**) Quantitative real time RT-PCR analysis of ERBB3 and Ptk2 expression in AR42J cells after transfection

miRNAs in ESC differentiation have been rapidly increasing. By directly binding to the 3'-UTR of pluripotency factors, miRNAs can regulate the state of ESCs. For example, miR-145 decreases ESC pluripotency by targeting Oct4, Sox2 and Klf4 in humans [43]. Additionally, the coding region of Nanog, Oct4 and Sox2 can also be targeted by miR-296, miR-470, and miR-134, which are involved in the regulation of mouse ESC differentiation [44]. More and more studies have shown that ESC-specific miRNAs, c-Myc-induced miRNAs, early embryonic miRNA clusters and p53-targeting miRNAs can regulate ESC self-renewal, reprogramming and differentiation [45–48]. MiRNAs also play a regulatory role in the early stages of embryonic development. Mouse Dicer−/− oocytes are deficient in spindle organization and chromosomal alignment. Even when they are fertilized with wild-type sperm, the mouse zygotes cannot divide into two daughter cells [49, 50]. Recent studies in zebrafish embryos under hypoxic conditions have shown that miR-125c plays an important role in inhibiting ESC proliferation, these actions lead to cell cycle arrest in G1 phase and the induction apoptosis, which helps embryos adapt to anoxic environments [51].

MiRNA expression profiling of ESCs at the single-cell level was first undertaken using a "real-time PCR-based 220-plex miRNA expression profiling method", which was essential to illuminate the diverse roles of miRNAs. The miRNA profiling data for the 220 miRNAs correlated well with cloning and northern blot data, which proved that this expression profiling method worked reliably for individual ESCs [52]. In another study, a dual-fluorescence reporter/sensor plasmid was imported into zebrafish blastomeres and mouse embryos via different methods to detect the dynamics of specific miRNAs at the single-cell level. The results revealed that both post-mitotic neurons and neuronal progenitor cells from developing mice express miR-124a, whereas miR-1 cannot be detected in the early stage of zebrafish development. This technique allows for the monitoring of miRNA appearance and disappearance during vertebrate development [53]. Another analysis was performed on 330 miRNAs using single-cell RNA-Seq analysis during the transition from the mouse blastocyst inner cell mass to pluripotent ESCs. The results indicated that pluripotency-related miRNAs, such as miR-290 and the miR-295 cluster, are expressed similarly in the inner cell mass cells and ESCs, while 51 miRNAs showed differential expression. For example, the let-7 family expression was reduced approximately 4–12-fold in ESCs compared to the inner cell mass cells, whereas the expression of miR-302c and miR-367 increased by 5- and 33-fold, respectively. Further study revealed that miRNAs could target the genes driving differentiation or ESC-specific pluripotency, thus contributing to maintaining the balance between pluripotency and the potential for rapid ESC differentiation [54].

A recent study performed in 2016 on single-cell miRNAs in embryo development described a new method for sequencing the small RNA transcriptome from individual naive and primed human ESCs. First, libraries were constructed by ligating adapters to whole RNA species possessing 5′-phosphate and 3′-hydroxyl groups regardless of their size. Then, designated masking oligonucleotides were used to evade the highly abundant 5.8S rRNAs followed by an enzymatic digestion step to reduce the formation of adaptor dimers. Finally, unique molecular identifiers were added to the 5′-adapters to counteract PCR stochasticity and enable RNA molecule counting. No experimental size-selection step was used, and the small RNAs were identified computationally [55, 56]. Sixty percent of 3800 captured miRNAs were differentially expressed between naive and primed human ESCs, which is consistent with miRNA profiling data from naive and primed cells in bulk. Notably, the miR-302 family, which regulates the cell cycle and apoptosis in human ESCs, displayed increased expression in primed human ESCs. The expression of the miR-371-3 cluster, which is crucial for the maintenance of human ESC pluripotency, was significantly high in naive cells [57]. In summary, miRNAs, which are considered the most important small non-coding RNAs, showed different expression profiles at various stages and play a highly important role in embryonic development.

3.4 Single-Cell circRNAs in Embryo Development

CircRNAs, a new special class of endogenous ncRNAs, are formed by back-splicing events via the circularization of an exon or intron [58]. CircRNAs are characterized by a covalently closed loop without a "5′-cap" or "3′-Poly A tail" [59]. They were thought to be the results of splicing errors for several decades after they were discovered in RNA viruses in 1976 [60]. The features of circRNAs, including abundance, stability, conservation and tissue-specific expression, were revealed by RNA-sequencing technologies and bioinformatics [61, 62]. CircRNAs can be classified into the following four categories: "exonic circRNAs", "circular intronic RNAs", "exon-intron circRNAs" and "intergenic circRNAs", based on their different biogenesis patterns from genomic regions [61, 63–66].

To date, many studies have shown that circRNAs have numerous biological functions, such as sequestering proteins from their native subcellular localizations, regulating alternative splicing, modulating the expression of parental

genes, functioning as miRNA sponges, and acting as scaffolds in the assembly of protein complexes and RNA–protein interactions [67–73]. Some reports have demonstrated that altered circRNAs are associated with human diseases, such as hsa-circRNA-103,636 in Major depressive disorder (MDD) and hsa-circ-002059 in gastric carcinoma [74, 75].

During embryonic development, the induction of circRNAs is a universal phenomenon; for example, the circRNA Sry is highly expressed in mice [76]. Another study showed that the introduction of circRNAs was important for directing human pluripotent stem cells to differentiate into cardiomyocytes. The developmental induction of circRNA from the NCX1 gene, which codes for a calcium transporter essential for cardiac development, is conserved between mouse and human, indicating that the developmental regulation of circRNAs may be evolutionarily conserved [77].

Single-cell universal poly(A)-independent RNA sequencing (SUPeR-seq), which utilizes random primers with fixed anchor sequences for cDNA synthesis, is a novel and highly robust single-cell transcriptome profiling method that can detect both poly(A+) and poly(A-) RNAs from individual cells. Fan et al. applied SUP-eR-seq to detect polyadenylated and non-polyadenylated RNAs from individual mouse oocytes and early embryos and discovered 2891 circRNAs. Further analysis showed that circRNAs began to express in mature oocytes, increased until the 8-cell stage, and then declined. They also selected 8 circRNAs for independent validation and verified 7 of them at single-base resolution by standard Sanger sequencing and RT-PCR. The results indicated that the most vital circRNAs were authentic and deciphered the regulation mechanisms for circRNAs in mouse early embryos [78].

3.5 Single-Cell snoRNAs in Embryo Development

SnoRNAs, which are approximately 60–300 nt in length, are another class of ncRNAs that are found in the cell nucleolus [79]. They are the targeting components of enzymes that methylate, cleave, or pseudouridylate ribosomal RNAs [80]. “RNA polymerase II” is involved in transcription of snoRNAs, which are divided into the following two types: “box H/ACA” and “box C/D”. “Box H/ACA” snoRNAs bind to the conserved core “box H/ACA” snoRNP proteins and can guide pseudouridylation of target rRNAs. In contrast, “boxC/D” snoRNAs (NOP56, NOP5/NOP58, NHP2L1, and fibrillarin) can direct 2′-O-ribose methylation of nucleotides by binding to conserved core “box C/D” snoRNP proteins [81, 82]. Both modifications, which are essential for RNA maturation, are introduced concurrently or immediately after rRNA operon transcription or before 45S rRNA cleavage. Some special snoRNAs, such as U3, can direct 45S rRNA pyrolysis steps, but not chemical modifications [83]. Small Cajal body-specific RNAs (scaRNAs), which are similar to snoRNAs and considered special snoRNAs, accumulate within the Cajal bodies, which are conserved subnuclear organelles present in the nucleoplasm. Small Cajal body-specific RNAs possess characteristic boxes for both the “box H/ACA” and “box C/D” regions as well as a “CAB box” (UGAG), so they are longer than the predominant classes of snoRNAs [84].

The location of snoRNAs in the nucleolus indicates their functions as guides for the post-transcriptional modification of ribosomal and some spliceosomal RNAs involved in the nucleolytic processing of original ribosomal RNA transcripts [85, 86]. These post-transcriptional modifications are essential for producing accurate and efficient ribosomes. Small Cajal body-specific RNAs can also target other RNAs, including tRNAs and mRNAs, and have some functions that remain to be clarified. In vertebrates, many small Cajal body-specific RNAs reside in the introns of protein-coding host genes and are processed out of the excised introns [80, 86, 87]. Furthermore, approximately 60% of snoRNAs can be processed into miRNAs, though most of their targets have not been confirmed [88].

The mutation of snoRNA genes has been demonstrated to be associated with various

human diseases. For example, losing paternally expressed imprinted genes, including the "HBII-52" and "HBII-85" snoRNAs, caused the neurogenetic disorder "Prader-Willi syndrome" (PWS) [89, 90]. Suppression of U26, U44 and U78 expression in zebrafish by disrupting host gene splicing or inhibiting snoRNA precursor processing reduced the snoRNA-guided methylation of target nucleotides. This could lead to severe morphological defects and zebrafish embryo lethality, which suggests that ribosomal RNA modifications play an important role in vertebrate development [91]. The dyskerin ribonucleoprotein complex, which regulates the expression of pluripotency genes that are critical for ESC self-renewal, has also been proven to be controlled by snoRNAs [92].

Using the new method described above, Faridani et al. detected the expression of snoRNAs in individual naive and primed human ESCs. On average, there were approximately 600 snoRNAs per cell. The snoRNA abundance profiles indicated cell-type specificity that corresponded with the variable expression of snoRNAs across cell types. However, the cell-type specificity was not as pronounced as that of miRNAs [57].

3.6 Single-Cell piRNAs in Embryonic Development

The last small ncRNA, piRNAs, is a class that is approximately 24–35 nt in length and is enriched in mammalian gonads and germ cells [93]. They are structurally characterized by 2′-O-methyl modification sites at their 3′-terminus and are processed from single-stranded precursor transcripts transcribed by intergenomic regions termed piRNA clusters via a Dicer-independent mechanism [94]. The first report of piRNAs as a class of RNAs was obtained from male mouse germ cells in 2006 by Aravin et al. [81]. Further studies showed that piRNAs can be divided into the following three classes depending on the genomic localization: simple repetitive, repeat-associated and non-repetitive piRNAs, which can be further classified into intergenic, intronic and exonic piRNAs [95, 96].

PiRNAs were first discovered to be associated with PIWI subfamily proteins (PIWI proteins), a subfamily of Argonaute proteins, and then to form piRNA complexes (pIRICS), which are silencing transposable elements in animal gonads [11]. The regulation of piRNAs is mainly focused on the activation of transposons because various transposons are located in piRNA clusters. There is a high risk of intracellular genome damaging during the transposition of transposons; therefore, the piRNA-mediated regulation of transposons is essential, especially for preserving normal gametogenesis and reproduction [97]. PiRNA-mediated regulation can achieve self and non-self recognition similar to the immune systems, and a complex mechanism is used to effectively select and regulate non-self genes [98]. Since the discovery of piRNAs, their functions as RNAs that ensure the repression of transposable element activity in germ cells have been verified comprehensively, meaning that they maintain genome integrity [99, 100]. It is has been demonstrated that piRNAs are necessary for successful spermatogenesis in the gonads of mammals [101]. Further studies confirmed the epigenetic regulatory role and gene silencing role of piRNAs, indicating that piRNAs in the testicles can regulate gene expression related to testicular development or conditions, even in embryonic development [102, 103].

XIWI, a Xenopus PIWI homologue, and piRNAs were characterized systematically from Xenopus eggs and oocytes. Small RNA analysis at the single-cell level showed abundant piRNAs and piRNA clusters located in the Xenopus tropicalis genome. However, in early embryos, XIWI staining was lost from the mitotic spindles and localized exclusively to germplasm islands that became restricted to a few blastomeres during the course of development, suggesting that piRNAs and XIWI were essential components of the germplasm [104]. The expression and function of piRNAs at the single-cell level in various stages of embryonic development needs further study.

3.7 Conclusions

Since the invention and development of single-cell RNA-seq transcriptome analysis technologies, they have been widely used in embryonic development research due to their powerful function. NcRNAs have been widely proven as key regulators of mammalian embryonic development. Using single-cell RNA-seq techniques, the differential abundance and expression of non-coding RNAs at different stages of embryonic development have been examined and identified. Some non-coding RNAs are essential for embryonic development, especially lncRNAs and miRNAs, which have been studied extensively. The reduction, deletion or overexpression of non-coding RNAs can result in abnormal embryonic development, even leading to congenital diseases or embryonic death. According to current non-coding RNA databases, single-cell studies focused on lncRNAs and miRNAs are more complicated than other non-coding RNAs. Moreover, single-cell circRNAs, piRNAs and snoRNAs need to be studied further in embryonic development.

Acknowledgements and Conflicting Declaration Thanks to Professor Xiang-dong Wang and Jian-qin Gu for their help and support. We declare that we have no financial relationships with other people or organizations that can inappropriately influence our work, Additionally, there are no professional or other personal interests of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled "Single-cell non-coding RNAs in embryonic development".

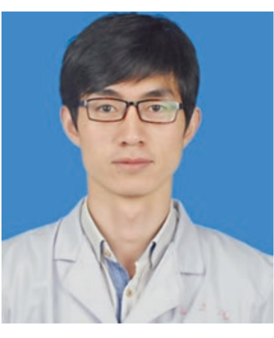
References

1. Kraushaar DC, Zhao K (2013) The epigenomics of embryonic stem cell differentiation. Int J Biol Sci 9(10):1134–1144
2. Bradley A, Evans M, Kaufman MH, Robertson E (1984) Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature 309(5965):255–256
3. Mattick JS, Makunin IV (2006) Non-coding RNA. Hum Mol Genet 15(1):R17–R29
4. Rinn JL, Chang HY (2012) Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145–166
5. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136(2):215–233
6. Seto AG, Kingston RE, Lau NC (2007) The coming of age for Piwi proteins. Mol Cell 26(5):603–609
7. Bouckenheimer J, Assou S, Riquier S, Hou C, Philippe N, Sansac C, Lavabre-Bertrand T, Commes T, Lemaître JM, Boureux A et al (2016) Long non-coding RNAs in human early embryonic development and their potential in ART. Hum Reprod Update 23(1):19–40
8. Tulay P, Sengupta SB (2016) MicroRNA expression and its association with DNA repair in preimplantation embryos. J Reprod Dev 62(3):225–234
9. Venø MT, Hansen TB, Venø ST, Clausen BH, Grebing M, Finsen B, Holm IE, Kjems J (2015) Spatio-temporal regulation of circular RNA expression during porcine embryonic brain development. Genome Biol 16:245
10. Skreka K, Schafferer S, Nat IR, Zywicki M, Salti A, Apostolova G, Griehl M, Rederstorff M, Dechant G, Hüttenhofer A (2012) Identification of differentially expressed non-coding RNAs in embryonic stem cell neural differentiation. Nucleic Acids Res 40(13):6001–6015
11. Huang Y, Bai JY, Ren HT (2014) PiRNAs biogenesis and its functions. Bioorg Khim 40(3):320–326
12. Saliba AE, Westermann AJ, Gorski SA, Vogel J (2014) Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res 42(14):8845–8860
13. Spizzo R, Almeida MI, Colombatti A, Calin GA (2012) Long non-coding RNAs and cancer: a new frontier of translational research? Oncogene 31(43):4577–4587
14. Zhu L, Zhu J, Liu Y, Chen Y, Li Y, Huang L, Chen S, Li T, Dang Y, Chen T (2015) Methamphetamine induces alterations in the long non-coding RNAs expression profile in the nucleus accumbens of the mouse. BMC Neurosci 16:18
15. Xu D, Yang F, Yuan JH, Zhang L, Bi HS, Zhou CC, Liu F, Wang F, Sun SH (2013) Long noncoding RNAs associated with liver regeneration 1 accelerates hepatocyte proliferation during liver regeneration by activating Wnt/β-catenin signaling. Hepatology 58(2):739–751
16. Perry RB, Ulitsky I (2016) The functions of long noncoding RNAs in development and stem cells. Development 143(21):3882–3894
17. Rivera RM, Ross JW (2013) Epigenetics in fertilization and preimplantation embryo development. Prog Biophys Mol Biol 113(3):423–432
18. Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N (1996) Requirement for Xist in X chromosome inactivation. Nature 379(6561): 131–137
19. Briggs SF, Dominguez AA, Chavez SL, Reijo Pera RA (2015) Single-Cell XIST expression in human preimplantation embryos and newly reprogrammed female induced pluripotent stem cells. Stem Cells 33(6):1771–1781

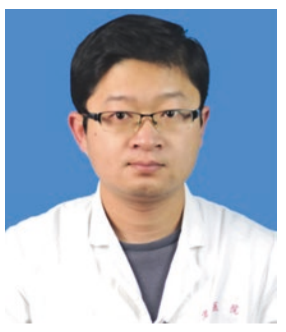
20. Chen LL, Zhao JC (2014) Functional analysis of long noncoding RNAs in development and disease. Adv Exp Med Biol 825:129–158
21. Bortolin-Cavaillé ML, Cavaillé J (2012) The SNORD115 (H/MBII-52) and SNORD116 (H/MBII-85) gene clusters at the imprinted Prader-Willi locus generate canonical box C/D snoRNAs. Nucleic Acids Res 40(14):6800–6807
22. Yan L, Yang M, Guo H, Yang L, Wu J, Li R, Liu P, Lian Y, Zheng X, Yan J et al (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 20(9):1131–1139
23. Xue Z, Huang K, Cai C, Cai L, Jiang CY, Feng Y, Liu Z, Zeng Q, Cheng L, Sun YE et al (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500(7464):593–597
24. Lv J, Liu H, Yu S, Liu H, Cui W, Gao Y, Zheng T, Qin G, Guo J, Zeng T et al (2015) Identification of 4438 novel lincRNAs involved in mouse pre-implantation embryonic development. Mol Gen Genomics 290(2):685–697
25. Liu R, Zhang W, Liu ZQ, Zhou HH (2017) Associating transcriptional modules with colon cancer survival through weighted gene co-expression network analysis. BMC Genomics 18(1):361
26. Qiu JJ, Ren ZR, Yan JB (2016) Identification and functional analysis of long non-coding RNAs in human and mouse early embryos based on single-cell transcriptome data. Oncotarget 7(38):61215–61228
27. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2):281–297
28. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–854
29. Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39(Database issue):D152–D157
30. Cai X, Hagedorn CH, Cullen BR (2004) Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. RNA 10(12):1957–1966
31. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Rådmark O, Kim S et al (2003) The nuclear RNase III Drosha initiates microRNA processing. Nature 425(6956):415–419
32. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN (2004) The Drosha-DGCR8 complex in primary microRNA processing. Genes Dev 18(24):3016–3027
33. Qin T, Fu Q, Pan YF, Liu CJ, Wang YZ, Hu MX, Tang Q, Zhang HW (2014) Expressions of miR-22 and miR-135a in acute pancreatitis. J Huazhong Univ Sci Technolog Med Sci 34(2):225–233
34. Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120(1):15–20
35. Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF (2006) Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. Science 312(5770):75–79
36. Petersen CP, Bordeleau ME, Pelletier J, Sharp PA (2006) Short RNAs repress translation after initiation in mammalian cells. Mol Cell 21(4):533–542
37. Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. Science 318:1931–1934
38. Altana V, Geretto M, Pulliero A (2015) MicroRNAs and physical activity. Microrna 4(2):74–85
39. Wei W, Hou J, Alder O, Ye X, Lee S, Cullum R, Chu A, Zhao Y, Warner SM, Knight DA et al (2013) Genome-wide microRNA and messenger RNA profiling in rodent liver development implicates mir302b and mir20a in repressing transforming growth factor-beta signaling. Hepatology 57(6):2491–2501
40. Melton C, Judson RL, Blelloch R (2010) Opposing microRNA families regulate self-renewal in mouse embryonic stem cells. Nature 463(7281):621–626
41. Kanellopoulou C, Muljo SA, Kung AL, Ganesan S, Drapkin R, Jenuwein T, Livingston DM, Rajewsky K (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19(4):489–501
42. Wang Y, Medvid R, Melton C, Jaenisch R, Blelloch R (2007) DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet 39(3):380–385
43. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS (2009) MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell 137(4):647–658
44. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 455(7216):1124–1128
45. Wang Y, Baskerville S, Shenoy A, Babiarz JE, Baehner L, Blelloch R (2008) Embryonic stem cell-specific microRNAs regulate the G1-S transition and promote rapid proliferation. Nat Genet 40(12):1478–1483
46. Lin CH, Jackson AL, Guo J, Linsley PS, Eisenman RN (2009) Myc-regulated microRNAs attenuate embryonic stem cell differentiation. EMBO J 28(20):3157–3170
47. Tata PR, Tata NR, Kühl M, Sirbu IO (2011) Identification of a novel epigenetic regulatory region within the pluripotency associated microRNA cluster, EEmiRC. Nucleic Acids Res 39(9):3574–3581
48. Neveu P, Kye MJ, Qi S, Buchholz DE, Clegg DO, Sahin M, Park IH, Kim KS, Daley GQ, Kornblum HI et al (2010) MicroRNA profiling reveals two distinct p53-related human pluripotent stem cell states. Cell Stem Cell 7(6):671–681

49. Tang F, Kaneda M, O'Carroll D, Hajkova P, Barton SC, Sun YA, Lee C, Tarakhovsky A, Lao K, Surani MA (2007) Maternal microRNAs are essential for mouse zygotic development. Genes Dev 21(6):644–648
50. Murchison EP, Stein P, Xuan Z, Pan H, Zhang MQ, Schultz RM, Hannon GJ (2007) Critical roles for Dicer in the female germline. Genes Dev 21(6):682–693
51. He Y, Huang CX, Chen N, Wu M, Huang Y, Liu H, Tang R, Wang WM, Wang HL (2017) The zebrafish miR-125c is induced under hypoxic stress via hypoxia-inducible factor 1α and functions in cellular adaptations and embryogenesis. Oncotarget 8:73846–73859
52. Tang F, Hajkova P, Barton SC, Lao K, Surani MA (2006) MicroRNA expression profiling of single whole embryonic stem cells. Nucleic Acids Res 34(2):e9
53. De Pietri Tonelli D, Calegari F, Fei JF, Nomura T, Osumi N, Heisenberg CP, Huttner WB (2006) Single-cell detection of microRNAs in developing vertebrate embryos after acute administration of a dual-fluorescence reporter/sensor plasmid. BioTechniques 41(6):727–732
54. Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, Surani MA (2010) Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. Cell Stem Cell 6(5):468–478
55. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, Taipale J (2011) Counting absolute numbers of molecules using unique molecular identifiers. Nat Methods 9(1):72–74
56. Koster MJ, Timmers HT (2015) Regulation of antisense transcription by Mot1p and NC2 via removal of TATA-binding protein (TBP) from the 3′-end of genes. Nucleic Acids Res 43(1):143–152
57. Faridani OR, Abdullayev I, Hagemann-Jensen M, Schell JP, Lanner F, Sandberg R (2016) Single-cell sequencing of the small-RNA transcriptome. Nat Biotechnol 34(12):1264–1266
58. Petkovic S, Müller S (2015) RNA circularization strategies in vivo and in vitro. Nucleic Acids Res 43(4):2454–2465
59. He J, Xie Q, Xu H, Li J, Li Y (2017) Circular RNAs and cancer. Cancer Lett 396:138–144
60. Sanger HL, Klotz G, Riesner D, Gross HJ, Kleinschmidt AK (1976) Viroids are single-stranded covalently closed circular RNA molecules existing as highly base-paired rod-like structures. Proc Natl Acad Sci U S A 73(11):3852–3856
61. Ebbesen KK, Kjems J, Hansen TB (2016) Circular RNAs: identification, biogenesis and function. Biochim Biophys Acta 1859(1):163–168
62. Li J, Yang J, Zhou P, Le Y, Zhou C, Wang S, Xu D, Lin HK, Gong Z (2015) Circular RNAs in cancer: novel insights into origins, properties, functions and implications. Am J Cancer Res 5(2):472–480
63. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M et al (2013) Circular RNAs are a large class of animal RNAs with regulatory potency. Nature 495(7441):333–338
64. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE (2013) Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA 19(2):141–157
65. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu S, Yang L, Chen LL (2013) Circular intronic long noncoding RNAs. Mol Cell 51(6):792–806
66. Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L et al (2015) Exon-intron circular RNAs regulate transcription in the nucleus. Nat Struct Mol Biol 22(3):256–264
67. Armakola M, Higgins MJ, Figley MD, Barmada SJ, Scarborough EA, Diaz Z, Fang X, Shorter J, Krogan NJ, Finkbeiner S et al (2012) Inhibition of RNA lariat debranching enzyme suppresses TDP-43 toxicity in ALS disease models. Nat Genet 44(12):1302–1309
68. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S (2014) circRNA biogenesis competes with pre-mRNA splicing. Mol Cell 56(1):55–66
69. Li F, Zhang L, Li W, Deng J, Zheng J, An M, Lu J, Zhou Y (2015) Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/β-catenin pathway. Oncotarget 6(8):6001–6013
70. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J (2013) Natural RNA circles function as efficient microRNA sponges. Nature 495(7441):384–388
71. Du WW, Fang L, Yang W, Wu N, Awan FM, Yang Z, Yang BB (2017) Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity. Cell Death Differ 24(2):357–370
72. Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB (2016) Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. Nucleic Acids Res 44(6):2846–2858
73. Meng S, Zhou H, Feng Z, Xu Z, Tang Y, Li P, Wu M (2017) CircRNA: functions and properties of a novel potential biomarker for cancer. Mol Cancer 16(1):94
74. Cui X, Niu W, Kong L, He M, Jiang K, Chen S, Zhong A, Li W, Lu J, Zhang L (2016) hsa_circRNA_103636: potential novel diagnostic and therapeutic biomarker in major depressive disorder. Biomark Med 10(9):943–952
75. Li P, Chen S, Chen H, Mo X, Li T, Shao Y, Xiao B, Guo J (2015) Using circular RNA as a novel type of biomarker in the screening of gastric cancer. Clin Chim Acta 444:132–136
76. Boyer TR, Erickson RP (1994) Detection of circular and linear transcripts of Sry in pre-implantation mouse embryos: differences in requirement for reverse transcriptase. Biochem Biophys Res Commun 198(2):492–496

77. Szabo L, Morey R, Palpant NJ, Wang PL, Afari N, Jiang C, Parast MM, Murry CE, Laurent LC, Salzman J (2015) Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. Genome Biol 16:126
78. Fan X, Zhang X, Wu X, Guo H, Hu Y, Tang F, Huang Y (2015) Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. Genome Biol 16:148
79. Kiss-László Z, Henry Y, Bachellerie JP, Caizergues-Ferrer M, Kiss T (1996) Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell 85(7):1077–1088
80. Kiss T (2002) Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. Cell 109(2):145–148
81. Jády BE, Kiss T (2001) A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. EMBO J 20(3):541–551
82. Ronchetti D, Todoerti K, Tuana G, Agnelli L, Mosca L, Lionetti M, Fabris S, Colapietro P, Miozzo M, Ferrarini M et al (2012) The expression pattern of small nucleolar and small Cajal body-specific RNAs characterizes distinct molecular subtypes of multiple myeloma. Blood Cancer J 2:e96
83. Atzorn V, Fragapane P, Kiss T (2004) U17/snR30 is a ubiquitous snoRNA with two conserved sequence motifs essential for 18S rRNA production. Mol Cell Biol 24(4):1769–1778
84. Marz M, Gruber AR, Höner Zu Siederdissen C, Amman F, Badelt S, Bartschat S, Bernhart SH, Beyer W, Kehr S, Lorenz R et al (2011) Animal snoRNAs and scaRNAs with exceptional structures. RNA Biol 8(6):938–946
85. Weinstein LB, Steitz JA (1999) Guided tours: from precursor snoRNA to functional snoRNP. Curr Opin Cell Biol 11(3):378–384
86. Decatur WA, Fournier MJ (2002) rRNA modifications and ribosome function. Trends Biochem Sci 27(7):344–351
87. Wan G, Liu Y, Han C, Zhang X, Lu X (2014) Noncoding RNAs in DNA repair and genome integrity. Antioxid Redox Signal 20(4):655–677
88. Taft RJ, Glazov EA, Lassmann T, Hayashizaki Y, Carninci P, Mattick JS (2009) Small RNAs derived from snoRNAs. RNA 15(7):1233–1240
89. Kishore S, Stamm S (2006) The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. Science 311(5758):230–232
90. Sahoo T, del Gaudio D, German JR, Shinawi M, Peters SU, Person RE, Garnica A, Cheung SW, Beaudet AL (2008) Prader-Willi phenotype caused by paternal deficiency for the HBII-85 C/D box small nucleolar RNA cluster. Nat Genet 40(6):719–721
91. Higa-Nakamine S, Suzuki T, Uechi T, Chakraborty A, Nakajima Y, Nakamura M, Hirano N, Suzuki T, Kenmochi N (2012) Loss of ribosomal RNA modification causes developmental defects in zebrafish. Nucleic Acids Res 40(1):391–398
92. Fong YW, Ho JJ, Inouye C, Tjian R (2014) The dyskerin ribonucleoprotein complex as an OCT4/SOX2 coactivator in embryonic stem cells. elife 3
93. Hirakata S, Siomi MC (2016) piRNA biogenesis in the germline: from transcription of piRNA genomic sources to piRNA maturation. Biochim Biophys Acta 1859(1):82–92
94. Siomi MC, Sato K, Pezic D, Aravin AA (2011) PIWI-interacting small RNAs: the vanguard of genome defence. Nat Rev Mol Cell Biol 12(4):246–258
95. Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, Morris P, Brownstein MJ, Kuramochi-Miyagawa S, Nakano T et al (2006) A novel class of small RNAs bind to MILI protein in mouse testes. Nature 442(7099):203–207
96. Houwing S, Kamminga LM, Berezikov E, Cronembold D, Girard A, van den Elst H, Filippov DV, Blaser H, Raz E, Moens CB et al (2007) A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish. Cell 129(1):69–82
97. Matsumoto N, Nishimasu H, Sakakibara K, Nishida KM, Hirano T, Ishitani R, Siomi H, Siomi MC, Nureki O (2016) Crystal structure of silkworm PIWI-clade Argonaute Siwi bound to piRNA. Cell 167(2):484–497.e9
98. Malone CD, Hannon GJ (2009) Small RNAs as guardians of the genome. Cell 136(4):656–668
99. Aravin AA, Hannon GJ, Brennecke J (2007) The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. Science 318(5851):761–764
100. Aravin AA, Sachidanandam R, Girard A, Fejes-Toth K, Hannon GJ (2007) Developmentally regulated piRNA clusters implicate MILI in transposon control. Science 316(5825):744–747
101. Lim SL, Qu ZP, Kortschak RD, Lawrence DM, Geoghegan J, Hempfling AL, Bergmann M, Goodnow CC, Ormandy CJ, Wong L et al (2015) HENMT1 and piRNA stability are required for adult male germ cell transposon repression and to define the Spermatogenic program in the mouse. PLoS Genet 11(10):e1005620
102. Watanabe T, Lin H (2014) Posttranscriptional regulation of gene expression by Piwi proteins and piRNAs. Mol Cell 56(1):18–27
103. Erwin AA, Galdos MA, Wickersheim ML, Harrison CC, Marr KD, Colicchio JM, Blumenstiel JP (2015) piRNAs are associated with diverse Transgenerational effects on gene and Transposon expression in a hybrid dysgenic syndrome of D. virilis. PLoS Genet 11(8):e1005332
104. Lau NC, Ohsumi T, Borowsky M, Kingston RE, Blower MD (2009) Systematic and single cell analysis of Xenopus Piwi-interacting RNAs and Xiwi. EMBO J 28(19):2945–2958



Qiang Fu is a resident physician who graduated from ZhengZhou University and majored in clinical medicine. During his master's period, he performed research on the function of pancreatic disease and micro-RNAs and their regulatory mechanisms, with Professor Zhang Hong Wei. He participated in two China national natural science fund projects. After graduation, he worked on surgeries of the liver, biliary tract and pancreas at the Henan people's hospital. He has a deep understanding of the diagnosis and treatment of common liver, biliary tract and pancreas surgical diseases. He has published more than 10 articles in domestic and foreign journals.



Chuan-Jiang Liu is a resident physician who graduated from Shandong University and majored in clinical medicine. After graduation, he worked on hepatobiliary and pancreatic surgery at Henan Province People's Hospital. He also has extensive experience in the diagnosis and treatment of liver, biliary tract, pancreas, spleen and other related diseases. He was involved in several national nature fund studies and successfully published 1 paper in SCI and 3 papers in Chinese periodicals as the first author.

Zhen-Sheng Zhai is a postgraduate studying at ZhengZhou University and majoring in clinical medicine. His research focuses on the function of micro-RNAs and its regulatory mechanisms under the direction of Professor Zhang HongWei. He has a certain understanding of miRNA function as well as the diagnosis, surgery, and treatment of common liver, biliary tract and pancreas diseases.

Xu Zhang is a postgraduate studying at ZhengZhou University and majoring in hepatobiliary surgery. His research focuses on the function of lncRNAs and its regulatory mechanisms under the direction of Professor Zhang HongWei. He also has a firm understanding of the diagnosis, surgery and treatment of common liver, biliary tract and pancreas diseases.

Tao Qin, MD, is a Chief physician and doctoral tutor. The deputy director of the first youth committee of the Henan provincial surgery society, the deputy head of the laparoscopic surgery committee of Henan provincial minimally invasive surgery committee, the secretary of the general surgical committee of Henan province, and committee member of the Henan epidemiology and Evidence medicine committee. In 2012, Qin Tao went to the University of Hamburg, Germany, ependov hospital and learned about liver transplantation. He can perform surgical operations on the liver, biliary tract and pancreas, specifically the diagnosis and treatment of diseases such as hepatic tumours, biliary tract tumours, pancreatic tumours, Budd-Chiari syndrome, PHT, SAP and other diseases. He is skilled in various complex surgeries of the liver, biliary tract and pancreas, including liver transplantation, giant liver tumour resection, left/ right hemihepatectomy, fine hepatic resection, radical resection of hilar cholangiocarcinoma, pancreatoduodenectomy, Budd -Chiari syndrome radical and cavity bypass surgery, portal hypertension disconnection after an operation, etc. He has completed a left hemihepatectomy, pancreatic tumour resections, and laparoscopic bile duct and biliary-enteric anastomosis. He is now responsible for two projects as the head of the China national natural science fund. He has participated in national and provincial research projects. Qin Tao has published more than 30 papers, five of in which are in SCI and have a cumulative impact factor of 7.591 points. The Chinese series includes 12 articles. He also has a patent for an artificial vascular bridge stapler from the national intellectual property office of China (patent number: ZL 2015 2 0237876.8).

Hong-Wei Zhang, MD, is a Professor at ZhengZhou University, Supervisor of Postgraduates, Chief physician, member of Henan provincial general surgery committee, member of the organ transplantation committee of Chinese medical association, member of the Henan provincial Anti-Cancer Association of Chinese medical association, leader of the Henan provincial Hepatobiliary Surgery group, chairman of the Henan provincial liver cancer union of prevention and cure of Chinese tumour, member of the Editorial board of several national journals and magazines. From 2002 to 2003, he majored in liver transplantation at the Université de Caen Basse-Normandien in France and received a master's degree. He has been engaged in basic and clinical research focused on liver, biliary tract and pancreas surgery, specifically focused on the diagnosis and treatment of diseases such as hepatic tumours, biliary tract tumours, pancreatic tumours, Budd-Chiari syndrome, PHT, SAP and other diseases. He is skilled in various complex surgeries on the liver, biliary tract and pancreas, such as liver transplantation, giant liver tumour resection, left/ right hemihepatectomy, fine hepatic resection, radical resection of hilar cholangiocarcinoma, pancreatoduodenectomy, Budd-Chiari syndrome radical and cavity bypass surgery, portal hypertension disconnection after operation, etc. In the last 10 years, the team for the comprehensive treatment of pancreatic diseases, which is organized by him, has committed themselves to studying pancreas disease, including basic and clinical aspects and the function of microRNAs in the pancreas. As the project leader, he has undertaken and completed the national natural science fund project and Provincial scientific research projects. The team members were awarded the provincial scientific progress prize for their remarkable job. He formed the Henan provincial liver cancer union for the prevention and cure of Chinese tumours, which has made major breakthroughs in the prevention, treatment and related research regarding liver cancer.

Check for updates

4 High Throughput Single Cell RNA Sequencing, Bioinformatics Analysis and Applications

Xiaoyun Huang, Shiping Liu, Liang Wu, Miaomiao Jiang, and Yong Hou

Abstract

Single cell sequencing (SCS) can be harnessed to acquire the genomes, transcriptomes and epigenomes from individual cells. Next generation sequencing (NGS) technology is the driving force for single cell sequencing. scRNA-seq requires a lengthy pipeline comprising of single cell sorting, RNA extraction, reverse transcription, amplification, library construction, sequencing and subsequent bioinformatic analysis. Computational algorithms are essential to fulfill many tasks of interest using scRNA-seq data. scRNA-seq has already enabled researchers to revisit long-standing questions in cancer biology, including cancer metastasis, heterogeneity and evolution. Circulating Tumor Cells (CTC) are not only an important mechanism for cancer metastasis, but also provide a possibility to diagnose and monitor cancer in a convenient way independent of surgical resection of the cancer.

Keywords

scRNA-seq · Single cell · Cancer · Bioinformatics

Authors Xiaoyun Huang and Shiping Liu have equally contributed with all other contributors.

X. Huang · S. Liu · L. Wu · M. Jiang · Y. Hou (✉)
BGI-Shenzhen, Shenzhen, China
e-mail: houyong@genomics.cn

4.1 Brief Introduction to Single Cell Sequencing

Single cell sequencing (SCS) can be harnessed to acquire the genomes, transcriptomes and epigenomes from individual cells. Next generation sequencing (NGS) technology is the driving force for single cell sequencing. Though various bias could be potentially introduced during the molecule amplification, It has been well recognized that SCS could help detect single nucleotide variations (SNVs) [1], copy number variations (CNVs) [2], structure variations (SVs) [3], gene expression and fusions [4–8], novel transcripts and alternative splicing [9], methylations [10] and chromatin patterns [11, 12] on single cell level with the help of mathematic algorithms and models. SCS also has great potential to reveal novel biological concepts, which have never been investigated. For example, researchers used single cell RNA-seq (scRNA-seq) to uncover new cell types in nervous system [13], immune system and hematopoietic system [14], as well as new insights into the clonal evolution of cancer [15]. Most recently, the accuracy and throughput of SCS have been increased dramatically to be capable of profiling more than thousands of single cells in parallel [5, 16].

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_4

4.2 Large-Scale scRNA-seq Library Preparation

scRNA-seq requires a lengthy pipeline comprising of single cell sorting, RNA extraction, reverse transcription, amplification, library construction, sequencing and subsequent bioinformatic analysis. As the key factor to increase the throughput of scRNA-seq study, high-throughput scRNA-seq library preparation technology developed very quickly these few years. FAC sorting-based scRNA-seq library preparation combined with automatic liquid instrument pipeline enable handling 96-well/384-well plate single cells per run [17]. Fluidigm C1 system based on microvalve microfluidic chip that was developed by Quake's lab enabled to prepare full-length transcripts of 96 single cells in parallel in 2012 [7], and a similar chip with higher throughput up to 800 single cells 3'end transcript preparation was released in 2015. Another type of microfluidic chip, microwell chip was also used in single-cell RNA amplification. Wu et al. developed an approach called MIRALCS [4], allowing massively parallel single-cell full-length transcripts amplification for 500–1000 single cells based on 5184-well chip. With the same chip, Wafergen Inc. released a single cell preparation system named ICell8, allowing the preparation of 3′ single cell transcripts with throughput up to 1800 cells per run [18]. Taking the advantage of barcode-bead technology, two different groups described microwell chip based method, with the capacity to obtain gene expression from thousands of cells at the single cell level, respectively [19, 20]. In addition, droplet microfluidic technique improves the throughput of single cell 3′ end RNA-seq to million level, and reduce the reaction volume to picoliter [21, 22]. A commercialized instrument with the same strategy was developed by 10X genomics, enabling the preparation of at most 48,000 single cells from eight different samples in parallel. With the development of high-throughput scRNA-seq library preparation technology, the cost has been reduced to less than one dollar per cell, which greatly promotes the studies at the single cell level.

4.3 Computational Analysis of scRNA-seq Data

Computational algorithms are essential to fulfill many tasks of interest using scRNA-seq data (Fig. 4.1). There is a general consensus that analyses of scRNA-seq data sets and conventional RNA-seq data have a lot in common. More specifically, the vast majority of the basic pipelines and tools established for sequencing data derived from bulk cell populations are applicable to that from single cells, following steps including read alignment, quality control and gene expression estimation. Whereas more dedicated softwares for use in fields such as identifying and characterizing cellular subpopulations, exploring evolution of cell groups and inferring the transcriptional kinetics are urgently needed owing to the zero-inflated nature and additional functions of scRNA-seq data sources.

- Quality Control: Single-cell datasets are expected to be extra messy, thus should undertake the quality control step before any downstream analysis. To begin with, FASTQC [23], Qualimap2 [24] and RSeQC [25] are commonly used for assessment of overall sequencing quality. After removal of adapters or noisy data with low quality, raw reads obtained from a well-designed experiment are firstly aligned to the reference genome using tools such as Tophat [26], HISAT [27] and STAR [28]. Subsequently, several indicators are calculated so as to discriminate cells with degraded RNA or substandard sequencing library, for instance, the number of expressed genes, the proportion of reads mapped to endogenous genes and the fraction of external spike-ins in mapped reads [9, 29, 30]. In addition, Treutlein considered normal expression of housekeeping genes a judgment factor of healthy cells [31].
- Expression estimation and normalization: Gene expression levels of qualified cells can be estimated as count from data without UMIs using HTSeq [32], WemIQ [33] or RSEM [34]. While relative expression including transcripts per million mapped reads (TPM)

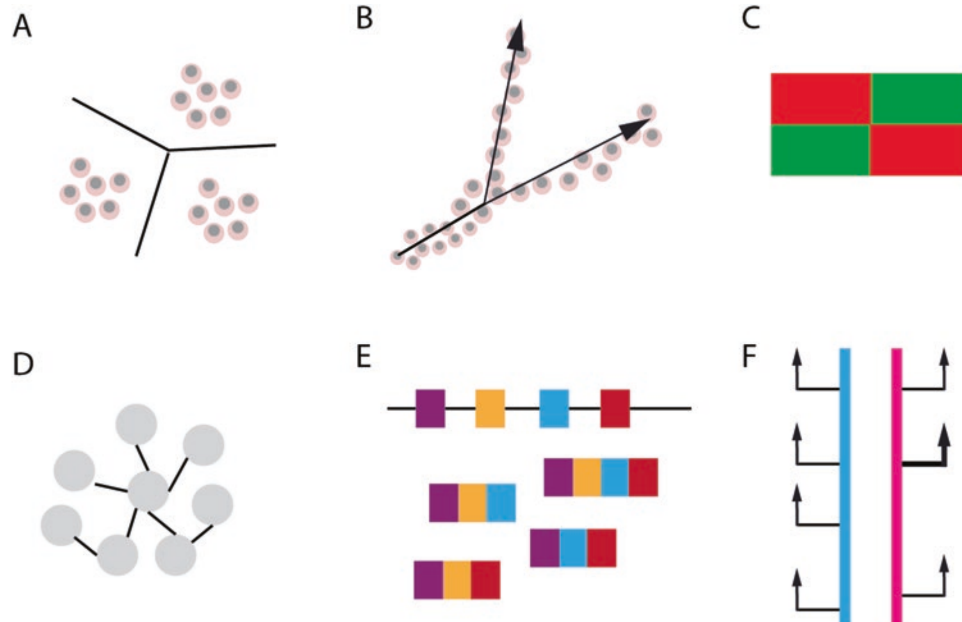


Fig. 4.1 Representative tasks enabled by scRNA-seq. (**a**) Subpopulation analysis can be performed with various unsupervised clustering algorithms; (**b**) Pseudotemporal ordering is essential to understand developmental trajectory or disease progression; (**c**) Differential gene expression analysis is important for the discovery of cell type specific biomarkers; (**d**) Network inference can be performed to learn regulatory intracellular and intercellular networks; (**e**) Analysis of alternative splicing offers a new perspective on biology and medicine; (**f**) Allele specific expression patterns can be addressed using scRNA-seq data

and reads/fragments per kilobase per million mapped reads (RPKM/FPKM) is widely adopted in downstream analysis. Besides, Islam et al. [29] and Hashimshony et al. [35] provide solutions to UMI-tagged reads. Normalization is essential due to the technical variability in comparison to expression levels between samples. Median normalization or a similar method are popular in many scRNA-seq studies without spike-ins or UMIs [30, 36–39]. In single cell experiments where spike-ins were applied, technical artifacts can be estimated by difference between their expected and observed expression. Nevertheless, instability arising from inconsistent detection of spike-ins brings about a more current notion of comparing absolute molecular counts of different cells with UMIs, which have greatly reduced the amplification noise by attaching random sequences to cDNA fragments ahead of PCR [21, 29, 40].

- Identification of subpopulations: Cellular subpopulation identification in heterogenous cells is one of the most exciting areas for exploration in the scRNA-seq experiments. Therefore, various algorithms for clustering have been developed to date. Pollen et al. [41] distinguished different types of cells along lung development using principal component analysis. The study by Li et al. [42] showed transcriptional heterogeneity in colorectal tumors with a novel strategy named reference component analysis (RCA). Following similar lines, self-organizing maps (SOMs) [43], circular a posteriori projection (CAP), ZIFA [44], t-SNE [45] and BackSPIN [46] clustering are approaches developed for differentiating between cells within a biological condition by dimensionality reduction of scRNA-seq data. In addition, RaceID [6] is a computationally efficient tool that relies on k-means clustering, whereas SNN-Cliq [47] clusters individual

cells by a graph-based algorithm based on shared nearest neighbor (SNN) similarity measurement. Guo et al. [48] further presented a pipeline for known cell type enrichment that is analogous to gene set enrichment analysis.

- Differential expression and transcript isoforms across conditions: Once subpopulations are distinguished, differential expression can be applied for cell type characterization. Researchers used to investigate differential expressed genes among cells of different types or stages with bulk RNA-Seq based strategy. However, an abundant zero values on expression matrix from single cells lead to potential fault sets of genes that may have expressed differently resulting from noise. As a consequence, plenty of mixture-model-based methods like MAST [49] and SCDE [50] have been created for accommodation of bimodality in expression levels. Similarly, D3E [51] identify DE genes by comparing two probability distributions on transcriptional bursting model. Korthauer et al. [52] have established a more accurate Bayesian modeling framework, scDD, for differential expression patterns detection under a wide range of circumstances recently. Unlike the traditional methods with a simple mean shift, the scDD model provides posterior probabilities differential distributions (DD) for each gene and classified gene as unimodal distributions (traditional DE), differential modes (DM), differential proportion (DP), or both DM and DE (abbreviated DB).
- Pseudotemporal ordering: Knowledge of the global expression profile in individual cells provides opportunities for the investigation of dynamic cellular processes such as normal tissue development, stem cell differentiation and tumor progression. A number of computational methods were built on the basis of the theory that differentiation paths can by constructed by reordering unsynchronized cells with gradual changes in gene expression levels at various stages. Similarities to cellular subpopulation identification approaches, most investigators perform pseudotemporal ordering by reducing the dimensionality of gene expression data. Take Monocle [53] as an example, which was the most effective tool to construct the differentiation paths in the infancy of single cell technology. Minimal spanning tree (MST) was built on data processed by independent component analysis (ICA) in Monocle, and the longest path through the MST was considered as a default setting for differentiation. Subsequently, Haghverdi L et al. [54] worked out a diffusion map based method that allows trajectory reconstruction in a single step. Rizvi et al. [55] presented a topology-based algorithm named single-cell topological data analysis (scTDA) for unbiased transcriptional regulation study through a nonlinear and unsupervised statistical framework. Furthermore, when it comes to oscillatory processes, Oscope [56] can be used for oscillatory trajectory reconstruction with co-regulation information among oscillators.
- Interrogation of spatial information: In spite of looking into the development of cell populations extending in time, scRNA-seq can be applied for spatial reconstruction via the integration of in situ RNA patterns with genome wide gene expression profiles. Satija R et al. [57] has established an accurate spatially resolved tools, Seurat, for mapping cellular localization, with which they inferred cellular localization of cells from dissociated zebrafish (Danio rerio) embryos and generated a transcriptome-wide map of spatial patterning. Meanwhile, another high-throughput approach by Kaia Achim [58] was published online by virtue of a reference gene expression database, which successfully allocates brain cells to precise locations from marine annelid Platynereis dumerilii by comparing specificity-weighted mRNA Profiles. Halpern K B et al. [59] reconstructed a genomic blueprint of mammalian liver by combining landmark genes expression and single-molecule fluorescence in situ hybridization.
- Network inference: Identification of co-regulated genes with data derived from single cell experiments is superior because it can

provide insight into regulatory networks that are hard to be noticed in bulk level. Understanding the transcriptional regulatory networks is of primary interest in a myriad of studies. For convenience, some statistical methods in bulk studies were reused when exploring scRNA-seq data. Weighted correlation network analysis (WGCNA) [60] can be used for gene clustering and summarizing with a comprehensive collection of functions for network construction, module detection, gene selection, calculations of topological properties, data simulation and visualization. Cell-centric statistics (CCs) [61] was invented to model transcriptome dynamics by analyzing aggregated cell-cell statistical distances within biomolecular pathways, for instance, differentially expressed pathways for a single cell of interest. While SCODE [62] inferred the co-regulatory network with ordinary differential equations(ODEs) by integrating the transformation of linear ODEs and linear regression.

- Differential Splicing: Experimental protocols with full-length transcript coverage to certain sequencing depth provides insight into alternative splicing isoform determination and quantification in scRNA-seq data analysis, which reflects heterogeneity among cells of a biological component from another perspective. A study of heterogeneity in immune cells in 2013 [9] was the first to reveal the dramatic diversity of splicing patterns in mouse bone-marrow-derived dendritic cells(BMDCs). Gokce O et al. [63] used fisher's exact test for differentially splicing junction definition among cell types and pointed out splice variant expressed in mouse striatum. SingleSplice [64] is the latest tailored method used to detect isoform usage differences in scRNA-seq data, which was applied to mouse embryonic stem cells and eventually shedded insight into the connection between alternative splicing and the cell cycle through a series of analysis.
- Allelic Expression Patterns: Another subtle point is that allele-specific expression can be accessed for in scRNA-seq to investigate the contribution of parental allele expression. Deng et al. [65] demonstrated an abundant random allele-specific gene expression using train-specific SNPs at single-cell resolution in mouse preimplantation embryos. Reinius B et al. [66] argued in an allele-sensitive scRNA–seq experiment that most patterns of random monoallelic expression of autosomal genes (aRME) are in a decentralized fashion rather than confined to clonally related cells according to previous hypothesis.

4.4 Application of High Throughput scRNA-seq

- Cancer Biology: scRNA-seq has already enabled researchers to revisit long-standing questions in cancer biology, including cancer metastasis, heterogeneity and evolution. Circulating Tumor Cells (CTC) are not only an important mechanism for cancer metastasis [67], but also provide a possibility to diagnose and monitor cancer in a convenient way independent of surgical resection of the cancer. One landmark study analysed CTC isolated from prostate cancer patients and revealed that the mechanism of resistance to androgen receptor inhibition in recurrent disease is partly due to noncanonical Wnt signaling [68].

A comprehensive picture of cancer heterogeneity is redefined by scRNA-seq. Several studies revealed the heterogeneity of cancer cells [69, 70]. A comprehensive profiling of melanoma using scRNA-seq is a classical example [70]. It was found that two distinct transcriptional signatures were present but they were not mutually exclusive. The melanoma characterized by activation of the transcription factor MITF also harbored a small subpopulation of cells with low MITF activity. The heterogeneity of cancer is not limited to the cell-to-cell variability among cancer cells. More importantly, cancer is itself a heterogeneous tissue comprised of malignant, immune, stromal and endothelial cells [71]. Recently, profiling of the immune cells within the

tumor microenvironment is attracting lots of attention [72–75]. Those studies covered various different cancers and single cell omic technologies. A recent study employed scRNA-seq to analyse T cells isolated from tumor tissues and adjacent normal tissues from hepatocellular carcinoma (HCC) patients, revealing the distinctive functional composition of T cells in HCC and the clonal enrichment of infiltrating Tregs and exhausted CD8 T cells [72].

The clonal evolution of cancer was proposed more than 40 years ago [76]. Longitudinal single cell analysis is now adding new evidence to this widely held concept [77]. Applying single nucleus sequencing to biopsy from primary breast cancer and its liver metastasis, it was suggested that tumor evolution might follow a punctuated expansion mode instead of a gradual progression path [78]. Single cell genome and exome sequencing enabled by MDA further increased the coverage of single cell genome sequencing and rendered the mutation and SNP calling at the single cell possible [79, 80]. The mutation and SNP information for individual cancer cells was valuable for population genetic analysis to understand the clonal evolution of tumor.

- Developmental Biology: Our understanding of developmental biology has also been dramatically enhanced by scRNA-seq. The identification of rare cell type was realized by the combination of organoid culture, scRNA-seq and development of novel algorithm [6]. This crystalized in the identification of Reg4 as a novel marker for enteroendocrine cells. New markers will then facilitate the investigation of rare cell types. Another recent study focused on the cells in the blood. New types of dendritic cells and monocytes were identified using scRNA-seq [14]. Our understanding of the cell types or subtypes constituting the brain was renewed by single-nucleus RNA sequencing [81] and scRNA-seq [82], while traditionally cell types were defined based on morphology, location and function.
- The Human Cell Atlas: With the development of high throughput single cell molecular profiling techniques, an international community or network is taking shape rapidly aiming to undertake the ambitious project to identify all cell types in the human body [83]. Single cell omic technologies are situated at the heart of the human cell altlas. Key efforts will be devoted to key organs, such as the liver, the heart, the kidney or the pancreas [84], as well key systems, such as the immune system and the central nervous system [85].

Our understanding of disease will also be greatly refined with the realization of the human cell atlas. In the future biopsy from patients will be routinely assayed with single cell techniques [70, 86] and compared to the normal reference in the human cell atlas. Specific abnormalities will be identified and used to inform both diagnosis and treatment.

The drug industry will benefit dramatically from the human cell atlas. Traditionally, drug discovery and development efforts have been hindered by the challenges that all healthy and diseased tissues are inherently heterogeneous [87]. The emergence and rapid application of single cell analysis tools will pave the way to eventually understand both health and disease at an unprecedented level so that medicine can finally ushers in a new era of personalized healthcare [88, 89].

References

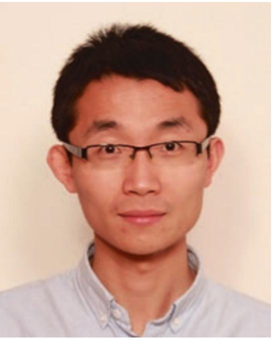
1. Zafar H, Wang Y, Nakhleh L, Navin N, Chen K (2016) Monovar: single-nucleotide variant detection in single cells. Nat Methods 13:505–507. https://doi.org/10.1038/nmeth.3835
2. Knouse KA, Wu J, Amon A (2016) Assessment of megabase-scale somatic copy number variation using single-cell sequencing. Genome Res 26:376–384. https://doi.org/10.1101/gr.198937.115
3. Voet T, Kumar P, Van Loo P, Cooke SL, Marshall J, Lin ML, Zamani Esteki M, Van der Aa N, Mateiu L, McBride DJ, Bignell GR, McLaren S, Teague J, Butler A, Raine K, Stebbings LA, Quail MA, D'Hooghe T, Moreau Y, Futreal PA, Stratton MR, Vermeesch JR, Campbell PJ (2013) Single-cell paired-end genome sequencing reveals structural variation per cell cycle. Nucleic Acids Res 41:6119–6138. https://doi.org/10.1093/nar/gkt345

4. Wu L, Zhang X, Zhao Z, Wang L, Li B, Li G, Dean M, Yu Q, Wang Y, Lin X, Rao W, Mei Z, Li Y, Jiang R, Yang H, Li F, Xie G, Xu L, Wu K, Zhang J, Chen J, Wang T, Kristiansen K, Zhang X, Li Y, Yang H, Wang J, Hou Y, Xu X (2015) Full-length single-cell RNA-seq applied to a viral human cancer: applications to HPV expression and splicing analysis in HeLa S3 cells. GigaSci 4:51. https://doi.org/10.1186/s13742-015-0091-4
5. Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA (2015) The technology and biology of single-cell RNA sequencing. Mol Cell 58:610–620. https://doi.org/10.1016/j.molcel.2015.04.005
6. Grun D, Lyubimova A, Kester L, Wiebrands K, Basak O, Sasaki N, Clevers H, van Oudenaarden A (2015) Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525:251–255. https://doi.org/10.1038/nature14966
7. Wu AR, Neff NF, Kalisky T, Dalerba P, Treutlein B, Rothenberg ME, Mburu FM, Mantalas GL, Sim S, Clarke MF, Quake SR (2014) Quantitative assessment of single-cell RNA-sequencing methods. Nat Methods 11:41–46. https://doi.org/10.1038/nmeth.2694
8. Saliba AE, Westermann AJ, Gorski SA, Vogel J (2014) Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res 42:8845–8860. https://doi.org/10.1093/nar/gku555
9. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498:236–240
10. Smallwood SA, Lee HJ, Angermueller C, Krueger F, Saadeh H, Peat J, Andrews SR, Stegle O, Reik W, Kelsey G (2014) Single-cell genome-wide bisulfite sequencing for assessing epigenetic heterogeneity. Nat Methods 11:817–820. https://doi.org/10.1038/nmeth.3035
11. Ramani V, Deng X, Qiu R, Gunderson KL, Steemers FJ, Disteche CM, Noble WS, Duan Z, Shendure J (2017) Massively multiplex single-cell Hi-C. Nat Methods 14:263–266. https://doi.org/10.1038/nmeth.4155
12. Buenrostro JD, Wu B, Litzenburger UM, Ruff D, Gonzales ML, Snyder MP, Chang HY, Greenleaf WJ (2015) Single-cell chromatin accessibility reveals principles of regulatory variation. Nature 523:486–490. https://doi.org/10.1038/nature14590
13. Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K, Martin-Villalba A (2015) Single-cell transcriptomics reveals a population of dormant neural stem cells that become activated upon brain injury. Cell Stem Cell 17:329–340. https://doi.org/10.1016/j.stem.2015.07.002
14. Villani AC, Satija R, Reynolds G, Sarkizova S, Shekhar K, Fletcher J, Griesbeck M, Butler A, Zheng S, Lazo S, Jardine L, Dixon D, Stephenson E, Nilsson E, Grundberg I, McDonald D, Filby A, Li W, De Jager PL, Rozenblatt-Rosen O, Lane AA, Haniffa M, Regev A, Hacohen N (2017) Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. Science 356. https://doi.org/10.1126/science.aah4573
15. Linnarsson S, Teichmann SA (2016) Single-cell genomics: coming of age. Genome Biol 17:97. https://doi.org/10.1186/s13059-016-0960-x
16. Vitak SA, Torkenczy KA, Rosenkrantz JL, Fields AJ, Christiansen L, Wong MH, Carbone L, Steemers FJ, Adey A (2017) Sequencing thousands of single-cell genomes with combinatorial indexing. Nat Methods 14:302–308. https://doi.org/10.1038/nmeth.4154
17. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, Mildner A, Cohen N, Jung S, Tanay A, Amit I (2014) Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science 343:776–779. https://doi.org/10.1126/science.1247651
18. http://www.wafergen.com/products/icell8-single-cell-system
19. Yuan J, Sims PA (2016) An automated microwell platform for large-scale single cell RNA-Seq. Sci Rep 6:33883. https://doi.org/10.1038/srep33883
20. Fan HC, Fu GK, Fodor SP (2015) Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. Science 347:1258367. https://doi.org/10.1126/science.1258367
21. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, Peshkin L, Weitz DA, Kirschner MW (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161:1187–1201. https://doi.org/10.1016/j.cell.2015.04.044
22. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA (2015) Highly parallel genome-wide expression profiling of individual cells using Nanoliter droplets. Cell 161:1202–1214. https://doi.org/10.1016/j.cell.2015.05.002
23. Andrews S (2010) https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
24. Okonechnikov K, Conesa A, García-Alcalde F (2015) Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. Bioinformatics 32:292–294
25. Wang L, Wang S, Li W (2012) RSeQC: quality control of RNA-seq experiments. Bioinformatics 28:2184–2185
26. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25:1105–1111
27. Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12:357–360
28. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29:15–21

29. Islam S, Zeisel A, Joost S, La Manno G, Zajac P, Kasper M, Lönnerberg P, Linnarsson S (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods 11:163–166
30. Brennecke P, Anders S, Kim JK, Kołodziejczyk AA, Zhang X, Proserpio V, Baying B, Benes V, Teichmann SA, Marioni JC (2013) Accounting for technical noise in single-cell RNA-seq experiments. Nat Methods 10:1093–1095
31. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH, Desai TJ, Krasnow MA, Quake SR (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509:371–375
32. Anders S, Pyl PT, Huber W (2015) HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31:166–169
33. Zhang J, Kuo C-CJ, Chen L (2014) WemIQ: an accurate and robust isoform quantification method for RNA-seq data. Bioinformatics 31:878–885btu757
34. Li B, Dewey CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC bioinf 12:323
35. Hashimshony T, Wagner F, Sher N, Yanai I (2012) CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep 2:666–673
36. Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, Teichmann SA, Marioni JC, Stegle O (2015) Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. Nat Biotechnol 33:155–160
37. Kim JK, Marioni JC (2013) Inferring the kinetics of stochastic gene expression from single-cell RNA-sequencing data. Genome Biol 14:R7
38. Dueck H, Khaladkar M, Kim TK, Spaethling JM, Francis C, Suresh S, Fisher SA, Seale P, Beck SG, Bartfai T (2015) Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. Genome Biol 16:122
39. Xue Z, Huang K, Cai C, Cai L, Jiang C-Y, Feng Y, Liu Z, Zeng Q, Cheng L, Sun YE (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA [thinsp] sequencing. Nature 500:593–597
40. Grün D, Kester L, Van Oudenaarden A (2014) Validation of noise models for single-cell transcriptomics. Nat Methods 11:637–640
41. Pollen AA, Nowakowski TJ, Shuga J, Wang X, Leyrat AA, Lui JH, Li N, Szpankowski L, Fowler B, Chen P (2014) Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat Biotechnol 32:1053–1058
42. Li H, Courtois ET, Sengupta D, Tan Y, Chen KH, Goh JJL, Kong SL, Chua C, Hon LK, Tan WS (2017) Reference component analysis of single-cell transcriptomes elucidates cellular heterogeneity in human colorectal tumors. Nat Genet 49(5):708–718
43. Kim DH, Marinov GK, Pepke S, Singer ZS, He P, Williams B, Schroth GP, Elowitz MB, Wold BJ (2015) Single-cell transcriptome analysis reveals dynamic changes in lncRNA expression during reprogramming. Cell Stem Cell 16:88–101
44. Pierson E, Yau C (2015) Dimensionality reduction for zero-inflated single cell gene expression analysis. Genome Biology 16:1–10
45. van der Maaten L, Hinton G (2008) Visualizing data using t-SNE. J Mach Learn Res 9:2579–2605
46. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, La Manno G, Juréus A, Marques S, Munguba H, He L, Betsholtz C (2015) Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347:1138–1142
47. Xu C, Su Z (2015) Identification of cell types from single-cell transcriptomes using a novel clustering method. Bioinformatics 31:1974–1980btv088
48. Guo M, Wang H, Potter SS, Whitsett JA, Xu Y (2015) SINCERA: a pipeline for single-cell RNA-Seq profiling analysis. PLoS Comput Biol 11:e1004575
49. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Prlic M (2015) MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. Genome Biol 16:278
50. Kharchenko PV, Silberstein L, Scadden DT (2014) Bayesian approach to single-cell differential expression analysis. Nat Methods 11:740–742
51. Delmans M, Hemberg M (2016) Discrete distributional differential expression (D 3 E)-a tool for gene expression analysis of single-cell RNA-seq data. BMC Bioinf 17:110
52. Korthauer KD, Chu L-F, Newton MA, Li Y, Thomson J, Stewart R, Kendziorski C (2016) A statistical approach for identifying differential distributions in single-cell RNA-seq experiments. Genome Biol 17:222
53. Trapnell C, Cacchiarelli D, Grimsby J, Pokharel P, Li S, Morse M, Lennon NJ, Livak KJ, Mikkelsen TS, Rinn JL (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32:381–386
54. Haghverdi L, Buettner F, Theis FJ (2015) Diffusion maps for high-dimensional single-cell analysis of differentiation data. Bioinformatics 31:2989–2998
55. Rizvi AH, Camara PG, Kandror EK, Roberts TJ, Schieren I, Maniatis T, Rabadan R (2017) Single-cell topological RNA-seq analysis reveals insights into cellular differentiation and development. Nat Biotechnol 35:551–560
56. Leng N, Chu L-F, Barry C, Li Y, Choi J, Li X, Jiang P, Stewart RM, Thomson JA, Kendziorski C (2015) Oscope identifies oscillatory genes in unsynchronized single-cell RNA-seq experiments. Nat Methods 12:947–950
57. Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33:495–502

58. Achim K, Pettit J-B, Saraiva LR, Gavriouchkina D, Larsson T, Arendt D, Marioni JC (2015) High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. Nat Biotechnol 33:503–509
59. Halpern KB, Shenhav R, Matcovitch-Natan O, Tóth B, Lemze D, Golan M, Massasa EE, Baydatch S, Landen S, Moor AE (2017) Single-cell spatial reconstruction reveals global division of labour in the mammalian liver. Nature 542:352–356
60. Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 4:1128
61. Schissler AG, Li Q, Chen JL, Kenost C, Achour I, Billheimer DD, Li H, Piegorsch WW, Lussier YA (2016) Analysis of aggregated cell–cell statistical distances within pathways unveils therapeutic-resistance mechanisms in circulating tumor cells. Bioinformatics 32:i80–i89
62. Matsumoto H, Kiryu H, Furusawa C, Ko MS, Ko SB, Gouda N, Hayashi T, Nikaido ISCODE (2017) An efficient regulatory network inference algorithm from single-cell RNA-Seq during differentiation. *Bioinformatics* 33:2314–2321btx194
63. Gokce O, Stanley GM, Treutlein B, Neff NF, Camp JG, Malenka RC, Rothwell PE, Fuccillo MV, Südhof TC, Quake SR (2016) Cellular taxonomy of the mouse striatum as revealed by single-cell RNA-seq. Cell Rep 16:1126–1137
64. Welch JD, Hu Y, Prins JF (2016) Robust detection of alternative splicing in a population of single cells. Nucleic Acids Res 44:e73
65. Deng Q, Ramsköld D, Reinius B, Sandberg R (2014) Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science 343:193–196
66. Reinius B, Mold JE, Ramsköld D, Deng Q, Johnsson P, Michaëlsson J, Frisén J, Sandberg R (2016) Analysis of allelic expression patterns in clonal somatic cells by single-cell RNA-seq. Nat Genet 48:1430–1435
67. Hong Y, Li Z, Zhang Q (2016) A circulating tumor cell cluster-based model for tumor metastasis (Hypothesis). Oncol Lett 12:4891–4895. https://doi.org/10.3892/ol.2016.5358
68. Miyamoto DT, Zheng Y, Wittner BS, Lee RJ, Zhu H, Broderick KT, Desai R, Fox DB, Brannigan BW, Trautwein J, Arora KS, Desai N, Dahl DM, Sequist LV, Smith MR, Kapur R, Wu CL, Shioda T, Ramaswamy S, Ting DT, Toner M, Maheswaran S, Haber DA (2015) RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 349:1351–1356. https://doi.org/10.1126/science.aab0917
69. Venteicher AS, Tirosh I, Hebert C, Yizhak K, Neftel C, Filbin MG, Hovestadt V, Escalante LE, Shaw ML, Rodman C, Gillespie SM, Dionne D, Luo CC, Ravichandran H, Mylvaganam R, Mount C, Onozato ML, Nahed BV, Wakimoto H, Curry WT, Iafrate AJ, Rivera MN, Frosch MP, Golub TR, Brastianos PK, Getz G, Patel AP, Monje M, Cahill DP, Rozenblatt-Rosen O, Louis DN, Bernstein BE, Regev A, Suva ML (2017) Decoupling genetics, lineages, and microenvironment in IDH-mutant gliomas by single-cell RNA-seq. Science 355. https://doi.org/10.1126/science.aai8478
70. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, Rotem A, Rodman C, Lian C, Murphy G, Fallahi-Sichani M, Dutton-Regester K, Lin JR, Cohen O, Shah P, Lu D, Genshaft AS, Hughes TK, Ziegler CG, Kazer SW, Gaillard A, Kolb KE, Villani AC, Johannessen CM, Andreev AY, Van Allen EM, Bertagnolli M, Sorger PK, Sullivan RJ, Flaherty KT, Frederick DT, Jane-Valbuena J, Yoon CH, Rozenblatt-Rosen O, Shalek AK, Regev A, Garraway LA (2016) Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science 352:189–196. https://doi.org/10.1126/science.aad0501
71. Mbeunkui F, Johann DJ Jr (2009) Cancer and the tumor microenvironment: a review of an essential relationship. Cancer Chemother Pharmacol 63:571–582. https://doi.org/10.1007/s00280-008-0881-9
72. Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, Kang B, Hu R, Huang JY, Zhang Q, Liu Z, Dong M, Hu X, Ouyang W, Peng J, Zhang Z (2017) Landscape of infiltrating T cells in liver cancer revealed by single-cell sequencing. Cell 169:1342–1356. e1316. https://doi.org/10.1016/j.cell.2017.05.035
73. Lavin Y, Kobayashi S, Leader A, Amir ED, Elefant N, Bigenwald C, Remark R, Sweeney R, Becker CD, Levine JH, Meinhof K, Chow A, Kim-Shulze S, Wolf A, Medaglia C, Li H, Rytlewski JA, Emerson RO, Solovyov A, Greenbaum BD, Sanders C, Vignali M, Beasley MB, Flores R, Gnjatic S, Pe'er D, Rahman A, Amit I, Merad M (2017) Innate immune landscape in early lung adenocarcinoma by paired single-cell analyses. Cell 169:750–765., e717. https://doi.org/10.1016/j.cell.2017.04.014
74. Chung W, Eum HH, Lee HO, Lee KM, Lee HB, Kim KT, Ryu HS, Kim S, Lee JE, Park YH, Kan Z, Han W, Park WY (2017) Single-cell RNA-seq enables comprehensive tumour and immune cell profiling in primary breast cancer. Nat Commun 8:15081. https://doi.org/10.1038/ncomms15081
75. Chevrier S, Levine JH, Zanotelli VRT, Silina K, Schulz D, Bacac M, Ries CH, Ailles L, Jewett MAS, Moch H, van den Broek M, Beisel C, Stadler MB, Gedye C, Reis B, Pe'er D, Bodenmiller B (2017) An immune atlas of clear cell renal cell carcinoma. Cell 169:736–749. e718. https://doi.org/10.1016/j.cell.2017.04.016
76. Nowell PC (1976) The clonal evolution of tumor cell populations. Science 194:23–28
77. Zhao Z, Goldin L, Liu S, Wu L, Zhou W, Lou H, Yu Q, Tsang SX, Jiang M, Li F, McMaster M, Li Y, Lin X, Wang Z, Xu L, Marti G, Li G, Wu K, Yeager M, Yang H, Xu X, Chanock SJ, Li B, Hou Y, Caporaso N, Dean M (2016) Evolution of multiple cell clones over a 29-year period of a CLL patient. Nat Commun 7:13765. https://doi.org/10.1038/ncomms13765

78. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, Muthuswamy L, Krasnitz A, McCombie WR, Hicks J, Wigler M (2011) Tumour evolution inferred by single-cell sequencing. Nature 472:90–94. https://doi.org/10.1038/nature09807
79. Xu X, Hou Y, Yin X, Bao L, Tang A, Song L, Li F, Tsang S, Wu K, Wu H, He W, Zeng L, Xing M, Wu R, Jiang H, Liu X, Cao D, Guo G, Hu X, Gui Y, Li Z, Xie W, Sun X, Shi M, Cai Z, Wang B, Zhong M, Li J, Lu Z, Gu N, Zhang X, Goodman L, Bolund L, Wang J, Yang H, Kristiansen K, Dean M, Li Y, Wang J (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148:886–895. https://doi.org/10.1016/j.cell.2012.02.025
80. Hou Y, Song L, Zhu P, Zhang B, Tao Y, Xu X, Li F, Wu K, Liang J, Shao D, Wu H, Ye X, Ye C, Wu R, Jian M, Chen Y, Xie W, Zhang R, Chen L, Liu X, Yao X, Zheng H, Yu C, Li Q, Gong Z, Mao M, Yang X, Yang L, Li J, Wang W, Lu Z, Gu N, Laurie G, Bolund L, Kristiansen K, Wang J, Yang H, Li Y, Zhang X, Wang J (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148:873–885. https://doi.org/10.1016/j.cell.2012.02.028
81. Lake BB, Ai R, Kaeser GE, Salathia NS, Yung YC, Liu R, Wildberg A, Gao D, Fung HL, Chen S, Vijayaraghavan R, Wong J, Chen A, Sheng X, Kaper F, Shen R, Ronaghi M, Fan JB, Wang W, Chun J, Zhang K (2016) Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. Science 352:1586–1590. https://doi.org/10.1126/science.aaf1204
82. Zeisel A, Munoz-Manchado AB, Codeluppi S, Lonnerberg P, La Manno G, Jureus A, Marques S, Munguba H, He L, Betsholtz C, Rolny C, Castelo-Branco G, Hjerling-Leffler J, Linnarsson S (2015) Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 347:1138–1142. https://doi.org/10.1126/science.aaa1934
83. Regev A, Teichmann S, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, Carninci P, Clatworthy M, Clevers H, Deplancke B, Dunham I, Eberwine J, Eils R, Enard W, Farmer A, Fugger L, Gottgens B, Hacohen N, Haniffa M, Hemberg M, Kim SK, Klenerman P, Kriegstein A, Lein E, Linnarsson S, Lundeberg J, Majumder P, Marioni J, Merad M, Mhlanga M, Nawijn M, Netea M, Nolan G, Pe'er D, Philipakis A, Ponting CP, Quake SR, Reik W, Rozenblatt-Rosen O, Sanes JR, Satija R, Shumacher T, Shalek AK, Shapiro E, Sharma P, Shin J, Stegle O, Stratton M, Stubbington MJT, van Oudenaarden A, Wagner A, Watt FM, Weissman JS, Wold B, Xavier RJ, Yosef N (2017) The human cell atlas. bioRxiv. https://doi.org/10.1101/121202
84. Muraro MJ, Dharmadhikari G, Grun D, Groen N, Dielen T, Jansen E, van Gurp L, Engelse MA, Carlotti F, de Koning EJ, van Oudenaarden A (2016) A single-cell transcriptome atlas of the human pancreas. Cell Syst 3:385–394. e383. https://doi.org/10.1016/j.cels.2016.09.002
85. Tasic B, Menon V, Nguyen TN, Kim TK, Jarsky T, Yao Z, Levi B, Gray LT, Sorensen SA, Dolbeare T, Bertagnolli D, Goldy J, Shapovalova N, Parry S, Lee C, Smith K, Bernard A, Madisen L, Sunkin SM, Hawrylycz M, Koch C, Zeng H (2016) Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. Nat Neurosci 19:335–346. https://doi.org/10.1038/nn.4216
86. Patel AP, Tirosh I, Trombetta JJ, Shalek AK, Gillespie SM, Wakimoto H, Cahill DP, Nahed BV, Curry WT, Martuza RL, Louis DN, Rozenblatt-Rosen O, Suva ML, Regev A, Bernstein BE (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 344:1396–1401. https://doi.org/10.1126/science.1254257
87. Marusyk A, Almendro V, Polyak K (2012) Intra-tumour heterogeneity: a looking glass for cancer? Nat Rev Cancer 12:323–334. https://doi.org/10.1038/nrc3261
88. Heath JR, Ribas A, Mischel PS (2016) Single-cell analysis tools for drug discovery and development. Nat Rev Drug Discov 15:204–216. https://doi.org/10.1038/nrd.2015.16
89. Hood L (2013) Systems biology and p4 medicine: past, present, and future. Rambam Maimonides Med J 4:e0012. https://doi.org/10.5041/RMMJ.10112



Xiaoyun Huang did his PhD in Systems Biology at German Cancer Research Center (DKFZ) in Heidelberg. His PhD thesis focused on signal transduction and cell fate decision of primary mouse hepatocytes at the single cell level. Currently, Xiaoyun is interested in the development of computational algorithms to harness the power of single cell omic technologies for unraveling the mechanisms underlying cancer heterogeneity and evolution.

Shiping Liu received his Ph.D. in genetics from South China University of Technology. He has published 13 SCI papers in the past five years, including six papers on the top journals (Nature, Science, and Cell), three of which he was the first or co-first author. Shiping is currently in charge of the scientific research of single cell genomics at BGI Research.

Yong Hou received his PhD. on bioinformatics at Copenhagen University, and work as Research Scientist and Associate Director of BGI-Research. He has strong background of next generation sequencing data analysis and interpretation, especially on single cell analysis and cancer research. Published more than 20 peer reviewed scientific papers on journals including Cell, Nature Biotechnology, and be the co-inventor of more than 30 of patents in related area.

5 Circulating Tumor Cells: The Importance of Single Cell Analysis

D. Heymann and M. Téllez-Gabriel

Abstract

Cancer cells that have shed from the primary tumor are able to invade into surrounding tissues, to intravasate into the bloodstream to become circulating tumor cells (CTCs), at least one part of that cells will be able to generate distant metastases. The discovery of CTCs has improved the study of cancer disease as it represents a non invasive biopsy that can be used as prognostic and prediction biomarkers. Tumour heterogeneity is a concept related to differences in tumor cells within the same tumor or between tumours in terms of genetic and phenotypic profiles, such as morphology, metabolic activity, proliferation rate, migration and metastatic abilities. Characterization of heterogeneity among CTCs at the single cell level may be useful to better understand the causes and progression of disease and for an accurate selection of molecular prognostic/prediction markers. In this chapter we aimed to describe methods for CTC enrichment and isolation as well as current methodologies for single cell analysis at different levels, including RNA, DNA, protein and epigenetic events. Finally we wanted to stress clinical and biological importance of single CTC analysis by reviewing some studies carried out in different cancer subtypes.

Keywords

Biomarkers · Circulating tumor cells · Liquid biopsy · Single cells · Tumor heterogeneity

5.1 Introduction

Circulating tumor cells constitute a heterogeneous population of cells derived from tumor which could contribute to spread from the primary to the metastasis sites or/and from metastasis to metastasis foci. This process, depicted in Fig. 5.1, is iniciated when CTCs invade the surrounding tissue through the lymphatic vessels or the bloodstream, survive in the circulation, are able to extravasate into a tissue and finally grow at the new site [76] (Fig. 5.1).

Several studies indicated that CTCs can be isolated at relatively early stages of the disease

D. Heymann
Department of Oncology and Metabolism, University of Sheffield, Medical School, INSERM, European Associated Laboratory "Sarcoma Research Unit", Sheffield, UK

Faculty of Medicine, University of Nantes, Nantes, France

Institut de Cancérologie de l'Ouest, site René Gauducheau, Saint-Herblain, France

M. Téllez-Gabriel (✉)
Laboratotio Hematologia Oncologica y de Transplantes, Institut Investigacions Biomèdiques (IBB) Sant Pau, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain
e-mail: mtellez@santpau.cat

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_5

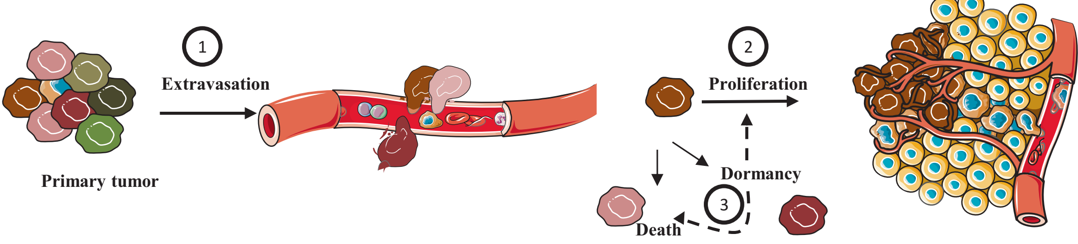


Fig. 5.1 Metastatic process and circulating tumor cells. Cells escaping from the primary tumour into the blood circulation (1) are carried by the flow, either blood stream or lymphatic system (2), to secondary sites where they grow if they find a favourable environment (3). Following the arrival of CTCs into a secondary organ only a subset will survive and generate metastases (clinically detectable) and remainder cells might either go into a state of dormancy/quiescence, or will die (clinically undetectable)

[30, 68], even before the detection of the primary tumor mass or metastatic foci by conventional methods. The absence of apparent relationship between CTCs and micrometastases is specifically explained by the current imaging technologies which are not sensitive enough to detect early tumor cell dissemination events or micrometastases, that represent key of cancer progression [39]. In this specific context, CTCs are very attractive as clinical biomarkers of recurrent and/or metastatic diseases.

This chapter aims to stress the existence of CTCs heterogeneity and remarks the importance of the analysis at single cell level to unravel the disease progression and to help in the therapeutic management and clinical follow up. The present chapter will summarize the main current methods available for the enrichment and isolation of CTCs as well as the state-of-the-art for studying cell heterogeneity at the single cell scale. The biological and therapeutic interest of single CTCs will be illustrated by various studies carried out in different cancer subtypes.

5.2 Methods for the Enrichment and Isolation of CTCs

Despite the study of circulating tumor cells is very promising for a better understanding of oncogenic events, there are certain issues intrinsic to the nature of CTCs that should be considered. CTCs are present at very low concentrations in the blood (i.e. CTCs can be find at rates as low as one in 10^9 erythrocytes and one in 10^7 leukocytes) [28]. Consequently, based on this very low frequency, the first challenge for isolating CTCs is their pre-enrichment and the depletion of blood cells [16]. While mature erythrocytes have distinct physicochemical and biological properties that allow them to be easily removed from blood, leukocytes present many properties which are similar to CTCs, resulting in a high level of leukocyte contamination in many separation methods. Therefore, the most important requirement of any separation process is an effective discrimination of CTCs from leukocytes. However, CTCs are heterogeneous and so they do not express universal specific tumor markers leading difficult the use of biochemical separation methods [16]. The clinical use of CTCs detection was preceded by the discovery that the expression of the EpCAMs were present on epithelial derived carcinomas and related cancers at different levels, but were absent in peripheral blood cells [62]. This finding promoted the research and development of different methodologies based on the EpCAM marker for the enrichment and isolation of CTCs [47] and led to the CellSearch® system (Janssen Diagnostics), the only method currently approved by the US Food and Drug Administration for clinical use which is based in EpCAM/cytokeratin separation. This system was considered as the gold standard method for CTC detection [2, 35]. However, the suitability of this method has been challenged by current studies that

demonstrate the presence of EpCAM-positive circulating epithelial cells in patients with benign colon diseases leading to a potential source of false-positive results [59]. In addition, during the epithelial–mesenchymal transition (EMT) in carcinoma cells has been observed a downregulation in epithelial markers, such as EpCAM and Cytokeratins (CK) [52], concomitant with an increased expression of mesenchymal markers (e.g. vimentin). The decrease in the expression of epithelial markers may therefore result in false-negative findings. Furthermore, remains unclear if all CK stained cells are CTCs as it has been observed that activated leukocytes can occasionally express epithelial markers [73, 74]. In this context, the EpCAM marker is not suitable to detect CTCs from carcinomas of mesenchymal origin or those cells that have undergone EMT. Consequently, these findings question the idea of considering EpCAM as a universal marker for CTC detection, highlighting the need to develop non–EpCAM-based technologies for isolating and detecting CTCs. This complexity is also enhanced by the heterogeneous functional properties of CTCs. Indeed, all CTCs do no exhibit an ability to generate metastases [87] and for clinical purposes it is mandatory to distinguish CTCs with metastatic and non-metastatic properties.

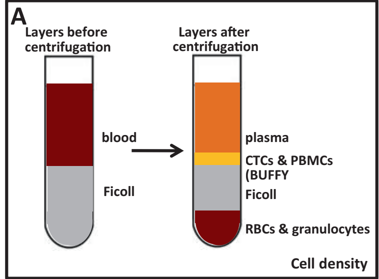
Current CTC separation methods can be classified in two categories: those based on biochemical properties (cell surface protein expression, invasive or migration capacity, viability) and those based on physical properties (size, shape, density, electric charges, deformability and magnetic susceptibility) (Fig. 5.1). Sects. 5.2.1. and 5.2.2. will then review some of the most used technologies for CTC isolation based on either physical or biochemical properties.

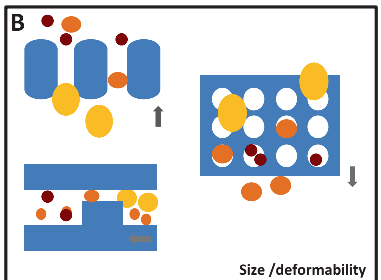
5.2.1 Methods of CTC Isolation Based on Physical Properties

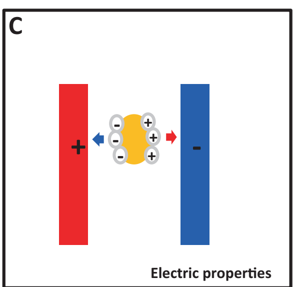
Density gradient centrifugation is widely-used technique for separating blood components. It is based in the migration of cellular components through a medium of higher or graded density when they are exposed to a elevated centrifugal force. Erythrocytes or polymorphonuclear leukocytes migrate to the bottom, whereas mononuclear leukocytes and CTCs remain at the top as a buffy coat (Fig. 5.2a) [41]. Percoll, Ficoll-Hypaque (GE Healthcare Life Sciences), and OncoQuick® (Greiner Bio-One) are the most commonly used density gradient media in preclinical and clinical studies. Pitfalls associated to these techniques include: (i) the loss of large CTCs and CTC clusters that fall to the bottom [26]; (ii) the cytotoxicity of the density medium which can affect the results of dowstream analysis [66]; (iii) the very low purity due to the contamination with leukocytes that are difficult to eliminate, making necessary to combine centrifugation with another enrichment method.

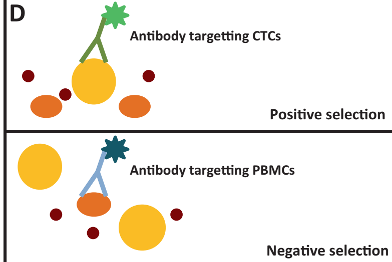
Microfiltration enrichment methods capture target cells according to their size and/or their deformability when they pass through an array of microscale constrictions. There are several types of microconstrictions: weir, pillar and pore (Fig. 5.2b). Several commercial devices utilizing the pore structure concept have been introduced, including the ISET®(RareCells) [50], ScreenCell® (ScreenCell) [16], CellSieve™(CreaTV Microtech), Flexible MicroSpring Array (FMSA) [32] or the FaCTChecker (Circulogix) [89]. Filtration allows processing large volumes of sample for rapid CTC enrichment with minimal sample manipulation. Recovery rates are around 90%, but with low final purity about 10%. The main limitation of the filtration approach is due to the high heterogeneity of CTC size, the possibility of membrane clogging, and troubles in the recovery of cells retained in the filter for further analysis. To minimize this last issue, Lim et al., designed a system based on a silicon micro-sieve that allows CTC antibody staining, separation and enumeration. They also demonstrated the feasibility of this method for fluorescence *in situ* hybridization (FISH) at the single cell level [48]. By the other hand, Zheng et al., developed a parylene membrane microfilter for capturing CTCs and allowing a genomic analysis based in polymerase chain reaction (PCR) [88].

Microfluidics comprises a variety of separation methods, which makes it possible to









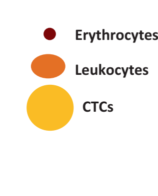


Fig. 5.2 Methods for the enrichment or isolation of CTCs from whole blood. Physical properties based: (**a**) cell density; (**b**) size/deformability; and (**c**) electrical properties. Immunoafinity – based techniques (**d**), either positive or negative selection of CTCs

manipulate very small volumes of biological fluids. There are microfluidic devices based on different size and deformability properties of CTCs and other blood cells, like the Parsortix system (Angle) [83], the ClearCell FX® (Clearbridge BioMedics) [45] or Vortex (Vortex Biosciences) [75]. Recent studies describing the important clinical value of CTC clusters [36] have impulsed the development of a microfluidic device based on the structural properties of CTC clusters, called Cluster-Chip (National Institute of Biomedical Imaging and Bioengineering (NIBIB)) [71]. Unfortunately, like mentioned for the microfiltration approaches, heterogeneity of CTC size leads to the loss of CTCs that have scaped through the white/red cell microchannels, and the contamination of CTC fraction with white blood cells. Dielectric properties, determined by differences in cell size, nuclear morphology, and membrane morphology is a property also used for CTC isolation in microfluidic devices. The ApoStream™ system (ApoCell) [29] or the DEPArray™ technology (Silicon Biosystems) benefit from dielectrophoretic (DEP) force to separate cells in basis on dielectric properties. The DEPArray™ device allows single-cell and high-fidelity recovery, cell viability and, it is possible the isolation of individual cells from paraffin-embedded samples in the last version of this device [63]. Despite the multiple advantages presented by DEPbased methods (Fig. 5.2.c) they present also some limitations, such as the low volumes of samples that can be processed in a non-continuous manner. Cells can suffer changes in their dielectric properties due to ion leakage and thus the isolation should be achieved within a short period of time ([82]). In addition, the medium used must have low conductivity, which sometimes is not achievable for studied samples.

5.2.2 Immunoafinity Based Methods

CTC isolation based on antibody use takes advantage of highly specific affinity reactions between antibodies and the target antigens present on the cells of interest. The antibodies used for cell capture are typically attached to either the device

surface or a magnetic substance for subsequent immunomagnetic capture [58]. This method is able to achieve high recovery and purity rates, with single-step detection and isolation of CTCs [77]. The performance of the immunomagnetic methods relays on the expression and specificity of the targeted antigen, as well as the binding ability of associated antibodies, the efficiency of the immunomagnetic labeling process and magnetic particles, and the separation method of isolated labeled cells. The lack of specificity of current tumor markers can be partially overcome using a mixture of antibodies targeting multiple antigens [27]. This section will give a brief overview of the devices available for direct (positive selection) or indirect CTC capture (negative selection) [21, 23] (Fig. 5.2d).

5.2.2.1 Positive CTC Selection

In positive selection, tumor antigens localized at the cell surface are targeted. Due to the high heterogeneity of CTC surface antigens it has been impossible to identify a universal CTC-specific marker. Currently there are multiple devices available in the market based on CTC positive selection. The most known and the only FDA approved despite their pitfalls previously mentioned, is CellSearch® (Janssen Diagnostics). In addition, they exist in other devices based on the same principle. The AdnaTest (Adnagen AG) uses a cocktail of antibodies specific to the type of cancer, currently this device is able for prostate, breast or ovarian carcinoma. The MagSweeper device, based on EpCAM selection, has the ability of processing whole blood without a previous step of red cell depletion [67]. Microfluidic adhesion-based devices consist in microchannels coated with an antibody against CTCs: the CEE™ chip (Biocept Laboratories) based on a cocktail of antibodies for CTC isolation and the CTC-Chip based on EpCAM selection that have incorporated 3D structures (microposts) to increase the effective surface, and then increasing cell adhesion. The CTC-iChip technology (D.A. Harber, Massachusetts General Hospital Cancer Center; M. Toner, Harvard Medical School, Boston, MA) combines cell size and EpCAM expression for CTC isolation. The IsoFlux™ is an automated, continuous flow microfluidic device based on the recognition of EpCAM expression at the cell surface of CTCs [31].

5.2.2.2 Negative CTC Selection

Another strategy is the indirect isolation of CTCs by successive steps of erythrocytes lysis and magnetically depletion of leukocytes targeting the CD45 antigen, as the most common marker. The RosetteSep™ (STEMCELL Tech) is based on a cocktail of antibodies that specifically cross-link red blood cells to each other and to white blood cells, forming high density cell complexes (rosettes) which can be easily separate from CTCs by a centrifugation step. In the batch separation approach, labeled cells (usually with CD45 antibody) migrate to the regions of higher magnetic frequency when the whole labeled sample is subjected to a magnetic field. The EasySep™ system (STEMCELL Tech), Dynabeads® (ThermoFisher) and MojoSort™ (Biolegend) are commercialized devices based on this principle. Continuous-flow separation devices were designed to increase the processed volume. The sample is continuously fed through within the module. This module can be provided with an activated filter that allow the retention of labeled cells, like the commercially available MACS® (Miltenyi Biotec) and MagniSort™ (eBioscience). There is a negative selection version of CTC-iChip based on the combination of CD45 and CD66b selection and cell size [44]. Negative selection methods are completely independent from the CTC phenotype, do not interact with CTCs which may result in their better viability. The most important requierement of this separation method is a very high specificity to remove all the leukocytes avoiding nonspecific CTC binding in order to achieve an acceptable degree of CTC purity.

5.3 Tumor Heterogeneity

Tumor heterogeneity refers to the fact that different tumor cells can show distinct genetic and phenotypic profiles, including their capacity to

generate metastasis. This phenomenon is observed both between tumors (inter-tumor heterogeneity) and within tumors (intra-tumor heterogeneity), and can be associated to genetic and non-genetic factors.

5.3.1 Clinical Implications of Tumor Heterogeneity and Single CTC Analysis as a Solution

Tumor heterogeneity has several clinical impacts: (i) it introduces significant challenges in using molecular prognostic markers as well as for classifying patients that might benefit from specific therapies; (ii) it must be one of the main causes of drug resistance; (ii) it limits the precision of histological diagnoses and consequently reduces the value of a biopsy; (ii) it is one of the main parameter explaining the lack of numerous therapies in oncology which have not been designed to address the tumor heterogeneity. Thus, for a better understanding of the pathogenesis of the disease would be necessary an accurate characterization of the tumor heterogeneity. Most of the methods used for the tumor characterization give a global overview of the cancer tissue, corresponding to an average picture of all the tumor clones and their microenvironment. In addition, metastatic tumor cells have divergent evolution in different microenvironments at the metastatic sites that may contribute to the change in the expression of biomarkers that were initially identified in the primary tumor [81]. Treating metastatic disease based on the biomarkers expressed in the primary tumor may thus be less than optimal. CTCs are a potential alternative for cancer diagnostic based on biopsies and offer the opportunity for monitoring serial changes in the evolution of tumor biology more accurately than a single biopsy of a primary or metastatic tumor [34, 79]. Latest advances in CTCs research area have offered the possibility of a reliable quantification and molecular characterization of CTCs at the single cell level that will help to do a better cancer diagnosis and therapeutic design [1, 23]. The heterogeneity of CTC has been shown in several studies (more details in Sect. 5.4.3. of this chapter). Dynamic changes or transitions in the biomarker status of CTCs may reflect the presence of the selective pressure that can be exerted by therapeutic interventions. Monitoring changes in the expression of longitudinal biomarkers displayed in CTCs over the course of multiple sequential therapies, may provide insight into the tumor evolution [55]. Furthermore, measuring CTC heterogeneity in time may help to identify the most effective drugs in individual cancer patients [86], especially those who already have or will soon have tumors that are resistant to anti-cancer treatments. Ideally, the identification of promoters or suppressors of genome instability in solid tumors, whose activation or inactivation is required to initiate intratumor heterogeneity and diversity, could provide a tractable means of ultimately attempting to limit tumors evolutionary processes [22].

5.3.2 Methods for Studying the Heterogeneity of Single Cell

In this section we will make a summary of the most used technologies for the identification of molecular differences at the single-cell level that are used nowadays. Advances in Next Generation Sequence (NGS), Whole Genome Analysis (WGA),, fluorescence-activated cell sorting (FACS) and many other methods allow for the analysis of multiple markers in single tumor cells isolated either from fresh or fixed primary tumors and metastases [56, 84]. These technical approaches have also been used for exploring tumor heterogeneity in individual CTCs isolated from blood or bone marrow [34].

From single cells, it is a necessary step of amplification to generate sufficient material for NGS minimizing artefacts. Whole-genome amplification (WGA) is used to analyze single-cell genomes, while Whole-Transcriptome Amplification (WTA) is the previous step for the subsequent analysis of the transcriptome. There are three main Whole Genome Amplification (WGA) methods [10, 37] (Fig. 5.3): (i) *degenerate oligonucleotide-primed polymerase chain reaction* (DOP-PCR) [3]. High amplification

bias, in which only certain regions of the genome are preferentially amplified and thus amenable to large-scale sequencing, results in relatively low coverage of the genome (~10%), making DOP-PCR useful for copy-number assessment in single cells but undesirable for single nucleotide variant (SNV) detection; (ii) *multiple-displacement amplification* (MDA) [15]. MDA works best for mutation detection but is not sufficient for copy number analysis due to moderate amplification bias and non-uniform genome coverage; and (iii) *multiple annealing and looping based amplification cycles* (MALBAC or PicoPLEX) [90]. This method provides high uniformity in coverage across the genome and is useful for detecting copy number variants (CNVs) in single cells. However, MALBAC has a high false positive error rate and it fails in the detection of point mutations. Depending on the objective of the study, there are different genomic interrogations of the amplified genomes. It is possible to query specific gene of interest either by PCR using a target-specific probe; or target capture through hybridization [25]; sequence all exomes (WES) [84]; or sequence the whole genome, called Whole Genome Sequencing (WGS) [56]. This last method provides information such as the single nucleotide variant (SNVs), copy number variants (CNVs) and non-coding and structural variants than the other two approaches don't. The main drawback is the high cost. For the determination of genetic heterogeneity between single cells after WGS must be applied model-based clustering, which allows to include false-negative errors [85].

If the aim of the study is the analysis of heterogeneity at the transcriptomic level, WTA must be the first step before applying single-cell RNA sequencing methods. Currently, there exist many methodologies available for WTA at single-cell level. RNA-seq [78], Quartz-seq [72], CEL-seq2 (cell expression by linear amplification and sequencing) [33], STRT-seq [40], Smart-seq2 (Switching mechanism at the 5′-end of the RNA transcript) [65] are the most common, each one having both assets and drawbacks [70]. The multiplexed error robust FISH (MERFISH) is a single-molecule imaging method that allows thousands of RNA species to be imaged in single cells by using combinatorial FISH labeling with encoding schemes capable of detecting and/or correcting errors [12]. Independently the performed single cell analysis, WGS or Whole-Transcriptome Sequencing (WTS), and due to the high number of technical errors that emerge during the process, the results must be validated, in order to avoid the interpretation of these technical errors as real biological variations. Despite the "omics" techniques initially presented some restrictions in the field of CTC research, many studies have now been described using these methodologies to explore heterogeneity in single CTCs.

Epigenetic events may be another cause of tumor cell heterogeneity, including histone modification, and DNA base modifications (such as methylation and hydroxymethylation). There have been developed robust technologies to provide genome-wide maps of most epigenetic marks, such as the ChIP-seq (Illumina), available for the study histone marks at the single-cell level [46]. To assess methylation levels, the whole-genome bisulphite sequencing (WGBS) offers the most comprehensive picture ([19] but precipitation techniques (methylated-DNA immunoprecipitation sequencing and methylated–DNA-binding domain sequencing) or reduced-representation bisulphite sequencing (RRBS) [17] are also used. Illumina's To minimize the DNA-damaging effects of bisulphite treatment the Infinium Human Methylation450 BeadChip platform has been developed, being an array-like alternative that has been found to provide acceptable DNA methylation profiles, [54].

Multiple methods for quantifying proteins in single cells have been developed, including single-cell Western blots [38], CyTOF–a mass cytometry in which antibodies are labeled with heavy metal ions- [7] and Proseek Multiplex [13], however these methods are limited by the quantity of proteins that can be analyzed (up to 40) and by the availability of highly-specific antibodies. To overcome some of these drawbacks, Budnik et al. designed the most recent high-throughput method that can quantify thousands of proteins in single cells: the Single Cell ProtEomics by Mass Spectrometry (SCoPE-MS) [9].

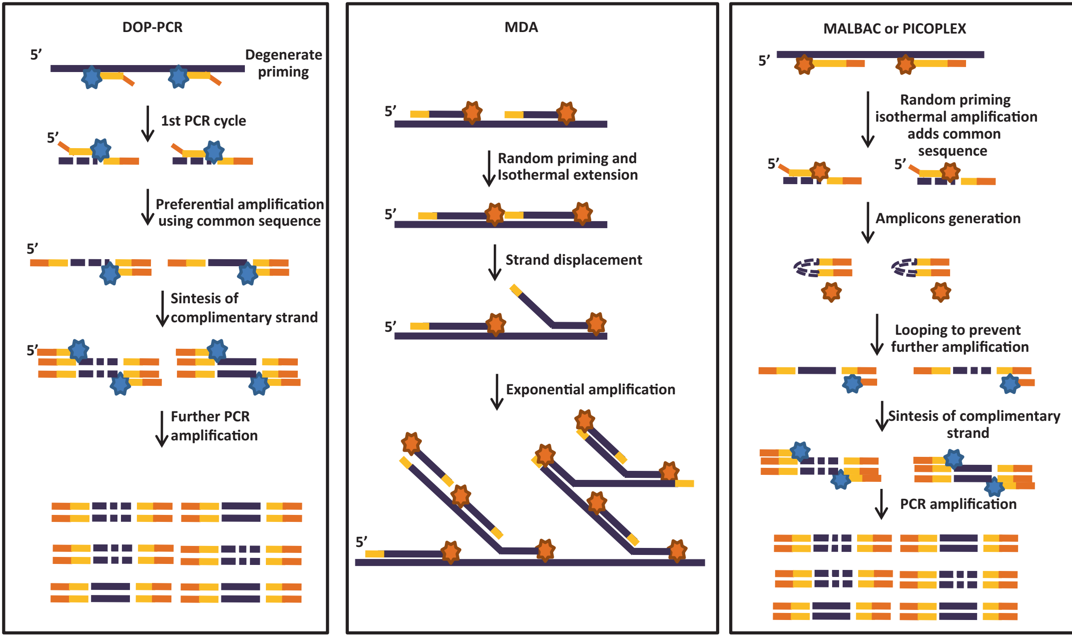


Fig. 5.3 Overview of the three main whole-genome amplification methods. Pure PCR methods such as degenerate oligonucleotide primed PCR (DOP-PCR) (**a**), Isothermal methods such as multiple displacement amplification (MDA) (**b**) and hybrid methods such as multiple annealing and looping based amplification cycles (MALBAC) and PicoPLEX

5.4 Clinical Value of CTCs

Historically, the therapeutic decision was made essentially from the analysis of the primary tumor. It is now admitted that the therapeutic strategy should integrate the analysis of metastatic lesions that may become detectable many years after primary tumor detection. However, taking a biopsy from a metastatic site is always an invasive procedure, thus the analysis of CTCs might be a noninvasive good alternative (liquid biopsy) for monitoring treatment efficacy in real-time and for detecting metastatic relapse and disease progression. In contrast to the characterization of primary tumors, that only provides a static view at the time of diagnosis, the analysis of CTCs may improve understanding of the different steps involved in the metastatic cascade, from invasion of tumor cells into the bloodstream to the formation of clinically detectable metastases [60].

5.4.1 Analysis of Single CTC in Clinical Practice

To implement single cell analysis in the clinic, blood collection, handling protocols require standardization (Standard Operating Procedures) and validation methods in order to minimize sample alterations and false positive or negative results. This section will review some of the most recent studies in the analysis of single CTCs heterogeneity with high clinical and scientific impact in different cancer types, using different technologies described in previous sections.

5.4.1.1 Colorectal Cancer (CRC)

Gasch C et al. used micromanipulation technique in order to obtain single CTC from previous enriched blood samples from CRC patients. Afterwards, these authors performed WGA for the sequencing of KRAS (v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog), BRAF (v-raf

murine sarcoma viral oncogene homolog B1) and PIK3CA (phosphatidylinositol-4,5-bisphosphate 3-kinase, catalytic subunit α) genes and also analyzed the EGFR amplification by quantitative PCR (qPCR). The results revealed a high intra- and inter-patient heterogeneity in EGFR expression and genetic alterations in EGFR, KRAS and PIK3A, which could explain differences in response rates to EGFR inhibition therapy in colorectal cancer patients [24]. One study conducted by Heitzer et al. compared the mutation spectrum in the primary tumors (PT), metastases (MET), and the corresponding CTCs of a panel of 68 CRC-associated genes by using array-CGH and next-generation sequencing. CTC from blood patients were enriched using the CellSearch® system, and individual CTCs were isolated with the help of micromanipulation. They found mutations on known driver genes, like APC, KRAS or PIK3CA in PT, MET and corresponding CTCs. This study supported the idea of using CTCs as a liquid biopsy in cancer patients, can be an additional option to monitor tumor genomes, which are prone to change during progression, treatment and relapse [34]. Using the DEPArray™ platform for the isolation of single CTCs from peripheral blood of patients with metastatic colon cancer and further WGA amplification, Frabbri et al., screened KRAS gene status mutation in CTCs and in the primary tumor tissue. They found more than 50% of discordance of KRAS gene mutation between CTCs and primary tissue, indicating again the importance of CTC monitoring to detect the changing genome along cancer progression [18].

5.4.1.2 Breast Cancer

Babayan et al. isolated CTCs by micromanipulation after a previous enrichment step of Ficoll from blood metastatic breast cancer patients, they assessed the expression of the estrogen receptor (ER) after WGA. Interestingly, they found that individual CTCs were both ER+ and ER- in patients with ER+ primary tumor, which may explain the 20% of ER-targeted therapy fails. These results suggest that ER expression monitoring on CTCs could optimize breast cancer treatment [6]. In a similar manner, Fernandez et al., isolated single CTC from metastatic triple negative breast cancer patients. The CellSearch™ system was used as a primary enrichment of CTC for further single-cell selection in the DEPArray™. After WGA, they analyze TP53 mutations in single isolated CTC, resulting in high heterogeneity and in some cases different to primary tumor status [20]. Pestrin et al. combined similar technology and found a high heterogeneity of PIK3CA mutational status in CTCs from metastatic breast cancer patients [64]. Such variations of PIK3CA mutational status have been confirmed during progressive and recurrent disease in breast cancer patients demonstrating the clinical utility to isolated CTCs combined with a PIK3CA allele-specific asymmetric rapid PCR and melting analysis of the amplified single cell DNA [51]. A recent study of De Luca et al., analyzed by NGS 50 cancer-related genes of the amplified DNA from CTC breast cancer patients, previously enriched by the combination of the CellSearh and the Deparray. They observed 51 sequence variants in 25 genes which demonstrated again the high heterogeneity in the mutational status of CTCs in cancer either between patients or within the same patient[14]. Jordan et al. identified dynamic functional states of HER2 within circulating breast cancer cells. Fresh whole blood was first processed through the microfluidic CTC-iChip for depletion of leukocytes and CTC isolation, individual CTCs were picked by micromanipulation. The RNAseq of individual CTCs documented the emergence of HER2+ CTCs in women initially diagnosed HER2-negative breast cancer. These results suggested the capacity of HER2 phenotype spontaneous interconversion in CTC may be responsible for the progression of breast cancer and acquisition of drug resistance [43]. Interestingly, Powell et al. were pioneers in the studies on CTCs focused on single-cell transcriptomic data. They introduced purified CTCs from the MagSweeper to a microfluidic platform (Fluidigm) to perform multiplexed quantitative PCR (qPCR) on 87 cancer genes in individual CTCs. Importantly, genes associated with metastatic process and EMT presented heterogeneous expression [67].

5.4.1.3 Prostate Cancer

In a study performed by Jiang et al., the authors combined the NanoVelcro chip and laser capture microdissection for the isolation of single CTCs. They amplified the single cell DNA by MDA and performed WGS. Importantly, they validated a high resolution protocol for the study of single CTC heterogeneity of the mutational landscape of somatic single nucleotide variations (SSNVs) and structural variations (SVs) [42]. Another single-cell analysis of prostate CTCs revealed an increased expression pattern of EMT-related genes in metastatic prostate cancer patients. CTC were enriched from blood by the ScreenCell® CC system and individual cells were recovered by micromanipulation. For the single-cell DNA analysis, Chen et al. used a microfluidic-based RT-PCR analysis (CellsDirect™) and the Fluidigm [11]. Miyamoto et al. isolated individual CTCs from prostate cancer by micromanipulation and previously enriched using the CTC-i-Chip. Afterwards they performed RNA amplification and WTA. They identified multiple differences in gene expression among CTCs isolated from the same patient and between CTCs and the primary tumor. Interestingly they found heterogeneity in the various signalling pathways that could contribute to the anti-androgen therapy failure [53]. A different strategy was used by Lohr et al. for the analysis of CTC isolated from metastatic prostate cancer patients. They enriched CTC fraction with the Magsweeper device, and then used two different strategies for single cell isolation before WES: a nanowell-based method with automated imaging and robotic retrieval of cells, and micromanipulation. Data supported the idea that the sequencing of can detect mutations in the early steps of tumor evolution as well as those shared with metastatic sites or trunk mutations [49].

5.4.1.4 Other Malignancies

In other types of cancer, the use of CTCs for clinical purposes has reported less studies or they are still ongoing. This section point out some of the studies already published in different malignancies.

A recent study in **pancreatic cancer** showed CTCs associated aberrations in the expression of extracellular matrix genes, then underlining the involvement of the tumor microenvironment for the spread of cancer to distant organ. CTCs were enriched by using the CTC-iChip previous to micromanipulation for single cell collection to further RNA amplification and WTS [80]. Ni et al. analyzed CTCs in **non-small-cell lung cancer** (NSCLC) and found a heterogeneity of single-nucleotide variations (SNVs) and insertions/deletions (INDELs), but with similar profiles than metastatic tumor tissues. These genes were involved in resistance to drugs and phenotypic transitions, indicating the utility of sequencing CTCs for cancer diagnosis. Authors performed an enrichment step of CTCs by CellSearch® system and isolated single cells by micropipetting before performing WES to analyze SNVs and INDELs [57]. Ramsköld, D et al. performed a Smart-Seq analysis in single CTCs from **melanoma** patients, isolated by micromanipulation previously enriched with the MagSweeper instrument. The transcriptomic analysis revealed high heterogeneity in gene expression patterns of melanoma CTCs highly informative for identifying candidate biomarkers [69]. The study performed by Blassl et al., found inter-cellular and intra/inter patient heterogeneity, single CTCs isolated from **ovarian** cancer patients bear the co-expression of epithelial, mesenchymal and stem cell transcripts. For single CTCs analysis, the first step was the enrichment of CTC by a combination of density gradient centrifugation and CD45 depletion, followed by an immunofluorescent labeling with EpCAM and Muc-1 antibodies and isolation of single cells via micromanipulation. Remarkably, they developed a gene expression analysis by reverse transcription and multiplex-RT-PCR analysis without the need of pre-amplification steps [8]. Recently Park S M et al. reported a high throughput, multiplexed strategy for single-cell mutation profiling of individual lung cancer CTCs for minimally invasive cancer therapy prediction and disease monitoring. Briefly, this nanoplatform consists in a magnetic sifter and single cell nanowell array capable of sorting up to 25,600 single CTCs from blood to capture, compartmentalize, and molecularly characterize them [61].

5.5 Discussion

Tumor heterogeneity is the result of the natural history of the tumor development and is a key issue of therapeutic approaches and its follow up is challenging. The identification of specific biomarkers predicting the therapeutic response or/and the metastatic process or/and the local recurrence would help to the therapeutic management. CTCs belong to the list of potential biomarkers and their heterogeneity could reflect those of the tumor.

The rarity of CTCs in the bloodstream of cancer patients made necessary the development of very accurately methodologies for their enrichment and isolation. Despite the numerous methodologies reported in the literature (reviewed in Sect. 5.2), there are still pitfalls that difficult the recovery of all the CTCs present in the sample. Moreover, some technologies are still at the proof-of concept stage only demonstrated using cultured cells. However, frequently set up on cell lines which do not effectively reflect the real state of CTCs in a natural biological fluid, especially in terms of heterogeneity [5], separation technologies over-predict their performance. The isolation of new cell lines exhibiting a large genomic and transcriptomic heterogeneity is mandatory. CTCs can be isolated and interrogated as a "potential biological window "into the genetics of a tumor by means of non-invasive sampling, making it possible to evaluate the tumor temporal heterogeneity during the clinical course of the disease. To overcome the averaging approach determined by using bulk CTC analysis, it is necessary to perform a single-cell analysis using sophisticated instrumentation. However, data mined from single-cell studies must be carefully studied with rigorous computational analyses, which are imperative for distinguishing pre-existing genetic alterations from amplification errors. Improper computation may result in bias and errors in data interpretation, especially in highly heterogeneous samples. Independent analyses from multiple single cells may also be used to reveal repeated and specific mutational patterns, which will be pivotal in distinguishing technical noise from biological signals. A complete picture of a cell state will often require measurements of different parameters in the same cells (i.e., transcriptome, genome and epigenome). Although it is usually possible to perform multiple assays on a bulk sample, this is only possible in certain cases for single-cell measurements. Further development of multi-model measurement methods will help understand different sources of heterogeneity. In addition, future studies on single cells with defined clinical endpoints need to address the question of how many CTCs should be profiled to account for heterogeneity. Another consideration that affects directly to clinical decision is that depending on the methodology used for their isolation, the number of recovered CTCs may be altered and thus the analysis of heterogeneity will be differentially reported. The relevance and significance of CTCs is by itself questionable. Do these cells have malignant properties themselves? It is more relevant to analyze CTC clusters? Indeed, recent findings have identified the importance of CTC clusters in tumor dissemination [4]. This may be a major source of heterogeneity at the metastatic point, and may imply more effort for developing methodologies to isolate and characterize CTC clusters, such as the ClusterChip.

Even of a lot of questions remain today without a clear answer, many studies have demonstrated the success of single CTCs characterization for clinical purposes, especially in colorectal cancer, prostate cancer and breast cancer. However, the development of new CTC isolation and single cell analysis technologies, that avoid current pitfalls, as well as the improvement on the data analysis could improve significantly their clinical value and thus increase their use in clinical practice.

References

1. Alix-Panabieres C, Pantel K (2014) The circulating tumor cells: liquid biopsy of cancer. Klin Lab Diagn 4:60–64
2. Andreopoulou E et al (2012) Comparison of assay methods for detection of circulating tumor cells in metastatic breast cancer: adnagen adnatest breastcancer select/detect versus veridex CellSearch system. Int J Cancer 130(7):1590–1597

3. Arneson N et al (2008, 2008) Whole-genome amplification by degenerate oligonucleotide primed PCR (DOP-PCR). CSH Protoc. pdb prot4919
4. Au SH et al (2016) Clusters of circulating tumor cells traverse capillary-sized vessels. Proc Natl Acad Sci U S A 113(18):4947–4952
5. Auman JT, McLeod HL (2010) Colorectal cancer cell lines lack the molecular heterogeneity of clinical colorectal tumors. Clin Colorectal Cancer 9(1):40–47
6. Babayan A et al (2013) Heterogeneity of estrogen receptor expression in circulating tumor cells from metastatic breast cancer patients. PLoS One 8(9):e75038
7. Bandura DR et al (2009) Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. Anal Chem 81(16):6813–6822
8. Blassl C et al (2016) Gene expression profiling of single circulating tumor cells in ovarian cancer–establishment of a multi-marker gene panel. Mol Oncol 10(7):1030–1042
9. Bogdan Budnik EL, Slavov N (2017) Mass-spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. bioRxIV: this article is a preprint and has not been peer-reviewed
10. Borgstrom E et al (2017) Comparison of whole genome amplification techniques for human single cell exome sequencing. PLoS One 12(2):e0171566
11. Chen CL et al (2013) Single-cell analysis of circulating tumor cells identifies cumulative expression patterns of EMT-related genes in metastatic prostate cancer. Prostate 73(8):813–826
12. Chen KH et al (2015) RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. Science 348(6233):aaa6090
13. Darmanis S et al (2016) Simultaneous multiplexed measurement of RNA and proteins in single cells. Cell Rep 14(2):380–389
14. De Luca F et al (2016) Mutational analysis of single circulating tumor cells by next generation sequencing in metastatic breast cancer. Oncotarget 7(18):26107–26119
15. Dean FB et al (2002) Comprehensive human genome amplification using multiple displacement amplification. Proc Natl Acad Sci U S A 99(8):5261–5266
16. Desitter I et al (2011) A new device for rapid isolation by size and characterization of rare circulating tumor cells. Anticancer Res 31(2):427–441
17. Doherty R, Couldrey C (2014) Exploring genome wide bisulfite sequencing for DNA methylation analysis in livestock: a technical assessment. Front Genet 5:126
18. Fabbri F et al (2013) Detection and recovery of circulating colon cancer cells using a dielectrophoresis-based device: KRAS mutation status in pure CTCs. Cancer Lett 335(1):225–231
19. Farlik M et al (2015) Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. Cell Rep 10(8):1386–1397
20. Fernandez SV et al (2014) TP53 mutations detected in circulating tumor cells present in the blood of metastatic triple negative breast cancer patients. Breast Cancer Res 16(5):445
21. Ferreira MM et al (2016) Circulating tumor cell technologies. Mol Oncol 10(3):374–394
22. Fisher R et al (2013) Cancer heterogeneity: implications for targeted therapeutics. Br J Cancer 108(3):479–485
23. Gabriel MT et al (2016) Circulating tumor cells: a review of non-EpCAM-based approaches for cell enrichment and isolation. Clin Chem 62(4):571–581
24. Gasch C et al (2013) Heterogeneity of epidermal growth factor receptor status and mutations of KRAS/PIK3CA in circulating tumor cells of patients with colorectal cancer. Clin Chem 59(1):252–260
25. Gawad C et al (2014) Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. Proc Natl Acad Sci U S A 111(50):17947–17952
26. Gerges N et al (2010) New technologies for the detection of circulating tumour cells. Br Med Bull 94:49–64
27. Ghazani AA et al (2013) Comparison of select cancer biomarkers in human circulating and bulk tumor cells using magnetic nanoparticles and a miniaturized micro-NMR system. Nanomedicine 9(7):1009–1017
28. Ghossein RA et al (1999) Molecular detection of micrometastases and circulating tumor cells in solid tumors. Clin Cancer Res 5(8):1950–1960
29. Gupta V et al (2012) ApoStream(), a new dielectrophoretic device for antibody independent isolation and recovery of viable cancer cells from blood. Biomicrofluidics 6(2):24133
30. Hamilton G, Rath B (2016) Detection of circulating tumor cells in non-small cell lung cancer. J Thorac Dis 8(6):1024–1028
31. Harb W et al (2013) Mutational analysis of circulating tumor cells using a novel microfluidic collection device and qPCR assay. Transl Oncol 6(5):528–538
32. Harouaka RA et al (2014) Flexible micro spring array device for high-throughput enrichment of viable circulating tumor cells. Clin Chem 60(2):323–333
33. Hashimshony T et al (2016) CEL-Seq2: sensitive highly-multiplexed single-cell RNA-Seq. Genome Biol 17:77
34. Heitzer E et al (2013) Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. Cancer Res 73(10):2965–2975
35. Hofman V et al (2011) Detection of circulating tumor cells as a prognostic factor in patients undergoing radical surgery for non-small-cell lung carcinoma: comparison of the efficacy of the CellSearch assay and the isolation by size of epithelial tumor cell method. Int J Cancer 129(7):1651–1660
36. Hong Y et al (2016) Circulating tumor cell clusters: what we know and what we expect (review). Int J Oncol 49(6):2206–2216

37. Hou Y et al (2015) Comparison of variations detection between whole-genome amplification methods used in single-cell resequencing. Gigascience 4:37
38. Hughes AJ et al (2014) Single-cell western blotting. Nat Methods 11(7):749–755
39. Husemann Y et al (2008) Systemic spread is an early step in breast cancer. Cancer Cell 13(1):58–68
40. Islam S et al (2012) Highly multiplexed and strand-specific single-cell RNA 5′ end sequencing. Nat Protoc 7(5):813–828
41. Jacob K et al (2007) Circulating tumor cells: detection, molecular profiling and future prospects. Expert Rev Proteomics 4(6):741–756
42. Jiang R et al (2015) A comparison of isolated circulating tumor cells and tissue biopsies using whole-genome sequencing in prostate cancer. Oncotarget 6(42):44781–44793
43. Jordan NV et al (2016) HER2 expression identifies dynamic functional states within circulating breast cancer cells. Nature 537(7618):102–106
44. Karabacak NM et al (2014) Microfluidic, marker-free isolation of circulating tumor cells from blood samples. Nat Protoc 9(3):694–710
45. Khoo BL et al (2014) Clinical validation of an ultra high-throughput spiral microfluidics for the detection and enrichment of viable circulating tumor cells. PLoS One 9(7):e99409
46. Lang JE et al (2015) Expression profiling of circulating tumor cells in metastatic breast cancer. Breast Cancer Res Treat 149(1):121–131
47. Liljefors M et al (2005) Clinical effects of a chimeric anti-EpCAM monoclonal antibody in combination with granulocyte-macrophage colony-stimulating factor in patients with metastatic colorectal carcinoma. Int J Oncol 26(6):1581–1589
48. Lim LS et al (2012) Microsieve lab-chip device for rapid enumeration and fluorescence in situ hybridization of circulating tumor cells. Lab Chip 12(21):4388–4396
49. Lohr JG et al (2014) Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat Biotechnol 32(5): 479–484
50. Ma YC et al (2013) Recent advances and prospects in the isolation by size of epithelial tumor cells (ISET) methodology. Technol Cancer Res Treat 12(4):295–309
51. Markou A et al (2014) PIK3CA mutational status in circulating tumor cells can change during disease recurrence or progression in patients with breast cancer. Clin Cancer Res 20(22):5823–5834
52. Mikolajczyk SD et al (2011) Detection of EpCAM-negative and Cytokeratin-negative circulating tumor cells in peripheral blood. J Oncol 2011:252361
53. Miyamoto DT et al (2015) RNA-Seq of single prostate CTCs implicates noncanonical Wnt signaling in antiandrogen resistance. Science 349(6254): 1351–1356
54. Morris TJ et al (2014) ChAMP: 450k chip analysis methylation pipeline. Bioinformatics 30(3):428–430
55. Nakazawa M et al (2015) Serial blood-based analysis of AR-V7 in men with advanced prostate cancer. Ann Oncol 26(9):1859–1865
56. Navin N et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472(7341):90–94
57. Ni X et al (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 110(52):21083–21088
58. Pamme N (2012) On-chip bioanalysis with magnetic particles. Curr Opin Chem Biol 16(3–4):436–443
59. Pantel K et al (2012) Circulating epithelial cells in patients with benign colon diseases. Clin Chem 58(5):936–940
60. Pantel K, Speicher MR (2016) The biology of circulating tumor cells. Oncogene 35(10):1216–1224
61. Park SM et al (2016) Molecular profiling of single circulating tumor cells from lung cancer patients. Proc Natl Acad Sci U S A 113(52):E8379–E8386
62. Patriarca C et al (2012) Epithelial cell adhesion molecule expression (CD326) in cancer: a short review. Cancer Treat Rev 38(1):68–75
63. Peeters DJ et al (2013) Semiautomated isolation and molecular characterisation of single or highly purified tumour cells from CellSearch enriched blood samples using dielectrophoretic cell sorting. Br J Cancer 108(6):1358–1367
64. Pestrin M et al (2015) Heterogeneity of PIK3CA mutational status at the single cell level in circulating tumor cells from metastatic breast cancer patients. Mol Oncol 9(4):749–757
65. Picelli S et al (2013) Smart-seq2 for sensitive full-length transcriptome profiling in single cells. Nat Methods 10(11):1096–1098
66. Posel C et al (2012) Density gradient centrifugation compromises bone marrow mononuclear cell yield. PLoS One 7(12):e50293
67. Powell AA et al (2012) Single cell profiling of circulating tumor cells: transcriptional heterogeneity and diversity from breast cancer cell lines. PLoS One 7(5):e33788
68. Rack B et al (2014) Circulating tumor cells predict survival in early average-to-high risk breast cancer patients. J Natl Cancer Inst 106(5):1–11
69. Ramskold D et al (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30(8):777–782
70. Saliba AE et al (2014) Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res 42(14):8845–8860
71. Sarioglu AF et al (2015) A microfluidic device for label-free, physical capture of circulating tumor cell clusters. Nat Methods 12(7):685–691
72. Sasagawa Y et al (2017) Erratum to: quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 18(1):9
73. Sleijfer S et al (2007) Circulating tumour cell detection on its way to routine diagnostic implementation? Eur J Cancer 43(18):2645–2650

74. Small AC et al (2012) The emerging role of circulating tumor cell detection in genitourinary cancer. J Urol 188(1):21–26
75. Sollier E et al (2014) Size-selective collection of circulating tumor cells using vortex technology. Lab Chip 14(1):63–77
76. Steeg PS (2006) Tumor metastasis: mechanistic insights and clinical challenges. Nat Med 12(8):895–904
77. Talasaz AH et al (2009) Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device. Proc Natl Acad Sci U S A 106(10):3970–3975
78. Tang F et al (2010) RNA-Seq analysis to capture the transcriptome landscape of a single cell. Nat Protoc 5(3):516–535
79. Tellez-Gabriel M et al (2016) Tumour heterogeneity: the key advantages of single-cell analysis. Int J Mol Sci 17(12):2142
80. Ting DT et al (2014) Single-cell RNA sequencing identifies extracellular matrix gene expression by pancreatic circulating tumor cells. Cell Rep 8(6):1905–1918
81. Tlsty TD, Coussens LM (2006) Tumor stroma and regulation of cancer development. Annu Rev Pathol 1:119–150
82. Vykoukal J et al (2008) Enrichment of putative stem cells from adipose tissue using dielectrophoretic field-flow fractionation. Lab Chip 8(8):1386–1393
83. Xu L et al (2015) Optimization and evaluation of a novel size based circulating tumor cell isolation system. PLoS One 10(9):e0138032
84. Xu X et al (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148(5):886–895
85. Yeung KY et al (2011) Construction of regulatory networks using expression time-series data of a genotyped population. Proc Natl Acad Sci U S A 108(48):19436–19441
86. Yu M et al (2014) Cancer therapy. Ex vivo culture of circulating breast tumor cells for individualized testing of drug susceptibility. Science 345(6193):216–220
87. Zhe X et al (2011) Circulating tumor cells: finding the needle in the haystack. Am J Cancer Res 1(6):740–751
88. Zheng S et al (2007) Membrane microfilter device for selective capture, electrolysis and genomic analysis of human circulating tumor cells. J Chromatogr A 1162(2):154–161
89. Zhou MD et al (2014) Separable bilayer microfiltration device for viable label-free enrichment of circulating tumour cells. Sci Rep 4:7392
90. Zong C et al (2012) Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. Science 338(6114):1622–1626

Dominique Heymann, PhD, is a Professor of Histology and Embryology at the University of Nantes (FR) and hospital prtaicioner in biology at the Integrated Cancer Centre (ICO, Saint-Herblain, FR). He is also honorary –professor of bone oncology at the University of Sheffield (UK). He heads the Inserm European Associated Laboratory "Sarcoma Research Unit". He has authored more that 240 articles in peer-review-journals, 23 book chapters, 2 books and 6 patents His main research is focused on the role of micro-environment in the pathogenesis of sarcomas and the isolation and characterization of rare cells in cancers.

Marta Téllez-Gabriel, PhD, is a postdoc at the Oncological Hematology and Transplants group at the Biomedical Research Institute of Sant Pau in Barcelona. She has authored 11 articles in peer-review-journals and 1 book chapter. Her main research is focused on the development of CART cells for the treatment of different hematological cancers. Also she is in collaboration with Pr Heymann for the isolation and characterization of rare cells in cancers.

6 Super-Resolution Fluorescence Microscopy for Single Cell Imaging

Han Feng, Xiaobo Wang, Zhiwei Xu, Xiaoju Zhang, and Yongju Gao

Abstract

In the past two decades, super-resolution fluorescence microscopy has undergone a dynamic evolution. Following proof-of-concept studies with stimulated emission depletion (STED) microscopy, several new approaches such as structured illumination microscopy (SIM), photoactivation localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM), have been developed for imaging of nanoscale structural details and fast cellular dynamics in biological research. In this chapter, after briefly explaining their principles, we will describe the recent application of these super-resolution techniques in single cell imaging. In addition, the extension of super-resolution microscopy to 3D, multicolor, live-cell imaging and multimodal imaging are also discussed, significantly improving the precision of single cell imaging. Combining with molecular biology, biochemistry and bio-computing algorithms, super-resolution fluorescence microscopy continues to expand its capabilities and provide comprehensive insights into the details of single cells.

Keywords

STED microscopy · SIM · PALM · STORM · Single cell imaging

H. Feng
Department of Pharmacy, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China

X. Wang (✉)
Department of Nuclear Medicine, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China

Clinical Research Service Center, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China
e-mail: xbwang@csu.edu.cn

Z. Xu
Clinical Research Service Center, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China

X. Zhang (✉)
Clinical Research Service Center, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China

Department of Respiratory Medicine, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China

Y. Gao (✉)
Department of Nuclear Medicine, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University, Zhengzhou, China

© Springer Nature Singapore Pte Ltd. 2018

J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_6

6.1 Introduction

The cellular structures and biological mechanisms undergoing fundamental human life processes often occur at a nanometer scale in single cells. It is critical to develop analytical methods for observing single cells at a nanoscale resolution, aiming at fully understanding the cellular processes such as apoptosis and carcinogenesis, addressing fundamental biological issues and ultimately providing unique strategies for the diagnosis and treatment of disease [1, 2]. Among them, various attempts have been focused on single cell imaging. Electron microscopy (EM) has atomic resolution of approximately 0.1–10 nm and is able to reveal the ultrastructural characterization of cellular organelles [3, 4]. However, the fixation procedures in sample preparation make it problematic for single live-cell imaging. Fluorescence microscopy offers a ubiquitous choice for single cell imaging in biological studies [5]. Unfortunately, the optical resolution of fluorescence microscopy has been limited by Abbe's law and has been restricted within 200 nm [6]. Thus, various biological processes that occur inside the single cell at nanometer scale are difficult to be resolved by fluorescence microscopy.

It is exceptionally challenging for achieving high spatiotemporal resolutions to observe nanoscale structural details and monitor fast cellular dynamics.

The emergence of super-resolution fluorescence microscopy has begun to attract notice because it can bypass the barrier of light diffraction and obtain nanoscale resolution in single cell imaging [7–10]. In the past two decades, super-resolution fluorescence microscopy has undergone a dynamic evolution. Following proof-of-concept studies with stimulated emission depletion (STED) microscopy, several new approaches such as structured illumination microscopy (SIM), photoactivation localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM), have been developed for super-resolution imaging in biological research. These innovative techniques have improved the resolution by an order of magnitude or more over conventional fluorescence microscopy and have empowered the visualization nanoscale structural details and fast cellular dynamics at the subcellular and molecular scale with unprecedented details. In 2014, the Nobel Prize in Chemistry has been awarded to Prof. Eric Betzig, Prof. William Esco Moerner, and Prof. Stefan Walter Hell for their ground-breaking work in super-resolution fluorescence microscopy [11–13].

In this chapter, after briefly explaining their principles, we describe the recent application of these super-resolution techniques in single cell imaging. In addition, the extension of super-resolution microscopy to 3D, multicolor, live-cell imaging and multimodal imaging are also discussed, significantly improving the precision of single cell imaging. Combining with molecular biology, biochemistry and bio-computing algorithms, super-resolution fluorescence microscopy continues to expand its capabilities and provide comprehensive insights into the details of single cells.

6.2 Super-Resolution Techniques for Single Cell Imaging

In general, super-resolution fluorescence microscopy is a type of far-field optical technique that relies on different principles for obtaining super-resolution (Fig. 6.1). Despite its relatively short history, super-resolution fluorescence microscopy has had a profound influence on almost all branches of life science. It is impossible to describe all of these applications in this chapter. Instead, after briefly explaining their principles, we will discuss some representative examples of these super-resolution techniques for single cell imaging in cell biology.

6.2.1 Stimulated Emission Depletion (STED) Microscopy

In 1994, Hell and his coworkers first proposed the schematic diagram of STED microscopy that was experimentally demonstrated as a sub-diffraction

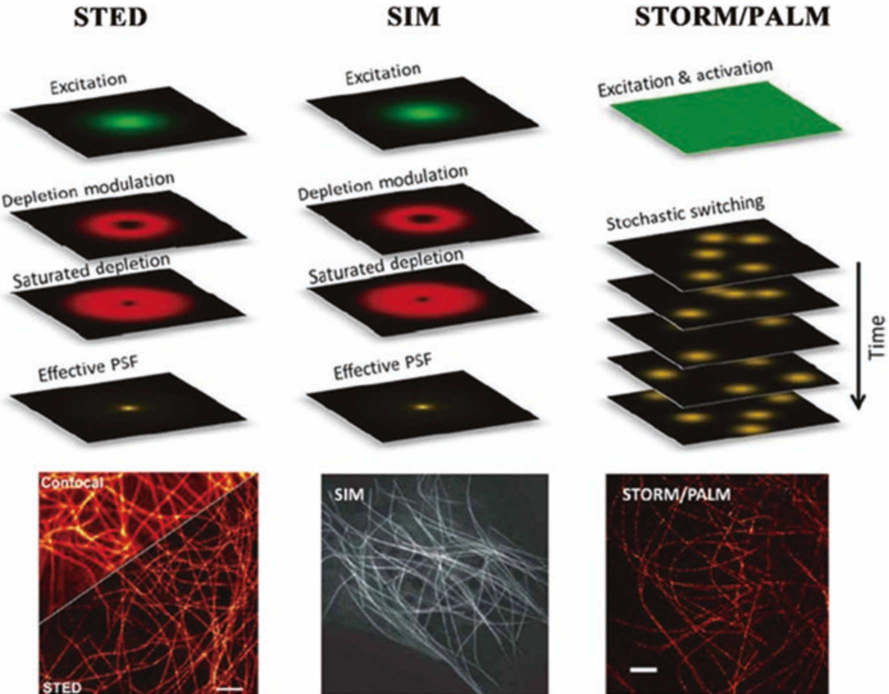


Fig. 6.1 Super-resolution fluorescence microscopy. Upper panel: Principles of super-resolution microscopy techniques. Lower panel: Confocal and super-resolution images of fluorescent protein labeled microtubules in living cells, showing confocal and STED microscopy of mCitrine-tubulin in a living PtK2 cell, SIM of EGFP-tubulin in a living Drosophila S2 cell, and STORM/PALM of mEos2-tubulin in a living Drosophila S2 cell, respectively. All images are shown with the same magnification. Scale bars: 2 μm. (Adapted with permission from Ref. [8], Copyright 2013 American Chemical Society)

imaging technique later [14, 15]. In STED microscopy, the fluorescence response is negatively modulated by a STED laser beam using the process of stimulated emission. Generally, when an excited fluorophore encounters an appropriate photon, it can jump to the ground state with emitting an identical photon. However, the stimulated emission can effectively deplete excited fluorophores without fluorescence emission using a strong STED laser. Because of the doughnut-shaped spatial pattern of laser beam, the fluorescence emission is generated from a smaller subset of fluorophores at the center of doughnut while suppressing the majority of fluorophores located in the overlapping region of the two laser beams. This effectively narrows the point spread functions (PSFs), minimizes blurring, and ultimately accomplishes resolution improvement beyond the diffraction limit.

Since first proposed in 1994, STED microscopy has seen its paradigm-shifting application in many aspects of cell biology [16–22]. For example, Hell et al. demonstrated that this technique has the ability to visualize individual vesicles in the synapse (approximately 40 nm in diameter). This study revealed synaptotagmin I, a protein in the vesicle membrane, remains isolated clusters on the presynaptic membrane [16]. Most impressively, video-rate STED microscopy imaging of synaptic vesicle movement in live hippocampal neurons was achieved by Westphal et al. with a 62 nm lateral resolution using ATTO647N-conjugated anti-synaptotagmin antibodies [23]. STED microscopy revealed the vesicle mobility within the highly confined space of synaptic boutons. However, the movement of vesicle was substantially faster in non-bouton regions, which might represent continuous transit through axons.

Along with STED microscopy technique, the development of new fluorescent probes has also attracted great attention of chemists and biochemists. Correspondingly, a surge of fluorescent probes including organic dyes, fluorescent protein and nanomaterials have been furnished for STED microscopy [24–27]. For example, super-resolution imaging of the Golgi apparatus structures and dynamics in single live-cells with STED microscopy using a bioorthogonal ceramide probe was achieved [28]. Schepartz et al. treated live-cells with a trans-cyclooctene-containing ceramide lipid (Cer-TCO) for targeting Golgi, and then utilize a tetrazine-tagged near-IR dye, SiR-Tz, to specifically recognize Cer-TCO *via* a "tetrazine-click" reaction. The assembling production of Cer-SiR, a highly photostable "vital dye", enabled the visualization of the Golgi apparatus by STED super-resolution microscopy in live-cells (Fig. 6.2). In addition, this "vital dye" was sufficiently safe and did not perturb the mobility of the Golgi-resident enzymes as well as the traffic of cargo through the Golgi. The commercially available quantum dots (QDs) with red-emittion have been employed in STED super-resolution microscopy by using an increasingly popular 775 nm STED laser light. Super-resolution imaging of fibroblasts with QDs-labeled vimentin filaments in 50 nm spatial resolution were obtained [29]. Specifically, the high photo-stability of QDs enables more than 1000 frames of superimposed STED scans without blinking; consequently, QDs hold promise for extended time-lapse imaging.

6.2.2 Structured Illumination Microscopy (SIM)

SIM, first reported by Gustafssonis et al., is another type of super-resolution techniques [30, 31]. In SIM, the fluorescent response is positively modulated by a sinusoidal pattern, which is generated by the interference of two excitation light beams through the excitation optics. This modulation pattern was used to reveal the hidden frequencies from unresolved sample structures that can be used to reconstruct a full SIM image. To achieve this, a set of images were captured by changing the defined angles and the phase of sinusoidal modulation pattern. Compared with conventional fluorescence microscope, SIM has a ~2-fold improvement in resolution, with an approximately 100–150 nm practical resolution. Furthermore, three dimension (3D) SIM has been achieved by using a 3D modulation pattern that is created by the interference of three excitation laser, resulting in a ~2-fold resolution improvement in all three dimensions [32]. In addition, saturated structural illumination microscopy (SSIM), a variation of SIM, has been developed for increasing the resolution up to ~ 5-fold by using nonlinear patterned excitation [33].

SIM improves its resolution by means of optics and is therefore compatible with all fluorophores and labeling protocols previously developed for conventional fluorescence microscopy. Second, it is based on wide-field microscopy techniques and needs very few frames for reconstruction of SIM images. Therefore, SIM is faster compared with other high-resolution methods. These features make it favorable for super-resolution imaging in single live-cells. SIM has demonstrated its capability for long term, live-cell imaging in subcellular dynamic structures such as microtubules [34, 35]. Furthermore, 3D-SIM has been achieved and opens new and facile possibilities for sub-diffraction multicolor imaging [36–38]. For example, Schermelleh and his colleagues performed multicolor imaging of the nuclear periphery including chromatin, nuclear lamina, and the nuclear pore complex (NPC) in single mammalian cells with 3D structured illumination microscopy [36]. Several features have been observed in SIM, which previously escaped from conventional microscopy and were detected only by electron microscopy [36].

The single NPCs were colocalized with channels in the lamin network and peripheral heterochromatin. The distinct NPC components were differentially localized. The double-layered invaginations of the nuclear envelope in prophase were detected. More interestingly, high-speed

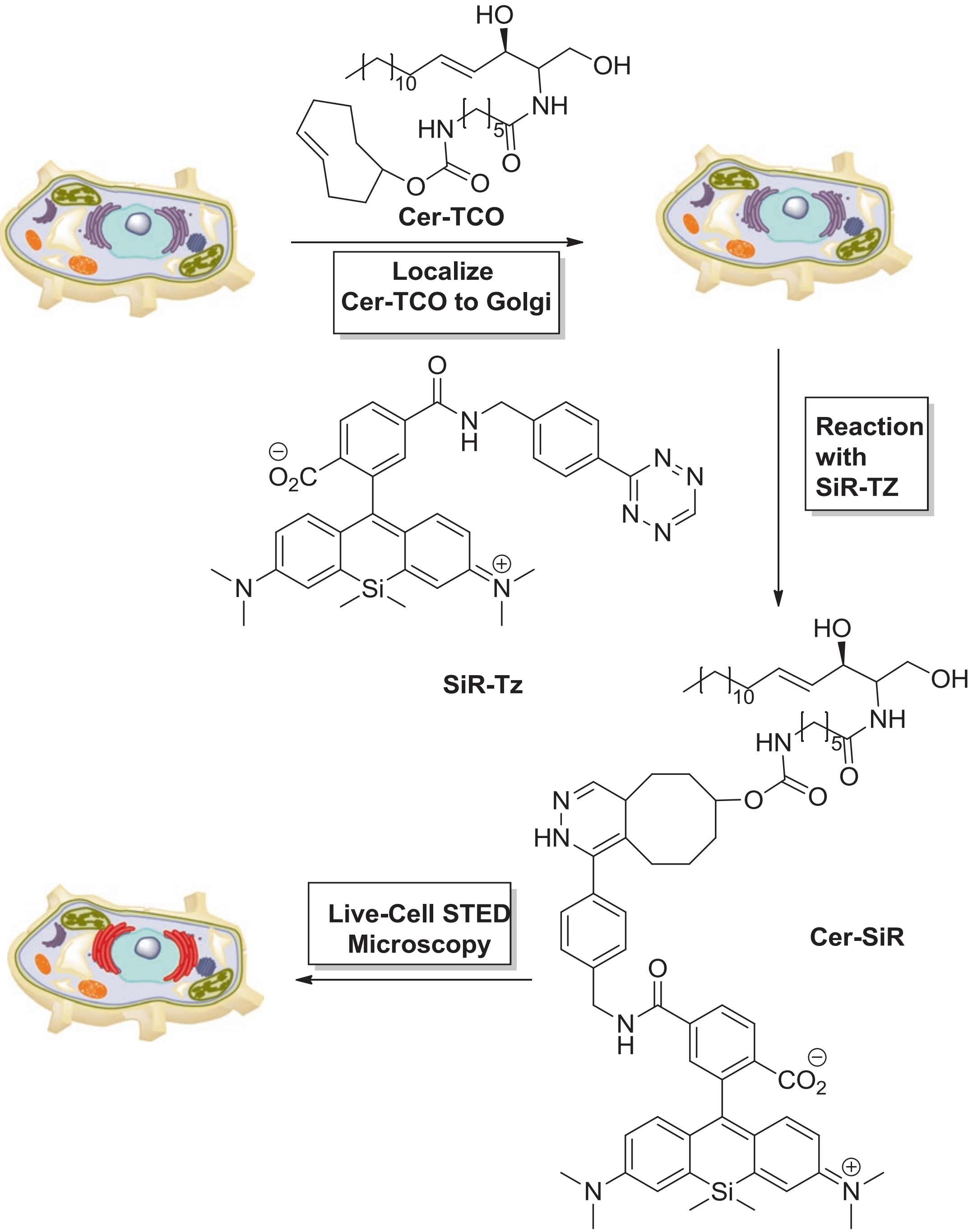


Fig. 6.2 Two-step procedure for high-density labeling of the Golgi in live cells. Cells are treated first with Cer-TCO, a trans-cyclooctenecontaining ceramide lipid, and then reacted with SiR-Tz, a tetrazine derivative of a highly photostable silicon rhodamine dye. The product of this reaction, Cer-SiR, allows extensive live-cell imaging by STED super-resolution microscopy

SIM imaging (11 Hz) has even been developed with 100-nm resolution [39]. With this technique, Kner et al. demonstrated the super-resolution video imaging of tubulin and kinesin dynamics by structured illumination in living Drosophila melanogaster S2 cells [39]. Additionally, SIM still can theoretically improve the resolution with the emission rate of fluorophores nonlinearly depending on the illumination intensity. As a realization of this idea, nonlinear SSIM has achieved 2D resolution of approximately 50-nm on a bead sample [40]. However, the high photostability of fluorescent probes and photo-damage required for SSIM are challenging for live-cells imaging. Fortunately, Rego et al. demonstrated reversible photoswitching of the fluorescent protein Dronpa with the required nonlinearity at six orders of magnitude lower light intensities for saturation [41]. With this fluorescent photoswitchable protein, cellular structures such as mammalian nuclear pore, microtubules, and actin cytoskeleton have been visualized in ~40 nm resolution.

6.2.3 Single-Molecule Localization Microscopy

Another approach to overcome the diffraction restriction is the single-molecule localization microscopy method, including PALM developed by Betzig et al. in 2006 [42], and STORM first developed by Zhuang's group almost at the same time [43]. These approaches rely on stochastic photoswitching of fluorophores between the fluorescent "ON" and "OFF" state. This stochastic photoswitching brings the possibility that only a random subset of activated fluorophores can be sparse enough to be optically resolved at any given time point. In each imaging cycle, a sparse subset of fluorophores were activated and optically resolved, which allows their positions to be accurately determined after algorithm analysis. Ultimately, the super-resolution images are reconstructed by the fusion of position information from thousands of frames. The localization precision of PALM and STORM techniques depends on sufficient photons collected from each activation event, and therefore on the reliability of the fluorophores used.

Although working with the same principle, PALM and STORM techniques have been reported to utilize different photoswitchable fluorophores independently. The fluorophores usually employed in PALM are genetically encoded fluorescent proteins or organic fluorophores that undergo only several cycles of photoconversion or photoactivation before being permanently photobleached. However, in STORM, the fluorescent probes are often reversibly photoswitchable organic fluorophores. In addition, variations of single-molecule localization techniques including fluorescence PALM (FPALM) and direct STORM (dSTORM) have also been embraced in this field [44, 45].

Together with fluorescent proteins such as Kaede and dEosFP, Betzig and his coworkers have successfully used PALM method to super-resolution imaging of intracellular proteins such as the lysosomal transmembrane protein CD63, cytochrome-C oxidase import sequence, vinculin at focal adhesions and actin within a lamellipodium [42]. Recently, Moerner et al. designed a family of photoactivatable push-pull fluorophores, HaloTag-based target-specific azido DCDHFs, to precisely locate cellular proteins in fixed and live single cell imaging [46]. Moreover, the cytoskeletal proteins (Popz, FtsZ, and AmiC) in live bacterial cells have been localized exactly by PALM imaging with these photoactivatable fluorophores. In addition, photoswitchable rhodamine fluorophore and hemicyanine dyes have also undergone extensive research in the field of photoactivated localization microscopy [47–51].

In 2007, Zhuang et al. developed a class of photoswitchable fluorescent probes and demonstrated multicolor STORM imaging [52]. These fluorescent probes are composed of a photoswitchable "reporter" fluorophore (e.g., Cy5) that can be switched between fluorescent "ON" and "OFF" state, and an "activator" (e.g., Alexa405, Cy2, Cy3) that facilitates photoactivation of the reporter. The different combination of reporters and activators generates a family of photoswitchable fluorescent probes for multicolor imaging. Using this approach, three-color STORM imaging of three different DNA constructs labeled with Alexa 405-Cy5, Cy2-Cy5, and Cy3-Cy5 was achieved. In addition, the authors further

extended this method to STORM imaging of microtubules labeled with Cy2-Alexa 647 and clathrin-coated pits (CPPs) labeled with Cy3-Alexa 647 in mammalian cells with 20–30 nm resolution. Furthermore, 3D STORM imaging was obtained by using optical astigmatism to determine the axial and lateral positions of each individual probe, which allows for resolving the three-dimensional morphology of nanoscopic cellular structures [53]. Additionally, Sauer and his coworkers reported a facile strategy for reversible photoswitching of Alexa Fluor and ATTO dyes under physiological conditions [54]. Subsequently, they demonstrated the potential of this method for STORM imaging of cytoskeletal network and mRNA with an approximately 20 nm resolution in fixed and living cells. Recently, a class of SiR fluorophores have been developed for STORM imaging of intracellular proteins in single live-cells [55, 56]. More interesting, a photoluminescence phenomenon termed aggregation-induced emission (AIE) has been employed in super-resolution imaging. As shown in Fig. 6.3, Tang and his co-workers have synthesized and demonstrated a new family of AIE-based bioprobes for super-resolution imaging of subcellular organelles in single cells with STORM [57, 58]. These results facilitate the development of AIE luminogens for super-resolution imaging in more fields.

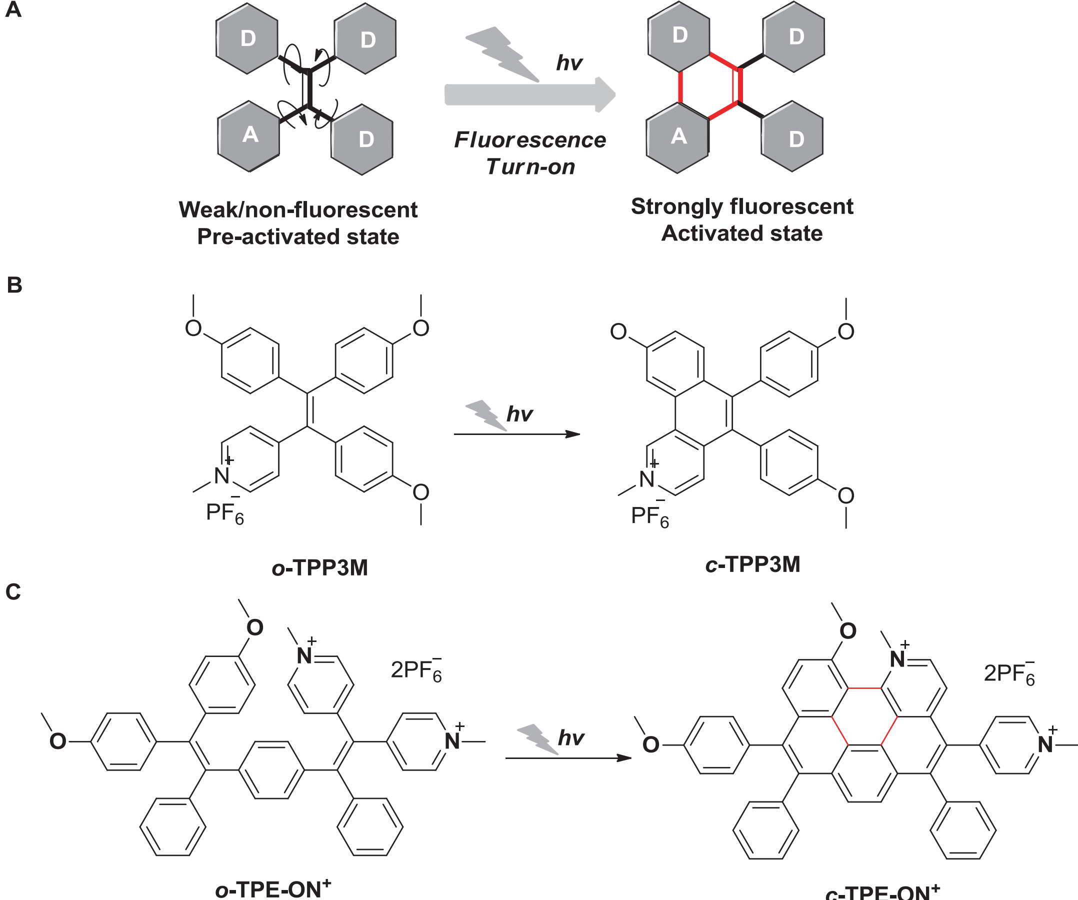


Fig. 6.3 Mitochondrion-specific photoactivatable fluorescence turn-on AIE-based bioprobes for localization super-resolution microscope. (**a**) Principle of design for the photoactivatable turn-on AIE-based probe; (**b**) Photocyclodehydrogenation of ***o*-TPP3M**; (**c**) Photocyclodehydrogenation process of ***o*-TPE-ON⁺**

6.2.4 Other Super-Resolution Microscopy

In 2009, Dertinger and his colleagues developed the super-resolution optical fluctuation imaging (SOFI) that is entirely different from the techniques mentioned above [59]. This approach depends on the higher-order statistical analysis of temporal fluctuations from each single fluorescent molecule recorded in a sequence of images to obtain resolution improvement. SOFI outfits the power of high-order cumulants to create resolution enhancement in all three-dimensions. The advantages of SOFI include the technical simplicity, use of off-the-shelf equipment, genetically encodable labels, simple and rapid data acquisition, and the robustness to significant background rejection and spatiotemporal resolution improvement [60, 61].

Compared with super-resolution techniques mentioned above, SOFI is a relatively young member of the super-resolution technique family. Nevertheless, SOFI has attracted substantial attention worldwide during the past few years for its rational balance in spatial and temporal resolution, imaging depth, and phototoxicity. Dertinger and his colleagues demonstrated super-resolution optical fluctuation imaging of β-tubulin networks labeled with Alexa-647-conjugated antibodies in fixed COS-7 cells under the same experimental conditions as in dSTORM [62]. Xu et al. developed and optimized a reversible switching fluorescent protein, Skylan-S for SOFI [63]. The photostability, contrast ratio, and averaged fluorescence intensity of Skylan-S in the fluctuation state are higher than these of Dronpa [61]. This fluorescent protein provided a 4-fold improvement of fluctuation range of pixels and a higher SOFI resolution. Furthermore, to demonstrate the capability of Skylan-S in SOFI, tubulin structures and clathrin-coated pits (CCPs) were separately labeled with Skylan-S in living U2OS cells and were observed using the SOFI technique. Recently, semiconductor polymer dots (Pdots) with high brightness, extraordinary photostability, and favorable biocompatibility, have been designed as fluorescent probes for SOFI technique (Fig. 6.4) [64].

Imaging of subcellular structures labelled with these small photoblinking Pdots was achieved with improved spatial resolution. Although promising, the feasibility of live-cell imaging is largely restricted by the labeling density of switched-on fluorophores. Xi and his coworkers improved the labeling density of fluorophores using a joint tagging super-resolution technique (JT-SOFI) [65]. To demonstrate the feasibility of JT-SOFI, commercially available QDs (525, QD625, and QD705) were employed to jointly label the tubulin in live COS7 cells, greatly improving their labeling density. After analyzing and combining the images obtained from all QDs, the microtubule networks could be visualized with high fidelity and remarkably enhanced sub-diffraction resolution. In addition, Landes et al. combined SOFI microscopy with fluorescence correlation microscopy (FCS) to develop a new technique termed "fcsSOFI" that allows to reveal nanometer dimensions and diffusion dynamics of cellular cytosol [66].

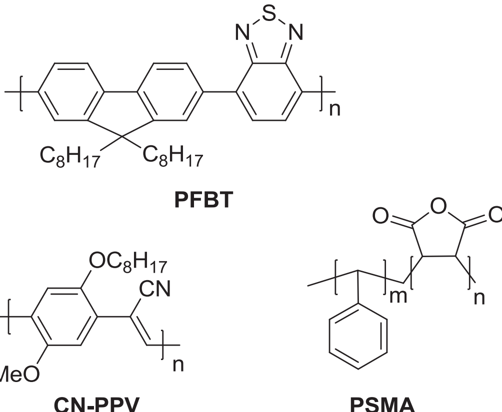


Fig. 6.4 Small photoblinking semiconductor polymer dots for super-resolution optical fluctuation imaging. Chemical structure of semiconductor polymer dots, poly[(9,9- dioctylfluorenyl-2,7-diyl)-co-(1,4-benzo-{2,1′,3}-thiadazole)](PFBT) and poly[2-methoxy-5-(2-ethylhexyloxy)-1,4-(1-cyanovinylene-1,4-phenylene)] (CN-PPV)

6.3 New Developments and Challenges

Although super-resolution approaches have well demonstrated their potentials in imaging previously unresolved details in single cells, they still need further development to become more routinely applicable. The extension of super-resolution microscopy to 3D, multicolor and live-cell imaging reveals the true details inside single cells, overcomes the limitations of light microscopy such as resolution, dimensionality, quantitation and imaging speed, and enables new applications in many biological systems. As illustrated multiple times in many reviews [7–9, 67], this chapter will not dwell on these techniques. However, each approach has its own strengths and weaknesses for single cell imaging [67]. In combination with two or more super-resolution techniques that were performed on the same sample, one can produce results that emphasize the advantages of each technique while offsetting their individual drawbacks. This multimodal approach, known as correlative super-resolution microscopy, adds new dimensions of information and provides new opportunities in this fast-growing field [68].

For example, Betzig et al. first demonstrated the high degree of correlation between TEM and PALM results of FP-tagged proteins on cryoprepared thin sections of fixed cells, helping to validate PALM as an accurate technique for imaging intracellular proteins with sub-diffraction-limit resolution [42]. Watanabe et al. described a correlative fluorescence electron microscopy that combines STED microscopy or PALM with electron microscopy for localization of tagged proteins in electron micrographs [69]. These multimodal approaches have been extended to virtually all super-resolution fluorescence microscopy and electron microscopy methods.

Simultaneously, these advances bring in new challenges such as the robustness, simplicity, and availability of optical instrumentation, fluorescent probes and fluorescent labeling methods, imaging routines and analysis algorithms. One obvious challenge is with fluorescent probes and fluorescent labeling methods that were originally designed for conventional fluorescence microscopy. As just one example, antibody-based fluorescent probes have been widely used in fluorescence microscopy. However, in super-resolution microscopy, there are problematic because of their relatively large size (~15 nm) and insufficient labeling density. In some cases, the clustering artifacts were introduced and easily misinterpreted in the super-resolution images [70]. Therefore, the development of small and bright fluorescent probes and new labeling methods for super-resolution imaging of single cells is highly in demand [71–74].

6.4 Conclusions and Perspectives

Because of the diffraction limit, the resolution of fluorescence microscopy seems to be insurmountable for a long time. However, super-resolution fluorescence microscopy breaks this seemingly impenetrable barrier and resolves complex mechanisms inside biological structures. Over the last 20 years, various super-resolution techniques have been developed, shattered the diffraction barrier and featured an order of magnitude higher resolution compared to that of conventional fluorescence microscopy. More importantly, these super-resolution techniques are beginning to provide a wealth of new insights in the life sciences including nanoscale architectures of cell organelles, organizations and heterogeneities of cellular components, biochemical reactions, cell-to-cell variations at a subcellular level. With the further developments outlined above, there exist a wide variety of optical enhancements of the standard super-resolution fluorescence microscopy that enable multidimensional, quantitative and holistic measurements, significantly improving the precision of single cell imaging. Combining with molecular biology, biochemistry and theoretical bio-computing algorithms, super-resolution fluorescence microscopy continues to expand its capabilities and provide comprehensive insights into the details of single cells that persist only within their native environments.

In the field of life and medicine, there is a gap between the dynamic fate of single molecules and cells and the overall behavior during development. Super-resolution fluorescence microscopy may be promising to bridge this gap by directly accessing molecular structural and functional information in tissue or even *in vivo* in whole organisms. Along with these new possibilities, extensive opportunities and challenges arise. In the future, super-resolution fluorescence microscopy is expected to substantially improve adaptive optics and fluorescent labels for the diagnosis and treatment of human diseases.

Acknowledgments The authors gratefully appreciate the support from the National Natural Science Foundation of China (No.81472835 and 81670091) and the National Key Clinical Specialist Construction Programs of China (2013-544).

Conflicts of Interest The authors declare that they have no conflict of interest.

References

1. Zheng XT, Li CM (2012) Single cell analysis at the nanoscale. Chem Soc Rev 41(6):2061–2071
2. Armbrecht L, Dittrich PS (2017) Recent advances in the analysis of single cells. Anal Chem 89(1):2–21
3. Sousa AA, Leapman RD (2012) Development and application of STEM for the biological sciences. Ultramicroscopy 123:38–49
4. Milne JL, Borgnia MJ, Bartesaghi A, Tran EE, Earl LA, Schauder DM et al (2013) Cryo-electron microscopy-a primer for the non-microscopist. FEBS J 280(1):28–45
5. Stephens DJ, Allan VJ (2003) Light microscopy techniques for live cell imaging. Science 300(5616):82–86
6. Abbe E (1873) Beitrage zur Theorie des mikroskops und der mikroskopischen wahrnehmung. Arch Mikroskop Anat 9:413–420
7. Huang B, Bates M, Zhuang X (2009) Super-resolution fluorescence microscopy. Annu Rev Biochem 78:993–1016
8. Stender AS, Marchuk K, Liu C, Sander S, Meyer MW, Smith EA et al (2013) Single cell optical imaging and spectroscopy. Chem Rev 113(4):2469–2527
9. Fornasiero EF, Opazo F (2015) Super-resolution imaging for cell biologists: concepts, applications, current challenges and developments. BioEssays 37(4):436–451
10. Stone MB, Shelby SA, Veatch SL (2017) Super-resolution microscopy: shedding light on the cellular plasma membrane. Chem Rev 117(11):7457–7477
11. Betzig E (2015) Single molecules, cells, and super-resolution optics (Nobel Lecture). Angew Chem Int Ed 54(28):8034–8053
12. Moerner WE (2015) Single-molecule spectroscopy, imaging, and photocontrol: foundations for super-resolution microscopy (Nobel Lecture). Angew Chem Int Ed 54(28):8067–8093
13. Hell SW (2015) Nanoscopy with focused light (Nobel Lecture). Angew Chem Int Ed 54(28):8054–8066
14. Hell SW, Wichmann J (1994) Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. Opt Lett 19(11):780–782
15. Klar TA, Hell SW (1999) Subdiffraction resolution in far-field fluorescence microscopy. Opt Lett 24(14):954–956
16. Willig KI, Rizzoli SO, Westphal V, Jahn R, Hell SW (2006) STED microscopy reveals that synaptotagmin remains clustered after synaptic vesicle exocytosis. Nature 440(7086):935–939
17. Kittel RJ, Wichmann C, Rasse TM, Fouquet W, Schmidt M, Schmid A et al (2006) Bruchpilot promotes active zone assembly, Ca^{2+} channel clustering, and vesicle release. Science 312(5776):1051–1054
18. Sieber JJ, Willig KI, Kutzner C, Gerding-Reimers C, Harke B, Donnert G et al (2007) Anatomy and dynamics of a supramolecular membrane protein cluster. Science 317(5841):1072–1076
19. Donnert G, Keller J, Medda R, Andrei MA, Rizzoli SO, Lührmann R et al (2006) Macromolecular-scale resolution in biological fluorescence microscopy. Proc Natl Acad Sci U S A 103(31):11440–11445
20. Kellner RR, Baier CJ, Willig KI, Hell SW, Barrantes FJ (2007) Nanoscale organization of nicotinic acetylcholine receptors revealed by stimulated emission depletion microscopy. Neuroscience 144(1):135–143
21. Lin W, Margolskee R, Donnert G, Hell SW, Restrepo D (2007) Olfactory neurons expressing transient receptor potential channel M5 (TRPM5) are involved in sensing semiochemicals. Proc Natl Acad Sci U S A 104(7):2471–2476
22. Schneider A, Rajendran L, Honsho M, Gralle M, Donnert G, Wouters F et al (2008) Flotillin-dependent clustering of the amyloid precursor protein regulates its endocytosis and amyloidogenic processing in neurons. J Neurosci 28(11):2874–2882
23. Westphal V, Rizzoli SO, Lauterbach MA, Kamin D, Jahn R, Hell SW (2008) Video-rate far-field optical nanoscopy dissects synaptic vesicle movement. Science 320(5873):246–249
24. Lukinavičius G, Reymond L, D'Este E, Masharina A, Göttfert F, Ta H et al (2014) Fluorogenic probes for live-cell imaging of the cytoskeleton. Nat Methods 11(7):731–733
25. Honigmann A, Mueller V, Ta H, Schoenle A, Sezgin E, Hell SW et al (2014) Scanning STED-FCS reveals

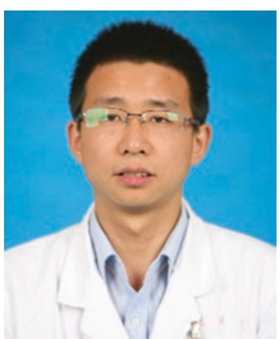
spatiotemporal heterogeneity of lipid interaction in the plasma membrane of living cells. Nat Commun 5:5412
26. D'Este E, Kamin D, Göttfert F, El-Hady A, Hell SW (2015) STED nanoscopy reveals the ubiquity of subcortical cytoskeleton periodicity in living neurons. Cell Rep 10(8):1246–1251
27. Liu Y, Lu Y, Yang X, Zheng X, Wen S, Wang F et al (2017) Amplified stimulated emission in upconversion nanoparticles for super-resolution nanoscopy. Nature 543(7644):229–233
28. Erdmann RS, Takakura H, Thompson AD, Rivera-Molina F, Allgeyer ES, Bewersdorf J et al (2014) Super-resolution imaging of the Golgi in live cells with a bioorthogonal ceramide probe. Angew Chem Int Ed 53(38):10242–10246
29. Hanne J, Falk HJ, Görlitz F, Hoyer P, Engelhardt J, Sahl SJ et al (2015) STED nanoscopy with fluorescent quantum dots. Nat Commun 6:7127
30. Gustafsson MG (2000) Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. J Microsc 198(Pt 2):82–87
31. Heintzmann R, Cremer CG (1999) Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating. Proc SPIE 3568:185–196
32. Gustafsson MG, Shao L, Carlton PM, Wang CJ, Golubovskaya IN, Cande WZ et al (2008) Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination. Biophys J 94(12):4957–4970
33. Heintzmann R, Jovin TM, Cremer C (2002) Saturated patterned excitation microscopy-a concept for optical resolution improvement. J Opt Soc Am A Opt Image Sci Vis 19(8):1599–1609
34. York AG, Parekh SH, Dalle Nogare D, Fischer RS, Temprine K, Mione M et al (2012) Resolution doubling in live, multicellular organisms *via* multifocal structured illumination microscopy. Nat Methods 9(7):749–754
35. Olshausen PV, Defeu Soufo HJ, Wicker K, Heintzmann R, Graumann PL, Rohrbach A (2013) Super-resolution imaging of dynamic MreB filaments in B. subtilis-a multiple-motor-driven transport? Biophys J 105(5):1171–1181
36. Schermelleh L, Carlton PM, Haase S, Shao L, Winoto L, Kner P et al (2008) Subdiffraction multicolor imaging of the nuclear periphery with 3D structured illumination microscopy. Science 320(5881):1332–1336
37. Shao L, Kner P, Rego EH, Gustafsson MG (2011) Super-resolution 3D microscopy of live whole cells using structured illumination. Nat Methods 8(12):1044–1046
38. Fiolka R, Shao L, Rego EH, Davidson MW, Gustafsson MG (2012) Time-lapse two-color 3D imaging of live cells with doubled resolution using structured illumination. Proc Natl Acad Sci U S A 109(14):5311–5315
39. Kner P, Chhun BB, Griffis ER, Winoto L, Gustafsson MG (2009) Super-resolution video microscopy of live cells by structured illumination. Nat Methods 6(5):339–342
40. Gustafsson MG (2005) Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution. Proc Natl Acad Sci U S A 102(37):13081–13086
41. Rego EH, Shao L, Macklin JJ, Winoto L, Johansson GA, Kamps-Hughes N et al (2012) Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution. Proc Natl Acad Sci U S A 109(3):135–143
42. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS et al (2006) Imaging intracellular fluorescent proteins at nanometer resolution. Science 313(5793):1642–1645
43. Rust MJ, Bates M, Zhuang X (2006) Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nat Methods 3(10):793–795
44. Hess ST, Girirajan TP, Mason MD (2006) Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. Biophys J 91(11):4258–4272
45. Heilemann M, van de Linde S, Schüttpelz M, Kasper R, Seefeldt B, Mukherjee A et al (2008) Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. Angew Chem Int Ed 47(33):6172–6176
46. Lee HL, Lord SJ, Iwanaga S, Zhan K, Xie H, Williams JC et al (2010) Super-resolution imaging of targeted proteins in fixed and living cells using photoactivatable organic fluorophores. J Am Chem Soc 132(43):15099–15101
47. Fölling J, Belov V, Kunetsky R, Medda R, Schönle A, Egner A et al (2007) Photochromic rhodamines provide nanoscopy with optical sectioning. Angew Chem Int Ed 46(33):6266–6270
48. Lee MK, Rai P, Williams J, Twieg RJ, Moerner WE (2014) Small-molecule labeling of live cell surfaces for three-dimensional super-resolution microscopy. J Am Chem Soc 136(40):14003–14006
49. Grimm JB, Sung AJ, Legant WR, Hulamm P, Matlosz SM, Betzig E et al (2013) Carbofluoresceins and carborhodamines as scaffolds for high-contrast fluorogenic probes. ACS Chem Biol 8(6):1303–1310
50. Tian Z, Li AD, Hu D (2011) Super-resolution fluorescence nanoscopy applied to imaging core-shell photoswitching nanoparticles and their self-assemblies. Chem Commun 47(4):1258–1260
51. Zhang H, Wang C, Jiang T, Guo H, Wang G, Cai X et al (2015) Microtubule-targetable fluorescent probe: site-specific detection and super-resolution imaging of ultratrace tubulin in microtubules of living cancer cells. Anal Chem 87(10):5216–5222
52. Bates M, Huang B, Dempsey GT, Zhuang X (2007) Multicolor super-resolution imaging with photo-switchable fluorescent probes. Science 317(5845):1749–1753
53. Huang B, Wang W, Bates M, Zhuang X (2008) Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy. Science 319(5864):810–813
54. Heilemann M, van de Linde S, Mukherjee A, Sauer M (2009) Super-resolution imaging with small organic fluorophores. Angew Chem Int Ed 48(37):6903–6908

55. Lukinavičius G, Umezawa K, Olivier N, Honigmann A, Yang G, Plass T et al (2013) A near-infrared fluorophore for live-cell super-resolution microscopy of cellular proteins. Nat Chem 5(2):132–139
56. Uno SN, Kamiya M, Yoshihara T, Sugawara K, Okabe K, Tarhan MC et al (2014) A spontaneously blinking fluorophore based on intramolecular spirocyclization for live-cell super-resolution imaging. Nat Chem 6(8):681–689
57. Gu X, Zhao E, Lam JW, Peng Q, Xie Y, Zhang Y et al (2015) Mitochondrion-specific live-cell bioprobe operated in a fluorescence turn-on manner and a well-designed photoactivatable mechanism. Adv Mater 27(44):7093–7100
58. Gu X, Zhao E, Zhao T, Kang M, Gui C, Lam JW et al (2016) A mitochondrion-specific photoactivatable fluorescence turn-pn AIE-based bioprobe for localization super-resolution microscope. Adv Mater 28(25):5064–5071
59. Dertinger T, Colyer R, Iyer G, Weiss S, Enderlein J (2009) Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI). Proc Natl Acad Sci U S A 106(52):22287–22292
60. Dertinger T, Colyer R, Vogel R, Enderlein J, Weiss S (2010) Achieving increased resolution and more pixels with super-resolution optical fluctuation imaging (SOFI). Opt Express 18(18):18875–18885
61. Dedecker P, Mo GC, Dertinger T, Zhang J (2012) Widely accessible method for super-resolution fluorescence imaging of living systems. Proc Natl Acad Sci U S A 109(27):10909–10914
62. Dertinger T, Heilemann M, Vogel R, Sauer M, Weiss S (2010) Super-resolution optical fluctuation imaging with organic dyes. Angew Chem Int Ed 49(49):9441–9443
63. Zhang X, Chen X, Zeng Z, Zhang M, Sun Y, Xi P et al (2015) Development of a reversibly switchable fluorescent protein for super-resolution optical fluctuation imaging (SOFI). ACS Nano 9(3):2659–2667
64. Chen X, Li R, Liu Z, Sun K, Sun Z, Chen D et al (2017) Small photoblinking semiconductor polymer dots for fluorescence nanoscopy. Adv Mater. https://doi.org/10.1002/adma.201604850
65. Zeng Z, Chen X, Wang H, Huang N, Shan C, Zhang H et al (2015) Fast super-resolution imaging with ultra-high labeling density achieved by joint tagging super-resolution optical fluctuation imaging. Sci Rep 5:8359
66. Kisley L, Brunetti R, Tauzin LJ, Shuang B, Yi X, Kirkeminde AW et al (2015) Characterization of porous materials by fluorescence correlation spectroscopy super-resolution optical fluctuation imaging. ACS Nano 9(9):9158–9166
67. Godin AG, Lounis B, Cognet L (2014) Super-resolution microscopy approaches for live cell imaging. Biophys J 107(8):1777–1784
68. Hauser M, Wojcik M, Kim D, Mahmoudi M, Li W, Xu K (2017) Correlative super-resolution microscopy: new dimensions and new opportunities. Chem Rev 117(11):7428–7456
69. Watanabe S, Punge A, Hollopeter G, Willig KI, Hobson RJ, Davis MW et al (2011) Protein localization in electron micrographs using fluorescence nanoscopy. Nat Methods 8(1):80–84
70. Ries J, Kaplan C, Platonova E, Eghlidi H, Ewers H (2012) A simple, versatile method for GFP-based super-resolution microscopy via nanobodies. Nat Methods 9(6):582–584
71. Opazo F, Levy M, Byrom M, Schäfer C, Geisler C, Groemer TW et al (2012) Aptamers as potential tools for super-resolution microscopy. Nat Methods 9(10):938–939
72. Deschout H, Cella Zanacchi F, Mlodzianoski M, Diaspro A, Bewersdorf J, Hess ST et al (2014) Precisely and accurately localizing single emitters in fluorescence microscopy. Nat Methods 11(3):253–266
73. Klein T, Löschberger A, Proppert S, Wolter S, van de Linde S, Sauer M (2012) Live-cell dSTORM with SNAP-tag fusion proteins. Nat Methods 8(1):7–9
74. Wombacher R, Heidbreder M, van de Linde S, Sheetz MP, Heilemann M, Cornish VW et al (2010) Live-cell super-resolution imaging with trimethoprim conjugates. Nat Methods 7(9):717–719

Han Feng was born in Shangqiu, China, in 1990. She received her MS degree in the Institute of Clinical

Pharmacology, Hunan Key Laboratory of Pharmacogenetics, Central South University, under the supervision of Prof. Dong-Sheng Ouyang (2015). In July 2015, she joined the Department of Pharmacy, Henan Province People's Hospital, and People's Hospital of Zhengzhou University. She has a range of research interests at clinical pharmacology and pharmacokinetics, and pharmacogenetics.

Xiaobo Wang achieved his MS degree in 2015, in the School of Pharmaceutical Science, Central South University, under the supervision of Prof. Wenbin Zeng. In July 2015, he joined the Department of Nuclear Medicine, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University as an assistant research fellow. His research interests include the development of small fluorescence probes and multifunctional nanoprobes and imaging agents based on nanomaterials.



Zhiwei Xu obtained his MS degree in Epidemiology and Health Statistics from Wuhan University (2012). Since then he joined Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University. In 2017, he relocated to Clinical Research Service Center. His research interest is epidemiology and medical statistics.

Xiaoju Zhang received her M.D. in Respiratory Medicine from Huazhong University of Science and Technology, China (2005). In 2008, she joined Department of Pulmonary Medicine, Henan Province People's Hospital, and People's Hospital of Zhengzhou University as an associate professor. After a postdoctoral training in University of California in 2009–2012, she was promoted as the director of Department of Pulmonary Medicine. Her current research interests include molecular mechanism and clinical translational research of lung cancer, early diagnosis of pulmonary nodules.

Yongju Gao received her M.D. in Imaging and Nuclear Medicine from Shandong University, China (2000). She is a visiting scholar of Johns Hopkins University in 2006–2007 and the director of Department of Nuclear Medicine, Henan Provincial People's Hospital, and People's Hospital of Zhengzhou University. She has a broad range of research interests at nuclear medicine and molecular imaging.

Check for updates

7 Single Cell Proteomics for Molecular Targets in Lung Cancer: High-Dimensional Data Acquisition and Analysis

Zheng Wang and Xiaoju Zhang

Abstract

In the proteomic and genomic era, lung cancer researchers are increasingly under challenge with traditional protein analyzing tools. High output, multiplexed analytical procedures are in demand for disclosing the post-translational modification, molecular interactions and signaling pathways of proteins precisely, specifically, dynamically and systematically, as well as for identifying novel proteins and their functions. This could be better realized by single-cell proteomic methods than conventional proteomic methods. Using single-cell proteomic tools including flow cytometry, mass cytometry, microfluidics and chip technologies, chemical cytometry, single-cell western blotting, the quantity and functions of proteins are analyzed simultaneously. Aside from deciphering disease mechanisms, single-cell proteomic techniques facilitate the identification and screening of biomarkers, molecular targets and promising compounds as well. This review summarized single-cell proteomic tools and their use in lung cancer.

Keywords

Lung cancer · Mass spectrometry · Single nucleotide polymorphism · Biomarker

7.1 Introduction

Reportedly, lung carcinoma is the leading cause of cancer-related mortality worldwide, which accounts for 27% of all cancer-related deaths in men and 25% in women in the United States of America [1]. The two primary types of lung cancer are small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Apparently, chemotherapeutic agents and radiotherapy techniques have undergone a significant progress in the last two decades. Novel chemotherapeutic agents were developed with higher efficiency and lower adverse effects, and which greatly improves the tolerance and survival in lung cancer patients. In addition, molecular targeted therapy has attracted considerable attention to enhance the efficacy of the lung cancer treatment, and has evolved as a milestone for lung cancer treatment. Through molecular target screening and selection, the overall survival (OS) and progression-free survival (PFS0 overwhelmingly increases with epidermal growth factor receptor-tyrosine

Z. Wang
Department of Respiratory and Critical Care Medicine, Zhengzhou University People's Hospital, Zhengzhou, China

X. Zhang (✉)
Department of Respiratory and Critical Care Medicine, Zhengzhou University People's Hospital, Zhengzhou, China

Biomedical Research Center, Zhengzhou University People's Hospital, Zhengzhou, China

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_7

kinase inhibitor (EGFR-TKIs) treatment compared with chemotherapy in selected unresectable NSCLC patients, supporting the latter as the first-line treatment regime for them [2].

Nonetheless, these advancements are insufficient for both reducing the mortality of NSCLC and effectively treat SCLC and some types of NSCLC. Despite all that medical science has to offer, the overall 5-year survival rate of all patients diagnosed with lung cancer remains approximately 15%~20%, with less than 7% of all patients surviving for 10 years [1]. Development of biomarkers, treatment targets and agents is thus urged for reducing the mortality of NSCLC or to effectively treating SCLC and some types of NSCLC [2, 3]. The complexity of molecular network in the pathogenesis of lung cancer awaits for further elucidations as well.

Proteins are involved in critical biological processes in lung cancer, including oncogenesis, proliferation, angiogenesis and metastasis. The complexity of deciphering the functioning proteins lies in their large number, previously unidentified origin, post-translational modifications and transformations, as well as protein-protein interactions that build networks. Hence, traditional protein analyzing methods, such as two-dimensional electrophoresis (2DE), are no longer inadequate for assessing the interaction networks of proteins in lung cancer. Conversely, proteomics offers a myriad of analytical procedures that when used in a large-scale might help in revealing the post-translational modification and biological activities of proteins dynamically and systematically [4, 5]. Using proteomic methods, protein abundance could be quantified or semi-quantified in lung or lung cancer tissues, which allows further identification of molecular interactions and signaling pathways in lung cancer [6]. In addition, protein biomarkers could be identified, protein biomarker panels could be verified, facilitating the validation of further clinical trials or translational studies [7]. Most importantly, novel molecular targets and promising compounds are identified with proteomic methods that benefit disease therapies in lung cancer [8].

7.2 The Usage of Proteomics in Lung Cancer

The usage of proteomics in lung cancer has been explored in several different aspects [9–11]. According to a preliminary study, histology, metastases and nodal involvement could be identified in frozen lung cancer sections with high accuracy using matrix-assisted laser desorption/ionization mass spectrometry [12]. Using an aptamer-based proteomic method, serum proteins are identified as markers that discriminate NSCLC from control subjects [13], or identify NSCLC from pulmonary nodules [14]. Liquid chromatography-tandem mass spectrometry (LC-MS/MS) was used to characterize the proteome profile of lung adenocarcinomas that is distinguishable from that of benign nodules or of normal lung tissues [15]. Using isotope labeling technique and high resolution mass spectrometer, the quantitative phosphoproteomic analysis revealed the phosphorylation status of multiple kinases, including crucial kinases controlling PI3K/AKT/mTOR and CDC42-PAK signaling pathways, are identified in human lung mucoepidermoid carcinoma cells chronically exposed to cigarette smoke [16].

A recent study revealed the differential value of protein profiles in exhaled breath condensate (EBC) samples of healthy subjects, healthy smokers, chronic obstructive pulmonary disease (COPD) and lung cancer patients, which offers a potential non-invasive diagnostic method for lung cancer [17]. The abundance of several extracellular matrix (ECM) proteins, including fibronectin and tenascin-C (Tnc), are significantly increased in primary lung tumors and associated lymph node metastases compared with normal tissue. A three-gene expression signature comprising TNC, S100A10, and S100A11 is strongly predictive of patient survival independent of age, sex, smoking history, and mutational load [18]. A signature of 25 mass spectrometry signals was associated with both relapse-free survival and overall survival (OS) in NSCLC patients using frozen resected tissue specimens [19]. An eight-peak profile signa-

ture was established in NSCLC patients that predict their responses to EGFR-tyrosine kinase inhibitor (TKI) treatment, as assessed by progression and overall survival [20]. VeriStrat, a MALDI mass spectrometry method, was evolved later on for prediction of the efficacy of monotherapy or combination therapy in NSCLC patients [21].

Basing on the proteomic profile, patients could be classified as either poorer outcome or better outcome by VeriStrat system, which is in accordance with the disease progression. Except for EGFR-TKI, VeriStrat system was also tried for assessing the efficacy of EGFR-TKI plus bevacizumab (an pan-EGFR monoantibody). Another study tried matrix-assisted laser desorption/ionization MS proteomic algorithm developed from a small dataset of erlotinib-bevacizumab treated patients to predict the clinical outcome of patients treated with erlotinib alone [22]. The prognostic value of VeriStrat was also demonstrated for treatment-naïve non-squamous advanced NSCLC patients treated with first-line chemotherapy, or erlotinib plus first-line chemotherapy [23, 24]. Prediction models have also been developed for etoposide chemoresistance using global proteomic profiling in NSCLC [25]. Both therapeutic efficacy and severe adverse effects are predictable by serum peptide profiling with MALDI-TOF mass spectrometry in NSCLC patients [26]. Multivariate protein predictive models have been developed for gefitinib-related interstitial lung diseases in pharmacoproteomic studies [27]. Recently, the value of VeriStrat system in treatment strategy selection was investigated in several clinical trials. Although it seems too early to make the conclusion for inconsistent clinical trial results, this proteomic method do show its potentials.

Beyond being a prognostic biomarker, proteomics is potentiated as a candidate for precise medicine in lung cancer [28, 29]. Proteomic methods are also used for drug development in lung cancer. OSU03013, a derivative of celecoxib, was found to induce cell cycle G1 phase arrest, apoptosis and expressive changes in proteomics in lung cancer cells, as revealed by 2DE, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) and MALDI-TOF-tandem mass spectrometry (MS/MS) [30]. The antitumor activity of Microcystis viridis lectin (MVL), a mannose-binding protein, has been tested in A549 lung cancer cells by flow cytometry (for apoptosis), 2DE and MALDI-TOF-MS [31]. Fifty protein spots were detected by 2DE, which are significantly differentially expressed protein spots after MVL treatment. Seven proteins were further identified by MALDI-TOF-MS In MVL-treated A549 cells, the two increased proteins were aldehyde dehydrogenase 1 family member A1 (ALDH1A1) and β-actin, and five decreased proteins include: heat shock protein 90 kDa beta member 1 (HSP90B1), heat shock 60 kDa protein 1 (HSPD1), plastin 3, tropomyosin 3 isoform 2 (TPM3), and β-tubulin (Figs. 7.1, 7.2, and 7.3).

Based on PANTHER classification system, the identified 7 proteins can be classified into 3 functional categories: protein folding (HSP90, HSP60), cell structure and motility (ACTB, PLS3, TUBB and TPM3), and other carbon metabolism (ALDH1A1). Successive interaction network analysis showed that four of the seven identified proteins proteins (HSP90, HSP60, β-tubulin, β-actin) are the direct interactors of 14-3-3 zeta, suggesting that 14-3-3 zeta may play a vital role in mediating MVL-induced apoptosis in lung cancer cells. Furthermore, a systemic review validated the usage of the proteomic platform in solid tumors, including lung cancer [32]. In recent years, proteomics studies in lung cancer have been emerging and reviewed elsewhere [33, 34]. Single-cell proteomics comprises fluorescence flow cytometry/fluorescence-activated cell sorting (FFC/FACS), mass cytometry, microfluidic chip, and chemical cytometry and has drawn overall research focus lately. In fact, the usage of single-cell proteomics has just begun and provoked research interests in lung cancer [35–40].

7.3 Single Cell Proteomic Methods

7.3.1 Flow Cytometry

For the last four decades, the flow cytometry usage has witnessed an evolution. Because

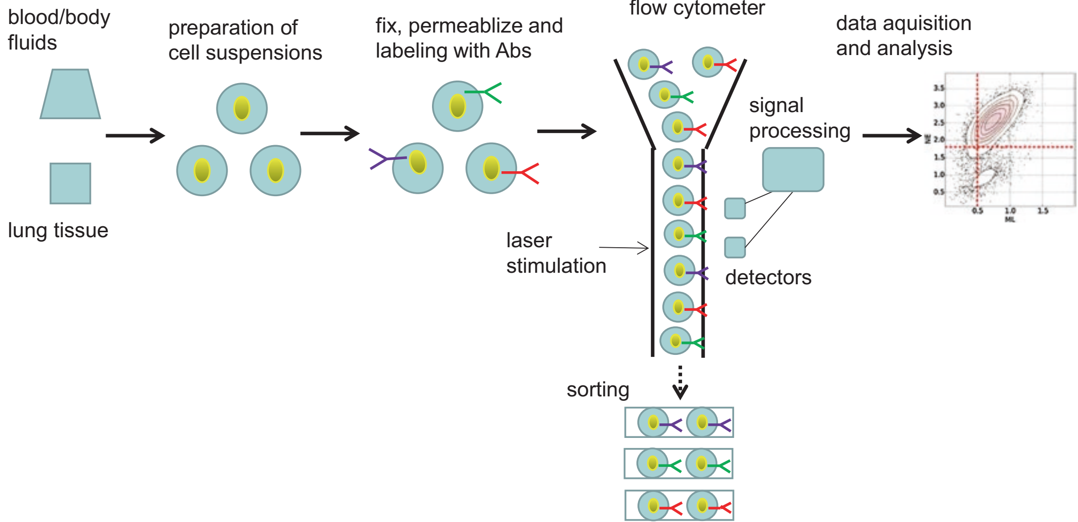


Fig. 7.1 Flow cytometry for single cell analysis. Cells are labeled by antibodies, and then the cell suspension is entrained in the center of a narrow, rapidly flowing stream of liquid. An electrical charging ring is placed just at the point where the stream breaks into droplets. A charge is placed on the ring based immediately prior to fluorescence intensity being measured, and the opposite charge is trapped on the droplet as it breaks from the stream. The charged droplets then fall through an electrostatic deflection system that diverts droplets into containers based upon their charge. In some systems, the charge is applied directly to the stream, and the droplet breaking off retains charge of the same sign as the stream. The stream is then returned to neutral after the droplet breaks off

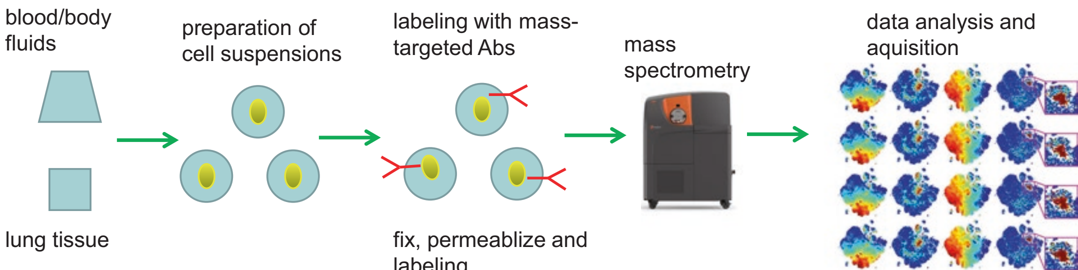


Fig. 7.2 The workflow of a typical mass cytometry analysis. Cells are first stained with elemental isotope-conjugated antibodies. After nebulization, atomization, and ionization, clouds of ions originating from single cells are filtered by the quadrupole and sent for time-of-flight mass analysis. Finally, high-dimensional datasets are obtained and processed to extract underlying biological information

conventional flow cytometry comprises promptly staining an antibody that targets a previously identified antigen, the identification of novel, post-translational modified proteins that might be involved in the pathobiology of diseases becomes challenging. Reportedly, multicolor FFC or FACS overcomes this problem [41]. In FFC, labeling of cellular proteins with fluorescent antibodies facilitates simultaneous assessments of up to 15 molecular markers per cell. By using spectrally distinct fluorophores and highly specialized equipment, this number could be further boosted. Intracellular staining (ICS) [42], which requires blocking protein secretion and cell fixation, can be coupled with cytometry to investigate the relative levels of functional proteins such as cytokines and phosphokinases.

One of the most prominent advantages of the cytometry methods (particularly FFC and FACS) is that large amounts of cells could be identified

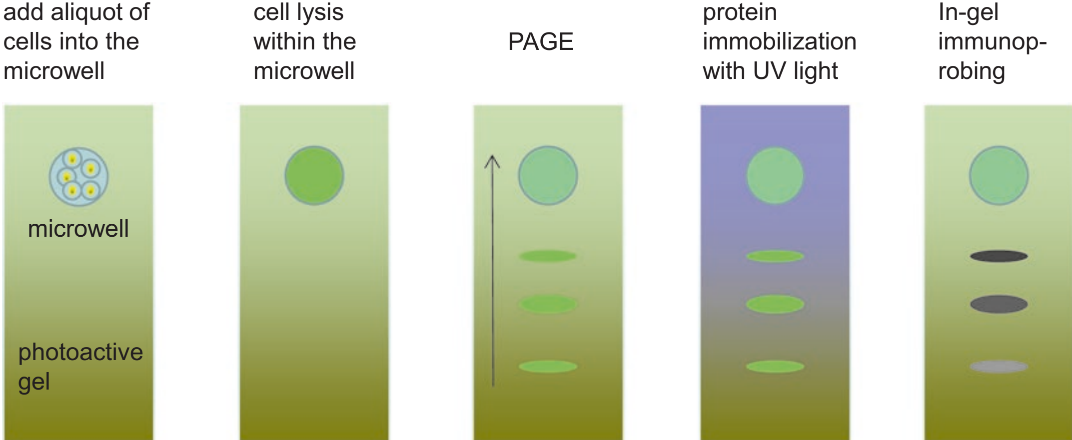


Fig. 7.3 Principles of single-cell western blotting. Firstly, cells are filled with the microwells. Then is chemical lysis of cells in each microwell. Then is polyacrylamide gel electrophoresis (PAGE) of each single-cell lysate. Optimization of PAGE depends on the molecular mass of targets, sample preparation, PAGE duration, PAG pore size and applied electric field strength. Then is the exposure of the gel to UV light to blot (immobilization) proteins to the gel matrix, which is also the immobilization of cells, and then is in-gel immunoprobing of immobilized proteins

at the same time, which enables practitioners to identify, sort and enumerate relatively rare cell types. The samples may be obtained from blood or tissues. The phenotyping and differentiation of a cell type is realized by FFC or FACS by its specific cellular markers. This is easily performed in hematological malignancies. Solid cancer cells are also distinguishable with cellular markers. Leelatian et al. characterized glioma, melanoma and small cell lung cancer cells with flow cytometry and mass cytometry to track cell subset abundance following different enzyme combinations and treatment times [43]. The results show that cell identity is well established by several key cellular markers, including CD45, CD3, CD4, CD8, CD19, CD64, HLA-DR, CD11c, CD56, CD44, GFAP, S100B, SOX2, nestin, vimentin, cytokeratin, and CD31. The authors also show that fluorescence cytometry and mass cytometry identifies comparable frequencies of cancer cell subsets, leukocytes, and endothelial cells in glioma and tonsil. These results provide a possible procedure for preparing viable single cell suspensions that preserve the cellular diversity of human tissue microenvironments. Another important study by Lin et al. explored the possible utility of single cell phospho-specific flow cytometry (SCPFC) in the investigation of signaling network interactions and dynamic changes of Tyrosine phospho-Stat1 (pStat1) in lung cancer cells [44]. The fluorescence intensity changes of pStat1 after IFN-γ stimulation were compatible to results obtained by Western blot analysis.

In metastatic animal models, cancer cells from subcutaneous tumors, malignant ascites, and peritoneal tumors responded to IFN-γ. Moreover, the association between cisplatin resistance and molecular characteristics was identified using SCPFC in lung cancer cells collected from malignant pleural effusions (MPE). MPE cancer cells with higher pStat1 changes after IFN-γ stimulation were more resistant to cisplatin. In all, this study provides the possible application of SCPFC in functional characterization of lung cancer cells and drug sensitivity analysis. Other studies revealed the usage of FFC/FACS in prognosis and disease monitoring of leukemia [45, 46], but this has not been tested in lung cancer yet. FFC/FACS has been used for isolation of circulating tumor cells (CTC) in solid organ tumors, which represents an important base for liquid biopsy [47]. FFC/FACS system could also be used in combination with other proteomic methods, being as the first step of target cell sorting in solid

organ tumors. A study isolated endothelial cells (ECs) from normal or cancer tissues by FACS using CD146 as a marker. Tissues were cultured ex vivo and then digested with trypsin and subjected to 2DE, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) and MALDI-TOF-tandem mass spectrometry (MS/MS) [48]. Differentially expressed peptides were quantified, verified using immunohistochemistry (IHC), and analyzed with in vitro siRNA functioning studies.

7.3.2 Mass Cytometry

Mass cytometry (MS) uses antibodies that are tagged with mass labels, which substantially provides higher level of multiplexing rather than fluorophore labels. In mass cytometry rare earth metals replace the fluorophores and coupled plasma mass spectroscopy replaces optical signal detection of conventional flow cytometry. Mass cytometry is capable of analyzing more than 50 cellular parameters in single cells, including markers for specific proteins, cell viability and DNA content [49]. Laser ablation electrospray ionization mass spectrometry (LAESI-MS) has also been used for metabolic analysis of small molecular metabolites [50].

Time-of-flight (TOF) mass spectrometry was developed for real time analysis of individual biological cells or other microparticles [51]. The instrument is based on inductively coupled plasma time-of-flight mass spectrometry and comprises a three-aperture plasma-vacuum interface, a dc quadrupole turning optics for decoupling ions from neutral components, an rf quadrupole ion guide discriminating against low-mass dominant plasma ions, a point-to-parallel focusing dc quadrupole doublet, an orthogonal acceleration reflection analyzer, a discrete dynode fast ion detector, and an 8-bit 1 GHz digitizer. A high spectrum generation frequency of 76.8 kHz provides capability for collecting multiple spectra from each particle-induced transient ion cloud, typically of 200–300 micros duration. This technique is sufficient for elemental immunoassay with up to 60 distinct available elemental tags.

Application of mass cytometry in cancer was firstly utilized for elucidating single cell biology of hematopoietic malignancies. Bone marrow samples obtained from leukemia patients as well as healthy individuals are characterized at diagnosis and relapse, basing on cell surface markers. Recently, a multi-national research group published the results of a clinical trial that use single cell mass cytometry to evaluate the treatment response to nilotinib, as well as sorting immune cells in CML [52]. Peripheral blood and bone marrow samples were collected from healthy individuals. As for CML patients, peripheral blood sample was collected before the first nilotinib dose, and 3 h, seven days and 28 days after the start of treatment. Mass cytometry was performed in mass cytometer devices using commercially-available kits. The results showed that single cell mass cytometry identifies the immunophenotype and signal transduction of cells, indicates the expressions and functioning status of proteins including phosphorylated Bcr-Abl, and detects hematologic remission in clinical trial samples. Cancer cells and their functional status in leukemia have been characterized by investigating signaling aberrations within signaling pathways, and shows the efficacy in identifying cell types and predicting treatment responses and prognosis [53].

Recently, mass cytometry and flow cytometry are integrated in the monitoring and investigation of CML. The complexity of the generated data requires advanced bioinformatics analyses, thereby complicating the introduction of mass cytometry in clinical settings for evaluating of prognosis and response to therapy [52, 54]. The combination of mass cytometry with highly multiplexed immunohistochemistry and laser ablation enables the researcher to study cells in the relevant context of tissue architecture. In this analysis the crucial spatial information on cells and their interactions is preserved [55]. The technology will enable the analysis of archived, fixed and paraffin-embedded tissues in an unprecedented high dimensional space that could lead to new discoveries in health and disease. As mentioned above, Leelatian et al. showed that fluorescence cytometry and mass cytometry

identifies comparable frequencies of cancer cell subsets, leukocytes, and endothelial cells in glioma and tonsil [43].

More recently, Rahman et al. reported their use of single cell mass cytometry in identifying the cellular markers of lung cells within active smokers or lung cancer cells [56]. They performed a CyTOF analysis on cell suspensions derived from matched blood, tumor lesion, and non-involved lung tissue from an active smoker undergoing surgical resection for early stage lung adenocarcinoma. The samples were stained with a 31-parameter antibody panel to allow a detailed characterization of the cellular heterogeneity of the samples. The data were visualized using viSNE, major immune subsets were identified based on canonical marker expression patterns, and single cell cerium levels were evaluated across each of these defined subsets. The results show that high levels of cerium were specifically associated with a phenotypically distinct subset of lung macrophages that were most prevalent in noninvolved lung tissue, whereas tumor associated macrophages showed relatively lower levels of cerium. These results demonstrate the first high-dimensional single cell characterization of environmental metal exposure associated with smoking, and offer a demonstration of the unique potential for applying mass cytometry to the field of environmental toxicology. Typical procedures of single cell mass cytometry are described in detail [56]. Many challenges remain however, especially in automated analysis, where the special perimeter of single cells in a histological tissue section needs to be accurately defined [57].

A novel imaging method, multiplexed ion beam imaging (MIBI), uses an ion beam to liberate metal ion reporters, which are quantified by mass spectrometry [58]. MIBI is capable of analyzing up to 100 targets simultaneously over a five-log dynamic range in formalin-fixed, paraffin-embedded breast cancer tissue sections. MIBI can quantitate HER2 expression on breast carcinoma tissue with known HER2 status and those values correlate with pathologist-determined immunohistochemistry (IHC) scores. This method is also capable of analyzing peripheral blood mononuclear cells (PBMC) immobilized on a poly-l-lysine-coated silicon wafer, which gains similar yield as mass cytometry does with PBMC suspensions [58]. MIBI analysis and immunofluorescence-based automated quantitative analysis (AQUA) technology, an industry-accepted quantitation system, showed strong correlation in the same blind study [59]. These data show MIBI analysis can quantitate protein expression with greater sensitivity and objectivity compared to standard pathologist interpretation, demonstrating its potential as a technology capable of advancing cancer and patient diagnostics. While requiring more specialized equipment, this latter method may offer increased speed, sensitivity and resolution.

7.3.3 Microfluidics and Chip Technologies

Microfluidics and its laboratory-on-a-chip technologies are novel single-cell proteomic approaches that could analyze specimens including blood, body fluids (eg. pleural effusions, ascites, CSF), fine needle aspiration of tumors or cancer-cell-containing tissue fluids. To analyze single cells, the generic microfluidic system has three modules for (1) isolation of single cells and loading onto the microfluidic chip device, (2) capture of single cells in a controlled environment for in situ stimulation and/or cell lysis and, (3) subsequent qualitative and quantitative proteomic analysis. With these methods, researches are also capable of measuring biophysical parameters of cells including motility, functions contextual to the microenvironment, and morphology. Furthermore, by using standard recombinant protein solutions or synthetic peptides, absolute single-cell protein copy numbers can be calculated.

Basing on the technique, the microfluidics and chip technologies could be classified into two categories, tools Using Cellular Staining Assays, or using surface-based immunoassays. The platform of microfluidics has been incorporated valves, mixers, microchannels, and microchambers onto the same microchip. Microfluid biochips have been developed for

detecting the microRNA profile in NSCLC serum samples [60]. Multicolor analysis of fluorescent antibody-labeled cytoplasmic proteins was realized to assay for six parameters from single cells from several human brain tumor biopsies [61]. Microfluidic platform-microfluidic image cytometry (MIC) is capable of quantitative, single-cell proteomic analysis of multiple signaling molecules using only 1000–2800 cells. Using such techniques, Sun et al. showed simultaneous measurement of four critical signaling proteins (EGFR, PTEN, phospho-Akt, and phospho-S6), which are associated with the phosphoinositide 3-kinase/Akt/mammalian target of rapamycin (mTOR) signaling pathway in cancer cells [61]. These MIC measurements were validated by clinical immunohistochemistry and confirmed the striking intertumoral and intratumoral heterogeneity characteristic of glioblastoma.

To further interpret the multiparameter, single-cell MIC measurements, the authors adapted bioinformatic methods including self-organizing maps that stratify patients into clusters that predict tumor progression and patient survival. These results indicate that the MIC platform represents a robust, enabling in vitro molecular diagnostic technology for systems pathology analysis and personalized medicine, especially when using together with bioinformatics analysis. Except for cancer cells, MIC may also be used in formalin-fixed, paraffin-embedded tissues. A sequential immunohistochemical staining procedure permits a high level (~60 parameters) of multiplexing analysis at the single-cell level [62]. This provides an opportunity to develop single-cell tools for diagnostics and therapeutic targets. Khoo et al. enriched and characterized putative CTCs from blood samples of patients with both advanced stage metastatic lung and breast cancers using a novel multiplexed spiral microfluidic chip [63]. This system detected putative CTCs under high sensitivity (100%; for Lung cancer samples: 10-1535 CTCs/ml) rapidly from clinically relevant blood volumes. Blood samples were completely separated into plasma, CTCs and PBMCs components and each fraction were characterized with immuno-phenotyping (Pan-cytokeratin/CD45, CD44/CD24, EpCAM), fluorescence in-situ hybridization (FISH) (EML4-ALK) or targeted somatic mutation analysis [63].

Tools using surface-based immunoassays are conceptually similar to ELISpot, but their results are significantly higher level of multiplexing, which allows functional assays of intracellular, membrane, and secreted proteins and protein-protein/cell-cell interactions. The technique of microengraved single-cell proteomics chips employs an array of nearly 105 microwells to isolate and culture single cells and quantized cell populations. A microengraved or antibody-coated substrate is used to cap the microwell array and to capture secreted proteins. The single-cell barcode chips (SCBCs) are versatile and information-rich tools in which single cells, or defined numbers of cells, are isolated within microchambers that each contain a manyelement antibody array [64, 65]. A microfluidics set up with all-optical manipulation has been used to separate cells, lyse and measure proteins, including p53, on surfaces micro-printed with antibodies [64]. Microfluidics technology enables highly multiplexed assays suitable for the investigation of secreted proteins and rare cells. This provides an opportunity to develop single-cell tools for future bedside diagnostics. Other chip-based technologies, with significant potential for the future, combine human tissue and organ engineering with microfluidic technologies, was extensively reviewed elsewhere [66].

7.3.4 Chemical Cytometry

Chemical cytometry is a single-cell technique that can be used to detect and quantify proteins, nucleic acids and small molecule metabolites [67]. The single cell is lysed and its components are separated by liquid chromatography or capillary electrophoresis before measurement of the cellular components by highly sensitive techniques such as laser-induced fluorescence or mass spectrometry. The cellular components are chemically modified with a fluorophore before separation and analysis. This activity-based probe technique has successfully been used to

identify low abundance proteins on the cell surface [68]. However, chemical cytometry is generally limited, as the capture and lysis of single cells is highly time-consuming, and the number of cells analyzed is often below one thousand in a data acquisition time of up to several hours. To address this, chemical cytometry is being merged with other platforms including microfluidics [69] and mass spectrometry [70]. Multiple proof-of-concept efforts have been published highlighting the potential of protein quantification and identification from samples of trypsin-digested cell lysates from single cells (corresponding to about 0.1 ng of protein) [71]. Targeted single-cell microchemical analysis has been performed by mass spectrometry peptidomics of paraformaldehyde-fixed and immuno-labelled neurons [72]. Like mass spectectomy, chemical cytometry have been employed in multiple/single reaction monitoring (MRM/SRM) to detect and quantify proteins, with abundances of less than 50 copies per cell. A latest review summarized the up-to-date use of chemical proteomic methods in protein modification (glycosylation, sialylation, phosphorylation), metabolism and epithelial-mesenchymal transition (EMT) of cancer stem cells [73].

7.3.5 Other Techniques

Single-cell Western blotting (scWestern or scWB) is based on Western blotting technique. Typically, scWB comprises the following five main stages: (a) gravity settling of cells into microwells; (b) chemical lysis of cells in each microwell; (c) polyacrylamide gel electrophoresis (PAGE) of each single-cell lysate; (d) exposure of the gel to UV light to blot (immobilization) proteins to the gel matrix; and (e) in-gel immunoprobing of immobilized proteins. One of the most significant advantages of scWB is that it overcomes the antibody cross-reactivity because proteins are first separated by molecular mass via electrophoresis before the antibody probing step, thereby enabling clear discrimination between on-target and off-target signals, including protein isoforms that lack selective antibodies [74, 75]. Previously, scWB was reported as exhibiting a linear dynamic range of 1.3~2.2 orders and detection thresholds of 27,000 molecules and may detect as much as 50% of proteins in the mammalian proteome [74, 75]. Although this technique is relevant when direct assessment of proteins in single cells is needed, with applications spanning from the fundamental biosciences to applied biomedicine, its analytical yields are limited by in-gel antibody probe concentration and combination [74, 75]. Furthermore, the non-specific binding of non-targeted proteins hinders the identification and quantification of target proteins [74, 75].

Single molecule array (SiMoA) detects proteins with single molecule resolution and, thus, leads to absolute quantification. SiMoA uses a large number of antibody-coated beads to capture a few proteins, resulting in the capture of single molecules on the beads. Sandwich-type immunoassay with enzyme amplification is utilized for signal readout of single molecules. SiMoA has investigated serum and other biofluids to demonstrate ultra-low detection limits and a large dynamic range compared to traditional ELISA [76]. Furthermore, the variation of a prostate-specific antigen across single prostate cancer cells has been examined with SiMoA to reveal the expression shifts with evaluated genetic drift. However, low multiplexing capacity, low throughput, and high cost for single-cell measurement limit the application of SiMoA.

7.4 Practical Considerations

As has been discussed elsewhere with mass cytometry, it's important to elaborately conceive the whole study plan, including the aims, the selection of the appropriate technique, with which kind of specimens, how to reduce the number of objective proteins, whether or how to perform successive functional studies or network studies [77].

In fact, proteomic techniques have becoming increasingly powerful and cheaper (as listed in Table 7.1), are qualified for cellular or surface proteins, and are able to deal with several sample

Table 7.1 Indices of single-cell proteomics tools

Technique	Numbers and types of poteins assayed	Through-put	Detection limit	Statistical accuracy and signal quantification	Notes and features
Flow cytometry					
Fluorescence flow cytometry	Around 15 proteins (mostly membrane proteins, a few cytoplasmic proteins)	10^4 cells/s	500 copies per cell	90% phenotyping accuracy; relative protein abundance	Standard for sorting and enumeration of cellular phenotypes. Secretion blocked and cells fixed for cytoplasmic proteins
Mass flow cytometry		10^3 cells/s	>103 copies per cell	Good cell counting statistics; relative protein abundance	Cells handled in bulk prior to analysis. Secretion blocked and cells fixed for cytoplasmic proteins
Microfluidics technologies					
Image cytometry	3–4 membrane or intracellular proteins and cell size	10^3–10^4 cells per chip	10^5 fluoro-phores per um^2	Good cell counting statistics; relative protein abundance	Cells are fixed and stained (in bulk) with fluorescent antibodies; protein assay and cell location spatially correlated
Cell array	1 intracellular protein	<10^3 cells per chip		Good cell counting statistics; relative protein abundance	Single cells separated and imaged on chip; continuous monitoring of cell physiology
Micro-droplet	1 membrane or intracellular protein	10^2 udrops s^{-1}	Not defined	Good cell sampling statistics	Cells entrained in microdroplets; microdroplet composition control permits screening cells
Micro-engraving	3 secreted or 3 membrane proteins	10^4–10^5 cells per chip	Not available	Very good cell number statistics; relative protein abundance	Cells isolated in miocrowells; surface immunoassays; proteins colorimetrically detected; secretome kinetics from single cells; proteomeic and functional assays from same cell
Single cell barcode chips	About 20 secreted, membrane, or cytoplasmic proteins, expandable	10^3–10^5 cells per chip	10^2 copies	Good cell counting statistics, absolute quantification, 10% measurement error per protein per cell	Cells isolated in microchambers, miniature antibody arrays yield spatial separation of specific protein assays; proteomeic and functional assays from same cell; single cells or defined small cell populations accessed.

types, including blood, pleural effusions, bronchoalveolar lavage fluid (BALF), cultured cell suspensions, fresh tissues (obtained by surgery or small biopsy) and formalin-fixed paraffin-embedded tissue [78–80]. So one important thing in technical selection is the availability of the devices and the technicians that operating them. Equally important in methodology is the combination of single cell proteomics with other-omics, such as genomics, transcriptomics, metabolomics and miRNomics [81–83]. This may allow us to decipher the modifications and interactions of nucleic acids and proteins.

7.5 Summary and Conclusions

The application of proteomics has been realized in lung cancer for elucidating its mechanisms and developing novel targets and biomarkers. Soon

FFC/FACS, mass cytometry, microfluidic chip, and chemical cytometry, representative of single-cell proteomics, will be increasingly used in lung cancer. Perhaps, the data integration obtained from single-cell proteomics, metabolomics, or genetics could enhance our understanding of the core events of lung oncogenesis [81, 84], thereby facilitating the development of therapeutic targets and molecular targeted drugs, as well as novel biomarkers for prevention, early detection, prognosis, and response to therapy In the clinical practice, the method selection might be based on the objectives, costs, availability, and other characteristics of these methods [85, 86]. The stability and repeatability of single cell proteomics as well as systemic validation of disease-, severity-, duration-, and subtype-specific values should be seriously considered before clinical application [87–99]. Thus, single cell proteomics will provide an alternative for the discovery and development of therapeutic targets and diagnostic biomarkers to understand molecular mechanism of lung carcinogenesis and the heterogeneity among lung cancer cells at protein levels.

References

1. Siegel RL, Miller KD, Jemal A (2017) Cancer statistics, 2017. CA Cancer J Clin 67(1):7–30
2. Greenhalgh J, Dwan K, Boland A, Bates V, Vecchio F, Dundar Y, Jain P, Green JA (2016) First-line treatment of advanced epidermal growth factor receptor (EGFR) mutation positive non-squamous non-small cell lung cancer. Cochrane Database Syst Rev 5:CD010383
3. Sabari JK, Lok BH, Laird JH, Poirier JT, Rudin CM (2017) Unravelling the biology of SCLC: implications for therapy. Nat Rev Clin Oncol 14(9):549–561
4. Cohen AA, Geva-Zatorsky N, Eden E, Frenkel-Morgenstern M, Issaeva I, Sigal A et al (2008) Dynamic proteomics of individual cancer cells in response to a drug. Science 322:1511–1516
5. Zimmer A, Amar-Farkash S, Danon T, Alon U (2017) Dynamic proteomics reveals bimodal protein dynamics of cancer cells in response to HSP90 inhibitor. BMC Syst Biol 11(1):33
6. Fujii K, Nakamura H, Nishimura T (2017) Recent mass spectrometry-based proteomics for biomarker discovery in lung cancer, COPD, and asthma. Expert Rev Proteomics 14(4):373–386
7. Birse CE, Lagier RJ, FitzHugh W, Pass HI, Rom WN, Edell ES et al (2015) Blood-based lung cancer biomarkers identified through proteomic discovery in cancer tissues, cell lines and conditioned medium. Clin Proteomics 12(1):18
8. Li L, Wei Y, To L, Zhu CQ, Tong J, Pham NA et al (2014) Integrated omic analysis of lung cancer reveals metabolism proteome signatures with prognostic impact. Nat Commun 5:5469
9. Bharti A, Ma PC, Salgia R (2007) Biomarker discovery in lung cancer-promises and challenges of clinical proteomics. Mass Spectrom Rev 26(3):451–466
10. Ocak S, Chaurand P, Massion PP (2009) Mass spectrometry-based proteomic profiling of lung cancer. Proc Am Thorac Soc 6(2):159–170
11. Soo RA, Adjei AA (2017) Predicting clinical outcomes using proteomics in non-small cell lung cancer-the past, present, and future. J Thorac Oncol 12(4):602–606
12. Yanagisawa K, Shyr Y, Xu BJ, Massion PP, Larsen PH, White BC et al (2003) Proteomic patterns of tumour subsets in non-small-cell lung cancer. Lancet 362:433–439
13. Ostroff RM, Bigbee WL, Franklin W, Gold L, Mehan M, Miller YE et al (2010) Unlocking biomarker discovery: large scale application of aptamer proteomic technology for early detection of lung cancer. PLoS One 5(12):e15003
14. Jung YJ, Katilius E, Ostroff RM, Kim Y, Seok M, Lee S, Jang S, Kim WS, Choi CM (2017) Development of a protein biomarker panel to detect non-small-cell lung cancer in Korea. Clin Lung Cancer 18(2):e99–e107
15. Codreanu SG, Hoeksema MD, Slebos RJC, Zimmerman LJ, Rahman SMJ, Li M, Chen SC, Chen H, Eisenberg R, Liebler DC, Massion PP (2017) Identification of proteomic features to distinguish benign pulmonary nodules from lung adenocarcinoma. J Proteome Res 16(9):3266–3276
16. Solanki HS, Advani J, Khan AA, Radhakrishnan A, Sahasrabuddhe NA, Pinto SM et al (2017) Chronic cigarette smoke mediated global changes in lung mucoepidermoid cells: a phosphoproteomic analysis. OMICS 21(8):474–487
17. López-Sánchez LM, Jurado-Gámez B, Feu-Collado N, Valverde A, Cañas A, Fernández-Rueda JL, Aranda E, Rodríguez-Ariza A (2017) Exhaled breath condensate biomarkers for the early diagnosis of lung cancer using proteomics. Am J Physiol Lung Cell Mol Physiol ajplung.00119.2017. https://doi.org/10.1152/ajplung.00119.2017. [Epub ahead of print]
18. Gocheva V, Naba A, Bhutkar A, Guardia T, Miller KM, Li CM, Dayton TL, Sanchez-Rivera FJ, Kim-Kiselak C, Jailkhani N, Winslow MM, Del Rosario A, Hynes RO, Jacks T (2017) Quantitative proteomics identify Tenascin-C as a promoter of lung cancer progression and contributor to a signature prognostic of patient survival. Proc Natl Acad Sci U S A 114(28):E5625–E5634
19. Yanagisawa K, Tomida S, Shimada Y, Yatabe Y, Mitsudomi T, Takahashi T (2007) A 25-signal proteomic signature and outcome for patients with resected non-small-cell lung cancer. J Natl Cancer Inst 99:858–867

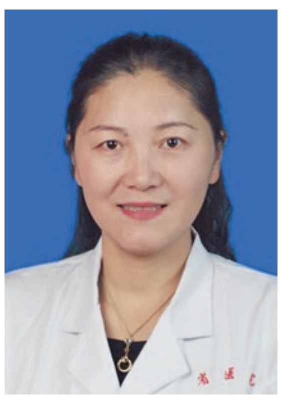
20. Taguchi F, Solomon B, Gregorc V, Rodr H, Gray R, Kasahara K et al (2007) Mass spectrometry to classify non-small-cell lung cancer patients for clinical outcome after treatment with epidermal growth factor receptor tyrosine kinase inhibitors: a multicohort cross-institutional study. J Natl Cancer Inst 99:838–846
21. Salmon S, Chen H, Chen S, Herbst R, Tsao A, Tran H et al (2009) Classification by mass spectrometry can accurately and reliably predict outcome in patients with nonsmall cell lung cancer treated with erlotinib-containing regimen. J Thorac Oncol 4(6):689–696
22. Gregorc V, Novello S, Lazzari C, Barni S, Aieta M, Mencoboni M et al (2014) Predictive value of a proteomic signature in patients with non-small-cell lung cancer treated with second-line erlotinib or chemotherapy (PROSE): a biomarker-stratified, randomised phase 3 trial. Lancet Oncol 15:713–721
23. Grossi F, Rijavec E, Genova C, Barletta G, Biello F, Maggioni C et al (2017) Serum proteomic test in advanced non-squamous non-small cell lung cancer treated in first line with standard chemotherapy. Br J Cancer 116(1):36–43
24. Lara PN Jr, Moon J, Hesketh PJ, Redman MW, Williamson SK, Akerley WL 3rd, Hirsch FR, Mack PC, Gandara DR (2016) SWOG S0709: Randomized phase II trial of erlotinib versus erlotinib plus carboplatin/paclitaxel in patients with advanced non-small cell lung cancer and impaired performance status as selected by a serum proteomics assay. J Thorac Oncol 11(3):420–425
25. Paul D, Chanukuppa V, Reddy PJ, Taunk K, Adhav R, Srivastava S, Santra MK, Rapole S (2016) Global proteomic profiling identifies etoposide chemoresistance markers in non-small cell lung carcinoma. J Proteome 138:95–105
26. Rovithi M, Lind JS, Pham TV, Voortman J, Knol JC, Verheul HM, Smit EF, Jimenez CR (2016) Response and toxicity prediction by MALDI-TOF-MS serum peptide profiling in patients with non-small cell lung cancer. Proteomics Clin Appl 10(7):743–749
27. Nyberg F, Ogiwara A, Harbron CG, Kawakami T, Nagasaka K, Takami S et al (2011) Proteomic biomarkers for acute interstitial lung disease in gefitinib-treated Japanese lung cancer patients. PLoS One 6(7):e22062
28. Peters S, Stahel RA, Dafni U, Ponce Aix S, Massutí B, Gautschi O et al (2017) Randomized phase III trial of erlotinib versus docetaxel in patients with advanced squamous cell non-small cell lung cancer failing first-line platinum-based doublet chemotherapy stratified by VeriStrat good versus VeriStrat poor. The European Thoracic Oncology Platform (ETOP) EMPHASIS-lung Trial. J Thorac Oncol 12(4):752–762
29. Gadgeel S, Goss G, Soria JC, Felip E, Georgoulias V, Lu S et al (2017) Evaluation of the VeriStrat serum protein test in patients with advanced squamous cell carcinoma of the lung treated with second-line afatinib or erlotinib in the phase III LUX-Lung 8 study. Lung Cancer 109:101–108
30. Tan YH, Lee KH, Lin T, Sun YC, Hsieh-Li HM, Juan HF, Wang YC (2008) Cytotoxicity and proteomics analyses of OSU03013 in lung cancer. Clin Cancer Res 14(6):1823–1830
31. Li Y, Zhang B, Wang X, Yan H, Chen G, Zhang X (2011) Proteomic analysis of apoptosis induction in human lung cancer cells by recombinant MVL. Amino Acids 41(4):923–932
32. Labots M, Schütte LM, van der Mijn JC, Pham TV, Jiménez CR, Verheul HM (2014) Mass spectrometry-based serum and plasma peptidome profiling for prediction of treatment outcome in patients with solid malignancies. Oncologist 19(10):1028–1039
33. Cheung CHY, Juan HF (2017) Quantitative proteomics in lung cancer. J Biomed Sci 24:37
34. Yang L, Tang C, Xu B, Wang W, Li J, Li X, Qin H, Gao H, He K, Song S, Liu X (2015) Classification of epidermal growth factor receptor gene mutation status using serum proteomic profiling predicts tumor response in patients with stage IIIB or IV non-small-cell lung cancer. PLoS One 10(6):e0128970
35. Heath JR, Ribas A, Mischel PS (2016) Single cell analytic tools for drug discovery and development. Nat Rev Drug Discov 15(3):204–216
36. Gavasso S, Gullaksen S, Skavland J, Gjertsen BT (2016) Single-cell proteomics: potential implications for cancer diagnostics. Expert Rev Mol Diagn 16(5):579–589
37. Lu Y, Yang L, Wei W, Shi Q (2017) Microchip-based single-cell functional proteomics for biomedical applications. Lab Chip 17(7):1250–1263
38. Wei W, Shin YS, Ma C, Wang J, Elitas M, Fan R, Heath JR (2013) Microchip platforms for multiplex single-cell functional proteomics with applications to immunology and cancer research. Genome Med 5(8):75
39. Beck TN, Chikwem AJ, Solanki NR, Golemis EA (2014) Bioinformatic approaches to augment study of epithelial-to-mesenchymal transition in lung cancer. Physiol Genomics 46(19):699–724
40. Marrugal Á, Ojeda L, Paz-Ares L, Molina-Pinelo S, Ferrer I (2016) Proteomic-based approaches for the study of cytokines in lung cancer. Dis Markers 2016:2138627
41. Herzenberg LA, Parks D, Sahaf B, Perez O, Roederer M, Herzenberg LA (2002) The history of the fluorescence activated cell sorter and flow cytometry: a view from Stanford. Clin Chem 48:1819–1827
42. Lamoreaux L, Roederer M, Koup R (2006) Intracellular cytokine optimization and standard operating procedure. Nat Protoc 1:1507–1516
43. Leelatian N, Doxie DB, Greenplate AR, Mobley BC, Lehman JM, Sinnaeve J et al (2017) Single cell analysis of human tissues and solid tumors with mass cytometry. Cytometry B Clin Cytom 92(1):68–78
44. Lin CC, Huang WL, Su WP, Chen HH, Lai WW, Yan JJ, Su WC (2010) Single cell phospho-specific flow cytometry can detect dynamic changes of phospho-Stat1 level in lung cancer cells. Cytometry A 77(11):1008–1019

45. Skavland J, Jorgensen KM, Hadziavdic K, Hovland R, Jonassen I, Bruserud O, Gjertsen BT (2011) Specific cellular signal transduction responses to in vivo combination therapy with ATRA, valproic acid and theophylline in acute myeloid leukemia. Blood Cancer J 1(2):e4
46. Kotecha N, Flores NJ, Irish JM, Simonds EF, Sakai DS, Archambeault S, Diaz-Flores E, Coram M, Shannon KM, Nolan GP, Loh ML (2008) Single-cell profiling identifies aberrant STAT5 activation in myeloid malignancies with specific clinical and biologic correlates. Cancer Cell 14(4):335–343
47. Lu Y, Liang H, Yu T, Xie J, Chen S, Dong H, Sinko PJ, Lian S, Xu J, Wang J, Yu S, Shao J, Yuan B, Wang L, Jia L (2015) Isolation and characterization of living circulating tumor cells in patients by immunomagnetic negative enrichment coupled with flow cytometry. Cancer 121(17):3036–3045
48. Mesri M, Birse C, Heidbrink J, McKinnon K, Brand E, Bermingham CL, Feild B, Fitzhugh W, He T, Ruben S, Moore PA (2013) Identification and characterization of angiogenesis targets through proteomic profiling of endothelial cells in human cancer tissues. PLoS One 8(11):e78885
49. Diercks AH, Ozinsky A, Hansen CL, Spotts JM, Rodriguez DJ, Aderem A (2009) A microfluidic device for multiplexed protein detection in nano-liter volumes. Anal Biochem 386:30–35
50. Shrestha B, Vertes A (2009) In situ metabolic profiling of single cells by laser ablation electrospray ionization mass spectrometry. Anal Chem 81:8265–8271
51. Bandura DR, Baranov VI, Ornatsky OI, Antonov A, Kinach R, Lou X, Pavlov S, Vorobiev S, Dick JE, Tanner SD (2009) Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. Anal Chem 81(16):6813–6822
52. Gullaksen SE, Skavland J, Gavasso S, Tosevski V, Warzocha K, Dumrese C et al (2017) Single cell immune profiling by mass cytometry of newly diagnosed chronic phase chronic myeloid leukemia treated with nilotinib. Haematologica 102(8):1361–1367
53. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir e-AD, Tadmor MD et al (2015) Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell 162(1):184–197
54. Giesen C, Wang HA, Schapiro D, Zivanovic N, Jacobs A, Hattendorf B (2014) Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods 11(4):417–422
55. Rahman AH, Lavin Y, Kobayashi S, Leader A, Merad M (2017) High-dimensional single cell mapping of cerium distribution in the lung immune microenvironment of an active smoker. Cytometry B Clin Cytom https://doi.org/10.1002/cyto.b.21545. [Epub ahead of print]
56. Brodie TM, Tosevski V (2017) High-dimensional single-cell analysis with mass cytometry. Curr Protoc Immunol 118:5.11.1–5.11.25
57. Schuffler PJ, Schapiro D, Giesen C, Wang HA, Bodenmiller B, Buhmann JM (2015) Automatic single cell segmentation on highly multiplexed tissue images. Cytometry A 87(10):936–942
58. Angelo M, Bendall SC, Finck R, Hale MB, Hitzman C, Borowsky AD, Levenson RM, Lowe JB, Liu SD, Zhao S et al (2014) Multiplexed ion beam imaging of human breast tumors. Nat Med 20:436–442
59. Rost S, Giltnane J, Bordeaux JM, Hitzman C, Koeppen H, Liu SD (2017 Aug) Multiplexed ion beam imaging analysis for quantitation of protein expresssion in cancer tissue sections. Lab Investig 97(8):992–1003
60. Roth C, Stückrath I, Pantel K, Izbicki JR, Tachezy M, Schwarzenbach H (2012) Low levels of cell-free circulating miR-361-3p and miR-625* as blood-based markers for discriminating malignant from benign lung tumors. PLoS One 7(6):e38248
61. Sun J, Michael Masterman-Smith M, Graham NA (2010) A microfluidic platform for systems pathology: multiparameter single-cell signaling measurements of clinical brain tumor specimens. Cancer Res 70(15):6128–6138
62. Huang L, Michael SA, Chen Y, Wu H (2017) Current advances in highly multiplexed antibody-based single-cell proteomic measurements. Chem Asian J 12(14):1680–1691
63. Khoo BL, Warkiani ME, Tan DS, Bhagat AA, Irwin D, Lau DP, Lim AS, Lim KH, Krisna SS, Lim WT, Yap YS, Lee SC, Soo RA, Han J, Lim CT (2014) Clinical validation of an ultra high-throughput spiral microfluidics for the detection and enrichment of viable circulating tumor cells. PLoS One 9(7):e99409
64. Fan B, Li X, Chen D, Peng H, Wang J, Chen J (2016) Development of microfluidic systems enabling high-throughput single-cell protein characterization. Sensors (Basel) 16(2):232
65. Yu J, Zhou J, Sutherland A, Wei W, Shin YS, Xue M, Heath JR (2014) Microfluidics-based single-cell functional proteomics for fundamental and applied biomedical applications. Annu Rev Anal Chem 7:275–295
66. Chen H, Lau MC, Wong MT, Newell EW, Poidinger M, Chen J (2016) Cytofkit: a bioconductor package for an integrated mass cytometry data analysis pipeline. PLoS Comput Biol 12(9):e1005112
67. Cohen D, Dickerson JA, Whitmore CD, Turner EH, Palcic MM, Hindsgaul O, Dovichi NJ (2008) Chemical cytometry: fluorescence-based single-cell analysis. Annu Rev Anal Chem 1:165–190
68. Xu F, Zhao H, Feng X, Chen L, Chen D, Zhang Y, Nan F, Liu J, Liu BF (2014) Single-cell chemical proteomics with an activity based probe: identification of low-copy membrane proteins on primary neurons. Angew Chem Int Ed Engl 53(26):6730–6733
69. Yan H, Zhang B, Wu H (2008) Chemical cytometry on microfluidic chips. Electrophoresis 29(9):1775–1786
70. Kleparnik K (2015) Recent advances in combination of capillary electrophoresis with mass spectrometry: methodology and theory. Electrophoresis 36(1):159–178

71. Mainz ER, Dobes NC, Allbritton NL (2015) Pronase E-based generation of fluorescent peptide fragments: tracking intracellular peptide fate in single cells. Anal Chem 87(15):7987–7995
72. Sun L, Zhu G, Yan X et al (2014) Capillary zone electrophoresis for analysis of complex proteomes using an electrokinetically pumped sheath flow nanospray interface. Proteomics 14(4–5):622–628
73. Jung HJ (2017) Chemical proteomic approaches targeting cancer stem cells: a review of current literature. Cancer Genomics Proteomics 14(5):315–327
74. Kang CC, Yamauchi KA, Vlassakis J, Sinkala E, Duncombe TA, Herr AE (2016) Single cell-resolution western blotting. Nat Protoc 11(8):1508–1530
75. Hughes AJ, Spelke DP, Xu Z, Kang CC, Schaffer DV, Herr AE (2014) Single-cell western blotting. Nat Methods 11(7):749–755
76. Streets AM, Huang Y (2014) Microfluidics for biological measurements with single-molecule resolution. Curr Opin Biotechnol 25:69–77
77. Spitzer MH, Nolan GP (2016) Mass cytometry: single cells, many features. Cell 165(4):780–791
78. Amann JM, Chaurand P, Gonzalez A, Mobley JA, Massion PP, Carbone DP, Caprioli RM (2006) Selective profiling of proteins in lung cancer cells from fine-needle aspirates by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Clin Cancer Res 12:5142–5150
79. Hmmier A, O'Brien ME, Lynch V, Clynes M, Morgan R, Dowling P (2017) Proteomic analysis of bronchoalveolar lavage fluid (BALF) from lung cancer patients using label-free mass spectrometry. BBA Clin 7:97–104
80. Carvalho AS, Cuco CM, Lavareda C, Miguel F, Ventura M, Almeida S et al (2017) Bronchoalveolar lavage proteomics in patients with suspected lung cancer. Sci Rep 7:42190
81. Fahrmann JF, Grapov DD, Wanichthanarak K, DeFelice BC, Salemi MR, Rom WN, Gandara DR, Phinney BS, Fiehn O, Pass H, Miyamoto S (2017) Integrated metabolomics and proteomics highlight altered nicotinamide- and polyamine pathways in lung adenocarcinoma. Carcinogenesis. pii: bgw205. https://doi.org/10.1093/carcin/bgw205. [Epub ahead of print]
82. Shen Q, Cheng F, Song H, Lu W, Zhao J, An X, Liu M, Chen G, Zhao Z, Zhang J (2017) Proteome-scale investigation of protein allosteric regulation perturbed by somatic mutations in 7000 cancer genomes. Am J Hum Genet 100(1):5–20
83. Backes C, Ludwig N, Leidinger P, Huwer H, Tenzer S, Fehlmann T, Franke A, Meese E, Lenhof HP, Keller A (2016) Paired proteomics, transcriptomics and miRNomics in non-small cell lung cancers: known and novel signaling cascades. Oncotarget 7(44):71514–71525
84. Xue M, Wei W, Su Y, Kim J, Shin YS, Mai WX, Nathanson DA, Heath JR (2015) Chemical methods for the simultaneous quantitation of metabolites and proteins from single cells. J Am Chem Soc 137(12):4066–4069
85. Wiwie C, Baumbach J, Röttger R (2015) Comparing the performance of biomedical clustering methods. Nat Methods 12:1033–1038
86. Weber LM, Robinson MD (2016) Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. Cytometry A 89(12):1084–1096
87. Shi L, Zhu B, Xu M, Wang X (2017) Selection of AECOPD-specific immunomodulatory biomarkers by integrating genomics and proteomics with clinical informatics. Cell Biol Toxicol https://doi.org/10.1007/s10565-017-9405-x
88. Hwang M, Park HH, Choi H, Lee KY, Lee YJ, Koh SH (2017) Effects of aspirin and clopidogrel on neural stem cells. Cell Biol Toxicol https://doi.org/10.1007/s10565-017-9412-y
89. Chu MP, Kriangkum J, Venner CP, Sandhu I, Hewitt J, Belch AR, Pilarski LM (2017 Apr) Addressing heterogeneity of individual blood cancers: the need for single cell analysis. Cell Biol Toxicol 33(2):83–97. https://doi.org/10.1007/s10565-016-9367-4
90. Wang W, Zhu B, Wang X (2017 Oct) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427. https://doi.org/10.1007/s10565-017-9400-2
91. Chen C, Shi L, Li Y, Wang X, Yang S (2016 Jun) Disease-specific dynamic biomarkers selected by integrating inflammatory mediators with clinical informatics in ARDS patients with severe pneumonia. Cell Biol Toxicol 32(3):169–184. https://doi.org/10.1007/s10565-016-9322-4
92. Gu J, Wang X (2016 Feb) New future of cell biology and toxicology: thinking deeper. Cell Biol Toxicol 32(1):1–3. https://doi.org/10.1007/s10565-016-9313-5.94
93. Amodio G, Sasso E, D'Ambrosio C, Scaloni A, Moltedo O, Franceschelli S, Zambrano N, Remondelli P (2016 Aug) Identification of a microRNA (miR-663a) induced by ER stress and its target gene PLOD3 by a combined microRNome and proteome approach. Cell Biol Toxicol 32(4):285–303. https://doi.org/10.1007/s10565-016-9335-z
94. Wang X (2016 Oct) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361. https://doi.org/10.1007/s10565-016-9350-0
95. Zhu D, Liu Z, Pan Z, Qian M, Wang L, Zhu T, Xue Y, Wu D (2016 Aug) A new method for classifying different phenotypes of kidney transplantation. Cell Biol Toxicol 32(4):323–332. https://doi.org/10.1007/s10565-016-9337-x
96. Xu M, Wang X (2017 Aug) Critical roles of mucin-1 in sensitivity of lung cancer cells to tumor necrosis factor-alpha and dexamethasone. Cell Biol Toxicol 33(4):361–371. https://doi.org/10.1007/s10565-017-9393-x
97. Wang W, Gao D, Wang X (2017) Can single-cell RNA sequencing crack the mystery of cells? Cell Biol Toxicol https://doi.org/10.1007/s10565-017-9404-y. [Epub ahead of print]

98. Wang X (2016 Aug) CBT profiles of cabozantinib approved for advanced renal cell carcinomas. Cell Biol Toxicol 32(4):259–261. https://doi.org/10.1007/s10565-016-9349-6
99. Bao L, Zhang Y, Wang J, Wang H, Dong N, Su X, Xu M, Wang X (2016 Oct) Variations of chromosome 2 gene expressions among patients with lung cancer or non-cancer. Cell Biol Toxicol 32(5):419–435. https://doi.org/10.1007/s10565-016-9343-z

Dr. Zheng Wang is specialized in respiratory medicine. In 2009 he received his medical doctor's degree at Tongji Medical College, Huazhong University Science and Technology, P. R. China. Then he has been serving in the respiratory clinics. He is now a deputy chief physician in Affiliated Henan Provincial People's Hospital, Zhengzhou University, and Vice Professor of Medicine and post-graduate supervisor of Zhengzhou University and Henan University. He is also a vice chair of Youth Academican Comittee of the Chinese Lung Cancer Association, and serves as an editor for Journal of Respiratory Medicine and Lung Disease. His major research interests are epithelial-mesenchymal transition in pulmonary fibrosis and lung cancer, bioinformatics in respiratory medicine and interventional bronchoscopy. Dr. Wang is sponsored by The National Natural Science Foundation of China, The Science and Technology Foundation of Henan Province and the Health Bureau of Henan Province. The amount of total grants exceed 50,000 Euros. Dr. Wang is the co-author of four books, holds 4 Chinese patents, and is the first or corresponding author of 10 peer-reviewed English articles and 15 Chinese articles.



Dr. Xiaoju Zhang is specialized in respiratory medicine and has been serving in the respiratory departments. In 2005 she received his medical doctor's degree at Tongji Medical College, Huazhong University Science and Technology, P. R. China. She received post-doctoral scientific training in University of California, San Francisco, between July 2009 and Jan 2012.1. Her doctoral and post-doctoral researches are about lung cancer and antibody preparation, respectively. She is now the director and chief physician in the Department of Respiratory and Critical Care Medicine, Affiliated Henan Provincial People's Hospital, Zhengzhou University. She is also serving as a professor of Medicine, post-graduate and medical doctors' supervisor of Zhengzhou University, Henan University and Xingxiang Medical College. She is a youth member of Chinese Respiratory Association, member of the Lung Cancer Expert Board of Chinese Respiratory Association, member of the Chinese Respiologists Association, member of the Chinese Society for Disease Prevention and Control, Vice chair of the Henan Respiratory Association, Vice chair of the Henan Respirologists Association. Dr. Zhang's major research interests are epithelial-mesenchymal transition in lung cancer, pulmonary nodules and bioinformatics in respiratory medicine. Dr. Zhang is sponsored by the National Natural Science Foundation of China, the Ministry of Human Resources and Social Security of China, the Science and Technology Foundation of Henan Province, the Health Bureau of Henan Province, the Science and Technology Foundation of Zhengzhou City. The amount of total grants exceed 800,000 US dollors. Dr. Zhang is the co-author of five books, holds 3 Chinese patents, and is the first or corresponding author of 15 peer-reviewed English articles and 30 Chinese articles.

8 Therapeutic Antibody Discovery in Infectious Diseases Using Single-Cell Analysis

Alexandria Voigt, Touyana Semenova, Janet Yamamoto, Veronique Etienne, and Cuong Q. Nguyen

Abstract

Since the discovery of mouse hybridoma technology by Kohler and Milstein in 1975, significant progress has been made in monoclonal antibody production. Advances in B cell immortalization and phage display technologies have generated a myriad of valuable monoclonal antibodies for diagnosis and treatment. Technological breakthroughs in various fields of 'omics have shed crucial insights into cellular heterogeneity of a biological system in which the functional individuality of a single cell must be considered. Based on this important concept, remarkable discoveries in single-cell analysis have made in identifying and isolating functional B cells that produce beneficial therapeutic monoclonal antibodies. In this review, we will discuss three traditional methods of antibody discovery. Recent technological platforms for single-cell antibody discovery will be reviewed. We will discuss the application of the single-cell analysis in finding therapeutic antibodies for human immunodeficiency virus and emerging Zika arbovirus.

Keywords

Single cell · Monoclonal antibodies · Hybridoma · Flow cytometry · Infectious diseases

A. Voigt · T. Semenova · J. Yamamoto · V. Etienne
Department of Infectious Diseases and Pathology, College of Veterinary Medicine, University of Florida, Gainesville, FL, USA

C. Q. Nguyen (✉)
Department of Infectious Diseases and Pathology, College of Veterinary Medicine, University of Florida, Gainesville, FL, USA

Department of Oral Biology, College of Dentistry, University of Florida, Gainesville, FL, USA

Center of Orphaned Autoimmune Diseases, University of Florida, Gainesville, FL, USA
e-mail: nguyenc@ufl.edu

8.1 Introduction

The humoral adaptive immunity elicits a protective immunological response against a pathogen. Unlike the innate response, humoral immunity is dependent on the extensive diversity of antigen recognition repertoires of the receptors expressed on B cells. Governed by allelic exclusion, each B cell should express solely one heavy chain and one light chain allele of immunoglobulin, therefore should produce an antibody that binds to one specific antigen. Some of these antigen-specific B cells differentiate into plasma cells to produce potent monoclonal antibodies (mAbs) and some develop into memory B cells to be reactivated for subsequent pathogen exposure. Plasma cells are short-lived, whereas memory

© Springer Nature Singapore Pte Ltd. 2018

J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_8

B cells are rate and difficult to reactivate in culture. Major technological advances have been attempted to harness the therapeutic power of adaptive immunity, specifically the effector function of mAbs. Since the discovery of mouse hybridoma technology by Kohler and Milstein in 1975, where immortalized myeloma and spleen cells were fused to produce anti-sheep red blood cell antibodies [1], the field of antibody discovery and application has evolved into a thriving industry comprised of basic research, diagnoses, and therapy. Global sale of monoclonal antibody products has increased dramatically from $39 billion in 2008 to $75 billion in 2013 [2]. Global sales revenue is expected to grow to $122.6 billion in 2019 [3]. The impetus for antibody development is the high specificity and affinity to an antigenic target that can either activate, inhibit, or block the target. Furthermore, the continued interest for antibody products is driven by the technological advancement of genomics, transcriptomics, proteomics, and metabolomics which identifies new targets of specific biological pathways that can be utilized to mitigate the disease process.

The Food and Drug Administration approved the first mouse mAb specific against CD3 (known as orthoclone OKT3; Ortho Biotech) for treatment of acute rejection of cadaveric renal transplantation. OKT3 was highly effective for acute renal-allograft rejection in a prospective randomized multicenter trial [4]. However, antibodies of mouse origin have not been successful due to human anti-mouse immune response in patients. To circumvent these challenges, a number of engineering approaches have been undertaken. For example, chimeric antibodies with mouse variable domain regions fused to human constant regions were tested [5, 6]. Another approach is antibody humanization in which by grafting mouse complementary determining regions (CDRs) that were evolved to bind to specific antigen into human immunoglobulin (Ig) backbone [7]. Other approaches have been attempted including human hybridoma technology and humanized transgenic animals in which the mouse Ig repertoires are replaced with human Ig repertoires. These technological variations have helped expand the therapeutic mAb product market in which 36 FDA-approved therapies constitute nearly 40% of the biologics market and 350 mAbs are currently in clinical trials [8–10].

Hybridoma technology and immortalization of antigen-specific B cells have been the traditional methods of mAb production. Sorting of desired B cell subset using fluorescence activated cell sorting (FACS), recombinant phage display technologies, and application of humanized transgenic mice have remarkably advanced the field. Some of these methods only capture average measurement from bulk or whole cell population undermines the heterogeneity or the autonomy of individual cells [11, 12]. Recent developments in microfluidic chamber devices and microfabrication of nanowells designed to identify antigen-specific single cell have revolutionized the process of antibody discovery. In this review, we will discuss the traditional methods of monoclonal antibody production, specifically immortalization of antigen-specific human B cell by Epstein-Barr virus, hybridoma generation, and phage display. We will focus on current platforms for single-cell antibody discovery including fluorescence activated cell sorting, microfluidic devices, and single-cell antibody nanowells. Lastly, we will discuss the application of the single-cell analysis in finding therapeutic antibodies for human immunodeficiency virus and emerging Zika arbovirus.

8.2 Traditional Methods of Antibody Discovery

8.2.1 Hybridoma Technology and Immortalization of Antigen-Specific Human B Cells

As mentioned, the hybridoma technique was first introduced to make mouse mAbs [1]. The technique, which has been refined over the years, took sheep red blood cell (SRBC) as immunogen and immunized BALB/c mice. The splenocytes of immunized mice were collected and fused with myeloma cells (Sp-1) to produce hybridoma cells. Immortalized hybridoma cells were selected in the

presence of hypoxanthine-aminopterin-thymidine (HAT) selection medium. Unfused cells lack the hypoxanthine-guanine phosphoribosyltransferase (HGPRT) gene which makes them sensitive to the HAT selection. The aminopterin blocks the de novo DNA nucleotide synthesis pathway, therefore cells must alternatively utilize the salvage pathway to replicate in the presence of hypoxanthine and thymidine. However, the myelomas deficient in HGPRT are unable to replicate. As a result, only fused cells inherit a functional HGPRT gene from B cells can proliferate and produce antibodies. Antibody-producing B cells are further cloned and expanded by limited dilution using 96- or 384-well plates. The cloning is typically performed in multiple rounds to possibly obtain expanded clones from a single cell. Supernatants are screened by enzyme-linked immunosorbent assay (ELISA) to identify antigen specific B cell clones (Fig. 8.1). The process is efficient, but it can be time-consuming and labor intensive. Additionally, the resultant antibodies are of mouse origin, thereby preventing direct therapeutic translation to humans. To avoid some of these obstacles, the Epstein-Barr virus (EBV) has also been utilized to help immortalize B cells. The transformation is achieved by the activation of EBV-encoded nuclear proteins (EBNA1, EBNA2, EBNA3A, EBNA3B, EBNA3C, EBNA-LP) and the latent membrane proteins (LMP1, LMP2A, LMP2B) in latently infected B cells. These proteins have multiple functions, but mainly induction of survival, proliferation, and inhibition of apoptosis by upregulating expression of the anti-apoptotic proteins [13]. The advantages of EBV-transformed B cells are the more rapid and efficient screening for antigen-specific B cells in comparison to the hybridoma method. Additionally, human B cells can be directly transformed to obtain antibodies, therefore there is little concern for anti-human antibody reaction. While the EBV-transformed B cells produce immunoglobulins, they yield lower quantities, which is sub-optimal for application purposes; these cells are notoriously difficult to clone and propagate [14]. The hybridoma technology and EBV transformation have shown promise and are methods that research and industry have adapted and improved, leading to several approved monoclonal antibodies [15–17]. Furthermore, significant advances in antibody engineering have been made to avoid adverse effects like acute anaphylaxis in patients, when treated with hybridoma derived mAbs [18].

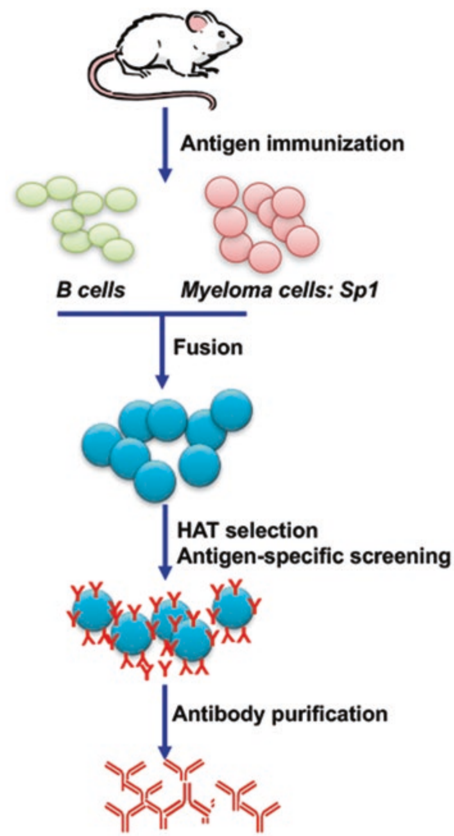


Fig. 8.1 Schematic of monoclonal antibody production by hybridoma technology. Laboratory mice are immunized with antigen of interest. Spleen cells are isolated and fused with immortalized myeloma cells such as SP1 cell line. Transformed fused cells are selected under HAT media. Antigen-specific B cells are screened using protein analytic methods (ELISA, Western blotting, flow cytometry). Once B cells of interest are identified, serial dilution will be performed to select for single cells which will be clonal expanded. Lastly, antigen-specific B cell clones will be cultured and antibodies will be purified

8.2.2 Phage Display

Phage display was first introduced in 1985 [19], and has proven an effective method for mAb production, in addition to quicker production

time when compared to hybridoma technology. The phage display technology is dependent on libraries naïve, immune, semi-synthetic, or synthetic of antibodies [20–22], which represent non-infected, cleared infection or immunized, random sequences paired with those naturally occurring, and purely generated sequences, respectively. Since a naïve library has antibodies which have not undergone a maturation process, many will have poor binding affinity [23], however high affinity mAbs have been generated [24, 25]. In contrast, the immune library is taken from individuals that are immune to the disease, so the library is inherently biased for antibodies that are mature and specific to the disease in question [20, 26]; hence, a higher frequency of high affinity antibodies can be obtained [20], but new immune libraries must be created for a specific infection or disease, resulting in a limited repertoire. Both semi-synthetic and synthetic libraries use synthetic oligonucleotides to generate diversity within the library, however may generate sequences that negative selection may have removed [20, 27]. While semi-synthetic libraries use some natural sequences, both types of libraries are devoid of a natural maturation process [22]. These libraries can be utilized to design phages which present antibodies or fragments of antibodies as part of their protein coat. Diversity is key to the success of a library, however frameshift mutations and transformation efficiency or the ability for cells to take up extracellular DNA and encode it can be a major concern [20]. Original cloning methods presented a multitude of technical difficulties, which resulted in several other innovations, including PCR-based assemblies [28]. Furthermore, in an effort to increase diversity, molecular methods, such as mutagenesis and sequence evolution [29–33], have been utilized to help increase the diversity in libraries; this can be done at various intervals in the development process.

After diverse libraries are created, the phages expressing various fragmented or entire antibodies must be selected [34] via a process called "panning". The antigen-antibody complex is put through a series of parameters –e.g. toxicity to host [35]– then the phages are eluted, often via ELISA, amplified, and sequenced [20, 36]. Panning can be performed against cells or nanoparticles (semi-automated), which ensures the reactivity/selectivity of the antibodies. However, both are labor intensive and the latter must be restricted to a lesser number of samples, to remain manageable [20, 37]. Nevertheless, semi-automated is a robust, reproducible, and efficient method of panning [38]. Panning can be enhanced via the use of next generation sequencing (NGS), which can help eliminate unwanted clones, identify frequent sequences, and the reveal the evolution of the phages; this helps to reduce the number of rounds of panning [39–41]. However, binding affinities are not taken into account in this process. These methods can be used to select desirable phages from diverse libraries, producing sequences that can be tested in a much shorter timeframe than hybridomas, but this method is still tedious and laborious [3].

8.3 New Platforms for Single-Cell Antibody Discovery

8.3.1 Fluorescence Activated Cell Sorting

Fluorescence activated cell sorting (FACS) is an engineering adaption of flow cytometry, in which, cells are obtained or "sorted" based on fluorescent markers. The markers are commonly fluorescent-labeled antibodies against cell-specific proteins/receptors. While flow cytometry has only developed since the 1960s, when the first Coulter counter was produced, it has become a standard for identification of cell subsets. The original Coulter counter was based on the principle that the movement of a cell could be detected via changes in electrical signals as it passed through a microchannel. This has evolved over time, to the current flow cytometer which detects diffraction of signal through a series of detectors when a fluorescently labeled cell passes through a microchannel [42]. The number of detectors available for a flow cytometer (and therefore the number of different fluorescent wavelengths it can distinguish) varies widely,

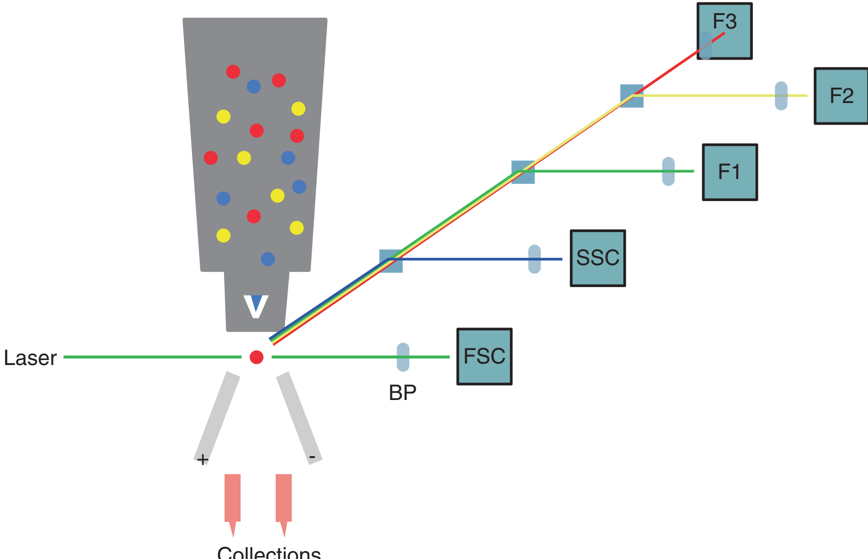


Fig. 8.2 Principles of fluorescence activated cell sorting (FACS). Fluorescence activated cell sorting (FACS) diagram. Yellow, blue, and red circles represent fluorescently-labelled cells in the flow cytometer, where, upon exiting, they are excited by a laser. Diffraction of the beam is measured through a series of detectors. Forward scatter (FSC) and side scatter (SSC) represent the size and complexity of the cell, respectively. Here, F1-F3 are the detectors and the tubes indicate user-defined positive and negative selection criteria

with the high end being able to distinguish nearly 20 signals. As shown in Fig. 8.2, as the cell passes through the microchannel, a laser excites the fluorescent molecule at its specific wavelength(s). The emitted signal passes through a series of bandpass filters (BP) and is able to be distinguished via the detectors. Sorting by FACS is initiated by applying a pulse of electricity to disrupt the droplet containing the cell, so that it is diverted into an appropriate receptacle. Sorting can be used restrict cell populations to those desired for subsequent experiments. Sorting for single cell is becoming a popular tool. For example, sorting of individual antigen-specific B cells were used to isolate HIV-specific antibodies. The high-throughput feature of FACS expedites the downstream applications, such as sequencing VDJ heavy and VJ light chains of single cell and, in this case, via transfection of these amplified DNA sequences into human kidney epithelial (HEK293) cells, produced monoclonal antibodies against an antigen [43]. This technique is heavily translational and can be used for a variety of diseases.

8.3.2 Microfluidic Devices

Microfluidic devices, or chips, started being used in biological applications in the late 1990s, where it was often used as a new immunoassay. The design of these early chips allowed for controlled release of specific reagents, in conjunction with the ability to mix reagents directly on the chip. Pore sizes/lengths and electrolytic buffers control flow rates of each of the components [44, 45] (Fig. 8.3). For example, if two reagents need to be mixed on the chip at a disproportionate ratio, the pore size of the lesser component can be made smaller, physically limiting the amount that can be mixed at a time. This protocol is useful when performing whole cell ELISAs, especially when a cell population is limited. While with a conventional ELISA, a significantly larger number of cells and more reagents are necessary. As it is common to replicate samples in duplicate or triplicate, the amounts of reagents can become astronomically higher, making this device an optimal example of high content cell screening (HSC) [46]. With microfluidic devices, the

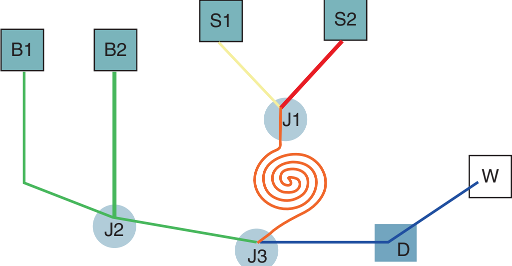


Fig. 8.3 Illustration of microfluidic device for antibody selection. Basic schematic of a microfluidic device. "B" indicates a buffer, "S" a sample, "J" a junction, "W" is waste, and "D" is a detector. Here, two samples join at junction 1, they then continue through a coil to mix. After exiting the coil, the cells pass through a separation microchannel where cells are further diluted by the buffers (at junction 3). Here they form a single cell suspension for detection before proceeding to waste

human error is removed; there is no pipetting error nor inconsistencies in plated densities, since these are significantly smaller sample sizes.

In the same fashion as the flow cytometer, this can be used to collect a pool of, or single, cells; it may also be used to detect the response or endogenous state of individual cells, e.g. cytokines produced. It should be noted that microfluidic devices encompass a wide range of processes and many different chips can be used. While the chip above is a basic schematic, many variations on this can be used based on the application. In fact, many laboratories design their own. For example, circulating tumor cells (CTCs) are rare, difficult to detect cells which are hypothesized to be the cause of metastatic cancers. A chip was designed for this which is composed simply of a series of posts coated with antibody against epithelial cell adhesion molecule (EpCAM), a common cancer marker. This antibody then captures any cancer cells which come into contact with them on the chip [47]. The CTC chip can be used to enumerate and evaluate this specific cell type, enabling purification from whole blood in a single step. Notably, while this is a fairly simple design for a chip, there were still many calculations and experiments necessary to optimize how cells can be adhered on this chip. The most important factors here were layout of the posts (including diameter and distance apart), the flow rate, where too high a rate would result in loss of cells, and shear stress, in which the cells would be lysed rendering them useless. By using this same layout, antibodies could be used for a myriad of capture chips. Once a device is optimized, a simple, streamlined process for isolation and/or characterization of single cells has been achieved. This technique allows the user to save both time and money, as the volumes of reagents necessary are quite small. The malleability of this technique to a specific protocol makes it one of the most useful devices, however the time and manpower necessary to establish a single technique may not be practical for some labs.

8.3.3 Single-Cell Antibody Nanowells (SCAN)

The technology was developed by Christopher Love and colleagues at MIT. SCAN is a soft lithographic technique that uses a dense array of nanowells (50 × 50 μm or 30 × 30 μm, holding a volume of 0.1–1 nl each) fabricated of polydimethylsiloxane (PDMS) to isolate individual cells for printing of corresponding molecules secreted by each cell. The array of nanowells is fabricated on standard 1″ × 3″ glass slides containing 84,672 wells for 50 μm size nanowells or 248,832 wells for 30 μm size. A capture slide coated, for example, with immunoglobulins (Igs) can be hybridized by placement on the top of the nanochip to capture the antibodies being secreted by the corresponding individual live B cells that are seeded in the nanowells (Fig. 8.4). Earlier works have shown that the nanowells with the rapid and high-throughput features were able to identify antigen-specific antibodies [48, 49]. With the capability that the single ex-vivo cell can be cultured and confined in each nanowell for an extended amount of time, it facilitates the recovery and clonal expansion of cells with specific engraved phenotype [50–52]. Using the arrays of nanowells with multiplexing capability, the Love group was able to examine the isotypes of the secreted antibodies, the specificity and relative affinity for HIV antigens, identify the reactive subset of B cells (memory and plasma B cells), and sequence/identify the genes encoding

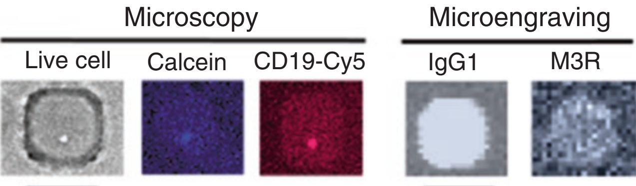


Fig. 8.4 An application of SCAN to identify anti-muscarinic acetylcholine receptor type- 3 (M3R) producing B cells. Representative array of nanowells with microscopic micrographs showing a live cell in bright field, calcium dye for live cell marker, and CD19 for B cell marker. Micro-engraving microarrays show the secretion of IgG1-isotypic anti-M3R antibody. Scale bar: 50 μm

the heavy and light chains [53]. Using this method, the group isolated HIV-specific neutralizing antibodies in colon biopsies [53]. This method is uniquely able to profile and isolate rare or low frequency B cells. The recent study by Tsioris et al. identified four novel West Nile virus (WNV) neutralizing antibodies in recently infected and post-convalescent subjects [54]. The most interesting aspect of the study was that given a low frequency of WNV-specific B cells (mean <24 events per 100,00 peripheral blood mononuclear cells), the group was able to identify some rare and potent neutralizing antibodies.

8.4 Application in Infectious Diseases

8.4.1 Human Immunodeficiency Virus (HIV)

According to UNAIDS/WHO, since the start of the HIV epidemic in the 1980s, worldwide 78 million people have become infected with HIV and 36 million people have died from HIV and AIDS-related diseases. As of 2016, 36.7 million people live with HIV. Combinations of highly active antiretroviral therapy (HAART) have been effective since their introduction in1996 and HIV-related mortality has been reduced since then, but remains above one million per year (1.1 in 2015 compared to 2 million in 2005), mainly due to insufficient access to screening and antiretroviral therapy in economically challenged countries which are often the most affected by the infection [55]. A recent meta-analysis shows that HIV-infected patients without access to HAART have a 2-year survival probability of progression from AIDS to AIDS-related death at 48% and the 6-year survival probability is 18%, whereas this life expectancy is 87% for the 2- year survival probability and 61% for 10-years survival probability for patients who received HAART [56]. An effective vaccine must achieve a production of protective antibodies against vaccine viral proteins. Due to an extensive genetic diversity of HIV, a prophylactic vaccine must provide global protection against all strains [57, 58]. Currently there are three principal research directions on HIV treatment and vaccine development using neutralizing antibodies: (1) activation of B cells by sequential immunogens for expression neutralizing antibodies, (2) development of novel neutralizing antibodies due to passive administration, and (3) vector-mediated gene transfer using adeno-associated virus vectors for delivery of HIV broadly neutralizing antibodies (bNAbs) and antibody-like proteins [59–61].

Human hybridoma, EBV transformation, FACS sorting of HIV-specific B cells, and combinatorial display technologies have been utilized in screening for single B cells that produce potent bNAbs. The interest in single cell antibody cloning has increased in the last few years due to advances in high-efficiency and throughput sequencing, which has reinvigorated studies on bNABs to obtain HIV-1 envelope-reactive antibodies [58, 62–64]. Initially, cloning from single cells was introduced to examine the development and silencing of autoreactive B cells [65]. This method was performed for identification of single B cells expressing antibodies [62, 65, 66] or to screen cultured B cells for the production of neutralizing activity [58, 67]. Single cells from HIV infected patients are isolated by FACS, then sequences of immunoglobulin genes isolated from each cell are cloned into a vector for protein expression. Obtained bNAbs are analyzed to understand their specificity, protective capacity, binding conformation, and reactivity breadth and potency. Usually, screening of monoclonal antibodies is utilized to elicit a clonal assessment of specificities present in HIV infected patients [68].

Passive administration of bNABs is advised for prevention and therapy of HIV infection. Studies on humans have proven safe and efficacious administration of monoclonal antibodies, yielding a promising approach of total control of HIV infection due to direct engagement in host immunity [69]. These bNABs must have high potency for HIV treatment with a capacity to reduce HIV viral load and minimize or prevent the risk of viral reactivation [59]. Pre-exposure prophylactic treatment has been studied in experiments with untreated non-human primate models infected with simian-human immunodeficiency virus (SHIV). Passive transfer or injection of HIV-1 bNABs protects host against viral infection [70–74]. A single bNABs infusion prevents virus acquisition with a single high dose [72, 75, 76] or repeated low doses SHIV infection; this protection can be up to 23 weeks depending on antibody potency and half-life [74]. Furthermore, introduction of a mutation in the fragment cytallizable (Fc) domain extends the antibody half-life median protection [74].

8.4.2 Emerging Arboviruses: Zika

Zika virus (ZIKV) infections are an emerging health pandemic of significant medical importance. The current outbreak has garnered attention by exhibiting unique characteristics of devastating neurodevelopmental defects in newborns of infected pregnant women [77, 78]. Over the past year, doctors in Brazil have documented over 4000 cases of microcephaly in which infants are born with abnormally small heads [79]. Detection of ZIKV in fetal brain tissues and anti-ZIKV antibodies in these mothers and/or infants established a possible causal link between ZIKV infection and this birth defect [80]. Typical symptoms of ZIKV infection include joint pain, fever, and rash. In addition, there is emerging a potential link to the dramatic increase in the reported cases of Guillain-Barré syndrome, another rare disorder of the peripheral nervous system characterized by muscle weakness and paralysis [81, 82]; in severe cases, Zika patients require life support. The spread of ZIKV has reached an alarming rate, particularly in the state of Florida. The influx of international travelers or tourists from ZIKV-infected areas, together with the warm tropical climate of the state, promotes the survival of the ZIKV-carrying mosquitoes, thus accelerating the spread of the virus. Responding to the Zika outbreak has been more than challenging. Unlike other well-known flaviviruses like Dengue, West Nile, Yellow Fever, and Japanese encephalitis viruses, there are no treatments or vaccinations, and diagnostic reagents are very limited. Although many investigations using immune-based therapies for arboviral infection have been pursued and have shown promise, there are no commercially available immune-based products for ZIKV. A better alternative would be to develop effective broadly neutralizing antibodies (bNAbs) as passive protection against ZIKV infection and more importantly prevent maternal-fetal transmission, reducing the likelihood of developing microcephaly in the newborns (Table 8.1) [67, 83–87].

As an emerging disease, there is a limited number of ZIKV monoclonal antibodies that are currently still at the testing phase (Table 8.2)

Table 8.1 Characterization of HIV broadly neutralizing antibodies

Epitope	Broadly neutralizing antibody	Breadth of neutralization with IC50<μg/mL[a]
CD4bs	VRC01	91% of 190 isolates[86]
		100% of 118 isolates representing major HIV-1 clades[87]
		93% of 162 isolates representing major HIV clades[67]
		89% of 178 isolates representing major HIV-1 clades[88]
	VRC02	91% of 190 isolates, representing major HIV-1 clades[86]
	VRC03	57% of 190 isolates, representing major HIV-1 clades[86]
	VRC-PG04 or PGV04	76% of 178 isolates, representing major HIV-1 clades[89]
		88% of 162 isolates representing major HIV clades[67]
	IgG1b12 or b12	41% of 190 isolates, representing major HIV-1 clades[86]
V1-V2 loops, quaternary structure	PG9	79% of 162 isolates representing major HIV-1 clades[88]
	ECD4-Ig	77% of 162 isolates representing major HIV clades[67]
V2-V3 loops, quaternary structure	PG16	73% of 162 isolates representing major HIV-1 clades[88]
	PGT145	78% of 162 isolates representing major HIV clades[67]
	PGT143	56% of 162 isolates representing major HIV clades[67]
gp120-gp41 quartenary interface	PGT151	66% of 117 cross-clade isolates[90]
V3 loop	PGT 131	40% of 162 isolates representing major HIV clades[67]

[a]Breadth of neutralization (percentage of viruses neutralized at IC50 > 1 μg/mL in a panel of 100–200 pseudoviruses. Data obtained from CATNAP, an online database hosted by the Los Alamos National Laboratory, USA (http://lanl.gov/catnap)

Table 8.2 Neutralizing antibodies against Zika virus

Neutralizing antibody	Viral unit	Epitope	Results
ZV-54[84]	Envelope subunit DIII	Lateral ridge	No cross-reactivity with DENV and/or JEV. Neutralization of 4 ZIKV strains *in-vitro.* Potency: 0.087-0.582 μg/mL
ZV-67[84]	Envelope subunit DIII	Lateral ridge	No cross-reactivity with DENV and/or JEV. Neutralization of 4 ZIKV strains *in-vitro.* Potency: 0.143-0.511 μl/mL
VH3-23/VK1-5[91]	Envelope subunit DIII	Lateral ridge	Recognition and neutralization of DENV-1 and ZIKV Potency: 0.7–4.6 ng/mL
ZV-64[84]	Envelope subunit DIII	C-C′ loop	No cross-reactivity with DENV and/or JEV. Reduced Inhibitory activity *in-vitro* against African and American strains
zEDIII[92]	Premembrane- envelope EDIII		Recognition and neutralization of ZIKV. No exacerbation of DENV infection
C10[93]	Envelope subunits DI, DII (near from fusion loop), DIII	Intradimer	Recognition and neutralization of ZIKV *in-vitro* and *in-vivo*
ZA10G6[94]	Envelope subunit DIII	Fusion loop	Recognition and neutralization of ZIKV *in-vitro*
ZV-2[84]	Envelope subunit DIII	ABDE sheet	No cross-reactivity with DENV and/or JEV. Neutralization of 4 ZIKV strains *in-vitro*
ZV-13[84]	Envelope subunit DI-II	Fusion loop	Cross-reactivity with DENV-1, 2, 3, 4, WNV, and JEV. No inhibitory activity *in-vitro*

Potency is the measure by the IC50 (μg/mL) in a panel of 100–200 pseudoviruses *DENV* Dengue virus, *JEV* Japanese encephalitis virus

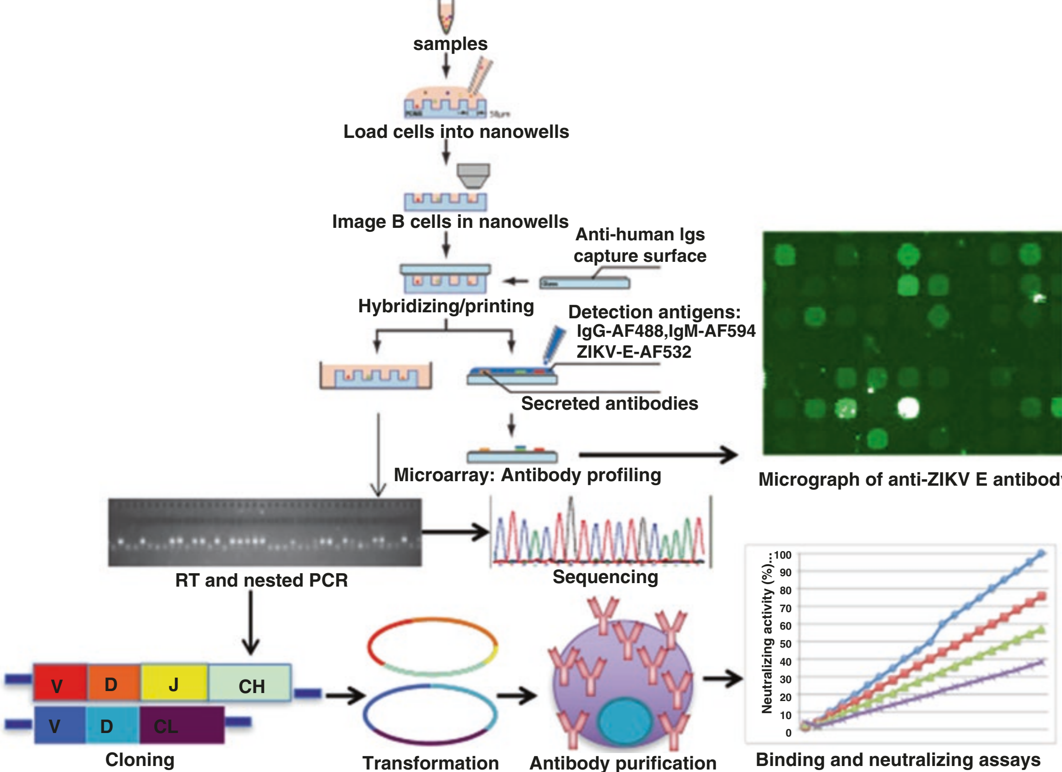


Fig. 8.5 Screening for ZIka virus neutralizing antibodies using SCAN. Peripheral blood cells of Zika infected patients will be isolated. Purified single-cell suspension will be labeled with anti-CD20-FITC and Calcein violet-405 for live cell and plated onto fabricated nanowells. Labeled cells in the nanowells will be imaged for surface markers and locations on the chips. Capture slide coated with anti-human immunoglobulins will be hybridized. Detection antibody mixture containing IgG-AF-488, IgM-AF594, and ZIKV E-AF532 will be added. Micrograph of anti-ZIKV E-secreting B cells will be generated. Individual ZIKV E-secreting B cells will be picked and performed RT/nested PCR for heavy/light chain sequences. Both chains will be cloned into an expression vector and expressed in 293 T cell line. Secreted antibodies will be purified and screened for binding and neutralizing activity against ZIKV. AF: Alexa Fluor

[88–92]. Using EBV–immortalized memory B cells that were reactive to ZIKV NS1 or E proteins, Stettler et al. have identified 119 bNAbs capable of neutralizing ZIKV. The authors have shown that the most potent neutralizing antibodies were ZIKV-specific and targeted EDIII or quaternary epitopes [93]. Using tradition hybridoma technology in the mouse, Zhao et al. isolated six mAbs that recognized ZIKV evelope (E) protein after screening more than 2000 hybridomas [88]. A recent study by Sapparapu et al. demonstrated that EBV-transformed ZIKV-specific B cells exhibited potent neutralizing capacity. Epitope mapping using X-ray crystallography indicated that the most effective bNAb recognized a unique quaternary epitope on the E protein dimer–dimer interface. Further studies showed the therapeutic efficacy in pregnant and non-pregnant mice in which mAb treatment markedly reduced tissue pathology, placental and fetal infection, and mortality in mice [94]. Future studies using single cell selection as proposed in Fig. 8.5 will generate a complete repertoire of ZIKVspecific antibodies, develop better bNAbs and reveal essential epitopes for future structure-based vaccine design.

8.5 Conclusion

Single-cell analysis is a powerful tool in examining a comprehensive repertoire of antigen-specific Abs from the most abundant to the least abundant B cells that are highly specific. Single-cell antibody discovery is critically important in selecting the few potent B cells with important capacity to produce the most competent therapeutic mAbs and broadly capable of neutralizing pathogens in infected individuals. Diseases in which vaccines are not readily available or effective, therapeutic mAbs can provide significant protection as passive immunity. The two quintessential examples are HIV and Zika as discussed. These technologies, while strong and important tools currently, have the potential to become widely utilized and even more powerful. They have the potential to be used in diagnostics and beyond that, these techniques are currently being used to develop treatments for other infectious diseases and cancer. In conjunction with shotgun mutagenesis and X-ray crystallography, antigenic epitopes can be mapped and the structural interactions between Abs and antigens can be examined. On a more fundamental level, single-cell analysis will be an essential player in creating immune therapeutics and eventually vaccines.

Acknowledgements This work was supported by the Florida Department of Health, Biomedical Research Program (JKY, CQN) and NIAID 1R21AI130561-01A1 (CQN).

References

1. Kohler G, Milstein C (1975) Continuous cultures of fused cells secreting antibody of predefined specificity. Nature 256:495–497
2. Ecker DM, Jones SD, Levine HL (2015) The therapeutic monoclonal antibody market. MAbs 7:9–14. https://doi.org/10.4161/19420862.2015.989042
3. Fitzgerald V, Leonard P (2017) Single cell screening approaches for antibody discovery. Methods 116:34–42. https://doi.org/10.1016/j.ymeth.2016.11.006
4. Ortho Multicenter Transplant Study, G (1985) A randomized clinical trial of OKT3 monoclonal antibody for acute rejection of cadaveric renal transplants. N Engl J Med 313:337–342. https://doi.org/10.1056/NEJM198508083130601
5. Brekke OH, Sandlie I (2003) Therapeutic antibodies for human diseases at the dawn of the twenty-first century. Nat Rev Drug Discov 2:52–62. https://doi.org/10.1038/nrd984
6. Sun LK et al (1987) Chimeric antibody with human constant regions and mouse variable regions directed against carcinoma-associated antigen 17-1A. Proc Natl Acad Sci U S A 84:214–218
7. Jones PT, Dear PH, Foote J, Neuberger MS, Winter G (1986) Replacing the complementarity-determining regions in a human antibody with those from a mouse. Nature 321:522–525. https://doi.org/10.1038/321522a0
8. Love KR, Bagh S, Choi J, Love JC (2013) Microtools for single-cell analysis in biopharmaceutical development and manufacturing. Trends Biotechnol 31:280–286. https://doi.org/10.1016/j.tibtech.2013.03.001
9. Aggarwal SR (2012) What's fueling the biotech engine-2011 to 2012. Nat Biotechnol 30:1191–1197. https://doi.org/10.1038/nbt.2437
10. Goodman M (2009) Market watch: sales of biologics to show robust growth through to 2013. Nat Rev Drug Discov 8:837. https://doi.org/10.1038/nrd3040
11. Urban PL et al (2011) Carbon-13 labelling strategy for studying the ATP metabolism in individual yeast cells by micro-arrays for mass spectrometry. Mol BioSyst 7:2837–2840. https://doi.org/10.1039/C1MB05248A
12. Nemes P, Knolhoff AM, Rubakhin SS, Sweedler JV (2011) Metabolic differentiation of neuronal phenotypes by single-cell capillary electrophoresis–electrospray ionization-mass spectrometry. Anal Chem 83:6810–6817. https://doi.org/10.1021/ac2015855
13. Kuppers R (2003) B cells under influence: transformation of B cells by Epstein-Barr virus. Nat Rev Immunol 3:801–812. https://doi.org/10.1038/nri1201
14. Wang S (2011) Advances in the production of human monoclonal antibodies. Antibody Technol J 2011:4. https://doi.org/10.2147/ANTI.S20195
15. Yamashita M, Katakura Y, Shirahata S (2007) Recent advances in the generation of human monoclonal antibody. Cytotechnology 55:55–60. https://doi.org/10.1007/s10616-007-9072-5
16. Reichert JM (2001) Monoclonal antibodies in the clinic. Nat Biotechnol 19:819–822. https://doi.org/10.1038/nbt0901-819
17. Reichert JM, Rosensweig CJ, Faden LB, Dewitz MC (2005) Monoclonal antibody successes in the clinic. Nat Biotechnol 23:1073–1078. https://doi.org/10.1038/nbt0905-1073
18. Hansel TT, Kropshofer H, Singer T, Mitchell JA, George AJ (2010) The safety and side effects of monoclonal antibodies. Nat Rev Drug Discov 9:325–338. https://doi.org/10.1038/nrd3003
19. Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228:1315–1317
20. Lim BN et al (2014) Principles and application of antibody libraries for infectious diseases. Biotechnol Lett 36:2381–2392. https://doi.org/10.1007/s10529-014-1635-x

21. Thie H, Meyer T, Schirrmann T, Hust M, Dubel S (2008) Phage display derived therapeutic antibodies. Curr Pharm Biotechnol 9:439–446
22. Miersch S, Sidhu SS (2012) Synthetic antibodies: concepts, potential and practical considerations. Methods 57:486–498. https://doi.org/10.1016/j.ymeth.2012.06.012
23. Pansri P, Jaruseranee N, Rangnoi K, Kristensen P, Yamabhai M (2009) A compact phage display human scFv library for selection of antibodies to a wide variety of antigens. BMC Biotechnol 9:6. https://doi.org/10.1186/1472-6750-9-6
24. Vaughan TJ et al (1996) Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. Nat Biotechnol 14:309–314. https://doi.org/10.1038/nbt0396-309
25. Willats WG, Gilmartin PM, Mikkelsen JD, Knox JP (1999) Cell wall antibodies without immunization: generation and use of de-esterified homogalacturonan blockspecific antibodies from a naive phage display library. Plant J 18:57–65
26. Ayat H et al (2013) Isolation of scFv antibody fragments against HER2 and CEA tumor antigens from combinatorial antibody libraries derived from cancer patients. Biologicals 41:345–354. https://doi.org/10.1016/j.biologicals.2013.05.004
27. Braunagel M, Little M (1997) Construction of a semisynthetic antibody library using trinucleotide oligos. Nucleic Acids Res 25:4690–4691
28. Andris-Widhopf J, Rader C, Steinberger P, Fuller R, Barbas CF 3rd. (2000) Methods for the generation of chicken monoclonal antibody fragments by phage display. J Immunol Methods 242:159–181
29. Low NM, Holliger PH, Winter G (1996) Mimicking somatic hypermutation: affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. J Mol Biol 260:359–368
30. Sblattero D, Bradbury A (2000) Exploiting recombination in single bacteria to make large phage antibody libraries. Nat Biotechnol 18:75–80. https://doi.org/10.1038/71958
31. Christensen PA et al (2009) Modifying antibody specificity by chain shuffling of V/V between antibodies with related specificities. Scand J Immunol 69:1–10. https://doi.org/10.1111/j.1365-3083.2008.02164.x
32. Fujii R, Kitaoka M, Hayashi K (2004) One-step random mutagenesis by error-prone rolling circle amplification. Nucleic Acids Res 32:e145. https://doi.org/10.1093/nar/gnh147
33. Holland EG et al (2013) AXM mutagenesis: an efficient means for the production of libraries for directed evolution of proteins. J Immunol Methods 394:55–61. https://doi.org/10.1016/j.jim.2013.05.003
34. Hammers CM, Stanley JR (2014) Antibody phage display: technique and applications. J Invest Dermatol 134:e17. https://doi.org/10.1038/jid.2013.521
35. Schirrmann T, Meyer T, Schutte M, Frenzel A, Hust M (2011) Phage display for the generation of antibodies for proteome research, diagnostics and therapy. Molecules 16:412–426. https://doi.org/10.3390/molecules16010412
36. Fukushi S et al (2012) Antigen-capture ELISA for the detection of Rift Valley fever virus nucleoprotein using new monoclonal antibodies. J Virol Methods 180:68–74. https://doi.org/10.1016/j.jviromet.2011.12.013
37. de Kruif J, Terstappen L, Boel E, Logtenberg T (1995) Rapid selection of cell subpopulation-specific human monoclonal antibodies from a synthetic phage antibody library. Proc Natl Acad Sci U S A 92:3938–3942
38. Turunen L, Takkinen K, Soderlund H, Pulli T (2009) Automated panning and screening procedure on microplates for antibody generation from phage display libraries. J Biomol Screen 14:282–293. https://doi.org/10.1177/1087057108330113
39. Dias-Neto E et al (2009) Next-generation phage display: integrating and comparing available molecular tools to enable cost-effective high-throughput analysis. PLoS One 4:e8338. https://doi.org/10.1371/journal.pone.0008338
40. Ravn U et al (2010) By-passing in vitro screening--next generation sequencing technologies applied to antibody display and in silico candidate selection. Nucleic Acids Res 38:e193. https://doi.org/10.1093/nar/gkq789
41. t Hoen PA et al (2012) Phage display screening without repetitious selection rounds. Anal Biochem 421:622–631. https://doi.org/10.1016/j.ab.2011.11.005
42. Fulwyler MJ (1965) Electronic separation of biological cells by volume. Science 150:910–911
43. Smith K et al (2009) Rapid generation of fully human monoclonal antibodies specific to a vaccinating antigen. Nat Protoc 4:372–384. https://doi.org/10.1038/nprot.2009.3
44. Chiem NH, Harrison DJ (1998) Microchip systems for immunoassay: an integrated immunoreactor with electrophoretic separation for serum theophylline determination. Clin Chem 44:591–598
45. Wheeler AR et al (2003) Microfluidic device for single-cell analysis. Anal Chem 75:3581–3586
46. Cheong R, Paliwal S, Levchenko A (2010) High-content screening in microfluidic devices. Expert Opin Drug Discov 5:715–720. https://doi.org/10.1517/17460441.2010.495116
47. Nagrath S et al (2007) Isolation of rare circulating tumour cells in cancer patients by microchip technology. Nature 450:1235–1239. https://doi.org/10.1038/nature06385
48. Ogunniyi AO, Story CM, Papa E, Guillen E, Love JC (2009) Screening individual hybridomas by microengraving to discover monoclonal antibodies. Nat Protoc 4:767–782. https://doi.org/10.1038/nprot.2009.40
49. Ronan JL, Story CM, Papa E, Love JC (2009) Optimization of the surfaces used to capture antibodies from single hybridomas reduces the time required for microengraving. J Immunol Methods 340:164–169. https://doi.org/10.1016/j.jim.2008.10.018
50. Love JC, Ronan JL, Grotenbreg GM, van der Veen AG, Ploegh HL (2006) A microengraving method for rapid selection of single cells producing antigen-specific antibodies. Nat Biotechnol 24:703–707. https://doi.org/10.1038/nbt1210

51. Esfandiary L et al (2016) Single-cell antibody nanowells: a novel technology in detecting anti-SSA/Ro60- and anti-SSB/La autoantibody-producing cells in peripheral blood of rheumatic disease patients. Arthritis Res Ther 18:107. https://doi.org/10.1186/s13075-016-1010-5
52. Nguyen CQ, Ogunniyi AO, Karabiyik A, Love JC (2013) Single-Cell analysis reveals Isotype-specific autoreactive B cell repertoires in Sjögren's syndrome. PLoS One 8:e58127. https://doi.org/10.1371/journal.pone.0058127
53. Ogunniyi AO et al (2014) Profiling human antibody responses by integrated single-cell analysis. Vaccine 32:2866–2873. https://doi.org/10.1016/j.vaccine.2014.02.020
54. Tsioris K et al (2015) Neutralizing antibodies against West Nile virus identified directly from human B cells by single-cell analysis and next generation sequencing. Integr Biol (Camb) 7:1587–1597. https://doi.org/10.1039/c5ib00169b
55. Mahy M et al (2017) Producing HIV estimates: from global advocacy to country planning and impact measurement. Glob Health Action 10:1291–1169. https://doi.org/10.1080/16549716.2017.1291169
56. Poorolajal J, Hooshmand E, Mahjub H, Esmailnasab N, Jenabi E (2016) Survival rate of AIDS disease and mortality in HIV-infected patients: a meta-analysis. Public Health 139:3–12. https://doi.org/10.1016/j.puhe.2016.05.004
57. Lelievre JD, Levy Y (2016) HIV-1 prophylactic vaccines: state of the art. J Virus Erad 2:5–11
58. Klein F et al (2013) Antibodies in HIV-1 vaccine development and therapy. Science 341:1199–1204. https://doi.org/10.1126/science.1241144
59. Ahmad M, Ahmed OM, Schnepp B, Johnson PR (2017) Engineered expression of broadly neutralizing antibodies against human immunodeficiency virus. Annu Rev Virol. https://doi.org/10.1146/annurev-virology-101416-041929
60. Brady JM, Baltimore D, Balazs AB (2017) Antibody gene transfer with adeno-associated viral vectors as a method for HIV prevention. Immunol Rev 275:324–333. https://doi.org/10.1111/imr.12478
61. Deal CE, Balazs AB (2015) Vectored antibody gene delivery for the prevention or treatment of HIV infection. Curr Opin HIV AIDS 10:190–197. https://doi.org/10.1097/COH.0000000000000145
62. Scheid JF et al (2009) A method for identification of HIV gp140 binding memory B cells in human blood. J Immunol Methods 343:65–67. https://doi.org/10.1016/j.jim.2008.11.012
63. Tiller T et al (2008) Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. J Immunol Methods 329:112–124. https://doi.org/10.1016/j.jim.2007.09.017
64. Wardemann H et al (2003) Predominant autoantibody production by early human B cell precursors. Science 301:1374–1377. https://doi.org/10.1126/science.1086907
65. Scheid JF et al (2009) Broad diversity of neutralizing antibodies isolated from memory B cells in HIV-infected individuals. Nature 458:636–640. https://doi.org/10.1038/nature07930
66. Klein F et al (2012) Broad neutralization by a combination of antibodies recognizing the CD4 binding site and a new conformational epitope on the HIV-1 envelope protein. J Exp Med 209:1469–1479. https://doi.org/10.1084/jem.20120423
67. Walker LM et al (2011) Broad neutralization coverage of HIV by multiple highly potent antibodies. Nature 477:466–470. https://doi.org/10.1038/nature10373
68. Ho IY et al (2016) Refined protocol for generating monoclonal antibodies from single human and murine B cells. J Immunol Methods 438:67–70. https://doi.org/10.1016/j.jim.2016.09.001
69. Caskey M et al (2015) Viraemia suppressed in HIV-1-infected humans by broadly neutralizing antibody 3BNC117. Nature 522:487–491. https://doi.org/10.1038/nature14411
70. Baba TW et al (2000) Human neutralizing monoclonal antibodies of the IgG1 subtype protect against mucosal simian-human immunodeficiency virus infection. Nat Med 6:200–206. https://doi.org/10.1038/72309
71. Hessell AJ et al (2009) Effective, low-titer antibody protection against low-dose repeated mucosal SHIV challenge in macaques. Nat Med 15:951–954. https://doi.org/10.1038/nm.1974
72. Moldt B et al (2012) Highly potent HIV-specific antibody neutralization in vitro translates into effective protection against mucosal SHIV challenge in vivo. Proc Natl Acad Sci U S A 109:18921–18925. https://doi.org/10.1073/pnas.1214785109
73. Mascola JR et al (1999) Protection of Macaques against pathogenic simian/human immunodeficiency virus 89.6PD by passive transfer of neutralizing antibodies. J Virol 73:4009–4018
74. Gautam R et al (2016) A single injection of anti-HIV-1 antibodies protects against repeated SHIV challenges. Nature 533:105–109. https://doi.org/10.1038/nature17677
75. Shingai M et al (2014) Passive transfer of modest titers of potent and broadly neutralizing anti-HIV monoclonal antibodies block SHIV infection in macaques. J Exp Med 211:2061–2074. https://doi.org/10.1084/jem.20132494
76. Saunders KO et al (2015) Sustained delivery of a broadly neutralizing antibody in nonhuman Primates confers long-term protection against simian/human immunodeficiency virus infection. J Virol 89:5895–5903. https://doi.org/10.1128/JVI.00210-15
77. Araujo AQ, Silva MT, Araujo AP (2016) Zika virus-associated neurological disorders: a review. Brain 139:2122–2130. https://doi.org/10.1093/brain/aww158
78. Baud D, Gubler DJ, Schaub B, Lanteri MC, Musso D (2017) An update on Zika virus infection. Lancet. https://doi.org/10.1016/S0140-6736(17)31450-2

79. Schuler-Faccini L et al (2016) Possible association between Zika Virus Infection and Microcephaly – Brazil, 2015. MMWR Morb Mortal Wkly Rep 65:59–62. https://doi.org/10.15585/mmwr.mm6503e2
80. Reagan-Steiner S et al (2017) Evaluation of placental and fetal tissue specimens for Zika Virus Infection – 50 States and District of Columbia, January–December, 2016. MMWR Morb Mortal Wkly Rep 66:636–643. https://doi.org/10.15585/mmwr.mm6624a3
81. Oehler E et al (2014) Zika virus infection complicated by Guillain-Barre syndrome—case report, French Polynesia, December 2013. Euro surveillance: bulletin Europeen sur les maladies transmissibles = Eur Commun Dis Bull 19:4–6
82. Frontera JA, da Silva IR (2016) Zika getting on your nerves? The association with the Guillain-Barre syndrome. N Engl J Med 375:1581–1582. https://doi.org/10.1056/NEJMe1611840
83. Wu X et al (2010) Rational design of envelope identifies broadly neutralizing human monoclonal antibodies to HIV-1. Science 329:856–861. https://doi.org/10.1126/science.1187659
84. Scheid JF et al (2011) Sequence and structural convergence of broad and potent HIV antibodies that mimic CD4 binding. Science 333:1633–1637. https://doi.org/10.1126/science.1207227
85. McLellan JS et al (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature 480:336–343. https://doi.org/10.1038/nature10696
86. Wu X et al (2011) Focused evolution of HIV-1 neutralizing antibodies revealed by structures and deep sequencing. Science 333:1593–1602. https://doi.org/10.1126/science.1207532
87. Bonsignori M et al (2012) Two distinct broadly neutralizing antibody specificities of different clonal lineages in a single HIV-1-infected donor: implications for vaccine design. J Virol 86:4688–4692. https://doi.org/10.1128/JVI.07163-11
88. Zhao H et al (2016) Structural basis of Zika Virus-specific antibody protection. Cell 166:1016–1027. https://doi.org/10.1016/j.cell.2016.07.020
89. Robbiani DF et al (2017) Recurrent potent human neutralizing antibodies to Zika Virus in Brazil and Mexico. Cell 169:597–609. e511. https://doi.org/10.1016/j.cell.2017.04.024
90. Yang M, Dent M, Lai H, Sun H, Chen Q (2017) Immunization of Zika virus envelope protein domain III induces specific and neutralizing immune responses against Zika virus. Vaccine 35:4287–4294. https://doi.org/10.1016/j.vaccine.2017.04.052
91. Zhang S et al (2016) Neutralization mechanism of a highly potent antibody against Zika virus. Nat Commun 7:13679. https://doi.org/10.1038/ncomms13679
92. Dai L et al (2016) Structures of the Zika Virus envelope protein and its complex with a Flavivirus broadly protective antibody. Cell Host Microbe 19:696–704. https://doi.org/10.1016/j.chom.2016.04.013
93. Stettler K et al (2016) Specificity, cross-reactivity, and function of antibodies elicited by Zika virus infection. Science 353:823–826. https://doi.org/10.1126/science.aaf8505
94. Sapparapu G et al (2016) Neutralizing human antibodies prevent Zika virus replication and fetal disease in mice. Nature. https://doi.org/10.1038/nature20564

Check for updates

9 Single Cell Genetics and Epigenetics in Early Embryo: From Oocyte to Blastocyst

Yubao Wei, Helong Zhang, Qian Wang, and Cuilian Zhang

Abstract

Single cell technology has enormously changed the landscape of biomedical science, including single cell omics, gene editing, single cell imaging, single cell (embryo) manipulate, or non-invasive micro-test. Single cell technology also leads the research area of early embryo from basic research to reproductive medical application. We got the knowledge of programming/reprogramming and the epigenetics dynamics in the cell lineage differentiation. In the reproductive medicine, the genomic sequencing of embryo or polar body and the preimplantation genetic diagnosis rely on the single cell techniques. Those discoveries will improve the assisted reproductive technologies, human health, and livestock husbandry. In the future, the comprehensive atlas of cell state and lineage information can be generated for cellular systems by single-cell multi-omics approaches.

Keywords

Single cell · Early embryo · Genetics · Epigenetics · Environment

Y. Wei · H. Zhang · Q. Wang · C. Zhang (✉)
The Reproductive Institute of Henan Provincial People's Hospital/Reproductive Hospital of Henan Province, Zhengzhou, China

9.1 Introduction

For the last two decades, we have got enormous knowledge of early embryo development. For early embryo, it has the dynamic character of morphological and genetic/epigenetic changes in the temporal and spatial. New investigation is performed in the embryo with a single cell resolution of morphology or "omics". The cell fate programming is the key question for embryology or cell biology, of which the totipotent is lost in the 2-cell to 4-cell stage [1]. The microstructure of embryo cell is "seen" more and more clearly. We can also study the physiology of early embryo: the osmotic pressure, the mechanic, the metabolism and so on. At the same time, we can also use mitochondrial transfer, genome editing, and other techniques to manipulate the early embryo.

Scientists verified single cell transcriptome is feasible in 1990; the cDNA was amplified in exponential rate [2]. Now even recombinant rate in sperm could be measured by single cell analysis [3]. Single cell genome sequencing including exome sequencing, or whole genome sequencing. Single cell transcriptome sequencing including single cell digital gene expression, single cell polyA sequencing, single cell lncRNA sequencing. 2012, the Smart-seq tech was discovered and has the power of covering the whole transcriptome almost without no gap [4].

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_9

Epigenetics has got more and more attention than before. Now we have got the landscape of methylation, acetylation, RNA, and so on. Every modification has different functions and structures, and evolutionary meaning such as genome imprinting. Many gene functions in early embryo are defined by Knockout/Knock in or gene editing technology. We could get the whole RNA numbers and classification by the single cell transcriptome. The embryo has fruitful RNAs since gamete formation.

To the preimplantation genetic diagnosis/Screening, little cells are captured and the DNA is isolated for DNA library construction, then was sequenced by next generation sequencing technology. We know more and more chromosome structure and DNA modification of early embryo. The chromosome structure of early embryo is very different from body cell such as condensed chromatin, different modification, methylation pattern, RNA content and so on. 2014 Sept. a preimplantation genetic diagnosis baby applied multiple annealing and looping-based amplification cycles (MALBAC) was born in Beijing, China which prevents the monogenic diseases [5]. Then in Mar. 2016, another MALBAC-based preimplantation genetic diagnosis baby was born in Wuxi, China [6].

The relationship between early embryo and the environmental factors, there is a “Barker’s hypothesis” or Developmental Origins of Health Disease theory. This theory was emerging at the 1980s as Barker and colleagues [7]. It proposed that under nutrition during gestation was the early origin of adult cardiovascular or metabolic disorders due to fetal programming that permanently shaped the body’s structure, function, and metabolism and contributed to adult disease [8] (Figs. 9.1, 9.2, and 9.3).

9.2 Single Cell Isolation Techniques

There are seven approaches of single cell isolation currently: micro-manipulation, laser capture microdissection, serial dilution, Fluorescence-activated cell sorting, microfluidics, Optical tweezers, and manual picking. After the single cell isolated, the DNA/RNA could be amplified by PCR/RT-PCR. Degenerate oligonucleotide-primed PCR or MALBAC are two methods avoiding amplifying bias. The PCR products could be used for next generation sequencing. Next generation sequencing gives the foundation of single cell sequencing. Microfluidic approach also prevalent in single cell research. It has the advantages as little sample amount required, high-throughput performance etc. The majorities techniques are droplet-based microfluidic, hydrodynamic trap, magnetic trap, acoustic trap, dielectrophoretic trap and optical trap. These techniques could be used for single cell isolation, cell culture, cell manipulation, cell mechanics and so on [9].

By MALBAC, scientists can detect single-nucleotide and copy-number variations on a single cell level. It was applied in the biopsy of preimplantation embryo or the polar body sequencing. MALBAC is a quasi-linear amplification which could reduce the bias associated with nonlinear amplification. Picograms of DNA fragments (~10 to 100 kb) from a single cell or few cells serve as templates [10]. In the last 5 years, MALBAC have made much progress in early embryo single-cell sequencing. For the single cell isolation techniques, fluorescence activated cell sorting, Flow cytometry, laser microdissection, manual cell picking, random, seeding/dilution, and microfluidics/lab-on-a-chip devices are mostly relevant technologies by the market surveys. At the same time, there are more than 20 parents in this field [11].

9.3 Single Cell Techniques for Early Embryo

Major epigenetic modifications such as are occurred during early development. Nucleocytoplasmic hybrids as transplanting a pronuclear into a recipient egg, the genotype did not consist with the phenotype in expectation as the methylation in the paternal or maternal genome is different, or partly because of the

	oocyte	1-cell	2-cell	4-cell	8-cell	morula	blastocyst
Optical	✓	✓	✓	✓	✓	✓	✓
Time-lapse		✓	✓	✓	✓	✓	✓
ICSI	✓						
PGD/PGS					✓		✓
PB genome	✓						
Mito transfer		✓					
Gene editing		✓					
Lamanspectrum		✓	✓	✓	✓	✓	✓
Microfluidics		✓	✓	✓	✓	✓	✓
Metabolome	✓	✓	✓	✓	✓	✓	✓

Fig. 9.1 Single cell techniques and its application for the preimplantation embryo. *ICSI* intracytoplasmic sperm injection, *PGD/PGS* preimplantation genetic diagnosis/screening, *PB genome* polar body genome

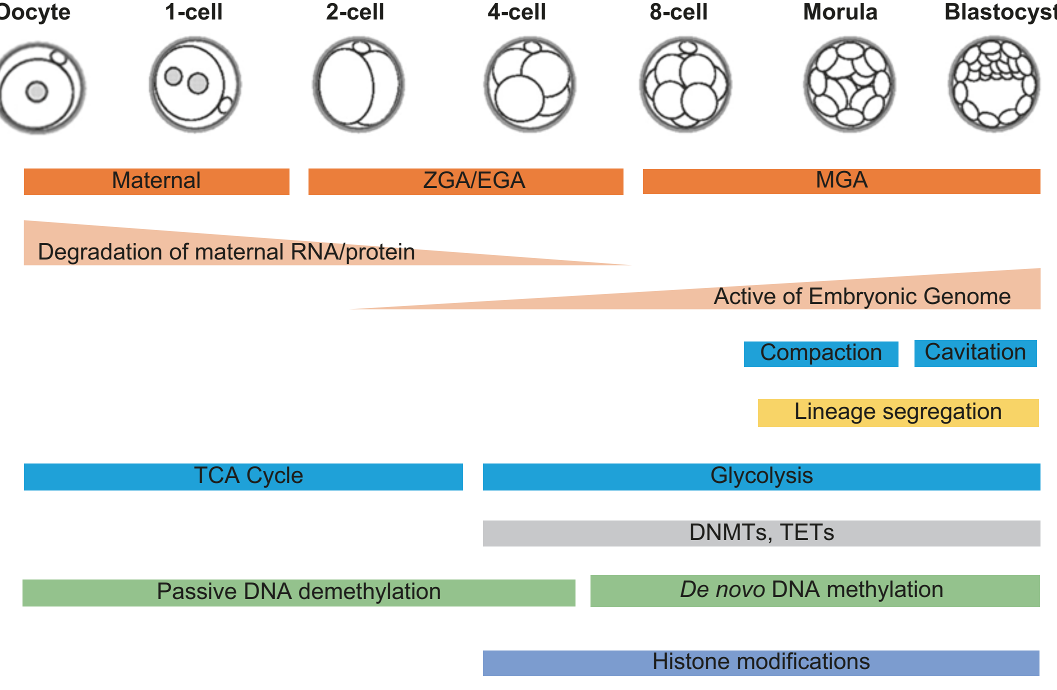


Fig. 9.2 Key development events in preimplantation embryos. *ZGA* zygotic genome activation, *EGA* embryonic genome activation, *MGA* mid-preimplantation gene activation, *TCA Cycle* tricarboxylic acid cycle, *DNMTs* DNA methyltransferases, *Tets* ten-eleven translocation enzymes

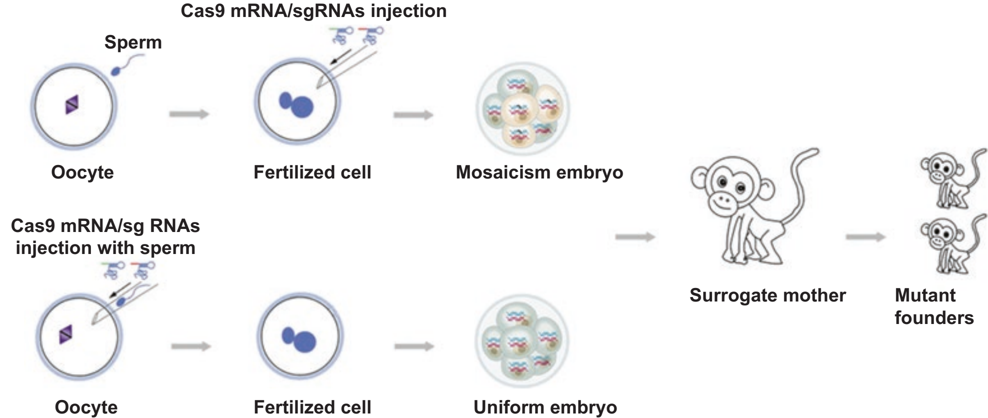


Fig. 9.3 Principle of gene editing on the early embryo. Gene editing could be carried out at fertilization stage or at zygote stage. Then it come to the mosaicism embryos or uniform embryos. The two type of embryos could be transferred to the surrogate and in gestation, then born the chimera founders or full mutant founders

nucleocytoplasmic interactions of the hybrid zygotes [12]. hMass cytometry could measure over 40 cellular parameters at single-cell resolution simultaneously. Now the cellular subsets could be discerned and analysis in a single experiment [13]. New techniques will give novel insights. Single cell RNA sequencing significantly improved transcriptome quantification of the preimplantation embryo at the individual transcript and also system levels. By bioinformatics analysis, stage-specific modules in core gene networks are operating in the transitional stage. It enabled base-resolution scrutiny without confounding effects as cell population heterogeneity [14].

Single cell quantitative PCR (qPCR), mRNA in situ hybridization, and single cell RNA sequencing are three common methods to measure single cell gene expression. Single cell transcriptomics is still in developing and have not yet reached the full potential [15]. The relation between transcriptome and genome could be detected by microfluidic single cell qPCR which could associate single nucleotide polymorphisms with gene expression phenotypes [16]. Carrier ChIP (CChIP) permit analysis few cells as 100 cells compared with the normal ChIP as a 10^7 cell. By this tech, the regulated gene such as “Yamanaka factor” could be detected in early embryo such as inner cell mass or trophectoderm cell lineages [17].

Using polyA-PCR amplification and microarray technology, the gene expression profiles of the human oocyte, 4-cell and blastocyst were investigated. The transcript profiles are stage-specific and variable within each stage. The pathway like cell cycle, nutrient metabolism, and apoptosis is variably expressed and could be applied as the marker of embryonic developmental competence. Global transcript profiles reflect the Heterogeneity of human embryo developmental competence [18]. By Single-cell mass spectrometry, scientists discovered dozens of small molecules for cell fates determination in 16-cell embryo [19]. Changing the metabolite concentration caused cell phenotype changes in cell movements at gastrulation or the tissue fates. There is a quantitative, single-cell based quantitative chromosome transmission fidelity assay for measuring chromosome transmission fidelity in yeast or other multi-cellular organisms in recent years [20]. Single cell DNA methylation analysis reveals epigenetic chimerism in preimplantation embryos. It has the diagnostic and potential therapeutic values, especially for the epimutation-based imprinting syndromes [21].

To multi-omics techniques, if use antibodies for protein detection or enrichment will be limited by the affinity characteristics of the reagents. In contrast, sequencing technology can parallelly interpret more than one analyze. Pacific Biosciences and nanopore sequencers can detect

DNA modifications [22, 23]. Both the two technologies are capable of getting the information of RNA sequences and protein modifications [24], and other analytics like microRNAs, peptides, and other small molecules [25]. Genome-wide survey of accessible chromatin in the preimplantation embryos using the assay for transposase-accessible chromatin using sequencing is available [26]. Gene activation of open chromatin could occur through different pathways from epigenetic modification reprogramming. Fully dissect epigenomic reprogramming of preimplantation embryo is are warranted in the near future [26]. 5hmC is very important for early embryo methylation pattern and genomic imprinting establishment. 5hmC and 5hC are the major methylation unit of DNA methylation modification. In the mammalian zygote, 5hmC and Tet3 are the key roles for DNA methylation which 5hC converted to 5hmC. If the nuclear acid base was protected by *PGC7/Dppa3/Stella*, then this mechanism could not take place [27].

The highly conserved noncoding elements in mammalian genomes show the principle that the initial epigenetic state in cells are defined by DNA sequence, lineage-specific gene expression and environmental cues give the subsequent alternation [28]. To highly conserved noncoding elements, it is with the feature of "bivalent domains,' consisting of large regions of H3K27 methylation and H3K4 methylation. Coincide with TF genes, Bivalent domains tend to expressed at low levels. The fundamental unit of chromatin structure in all eukaryotes is the nucleosome. Even Semiotics are used to study the epigenetic code. Semiotics system generally consisted by a sign, its meaning (production) and the code. Such as DNA methylation on imprinting control region of certain imprinted genes is the sign, Gene expression or not is the meaning, and it belongs to epigenetics code which the adapter is the DNMTs or Tets. The genetic/epigenetics code meets fundamental requirements of a semiotic system [29]. This will give a paradigm shift for the epigenetic research.

Chromatin dynamics could give the full picture of the chromosome changing in a cell's life. For early embryo or germ cell lineage, it reflected the tot potency status conversion. It has two steps: nascent primordial germ cells build a distinctive chromatin signature at E8.5, then several histone modifications were erased and histone variants were exchanged when primordial germ cells reside in the Gonads [30]. Methylation reprogramming is a conserved mechanism in mammals. In cloned embryos, there are aberrant epigenetic reprogramming which contributes to the low efficiency of clones [31]. Somatic nuclei of cloned embryos undergo some genome-wide reprogramming events such as methylation and demethylation. 5mC was regarded as a minor constituent of mammalian genomes six decades ago. Nowadays the genome-wide chromatin maps studies showed that the transcriptional regulation roles and the role of epigenetic modifications across many cell types [32]. International Human Epigenome Consortium, has drawn a comprehensively map the entire human epigenome [33].

Increasing knowledge of the epigenetic regulation of early embryo development would affect a wider biomedical area. Major epigenetic reprogramming and their intersections in the early embryo creates the plasticity followed the principal cell lineages differentiation. Lineage-determining transcription factors, such as *Elf5,* functioned in the trophoblast-specific transcriptional circuit, or 'gatekeeper' genes by the epigenetic marking. Another similar gatekeeper function is the methylation of *Stella*, participated in the transition of embryo stem cells to epiblast stem cells [34].

Preformation and epigenesis are two modes, for the specification of germ cell fate during the embryo development [35]. The preformation is seen in Drosophila and so on, involving localization of maternal determinants of the egg, or germ plasm. By contrast, the 'Epigenesis', which could be seen in mammals, either germ cells or somatic mesoderm are formation response to signaling molecules from adjacent tissues. This implies that cells recruited for the germline will undergo 'epigenetic reprogramming' from a somatic to a potentially totipotent germline phenotype [36]. Apart from embryo research, tumor heterogeneity is another import application for the single

cell techniques. The phenotype, genetic and non-genetic factors difference between tumor and its niches are illuminated by the single cell techniques in the last decades, which give many inspirations to early embryo research.

9.4 The Bioinformatics Tool for Single Cell Embryo Research

The bioinformatics approaches are essential to single cell analysis. Single cell analysis presents an obvious knowledge the cells not behavior in a homogeneous even in the same tissue. From experimental techniques to the data mining techniques, there are many bioinformatics algorithms developed for single-cell analysis [37]. The omics technologies such as genomics, transcriptomics, proteomics and metabolomics, and epigenomic produce more and more knowledge in this field, the enormous information would give us a solid understanding of the biological processes involved in the reproductive system like oocyte, embryo, culture medium, Endometrium, Spermatozoa and testis, even placenta. Also, we can get a broader view of complex biological systems with a relatively low cost which could improve the success of assisted reproductive technology [38].

There are DBTMEE (http://dbtmee.hgc.jp) for early embryo bioinformatics research. DBTMEE combines gene expression profiles with various public resources, including RNA-seq data of embryonic stem cells and induced pluripotent stem cell (iPS), or the expression genes related to totipotent, pluripotent and differentiated cells while genetic and epigenetic characteristics taking into consideration [39]. There were dozens of bioinformatics tools. Such as Dr. seq2, a Quality Control and analysis pipeline tool, used for single cell transcriptome and epigenome data, including scATAC-seq and Drop-ChIP data [15]. Single cell omics datasets are with high-dimension and increasing complexity. It requires new computational tools and analytical strategies to mining the biological insights from these data. To understand the relationships among the cell population is the basis, then it should generate testable hypotheses aiming to illuminate how the heterogeneous cell population would respond or adapt to cellular niche or environmental cues [40].

'Omics' and microarrays have already been used in the IVF/ICSI (in vitro fertilization/Intracytoplasmic Sperm Injection) cycles. Specific challenges of microarray technology here is picogram levels of mRNA in a single cell/embryo, and the high degree of expression plasticity or high dynamic changes in early stage embryos [41]. The embryologist pays great attention to how to correlate gene/protein/metabolite to its regulatory function in early embryonic development. Ultimately, we could use non-invasive tests to get high-quality oocytes and embryos, leading to increased implantation rates and higher success in elective Single Embryo Transfer. By our knowledge, only a small fraction of the human genome encodes proteins, so biologist put many strengths on the function and evolution of non-coding regulatory elements on a genome-wide scale [42]. It is better to leverage information from coevolution sequences to study protein-coding genes and their interactions. Phylogenetic profiling would mature as a powerful tool for gene function discovery [43].

9.5 The Epigenetics Mechanism in Early Embryo

Waddington explained cellular plasticity using epigenetics: Cells are residents on "landscape" of many potential states 60 years ago, they traveled to certain states during development and in disease. Waddington considered inheritance and development to be the same problem and change the paradigm by introducing the notion of epigenetics, which is defined as the unfolding of the genetic program for development [44]. How the H3.3 functions to support oogenesis and early development? Using a mosaic mouse model and imaging system. The mosaic mouse models express fluorescently tagged histones which H3.3 modifications were inhibited, demonstrated that the histones involved cell-cell competition between both germ and embryonic cells. And the

imaging system showed male-derived genome characters and H3.3 could regulate de novo chromatin assembly in a residue-dependent manner [45]. The genetic mosaic model with the histone rescue assay is an edged tool for investigating chromatin dynamics during early embryo development. It is generally stable for the epigenetic modifications of the genome in somatic cells of multicellular organisms. But to germ cells and early embryos, it changes dramatically on a genome-wide scale. Histone modification and small RNAs also contribute to the epigenetic inheritance and reprogramming.

Genetic diversity and ensure homologous chromosomes segregation are guarded by meiotic recombination. The sperm cells are the very experiment material for this topic. By sequencing 99 sperm using MALBAC, recombination near transcription start sites is very low. And a decreased crossover frequency comes with an increase of autosomal aneuploidy in the whole sperm genome. To sperm autosomal chromosomes, the segregation errors are not randomly generated during spermatogenesis but with relatively repressed crossover activity and high error rate [46]. In early embryogenesis, the protein G9a is an important regulator to inactivate *Oct-3/4*, the totipotent factors. Epigenetic changes could activate or suppress the gene expression or not even in differentiated cells [46]. Germ cells and early embryos seem to have an innate mechanism for recognizing and demethylating *Oct-3/4*. Once it was methylation or undergoes inactivation post-implantation, irreversibly it was locked in a repressed state [47].

From oocyte to blastocyst stages, the preimplantation embryos, the major and essential objective is how the embryonic and extra-embryonic cell fates determined. By live-cell tracking, we know cell-fate decisions heterogeneous gene expression initiates at the 4-cell stage by heterogeneous gene expression. *Oct4*, *Sox2* and the targets like *Sox21* are highly heterogeneous in the 4-cell embryo [1]. Heterogeneous expression of *Sox21* and other totipotent factors are the mechanisms directing embryo cell fate. Several target genes of *Oct4* and *Sox2* are co-regulated in the subset genes at 4-cell stage; *Sox21* is co-regulated with other transcription factors like *Nanog* and *Esrr*, supports this hypothesis [1]. The morphological dynamics also have enormous value for classifying the embryo which is applied in time-lapse microscopy [48]. Transcriptional architecture in mammalian preimplantation development is a complex process. By single-cell RNA sequencing, we can see a single-nucleotide resolution of variants of mRNAs and paternal-specific single nucleotide polymorphisms. The majority of stage-specific gene network modules are notably preserved, but developmental specificity and timing are different between human and mouse [49]. By single cell bioluminescence imaging, the circadian clock oscillation was measured in the mouse embryo from E10 to E19 [50]. This approach could use to preimplantation embryos.

Cell fate determination or cell reprogramming is always the holy grail in the biomedicine area. More and more data indicate that the cell fate is differentiated between the 2- to 4- cell stage and associate the differences with inner cell mass or trophectoderm differentiation, especially co-expression networks give us the evidence for molecular asymmetry [51]. Cell fate decisions mechanism are the biased expression of key transcription factors in a guiding way rather in a determined way. A higher level of *Cdx2* prone to form trophectoderm, a higher level of *NANOG l*eads a cell to form epiblast, and a higher level of *GATA6 e*xpression will form PE [52]. Once expression differences generated, they are self-reinforcing. Development is flexible until a certain point and reflected changing circumstances [53]. Single-cell technology not only tells us the locate cells on this landscape but illuminates the molecular mechanisms how the landscape shaped itself [54]. This idea could trace to Robert Hooke's observation in 1665, the founder of modern biology, who aimed to classify cells by form and function. In the last 400 years, biologists followed this idea and get many prominent discoveries in cells.

DNA Methylation is an essential epigenetic mechanism controlling transcription in mammals. Which evolution clue could be traced to *Neurospora crassa* [55], a fungi species. DNA

Methylation plays a pivotal role in cell lineage differentiation, sex chromosome dosage compensation, repression of retrotransposons, maintenance of genome stability, coordinated expression of imprinted genes [55, 56]. The drastic reprogramming of histone modifications calls for its functions and the molecular mechanisms research. For histone modification, H3K27me3 silenced multiple key regulators and the cellular memory. Pervasive H3K27me3 are usually found in regions depleted of transcription or DNA methylation [57]. Fertilized zygotes undergo dramatic epigenomic reprogramming. By highly sensitive ChIP-seq (STAR ChIP-seq), we would illuminate how histone modifications are inherited and reprogrammed.

The external environment or developmental cues will have affections on cells. A large part of the affections is through epigenetic signals, which mean the establishment, maintenance, and the metastable transcriptional states [58]. From two evolutional models *Schmidtea mediterranea* and social insects, we found incredible epigenetic flexibility in tissue regenerating [59] and why one genome could display dramatic physiological, morphological, and behavioral differences as the epigenetic casting [60]. Actually, embryonic stem cells are not uniform cell group but could be in a metastable state and shift pluripotency between inner cell mass- or epiblast-like states. The Stella expression level was used to determine the precise phenotypic state of different embryonic stem cell lines. Many genetic factors and epigenetic regulators affect its expression [61]. Chromatin modifiers, 24 histone demethylases and 41 histone methyltransferases, expression analysis showed that they co-expressed at the same histone residues. The "writer" and the "eraser" maintain a highly dynamic methylation state of the chromatin [62]. In this study, more than half of the chromatin modifiers expressed in either maternal or zygotic manner. *Prdm14*, *Ezh1*, *Scmh1* and *Tet1* take certain roles in embryo cell fate decisions.

TRIM28 is a core component of epigenetic modifier complex for specific genomic loci. It formed the repressive chromatin by recruiting chromatin-modification and remodeling factors [63]. It was required for maintaining genomic imprints and proper epigenetic environment during oocyte-to-embryo transition. Loss of maternal Trim28 results in pleiotropic, lethal phenotype [64]. There is dynamic epigenetic reprogramming in human embryonic stem cell lines ex vivo. The embryonic stem cell is with the complexity of gene regulatory networks, by transcriptomics inferring. From naive to primed pluripotent state, the key gene sets show different expression patterns, demonstrating distinct physiological properties. In primed status, *Sox2* and *Oct4* destabilized and *Klf4* levels drop [65]. *BMP* and *LIF* pathway is the major role in keeping the pluripotent cell state. The heterogeneity arising from timing variability is amplified in response to fluctuations in gene expression level, differentiation occurs in a proportion cells way rather than through changes in the cell states themselves [65].

X-chromosome inactivation and cell differentiation are tightly linked [66]. Dosage compensation of X chromosome RNA in blastocyst including TE, inner cell mass, EPI was investigated by single cell transcriptome, 1529 individual cells derived from 88 human preimplantation embryos [67]. Dosage compensation of X chromosome is very common in animal development and XIST is pivotal for the mechanism. X-chromosome inactivation occurred by imprinted inactive as early as cleavage stage [68]. For the sperm development, the DNA methylation is also important to male health. Oligozoospermia or azoospermia is very common in male clinics. The methylation status of imprinted genes *GTL2* and *PEG3* are aberrant compared to control [69]. Imprint methylation errors in sperm are associated with the environmental factors like polychlorinated biphenyl, or lifestyle like smoking, alcohol drinking. So improved particular biological or environmental factors would get a better epigenetics quality of sperm.

The DNase I-hypersensitive sites of preimplantation embryos reveal the chromatin regulatory landscape such as packaged sperm and egg genomes with key transcription factors [70]. It indicated paternal chromatin accessibility is

swiftly reprogrammed after fertilization. Allelic differences are established and maintained prior to the onset of parental allelic gene expression, an allelic difference of chromatin accessibility might be part of the mechanism contributing to allelic gene expression.

9.6 Single Cell Research on ES Cell

The cellular organelle could also be studied by single cell techniques. Such as the exosome are analyzed using antibody-modified glass slides [71]. Exosome properties have a closed relation to cell behaviors as cell-cell signaling, drug resistance, tumorigenesis, metastasis, and many other biological processes. Correlating the exosomes to the cell source is not very easy, given the high diversity and homogeneity of biological samples. So, it is highly desirable to quantify the single cell exosome, finding the changes in exosome properties when cells are affected by stresses, environmental changes or stimuli [71]. The work of Yamanaka lab, which the iPS could be induced by four totipotent factors *Oct3/4*, *Sox2, c-Myc*, and *Klf4*, is a classic. It demonstrated fibroblast could be directly generated by pluripotent stem cells by regulated only a few factors [72]. Now after more than 10 years, the iPS area get more and more solid clinic trial knowledge for its application.

After Yamanaka' s work, human iPS were cultured by ectopic expression of *Oct4*, *Sox2*, *Klf4* and *Myc*, the four totipotent transcription factors, *Oct4* and *Sox2* are essential [73]. Human pluripotent cell lines could be cultured from patients with certain diseases. This gives promise to the clinical success with human iPS cells, which shut down the potential harm of genetic modification. Demethylation in primordial germ cells is necessary for erasure of imprints and then the new imprints could be established properly later. Plants do not erase imprints; instead, they establish the imprints by demethylating the maternal genome of endosperm after fertilization. Epigenetic regulation between the two primary lineages (embryonic, extraembryonic) is fundamentally different from animals and plants, as genome-wide demethylation in endosperm but not the embryo of plants [74, 75]. This difference is apparently being a conserved mechanism in plants and animals [76].

Mouse germ cell and embryo is the most intended investigated experimental materials for address the developmental questions. The activation of genome-wide DNA demethylation is associated with the activation of *BER*, suggesting that the present ssDNA breaks trigger extensive chromatin remodeling and histone exchange in primordial germ cells [51]. Germ cell fate specification as oocytes and spermatozoa is fundamental events in development and genetics. *Prdm14* and *Prdm1* seem to be critical for the Re-acquisition of potential pluripotency and Genome-wide epigenetic reprogramming. The *Bmp4-Smad1* pathway participates in the initial activation of both *Prdm1* and *Prdm14* [77]. The relationship and methylation pattern of germ cells and the gonadal environment are becoming more and more clear. Many proteins involved in germline methylation, preferring imprinting control regions over other sequences. As the single cell techniques evolution, we will get the full picture of the epigenetic mechanisms that govern imprinting [42].

Stochastic and deterministic are the two models to explain the mechanism of reprogramming. iPS experiments are most consistent with the stochastic model [78], nuclear transfer or cell fusion consistent with a deterministic process [79]. There is data agreed with the stochastic model gene activation early in the process but also the deterministic models which sequence of gene activation at later stages [80]. The single-cell transcriptome is used for gene expression during human preimplantation development and the derivation of human embryonic stem cells from blastocysts. The transcriptomes of human EPI and the primary human embryonic stem cells showed dramatic global changes [81]. Single cell transcriptomes in model animals reveal the *Hippo* signaling participated in the embryo pattern formation [82].

9.7 Single Cell Techniques in Assisted Reproductive Technology

Now in IVF center, embryo evaluation is usually according to morphological metric- the uniformity and mitosis dynamic of cells, the cell connection, and the size/shape of the blastocyst [83]. At the same time, other single cell techniques like Raman microspectroscopy, metabolome, lipidome offer a reliable alternation for embryo grading. In reproductive medicine, the embryo manipulation like IVF or ICSI procedure or IVF culture would change the epigenetic status, also inflicted on transcriptional asymmetries such as genomic imprinting [84]. This will make a profound affection on the epigenetics of the body in the later life, will sometimes cause diseases. The hypothesis was verified by model animals with defined genotypes under strictly experimental conditions or procedures [85]. Interestingly, ICSI has a differential impact on the gene expression of the paternal and maternal alleles of imprinted genes. DNA methylation changes are easily inherited to the next generation in plants and also there is more and more evidence found in model animals, although the phenotype only can pass to a few generations. There is also evidence that the trans-generational heredity of epi-mutation does not rely on chromatin, but might involve RNA or cytoplasmic factors [86, 87].

The sperm genome and epigenome is critical for healthy offspring. Histone-bound DNA of sperm is highly susceptible to environmental affection such as oxidation damage. When a zygote transited to an embryo, not only the genome of the gametes is integrated, but also the cytoplasm content like mRNAs are fuse together. In the spermatogenesis course, sperm loses the majority of cytosolic antioxidants. Sperm is very vulnerable to free radical which cause DNA damage. DNA repair enzymes also be found in lower levels which give an explanation for the persistence of DNA damage in sperm [88]. Various forms of assisted reproductive technologies stressed the gametes and early embryo in the course of fertilization or culture. These stresses would change the patterns of gene expression and perturb the epigenetic reprogramming of the early embryo. The long-term effect of the stresses is also being illustrated in the last decades such as “large offspring syndrome” [89], changes in placental structure and function [90, 91].

When haploid genome undergoes endoduplication or if diploid pronuclear fails to develop, uniparental diploid may occur sometimes. In general, biparental diploid embryo appears higher for in vitro fertilization than intracytoplasmic sperm injection. Culture embryo up to blastocyst is a screening since most haploid embryos stop reaching this stage. Comprehensive counseling is in need for patients with potential risks and preimplantation genetic diagnosis could be offered [92]. Gene editing on human embryo is controversy in the scientific realm. Chinese Scientist Huang Junjiu and Liu Jianqiao put the advance of gene editing to the human early embryo. The work of Liu Jianqiao and his colleagues corrected the point mutations in *HBB* and *G6PD* in human zygote [93]. Huang Junjiu and his colleagues using gene editing to correct the hemophilia β gene [94]. Fan Yong and his colleagues using CRISPR/Cas9 to edit the *CCR5* gene in early embryo aiming to cure the HIV infection [95]. CRISPR/Cas9 could be used to correct the mosaicism in preimplantation embryo [96]. Techniques evolution but the accuracy efficiency and safety are the key principles we should observe strictly.

Epigenome sequencing the different layers of an individual cell is a difficult task. Single-cell multi-omics sequencing technology can analyze DNA methylation, chromatin state, copy number variation, ploidy simultaneously from an individual cell. When it applies to trace preimplantation developmental transition, the first map of allele-specific chromatin state and DNA methylation in early embryos is draw. It provides new insights in epigenomic reprogramming [97].

Next generation sequencing succeeds in detecting mosaicism in trophectoderm biopsies. Mosaic embryos may categorize as aneuploidy, leading to potentially viable embryos but being discarded, or entirely normal, carrying the risk of aneuploid pregnancy. The false positives and negatives frequency appears to be low, but it will

require additional works to verify. Then we can get an improved understanding of the causes and consequences of mosaicism and enhanced clinical outcome [98].

9.8 Conclusions

We are embracing an integrated single-cell era combined genomic, epigenomic, transcriptomic, and proteomic analysis etc. that will revolutionize the biomedicine realm [99]. Enabling single-cell genomic and transcriptomic approaches, which are also compatible with increasingly sensitive mass spectrometry-based approaches Analyze many genes in parallel could be at the single-cell level over developmental time. This provides a more accurate view of cellular phenotypes. The cellular level and are likely affected by the relative expression levels of many genes. The single-cell analysis offers intriguing new insights into the formation of the mammalian blastocyst. Cell signaling events appear to precede the segregation of lineage-specific transcriptional programs in the cell fate decisions [100].

Single cell technology has revolutionized almost every area of biomedical sciences. For the Challenges and emerging directions of single cell analysis, there are in situ transcriptomic analysis, lineage tracing, live imaging transcriptome, single cell multi-omics, modeling, functional validation, scRNA-seq with CRISPR/Cas9, etc. [101]. The advantage of single-cell analyses is obvious. It helps us understand the intercellular variability or heterogeneity. It contributes to defining cell types by the cellular character in the molecular level. Third, single-cell techniques could analyze rare cell events. Single cell analysis already is largely used for the Day3 embryo or blastocyst biopsy in the IVF clinic. Finally, single-cell analyses will contribute to unraveling the somatic mutations [102].

The Interdisciplinary research will introduce many opportunities and also many challenges. The past decades witness so many milestones and there will be more progress in the next decades. The far-reaching implications of these discoveries are improvements in assisted reproductive technologies, human health, animal cloning, livestock husbandry and so on. In the future, refine single-cell multi-omics approaches to the point where comprehensive atlases of cell state and lineage can be generated for cellular systems.

Conflicting Declaration No interests to declaration.

References

1. Goolam M, Scialdone A, Graham SJ, Macaulay IC, Jedrusik A, Hupalowska A et al (2016) Heterogeneity in Oct4 and Sox2 targets biases cell fate in 4-cell mouse embryos. Cell 165(1):61–74. [PubMed: 27015307]
2. Brady G, Barbara M, Iscove NN (1990) Representative in vitro cDNA amplification from individual hemopoietic cells and colonies. Methods Mol Cell Biol 2:17–25. [PubMed: none]
3. Wang J, Fan HC, Behr B, Quake SR (2012) Genome-wide single-cell analysis of recombination activity and de novo mutation rates in human sperm. Cell 150(2):402–412. [PubMed: 22817899]
4. Goetz JJ, Trimarchi JM (2012) Transcriptome sequencing of single cells with smart-Seq. Nat Biotechnol 30(8):763–765. [PubMed: 22871714]
5. Yan L, Huang L, Xu L, Huang J, Ma F, Zhu X et al (2015) Live births after simultaneous avoidance of monogenic diseases and chromosome abnormality by next-generation sequencing with linkage analyses. Proc Natl Acad Sci U S A 112(52):15964–15969. [PubMed: 26712022]
6. Xu J, Fang R, Chen L, Chen D, Xiao JP, Yang W et al (2016) Noninvasive chromosome screening of human embryos by genome sequencing of embryo culture medium for in vitro fertilization. Proc Natl Acad Sci U S A 113(42):11907–11912. [PubMed: 27688762]
7. Barker DJ, Osmond C (1986) Infant mortality, childhood nutrition, and ischaemic heart disease in England and Wales. Lancet 1(8489):1077–1081. [PubMed: 2871345]
8. Wadhwa PD, Buss C, Entringer S, Swanson JM (2009) Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. Semin Reprod Med 27(5):358–368. [PubMed: 19711246]
9. Lo SJ, Yao DJ (2015) Get to understand more from single-cells: current studies of microfluidic-based techniques for single-cell analysis. Int J Mol Sci 16(8):16763–16777. [PubMed: 26213918]
10. Zong C, Lu S, Chapman AR, Xie XS (2012) Genome-wide detection of single-nucleotide

and copy-number variations of a single human cell. Science 338(6114):1622–1626. [PubMed: 23258894]
11. Gross A, Schoendube J, Zimmermann S, Steeb M, Zengerle R, Koltay P (2015) Technologies for single-cell isolation. Int J Mol Sci 16(8):16897–16919. [PubMed: 26213926]
12. Reik W, Romer I, Barton SC, Surani MA, Howlett SK, Klose J (1993) Adult phenotype in the mouse can be affected by epigenetic events in the early embryo. Development 119(3):933–942. [PubMed: 8187648]
13. Spitzer MH, Nolan GP (2016) Mass cytometry: single cells. Many Features Cell 165(4):780–791. [PubMed: 27153492]
14. Xue Z, Huang K, Cai C, Cai L, Jiang CY, Feng Y et al (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500(7464):593–597. [PubMed: 23892778]
15. Zhao C, Hu S, Huo X, Zhang Y (2017) Dr.seq2: a quality control and analysis pipeline for parallel single cell transcriptome and epigenome data. PLoS One 12(7):e0180583. [PubMed: 28671995]
16. Wills QF, Livak KJ, Tipping AJ, Enver T, Goldson AJ, Sexton DW et al (2013) Single-cell gene expression analysis reveals genetic associations masked in whole-tissue experiments. Nat Biotechnol 31:748–752. [PubMed: 23873083]
17. O'Neill LP, VerMilyea MD, Turner BM (2006) Epigenetic characterization of the early embryo with a chromatin immunoprecipitation protocol applicable to small cell populations. Nat Genet 38(7):835–841. [PubMed: 16767102]
18. Shaw L, Sneddon SF, Zeef L, Kimber SJ, Brison DR (2013) Global gene expression profiling of individual human oocytes and embryos demonstrates heterogeneity in early development. PLoS One 8(5):e64192. [PubMed: 23717564]
19. Onjiko RM, Moody SA, Nemes P (2015) Single-cell mass spectrometry reveals small molecules that affect cell fates in the 16-cell embryo. Proc Natl Acad Sci U S A 112(21):6545–6550. [PubMed: 25941375]
20. Zhu J, Heinecke D, Mulla WA, Bradford WD, Rubinstein B, Box A et al (2015) Single-cell based quantitative assay of chromosome transmission fidelity. G3 (Bethesda) 5(6):1043–1056. [PubMed: 25823586]
21. Lorthongpanich C, Cheow LF, Balu S, Quake SR, Knowles BB, Burkholder WF et al (2013) Single-cell DNA-methylation analysis reveals epigenetic chimerism in preimplantation embryos. Science 341(6150):1110–1112. [PubMed: 24009393]
22. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM et al (2013) Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. Proc Natl Acad Sci U S A 110(47):18904–18909. [PubMed: 24167255]
23. Schreiber J, Wescoe ZL, Abu-Shumays R, Vivian JT, Baatar B, Karplus K et al (2013) Error rates for nanopore discrimination among cytosine, methylcytosine, and hydroxymethylcytosine along individual DNA strands. Proc Natl Acad Sci U S A 110(47):18910–18915. [PubMed: 24167260]
24. Rosen CB, Rodriguez-Larrea D, Bayley H et al (2014) Single-molecule site-specific detection of protein phosphorylation with a nanopore. Nat Biotechnol 32(2):179–181. [PubMed: 24441471]
25. Feng Y, Zhang Y, Ying C, Wang D, Du C et al (2015) Nanopore-based fourth-generation DNA sequencing technology. Genomics Proteomics Bioinformatics 13(3):4–16. [PubMed: 25743089]
26. Wu J, Huang B, Chen H, Yin Q, Liu Y, Xiang Y et al (2016) The landscape of accessible chromatin in mammalian preimplantation embryos. Nature 534(7609):652–657. [PubMed: 27309802]
27. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M et al (2011) 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun 2:241. [PubMed: 21407207]
28. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J et al (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125(2):315–326. [PubMed: 16630819]
29. Turner BM (2007) Defining an epigenetic code. Nat Cell Biol 9(1):2–6. [PubMed: 17199124]
30. Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F et al (2008) Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature 452(7189):877–881. [PubMed: 18354397]
31. Dean W, Santos F, Stojkovic M, Zakhartchenko V, Walter J, Wolf E et al (2001) Conservation of methylation reprogramming in mammalian development: aberrant reprogramming in cloned embryos. Proc Natl Acad Sci U S A 98(24):13734–13738. [PubMed: 11717434]
32. Meissner A (2010) Epigenetic modifications in pluripotent and differentiated cells. Nat Biotechnol 28(10):1079–1088. [PubMed: 20944600]
33. Rivera CM, Ren B (2013) Mapping human epigenomes. Cell 155(1):39–55. [PubMed: 24074860]
34. Hemberger M, Dean W, Reik W (2009) Epigenetic dynamics of stem cells and cell lineage commitment: digging Waddington's canal. Nat Rev Mol Cell Biol 10(8):526–537. [PubMed: 19603040]
35. Extavour CG, Akam M (2003) Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. Development 130(24):5869–5884. [PubMed: 14597570]
36. Seki Y, Yamaji M, Yabuta Y, Sano M, Shigeta M, Matsui Y et al (2007) Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. Development 134(14):2627–2638. [PubMed: 17567665]

37. Yalcin D, Hakguder ZM, Otu HH (2016) Bioinformatics approaches to single-cell analysis in developmental biology. Mol Hum Reprod 22(3):182–192. [PubMed: 26358759]
38. Egea RR, Puchalt NG, Escriva MM, Varghese AC (2014) OMICS: current and future perspectives in reproductive medicine and technology. J Hum Reprod Sci 7(2):73–92. [PubMed: 25191020]
39. Park SJ, Shirahige K, Ohsugi M, Nakai K (2015) DBTMEE: a database of transcriptome in mouse early embryos. Nucleic Acids Res 43(Database issue):D771–D776. [PubMed: 25336621]
40. Su Y, Shi Q, Wei W (2017) Single cell proteomics in biomedicine: High-dimensional data acquisition, visualization, and analysis. Proteomics 17(3–4). [PubMed: 28128880]
41. Weaver JR, Susiarjo M, Bartolomei MS (2009) Imprinting and epigenetic changes in the early embryo. Mamm Genome 20(9–10):532–543. [PubMed: 19760320]
42. Boyle AP, Araya CL, Brdlik C, Cayting P, Cheng C, Cheng Y et al (2014) Comparative analysis of regulatory information and circuits across distant species. Nature 512(7515):453–456. [PubMed: 25164757]
43. Dey G, Meyer T (2015) Phylogenetic profiling for probing the modular architecture of the human genome. Cell Syst 1(2):106–115. [PubMed: 27135799]
44. Holliday R (2014) Epigenetics: a historical overview. Epigenetics 1(2):76–80. [PubMed: 17998809]
45. Zhou L, Baibakov B, Canagarajah B, Xiong B, Dean J (2017) Genetic mosaics and time-lapse imaging identify functions of histone H3.3 residues in mouse oocytes and embryos. Development 144(3):519–528. [PubMed: 27993980]
46. Lu S, Zong C, Fan W, Yang M, Li J, Chapman AR et al (2012) Probing meiotic recombination and aneuploidy of single sperm cells by whole-genome sequencing. Science 338(6114):1627–1630. [PubMed: 23258895]
47. Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y et al (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. Nat Cell Biol 8(2):188–194. [PubMed: 16415856]
48. Wong CC, Loewke KE, Bossert NL, Behr B, De Jonge CJ, Baer TM et al (2010) Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. Nat Biotechnol 10:1115–1121. [PubMed: 20890283]
49. Hajkova P, Jeffries SJ, Lee C, Miller N, Jackson SP, Surani MA (2010) Genome-wide reprogramming in the mouse germ line entails the base excision repair pathway. Science 329(5987):78–82. [PubMed: 20595612]
50. Umemura Y, Koike N, Ohashi M, Tsuchiya Y, Meng QJ, Minami Y et al (2017) Involvement of posttranscriptional regulation of Clock in the emergence of circadian clock oscillation during mouse development. Proc Natl Acad Sci U S A. pii: 201703170. Epub ahead of print. [PMID: 28827343]
51. Biase FH, Cao X, Zhong S (2014) Cell fate inclination within 2-cell and 4-cell mouse embryos revealed by single-cell RNA sequencing. Genome Res 24(11):1787–1796. [PubMed: 25096407]
52. Zernicka-Goetz M, Morris SA, Bruce AW (2009) Making a firm decision: multifaceted regulation of cell fate in the early mouse embryo. Nat Rev Genet 10(7):467–477. [PubMed: 19536196]
53. Trapnell C (2015) Defining cell types and states with single-cell genomics. Genome Res 25(10):1491–1498. [PubMed: 26430159]
54. Messerschmidt DM, Knowles BB, Solter D (2014) DNA methylation dynamics during epigenetic reprogramming in the germline and preimplantation embryos. Genes Dev 28(8):812–828. [PubMed: 24736841]
55. Zhao Y, Shen Y, Yang S, Wang J, Hu Q, Wang Y et al (2010) Ubiquitin ligase components Cullin4 and DDB1 are essential for DNA methylation in Neurospora crassa. J Biol Chem 285(7):4355–4365. [PubMed: 19948733]
56. Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16(1):6–21. [PubMed: 11782440]
57. Zheng H, Huang B, Zhang B, Xiang Y, Du Z, Xu Q et al (2016) Resetting epigenetic memory by reprogramming of histone modifications in mammals. Mol Cell 63(6):1066–1079. [PubMed: 27635762]
58. Bonasio R, Tu S, Reinberg D (2010) Molecular signals of epigenetic states. Science 330(6004):612–616. [PubMed: 21030644]
59. Sánchez AA (2006) Planarian regeneration: its end is its beginning. Cell 124(2):241–245. [PubMed: 16439195]
60. Bonasio R, Zhang G, Ye C, Mutti NS, Fang X, Qin N et al (2010) Genomic comparison of the ants Camponotus floridanus and Harpegnathos saltator. Science 329(5995):1068–1071. [PubMed: 20798317]
61. Hayashi K, Lopes SM, Tang F, Surani MA (2008) Dynamic equilibrium and heterogeneity of mouse pluripotent stem cells with distinct functional and epigenetic states. Cell Stem Cell 3(4):391–401. [PubMed: 18940731]
62. Nestorov P, Hotz HR, Liu Z, Peters AH (2015) Dynamic expression of chromatin modifiers during developmental transitions in mouse preimplantation embryos. Sci Rep 5:14347. [PubMed: 26403153]
63. Iyengar S, Farnham PJ (2011) KAP1 protein: an enigmatic master regulator of the genome. J Biol Chem 286(30):26267–26276. [PubMed: 21652716]
64. Messerschmidt DM, de Vries W, Ito M, Solter D, Ferguson-Smith A, Knowles BB (2012) Trim28 is required for epigenetic stability during mouse oocyte to embryo transition. Science 335(6075):1499–1502. [PubMed: 22442485]
65. Blauch LR, Gai Y, Khor JW, Sood P, Marshall WF, Tang SKY (2017) Microfluidic guillotine for

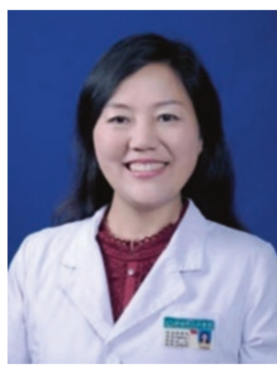
single-cell wound repair studies. Proc Natl Acad Sci U S A. [PubMed: 28652371]
66. Silva SS, Rowntree RK, Mekhoubad S, Lee JT (2008) X-chromosome inactivation and epigenetic fluidity in human embryonic stem cells. Proc Natl Acad Sci U S A 105(12):4820–4825. [PubMed: 18339803]
67. Petropoulos S, Edsgard D, Reinius B, Deng Q, Panula SP, Codeluppi S et al (2016) Single-cell RNA-seq reveals lineage and X chromosome dynamics in human preimplantation embryos. Cell 165(4):1012–1026. [PubMed: 27662094]
68. Okamoto I, Otte AP, Allis CD, Reinberg D, Heard E (2004) Epigenetic dynamics of imprinted X inactivation during early mouse development. Science 303(5658):644–649. [PubMed: 14671313]
69. Kobayashi N, Miyauchi N, Tatsuta N, Kitamura A, Okae H, Hiura H et al (2017) Factors associated with aberrant imprint methylation and oligozoospermia. Sci Rep 7:42336. [PubMed: 28186187]
70. Inoue A, Jiang L, Lu F, Suzuki T, Zhang Y (2017) Maternal H3K27me3 controls DNA methylation-independent imprinting. Nature. https://doi.org/10.1038/nature23262. [PubMed: 28723896]
71. Chiu YJ, Cai W, Lee T, Kraimer J, Lo YH (2017) Quantitative analysis of exosome secretion rates of single cells. Bio Protoc 7(4). [PubMed: 28603750]
72. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126(4):663–676. [PubMed: 16904174]
73. Park IH, Zhao R, West JA, Yabuuchi A, Huo H, Ince TA et al (2008) Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451(7175):141–146. [PubMed: 18157115]
74. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL et al (2009) Genome-wide demethylation of Arabidopsis endosperm. Science 324(5933):1451–1454. [PubMed: 19520962]
75. Gehring M, Bubb KL, Henikoff S (2009) Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science 324(5933):1447–1451. [PubMed: 19520961]
76. Feng S, Jacobsen SE, Reik W (2010) Epigenetic reprogramming in plant and animal development. Science 330(6004):622–627. [PubMed: 21030646]
77. Yamaji M, Seki Y, Kurimoto K, Yabuta Y, Yuasa M, Shigeta M et al (2008) Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet 40(8):1016–1022. [PubMed: 18622394]
78. Jaenisch R, Young R (2008) Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. Cell 132(4):567–582. [PubMed: 18295576]
79. Hanna JH, Saha K, Jaenisch R (2010) Pluripotency and cellular reprogramming: facts, hypotheses, unresolved issues. Cell 143(4):508–525. [PubMed: 21074044]
80. Buganim Y, Faddah DA, Cheng AW, Itskovich E, Markoulaki S, Ganz K et al (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 150(6):1209–1222. [PubMed: 22980981]
81. Yan L, Yang M, Guo H, Yang L, Wu J, Li R et al (2013) Single-cell RNA-Seq profiling of human preimplantation embryos and embryonic stem cells. Nat Struct Mol Biol 20(9):1131–1139. [PubMed: 23934149]
82. Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C et al (2017) The Drosophila embryo at single-cell transcriptome resolution. Science. pii: eaan3235. Epub ahead of print. [PMID: 28860209]
83. Brayboy LM, Wessel GM (2016) The double-edged sword of the mammalian oocyte – advantages, drawbacks and approaches for basic and clinical analysis at the single cell level. Mol Hum Reprod 22(3):200–207. [PubMed: 26590170]
84. Kohda T, Ishino F (2013) Embryo manipulation via assisted reproductive technology and epigenetic asymmetry in mammalian early development. Philos Trans R Soc Lond Ser B Biol Sci 368(1609):20120353. [PubMed: 23166403]
85. Kohda T (2013) Effects of embryonic manipulation and epigenetics. J Hum Genet 58(7):416–420. [PubMed: 23739123]
86. Probst AV, Dunleavy E, Almouzni G (2009) Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol 10(3):192–206. [PubMed: 19234478]
87. Morgan DK, Whitelaw E (2008) The case for transgenerational epigenetic inheritance in humans. Mamm Genome 19(6):394–397. [PubMed: 18663528]
88. Kumar M, Kumar K, Jain S, Hassan T, Dada R (2013) Novel insights into the genetic and epigenetic paternal contribution to the human embryo. Clinics 68(S1):5–14. [PubMed: 23503950]
89. Behboodi E, Anderson GB, BonDurant RH, Cargill SL, Kreuscher BR, Medrano JF et al (1995) Birth of large calves that developed from in vitro-derived bovine embryos. Theriogenology 44(2):227–232. [PubMed: 16727722]
90. Chavatte-Palmer P, Camous S, Jammes H, Le Cleac'h N, Guillomot M, Lee RS (2012) Review: placental perturbations induce the developmental abnormalities often observed in bovine somatic cell nuclear transfer. Placenta 33(Suppl):S99–S104. [PubMed: 22000472]
91. Wei Y, Chen S, Huang X, Lam SM, Shui G, Sun F (2015) Assisted reproduction causes reduced intrauterus fetal growth by disrupting placental lipid metabolism. bioRxiv 11. https://doi.org/10.1101/030965. [PubMed: none]
92. Macaulay IC, Ponting CP, Voet T (2017) Single-cell multiomics: multiple measurements from single cells. Trends Genet 33(2):155–168. [PubMed: 28089370]
93. Tang L, Zeng Y, Du H, Gong M, Peng J, Zhang B et al (2017) CRISPR/Cas9-mediated gene editing

in human zygotes using Cas9 protein. Mol Gen Genomics 292(3):525–533. [PubMed: 28251317]
94. Liang P, Xu Y, Zhang X, Ding C, Huang R, Zhang Z et al (2015) CRISPR/Cas9-mediated gene editing in human tripronuclear zygotes. Protein Cell 6(5):363–372. [PubMed: 25894090]
95. Kang X, He W, Huang Y, Yu Q, Chen Y, Gao X et al (2016) Introducing precise genetic modifications into human 3PN embryos by CRISPR/Cas-mediated genome editing. J Assist Reprod Genet 33(5):581–588. [PubMed: 27052831]
96. Ma H, Marti-Gutierrez N, Park SW, Wu J, Lee Y, Suzuki K et al (2017) Correction of a pathogenic gene mutation in human embryos. Nature 548(7668):413–419. [PMID: 28783728]
97. Guo F, Li L, Li J, Wu X, Hu B, Zhu P et al (2017) Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. Cell Res 27(8):967–988. [PubMed: 28621329]
98. Munné S, Wells D (2017) Detection of mosaicism at blastocyst stage with the use of high-resolution next-generation sequencing. Fertil Steril 107(5):1085–1091. [PubMed: 28390692]
99. Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14(9):618–630. [PubMed: 23897237]
100. Guo G, Huss M, Tong GQ, Wang C, Li Sun L, Clarke ND et al (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell 18(4):675–685. [PubMed: 20412781]
101. Yuan GC, Cai L, Elowitz M, Enver T, Fan G, Guo G et al (2017) Challenges and emerging directions in single-cell analysis. Genome Biol 18(1):84. [PubMed: 28482897]
102. Speicher MR (2013) Single-cell analysis: toward the clinic. Genome Med 5(8):74. [PubMed: 23998189]

Yubao Wei, PhD. He is a researcher and director of the molecular lab of The Reproductive Institute of Henan Provincial People's Hospital. He has a ten year's research experiences and his research interests are genetics and epigenetics of early embryo and placenta development. He is funded by the "23456" talent project of the hospital. Now he is focusing on Genome imprinting, Polycystic Ovary Syndrome and azoospermia. His has dozens of research articles and popular science articles. He is the member of Genetics Society of China, China Zoological Society, Chinese Society for Biochemistry and Molecular Biology.

Helong Zhang, MMed. He is a senior doctor in Reproductive hospital of Henan Province with 10 years' clinic experiences. He got his Master degree from a renounced medical university, the Chinese Medical University, Shenyang City. His strength is at Artificial Reproductive Technology, Gynecological Operation, Hysteroscopy and Controlled Ovarian Stimulation. He was invited to many international conferences and have visited to numbers of IVF center around the world.

Qian Wang, MMed. Resident doctor at Henan Provincial Reproductive Hospital. Her Master of Medical degree was issued in 2016 by Zhengzhou University. She is excellent at assisted reproductive technologies and familiar with the clinic works such as diagnosis and treatment of various infertility. Her research strength is mainly on the epigenetics of oocytes and embryos of PCOS.



Cuilian Zhang, MD, PhD. Professor, Doctoral advisor, Chief Physician, Dean of Henan Provincial Reproductive Hospital. She is Chairman of Reproductive Medical Doctor Branch of Henan Medical Doctor Association, Vice-chairman of Henan Society of Reproductive Medicine, Standing Committee member of Chinese Medical Doctor Association, Deputy director of the precision assisted reproductive medicine research unit of Chinese Medical Doctor Association, the Chinese youth committee of Chinese Medical Association's Reproductive Medicine Branch, Member of the clinical division of Chinese Medical Association's Reproductive Medicine Branch, Member of the clinical reproductive genetics committee of Chinese Physicians Association, and so on. She has more than 100 publications and 9 research projects. She is the Editor of Chinese Journal of Reproduction and Contraception, Editor of Journal of Practical Diagnosis and Therapy. She was got several professional awards.

10 The Potential Roles and Advantages of Single Cell Sequencing in the Diagnosis and Treatment of Hematological Malignancies

Mingyue Shi, Xiaoyan Dong, Lei Huo, Xiaobin Wei, Fang Wang, and Kai Sun

Abstract

Hematological malignancies (HM) are a heterogeneous group of life-threatening hematological diseases. The heterogeneity and clonal evolution of HM subpopulations are the main obstacles for precise diagnoses, risk stratification, and even targeted therapies. Standard bulk-sample genomic examinations average total mutations from multiple subpopulations and conceal the clonal diversity that may play a significant role in HM progression. Therefore, the development of novel methods that detect intra-tumor heterogeneity is critical for the discovery of novel potential therapeutic targets. The recently developed single cell sequencing (SCS) technologies can analyse genetic polymorphisms at a single cell level. SCS requires the precise isolation of single cells and amplification of their genetic material. It allows the analysis of genomic, transcriptomic, and epigenomic information in single cancer cells. SCS may also be able to monitor minimal residual disease (MRD) of HM by sequencing circulating tumor cells (CTCs) from peripheral blood. Functional heterogeneity and clonal evolution exist in acute leukemia, multiple myeloma (MM) and chronic myeloid leukemia (CML) subpopulations and have prognostic value. In this thesis, we provide an overview of SCS technologies in HM and discuss the heterogeneous genetic variation and clonal structure among subpopulations of HM. Furthermore, we aimed to shed light on the clinical applications of SCS technologies, including the development of new targeted therapies for drug-resistant or recurrent HM.

Keywords

Hematological malignancies · Single cell sequencing · Heterogeneity · Clonal evolution · Precise therapy

M. Shi
Department of Hematology, Institute of Hematology, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, Henan, China

Zhengzhou University the Academy of Medical Sciences, Zhengzhou University, Zhengzhou, China

X. Dong · L. Huo · X. Wei · F. Wang · K. Sun (✉)
Department of Hematology, Institute of Hematology, Henan Provincial People's Hospital, Zhengzhou University People's Hospital, Zhengzhou, Henan, China
e-mail: sunkai@cellscience.org

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_10

Abbreviations

ALL	acute lymphoblastic leukemia
AML	acute myeloid leukemia
CML	chronic myeloid leukemia
CTCs	circulating tumor cells
HM	hematological malignancies
HSCs	hematopoietic stem cells
LSCs	leukemic stem cells
MM	multiple myeloma
MPN	myeloproliferative neoplasms
MRD	minimal residual disease
RNA-Seq	RNA sequencing
SCS	single cell sequencing
TKI	tyrosine kinase inhibition

10.1 Introduction

Hematological diseases are disorders that primarily affect the blood, including malignant blood diseases, such as leukemia, lymphoma, MM; and non-malignant blood diseases, such as anemia, thrombocytopenia, and hereditary blood diseases. Since HM are the major group of heterogeneous life-threatening malignant hematological diseases that need chemotherapy or even hematopoietic stem cell transplantation [40], this article mainly focuses on explaining the heterogeneity and clonal evolution of HM using the newly developed single cell technologies. Cells are the smallest unit of life, and each cell varies widely. A normal hematopoietic cell may develop into a potentially malignant cancer stem cell that ultimately leads to HM. Recent studies have already demonstrated that cancer stem cells are not monoclonal population but are composed of different subpopulations of cells, which give rise to intra-tumor heterogeneity. Clonal diversity is a prominent trait of HM and is expected to play a significant part in relapse and resistance to therapy. To date, however, our understanding of intra-tumor heterogeneity is limited by the universally used next generation sequencing, which averages the genetic information of complex cell subpopulations. To address this dilemma, single cell approaches were developed and have been used to better understand HM biology, such as cancer evolution and diversity. Additionally, single cell methods can provide clues about the role of cancer stem cells in tumor progression.

Within an individual blood cancer, there coexists multiple genetically heterogeneous subclones that undergo multiple genetic mutations over time or treatment [4]. Such heterogeneity and clonal evolution pose a great challenge to disease diagnosis, risk stratification, prognosis, drug resistance and targeted therapy. Although standard bulk-sample genomic studies have significantly contributed to the understanding of HM biology, they cannot completely delineate clonal diversity and evolution among heterogeneous HM tumor cells. Clonal evolution in HM can be attributed to the emergence of malignant clones during disease progression and therapy, resulting in drug resistance or relapse. Generally, researchers use next generation sequencing to monitor mutant alleles and calculate clonality from HM diagnosis to relapse. Such techniques evaluate clonal frequencies via allele frequencies in DNA sequencing information from bulk samples. However, they are inaccurate for all mutant loci in HM samples.

Due to extensive advances in the past few years, single-cell genomic techniques have enabled the identification of tumor heterogeneity and putative subclonal populations [47]. By precisely isolating single cells and amplifying their genetic material, SCS technologies allow the analysis of clonal diversity using single-cell genomic, transcriptomic and even epigenomic modalities, such as DNA methylation [23, 32]. SCS technologies have a broad impact on diverse medical fields, as they enabled the elucidation of molecular mechanisms involved in early embryonic development, carcinogenesis, tumor heterogeneity and clonal evolution, as well as CTCs and immunology [30, 46]. Moreover, SCS was picked as the technique in 2013 by the Nature Publishing Group.

HM, which include various forms of leukemia, lymphoma, and myeloma, are mostly caused by a variety of gene mutations that induce clinically detectable abnormal gene expression [39]. HM are highly heterogeneous and display clonal evolution after chemotherapy or disease relapse [9, 11]. Since individual cells are the basic units of all

HM, it is critical to gain more knowledge on cellular evolution and the genomic variability of blood tumor subtypes at a single-cell level. Indeed, functional heterogeneity with prognostic value has already been illustrated by recent studies that have applied SCS techniques to investigate acute myeloid leukemia (AML), myelodysplastic syndromes (MDS), MM, CML and chronic lymphoid leukemia subpopulations [6, 34, 41]. Genomic information from single cells is presently technically available. Employing flow cytometry to immunophenotype single cells is currently a routine method for hematological diagnosis. Therefore, single-cell genomic profiling has, in theory, clinical utility in the diagnosis of HM. For example, a research group recently utilized RNA SCS (RNA sequencing, RNA-Seq) to interrogate 20 cells that were isolated from the bone marrow of an AML patient [47]. In this study, they analyzed single-cell transcriptomic profling of clinical patients. Using single-cell RNA-Seq and routine flow cytometry, they identified gene expression profile of putative blasts. The genetic data showed that single-cell genomic analyses might be helpful for individual cell classification, diagnosis, prognosis, and the ideal treatment modality for various kind of HM [47].

With the rapid development of SCS methods, we believe that single-cell genomics could provide a more rigorous way to determine the heterogeneity and phenotypic/genetic evolution of HM subpopulations. DNA SCS was used in one study to analyze single leukemia cells separated from three patients diagnosed with secondary AML that originated from MDS. Combined analysis of the genetic variants in bulk samples and single cells helped locate mutations within clonal hierarchies [16]. In another examination, 1479 single cells were isolated from six acute lymphocytic leukemia (ALL) patients and then targeted sequenced. This study confirmed that various clonal heterogeneous subpopulations existed in ALL patients and might include potential clonal origins or evolution of leukemia cells [12].

Moreover, analysis of CTCs with RNA SCS methods might provide the same or even more sensitive genetic information on myeloma cells in bone marrow [24]. The result of this study suggested that RNA SCS of CTCs could help classify and quantitatively assess MM stages and prognostic genes [24]. Recent advancements in SCS technologies allow for a more detailed understanding of the genomic architecture and clonal evolution of blood cancer cell subpopulations. Thus, they might be able to improve targeted therapies for complex subpopulations, which may lead to improvements in patient outcomes. Future use of single cell analysis is likely to have a profound clinical potential.

SCS methods for the targeted sequencing of CTCs can also be used to noninvasively monitor of MRD and accurately identify of clonal heterogeneity among patients with HM, which will lead to significant advances in personalized management [8]. As we all know, bone marrow puncture and biopsy are essential invasive procedures in HM diagnosis and MRD monitoring, which are associated with pain, inconvenience, and expense. Furthermore, flow cytometry and quantitative real-time PCR are routine diagnostic tools for all HM patients but are limited by the inability to identify clonal heterogeneity among single cells [29]. Thus, by sequencing CTCs at different time points during therapeutic process, hematologists might be able to trace the progress of mutations and quickly alter treatment strategies before drug resistance occurs [25]. In the near future, instead of bone marrow biopsy, targeted sequencing of single peripheral blood cells may provide reliable information on disease evolution and resistance to treatment.

10.2 SCS Techniques and Potential Advantages

10.2.1 Isolation of Single Cells

To use DNA or RNA SCS to analyze a single cell, we must first separate one single cell from the heterogeneous samples. To randomly isolate a single cell from a plentiful population, several methods can be employed: mouth pipetting, micromanipulation, flow-assisted cell sorting, laser-capture-microdissection, serial dilution and microfluidics [23].

10.2.2 Single-Cell Genome/DNA Sequencing (DNA SCS/DNA-Seq) Methods

DNS SCS has revealed remarkable cellular heterogeneity inside HM and has identified significant clonal evolution. Before sequencing the genome of a single cell, DNA needs to be amplified by whole-genome-amplification. The most prevalent whole-genome-amplification approaches are multiple displacement amplification and degenerative oligonucleotide PCR (DOP-PCR) [30]. Other DNA-Seq methods are multiple annealing and looping based amplification cycles (MALBAC) and single nucleus exome sequencing (SNES) [44]. Then the expanded DNA is utilized to create libraries for next generation sequence.

10.2.3 Single-Cell RNA Sequencing (RNA SCS/RNA-Seq) Methods

RNA SCS has shed new light on the role of tumor microenvironments in disease progression and drug resistance [49]. It can analyze the transcriptome in a single cell and promote the study of clonal heterogeneity within temporally and spatially complex cell populations. To sequence the single-cell transcriptome, RNA must first be amplified by whole-transcriptome amplification. Commonly used single-cell transcriptome sequencing methods include the following: Switching Mechanism At the end of the 5′ end of the RNA Transcript (SMART-Seq), SMART-Seq2, Cell Expression by Linear amplification and Sequencing (CEL-Seq), CEL-Seq2, Single-cell tagged reverse transcription sequencing (STRT-Seq), Gene expression cytometry (Cyto-Seq), Drop-seq, in-Drop [36]. Smart-Seq and Smart-Seq2 are currently the most commonly used techniques [23].

10.2.4 Single-Cell Epigenomic Sequencing Methods

There is evidence that genome and transcriptome sequencing does not provide complete information on cells, and that cell properties are also determined by DNA methylation, histone modification, non-coding RNA regulation, etc. At present, single-cell reduced-represent bisulphite sequencing (scRRBS) and single cell bisulfite sequencing are the main single-cell DNA methylation sequencing techniques (scBS-Seq) [10].

10.2.5 Potential Applications and Advantages of SCS

Traditional methods for the examination of genetic information depend on measuring the bulk samples, which mix the DNA of thousands or millions of cells; however, these methods obscure intra-tumor heterogeneity. Unlike bulk-sample genotyping, SCS methods enhance the understanding of heterogeneous genetic changes and provide more accurate methods to reveal the clonal substructure in heterogeneous samples [16]. DNA/RNA SCS techniques are potent and novel tools for the elucidation of uncommon stem cell functions throughout HM progression and for explaining clonal diversity (Fig. 10.1).

During the HM diagnosis or relapse, blood tumor cells frequently undergo clonal evolution and acquire genetic mutations. Thus, non-invasive and precise monitoring of MRD could lead up to improved therapeutic choices and extend HM patients survival. Since CTCs in the blood are easily accessible, they are ideal for use in the non-invasively monitoring of each patient's molecular traits. Several methods have been developed for the evaluation of treatment reactions and non-invasive detection of MRD, including flow cytometry and quantitative real-time PCR methods, which concurrently analyze millions or billions cells from HM subpopulations; thus, they generally ignore intra-tumor heterogeneity. On the other hand, SCS methods can be used to non-invasively observe heterogeneous genetic changes of single cells in heterogeneous samples (Fig. 10.1). The capacity to accurately detect CTCs and relevant genomic information could improve the prediction of clinical outcomes [5]. Improvements in SCS technologies for CTCs detection in recent years have led to augmented sensitivity with diminished costs.

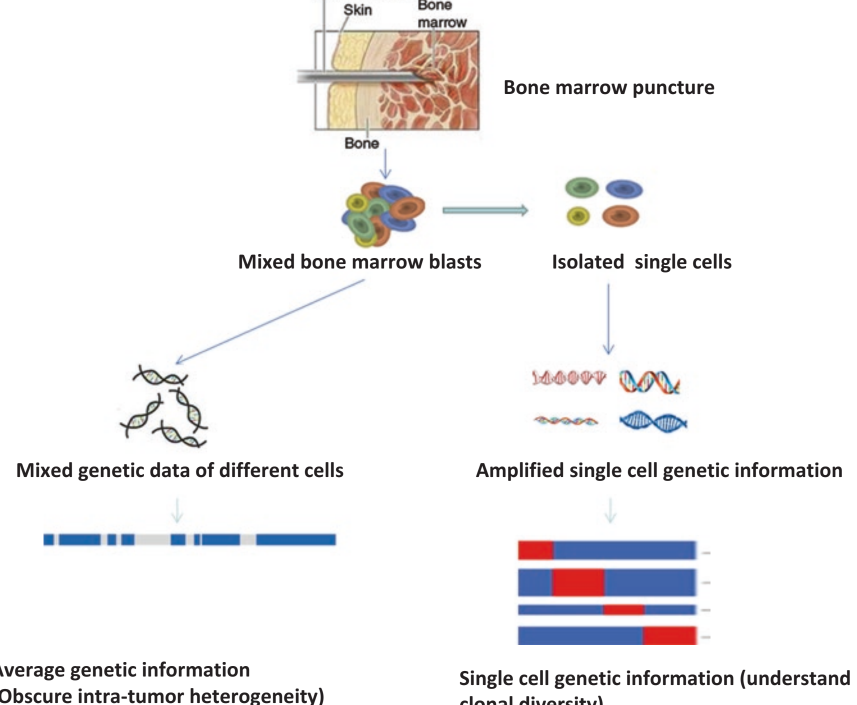


Fig. 10.1 Advantages of single-cell sequencing Traditional bulk-sample examinations mix the DNA of thousands or millions of cells, but these procedures obscure intra-tumor heterogeneity, SCS methods can provide a non-invasive way to precisely dissect the genetic information of heterogeneous subpopulations at a single cell level

One of the most recent studies has shown that SCS technologies could be directly applied for the non-invasive monitoring of myeloma cells in the peripheral blood. Subsequently, these researchers were able to verify mutations in primary myeloma derived from bone marrow [24]. Thus, peripheral blood can act as a tumor source when using SCS methods instead of BM biopsy and may provide reliable information on disease heterogeneity, evolution and treatment.

10.3 The Applications of SCS for HM

HM, including both myeloid and lymphoid/plasmacytoid cancers, such as MDS, CML, AML, lymphomas, chronic lymphoid leukemia and MM et al. are known to be highly heterogeneous and evolve over time [21, 26]. Within an individual HM, there co-exists multiple subclones that undergo genetic mutations to cause emergence, expansion, and extinction. This intra-tumor complicates the clinical diagnosis and therapeutic treatment of HM.

The classical hematological malignant cell types are mostly defined by surface markers through flow cytometry. To date, SCS technologies have verified heterogeneity in many of the previously determined classical hematopoietic cell types [48]. Thus, traditional methods that define cell types by surface makers are increasingly being challenged. By analyzing the genome, transcriptome and epigenome at the single cell level, SCS methods can be used to explain HM clonal diversity and evolution during diagnosis and treatment. SCS may also be able to trace clonal origins of single cells and the locations of genetic mutations. Here, we will

Table 10.1 Table of single cell sequencing publications in hematology research

Disease/Cell type	Method	References	Doi	Description
HSC	RNA-SCS	[51]	10.1038/nature17997	Tracing HSCs formation at a single cell resolution
AML	DNA-SCS	[17]	10.1126/scitranslmed.3004315	Clonal evolution of pre leukemic HSCs precedes AML
AML	DNA-SCS	[16]	10.1371/journal.pgen.1004462	Clonal architecture of secondary AML as defined by SCS
AML	DNA/RNA-SCS	[47]	10.3892/ol.2017.5669	Single-cell genomic profiling of AML for clinical use
AML	DNA –SCS	[34]	10.1126/scitranslmed.aaa0763	Single-cell genotyping demonstrates complex clonal diversity in AML
AML	RNA-SCS	[27]	10.1038/nbt.3637	Analyzing tumor heterogeneity and driver genes in AML by SBCapSeq
ALL	DNA-SCS	[12]	10.1073/pnas.1420822111	Dissecting the clonal origins of childhood ALL by single-cell genomics
ALL	RNA-SCS	[8]	10.1016/j.ccell.2016.11.002	Characterization of rare, dormant therapy-resistant cells in ALL
ALL	DNA-SCS	[2]	10.1186/s13059-016-0971-7	SCS reveals karyotype heterogeneity in ALL
CML	RNA-SCS	[13]	10.1038/nm.4336	Single-cell transcriptomics uncovers distinct molecular signatures of CML LSCs
MPN	DNA-SCS	[15]	10.1016/j.cell.2012.02.028	Single-cell exome sequencing to detect clonal evolution of a JAK2-negative MPN
MM	RNA-SCS	[28]	10.1038/leu.2015.361	Single-cell analysis of targeted transcriptome predicts drug sensitivity of MM
MM	RNA-SCS	[24]	10.1126/scitranslmed.aac7037	Genetic interrogation of circulating MM cells at single-cell resolution

discuss the impact of SCS technologies on both basic research on and clinical applications for HM. In this article, we summarized the most recent studies that have used SCS methods to observe the clonal diversity, especially in HM (Table 10.1).

10.3.1 SCS for AML

Serial mutations and/or epigenetic events in self-renewal hematopoietic stem cells (HSCs) can lead to the development of HM [20]. Although the clonal evolution of leukemia has been difficult to study, it has been shown that some relapsed pediatric acute lymphoblastic leukemia (ALL) and AML cases were derived from a preleukemic HSC clone that acquired additional new mutations, thereby resulting in leukemic relapse [7, 51].

Improvements in SCS approaches have permitted single cell genomic analysis, and these methods are instantly being used to observe clonal evolution and heterogeneity in leukemia. The RNA SCS methods allow us to identify preleukemic HSCs that might precede the onset of leukemia. Exome sequencing techniques for single cells have already been used to identify diverse mutations within preleukemic subclones in one study [17]. The study revealed cellular and genomic changes from HSCs to the dominant presenting leukemic clone. After the establishment of malignant leukemic stem cells (LSCs), they underwent subsequent evolution, which resulted in the formation of diverse subclones

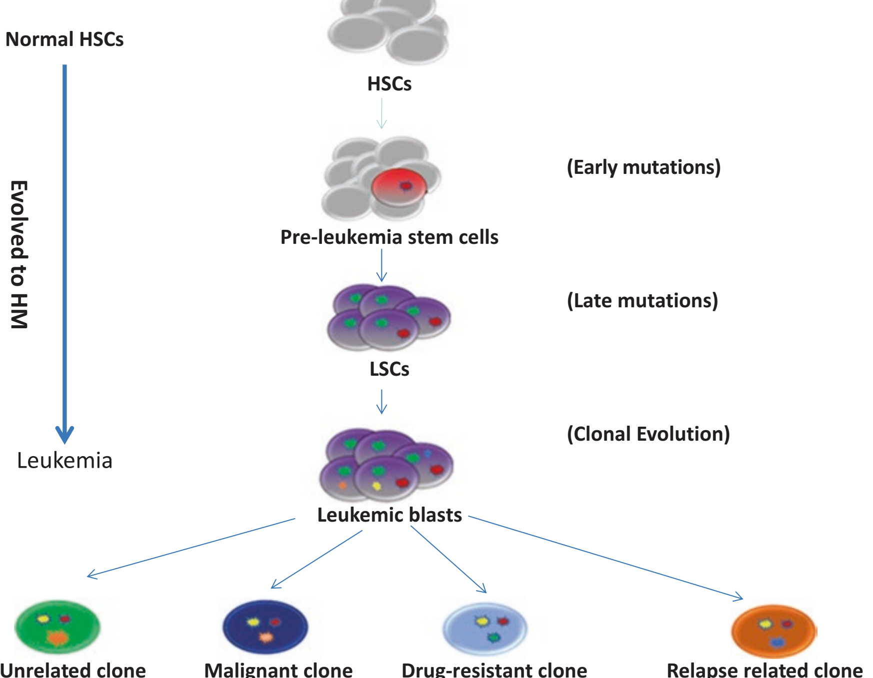


Fig. 10.2 The potential clonal origin and clonal evolution among HM subpopulations
First, genetic mutations lead to establishment of the initial malignant leukemic stem cells (LSCs). Then, LSCs undergo multiple distinct mutations, resulting in the formation of diverse subpopulations with increasing malignancy and refractoriness to therapy

with increasing malignancy and refractoriness to therapy (Fig. 10.2).

Meanwhile, LSCs are considered to originate from hematopoietic stem or progenitor cells. LSCs are capable of initiating and maintaining leukemia [37]. The existence of LSCs is associated with poor prognosis, treatment failure and leukemia relapse [50]. Leukemia can only be curved by the removal of LSCs. However, at present, separation of rare LSCs from the heterogenous leukemic subpopulation is still the main challenge. In the heterogeneous process of leukemia, averaged measurements of the total cell population may miss the LSC subpopulations among leukemic clones. AML is a clonal neoplastic disorder, and most patients die from disease relapse, which is associated with clonal evolution at a cytogenetic level [19]. Single-cell gene expression was implemented to investigate AML intra-tumor heterogeneity in an MLL-AF9 mouse model. The hereditary information supported that the 'stemness' of LSCs was different from that of general stem cells [38] .

Cellular heterogeneity and clonal evolution are fundamental characteristics of HM. Single-cell phenotypes are time and space dependent and involve multiple mutations and genetic changes. SCS methods can be used to analyze the genome, transcriptome and epigenome at a single cell resolution, which could help identify mutated subpopulations that might subsequently contribute to drug resistant and disease relapse. RNA-Seq was applied to investigate 20 cells that were isolated from the bone marrow of an AML patient. The result demonstrated that single-cell genomic profiling might be useful for individual cell categorization and might improve the accuracy of diagnosis and improve targeted therapies for AML [47].

Other works have further shown the prognostic value of SCS methods in analyzing heterogeneity and clonal evolution in AML, MDS and CML [13, 19, 41]. Researchers have combined multiple displacement amplification and Sanger sequencing to assess genetic mutations of single cells in inv. (16) AML and have verified the co-occurrence of *CAND1*, *PTPRT*, and *DOCK6* mutations within the same AML clone [33]. These researchers suggested that the *PTPRT* mutation, which is potentially associated with increased proliferation during AML development, occurred after *CAND1* and *DOCK6* mutations [33]. Moreover, DNA SCS was applied to analyze single leukemic cells that were separated from three patients with secondary AML derived from MDS in one study [16]. The result showed the advantages of integrated analyses of genetic variants found in bulk samples and single cells. This might help locate mutations within clonal hierarchies and provide insight into the clonal hierarchy in heterogeneous samples.

A central challenge in curing leukemia is the elimination of leukemia cells or MRD involving heterogeneous subpopulations that are defined by different mutated genes. The determination of how these genes mutated and how the single cell evolved requires single-cell analysis. Recently, the sleeping beauty capture hybridization sequencing technique was developed, and it simplified the transposon-based sequencing of mutations in a myeloid leukemia mouse model [27]. By sequencing transposon insertions of single leukemia cells, sleeping beauty capture hybridization sequencing was able to identify leukemogenesis genes and possible cooperating cancer genes.

10.3.2 SCS for ALL

Recent studies that have employed SCS to observe single-cell genomics have permitted researchers to understand clonal diversity and the evolution of leukemia cell genome with higher resolution than before. Single-cell whole genome sequencing [42] and AneuFinder were implemented to detect heterogeneity of the chromosome karyotype in untreated mouse T-ALL/lymphoma and human B-ALL samples, respectively [2]. These methods revealed severe intratumor karyotype heterogeneity in the B-ALL samples from children and may become important tools for therapy stratification and the prediction of disease outcome [2].

In another study, targeted deep-exome sequencing of 1479 single cells from six patients was performed to dissect clonal origins among children ALL. This study observed sequential deletions, single-nucleotide variants and IgH sequence rearrangements in pediatric ALL patients. The genetic information indicated that IgH rearrangements could occur prior to the acquisition of mutations and might continue to occur in even the most evolved clones. This study confirmed the temporal sequence of mutations in ALL development and offered a better understanding of the pathophysiological mechanism of pediatric ALL [12]. Thus, from the clonal evolution perspective, SCS technologies will likely help clinicians better understand the pathogenesis of diseases at a high-resolution and provide individualized treatments that are specifically optimized for their patients.

Despite efficient chemotherapy, many leukemia patients still ultimately relapse. Therefore, the development of targeted strategies that obviate neoplasm cells in MRD is expected to improve the prognosis of ALL patients. To considerably comprehend MRD biology in leukemia, one study established some preclinical mouse models to mimic MRD and relapse in patients. Then, the study merged single cell and bulk sample information to discover distinctly expressed genes in ALL. RNA-Seq results illustrated that the leukemia-like cells from ALL patients were different during MRD. Furthermore, they identified a rare subclone of leukemia cells with characteristics of leukemia initiation, long-term dormancy and therapeutic tolerance [8]. In conclusion, we found that, like solid neoplasm, the heterogeneous cells of ALL can display properties of clonal evolution and therapeutic resistance.

10.3.3 SCS for CML

By exploiting genetic mutations of particular cancer cell subpopulations, researchers can develop highly effective targeted regimen. For instance, tyrosine kinase inhibition (TKI) has shown potent efficacy for CML treatment [22]. Nonetheless, patients in the chronic phase of CML possessed rare CML LSCs that were selectively tolerant to TKI therapy and frequently induced relapse following treatment cessation [13]. Moreover, approaches for analyzing *BCR-ABL*-positive LSCs throughout disease progression are not presently available.

In one study, Smart-seq2, a whole transcriptome approach, was applied to detect the *BCR-ABL* gene in single cells isolated from the K562 cell line. This study identified that the transcriptome of CML LSCs that were tolerant to TKI was distinct from that of silent, normal HSCs. Furthermore, dysregulation of drug resistance specific mutations and relevant pathways was observed in this type of CML LSCs. These findings may benefit the development of selectively targeted therapy for drug-resistant CML [13]. Single-cell genomic analysis further demonstrated that TKI therapy induced heterogeneity and genetic changes in CML LSCs. Another study observed many genes from single-cell analyzes that may define LSC heterogeneity in CML patients at diagnosis and following TKI treatment. The results illustrated differencess in the reactions to TKI therapy among different subclones [45].

10.3.4 SCS for Myeloproliferative Neoplasms (MPN)

MPN are hematopoietic neoplasms that stem from hereditary mutations acquired in HSCs or their progenitors that result in abnormal differentiation and an increase in the proliferation of erythroid, megakaryocytic, and granulocytic lineages. MPN patients frequently express the *JAK2-V617F* gene. One study employed a multiple displacement amplification based high-throughput DNA SCS method to analyze tumor evolution in *JAK2-V617F* negative MPN patients. Based on the genomic information of 58 single cell exomes, the researchers demonstrated that essential thrombocythemia was, in all likelihood, of monoclonal genesis and discovered several specific genes that might play roles in essential thrombocythemia initiation and progression [15]. The study suggested that SCS methods contributed to the characterization of the genetic architecture and clonal evolution of HM.

10.3.5 SCS for MM

MM is a type of plasma cell malignancy attributed to the clonal proliferation of myeloma cell subpopulations which exhibit significant complex genetic heterogeneity in bone marrow [35]. Despite recent improvements in therapeutic strategies, including proteasome inhibitors and immunomodulators, MM, nevertheless, continues to be a predominantly incurable disease due to relapse. MM is characterized by considerable genetic variants among myeloma cell that play a significant role in the heterogeneity of myeloma evolution, clinical progression and drug resistance. Not every patient is equally sensitive to therapy, and most patients develop treatment resistance during therapy. Intra-tumor heterogeneity and changes in dominant clones are crucial causes of myeloma progression and resistance to treatments [18]. Recent SCS studies have shown that MM subpopulations are especially heterogeneous and may lead to relapse. In one study, Single Cell Analysis of Targeted Transcriptome was employed to detect the genomic information of single cells isolated from myeloma cell lines and untreated MM patients [28]. The adoption of Single Cell Analysis of Targeted Transcriptome could contribute to the analysis of intra-tumor heterogeneity in the clinic and improve drug selection based on subclonal cellular reactions.

MM remains an incurable disease, with almost all patients ultimately reaching a treatment-refractory state. Constant clonal evolution and

genetic heterogeneity of MM are a likely explanations for the emergence of drug resistance [3]. The monitoring of MM genomic evolution and its effect on therapy by serial bone marrow biopsy is invasive and painful. Using RNA SCS, Lohr et al. revealed that CTCs provided genetic information at similar or increased sensitivity compared to bone marrow MM cells [24]. Single CTCs RNA-Seq facilitates the categorization of MM and the quantitative evaluation of genes related to prognosis. By dissecting CTCs by SCS methods at different time points during diagnosis and treatment, hematologists can track clonal evolution and precisely alter therapeutic strategies before drug tolerance occurs.

SCS plays an important role in the measurement of intra-tumor heterogeneity, which is closely associated with HM genome mutation, evolution, and drug resistance. Furthermore, efficacy of and cell responses to targeted drugs can be identified by SCS technologies [43]. Over all, studies performed by SCS technologies to detect LSCs, CTCs, or monitor MRD in HM subpopulations shed light on precise therapies that can target complex heterogeneous subpopulations and improve patient outcomes (Fig. 10.3).

10.4 The Potential Therapeutic Value of Heterogeneity and Clonal Evolution in HM

The past decade has been a noteworthy period of advancement, particularly in HM therapy. ABL inhibitors (for example, imatinib) have illustrated therapeutic benefits, which is a significant aspect of precision medicine [14]. However, drug resistance has still been observed in a number of patients who were expected to exhibit a better reaction. Moreover, targeted therapies have already achieved highly effective responses; however, they have only provided short-term relief, as genetic changes within multiple subpopulations eventually lead to relapse.

Clonal heterogeneity is thought to underlie cancer plasticity. Certain genetic marks may lead to the accumulation of extra somatic mutations, and previous subclones may show drug resistance or evolve to become dominant clones [22]. Traditional methods of sequencing bulk samples reveal mutations from the whole subpopulations together and mask the clonal heterogeneity, which may have significant functions in the progression of HM [31]. Thus, only single cell analysis can accurately identify subpopulations that may contribute to drug resistance and disease relapse. Understanding the heterogeneity and clonal evolution of HM are key elements and obstacles to improving targeted therapies in the context of precision medicine. Thus, SCS methods may provide reliable information with respect to disease heterogeneity, evolution, and treatment (Fig. 10.4).

10.5 Conclusion

Precision medicine involves personalizing treatment based on a patient's genetic information, which requires a better understanding of the precise diagnosis and treatment at a single cell level [1]. Researchers believe that heterogeneity and clonal evolution are prevalent in HM and inhibit precise diagnosis and treatment. Moreover, with the accumulation of multiple mutations, clonal evolution of HM subpopulations may lead to the emergence of more aggressive and resistant phenotypes, which are associated with adverse clinical outcomes.

In this article, we summarized the most recent studies that have used SCS methods to observe the clonal diversity, especially in HM. In addition, we concluded that advances in SCS technologies may enable the dissection of genetic heterogeneity and the clonal evolution of HM at a single cell resolution during diagnosis, treatment, MRD monitoring, or even at disease relapse. The development and application of SCS technologies can improve targeted treatment and personalized medicine by delineating the clonal substructure and by identifying molecular targets that are present in the most malignant subpopulations.

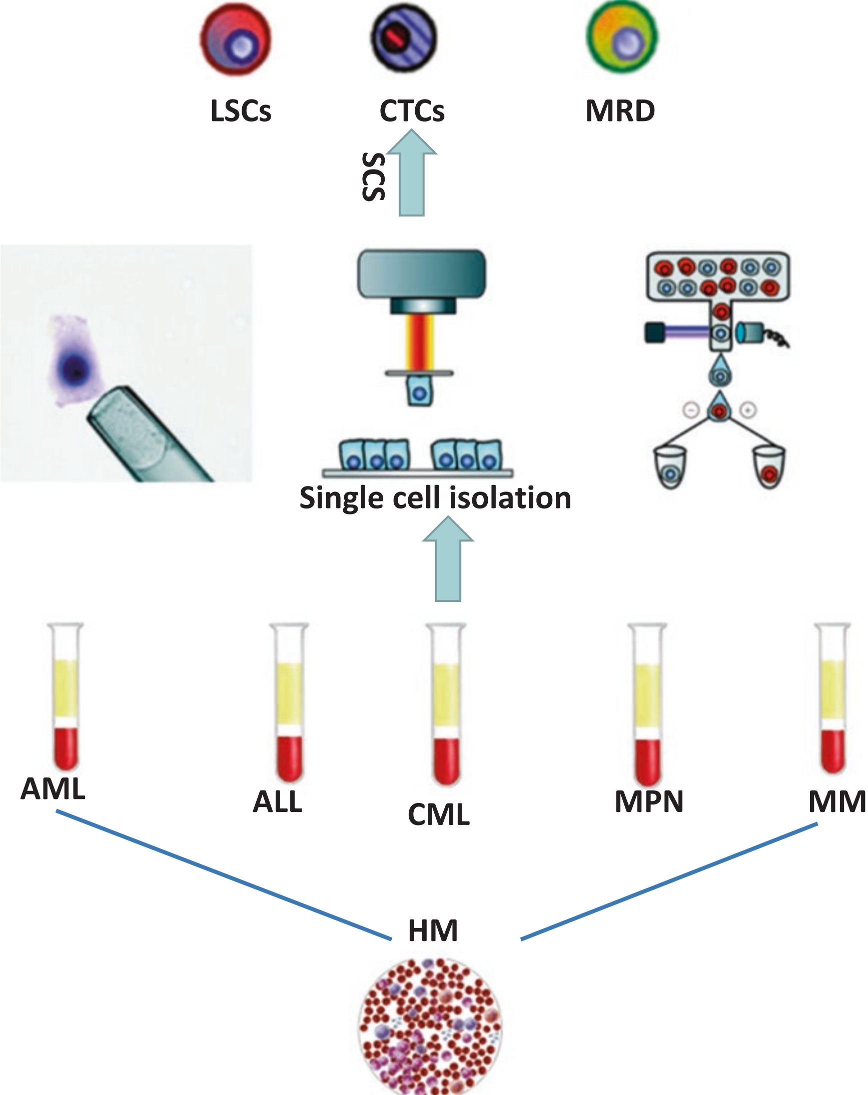


Fig. 10.3 SCS techniques in HM diagnosis, relapse monitoring and drug-resistance detection
SCS techniques can detect rare leukemia subpopulations in the form of leukemic stem cells (LSCs), circulating tumor cells (CTCs) or minimal residual disease (MRD) that may contribute to HM formation, progression, relapse or drug-resistance at a single cell level

Due to rapid development, SCS methods will have numerous translational application in the clinic in the near future, particularly for early detection, prognostics, non-invasive monitoring, and guiding targeted therapy [31]. Following the transition of these tools into the clinic, we expect that SCS will have a large impact on reducing morbidity and improving the quality of life for HM patients.

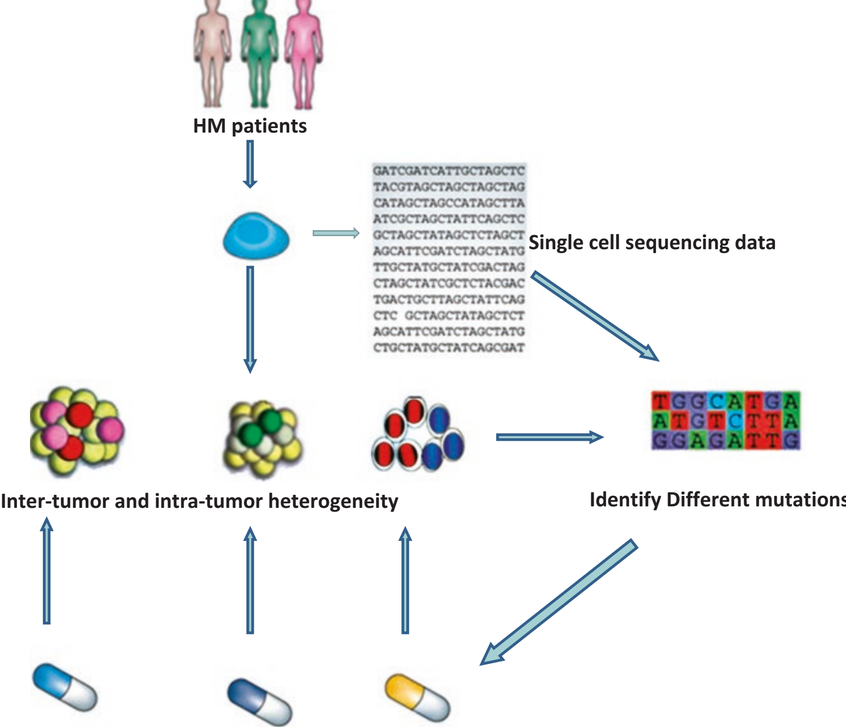


Fig. 10.4 SCS applications in analyzing heterogeneity and clonal evolution HM
Potential employment of DNA/RNA SCS techniques in analyzing heterogeneity and clonal evolution, which are the main obstacles to precise diagnosis of and personalized therapies for HM

Acknowledgements This study was partially supported by the National Natural Science Foundation of China (No. 81273259,No. 81471589), the Health Bureau of Henan Province, P.R. China (No. 201201005) and the Foundation and Frontier Research Grant from the Henan Provincial Science and Technology Bureau, P.R. China (No.142300410078).

Conflicts of Interest The authors disclose that they have no relevant conflicts of interest.

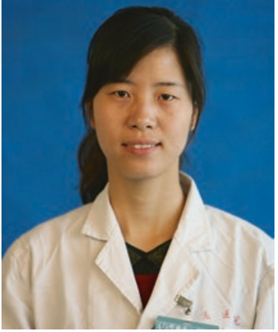
References

1. Aronson SJ, Rehm HL (2015) Building the foundation for genomics in precision medicine. Nature 526:336–342
2. Bakker B, Taudt A, Belderbos ME, Porubsky D, Spierings DC, de Jong TV, Halsema N, Kazemier HG, Hoekstra-Wakker K, Bradley A et al (2016) Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. Genome Biol 17:115
3. Binder M, Rajkumar SV, Ketterling RP, Dispenzieri A, Lacy MQ, Gertz MA, Buadi FK, Hayman SR, Hwa YL, Zeldenrust SR et al (2016) Occurrence and prognostic significance of cytogenetic evolution in patients with multiple myeloma. Blood Cancer J 6:e401
4. Chu MP, Kriangkum J, Venner CP, Sandhu I, Hewitt J, Belch AR, Pilarski LM (2017) Addressing heterogeneity of individual blood cancers: the need for single cell analysis. Cell Biol Toxicol 33:83–97
5. Court CM, Ankeny JS, Sho S, Hou S, Li Q, Hsieh C, Song M, Liao X, Rochefort MM, Wainberg ZA et al (2016) Reality of single circulating tumor cell sequencing for molecular diagnostics in pancreatic Cancer. J Mol Diagn: JMD 18:688–696
6. Del Giudice I, Marinelli M, Wang J, Bonina S, Messina M, Chiaretti S, Ilari C, Cafforio L, Raponi S, Mauro FR et al (2016) Inter- and intra-patient clonal and subclonal heterogeneity of chronic lymphocytic leukaemia: evidences from circulating and lymph nodal compartments. Br J Haematol 172:371–383

7. Ding L, Ley TJ, Larson DE, Miller CA, Koboldt DC, Welch JS, Ritchey JK, Young MA, Lamprecht T, McLellan MD et al (2012) Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. Nature 481:506–510
8. Ebinger S, Ozdemir EZ, Ziegenhain C, Tiedt S, Castro Alves C, Grunert M, Dworzak M, Lutz C, Turati VA, Enver T et al (2016) Characterization of rare, dormant, and therapy-resistant cells in acute lymphoblastic leukemia. Cancer Cell 30:849–862
9. Fakhri B, Vij R (2016) Clonal evolution in multiple myeloma. Clin Lymphoma, Myeloma Leuk 16(Suppl):S130–S134
10. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schonegger A, Klughammer J, Bock C (2015) Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. Cell Rep 10:1386–1397
11. Garg M, Nagata Y, Kanojia D, Mayakonda A, Yoshida K, Haridas Keloth S, Zang ZJ, Okuno Y, Shiraishi Y, Chiba K et al (2015) Profiling of somatic mutations in acute myeloid leukemia with FLT3-ITD at diagnosis and relapse. Blood 126:2491–2501
12. Gawad C, Koh W, Quake SR (2014) Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. Proc Natl Acad Sci U S A 111:17947–17952
13. Giustacchini A, Thongjuea S, Barkas N, Woll PS, Povinelli BJ, Booth CAG, Sopp P, Norfo R, Rodriguez-Meira A, Ashley N et al (2017) Single-cell transcriptomics uncovers distinct molecular signatures of stem cells in chronic myeloid leukemia. Nat Med 23:692–702
14. He M, Xia J, Shehab M, Wang X (2015) The development of precision medicine in clinical practice. Clin Transl Med 4:69
15. Hou Y, Song L, Zhu P, Zhang B, Tao Y, Xu X, Li F, Wu K, Liang J, Shao D et al (2012) Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. Cell 148:873–885
16. Hughes AE, Magrini V, Demeter R, Miller CA, Fulton R, Fulton LL, Eades WC, Elliott K, Heath S, Westervelt P et al (2014) Clonal architecture of secondary acute myeloid leukemia defined by single-cell sequencing. PLoS Genet 10:e1004462
17. Jan M, Snyder TM, Corces-Zimmerman MR, Vyas P, Weissman IL, Quake SR, Majeti R (2012) Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. Sci Transl Med 4:149ra118
18. Keats JJ, Chesi M, Egan JB, Garbitt VM, Palmer SE, Braggio E, Van Wier S, Blackburn PR, Baker AS, Dispenzieri A et al (2012) Clonal competition with alternating dominance in multiple myeloma. Blood 120:1067–1076
19. Klco JM, Spencer DH, Miller CA, Griffith M, Lamprecht TL, O'Laughlin M, Fronick C, Magrini V, Demeter RT, Fulton RS et al (2014) Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. Cancer Cell 25:379–392
20. Kowalczyk MS, Tirosh I, Heckl D, Rao TN, Dixit A, Haas BJ, Schneider RK, Wagers AJ, Ebert BL, Regev A (2015) Single-cell RNA-seq reveals changes in cell cycle and differentiation programs upon aging of hematopoietic stem cells. Genome Res 25:1860–1872
21. Kriangkum J, Motz SN, Mack T, Beiggi S, Baigorri E, Kuppusamy H, Belch AR, Johnston JB, Pilarski LM (2015) Single-cell analysis and next-generation Immuno-sequencing show that multiple clones persist in patients with chronic lymphocytic leukemia. PLoS One 10:e0137232
22. Landau DA, Carter SL, Getz G, Wu CJ (2014) Clonal evolution in hematological malignancies and therapeutic implications. Leukemia 28:34–43
23. Liang J, Cai W, Sun Z (2014) Single-cell sequencing technologies: current and future. J Genet Genomics = Yi chuan xue bao 41:513–528
24. Lohr JG, Kim S, Gould J, Knoechel B, Drier Y, Cotton MJ, Gray D, Birrer N, Wong B, Ha G et al (2016) Genetic interrogation of circulating multiple myeloma cells at single-cell resolution. Sci Transl Med 8:363ra147
25. Lohr JG, Stojanov P, Carter SL, Cruz-Gordillo P, Lawrence MS, Auclair D, Sougnez C, Knoechel B, Gould J, Saksena G et al (2014) Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. Cancer Cell 25:91–101
26. Manier S, Salem KZ, Park J, Landau DA, Getz G, Ghobrial IM (2017) Genomic complexity of multiple myeloma and its clinical implications. Nat Rev Clin Oncol 14:100–113
27. Mann KM, Newberg JY, Black MA, Jones DJ (2016) Analyzing tumor heterogeneity and driver genes in single myeloid leukemia cells with SBCapSeq. Nat. Biotech 34:962–972
28. Mitra AK, Mukherjee UK, Harding T, Jang JS, Stessman H, Li Y, Abyzov A, Jen J, Kumar S (2016) Single-cell analysis of targeted transcriptome predicts drug sensitivity of single cells within human myeloma tumors. Leukemia 30:1094–1102
29. Mosna F, Capelli D, Gottardi M (2017) Minimal residual disease in acute myeloid leukemia: still a work in progress? J Clin Med 6:57
30. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472:90–94
31. Navin NE (2015a) Delineating cancer evolution with single-cell sequencing. Sci Transl Med 7:296fs229
32. Navin NE (2015b) The first five years of single-cell cancer genomics and beyond. Genome Res 25:1499–1507
33. Niemoller C, Renz N, Bleul S, Blagitko-Dorfs N, Greil C, Yoshida K, Pfeifer D, Follo M, Duyster J, Claus R et al (2016) Single cell genotyping of exome sequencing-identified mutations to characterize the clonal composition and evolution of inv(16) AML in a CBL mutated clonal hematopoiesis. Leuk Res 47:41–46

34. Paguirigan AL, Smith J, Meshinchi S, Carroll M, Maley C, Radich JP (2015) Single-cell genotyping demonstrates complex clonal diversity in acute myeloid leukemia. Sci Transl Med 7:281re282
35. Rajkumar SV (2016) Multiple myeloma: 2016 update on diagnosis, risk-stratification, and management. Am J Hematol 91:719–734
36. Rato S, Golumbeanu M, Telenti A, Ciuffi A (2016) Exploring viral infection using single-cell sequencing. Virus Res
37. Riether C, Schurch CM, Ochsenbein AF (2015) Regulation of hematopoietic and leukemic stem cells by the immune system. Cell Death Differ 22:187–198
38. Saadatpour A, Guo G, Orkin SH, Yuan GC (2014) Characterizing heterogeneity in leukemic cells using single-cell gene expression analysis. Genome Biol 15:525
39. Saez B, Walter MJ (2017) Splicing factor gene mutations in hematologic malignancies. Blood 129:1260–1269
40. Sant M, Minicozzi P, Mounier M, Anderson LA, Brenner H, Holleczek B, Marcos-Gragera R, Maynadie M, Monnereau A, Osca-Gelis G et al (2014) Survival for haematological malignancies in Europe between 1997 and 2008 by region and age: results of EUROCARE-5, a population-based study. Lancet Oncol 15:931–942
41. Shastri A, Will B, Steidl U, Verma A (2017) Stem and progenitor cell alterations in myelodysplastic syndromes. Blood 129:1586–1594
42. van den Bos H, Spierings DC, Taudt AS, Bakker B, Porubsky D, Falconer E, Novoa C, Halsema N, Kazemier HG, Hoekstra-Wakker K et al (2016) Single-cell whole genome sequencing reveals no evidence for common aneuploidy in normal and Alzheimer's disease neurons. Genome Biol 17:116
43. Wang W, Zhu B, Wang X (2017) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33:423
44. Wang Y, Waters J, Leung ML, Unruh A, Roh W, Shi X, Chen K, Scheet P, Vattathil S, Liang H et al (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 512:155–160
45. Warfvinge R, Geironson L, Sommarin MNE (2017) Single-cell molecular analysis defines therapy response and immunophenotype of stem cell subpopulations in CML. Blood 129:2384–2394
46. Xue Z, Huang K, Cai C, Cai L, Jiang CY, Feng Y, Liu Z, Zeng Q, Cheng L, Sun YE et al (2013) Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. Nature 500:593–597
47. Yan B, Hu Y, Ban KHK, Tiang Z, Ng C, Lee J, Tan W, Chiu L, Tan TW, Seah E et al (2017) Single-cell genomic profiling of acute myeloid leukemia for clinical use: a pilot study. Oncol Lett 13:1625–1630
48. Ye F, Huang W, Guo G (2017) Studying hematopoiesis using single-cell technologies. J Hematol Oncol 10:27
49. Yuan GC, Cai L, Elowitz M, Enver T, Fan G, Guo G, Irizarry R, Kharchenko P, Kim J, Orkin S et al (2017) Challenges and emerging directions in single-cell analysis. Genome Biol 18:84
50. Zhou J, Chng WJ (2014) Identification and targeting leukemia stem cells: the path to the cure for acute myeloid leukemia. World J Stem Cells 6:473–484
51. Zhou Y, Kanagal-Shamanna R, Zuo Z, Tang G, Medeiros LJ, Bueso-Ramos CE (2016) Advances in B-lymphoblastic leukemia: cytogenetic and genomic lesions. Ann Diagn Pathol 23:43–50

Dr. Mingyue Shi got her M.D. degree in internal medicine at the University of Zhengzhou in 2016, with the thesis on lymphocyte subsets and cytokines in myeloma patients. During her postgradual studies she spent 3 years in the department of hematology in Henan Provincial People's Hospital. In 2016 she also obtained a doctor's admission in the Academy of Medical Sciences, Zhengzhou University. Currently Dr. Mingyue Shi is studying at the Central Laboratory of Henan Provincial People's Hospital. And she has been in one of the national natural science fund surface projects of Prof. Kai Sun. Her main research interest focuses on immune mechanism and potential targeted therapy of hematologic malignancies. Dr. Mingyue Shi has published 3 scientific papers in both international and national well-recognized journals.

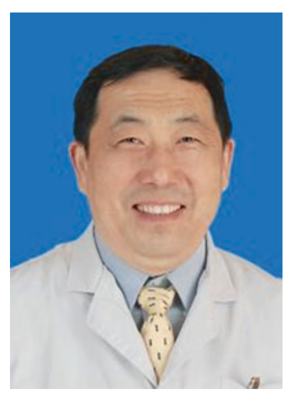


Dr. Xiaoyan Dong got her M.D. degree in hematology at Jilin University in 2011, with the professional direction for basic and clinical research in leukemia. During her postgradual studies she spent 3 years in the Cancer Center at The First Hospital of Jilin University. Since 2011 she has been engaged in the pathological diagnosis of hematologic diseases at the Institute of Hematology/Department of Hematology, Henan Provincial People's Hospital. Her main research interest focuses on the molecular biology research of hematologic malignancies. Dr. Xiaoyan Dong has published several scientific papers in both international and national well-recognized journals.

Dr. Lei Huo got his master admission in hematology at Zhengzhou University People's Hospital in 2015. During his postgradual studies he spent 2 years in central laboratory of Henan Provincial People's Hospital. And he has participated in Prof. Sun's national natural science fund surface projects with responsible for animal experiments. His main research interest focuses on immune mechanism of HSCT.

Dr. Xiaobin Wei received a BSc degree From Henan University of Science and Technology in china, a MSc and a PhD from Northwest A&F University in china. He previously studied the structure and function of proteins associated with organism's genome stability such as helicase and telomerase. At present, his research focuses on diagnosis and treatment of hematologic malignancies by using next-generation sequencing and genome-editing technology. Dr. Xiaobin Wei has published 1 scientific paper in international well-recognized journal.

Dr. Fang Wang got her M.D. degree in internal medicine at the University of Zhengzhou in 2017, with the thesis on the basic and clinical study of hematologic diseases. During her postgradual studies she spent 2 years in the department of hematology in Henan Provincial People's Hospital. Currently she is working at the Henan Province People's Hospital as a research scientist. Her main research interest focuses on single – cell sequencing technology and its application. Dr. Fang Wang has published two scientific papers in national well-recognized journal.



Dr. Kai Sun is the director of the Department of Hematology, master and doctoral mentor, Henan Provincial People's Hospital/Zhengzhou University People's Hospital, China. From 1998 to 2000, he received postdoctoral training at the postdoctoral workshop of clinical medicine in Beijing Medical University (instructor: Dr Daopei Lu academician). Later, he worked with William J. Murphy, a famous American immunologist, for 11 years in basic research work on transplantation immunization and tumor immunity. His current research is focused on the improvement of hematopoietic stem cell transplantation and better understanding of the relevant immune mechanism. Dr. Kai Sun has more than 10 publications with h-index: 13, including the first author of the studies in BLOOD and PNAS on the separation of GVHD and GVL. Dr. Kai Sun's topic in the PNAS journal reported that bortezomib could effectively separate GVHD and GVL effect, and the paper has been cited more than 200 times. Dr. Kai Sun has already completed a national natural science fund surface project to explain the role and immune mechanism of recipient "unlicensed" Ly49+ NK cell subsets for specific regulation in allograft rejection. And another national natural science fund surface project about the immune mechanism of the synergistic effect of SAHA and low-dose bortezomib in the separation of GVHD/GVL responses is on the way.

11 Application of Single Cell Sequencing in Cancer

Lan Yu, Hua Zhao, Li Meng, and Cuilian Zhang

Abstract

Cancer is a heterogenetic disease at both the level of clinical manifestation and the level of the genome. Single-cell sequencing provides an unprecedented means of characterizing the intra-tumor heterogeneity and detecting and analyzing the genomes of cancer cells. These data will help to reconstruct the understanding of the evolutionary lineage of cancer cells. In the future, single-cell technology is believed to be a useful tool in diagnostic and prognostic application in oncology. The application of single cell technology in clinics will make it possible to detect cancer non-invasively at early stages and to develop precision medicine. In this chapter, we review the research and application status of the single cell technology in cancer.

Keywords

Single cell sequencing · Cancer · Heterogeneity · Circulation tumor cells

L. Yu · H. Zhao · L. Meng · C. Zhang (✉)
The Reproductive Institute of Henan Provincial People's Hospital/Reproductive Hospital of Henan Province, Zhengzhou, China

11.1 Introduction

The human body is estimated to have approximately 3.72×10^{13} cells [1]. A single cell is the smallest structural and functional unit of the body. Each cell varies, having a different origin and background. In addition, each single cell possesses distinct DNA, RNA, and protein conversions [2]. Scientists started to notice and study single cells due to the invention of the microscope. The concept of cellular heterogeneity was proposed in 1957 and gradually gained acceptance [3]. Nicholas Navin, who was called the father of single-cell sequencing, proposed the pioneering theory that the complexity of intratumoral heterogeneity can be explained by the genome of a single tumor cell [4]. In the past, it was not an easy task to isolate a single cell, due to limitations to the existing techniques. It was only recently that single cellular biology and heterogeneity have been studied because of the exploitation of genome-scale approaches and methods for obtaining intact single cells, such as laser-capture microdissection (LCM), flow cytometry using fluorescence-activated cell sorting (FACS), and micromanipulation. The human genome contains at least 3.2 billion base pairs; thus, it was a huge challenge to profile the whole genome in a short time with an acceptable cost. However, the breakthrough of next-generation sequencing and RNA-sequencing make it possible to assess the human genome in a rapid and reliable way [5]. In

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_11

addition to the notable techniques mentioned above, whole genome amplification (WGA) is another milestone that makes single-cell sequencing a reality. WGA is a bridge between the isolation of a single cell and sequencing.

Heterogeneity exists not only in different organisms but also in different organs or tissue within one organism. Furthermore, abundant studies have demonstrated that heterogeneity exists generally in the same organ and tissue from one organism [1]. This heterogeneity is extremely prominent in disease.

Cancer is one of the most complex and heterogeneous diseases, where genomic, transcriptomic, and proteomic changes come up during tumorigenesis, mutation, stochastic variation, or environmental changes. Moreover, the polyclonal nature of cancer makes the genomic heterogeneity more complicated, resulting in significant heterogeneity among different cells within each tumor. However, the study of genomics, proteomics, transcriptomics, and biomarkers analysis in cancers is based on a mass of cells or tissue samples. Current experimental and clinical regiments assume that all of the cells in tumors are homogeneous. Traditional research and cancer treatment has a limited ability to characterize the nature of the cancer because it ignores the heterogeneity of different cells. Therefore, diagnosis based on biopsy (bulk cells) could underestimate the extent to which there is heterogeneity within the tumor and could fail to identify all of the possible solutions [6]. This underestimation leads directly to improper clinical decision making. Consequently, overlooking the heterogeneity of the tumor is considered the reason for cancer treatment failure and relapse because a treatment that is effective for one type of tumor cell may be useless for another, driving further evolution and eventually leading to drug resistance [7].

In recent years, there has been a boom in the development of the single cell methods of study, such as the microfluidic-based single-cell sorting methods [8, 9], high-throughput multiplexed quantitative PCR (qPCR) [10–14], mass cytometry-based proteomic strategies [15, 16], single-cell transcriptome sequencing [17], single-cell whole genome sequencing, single-cell epigenomic sequencing [18], next generation sequencing, and computational methods for analyzing single cell data. This boom has led to an unprecedented revolution in the field of cancer research and opened the door to a vast number of possibilities [19–22]. For example, the newly developed single-cell sequencing techniques help us to understand the intra-tumoral heterogeneity and map different clones in tumors, as well as enhancing the capacity for analyzing tumor stem cells and circulating tumor cells that are rare but informative for cancer research [23].

11.2 Current Studies of Single-Cell Sequencing on Cancer Research

Genome instability is a hallmark of cancer. With the development of cancer, it is believed that the primal mutated cells experience clonal evolution. Meanwhile, cancer stem cells develop in a hierarchical way [24]. Knowledge of the evolution of cancer and the cancer genome in a spatial and temporal manner will have a significant impact on the understanding of cancer initiation, evolution, metastasis and relapse and will therefore lead to the development of patient-specific treatment strategies [25]. See Fig. 11.1 for details.

11.2.1 Breast Cancer

Single-cell sequencing has been used in cancer study. The first report was single-cell sequencing in two triple-negative breast cancers. The study showed that one tumor was highly monogenomic, but the other was heterogeneous and contained distinct subpopulations [26]. Studies on several breast cancer groups indicated that copy number variation, duplication or deletion of the chromosome may occur at an early stage during tumorigenesis. However, point mutations evolve gradually in tumor development, which could drive clonal diversity [27–29]. In different breast cancer cell lines, single cell RNA sequencing indicated that RNA transcripts have high variability, which was also evident at the protein

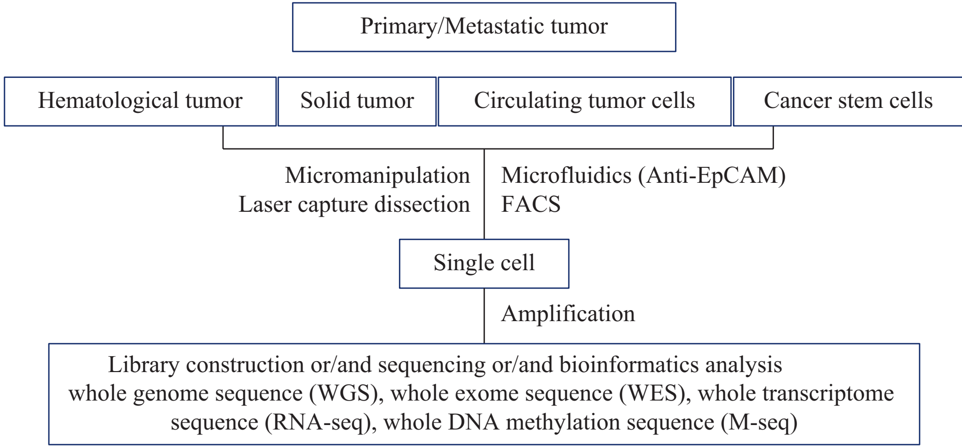


Fig. 11.1 The skeleton diagram of this chapter

level [30]. The protein differences could explain the increased metastatic capacity and resistance to chemotherapy. Whole transcriptome sequencing of a single MDA-MB-231 metastatic breast cancer cell showed that after 5 days of exposure to paclitaxel, the surviving cells expressed unique RNA variants that regulated cell adhesion, cell surface signaling, and microtubule organization [31]. This expression indicated that screening out the suspicious candidates by RNA sequencing could guide clinical drug application.

11.2.2 Ovarian Cancer

High grade serous ovarian cancer samples were used to study the cancer's heterogeneity at the single cell level. The study demonstrated the heterogeneity in the ovarian tumor. Meanwhile, this study classified the tumor based on bulk cells or single cell RNA sequencing. The results indicated that classifying the cancer by bulk cells or tumor molecular information could mask the gene expression pattern of the large groups of cells [32]. Malecki et al. developed methods to isolate single cells from the peritoneal fluid, blood, lymph, and ovarian cancer biopsies for sequencing and *in situ* hybridization. This novel work-flow and single cell analysis opened a new route to profile the whole spectrum of cancer clones at the molecular level, one cell clone at a time [33].

11.2.3 Lung Cancer

Adenocarcinoma of the lung is the most histological subtype of lung cancer. Single cells extracted from the lung adenocarcinoma patient-derived xenografts are analyzed by RNA-sequencing. Sixty-four genes were identified to be associated with poor prognosis [34]. Four groups were classified based on these gene changes and were exposed to chemotherapy. The results suggested that the group with cell-cycle quiescent and high ion channel transport genes showed a high survival rate [35]. Another study with lung adenocarcinoma cell lines by RNA sequencing also showed that cells with a high diversity in gene expression may be the reason for the acquisition of drug resistance in cancer.

11.2.4 Hematopoietic Tumors

Whole-genome sequencing (WGS) was used in single cancer cells isolated from patients diagnosed with myelodysplastic syndrome. These

patients all progressed to secondary acute myeloid leukemia [36]. The sequencing successfully identified genomic complexity, which was not evident in the analysis of the whole tumor. Sequencing of the single nucleotide variants (SNVs) was used to draw the clonal structure and evolutionary history of the acute lymphoblastic leukemia by analyzing 1479 single cells from six children diagnosed with acute lymphoblastic leukemia. Compared to other cancers, acute lymphoblastic leukemia was characterized by distinct clonal populations of cells where alterations in copy number preceded the occurrence of single nucleotide variants [37]. Studies on chromosomally unstable B cell acute lymphoblastic leukemia showed that single-cell sequencing could identify the subpopulations with copy number variants that were not detected in the whole tumor analysis [38]. The copy number variants were also observed when the tumor was engrafted into immunodeficient mice. These findings suggested that copy number variants in single cells may evolve in response to the microenvironment, such as in chemotherapy.

11.2.5 Renal Cell Cancer

Renal cell cancer accounts for more than 200,000 new cases and over 100,000 deaths worldwide each year [6]. Clear cell renal cell carcinoma is the most common renal cell and is considered to be a disease with a low mutation rate and few mutations shared among patients. Whole-genome sequencing was applied on 20 single cells from one clear cell renal cell carcinoma patient to study the intra-tumoral heterogeneity. The results showed that there were no significant sub-clonal populations detected, but many rare mutations were found in a small number of cancer cells [39]. It is not easy to detect these mutations with bulk-cell sequencing. The study displayed the intra-tumor genetic landscape of clear cell renal cell carcinoma at a single cell level and suggested that clear cell renal cell carcinoma may be more complex at the genetic level than expected.

11.2.6 Glioblastoma

Glioblastoma is the most common brain and central nervous system malignancy and has a poor prognosis. Glioblastoma is biologically aggressive because of its rapid growth and high invasiveness. Glioblastoma is characterized by extensive cellular and molecular heterogeneity. Epidermal growth receptor factor (EGFR) is one of the genes which is thought to affect the development and treatment of glioblastoma. Single-nucleus whole-genome sequencing studies demonstrated that the EGFR copy number was highly variable because of EGFR amplification [40]. Single cell RNA sequencing on 430 cells from five glioblastoma patients detected several oncogenic variants of EGFR at the single cell level [41]. Moreover, the study also showed high intra-tumoral cell heterogeneity in the molecular markers related to hypoxia and the immune response. In addition to EGFR and the genes mentioned above, this study also detected variation of several signaling molecules and cell surface receptors, which may contribute to targeted-therapy resistance [41].

The evidence suggests that the study of molecular heterogeneity at the individual cellular level may reveal the aggressive or drug-resistant subclones that are masked when analyzing large sections of the tumor. Meanwhile, study at single cell level provides a powerful approach, which may have significant impacts on the prediction of disease outcomes and provide important information for effective chemotherapy.

11.2.7 Circulating Tumor Cells

Clinical therapeutic regimens based on primary tumors often end with failure because of metastases and drug resistance. Metastasis is highly related to death and is difficult to detect in the early stages of the disease. Circulating tumor cells (CTCs) are a population of tumor cells that could mediate metastasis. CTCs were first observed a cancer patient's peripheral blood in

1869 [42]. To invade the blood vessels, the tumor cells first lose their epithelial features and cell polarity to become mesenchymal-like cells, which have increased invasiveness and resistance to apoptosis. This process is called epithelial to mesenchymal transition (EMT). The EMT cells also need to acquire the ability to be motile and the enzymatic activity to break the basement membrane and vascular wall. The EMT process has been studied for years but it is still not conclusive in terms of its involvement in metastases progress. Study on CTCs is an ideal approach for understanding this issue.

CTCs can be separated from white blood cells (WBCs) according to surface markers by using flow cytometry or micromanipulation to exclude contamination from WBCs. However, most CTCs undergo anoikis because most of them cannot survive in the blood stream. The enrichment of CTCs is a key step in our ability to utilize the aforementioned techniques because there is only one CTC in 10^9 normal cells [43]. This situation is a critical problem because the remaining CTCs are insufficient to provide DNA or RNA for next generation sequencing. Given the progress in developing whole genome amplification and whole transcriptome library construction, sequencing-based analyses for CTCs can be scaled down to the single-cell level [42]. CTCs are supposed to carry stem-like cell characteristics and genetic alteration has been observed in primary tumor. Therefore, the development of single-cell sequencing helps us to understand the clonal evolution and the role of CTCs in metastatic cancer. Although studies have shown that there are common or similar mutational profiles in primary carcinomas, metastases, and CTCs, we cannot ignore the important molecular heterogeneity that has been observed by next-generation sequencing because of the potential utility of CTCs in treatment [44]. This finding is observed because the CTCs are bridges between the primary tumor and metastasis; therefore, mutations in CTCs can affect the disease outcomes. In lung and prostate cancer, the most mutated genes or copy number variants are concurrently observed in the primary tumor, CTCs and the metastatic site [26, 45–47]. This finding suggests that the study of CTCs may provide primary tumor information for clinical decision making.

Single cell sequencing studies have demonstrated that CTCs exhibit tremendous cell-to-cell heterogeneity within the patient [48–51]. For example, the mutational heterogeneity in the TP53, PDGF, PI3K and catalytic subunit alpha genes have been observed among individual CTCs from women with metastatic breast cancer [52, 53]. Meanwhile, the TP53 mutational rate was different among the CTCs in breast cancer patients. Certain CTCs shared an identical mutation with the corresponding primary carcinoma, while other CTCs carry different mutations [54]. Although the mutations detected in CTCs could reflect the characteristics of the primary tumor, there are genomic changes that induce metastasis, which may occur primarily in CTCs [55]. For example, immune escape pathways are up-regulated in the CTCs of colorectal cancer patients, thereby implying a mechanism for CTC survival in the blood stream [56]. Previous studies that were verified by targeted deep sequencing showed that most of the seemingly exclusive mutations in CTCs are present in the primary tumor at a relatively low frequency in the minor subclones [42]. This phenomenon indicates that the mutations of cancer genome are dynamic changes that occur during disease progression. The progression may be consistent pretty much across the disease. Therefore, the genomic signatures of many individual CTCs from cancer patients may be more informative than traditional biopsies of the primary tumor for designing targeted therapies and monitoring therapeutic response [44].

Considering the constantly evolving nature of cancer, researchers have realized that certain cellular biomarkers also change dynamically, which reminds us to establish a reliable biomarker system for cancer progression [27, 57]. Other studies have demonstrated that the selective pressures exerted during systemic treatment drive clonal evolution and can result in the subclonal expansion of initially rare variants [58]. For example, EGFR T7900 mutation was increased in NSCLS patients after anti-EGFR treatment [59] and MET amplification was higher, as well, because of the

selection pressure during treatment [60]. Temporal heterogeneity is a notable issue which should be focused on; however, the application of longitudinal tumor sampling was previously impractical. Single-cell sequencing of CTCs is superior to biopsies because it is non-invasive and easily repeatable. Serial sequencing of CTCs provides us with a chance to monitor the tumor evolution at the longitudinal level.

Since the emergence of single-cell sequencing, less DNA or RNA is required for the techniques that are now used for genetic analyses, providing an earlier and more convenient opportunity for clinicians and researchers to obtain fundamental information regarding the metastatic process or for early tumor detection [4]. Furthermore, single-cell sequencing of CTCs may provide significant information to study whether tumor cells will evolve and become resistant to chemotherapy or targeted drugs [61]. However, the application of single CTC analysis in the clinic is currently far from developed because of the limitations in CTC detection and isolation. Rare platforms could be adapted to all types of cancer, and the situations among the different patients are highly diverse, making it difficult generally to evaluate the efficiency or accuracy of single CTC analysis in the clinic for now [42]. Although CTC research is still in its infancy, it is a promising field of study because of its scientific significance in elucidating cancer metastasis and clinical value in non-invasive cancer detection, prognosis, and diagnosis.

11.2.8 Cancer Stem Cell

Stem cells have the ability to perpetuate their lineage and differentiate to oriented cells under specific stimulation. Stem cells can also interact with the environment to maintain a balance between quiescence, proliferation, and regeneration [62]. Normal stem cells have the capacity to self-renew and maintain their homeostasis, while the homeostasis is disruptive in cancer stem cells. Except for homeostasis, cancer stem cells keep all aspect of stemness, such as the enhanced capacities for self-renewal, cloning, growing, metastasizing, homing, and proliferating, which sustains the progression of cancer [62]. Cancer stem cells are rare but they can recruit nutrients from their neighbor cells and escape from the immune system. So far, the origin of cancer stem cells is far from clear, but this is probably due to the accumulation of mutations over time. Current knowledge indicates that the therapeutic elimination of all cancer stem cells is not only required but might also be sufficient to cure cancer.

Currently, the study of cancer stem cells has become a prime research subject because the character of the cancer stem cells has been associated with tumor progression, metastasis, relapse, and drug resistance. The breakthrough in cancer stem cells research is considered a means for the development of anti-cancer therapeutics. The rarity, low immunogenicity, cellular heterogeneity, and small amount of available RNA in cancer stem cells has limited the understanding of tumor biology. However, the development of methods for the detection and isolation of stem cells, as well as single-cell sequencing, such as whole genome and RNA sequencing, provide a good opportunity to study the complex intra-tumoral heterogeneity at the single cell level [63]. Lawson et al. developed a workflow to enrich breast cancer stem cells from MCF7 and T47D cell lines. Single cell gene analysis was later applied to study the biomarkers related to differentiation, pluripotency, EMT, and proliferation. By comparing the cellular subpopulations and different stages, these researchers further demonstrated that there were progression and transition between differentiation stages in breast cancer at single cell level. For example, T47D cells transitioned from a quiescent state to a more differentiated phenotype while MCF7 cells gradually differentiated via a progenitor-like state [64]. The single cell analysis enabled us to better understand the sequential order of events during cancer stem cell differentiation at the transcriptional level.

Chronic myeloid leukemia (CML) is considered a less genetically complex disease. Tyrosine kinase inhibitors (TKI) are an effective drug for targeting the BCR-ABL gene and dramatically improving its outcome [65]. However, most patients relapse following treatment because

CML stem cells are resistant to TKI treatment, even those these cells are rare. Further study on CML stem cells will provide a more theoretical basis for clinical treatment. So far, the presence of the BCR-ABL fusion gene remains the only unequivocal marker of CML stem cells. Single cell analysis provides a feasible approach to analyze aberrant gene expression in normal, BCR-ABL positive and negative stem cells. Novel candidate gene expression, such as RXFP1, which is the receptor for the hormone relaxin, the small GTPase RAB31, SRSF2, and LGALS1, showed differences in expression among the above groups. Moreover, studies on BCR-ABL negative stem cells showed that TNF-alpha and TGF-alpha, which were correlated with poor treatment response in CML, were also dysregulated in CML [66–68]. These findings indicated that cell-extrinsic factors disrupt normal stem/progenitor cells in CML [69–71]. The authors also suggested the exploration of gene expression markers in BCR-ABL negative stem cells for developing clinical predictive biomarkers. This study suggested that single cell studies could unravel heterogeneity in clonal cancer stem cells, as well as in coexisting and frequently suppressed normal stem cells, thereby providing insights into the molecular mechanism of therapy resistance.

Single-cell sequencing studies of bladder cancer stem cells and non-stem cells, bladder epithelial stem cells and non-stem cells showed that bladder cancer stem cells were relatively homogeneity. The data also suggests that bladder cancer stem cells may originate from mutated bladder epithelial stem cells or alternated bladder cancer non-stem cells. There were 21 key altered genes that were identified in bladder cancer stem cells, including six genes that were not previously described in bladder cancer (ETS1, GPRC5A, MKL1, PAWR, PITX2 and RGS9BP). Meanwhile, bladder cancer stem cells showed a higher frequency of nonsynonymous mutation compared to bladder cancer non-stem cells, which underscored the critical function of specific mutations on bladder cancer stem cells acquiring stemness [72].

11.3 Application of Single-Cell Sequencing on Cancer Diagnosis and Biomarker

Single-cell sequencing has spread various research fields, such as cancer, immunology, microbiology, and embryogenesis. The transformation of single-cell sequencing from the bench to bedside will dramatically benefit patients. Preimplantation diagnosis using single-cell sequencing at the assistant reproductive area is a very successful example. Precise diagnosis of embryos before implantation effectively decreases the birth defects. The same technique is also applied in non-invasive pre-natal diagnosis. Studies on cancer genomic, transcriptomic, and epigenomic changes have dramatically improved the understanding of tumorigenesis, metastases, and relapse. Traditional investigations were based on bulk cells or tissues, which masked the heterogeneity of the disease. Current clinical diagnosis and therapy is also based on bulk samples. However, precise diagnosis and personalized therapeutic regimens that aim to remit or even cure cancer should focus on this heterogeneity because the disease outcome is mostly decided by only certain subpopulations of cancel cells. In clinical treatment of cancer, the lack of adequate biomarkers for diagnosis and staging confound clinical decision-making and delay potentially effective therapies. Development of single-cell sequencing of the cancer cells will improve the discovery of effective biomarkers to re-staging the disease or evaluating the cancer's prognosis. For example, single cell DNA cytophotometry has been used to analyze malignant melanoma at the clinical stage, and the mean nuclear area has been shown to be a prognostic factor [73]. Single-cell sequencing of prostate cancer cells suggested that the loss of PTEN is associated with poor prognosis [74]. In early gastric cancer, clustered-cell micrometastases of the lymph nodes have been used to predict poor survival compared with single-cell micrometastases [75].

In recent years, understanding of cancer at the molecular level has indicated that SNV, CNV,

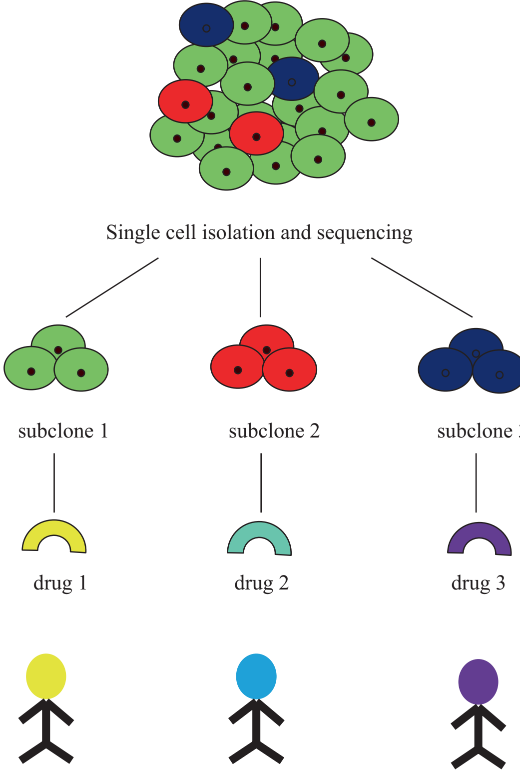


Fig. 11.2 Tumor cells are heterogeneous. Based on the single cell sequencing, tumor cells could be subdivided into different clones. Further study of tumor cell heterogeneity would help to characteristic each subclone and help to develop specific drugs improving the current knowledge on cancer, and ultimately improve tumor diagnosis and treatment

aneuploidy, and genomic rearrangements promote factors of cancer. However, current studies are mostly based on bulk cells, which mask the detection of cancer cell heterogeneity, cancer stem cells, and CTCs. Currently, new methods, such as fine-needle aspiration (FNA), FACS, and CellSearch, have been developed to obtain single cancer cells, cancer stem cells, or CTCs. Sequencing and identification of these single cells will provide precise information for diagnosis and clinical decision making. Meanwhile, single-cell sequencing could be used for early screening and therapy monitoring by extracting single cells from liquid biopsy. For example, early detection of CTCs and identification of possible key mutations may suggest the high risk of metastasis, while detection and sequencing of cancer stem cells may suggest the possibility of drug resistance and cancer recurrence.

Analysis of genomics, proteomic and transcriptomic changes and clinical outcomes will also provide novel specific candidates molecular biomarkers to predict disease prognosis. Also, patients with a low risk of relapse or metastasis could avoid overtreatment following this analysis.

In addition, in many clinical cases, it is not possible to obtain bulk cells from, for instance, urine, prostate fluid, vaginal fluid, or small biopsies. Using single-cell sequencing, these important cancer cells present in fluids, as well as certain solid tumor cells, could be identified and analyzed [76, 77]. See Fig. 11.2 for details.

11.4 Application of Single-Cell Sequencing on Cancer Treatment and Drug Development

A prime objective to understanding and dissecting the heterogeneous biology of tumors is to terminate the disease using effective approaches. Clinical failure of cancer treatment often ends with drug resistance. Heterogeneous responses triggered by chemotherapy pressure at the cellular

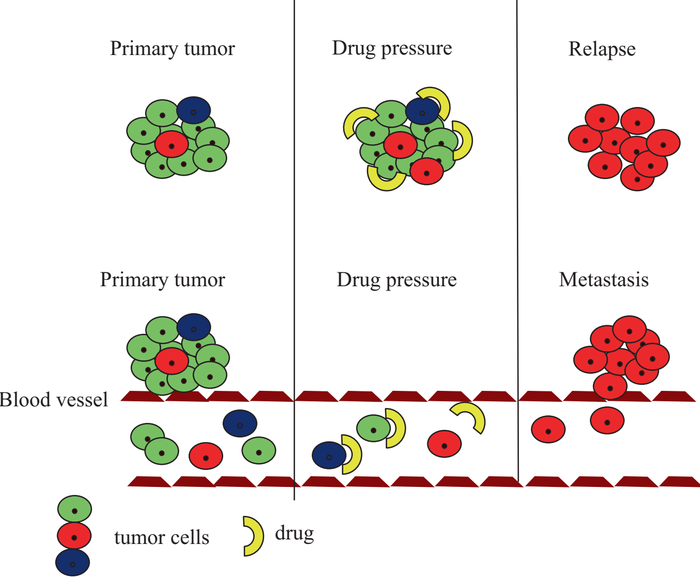


Fig. 11.3 The development of drug resistant during chemotherapy for tumor

level are one of the main reasons for drug resistance. See Fig. 11.3 for details. The deep understanding regarding cellular heterogeneity will provide fundamental insights into the development of drug resistance and clues for solution. Single-cell analysis at the DNA, RNA, and protein levels offers great promise to understanding cellular heterogeneity and tools for personalized treatment. Drug resistance can be classified as intrinsic and acquired resistance. Intrinsic resistance is present prior to chemotherapy exposure, and the tumor fails to respond to initial treatment. Acquired resistance occurs only during or after the course of treatment [5, 78]. Intrinsic resistance is not related to chemotherapy, as the tumor is not responsive to initial treatment. Acquired resistant often occurs after the exposure to drugs [5]. One reason for drug resistance is the individual differences in absorption and metabolism [5, 79]. Another reason is due to heritable genetic and epigenetic alterations and phenotypic plasticity in the cancer cell. Acquired drug resistance usually occurs in a small population, such as cancer stem cells. Deep single-cell sequencing of these subpopulations is not only able to provide a clue to understanding the mechanism of drug resistance but also offers a chance to benefit the patient through personalized chemotherapy.

It has been agreed that every tumor is unique from the others. Therefore, precision medicine as a new emerging therapeutic plan is considered to have the potential to improve cancer treatment. The primary principle of precision medicine in clinical oncology is to treat patients on the basis of their individual genetic mutations [80]. Single-cell sequencing may help to realize the dream of precision medicine in cancer pharmacology. Currently, single-cell RNA sequencing has been used in optimizing the treatment strategy in metastatic renal cell carcinoma [81]. Single-cell targeted transcriptome analysis has been performed to predict drug sensitivity in multiple myeloma [82].

11.4.1 Single-Cell Sequencing in the Future

Single-cell sequencing is a rapidly evolving approach to characterize an individual cell at the molecular level. With the help of computational methods, single-cell technologies have developed rapidly and have created a new avenue for investigating cancer heterogeneity at the cellular level systematically.

Single-cell sequencing directly promotes the development of next generation genomic medicine.

This method's high resolution and sensitivity can reveal the molecular basis of complex systems, such as aneuploidy, single-gene disorders, and chromosomal translocation. Well-developed single-cell sequencing will help clinicians to detect early germline- or somatic-mutation-based diseases and make a precise diagnosis, which has a decisive influence on the therapeutic regimen and outcomes. Sequencing of single cells will fundamentally improve current oncology. Specifically, the technique could enable physicians to detect rare tumor cells, monitor circulating tumor cells, measure intra-tumoral heterogeneity, and guide chemotherapy. Single-cell techniques will provide an unprecedented methodology for cancer research and help us to understand the process of tumor metastasis.

Currently, the clinical application of single-cell sequencing for diagnosis and guiding the treatment of cancer is a rising field, but it has shown its potential power in the future clinical settings. For example, this technique's application in genomic sequencing could provide a holographic view of individual disease, which improves the development of individualized treatment. It equips clinicians to identify rare clones and helps to explain how clonal diversity results in metastasis, and chemo- or radio-resistance in cancers. Moreover, its application in the CTC field has provided clues to understanding the evolution of primary tumor to metastatic tumors in a genetic manner [83].

Using the microscope, traditional pathologists could screen thousands of cells and find abnormity at the cytological level, but they cannot identify the genomic lesion, such as the copy number variation. Advances in whole genome amplification make it possible to acquire enough genetic material and next-generation sequencing enables us to detect genomic lesions at a specific locus in individual cancer cells. Genomic profiling of single cells can eliminate noise and provide an absolute picture of the genome; this allows researchers to identify thousands of potential carcinogenesis-related genes, thereby providing the evidence to oncologist for effective treatment decisions. In future medicine, single-cell sequencing may be applied in early cancer detection, CTC monitoring during the treatment of metastatic patients, and identification of the genomic diversity of solid tumors. It was reported that epithelial cells number in the blood is useful for disease treatment [84, 85]. Compared to simple counting, copy number profiling of CTCs may provide a more powerful tool to identify genomic amplifications of oncogenes and deletions of tumor suppressors. Such methods will also allow clinicians to monitor CTCs during adjuvant- or chemo-therapy and to determine whether the tumor is likely to relapse. The new achievements derived from single cell study will help to develop precision medicine and drug development. Although it is promising to apply the single-cell sequencing to the clinic, there are numerous challenges ahead for translating this application to a clinical setting.

At the present, single-cell sequencing technology has strong prospects, but this method still needs to overcome the numerous limitations that currently exist, such as relatively low efficiency and high cost. Significant challenges are still posed by contamination [86], amplification bias (DNA and RNA) [87, 88], efficiency of sequencing [89], and algorithms sufficiently advanced to link the sequencing data [90]. In addition to the sequencing, single-cell data also faces a number of intrinsic challenges, including systematic noise and features of biological systems, as well as the assessment of sparse and complex of the data. A positive development is that in the past few years, this technology has been improved dramatically and will be further developed. With rapid technological progress, we believe that in the near future, improved single-cell sequencing technologies will overcome these challenges such that more precise sequencing data will be obtained by pathologists and clinicians for the development of therapeutic regimens.

In short, single-cell sequencing is a rapidly expanding field that surprises us with daily updates that refresh our understanding of fundamental cancer biology. This method exhibits strong potential for improving our understanding of the nature and complexity of cancer. Although facing great challenges, this technique is still worth promoting and developing because of its

great practical value. When future innovations are able to overcome the current challenges, oncologists will be able to obtain sufficient genomic information for diagnosis from limited clinical samples and make precise therapeutic regimens. Single cell sequencing will improve not only the development of diagnostics but also cancer therapy regimens. The advances will ultimately elevate the survival of cancer patients.

There is no conflict of interest to declare.

Acknowledgments We are grateful to the president of the hospital, Jianqin Gu, Professor Xiangdong Wang, Professor Cuilian Zhang, and Professor Li Meng for providing valuable guidance in every stage of the writing of this chapter and to our colleagues for providing support and advice.

References

1. Wang J, Song Y (2017) Single cell sequencing: a distinct new field. Clin Transl Med 6(1):10
2. Coskun AF, Eser U, Islam S (2016) Cellular identity at the single-cell level. Mol BioSyst 12(10):2965
3. Novick A, Weiner M (1957) Enzyme induction as an all-or-none phenomenon. Proc Natl Acad Sci U S A 43(7):553
4. Navin N, Hicks J (2011) Future medical applications of single-cell sequencing in cancer. Genome Med 3(5):1–12
5. Schmidt F, Efferth T (2016) Tumor heterogeneity, single-cell sequencing, and drug resistance. Pharmaceuticals 9(2):33
6. Ellsworth DL, Blackburn HL, Shriver CD et al (2017) Single-cell sequencing and tumorigenesis: improved understanding of tumor evolution and metastasis. Clin Transl Med 6(1):15
7. Bedard PL, Hansen AR, Ratain MJ et al (2013) Tumor heterogeneity in the clinic. Nature 501(7467):355–364
8. Nicholas CR, Gaur M, Wang S et al (2007) A method for single-cell sorting and expansion of genetically modified human embryonic stem cells. Stem Cells Dev 16(1):109–117
9. Thorsen T, Maerkl SJ, Quake SR (2002) Microfluidic large-scale integration. Science 298(5593):580–584
10. Guo G, Luc S, Marco E et al (2013) Mapping cellular hierarchy by single-cell analysis of the cell surface repertoire. Cell Stem Cell 13(4):492
11. Guo G, Huss M, Guo QT et al (2010) Resolution of cell fate decisions revealed by single-cell gene expression analysis from zygote to blastocyst. Dev Cell 18(4):675
12. Buganim Y, Faddah DA, Cheng AW et al (2012) Single-cell expression analyses during cellular reprogramming reveal an early stochastic and a late hierarchic phase. Cell 150(6):1209
13. Dalerba P, Kalisky T, Sahoo D et al (2011) Single-cell dissection of transcriptional heterogeneity in human colon tumors. Nat Biotechnol 29(12):1120
14. Moignard V, Macaulay IC, Swiers G (2013) Characterization of transcriptional networks in blood stem and progenitor cells using high-throughput single-cell gene expression analysis. Nat Cell Biol 15(4):363–372
15. Bendall SC, Simonds EF, Qiu P et al (2011) Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. Science 332(6030):687–696
16. Behbehani GK, Bendall SC, Clutter MR et al (2012) Single-cell mass cytometry adapted to measurements of the cell cycle. Cytometry A 81A(7):552
17. Tang F, Barbacioru C, Wang Y et al (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6(5):377
18. Nagano T, Lubling Y, Stevens TJ et al (2013) Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. Nature 502(7469):59–64
19. Elad A, Davis KL, Tadmor MD et al (2013) viSNE enables visualization of high dimensional single-cell data and reveals phenotypic heterogeneity of leukemia. Nat Biotechnol 31(6):545
20. Marco E, Karp RL, Guo G et al (2014) Bifurcation analysis of single-cell gene expression data reveals epigenetic landscape. Proc Natl Acad Sci U S A 111(52):E5643
21. Trapnell C, Cacchiarelli D, Grimsby J et al (2014) The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. Nat Biotechnol 32(4):381–386
22. Moignard V, Woodhouse S, Haghverdi L et al (2015) Decoding the regulatory network of early blood development from single-cell gene expression measurements. Nat Biotechnol 33(3):269
23. Tsoucas D, Yuan GC (2017) Recent progress in single-cell cancer genomics. Curr Opin Genet Dev 42:22–32
24. Clevers H (2011) The cancer stem cell: premises, promises and challenges. Nat Med 17(3):313–319
25. Saadatpour A, Lai S, Guo G et al (2015) Single-cell analysis in Cancer genomics. Trends Genet Tig 31(10):576
26. Navin N, Kendall J, Troge J et al (2011) Tumour evolution inferred by single-cell sequencing. Nature 472(7341):90
27. Wang Y, Waters J, Leung ML et al (2014) Clonal evolution in breast cancer revealed by single nucleus genome sequencing. Nature 512(7513):155
28. Gao R, Davis A, Mcdonald TO et al (2016) Punctuated copy number evolution and clonal stasis in triple-negative breast Cancer. Nat Genet 48(10):1119

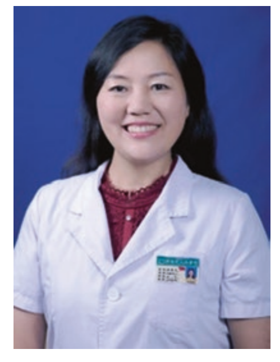
29. Baslan T, Kendall J, Ward B et al (2015) Optimizing sparse sequencing of single cells for highly multiplex copy number profiling. Genome Res:25(5)
30. Nguyen A, Yoshida M, Goodarzi H et al (2016) Highly variable cancer subpopulations that exhibit enhanced transcriptome variability and metastatic fitness. Nat Commun 7:11246
31. Lee MC, Lopez-Diaz FJ, Khan SY et al (2014) Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. Proc Natl Acad Sci U S A 111(44):E4726
32. Winterhoff BJ, Maile M, Mitra AK et al (2017) Single cell sequencing reveals heterogeneity within ovarian cancer epithelium and cancer associated stromal cells. Gynecol Oncol 144(3):598–606
33. Malecki M, Szybalski W (2012) Isolation of single, intact chromosomes from single, selected ovarian cancer cells for in situ hybridization and sequencing. Gene 493(1):132
34. Min JW, Kim WJ, Han JA et al (2015) Identification of distinct tumor subpopulations in lung adenocarcinoma via single-cell RNA-seq. PLoS One 10(8):e0135817
35. Kim KT, Lee HW, Lee HO et al (2015) Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. Genome Biol 16(1):1–15
36. Hughes AE, Magrini V, Demeter R et al (2014) Clonal architecture of secondary acute myeloid leukemia defined by single-cell sequencing. PLoS Genet 10(7):e1004462
37. Gawad C, Koh W, Quake SR (2014) Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. Proc Natl Acad Sci U S A 111(50):17947–17952
38. Bjorn B, Aaron T, Belderbos ME et al (2016) Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. Genome Biol 17(1):115
39. Xu X, Hou Y, Yin X et al (2012) Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. Cell 148(5):886–895
40. Francis JM, Zhang CZ, Maire CL et al (2014) EGFR variant heterogeneity in glioblastoma resolved through single-nucleus sequencing. Cancer Discov 4(8):956
41. Patel AP, Tirosh I, Trombetta JJ et al (2014) Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. Science 344(6190):1396–1401
42. Chen XX, Bai F (2015) Single-cell analyses of circulating tumor cells. Cancer Biol Med 12(3):184–192
43. Yu M, Stott S, Toner M et al (2011) Circulating tumor cells: approaches to isolation and characterization. J Cell Biol 192(3):373–382
44. Tan SJ, Yeo T, Sukhatme SA et al (2017) Personalized treatment through detection and monitoring of genetic aberrations in single circulating tumor cells. Adv Exp Med Biol 994:255
45. Ni X, Zhuo M, Su Z et al (2013) Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. Proc Natl Acad Sci U S A 110(52):21083–21088
46. Jiang R, Lu YT, Hao H et al (2015) A comparison of isolated circulating tumor cells and tissue biopsies using whole-genome sequencing in prostate cancer. Oncotarget 6(42):44781–44793
47. Lohr JG, Adalsteinsson VA, Cibulskis K et al (2014) Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. Nat Biotechnol 32(5):479–484
48. Solakoglu O, Maierhofer C, Lahr G et al (2002) Heterogeneous proliferative potential of occult metastatic cells in bone marrow of patients with solid epithelial tumors. Proc Natl Acad Sci U S A 99(4):2246–2251
49. Kraus J, Pantel K, Pinkel D et al (2003) High-resolution genomic profiling of occult micrometastatic tumor cells. Genes Chromosomes Cancer 36(2):159–166
50. Gangnus R, Langer S, Breit E et al (2004) Genomic profiling of viable and proliferative micrometastatic cells from early-stage breast cancer patients. Clin Cancer Res 10(10):3457
51. Heitzer E, Auer M, Gasch C et al (2013) Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. Cancer Res 73(10):2965–2975
52. Luca FD, Rotunno G, Salvianti F et al (2016) Mutational analysis of single circulating tumor cells by next generation sequencing in metastatic breast cancer. Oncotarget 7(18):26107–26119
53. Pestrin M, Salvianti F, Galardi F et al (2015) Heterogeneity of PIK3CA mutational status at the single cell level in circulating tumor cells from metastatic breast cancer patients. Mol Oncol 9(4):749–757
54. Fernandez SV, Bingham C, Fittipaldi P et al (2014) TP53 mutations detected in circulating tumor cells present in the blood of metastatic triple negative breast cancer patients. Breast Cancer Res 16(5):445
55. Deng G, Krishnakumar S, Powell AA et al (2014) Single cell mutational analysis of PIK3CA in circulating tumor cells and metastases in breast cancer reveals heterogeneity, discordance, and mutation persistence in cultured disseminated tumor cells from bone marrow. BMC Cancer 14(1):456
56. Steinert G, Schölch S, Niemietz T et al (2014) Immune escape and survival mechanisms in circulating tumor cells of colorectal cancer. Cancer Res 74(6):1694
57. Gerlinger M, Rowan AJ, Horswell S et al (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366(10):883–892

58. Hiley C, Bruin ECD, Mcgranahan N et al (2014) Deciphering intratumor heterogeneity and temporal acquisition of driver events to refine precision medicine. Genome Biol 15(8):453
59. Su KY, Chen HY, Li KC et al (2012) Pretreatment epidermal growth factor receptor (EGFR) T790M mutation predicts shorter EGFR tyrosine kinase inhibitor response duration in patients with non-small-cell lung cancer. J Clin Oncol 30(4):433
60. Turke AB, Zejnullahu K, Wu YL et al (2010) Preexistence and clonal selection of MET amplification in EGFR mutant NSCLC. Cancer Cell 17(1):77
61. Patel LR, Nykter M, Chen K et al (2013) Cancer genome sequencing: understanding malignancy as a disease of the genome, its conformation, and its evolution. Cancer Lett 340(2):152–160
62. Caicedo AS, Aponte PM (2017) Stemness in cancer: stem cells, Cancer stem cells, and their microenvironment. Stem Cells Int 2017:5619472
63. Wen L, Tang F (2016) Single-cell sequencing in stem cell biology. Genome Biol 17:71
64. Lawson DA, Bhakta NR, Kessenbrock K et al (2015) Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature 526(7571):131
65. Longo DL (2017) Imatinib changed everything. N Engl J Med 376(10):982–983
66. Liu JX, Zhang D, Xie X et al (2013) Eaf1 and Eaf2 negatively regulate canonical Wnt/β-catenin signaling. Development 140(5):1067–1078
67. Goardon N, Marchi E, Atzberger A et al (2011) Coexistence of LMPP-like and GMP-like leukemia stem cells in acute myeloid leukemia. Cancer Cell 19(1):138
68. Zhao K, Yin LL, Zhao DM et al (2014) IL1RAP as a surface marker for leukemia stem cells is related to clinical phase of chronic myeloid leukemia patients. Int J Clin Exp Med 7(12):4787–4798
69. Schepers K, Campbell TB, Passegue E (2015) Normal and leukemic stem cell niches: insights and therapeutic opportunities. Cell Stem Cell 16(3):254–267
70. Welner RS, Amabile G, Bararia D et al (2015) Treatment of chronic myelogenous leukemia by blocking cytokine alterations found in normal stem and progenitor cells. Cancer Cell 27(5):671–681
71. Reynaud D, Pietras E, Barryholson K et al (2011) IL-6 controls leukemic multipotent progenitor cell fate and contributes to chronic myelogenous leukemia development. Cancer Cell 20(5):661
72. Yang Z, Li C, Fan Z et al (2017) Single-cell sequencing reveals variants in ARID1A, GPRC5A and MLL2 driving self-renewal of human bladder Cancer stem cells. Eur Urol 71(1):8
73. Lindholm C, Hofer PA, Jonsson H (1990) Single cell DNA cytophotometry in clinical stage I malignant melanoma. Relationship to prognosis. Acta oncol 29(2):147–150
74. Heselmeyerhaddad KM, Berroa Garcia LY, Bradley A et al (2014) Single-cell genetic analysis reveals insights into clonal development of prostate cancers and indicates loss of PTEN as a marker of poor prognosis. Am J Pathol 184(10):2671–2686
75. Cao L, Hu X, Zhang Y et al (2011) Adverse prognosis of clustered-cell versus single-cell micrometastases in pN0 early gastric cancer. J Surg Oncol 103(1):53–56
76. Ren SC, Qu M, Sun YH (2013) Investigating intratumour heterogeneity by single-cell sequencing. Asian J Androl 15(6):729
77. Adalsteinsson VA, Love JC (2014) Towards engineered processes for sequencing-based analysis of single circulating tumor cells. Curr Opin Chem Eng 4:97–104
78. Yardley DA (2013) Drug resistance and the role of combination chemotherapy in improving patient outcomes. Int J Breast Cancer 2013:137414
79. Jain RK (2001) Delivery of molecular and cellular medicine to solid tumors. Microcirculation 46(1–3):49–67
80. Wang X (2016) Gene mutation-based and specific therapies in precision medicine. J Cell Mol Med 20(4):577–580
81. Kim KT, Lee HW, Lee HO et al (2016) Application of single-cell RNA sequencing in optimizing a combinatorial therapeutic strategy in metastatic renal cell carcinoma. Genome Biol 17(1):80
82. Mitra AK, Mukherjee UK, Harding T et al (2016) Single-cell analysis of targeted transcriptome (SCATTome) predicts drug sensitivity of single cells within human myeloma tumors. Leukemia 30(5):1094
83. Navin NE (2015) Delineating cancer evolution with single-cell sequencing. Sci Transl Med 7(296):296fs29
84. Bidard FC, Mathiot C, Delaloge S et al (2010) Single circulating tumor cell detection and overall survival in nonmetastatic breast cancer. Ann Oncol 21(4):729–733
85. Rao C, Bui T, Connelly M et al (2011) Circulating melanoma cells and survival in metastatic melanoma. Int J Oncol 38(3):755
86. Blainey PC (2013) The future is now: single-cell genomics of bacteria and archaea. FEMS Microbiol Rev 37(3):407
87. Ning L, Liu G, Li G et al (2014) Current challenges in the bioinformatics of single cell genomics. Front Oncol 4:7
88. Hodzic E (2016) Single-cell analysis: advances and future perspectives. Bosn J Basic Med Sci 16(4):313–314
89. Schadt EE, Turner S, Kasarskis A (2010) A window into third-generation sequencing. Hum Mol Genet 19(2):227–240
90. Kim KI, Simon R (2014) Using single cell sequencing data to model the evolutionary history of a tumor. Bmc Bioinf 15(1):1–13

Lan Yu M.D., Ph.D. Attending physician, graduated from Huazhong University of Science and Technology. In 2011–2016, she was funded by China Scholarship Council to study as a Ph.D candidate in Turku University, Finland. The study was about the role of FGF and FGFR in prostate cancer. Her PhD study was also funded by Ida Montinin Säätiö, Lounaissuomalaiset Syöpäjärjestöt, and K.Abin Johanssons stiftelse.

Hua Zhao M.D., Ph.D. Medical doctor, attending physicians. Ph.D. conferred on her by Huazhong University of Science and Technology. She takes part in scientific research projects and has more than 12 publications. She has been in reproductive medicine for 8 years. She was good at the diagnosis and treatment of various infertility and reproductive endocrine diseases and the clinical application of assisted reproductive technologies such as in vitro fertilization-embryo transfer (IVF-ET) and artificial insemination (IUI).

Li Meng, Ph.D. President of BioBest IVF Consultant, Inc., USA. High-complexity Clinical Laboratory Director (HCLD). Adjunct Professor of Reproductive Medical Center, Shang Dong TCM University, P.R. China. Senior scientific consultant of Reproductive Medicine Center, Changhua Christian Hospital, Taiwan. Senior scientific consultant in Mariposa Biotechnology, Inc; San Clemente. Adjunct Professor of Harbin Medical University, Harbin, HeilongJiang, P.R.China. He is in editorial board of Reproductive and Developmental Medicine.



Cuilian Zhang M.D., Ph.D. Professor of reproductive sciences, Doctoral advisor, Chief Physician, Vice Dean of Henan Provincial People's Hospital, Dean of Henan Provincial Reproductive Hospital. She enjoys the special allowance of the State Council. Her current positions also include Scientific and Technological innovation and outstanding talents in Henan province, Chairman of Reproductive Medical Doctor Branch of Henan Medical Doctor Association, Vice-chairman of Henan Society of Reproductive Medicine, Standing Committee member of Chinese Medical Doctor Association, Deputy director of the first precision assisted reproductive medicine research unit of Chinese Medical Doctor Association, the first Chinese youth committee of Chinese Medical Association's Reproductive Medicine Branch, Member of the clinical division of Chinese Medical Association's Reproductive Medicine Branch, Member of the first clinical reproductive genetics committee of Chinese Physicians Association, Reproductive Medicine Group of Fertility & Health Care Committee of Maternal and Child Health Association of China, Reproductive Medicine Branch of China International Exchange and Promotive Association of Medical and Health Care, Standing committee member of Maternal and child health institute in reproductive endocrinology professional committee, Editor of Chinese Journal of Reproduction and Contraception, Editor of Journal of Practical Diagnosis and Therapy. She has more than 100 publications and 9 projects. She was awarded the First prize of provincial bureau of health scientific research achievements, the second prize of Henan scientific and technological achievements, and the third prize of Henan scientific and technological achievements.

12 Emergence of Bias During the Synthesis and Amplification of cDNA for scRNA-seq

Qiankun Luo and Hui Zhang

Abstract

The advent of single-cell omics technology has promoted our understanding of the genomic, epigenomic, and transcriptomic heterogeneity in individual cells. Compared to traditional sequencing studies using bulk cells, single-cell transcriptome technology is naturally more dynamic for in depth analysis of genomic variation resulting from cell division and is useful in unraveling the regulatory mechanisms of gene networks in many diseases. However, there are still some limitations of current single-cell RNA sequencing (scRNA-seq) protocols. Biases that arise during the RNA reverse transcription and cDNA pre-amplification steps are the most common problems and play pivotal roles in limiting the quantitative accuracy of scRNA-seq. In this review, we will describe how these biases emerge and impact scRNA-seq protocols. Moreover, we will introduce several current and convenient modified scRNA-seq methods that allow for bias to be decreased and estimated.

Keywords

Amplification · Single-cell · Transcriptomic · Technical noise

12.1 Introduction

The molecular mechanisms involved in the normal development of organisms and in tumor evolution have traditionally been studied using large populations of cells [1]. While transcriptome sequencing of RNAs extracted from bulk cells is informative for some questions, cell-to-cell signal variations are concealed. Some cell types, such as stem cells or circulating tumor cells, cannot be studied in large populations, and individual cells may have independent or special roles. Studies of those cell types significantly benefit from the use of single cell RNA sequencing (scRNA-seq) technology. The development of next-generation gene sequencing and scRNA-seq technology provide to analyse transcriptional variations and intercellular heterogeneity at high resolution and to explore the interaction between intrinsic cellular activation and extrinsic stimuli, such as internal cell environments, drugs, or infections. scRNA-seq methods with different characteristics recently developed renew our understanding of transcriptomic diversity and

Authors Qiankun Luo and Hui Zhang have equally contributed to this chapter as the first author.

Q. Luo (✉)
Department of General Surgery, Second Affiliated Hospital, Zhengzhou University, Zhengzhou, China

H. Zhang
Department of Science Research and Discipline Construction, Henan Provincial People's Hospital, Jinshui District, Zhengzhou, China

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_12

heterogeneity. scRNA-seq has been applied to study stem cell differentiation, early embryos, sensitivity to drugs, dynamic phenotypes, cell sub-population identification, or heterogeneity [2–12].

High-throughput scRNA-seq technologies, microfluidic systems, and droplet-based sequencing have enabled the throughput of scRNA-seq to allow for tens of thousands of cells to be assayed at a time [13–15]. The general processing strategies include the single-cell isolation and lysis, reverse transcription of mRNA, cDNA pre-amplification, and sequencing library construction [16]. Despite of significant improvements in the throughput, sensitivity and scalability of scRNA-seq, there are still a large number of limitations in the quantitative accuracy of scRNA-seq technology. The initial step of scRNA-seq is to isolate cells from tissues using trypsin or collagenase, which may potentially affect cell viability and transcripts [16], subsequently impacting the quality and efficiency of scRNA-seq protocols. The targeted cells are isolated from suspended cells using flow-activated cell sorting platforms, and captured through surface markers identified by fluorescent-labeled antibodies. However, the need for antibodies and large basal cell populations limits its application [17, 18]. Another potential limitation of scRNA-seq protocols is that the cell lysis procedure may damage the natural context and cellular environment of individual cells, resulting in the loss and impair of the spatial information on transcripts and the complete detection of the biological information present in cells. Novel techniques developed in recent years, include the in situ sequencing and RNA-seq on single nuclei, maintain the spatial information, and avoid the harsh cell isolation procedures [19–21], although a number of the drawbacks still exist. The technical noise from reverse transcription and cDNA pre-amplification steps are currently the most important limitation for the quantitative accuracy of scRNA-seq [22]. *Kolodziejczyk et al.* have reviewed the estimates of technical variability by the application of spike-ins of an external RNA control consortium (ERCC) and unique molecular identifiers (UMIs). The computational method to process data has been improved to normalize and reduce the noise [16, 23]. Despite that modifications to improve the efficiency and decrease the technical noise have been made to the approaches, biases that occur during the synthesis and amplification of cDNA are still the key limitation of scRNA-seq protocols. In this review, we will discuss how these biases arise and the improvement in techniques that facilitate their reduction.

12.2 Factors Affecting Bias

12.2.1 Synthesis of First Strand cDNA

Single-cell transcriptome analysis platforms generally require enough materials for hybridization to perform gene expression microarray or sequencing. For example, RNA cannot be sequenced directly if a minute amount of RNA can be captured from a single cell (10 pg), so the amplification of transcripts seems necessary for scRNA-seq. The poly(A) tailing induces large numbers of transcripts to be lost, resulting in the reverse transcription of only approximately 10–20% of transcripts. These factors have a strong impact on the sensitivity of scRNA-seq and are the primary reasons for the emergence of bias [24]. Current scRNA-seq protocols face two challenges, i.e. the reverse transcription and cDNA amplification steps, during which the undesirable bias emerges.

The low amount of RNA molecules extracted from cells contains tRNA and rRNA, which accounts for a large proportion of cellular RNA and may have immediate impacts on the mRNA sequencing reads seriously [25]. The reverse transcription (RT) primers are used to exclude rRNA and tRNA. Because mRNAs have a poly(A) tail, current methods synthesize the first strand cDNA and capture polyadenylated RNA using a ploy(dT) primer. The poly(A) tailing selection approach can isolate mRNAs and most lncRNAs effectively and conveniently within

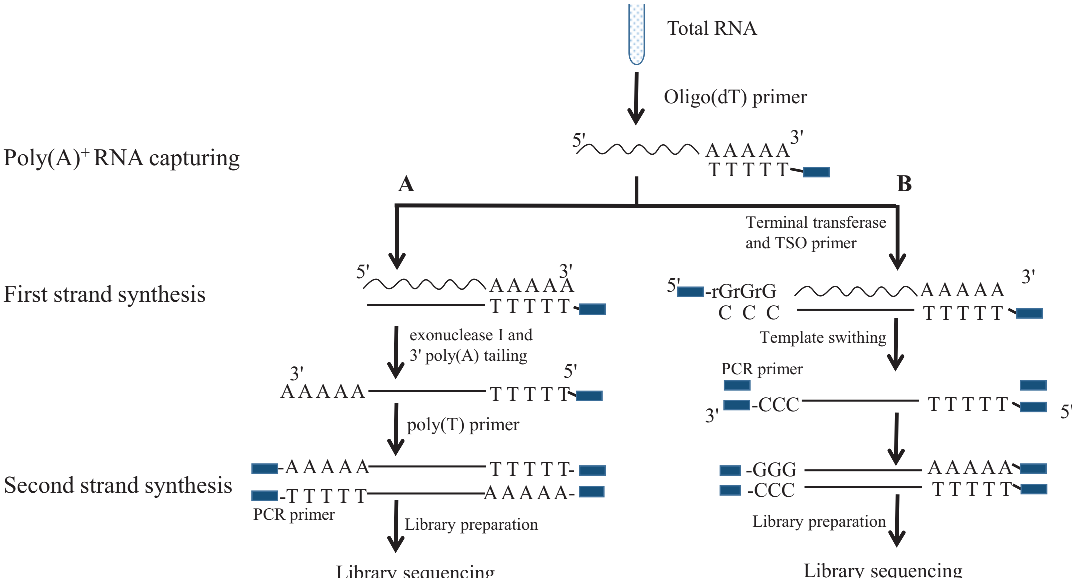


Fig. 12.1 Two methods for the synthesis of double-stranded cDNA both rely on 3′ end poly(A) tailing to capture polyadenylated RNA. An oligo(dT) primer is used for reverse transcription, causing a strong poly(A)-RNA coverage bias. The second cDNA strand synthesis involves either the homopolymer dependent method or the template-switching dependent method. (**a**) Exonucleaseis used to digest the RT primer. Poly(A) tails are added to 3′ ends of first strand cDNA molecules to synthesize the second strand cDNA. (**b**) During the reverse transcription, the terminal transferase is used to add non-template cytosine to the 3′ end of first strand cDNA as a primer for template-switching, then a poly(G) template is added as an adapter to perform the switching and synthesis of full-length cDNA (This figure was modified from [30])

single cells. This method is the most widely used in current scRNA-seq methods (Fig. 12.1). For instance, the Quartz-Seq and SC3-seq approaches both use oligo-dT24 as a reverse transcription (RT) primer to perform the first strand synthesis of cDNA [26, 27].

The current approaches heavily rely on polyadenylated mRNAs and only capture RNA with a poly(A) tail. This inevitably precludes RNA species lacking poly(A) tails, including microRNAs, non-polyadenylated lncRNAs, and circular RNAs, as the primary factor for the apparent emergence of bias. There is only have a small fraction of polyadenylated RNA in prokaryotic cells, such as bacteria, and be hardly extracted with a poly T primer [28].To rectify such problem, another approach using 5′-monophosphate dependent exonuclease was proposed to exclude tRNAs and rRNAs [29]. The technique could deplete some relevant mRNAs and small RNA species, and introduce some noise to the transcripts [30]. The novel SUPeR-seq protocol was proposed to substitute random primers with a fixed anchor sequence for an oligo(dT) primer and to achieve first strand cDNA synthesis [31]. This method can detect both polyadenylated and non-polyadenylated transcripts simultaneously in a single cell. After the digestion of the excess primers, a poly(A) tail is added to 3'end of the first strand cDNA, and then another anchor sequence is added to synthesize the second strand cDNA. However, SUPeR-seq cannot completely break down the rRNAs secondary structures present in the transcripts, with the rRNA contamination about 15% of the overall cDNA. The scRNA-seq method is able to detect whole target

transcripts without missing transcripts or contamination. Due to the low amount of RNA molecules present in single cells, polyadenylated RNA is even less abundant (approximately 0.1 pg), substantial technical noise during RNA reverse transcription can produce strong biases on the synthesis of cDNA, which is amplified in the amplification process.

12.2.2 Synthesis of Second Strand cDNA

After the synthesis of first strand cDNA, the synthesis of the second cDNA can be accomplished by two different methods (Fig. 12.1a), including the poly(A) tail and terminal deoxynucleotidyl transferase. The reverse transcription primer is digested by exonuclease I to reduce the byproducts from the previous step, and the terminal deoxynucleotidyl transferase is used to achieve coverage of the 3′ ends of the first strand cDNA with a poly(A) tag. The synthesis of second strand cDNA can be carried out using the Tang protocol or Quartz-Seq [26, 32]. The transcripts are also extracted from different cells to maintain the strand information by the integration of DNA barcodes, performed in STRT-seq by the application of a template-switching mechanism and upgraded by Smart-seq and Smart-seq2 [33]. The template-switching oligos, containing 2–5 non-template cytosines, are added to the 3′ end of the first strand cDNA. The adapter sequence poly(G) template is then added to 5′ end of the RNA transcript. For example, the Moloney Murine Leukemia Virus with the template-switching activity is used to switch the template site located at the 5′ end of the RNA molecule and synthesize the second strand cDNA. The template-switching oligo method ensures the efficient and accurate dissection of full-length cDNA by a single primer. Each full-length cDNA consists of a natural 5′ poly(T) end transcribed from the 3’end of the RNA and an artificial poly(C) sequence [33] (Fig. 12.1b).

It is difficult to remove the reverse transcription primer completely by exonuclease I which can degrade the unreacted primer after reverse transcription and cause undue deletion on sections from the 5′ end of transcripts. The additional artificial poly(A) tail on the 3′ end of the first strand cDNA, apart from the intrinsic 3′ end poly(A) tail of the input RNA, will impact the strand information of double-strand cDNA [30]. The improper length of the synthetic poly(A) tail could diminish the sequencing quality or the efficiency of the second strand cDNA synthesis [31]. The 5’end of the polyadenylated mRNA is tagged with a poly(G) sequence using Moloney Murine Leukemia Virus or terminal transferase via the template-switching activity. The first strand should cover the 5′ end of mRNA transcripts as a strong 5′ end coverage bias of the transcripts emerges, if the full-length cDNA is captured [26, 34]. Therefore, both the poly(A) tailing-dependent method and the template-switching mechanism in the second strand cDNA synthesis can produce the technical noise to impair the efficiency or quality of scRNA-seq.

12.2.3 cDNA Amplification

12.2.3.1 Polymerase Chain Reaction-Based Amplification

The double-strand cDNA is purified and pre-amplified to construct the cDNA library prior to sequencing or microarray analysis after the reverse transcription. The large amounts of transcripts are essential for current high-throughput technologies to perform whole transcript amplification, generally performed by polymerase chain reaction (PCR) or in vitro transcription (IVT). PCR is an exponential and non-linear method with a higher amplification efficiency than IVT. The PCR application for the synthesis of the second strand of cDNA depends upon the addition of a poly(A) tail to the 3′ end of first strand cDNA as a template-switching dependency

method (Fig. 12.1). The length of synthetic poly(A) sequence is controlled to ensure the efficiency and quality of PCR. Fan X. *et al.* added a mixture of dATP and ddATP with terminal deoxynucleotidyl transferase when adding the poly(A) tail to first strand cDNA and showed that it was helpful to guarantee the length of the poly(A) tail [31]. It is essential to initially process several rounds of PCR using a poly(T) primer with a 5′-amine-terminated anchor sequence to avoid the primer portion of cDNA molecules binding to adaptors. Sasagawa et al. suppressed PCR primer to enrich the short DNA fragments and decreased the byproduct, forming a 'pan-like' structure in those DNA fragments with complementary sequences [26]. Such method shortened the length of non-reacted primers by limiting the terminal transferase activity time to suppress the synthesis of byproducts. However, the use of terminal deoxynucleotidyl transferase for reverse transcription could result in the coverage bias of 5′ end of transcripts. The addition of a poly(A) tail at the 3′ end of the first strand of cDNA can impact the strand information of double-stranded cDNA [35].

12.2.3.2 IVT-Based Amplification

The IVT is a linear amplification method that requires at least 400 pg of total RNA molecules as the initial input material and a physically labor-intensive process that requires three round of amplification to strongly limit the application

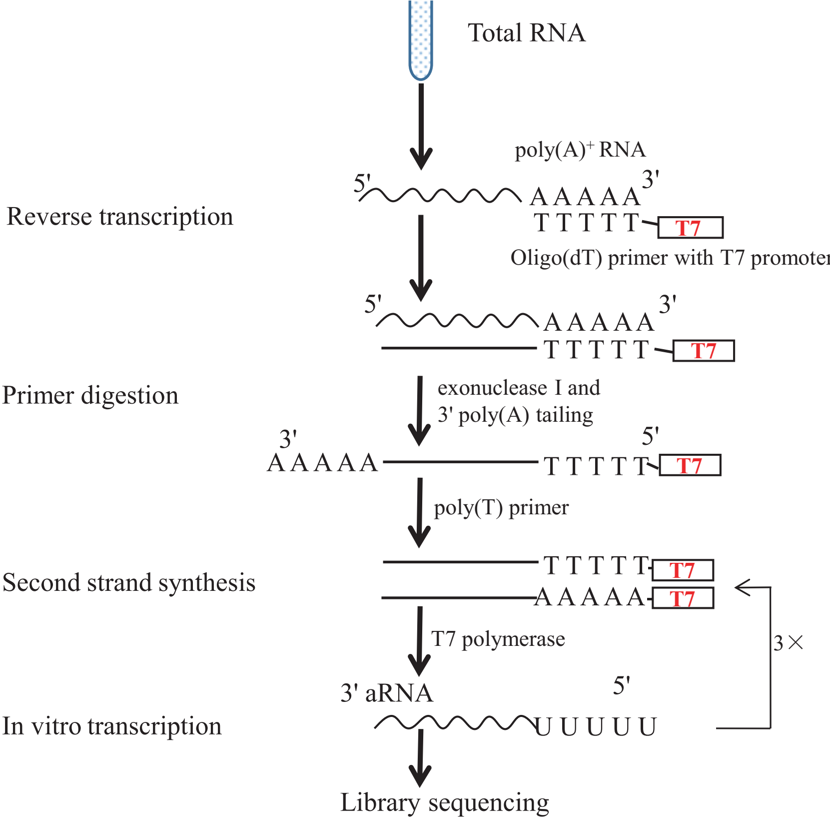


Fig. 12.2 T7 promoter was added to first-strand cDNA during reverse transcription. It facilitated the amplification of transcripts in vitro (This figure was modified from [30])

of IVT for single-cell RNA amplification [36] (Fig. 12.2). IVT is currently integrated with CEL-Seq method to reach obtain sufficient amounts of RNA and complete the linear amplification [37]. Such strategy relies on the T7 promoter located at the 5′ end of the cDNA and has a higher accuracy and specificity with a relatively good clearance of non-specific byproducts, as compared to PCR. The IVT protocol for single cell transcriptome analysis has a tedious program to require 3 rounds of IVT procedures or 2 rounds of IVT with an additional PCR procedure, essential to compile samples for a single IVT [36]. The additional round of reverse transcription leads to a significantly increased 3′ end bias [16]. scRNA-seq has a great potential to obtain biological information for special samples, including embryo cells, tumor cells, neuronal cells and other rare cells. The synthesis and amplification of cDNA are the main challenges of scRNA-seq, since the bias emerges during these steps and greatly limits the sequencing efficiency and accuracy. The reverse transcription and pre-amplification become even more important and essential to sequence the low amount of RNA molecules obtained from single cells.

12.3 Modifications of Protocols

12.3.1 ERCC Reduces Technical Noise

Many protocols have been greatly improved in recent years to improve the efficiency and coverage of transcripts for scRNA-seq technology and to decrease the technical noise generated from the synthesis and amplification of cDNA. For instance, SUPeR-seq utilizes a random primer with a fixed anchor sequence to synthesize cDNA and capture poly(A)- and poly(A) + RNA simultaneously in single cell [31]. Although the template-switching and IVT methods can theoretically cover the full-length and strand specificity, the strand specificity can be lost due to the transcript amplification. The directional features of the strands are retained [30]. The methods have been modified to analyze large numbers of cells in parallel. Current protocols have incorporated traditional methods that, depending on 3′ or 5′ ends of RNA, utilize a template-switching oligo, greatly increasing the number of input cells to thousands of single cells [9, 38]. The generation of technical noise is inevitable with the current scRNA-seq protocols, seriously impacting the reliability of scRNA-seq technologies. The estimation of technical noise is a vital precondition to assess biological variability using scRNA-seq technology. Some strategies were designed to compute and ameliorate technical noise, such as ERCC and UMIs.

The extrinsic spike-in RNA is a highly recommended modification of scRNA-seq protocols to measure the technical noise in scRNA protocols and can theoretically identify any poly(A)-tailed RNAs lost from the single cell library. The Grün's protocol used an artificial mixture of 92 synthetic species to quantify the technical noise [39]. This is the mostly commonly used ERCC with an accurately calibrated quantity of each single cell sample. The artificial spike-in RNAs are reverse transcribed using the same procedure as the natural cell transcripts. It is worth noting that the spike in RNAs is vulnerable to degradation in the reverse transcription procedure and that the amount of artificial RNA is multivariate by hundreds of folds [36]. The technical noise can be captured through the dissection of the externally spiked in RNA species [40]. The gene expression of the transcripts can be estimated by calculating the spike-in RNA reads, and the number of molecule reads examined with spike-in transcripts are similar to the result detected by Poisson distribution (Fig. 12.2c) [24, 39]. ERCC spike-ins control RNAs can also be used in conjunction with UMIs to model the technical variability. The capture efficiency can be calculated through the proportion of the read number mapped to the known number of spike-ins. The number of RNA

molecules can be calculated with the same ratio mapping to the number of UMIs [23]. It is helpful to quantify the amplification bias and imporve the measured capture efficiency (Fig. 12.2d). There are still some limitations in the widespread use of ERCC as it depends upon the poly(A) tail, such as the 3′ end coverage bias in scRNA protocols, to hardly estimate the expression of external RNA. The amount of artificial spike-in RNAs should be strictly consistent with the total number of natural RNAs in single cells to ensure that the variation of the spike-ins is not obstructed by the degradation during manipulations [16]. Rhonda and Kendziorski showed that the high level of spike-in species added to transcripts take up a large fraction of sequencing reads, which notably increases the physical labor [40]. Another challenge is to normalize the data generated from scRNA-seq data, different from those from bulk RNA-seq. SCnorm was recently suggested for accurate and efficient normalization of scRNA-seq data [41]. The quantile regression is used in SCnorm to estimate the dependence of transcript expression on sequencing depth for every gene. On basis of the dependence and quantile regression, genes and scale factors can be grouped and data can be normalized using the adjustment for sequencing depth with or without spike-ins.

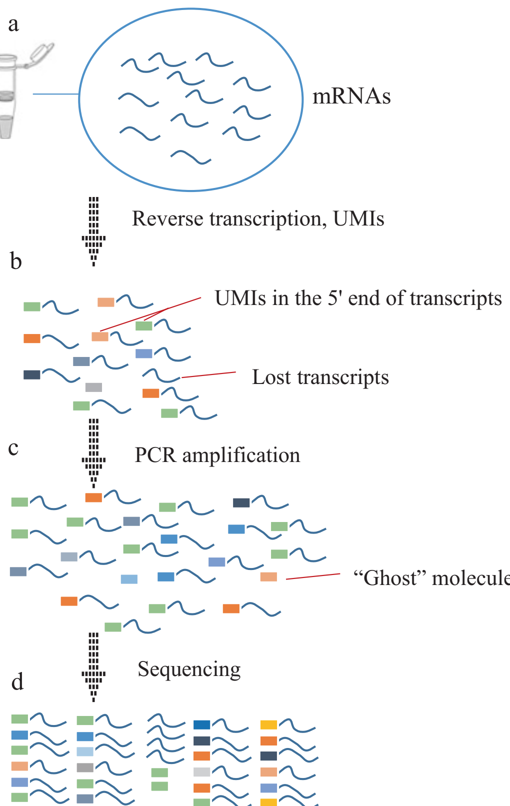


Fig. 12.3 UMIs application in sequencing. (**a**) UMIs (colored boxs) was added to 5′ end of cDNA, mRNA untagged will be lost in (**b**). Remove singleton barcodes to avoid errors caused by "ghost" molecules (This figure was modified from [24])

12.3.2 UMIs Simplify the Counting of Molecules

To remedy the PCR amplification bias during the cDNA synthesis, fragments were sequenced from the 3′ or 5′ end and the UMIs were developed and used to barcode single cells [39]. UMIs are DNA fragments that consist of four to ten random nucleotides. During the reverse transcription, each individual cDNA is tagged with a random sequence (Fig. 12.3). The cDNA library can be sequenced at a sufficient depth to ensure that the altered cDNA molecules can be observed [24]. The number of cDNA molecules can be directly calculated by counting the number of distinct UMIs. This approach has recently been widely used as an internal barcode to alleviate the technical noise. The number of molecules counted with UMIs fluctuated to a lesser degree than that with the conventional method when comparing the increasing number of input transcripts [24]. The UMIs can reduce the technical noise generated during reverse transcription and cDNA amplification. It is important that the cDNA molecules are done under saturated sequencing conditions to ensure that each cDNA incorporated with barcode

molecule can be measured. Another caution is that a "ghost" molecule may arise due to sequencing errors when using UMIs [23] (Fig. 12.3). The amplification bias can be corrected by the integrated UMIs even though there are still a number of risks in UMI methods. The methods that rely on UMIs for scRNA-seq can only be used for sequencing single tagged transcripts, such as sequencing only the 5′ or 3′ end of the transcript, making them unsuitable for studies of isoforms or allele-specific expression [34, 40].

12.4 Conclusions and Perspectives

With great improvement in throughput, quantitation, and sensitivity, scRNA-seq become more important tools to explore new cell differentiation routes, cell to cell variability, whole tissue-analysis, identification of cell subgroups, and tumor heterogeneity. It is important to remember a large number of limitations to restrict the quantitative accuracy of scRNA-seq protocols during the application, such as the destruction of spatial context for single-cell isolation, cell viability, reverse transcription, cDNA pre-amplification, mRNA capture efficiency, technical noise, or normalization of amplification bias. We believe that scRNA-seq can be a new approach to identify and validate single cell biomarkers for the definition of cell subtypes, heterogeneity, and progenitors [42].

It is even more exciting if scRNA-seq can be applied for the discovery of therapeutic targets and provide the potential to develop and improve therapeutic strategy for patients. Recent preclinical evidence demonstrated that scRNA-seq was used to define innate lymphoid cell precursor subsets, distinct cell development stages and pathways, and high expression of programmed death 1 and interleukin 25 receptor as the early checkpoint for immunotherapy [43]. It is also curious to compare and address the significance of biomarkers from the scRNA-seq and bulk cell RNA-seq. It will be more valuable if single cell biomarkers have the specificity of disease severity, duration, and responses to therapy, as expected from bulk cells [44–50].

Conflicts of Interest Statement Qiankun Luo and Hui Zhang have no conflicts of interest.

References

1. Shapiro E, Biezuner T, Linnarsson S (2013) Single-cell sequencing-based technologies will revolutionize whole-organism science. Nat Rev Genet 14(9):618–630
2. Llorens-Bobadilla E, Zhao S, Baser A, Saiz-Castro G, Zwadlo K, Martin-Villalba A (2015) Single-cell transcriptomics reveals a population of dormant neural stem cells that become activated upon brain injury. Cell Stem Cell 17(3):329–340
3. Chu MP, Kriangkum J, Venner CP, Sandhu I, Hewitt J, Belch AR, Pilarski LM (2017) Addressing heterogeneity of individual blood cancers: the need for single cell analysis. Cell Biol Toxicol 33(2):83–97. PMID: 27761761
4. Wang W, Zhu B, Wang X (2017 Oct) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427. https://doi.org/10.1007/s10565-017-9400-2
5. Deng Q, Ramskold D, Reinius B, Sandberg R (2014) Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. Science 343(6167):193–196
6. Wang W, Wang X (2017) Single-cell CRISPR screening in drug resistance. Cell Biol Toxicol 33(3):207–210. https://doi.org/10.1007/s10565-017-9396-7. PMID: 28474250
7. Treutlein B, Brownfield DG, Wu AR, Neff NF et al (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509(7500):371–375
8. Grün D, Lyubimova A, Kester L, Wiebrands K et al (2015) Single-cell messenger RNA sequencing reveals rare intestinal cell types. Nature 525(7568):251–255
9. Wang W, Gao D, Wang X (2018 Feb) Can single-cell RNA sequencing crack the mystery of cells? Cell Biol Toxicol 34(1):1–6. https://doi.org/10.1007/s10565-017-9404-y
10. Shalek AK, Satija R, Shuga J, Trombetta JJ et al (2014) Single-cell RNA-seq reveals dynamic paracrine control of cellular variation. Nature 510(7505):363–369
11. Patel AP, Tirosh I, Trombetta JJ, Shalek AK et al (2014) Single-cell RNA-seq highlights intratu-

moral heterogeneity in primary glioblastoma. Science 344(6190):1396–1401
12. Wang DC, Wang W, Zhu B, Wang X (2018 Jan 6) Lung cancer heterogeneity and new strategies for drug therapy. Annu Rev Pharmacol Toxicol 58:531–546. https://doi.org/10.1146/annurev-pharmtox-010716-104523
13. Streets AM, Zhang X, Cao C, Pang Y et al (2014) Microfluidic single-cell whole-transcriptome sequencing. Proc Natl Acad Sci U S A 111(19):7048–7053
14. Zhu Z, Qiu S, Shao K, Hou Y (2017) Progress and challenges of sequencing and analyzing circulating tumor cells. Cell Biol Toxicol https://doi.org/10.1007/s10565-017-9418-5
15. Klein AM, Mazutis L, Akartuna I, Tallapragada N et al (2015) Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. Cell 161(5):1187–1201
16. Kolodziejczyk AA, Kim JK, Svensson V et al (2015) The technology and biology of single-cell RNA sequencing. Mol Cell 58(4):610–620
17. Ponten F, Gry M, Fagerberg L, Lundberg E, Asplund A et al (2009) A global view of protein expression in human cells, tissues, and organs. Mol Syst Biol 5:337
18. Chen Z, Pan J, Jia Y, Li D, Min Z, Su X, Yuan H, Shen G, Cao S, Zhu L, Wang X (2017 Feb) Effect of memory CD4+ T cells' signal transducer and activator of transcription (STATs) functional shift on cytokine-releasing properties in asthma. Cell Biol Toxicol 33(1):27–39. https://doi.org/10.1007/s10565-016-9357-6
19. Lee JH, Daugharthy ER, Scheiman J, Kalhol R, Ferrante TC et al (2015) Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. Nat Protoc 10(3):442–458
20. Satija R, Farrell JA, Gennert D, Schier AF, Regev A (2015) Spatial reconstruction of single-cell gene expression data. Nat Biotechnol 33(5):495–502
21. Krishnaswami SR, Grindberg RV, Novotny M, Venepally P et al (2016) Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. Nat Protoc 11(3):499–524
22. Brennecke P, Anders S, Kim JK, Kolodziejczyk AA, Zhang X et al (2013) Accounting for technical noise in single-cell RNA-seq experiments. Nat Methods 10(11):1093–1095
23. Stegle O, Teichmann SA, Marioni JC (2015) Computational and analytical challenges in single-cell transcriptomics. Nat Rev Genet 16(3):133–145
24. Islam S, Zeisel A, Joost S, La Manno G, Zajac P et al (2014) Quantitative single-cell RNA-seq with unique molecular identifiers. Nat Methods 11(2):163–166
25. He S, Wurtzel O, Singh K, Froula JL, Yilmaz S, Tringe SG et al (2010) Validation of two ribosomal RNA removal methods for microbial metatranscriptomics. Nat Methods 7(10):807–812
26. Sasagawa Y, Nikaido I, Hayashi T et al (2013) Quartz-seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity. Genome Biol 14(4):R31
27. Nakamura T, Yabuta Y, Okamoto I, Aramaki S, Yokobayashi S et al (2015) SC3-seq: a method for highly parallel and quantitative measurement of single-cell gene expression. Nucleic Acids Res 43(9):e60
28. Güell M, Yus E, Lluch-Senar M, Serrano L (2011) Bacterial transcriptomics: what is beyond the RNA horiz-ome? Nat Rev Microbiol 9(9):658–669
29. Kang Y, Norris MH, Zarzycki-Siek J, Nierman WC, Donachie SP et al (2011) Transcript amplification from single bacterium for transcriptome analysis. Genome Res 21(6):925–935
30. Saliba AE, Westermann AJ, Gorski SA, Vogel J (2014) Single-cell RNA-seq: advances and future challenges. Nucleic Acids Res 42(14):8845–8860
31. Fan X, Zhang X, Wu X, Guo H, Hu Y et al (2015) Single-cell RNA-seq transcriptome analysis of linear and circular RNAs in mouse preimplantation embryos. Genome Biol 16(1):148
32. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C et al (2009) mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods 6(5):377–382
33. Picelli S, Faridani OR, Björklund AK et al (2014) Full-length RNA-seq from single cells using Smart-seq2. Nat Protoc 9(1):171–181
34. Grün D, van Oudenaarden A (2015) Design and analysis of single-cell sequencing experiments. Cell 163(4):799–810
35. Ramskold D, Luo S, Wang YC, Li R, Deng Q, Faridani OR et al (2012) Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. Nat Biotechnol 30(8):777–782
36. Tang F, Lao K, Surani MA (2011) Development and applications of single cell transcriptome analysis. Nat Methods 8(4 Suppl):S6–S11
37. Hashimshony T, Wagner F, Sher N, Yanai I (2012) CEL-seq: single-cell RNA-seq by multiplexed linear amplification. Cell Rep 2(3):666–673
38. Fan HC, Fu GK, Fodor SP (2015) Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. Science 347(6222):1258367
39. Grün D, Kester L, van Oudenaarden A (2014) Validation of noise models for singlecell transcriptomics. Nat Methods 11(6):637–640
40. Bacher R, Kendziorski C (2016) Design and computational analysis of single-cell RNA-sequencing experiments. Genome Biol 17:63
41. Bacher R, Chu LF, Leng N, Gasch AP, Thomson JA, Stewart RM et al (2017 Jun) SCnorm: robust normalization of single-cell RNA-seq data. Nat Methods 14(6):584–586. https://doi.org/10.1038/nmeth.4263
42. Niu F, Wang DC, Lu JP, Wu W, Wang XD (2016 Sep) Potentials of single-cell biology in identification and validation of disease biomarkers. J Cell Mol Med 20(9):1789–1795

43. Yu Y, Tsang JC, Wang C, Clare S, Wang J, Chen X et al (2016 Nov 3) Single-cell RNA-seq identifies a PD-1hi ILC progenitor and defines its development pathway. Nature 539(7627):102–106. https://doi.org/10.1038/nature20105
44. Wang X (2016 Oct) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361
45. Shi L, Zhu B, Xu M, Wang X (2017 Aug 4) Selection of AECOPD-specific immunomodulatory biomarkers by integrating genomics and proteomics with clinical informatics. Cell Biol Toxicol 34:109. https://doi.org/10.1007/s10565-017-9405-x
46. Chen C, Shi L, Li Y, Wang X, Yang S (2016) Disease-specific dynamic biomarkers selected by integrating inflammatory mediators with clinical informatics in ARDS patients with severe pneumonia. Cell Biol Toxicol 32(3):169–184. PMID: 27095254
47. Gu J, Wang X (2016) New future of cell biology and toxicology: thinking deeper. Cell Biol Toxicol 32(1):1–3. PMID: 26874518
48. Zhu D, Liu Z, Pan Z, Qian M, Wang L, Zhu T, Xue Y, Wu D (2016) A new method for classifying different phenotypes of kidney transplantation. Cell Biol Toxicol 32(4):323–332. PMID: 27278387
49. Wang L, Zhu B, Zhang M, Wang X (2017 Apr) Roles of immune microenvironment heterogeneity in therapy-associated biomarkers in lung cancer. Semin Cell Dev Biol 64:90–97
50. Xu M, Wang X (2017 Aug) Critical roles of mucin-1 in sensitivity of lung cancer cells to tumor necrosis factor-alpha and dexamethasone. Cell Biol Toxicol 33(4):361–371. https://doi.org/10.1007/s10565-017-9393-x

Qiankun Luo is a Resident Doctor of General Surgery at the Second Affiliated Hospital of Zhengzhou University. He has performed basic research at the Center Laboratory of Henan Provincial People's Hospital. Acute necrotic pancreatitis and chronic pancreatitis are the primary focus of his research. He is also interested in genomics and proteomics and has a keen interest in the use of targeted gene therapy to treat cancer. He has published 2 scientific papers in a well-known local journal.

Hui Zhang is an editor-in-chief assistant at Henan Provincial People's Hospital. She has worked at the Department of Hepatobiliary Surgery since graduating from Henan University in 2012.She has read a large number of articles on topics such as single-cell analysis, gene editing, molecular drug targeting, stem cells, and cell heterogeneity.

13 Detection and Application of RNA Editing in Cancer

Mengjia Qian, Claudio Spada, and Xiangdong Wang

Abstract

RNA editing is the process which happened in the post-transcriptional stage that the genetic information contained in an RNA molecule will be changed. RNA editing has been found to be related with many cancers, so through identifying RNA editing sites, we can find useful information on the process of carcinogenesis. In this review, we will discuss the main types of RNA editing and their role in cancers, as well as the current detection methods of RNA editing and the challenges we should overcome.

Keywords

RNA editing · Cancer · Detection

13.1 Introduction

RNA editing is the process which happened in the post-transcriptional stage that the genetic information contained in an RNA molecule will be changed through chemical changes in the base makeup. Although RNA editing is tissue dependent, brain is the most edited tissue while heart and muscle was the least ones, it was likely to have important function in non-brain tissues and was found to be related with many cancers [1]. Through identifying RNA editing sites and types, we can find useful information on the process of carcinogenesis, thus improving the diagnosis and treatment in cancers. Here, we will introduce some applications of RNA editing in the diagnosis and treatment in some cancers including lung cancer, hepatocellular cancer (HCC), breast cancer, and so on. Especially, they all applied by next generation sequencing (NGS). However, using NGS to detect RNA editing still has many challenges need to be overcome. In this review, we will summarize the problems and give proper suggestions on how to avoid and solve them.

13.2 RNA Editing Types

13.2.1 A-to-I vs C-to-U

Adenosine to inosine (A-to-I) and cytidine to uracil (C-to-U) are the main two types of RNA editing in mammals (Fig. 13.1) [2]. A-to-I RNA editing has the function on pre-mRNA splicing, translation and gene regulation [3], thus further affecting the stability and biogenesis of RNAs [4]. It has been reported to be discovered in intronic regions, exonic regions and 5' and 3' untranslated regions (UTRs). There are two types of A-to-I RNA editing, site-specific and

M. Qian · C. Spada · X. Wang (✉)
Zhongshan Hospital Institute of Clinical Science, Fudan University Shanghai Medical College, Shanghai, China
e-mail: Xiangdong.wang@clintransmed.org

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_13

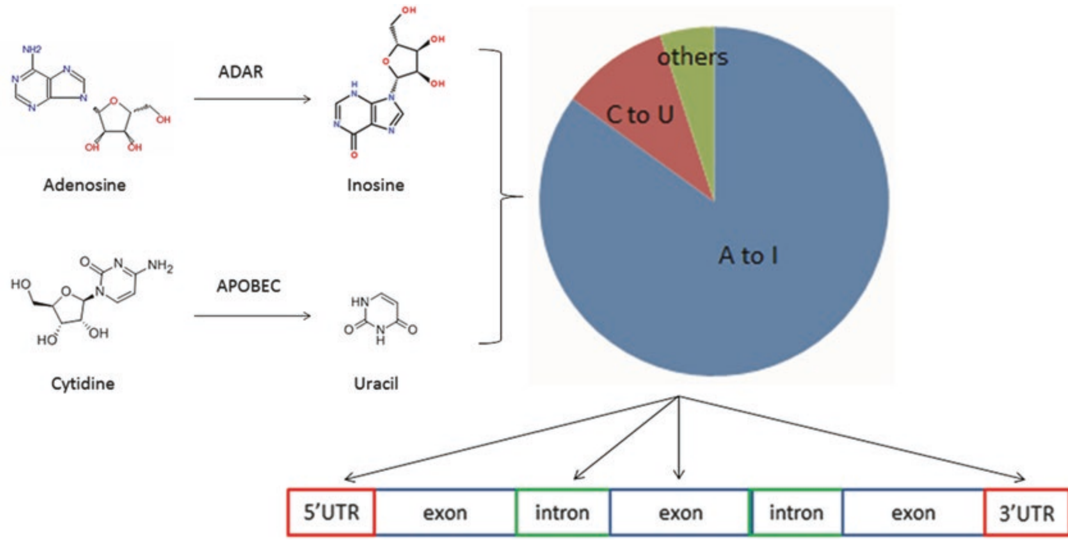


Fig. 13.1 Brief introduction of RNA editing types RNA editing can occur in intronic regions, exonic regions and 5' and 3' untranslated regions of the genes. There are mainly two types of RNA editing in humans: A to I and C to U. A to I editing is abundant and is catalyzed by ADARs while C to U editing is rare and catalyzed by APOBECs. Other types of RNA editing such as ADATs is also existed in humans that still need to be further explored and studied

promiscuous editing [5]. The former one occurred in short duplex regions will cause recoding of open reading frames that will alter the functions of the generated proteins. The latter one occurred in longer duplex such as Alus, in which nearly 50% adenosines are targeted by adenosine deaminase acting on RNA (ADAR) family of enzymes (ADARs) [6] and its frequency is varied with ADAR expression levels [7]. There are totally three ADAR genes in humans including ADAR1, ADAR2 (or ADARB1) and ADAR3 (or ADARB2). Despite ADAR3, which has no editing ability, ADAR1 and ADAR2 are the most common editing enzymes and are important in alternative splicing regulation since they have been discovered to be associated with spliceosome subunits [8]. A-to-I editing also occurred in long non-coding RNA (lncRNA) when the repetitive elements such as Alus are present [9].

C-to-U editing is catalyzed by a family of cytidine deaminases called APOBECs [10]. Currently, APOBECs include APOBEC-1, -2, -3A, -3B, -3C, -3D, -3E, -3F, -3G, -3H, -4 and activation-induced cytidine deminase (AID) [11]. Among them, APOBEC1 is the best studied one. APOBEC1 mediated RNA editing can regulate the mRNA stability of many cancer-related genes through 3' UTR and also contribute to the change of target microRNA (miRNAs) [12]. Different from APOBEC1, most RNA editing by APOBEC3A is site-specific and only targets the coding regions of genes which usually cause gene missense and nonsense alterations [13]. It also has been found its close relationship with cancer genome [14] such as BARD1, BRAC1, BRAC2, PTEN, SF3B1, TSC2, MSH2, KMT2A, MDM2, ATM and BRIP1 [15]. Moreover, APOBEC2 and APOBEC3B are also reported related with cancer. APOBEC2 has been discovered involved in RNA editing of phosphatase and tensin homolog (PTEN) and related with tumorigenesis [16]. APOBEC3B has been reported to be correlated with increased DNA damage and act as an enzymatic source of mutation in cancer [17].

13.2.2 mRNA vs Non-coding RNA

RNA editing has been found occurred in protein-coding regions of messenger RNA (mRNA)

which alters protein expression. RNA editing for mRNAs mainly includes two types, A-to-I and C-to-U [18]. The A-to-I type is the most abundant one [19] which occurs nearly 85% [20] and has been reported to be related with tumor malignancy [21]. mRNA editing through ADAR has been related with many genes such as NEIL1, BLCAP, NARF and FLNA [22]. C-to-U type is rare and only a few sites have been discovered till now [23]. Since then, RNA editing are not only been found in mRNA, but also in non-coding RNA such as miRNA, ribosomal RNA (rRNA), lncRNA and transfer RNA (tRNA) which contain inverted Alu repeats [24, 25] which until now we still do not have a clear idea of it [26].

The function of miRNA that can mediate post-transcriptional regulation such as gene silencing has been reported to be related with RNA editing [27]. Many studies have discovered that nearly 20% of the editing sites were located in 3' UTR [9] which will change the translational efficiency and stability, thus affecting the function of genes [28]. For example, miR-376 has been widely studied for its editing role in targeting and maturation [29]. However, in miRNA sequences, though nearly 10-20% are A-to-I editing [30], they are less found overrepresented. This implicated that it may have some other mechanisms existed in this context despite the role of ADARs [9]. Additional nucleotides at the 5' end may be another editing type that needs to be further and carefully studied [31].

LncRNAs are also the potential substrates for ADARs. Recent NGS approaches have showed that in human transcriptomes, lncRNAs such as Malat 1 and Jpx can be edited at several sites [9]. So are tRNA and rRNA. Different from mRNA, tRNA editing events are usually affected by adenosine deaminases acting on tRNA enzyme family (ADAT) [32]. However, we still know little about tRNA editing which still need to be further studied. rRNA plays a key role in translational regulation and the modification of it should also play a crucial role in it. rRNA editing has been found related with cancer which also should be deeper studied [33].

13.3 Current Detection Methods of RNA Editing

13.3.1 Detection Methods

The direct way to identify RNA editing sites is to compare both genomic and transcriptomic sequencing data from the same individual sample to separately call mismatches in both data and finally identify the RNA specific variant (Fig. 13.2) [34]. Since the prevalence of NGS, more than 2 million RNA editing sites in the human transcriptome have been identified [35]. Whole transcriptome deep sequencing (RNA-seq) has been regarded as the best tools to perform RNA-editing with the ability to simultaneously analyze the entire transcriptome [36]. Compared with whole exome sequencing or whole genome sequencing, some mutations edited at RNA level is more likely to be detected and focused through RNA-seq [37]. However, to identify RNA editing site more accurately, molecular features of RNA editing sites such as RNA folding changes or tissue-preferred distribution should also be considered in the analysis of complex or functionally relevant data [9]. Furthermore, deep sequencing depth is recommended to be performed in a larger extent of transcriptome samples.

Furthermore, despite NGS tool, scientists have developed another method called inosine chemical erasing (ICE). ICE is designed to identify inosine sites on RNA molecules through reverse transcription, PCR amplification and direct sequencing without comparing the sequence with cellular gene expression or genomic DNA reference. Recently, this method was further improved by combing with NGS which make it able to identify novel editing sites more precisely [38].

13.3.2 Challenges

However, errors will occur during the whole sequencing process including library preparation,

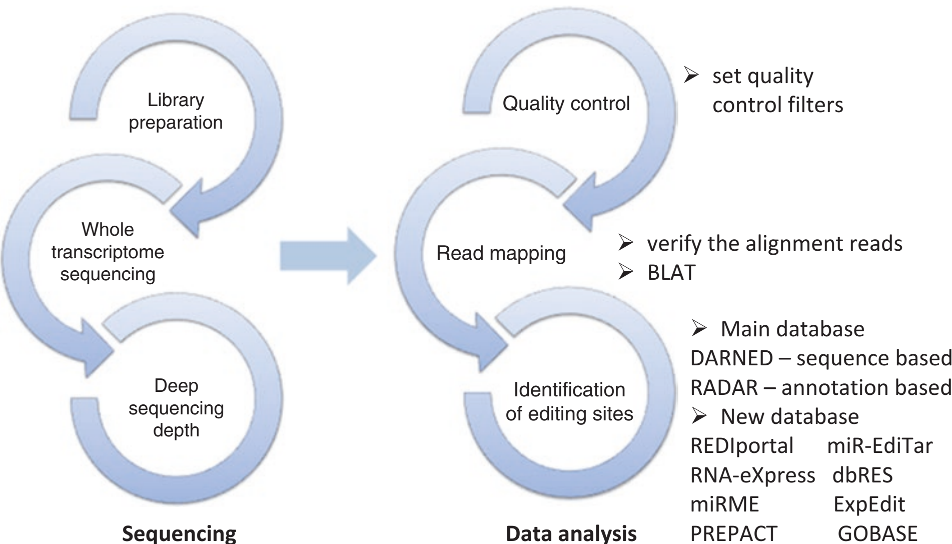


Fig. 13.2 Detection flow of RNA editing
The direct detection flow to identify RNA editing sites is to compare both genomic and transcriptomic sequencing data from the same sample to separately call mismatches in both data and finally identify the RNA specific variant. RNA-seq and deep sequencing depth is recommended during the sequencing method choose. Data analysis steps include setting quality control filters, verification of reads mapping alignment and identification of editing sites. BLAT is recommended as a highly sensitive aligner tool while DARNED and RADAR are recommended as two main databases of RNA editing, as well as some new ones specific on the special fields

sequencing and data analysis. For example, sequencing errors, mismapping errors, insufficient coverage of sequences, anomalous alignment of short-read sequences as well as single nucleotide polymorphisms (SNP) will affect the detection of RNA editing in humans [34, 39]. Moreover, false positives are also the problem that will affect the result of editome detection. They usually occur close to the start or end of reads while true positives do not show this positional bias [40]. They are often caused by technical artifacts such as reverse transcription errors or [41] caused by random hexamer used in the process of library preparation [42].

There are several ways to avoid such problems and to identify RNA editing sites accurately. Firstly, it is suggested to set quality control filters when perform data analysis and to filter out mismatches in the first six base pairs of each sequencing reads and the mismatches in introns which locate closely to the intron-exon boundary since many reads in such region are always mismapped onto processed pseudogenes or the reads ends are mismapped into adjacent introns [43]. For example, a software package, RES-Scanner, removes sites in homopolymer runs of five or more base pairs to avoid sequencing errors which has been validated its precision [44]. Another software package, REDItools, has also implemented a variety of filters to overcome the problem of sequencing biases [45]. However, currently most pipelines use ad hoc filtering methods to avoid this problem, they still cannot remove all false positive calls which still need to be further explored [46]. Secondly, it is recommended to further verify the correct alignment of mismatched reads through a highly sensitive aligner such as BLAT, a BLAST-like alignment tool with a basically different algorithm from most short-read aligners [47]. Through BLAT, it will realign all the reads which support RNA editing and then the reads are defined as qualifying reads [44]. Thirdly, to solve the position bias problem, a new software package, JACUSA, which is based on the JAVA framework for accurate SNV assessment, is developed to detect the position-specific editing events in both RNA-DNA or RNA-RNA comparisons [40]. Finally, since the existence of copy number varia-

tion, it is recommended to discard candidate editing sites with DNA read depths twice more than the peak or mean depth [44].

13.3.3 Database and Software of RNA Editing

DARNED and RADAR are the two main databases of RNA editing. They both allow people to search and identify RNA editing sites in a specific genomic location or gene. DARNED is featured on sequence based searches and dbSNP identifiers while RADAR is featured on annotation based searches including genic location, repetitive elements and editing conservation [48, 49]. Recently, there are many new databases emerged. REDIportal is a new and comprehensive database of A-to-I RNA editing in humans [50]. Samely, miR-EdiTar is also a databased focusing on A-to-I editing, but specific at miRNA binding sites. miRME is also specifically focus on miRNA editing sites using a progressive sequence alignment approach [51]. dbRES is a database contains known RNA editing sites curated from the literature and GenBank. RNA-eXpress and ExpEdit offer the function of annotate RNA editing prediction for RNA-seq data. PREPACT and GOBASE offer the interface for RNA editing data of mitochondrion and chloroplast-encoded sequences [22]. All these databases are aiming at helping to search editing sites which may have biological and functional significances.

13.3.4 Future Direction

On one hand, in the experimental field, sing cell RNA editing can be a hotspot in the future. Recently, a study has successfully demonstrated genome-wide A-to-I RNA editing at single cell level which implicated that single cell RNA editing research will be a powerful future direction [52]. On the other hand, introducing machine learning into the analysis of RNA editing is the field needs to be further explored. RED-ML, a software tool based on machine learning, has been developed to detect RNA editing sites. It can allow people to input even a single BAM file which makes it very simple to use [53]. Establishing joint models with machine learning on RNA editing analysis will be a promising direction in the future research.

13.4 Application of RNA Editing in Cancer

RNA editing will lead to post-transcriptional modification and may vary over time and tumor stage which make it potentially been related with the process of carcinogenesis [54]. Over-editing may cause some important proteins to change or lose their functions in specific tissues while normal ones do not get edited, thus it has been proved to be related with many cancers such as hepatocellular carcinoma, prostate cancer, breast cancer, colorectal cancer and lung cancer [55–59], especially in the occurrence of A-to-I editing [60]. For example, A-to-I RNA editing in colorectal cancer will cause N136S amino acid change in RHQQ which will increase the enzyme activity of RhoQ GTPase, thus increasing cancer invasion potential and cancer recurrence [57]. Here we will discuss some studies of the roles of RNA editing in some cancers (Fig. 13.3).

13.4.1 Lung Cancer

From analyzing RNA editing results, more RNA editing sites were found in tumors than in normal adjacent tissues in lung cancer patients. Among them, most RNA editing which modulated splicing, RNA structure and gene expression occurred in non-coding regions, especially in introns and UTRs. Editing in UTRs will result in differential gene expression because UTRs contain binding sites for regulatory proteins and microRNAs. Moreover, hyper-edited genes which may promote metastasis were also discovered though comparing RNA editome of tissue samples from both primary and metastatic lung cancer patients [61]. For example, ADAR1 has been reported to be significantly higher in lung adenocarcinoma

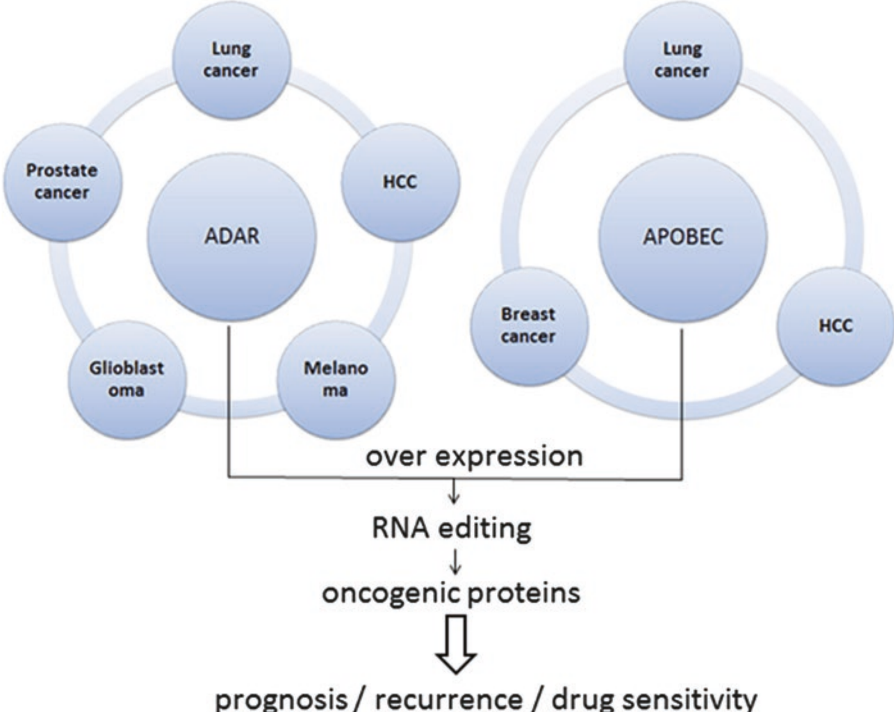


Fig. 13.3 Application of RNA editing in cancer
RNA editing will lead to post-transcriptional modification and may vary over time and tumor stage which make it potentially been related with the process of carcinogenesis. Overexpression of ADARs or APOBECs will increase the editing frequencies of target transcripts in tumors and over-editing may cause some important proteins to change or lose their functions. Thus, RNA editing may regulate the sensitivity of tumor cells to the drugs and affect the prognosis and cancer recurrence which we should further studied to explore its role in cancer diagnosis and treatment

[54] and proved to have the effect in downstream RNA editing patterns. Overexpression of ADAR1 will increase the editing frequencies of target transcripts such as miR-381 and NEIL1. Moreover, overexpression of ADAR1 will affect the patient prognosis in non-small-cell lung cancer [62]. APOBEC1 has also been detected in high levels in lung cancer patients. However, the question that whether the C-to-U editing arises during tumor formation or arises as the response to tumor has to be further studies [63].

13.4.2 Hepatocellular Carcinoma

Up regulated ADAR1 in HCC has been found to be related with the increased expression of downstream oncogenic proteins, postoperative recurrence and poor prognosis [64]. High expression of ADAR1 can edit and stabilize the mRNA of antienzyme inhibitor 1 (AZIN1) which will cause S367G amino acid substitution. S367G RNA editing site of AZIN1 is related with conformational change of the protein and the protein affinity with antizyme which will further prevent the ubiquitin-independent degradation of ornithine decarboxylase and cyclin D1, finally result in the cell proliferation. So, it is suggested that higher level of AZIN1 editing may be a potential marker in cancer diagnosis and impact cancer therapies [55]. Moreover, edited AZIN1 has recently been found to be related with the high drug resistance of an IGF-1R inhibitor, BMS536924 and IC_{50} value of many chemotherapy agents such as topo-

tecan, paclitaxel and irinotecan [65]. These results suggest that RNA editing may regulate the sensitivity of tumor cells to the drugs. ADAR2 can also elevate in HCC and mediate pre-miR-214 RNA editing that will cause the decrease of mature miR-214 and increase its downstream target Rab15 which is an important member of RAS oncogene family [65]. Moreover, APOBEC transgene expression in the liver was also reported to be the driver of HCC [66].

13.4.3 Melanoma

In melanoma, an aggressive skin cancer, ADAR1 has been found to have the ability to inhibit tumor growth and metastasis. ADAR1 mediated RNA editing could decrease the mature number of pri-miR-455 and increase the levels of its target, the tumor suppressor, cytoplasmic polyadenylation element-binding protein 1 (CPEB1) in cancer progression [67] (Fig. 13.4).

13.4.4 Prostate Cancer

Prostate cancer antigen 3 (PCA3) is an antisense intronic long noncoding RNA which is the most specific prostate cancer biomarker. PCA3 can control PRUNE2 levels through RNA editing via ADARs in prostate cancer. However, PCA3 and PRUNE2 have opposite effects on tumor growth. PCA3 may play a dominant-negative role in prostate cancer while PRUN32 may act as a suppressor role [68].

13.4.5 Breast Cancer

APOBEC1, which belongs to APOBEC cytidine deaminase and can convert cytosine to uracil during RNA editing, has recently been reported to be related with many cancers through NGS [69]. For example, the high level of APOBEC mutation was related with high level of HER2E subtype in breast cancer, which suggested that it was signifi-

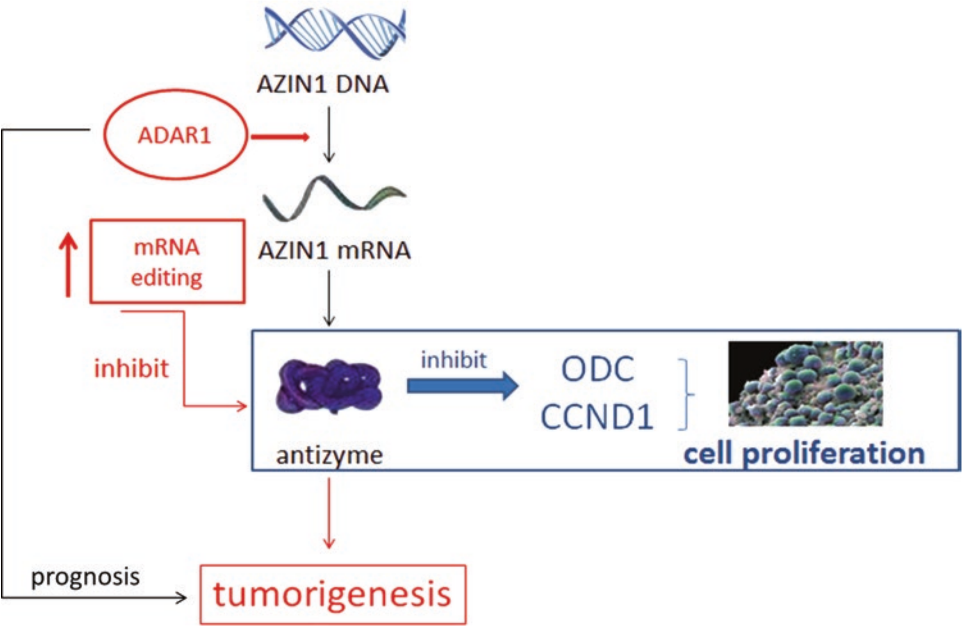


Fig. 13.4 ADAR1 mediated RNA editing in HCC

cantly linked with cancer development [70]. Moreover, reduced RNA editing of GABRA3 has been reported to have the ability to promote tumor progression, invasion and metastatic potential [71].

13.4.6 Glioblastoma

As the A-to-I editing of glutamate receptor subunit B is good to glioblastoma, A-to-I editing of miR-376a in glioblastoma has also been reported to have the ability to suppress the migration and invasiveness of glioblastoma [60]. It is edited by ADAR2, but the editing is reduced in glioblastoma. So, the number of non-edited miR-376a will increase which will finally suppress RAP2A, an important member of RAS family [72]. Moreover, ADAR2 mediated onco-miRNAs such as pri-miR-21 and pri-miR-221/222 can also contribute to the reduced number of mature miRNA level [73]. However, reduced RNA editing of AMPA receptor component GluR2 will increase the potential of tumor cell proliferation and cancer invasion [74]. All these findings indicated that RNA editing may contribute to cancer progression whether it was up-regulated or down-regulated. There is an urgent need to identify and development of biology-specific biomarkers to dynamically monitor the outcome of RNA editing in the in vitro and in vivo systems [75–79].

13.5 Conclusion

Nowadays, transcriptomics studies mainly focus on three aspects, the RNA species (mRNA and non-coding RNA), the RNA structure (start sites, splicing patterns and post-transcriptional process) and the expression levels of RNA. Among them, the RNA structure research tool, RNA editing, remains the least popular one which we still have more to explore on the role of it in cancer research [28]. RNA editing enzymes such as ADARs and APOBECs are all promising targets in cancer therapeutic strategy. Here we listed several examples of RNA editing studies in some cancers. However, their pathways are differentially regulated in cancers which should be further clearly studied. The best tool to study RNA editing is NGS. Here, we also discussed the challenges and the possible ways to overcome them. We are sure to believe that RNA editing performed by NGS has the ability in studying transcriptomes, even at single cell level. It will be sure to help a lot in cancer diagnosis and treatment in the near future.

Acknowledgements The work was supported by Zhongshan Distinguished Professor Grant (XDW), The National Nature Science Foundation of China (91230204, 81270099, 81320108001, 81270131, 81300010), The Shanghai Committee of Science and Technology (12JC1402200, 12431900207, 11410708600, 14431905100), Operation funding of Shanghai Institute of Clinical Bioinformatics, Ministry of Education for Academic Special Science and Research Foundation for PhD Education (20130071110043), and National Key Research and Development Program (2016YFC0902400, 2017YFSF090207).

References

1. Han L, Diao L, Yu S, Xu X, Li J, Zhang R et al (2015) The genomic landscape and clinical relevance of A-to-I RNA editing in human cancers. Cancer Cell 28:515–528. [PMID:26439496]
2. Licht K, Jantsch MF (2016) Rapid and dynamic transcriptome regulation by RNA editing and RNA modifications. J Cell Biol 213:15–22. [PMID:27044895]
3. Tajaddod M, Jantsch MF, Licht K (2016) The dynamic epitranscriptome: A to I editing modulates genetic information. Chromosoma 125:51–63. [PMID:26148686]
4. Yang Y, Zhou X, Jin Y (2013) ADAR-mediated RNA editing in non-coding RNA sequences. Sci China Life Sci 56:944–952. [PMID:24008387]
5. Mannion N, Arieti F, Gallo A, Keegan LP, O'Connell MA (2015) New insights into the biological role of mammalian ADARs; the RNA editing proteins. Biomolecules 5:2338–2362. [PMID:26437436]
6. Bazak L, Levanon EY, Eisenberg E (2014) Genome-wide analysis of Alu editability. Nucleic Acids Res 42:6876–6884. [PMID:24829451]
7. Bahn JH, Ahn J, Lin X, Zhang Q, Lee JH, Civelek M et al (2015) Genomic analysis of ADAR1 binding and its involvement in multiple RNA processing pathways. Nat Commun 6:6355. [PMID:25751603]
8. Fang H, Wang W (2016) Could CRISPR be the solution for gene editing's Gordian knot? Cell Biol Toxicol 32(6):465–467. [PMID: 27614448]
9. Peng Z, Cheng Y, Tan BC, Kang L, Tian Z, Zhu Y et al (2012) Comprehensive analysis of RNA-Seq data

reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30:253–260. [PMID:22327324]
10. Smith HC, Bennett RP, Kizilyer A, McDougall WM, Prohaska KM (2012) Functions and regulation of the APOBEC family of proteins. Semin Cell Dev Biol 23:258–268. [PMID:22001110]
11. Meier JC, Kankowski S, Krestel H, Hetsch F (2016) RNA Editing-Systemic Relevance and Clue to Disease Mechanisms? Front Mol Neurosci 9:124. [PMID:27932948]
12. Blanc V, Park E, Schaefer S, Miller M, Lin Y, Kennedy S et al (2014) Genome-wide identification and functional analysis of Apobec-1-mediated C-to-U RNA editing in mouse small intestine and liver. Genome Biol 15:R79. [PMID:24946870]
13. Sharma S, Patnaik SK, Taggart RT, Kannisto ED, Enriquez SM, Gollnick P et al (2015) APOBEC3A cytidine deaminase induces RNA editing in monocytes and macrophages. Nat Commun 6:6881. [PMID:25898173]
14. Swanton C, McGranahan N, Starrett GJ, Harris RSAPOBEC (2015) Enzymes: mutagenic fuel for cancer evolution and heterogeneity. Cancer Discov 5:704–712. [PMID:26091828]
15. Sharma S, Patnaik SK, Kemer Z, Baysal BE (2017) Transient overexpression of exogenous APOBEC3A causes C-to-U RNA editing of thousands of genes. RNA Biol 14:603–610. [PMID:27149507]
16. Okuyama S, Marusawa H, Matsumoto T, Ueda Y, Matsumoto Y, Endo Y et al (2012) Excessive activity of apolipoprotein B mRNA editing enzyme catalytic polypeptide 2 (APOBEC2) contributes to liver and lung tumorigenesis. Int J Cancer 130:1294–1301. [PMID:21469143]
17. Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B et al (2013) APOBEC3B is an enzymatic source of mutation in breast cancer. Nature 494:366–370. [PMID:23389445]
18. Srivastava PK, Bagnati M, Delahaye-Duriez A, Ko JH, Rotival M, Langley SR et al (2017) Genome-wide analysis of differential RNA editing in epilepsy. Genome Res 27:440–450. [PMID:28250018]
19. Lee SY, Joung JG, Park CH, Park JH, Kim JH (2015) RCARE: RNA sequence comparison and annotation for RNA editing. BMC Med Genet 8(Suppl 2):S8. [PMID:26043858]
20. Wang W, Gao D, Wang X (2017) Can single-cell RNA sequencing crack the mystery of cells? Cell Biol Toxicol 34(1):1–6. https://doi.org/10.1007/s10565-017-9404-y. [PMID: 28733864
21. Wang W, Zhu B, Wang X (2017) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427. [PMID: 28638956]
22. Wang W, Wang X (2017) Single-cell CRISPR screening in drug resistance. Cell Biol Toxicol 33(3):207–210. https://doi.org/10.1007/s10565-017-9396-7. [PMID: 28474250]
23. Blanc V, Davidson NO (2010) APOBEC-1-mediated RNA editing. Wiley Interdiscip Rev Syst Biol Med 2:594–602. [PMID:20836050]
24. Nigita G, Veneziano D, Ferro A, A-to-I RNA Editing (2015) Current knowledge sources and computational approaches with special emphasis on non-coding rna molecules. Front Bioeng Biotechnol 3:37. [PMID:25859542]
25. Li S, Mason CE (2014) The pivotal regulatory landscape of RNA modifications. Annu Rev Genomics Hum Genet 15:127–150. [PMID:24898039]
26. Nishikura K (2016) A-to-I editing of coding and non-coding RNAs by ADARs. Nat Rev Mol Cell Biol 17:83–96. [PMID:26648264]
27. Borchert GM, Gilmore BL, Spengler RM, Xing Y, Lanier W, Bhattacharya D et al (2009) Adenosine deamination in human transcripts generates novel microRNA binding sites. Hum Mol Genet 18:4801–4807. [PMID:19776031]
28. Liu H, Ma CP, Chen YT, Schuyler SC, Chang KP, Tan BC (2014) Functional Impact of RNA editing and ADARs on regulation of gene expression: perspectives from deep sequencing studies. Cell Bioscience 4:44. [PMID:25949793]
29. Nachmani D, Zimmermann A, Oiknine Djian E, Weisblum Y, Livneh Y, Khanh Le VT et al (2014) MicroRNA editing facilitates immune elimination of HCMV infected cells. PLoS Pathog 10:e1003963. [PMID:24586166]
30. Kawahara Y, Megraw M, Kreider E, Iizasa H, Valente L, Hatzigeorgiou AG et al (2008) Frequency and fate of microRNA editing in human brain. Nucleic Acids Res 36:5270–5280. [PMID:18684997]
31. Vitsios DM, Enright AJ (2015) Chimira: analysis of small RNA sequencing data and microRNA modifications. Bioinformatics 31:3365–3367. [PMID:26093149]
32. Torres AG, Pineyro D, Filonava L, Stracker TH, Batlle E, Ribas de Pouplana L (2014) A-to-I editing on tRNAs: biochemical, biological and evolutionary implications. FEBS Lett 588:4279–4286. [PMID:25263703]
33. Penzo M, Galbiati A, Trere D, Montanaro L (2016) The importance of being (slightly) modified: the role of rRNA editing on gene expression control and its connections with cancer. Biochim Biophys Acta 1866:330–338. [PMID:27815156]
34. Picardi E, Manzari C, Mastropasqua F, Aiello I, D'Erchia AM, Pesole G (2015) Profiling RNA editing in human tissues: towards the inosinome Atlas. Sci Rep 5:14941. [PMID:26449202]
35. Ramaswami G (2016) Li JB. Identification of human RNA editing sites: A historical perspective. Methods 107:42–47. [PMID:27208508]
36. Bao L, Zhang Y, Wang J, Wang H, Dong N, Su X, Xu M, Wang X (2016) Variations of chromosome 2 gene expressions among patients with lung cancer or non-cancer. Cell Biol Toxicol 32(5):419–435. [PMID: 27301951]
37. O'Brien TD, Jia P, Xia J, Saxena U, Jin H, Vuong H et al (2015) Inconsistency and features of single nucleotide variants detected in whole exome sequenc-

ing versus transcriptome sequencing: a case study in lung cancer. Methods 83:118–127. [PMID:25913717]
38. Sakurai M, Ueda H, Yano T, Okada S, Terajima H, Mitsuyama T et al (2014) A biochemical landscape of A-to-I RNA editing in the human brain transcriptome. Genome Res 24:522–534. [PMID:24407955]
39. Wang XD (2015) In: Wang XD (ed) Single cell sequencing and systems immunology, Translational bioinformatics, vol 5. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-017-9753-5
40. Piechotta M, Wyler E, Ohler U, Landthaler M, Dieterich C (2017) JACUSA: site-specific identification of RNA editing events from replicate sequencing data. BMC Bioinf 18:7. [PMID:28049429]
41. Giromini C, Rebucci R, Fusi E, Rossi L, Saccone F, Baldi A (2016) Cytotoxicity, apoptosis, DNA damage and methylation in mammary and kidney epithelial cell lines exposed to ochratoxin A. Cell Biol Toxicol 32(3):249–258. [PMID: 27154019]
42. Li JB, Church GM (2013) Deciphering the functions and regulation of brain-enriched A-to-I RNA editing. Nat Neurosci 16:1518–1522. [PMID:24165678]
43. Ramaswami G, Lin W, Piskol R, Tan MH, Davis C, Li JB (2012) Accurate identification of human Alu and non-Alu RNA editing sites. Nat Methods 9:579–581. [PMID:22484847]
44. Wang Z, Lian J, Li Q, Zhang P, Zhou Y, Zhan X et al (2016) RES-Scanner: a software package for genome-wide identification of RNA-editing sites. GigaScience 5:37. [PMID:27538485]
45. Picardi E, D'Erchia AM, Gallo A, Montalvo A, Pesole G (2014) Uncovering RNA editing sites in long non-coding RNAs. Front Bioeng Biotechnol 2:64. [PMID:25538940]
46. Toung JM, Lahens N, Hogenesch JB, Grant G (2014) Detection theory in identification of RNA-DNA sequence differences using RNA-sequencing. PLoS One 9:e112040. [PMID:25396741]
47. Wang XD, Baumgartner C, Shields DS, Deng HW, Beckmann JS (2016) In: Wang XD (ed) Application of clinical bioinformatics, Translational bioinformatics, vol 11. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-017-7543-4
48. Kiran AM, O'Mahony JJ, Sanjeev K, Baranov PV (2013) Darned in 2013: inclusion of model organisms and linking with Wikipedia. Nucleic Acids Res 41:D258–D261. [PMID:23074185]
49. Ramaswami G, Li JB (2014) RADAR: a rigorously annotated database of A-to-I RNA editing. Nucleic Acids Res 42:D109–D113. [PMID:24163250]
50. Picardi E, D'Erchia AM, Lo Giudice C, Pesole G (2017) REDIportal: a comprehensive database of A-to-I RNA editing events in humans. Nucleic Acids Res 45:D750–D7D7. [PMID:27587585]
51. Zheng Y, Ji B, Song R, Wang S, Li T, Zhang X et al (2016) Accurate detection for a wide range of mutation and editing sites of microRNAs from small RNA high-throughput sequencing profiles. Nucleic Acids Res 44:e123. [PMID:27229138]
52. St Laurent G, Tackett MR, Nechkin S, Shtokalo D, Antonets D, Savva YA et al (2013) Genome-wide analysis of A-to-I RNA editing by single-molecule sequencing in Drosophila. Nat Struct Mol Biol 20:1333–1339. [PMID:24077224]
53. Xiong H, Liu D, Li Q, Lei M, Xu L, Wu L et al (2017) RED-ML: a novel, effective RNA editing detection method based on machine learning. GigaScience. 6:1–8. [PMID:28328004]
54. Paz-Yaacov N, Bazak L, Buchumenski I, Porath HT, Danan-Gotthold M, Knisbacher BA et al (2015) Elevated RNA editing activity Is a major contributor to transcriptomic diversity in tumors. Cell Rep 13:267–276. [PMID:26440895]
55. Chen L, Li Y, Lin CH, Chan TH, Chow RK, Song Y et al (2013) Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. Nat Med 19:209–216. [PMID:23291631]
56. Wang DC, Wang X (2017 Apr) Systems heterogeneity: an integrative way to understand cancer heterogeneity. Semin Cell Dev Biol 64:1–4
57. Han SW, Kim HP, Shin JY, Jeong EG, Lee WC, Kim KY et al (2014) RNA editing in RHOQ promotes invasion potential in colorectal cancer. J Exp Med 211:613–621. [PMID:24663214]
58. Fumagalli D, Gacquer D, Rothe F, Lefort A, Libert F, Brown D et al (2015) Principles governing A-to-I RNA editing in the breast cancer transcriptome. Cell Rep 13:277–289. [PMID:26440892]
59. Qin YR, Qiao JJ, Chan TH, Zhu YH, Li FF, Liu H et al (2014) Adenosine-to-inosine RNA editing mediated by ADARs in esophageal squamous cell carcinoma. Cancer Res 74:840–851. [PMID:24302582]
60. Watanabe K, Takai D (2013) Disruption of the expression and function of microRNAs in lung cancer as a result of epigenetic changes. Front Genet 4:275–[PMID:24348521]
61. Peng L, Lee LJ, Xiong H, Su H, Rao J, Xiao D et al (2017) Characterization of RNA editome in primary and metastatic lung adenocarcinomas. Oncotarget 8:11517–11529. [PMID:28009993]
62. Anadon C, Guil S, Simo-Riudalbas L, Moutinho C, Setien F, Martinez-Cardus A et al (2016) Gene amplification-associated overexpression of the RNA editing enzyme ADAR1 enhances human lung tumorigenesis. Oncogene 35:4407–4413. [PMID:26640150]
63. Valdmanis PN, Roy-Chaudhuri B, Kim HK, Sayles LC, Zheng Y, Chuang CH et al (2015) Upregulation of the microRNA cluster at the Dlk1-Dio3 locus

in lung adenocarcinoma. Oncogene 34:94–103. [PMID:24317514]
64. Chan TH, Lin CH, Qi L, Fei J, Li Y, Yong KJ et al (2014) A disrupted RNA editing balance mediated by ADARs (Adenosine DeAminases that act on RNA) in human hepatocellular carcinoma. Gut 63:832–843. [PMID:23766440]
65. Nakano M, Nakajima M (2017) Significance of A-to-I RNA editing of transcripts modulating pharmacokinetics and pharmacodynamics. Pharmacol Ther 181:13–21. [PMID:28716651]
66. Yamanaka S, Balestra ME, Ferrell LD, Fan J, Arnold KS, Taylor S et al (1995) Apolipoprotein B mRNA-editing protein induces hepatocellular carcinoma and dysplasia in transgenic animals. Proc Natl Acad Sci U S A 92:8483–8487. [PMID:7667315]
67. Shoshan E, Mobley AK, Braeuer RR, Kamiya T, Huang L, Vasquez ME et al (2015) Reduced adenosine-to-inosine miR-455-5p editing promotes melanoma growth and metastasis. Nat Cell Biol 17:311–321. [PMID:25686251]
68. Salameh A, Lee AK, Cardo-Vila M, Nunes DN, Efstathiou E, Staquicini FI et al (2015) PRUNE2 is a human prostate cancer suppressor regulated by the intronic long noncoding RNA PCA3. Proc Natl Acad Sci U S A 112:8403–8408. [PMID:26080435]
69. Nik-Zainal S, Alexandrov LB, Wedge DC, Van Loo P, Greenman CD, Raine K et al (2012) Mutational processes molding the genomes of 21 breast cancers. Cell 149:979–993. [PMID:22608084]
70. Roberts SA, Lawrence MS, Klimczak LJ, Grimm SA, Fargo D, Stojanov P et al (2013) An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet 45:970–976. [PMID:23852170]
71. Gumireddy K, Li A, Kossenkov AV, Sakurai M, Yan J, Li Y et al (2016) The mRNA-edited form of GABRA3 suppresses GABRA3-mediated Akt activation and breast cancer metastasis. Nat Commun 7:10715. [PMID:26869349]
72. Choudhury Y, Tay FC, Lam DH, Sandanaraj E, Tang C, Ang BT et al (2012) Attenuated adenosine-to-inosine editing of microRNA-376a* promotes invasiveness of glioblastoma cells. J Clin Invest 122:4059–4076. [PMID:23093778]
73. Tomaselli S, Galeano F, Alon S, Raho S, Galardi S, Polito VA et al (2015) Modulation of microRNA editing, expression and processing by ADAR2 deaminase in glioblastoma. Genome Biol 16:5. [PMID:25582055]
74. Ishiuchi S, Yoshida Y, Sugawara K, Aihara M, Ohtani T, Watancabe T et al (2007) Ca2+-permeable AMPA receptors regulate growth of human glioblastoma via Akt activation. J Neurosci Off J Soc Neurosci 27:7987–8001. [PMID:17652589]
75. Shi L, Zhu B, Xu M, Wang X (2017) Selection of AECOPD-specific immunomodulatory biomarkers by integrating genomics and proteomics with clinical informatics. Cell Biol Toxicol 34(2):109–123. https://doi.org/10.1007/s10565-017-9405-x. [PMID: 28779230]
76. Wang X (2016) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361. [PMID: 27405768]
77. Chen C, Shi L, Li Y, Wang X, Yang S (2016) Disease-specific dynamic biomarkers selected by integrating inflammatory mediators with clinical informatics in ARDS patients with severe pneumonia. Cell Biol Toxicol 32(3):169–184. [PMID: 27095254]
78. Wang L, Zhu B, Zhang M, Wang X (2017) Roles of immune microenvironment heterogeneity in therapy-associated biomarkers in lung cancer. Semin Cell Dev Biol 64:90–97
79. Bao L, Diao H, Dong N, Su X, Wang B, Mo Q, Yu H, Wang X, Chen C (2016) Histone deacetylase inhibitor induces cell apoptosis and cycle arrest in lung cancer cells via mitochondrial injury and p53 up-acetylation. Cell Biol Toxicol 32(6):469–482. [PMID: 27423454]

Mengjia Qian, BS, Researcher of Clinical Science Institute of Fudan University Zhongshan Hospital. Her main research is focused on next generation sequencing technology and lung diseases. She is the author of more than 10 scientific publications with the impact factor about 60.

Claudio Spada serves as President of the Scientific Society Healthcare Safety and Accountability. Is Associated to The First Hospital of Wenzhou Medical University, Zhejiang, P.R. China; to the Healthcare Accountability Lab, Department of Biomedical Sciences for Health, University of Milan, Italy. His main fields of research include Biomedical Science and Technology application in Healthcare; Quality of Care and Patient's Safety; Continuing; Advanced Medical Simulation Systems, Bioethics; Medical Education. He also serves as member of the Editorial Board and as Reviewer for International Scientific Journals.

Xiangdong Wang, MD, PhD, is a Distinguished Professor of Medicine, Director of Shanghai Institute of Clinical Bioinformatics, Executive Director of Clinical Science Institute of Fudan University Zhongshan Hospital, Director of Fudan University Center of Clinical Bioinformatics, Deputy Director of Shanghai Respiratory Research Institute, and visiting professor of King's College of London. His main research is focused on clinical bioinformatics, disease-specific biomarkers, lung chronic diseases, cancer immunology, and molecular & cellular therapies. He is the author of more than 200 scientific publications with the impact factor about 900, citation number about 6918, h-index 46, i10-index 181, and cited journal impact factor about 7000.

Is Pooled CRISPR-Screening the Dawn of a New Era for Functional Genomics

14

Jufang Yao and Hui-Li Dai

Abstract

Functional genomics aims to develop an in-depth understanding of how specific gene dysfunctions are related to diseases. A common method for investigating the genome and its complex functions is via perturbation of the interactions between the DNA, RNA and their protein respective protein derivatives. Commonly, arrayed and pooled genetic screens are utilized to achieve this and in recent years have been fundamental in achieving the current level of understanding for gene dysfunctions. However, they are limited in specific aspects which scientists have attempted to address. Clustered regularly palindromic repeats (CRISPR)-based methods for genetic screens have in recent years become more prevalent but crucially shared similar properties to previous methods and failing to provide a distinct advantage over previous methods. CROP-seq, Perturb-seq, and CRISPR-seq have combined CRISPR and single-cell RNA-sequencing (scRNA-seq) and is the newest addition to the geneticist's arsenal, providing scientists with methods to edit DNA with improved speed, accuracy, and efficiency which could usher us into a new era of study methods for functional genomics. We briefly overview the CRISPR-Cas9 systems, the evolution of genetic screening in recent years, and evaluate and discuss the significance of CROP-seq, Perturb-seq, and CRISPR-seq.

Keywords

CRISPR · Genomics · Gene editing · Screening · Single cell

J. Yao
Department of Animal Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China

H.-L. Dai (✉)
Renji Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

14.1 Introduction

Various tools are available to investigate and explore complex biological processes such as the cellular circuit involved in immune cell generation and diversification. Throughout recent years genetic screening, is often utilized to analyse mammalian gene functions in a systematic way. Studies covering phenotypes and genotypes of humans have extensively been carried out, yet little is understood regarding the complex mechanisms linking phenotypic expression to specific genotypes. The lack of results is associated with a variety of reasons such as unreliable readouts, costs, excessive workload and lack of optimal analysing methods.

The clustered regularly palindromic repeats (CRISPR)-Cas9 systems are more efficient in

© Springer Nature Singapore Pte Ltd. 2018

J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_14

comparison to older methods, such as zinc finger and transcription activator-like effector (TALEN) due to its efficiency in assembling new test protein samples which has been a major contributor in its establishment as a key method in function genomics [1]. The specificity of the system is dependent on the marker sgRNA resulting in a dependence on the innate properties of the tested sgRNA. Thus, the results are potentially unreliable as sgRNAs either express high specificity to a target site but also provide hundreds of off-target sites. A major challenge is to identify the optimal sgRNAs required for a study which will target the specific target sites required resulting in high expression activity while simultaneously ignoring unwanted off-target effects. In addition, multiple Cas9 endonucleases alterations are irreversible at off-targets and can cause harmful deleterious effects, highlighting the importance of high specificity with low off-targets. Some scientists argue that even a single off-target effect could potentially be excessively detrimental as it can permanently alter the genome.

Four recently published studies [2–5] merges two methods, CRISPR-Cas9 and single-cell sequencing, by utilizing individual cells as platforms for perturbations which addresses various limitations that currently exist in screening approaches. This chapter aims to provide a framework on how these methods handle drawbacks frequently associated with CRISPR-Cas9 and what implications this means for future therapeutic applications.

14.2 CRISPR-Cas9 Structure and Function

Classification is dependent on the Cas protein with three main categories, type I-Cas3, type II-Cas9, and type III-Cas10. The CRISPR-Cas9 system is commonly referred to type II, derived from *Streptococcus pyogenes*, and was first implemented in mammalian cells by various groups [6]. Cas9 is a large multifunctional protein and is vital in the defence against viral pathogens. In contrast to TALENs and ZFNs, CRISPR-Cas9 does not require the modification of individual target sites to be specifically engineered. The system, consisting of a CRISPR RNA, Cas9, and a single guide RNA (sgRNA) has two key features, the nucleotide sequence present at 5′ end which through using the fundamentals of Watson-Crick base pairing enables the targeting of DNA sites and a double strand structure which binds to Cas9 at the 3′ end[5]. Thus, CRISPR-Cas9 system only requires a modification in the guide RNA sequence, resulting in a much more efficient and optimal process as it can cleave specific targets, resulting in DNA double-strand breaking (DSB), which can then be repaired via non-homologous end joining (NHEJ) and homology-directed recombination, preventing cell death. The protospacer adjacent motif (PAM) is a sequence motif beside the target sequence and is crucial for initial DNA binding, absence of PAM results in the inability to identify targeting sequences by the Cas9 [7, 8]. In addition, catalytically dead Cas9 (dCas9) has a versatile use, as it is acts as a DNA-binding platform that can be reprogrammed for applications such as transcriptional activation (CRISPRa) [9] or transcriptional repression (CRISPRi) [10]. High throughput screening followed by chromatin immunoprecipitation (ChIP-seq) was carried out by multiple groups to determine the binding specificity of dCas9 in the whole genome [11, 12], as nuclease activity is not present in dCas9 regulators. The results showed that the highest intensity binding was experienced at the targeted site, however varying amounts of less-intense bindings were noted at off-targeted sites, possibly due to the off-target sites having motifs similar to PAM proximal target sequences. The ChIP-seq results highlighted the binding of dCas9 to targeted genomic sites with transient binding to other regions that express a similar PAM sequence.

14.3 History of Genetic Screening and Integration of Single-Cell with CRISPR-Cas9

RNA interference (RNAi) genetic screening was initially used by the cleavage of complementary mRNAs by an exogenous RNA. This process was triggered by activating an endogenous pathway via the introduction of a 21-nucleotide double stranded RNA sequence resulting in protein activation of the Argonaut family ultimately resulting in the degradation of the cleaved mRNA. Throughout the years effective knockout methods have been developed and optimized and the technique has been considerably successful [13]. However, RNAi results were difficult to interpret due to the high possibility of off-target effects which, although widely accepted could result in faulty conclusions as they relied on a single RNA sequence to act as a surrogate for a specific gene [14–16]. Thus, multiple trials were mandatory using different sequences which targeted the specific gene and the same phenotype as previous trials. This issue, remains even in CRISPR-Cas9 screening resulting in a time consuming and costly process [17, 18].

CRISPR is frequently utilized simultaneously with a genetic screen to identify abnormal mammalian gene functions to alter. These screens, are separated into two categories; arrayed, where the perturbations are delivered to each individual cell and then observed and noted individually or, pooled, where the perturbations are delivered to the whole cell group allowing screening of large amounts of parallel perturbations in a single sample population. Arrayed preparations offer detailed high-resolution phenotype imaging including protein and transcriptional-based measurements. Limitations of arrayed platforms include the requirement of access to robolytics, its high costs, and that it's significantly more labour intensive. Additionally, it also requires a large sample pool of cells to analyse, limiting the cell types to cells that can propogate *in vitro*. Pooled readouts often measure a larger variety of parameters such as biomarker expression and drug resistances while benefiting from better scaling and efficiency, however due to the has poor content readouts and are unable to achieve high quality phenotype readout. In many cases further specifications are required after the initial screen to develop reliable results, which are then required to be re-validated due to the frequency of false positive results. Identifying and pinpointing the specific molecular mechanisms that result in desirable phenotypes can be difficult and time-consuming. Single-cell analyses are optimal when analysing individual cells at a molecular resolution which bulk sequencing measurements were unable to achieve. However, due to the time consuming and expensive process many scientists opted for other methods due to efficiency.

The creation of pooled libraries with large amounts of perturbations allowed a new avenue for genetic screening resulting in cheaper and more results produced which addressed two major limitations of RNAi and CRISPR-Cas9 screening at the expense of depth. CRISPR-Cas9 screening were first suggested by two reports [19, 20] and investigated biological pathways which were already well established from previous investigations using RNAi. Wang *et al.* created a library with approximately 7000 human genes with each gene consisting with a multitude of sgRNAs. Their aim was to identify specific parts of an altered repair pathway that causes death in KBM7 cell lines by using the purine analogue 6-thioguanine. The results highlighted the efficacy of CRISPR methods where the predicted sgRNAs targeted expected genes. Shalem *et al.* approach was to investigate the vemurafenib resistance using a genome-wide library. Multiple genes were identified to contribute to vemurafenib resistance such as NF1, MED12, NF2, or others [21, 22]. Crucially this investigation demonstrated reduced false-positive rates using CRISPR methods instead of shRNA screens as an increased fraction of perturbations were identified to target the same gene. Further studies integrated the use of dCas9 via CRISPRa or CRISPRi instead of inactivating genes through indels and DSBs which can either examine spliced isoforms through endogenous transcript level modulation by generating knockdowns or creating phenotypes identical to RNAi studies resulting in more detailed data if required to

identify the extents of phenotypic effects in contrast to just complete loss of function.

14.4 Integration of Single-Cell and CRISPR

CROP-seq, PERTURB-seq, and CRISP-seq [2–5] can simultaneously read the perturbation and phenotype of investigated cells. Thus, it can be used to derive large amounts of parallel genetic perturbations from a single sample via single cell RNA sequencing (scRNA-seq) readouts. Crucially, the entire transcriptome is recorded in comparison to just measuring cell survival resulting in accurate estimates of initial points to analyse interaction between function and gene in cells. These methods addressed the fundamental issue of how expression of sgRNA via the U6 promotor of an RNA polymerase III is unable to be read via the scRNA-seq methods as a required poly (A) tail does not exist on sgRNA. The proposed method was to generate vectors that are carriers for Pol III:sgRNA and a Pol II-driven marker that is selectable or fluorescent where a specific sgRNA sequence is included in the 3′ UTR. CROP-seq utilizes a unique solution in comparison to PERTURB-seq and CRISP-seq which are reliant on pre-generated libraries using arrayed cloning which links specific sgRNAs to its respective expressed guide barcode (GBC) that was perturbed in the transcript. CROP-seq places the Pol III:sgRNA cassette into a reporter transcript and deposited in the lentiviral vector which undergoes duplication in the viral integration phase. This solution is important for CROP-seq compatibility with established sgRNA libraries for future use in pooled CRISPR screens. Additionally, perturbations in complex cell populations and cells which were unable to be cultured could be used in a reporter-free screen and identification of many perturbations were possible within each cell. Full transcriptome sequencing in CRISP-seq crucially produces an assay that does not rely on biomarkers and can record a spectrum of phenotypes. The assay is potentially applicable in both CRISPRi and CRISPRa and highlights the flexibility of the CRISPR-CAS9 system. Adamson *et al*. and Dixit *et al*. both demonstrated the scalability of using CRISPRi by targeting libraries which contained approximately 18,900 genes and over 200,000 individual cells respectively. These methods could play a pivotal role in the understanding of complex mechanisms not only limited to the genome but also drug toxicology. Currently, the mechanisms of drug resistances, toxicity and side-effects are not fully understood and these methods could provide an important alternative to the methods we are currently performing. One potential application is the use of CRISP-seq in the probing of regulatory circuits within host genes that are responsible for drug responses which could highlight the interactions between drugs and genes [23].

A few important obstacles remain to be addressed, although a number of discussions have occurred [24–30]. Although Perturb-seq could be increased to a genome-scale the costs would steeply increase. Further optimization and maturation of this method is required to reduce the costs for individual cells and to remove unrelevant genes that are abundant. Additionally, the generation of intrinsically noisy data in these methods is another obstacle, while efforts have been made to circumvent this, such as decoupling UPR branches, predicting potential interactions between gene expression its effects, etc. it will need to be addressed prior to larger scaling. As cell-autonomous processes are the only informative phenotype that these methods can study due to inherent limitations, other processes such as stem cell differentiation would still rely on other methods such as bulk arrayed screens. Finally, a reliable way to identify successful perturbations in target sites is required due to the inability for these methods to identify the action of sgRNAs.

14.5 Conclusion and Future Perspective

CROP-seq, PERTURB-seq, and CRISP-seq are promising new methods to unravel the mysteries within the genome. However, these methods are

in their infancy and further optimisation and research are required to understand its full potential and to achieve optimal results. Mutations and gene changes are pivotal in developing resistances to drugs and thus crucial to understand. In tandem with these techniques, pooled screening can be useful to investigate the regulatory mechanisms for cell responses to stimulus, identifying host factors that are responsible for viral and bacterial protection, and could contribute further to our understanding in molecular mechanisms such as side-effects of drugs and toxicology. Limitations including the restriction of phenotype selection to cell-autonomous processes, the challenge of verifying and confirming disruptions in gene function, and high costs are points that need to be addressed and solved however it is undeniable that CROP-seq, PERTURB-seq, and CRISP-seq are simple but exciting methods that could change the landscape of functional genomics.

References

1. Fang H, Wang W (2016) Could CRISPR be the solution for gene editing's Gordian knot? Cell Biol Toxicol 32(6):465–467
2. Jaitin DA, Weiner A, Yofe I, Lara-Astiaso D, Keren-Shaul H, David E et al (2016) Dissecting immune circuits by linking CRISPR-pooled screens with single-cell RNA-seq. Cell 167(7):1883–1896.e15. https://doi.org/10.1016/j.cell.2016.11.039
3. Dixit A, Parnas O, Li B, Chen J, Fulco CP, Jerby-Arnon L et al (2016) Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. Cell 167(7):1853–1866.e17. https://doi.org/10.1016/j.cell.2016.11.038
4. Adamson B, Norman TM, Jost M, Cho MY, Nuñez JK, Chen Y et al (2016) A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. Cell 167(7):1867–1882.e21. https://doi.org/10.1016/j.cell.2016.11.048
5. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J et al (2017) Pooled CRISPR screening with single-cell transcriptome readout. Nat Methods 14(3):297–301. https://doi.org/10.1038/nmeth.4177. Epub 2017 Jan 18
6. Makarova KS, Haft DH, Barrangou R, Brouns SJ, Charpentier E, Horvath P et al (2011) Evolution and classification of the CRISPR-Cas systems. Nat Rev Microbiol 9(6):467–477. https://doi.org/10.1038/nrmicro2577
7. Sakuma T, Yamamoto T (2017) Magic wands of CRISPR-lots of choices for gene knock-in. Cell Biol Toxicol 33(6):501–505. https://doi.org/10.1007/s10565-017-9409-6
8. Wu X, Scott DA, Kriz AJ, Chiu AC, Hsu PD, Dadon DB et al (2014) Genome-wide binding of the CRISPR endonuclease Cas9 in mammalian cells. Nat Biotechnol 32(7):670–676. https://doi.org/10.1038/nbt.2889. Epub 2014 Apr 20
9. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, Whitehead EH et al (2014) Genome-scale CRISPR-mediated control of gene repression and activation. Cell 159(3):647–661. https://doi.org/10.1016/j.cell.2014.09.029
10. Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE et al (2013) CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell 154(2):442–451. https://doi.org/10.1016/j.cell.2013.06.044
11. O'Geen H, Henry IM, Bhakta MS, Meckler JF, Segal DJ (2015) A genome-wide analysis of Cas9 binding specificity using ChIP-seq and targeted sequence capture. Nucleic Acids Res 43(6):3389–3404. https://doi.org/10.1093/nar/gkv137
12. Kuscu C, Arslan S, Singh R, Thorpe J, Adli M (2014) Genome-wide analysis reveals characteristics of off-target sites bound by the Cas9 endonuclease. Nat Biotechnol 32(7):677–683. https://doi.org/10.1038/nbt.2916
13. Guo L, Zhang S, Zhang B, Chen W, Li X, Zhang W et al (2016) Silencing GTSE-1 expression inhibits proliferation and invasion of hepatocellular carcinoma cells. Cell Biol Toxicol 32(4):263–274. https://doi.org/10.1007/s10565-016-9327-z
14. biology KWGJM (2012) Use and abuse of RNAi to study mammalian gene function. Science 337(6093):421–422. https://doi.org/10.1126/science.1225787
15. Paes BCMF, Moço PD, Pereira CG, Porto GS, de Sousa Russo EM, Reis LCJ et al (2017) Ten years of iPSC: clinical potential and advances in vitro hematopoietic differentiation. Cell Biol Toxicol 33(3):233–250. https://doi.org/10.1007/s10565-016-9377-2
16. Lin C, Chen S, Li Y (2017) T cell modulation in immunotherapy for hematological malignancies. Cell Biol Toxicol 33(4):323–327. https://doi.org/10.1007/s10565-017-9397-6
17. Xu M, Wang X (2017) Critical roles of mucin-1 in sensitivity of lung cancer cells to tumor necrosis factor-alpha and dexamethasone. Cell Biol Toxicol 33(4):361–371. https://doi.org/10.1007/s10565-017-9393-x
18. Echeverri CJ, Beachy PA, Baum B, Boutros M, Buchholz F, Chanda SK et al (2006) Minimizing the risk of reporting false positives in large-scale RNAi screens. Nat Methods 3(10):777–779
19. Wang T, Wei JJ, Sabatini DM, Lander ES (2014) Genetic screens in human cells using the CRISPR-Cas9 system. Science 343(6166):80–84. https://doi.org/10.1126/science.1246981

20. Zhu Z, Wang DC, Popescu LM, Wang X (2014) Single-cell transcriptome in the identification of disease biomarkers: opportunities and challenges. J Transl Med 12:212. https://doi.org/10.1186/s12967-014-0212-3
21. Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD, Sanjana NE, Zhang F (2017) Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. Nat Protoc 12(4):828–863. https://doi.org/10.1038/nprot.2017.016
22. Whittaker SR, Theurillat JP, Van Allen E, Wagle N, Hsiao J, Cowley GS et al (2013) A genome-scale RNA interference screen implicates NF1 loss in resistance to RAF inhibition. Cancer Discov 3(3):350–362. https://doi.org/10.1158/2159-8290.CD-12-0470
23. Wang W, Wang X (2017) Single-cell CRISPR screening in drug resistance. Cell Biol Toxicol 33(3):207–210. https://doi.org/10.1007/s10565-017-9396-7
24. Kumar D, Anand T, Kues WA (2017) Clinical potential of human-induced pluripotent stem cells: perspectives of induced pluripotent stem cells. Cell Biol Toxicol 33(2):99–112. https://doi.org/10.1007/s10565-016-9370-9
25. Wang W, Gao D, Wang X (2017) Can single-cell RNA sequencing crack the mystery of cells? Cell Biol Toxicol 34(1):1–6. https://doi.org/10.1007/s10565-017-9404-y
26. Devine H, Patani R (2017) The translational potential of human induced pluripotent stem cells for clinical neurology: the translational potential of hiPSCs in neurology. Cell Biol Toxicol 33(2):129–144. https://doi.org/10.1007/s10565-016-9372-7
27. Wang W, Zhu B, Wang X (2017) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427. https://doi.org/10.1007/s10565-017-9400-2
28. Chu MP, Kriangkum J, Venner CP, Sandhu I, Hewitt J, Belch AR, Pilarski LM (2017) Addressing heterogeneity of individual blood cancers: the need for single cell analysis. Cell Biol Toxicol 33(2):83–97. https://doi.org/10.1007/s10565-016-9367-4. Epub 2016 Oct 19
29. Zhu Z, Qiu S, Shao K, Hou Y (2017) Progress and challenges of sequencing and analyzing circulating tumor cells. Cell Biol Toxicol. https://doi.org/10.1007/s10565-017-9418-5
30. Wang X (2016) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361. https://doi.org/10.1007/s10565-016-9350-0

Jufang Yao, Director of Laboratory Animal Center in Renji Hospital Affiliated to Shanghai Jiao Tong University School of Medicine. Her main research is focused on animal genetics and gene modifiications. She has been undertaken research projects from Science and Technology Commission of Shanghai Municipality. She has over 10 scientific publications.

Huili Dai, Distinguished Professor of Internal Medicine, Vice President of Renji Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Director of Shanghai Cancer Institute, Vice Chairman of Shanghai Medical Association Medical Scientific Research Management Branch, Member of Clinical Research Group of Chinese Medical Association Scientific Research Management Branch, and Committee member of the Huaxia Medical Science Awards. Her main research is focused on the treatment and prevention of renal injury. She has undertaken more than 10 research projects, including projects from National Natural Science Foundation of China as well as the Science and Technology Commission of Shanghai Municipality. She was awarded many Scientific and Technological Awards, including National and Ministry of Education, and has over 60 scientific publications.

15 Roles of Single Cell Systems Biomedicine in Lung Diseases

Yiming Zeng, Xiaoyang Chen, and Xiangdong Wang

Abstract

Single cell sequencing is important to detect the gene heterogeneity between cells, as the part of single-cell systems biology which combines computational science, mathematical modelling and high-throughput technologies with biological function and organization in the cell. We initially arise the question how to integrate the outcomes of single-cell systems biology with clinical phenotype, interpret alterations of single-cell gene sequencing and function in patient response to therapies, and understand the significance of single-cell systems biology in the discovery and development of new molecular diagnostics and therapeutics. The present review furthermore focuses the significance of singe cell systems biology in respiratory diseases and calls the special attention from scientists who are working on single cell systems biology to improve the diagnosis and therapy for patients with lung diseases.

Keywords

Single cell · Lung · Cancer · COPD · Diagnosis

15.1 Introduction

Respiratory diseases currently attract the special attention more than any time, due to the increasing incidence, morbidity, and mortality of the disease. Of those, lung cancer becomes the leading cause of mortality in multiple cancers and the number of new diagnostic lung cancer is still increasing [1–3]. The high incidence of lung cancer is associated with a large number of risk factors, including viral infection, smoking, or chronic lung diseases [1, 3]. The potential transits from smoking into chronic lung diseases or from chronic lung diseases into lung cancer are considered as one of important pathogeneses in the development of lung cancer, evidenced by the epidemiological findings that about 90% patients with chronic obstructive pulmonary diseases (COPD) and/or lung cancer had the long-term history of smoking and about 80% of patients with lung cancer were accompanied with COPD [4]. Of those, genetic backgrounds and heterogeneity are the most important factors to be considered and explored, since the small population of smokers or patients with COPD developed into lung cancer.

Authors Yiming Zeng and Xiaoyang Chen have equally contributed to this chapter as the first author.

Y. Zeng (✉) · X. Chen · X. Wang (✉)
Department of Respiratory Pulmonary and Critical Care Medicine, The Second Hospital of Fujian Medical University,
Quanzhou, Fujian Province, China
e-mail: Xiangdong.wang@clintransmed.org

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_15

Single cell RNA and DNA sequencing becomes a critical tool to detect the gene heterogeneity between cells, as the part of single-cell systems biology which combines computational science, mathematical modelling and high-throughput technologies with biological function and organization in the cell [5]. We initially arise the question how to integrate the outcomes of single-cell systems biology with clinical phenotype, interpret alterations of single-cell gene sequencing and function in patient response to therapies, and understand the significance of single-cell systems biology in the discovery and development of new molecular diagnostics and therapeutics [5]. The present review furthermore focuses the significance of singe cell systems biology in respiratory diseases and calls the special attention from scientists who are working on single cell systems biology to improve the diagnosis and therapy for patients with lung diseases.

15.2 Deep Understanding of Single Cell Systems Biology

Single cell systems biology is an emerging and exploratory science to understand the single cell from multi-directional aspects, including gene expression and sequencing, signal functioning, metabolism, proteomic profiling, imaging, and computational model. Single cell systems biology provides a full picture of the single cell phenotypes and multi-dimensional, multi-layer, multi-crossing and stereoscopic single-cell biology to deeply understand pathophysiological roles of each cell in the initiation and progression of the disease [5]. It may be difficult for clinicians to define and differ the concepts of single cell sequencing, biology, systems biology, and biomedicine and clarify the values of single cell systems biology in clinical practice.

One of the important elements in single cell systems biology is to figure out the "single-cell dynamic phenotypes", which was defined as time-dependent observable characteristics of single cells, e.g., morphology, biological properties, bio-behaviors, genetic changes, and productions [6]. The alterations of single-cell dynamic phenotypes in the disease may contribute to disease pathological categories, progression, constraints, and mutations. The single cell dynamic phenotypes can be measured and monitored in multiple aspects according to cell function and morphology. Lawson et al. investigated the genotypes for individual cells in situ after a detailed characterization of the phenotype by establishing the scale of pool-generated strain libraries [7]. Using single-molecule fluorescence time-lapse imaging, it is possible to monitor the roles of regulatory or coding sequences in the temporal expression, location, or function of a gene product, and in intracellular dynamics of a labeled reporter. Single-molecule imaging in single cells can trace intra-single-cell molecule expression, signal interactions, and locations. Liu et al. described a new approach to image dynamics of synaptic vesicle transport polarity and transcription factor hops between clustered binding sites in spatially restricted subnuclear regions [8]. It will be more important than those observations to understand molecular mechanisms by which single cell dynamic phenotypes are initiated, formed, and influenced by various microenvironments.

Changes of single cell genome and sequences are also one of critical dynamic phenotypes, and vary with cell microenvironments, e.g. the formation of gene heterogeneity and mutations. Single cell dynamic phenotypes can be the characters of organ dysfunction in the disease and have the specificity of organ and diseases. We speculate that single epithelial dynamic phenotypes vary among trachea, bronchia, small airways, and terminal airways, alter after the stimulations of different pathogens (e.g. antigen, bacteria, virus, or toxins), and are controlled by various pathologies. We demonstrated that epithelial proteomic profiles can be changed in different conditions and vary with organ-, tissue-, type-, and function-specific patterns [9]. Although the similarities and differences of epithelial proteomics between different cells, locations, and diseases were discussed, the difference between cells in the same location and disease should be further investi-

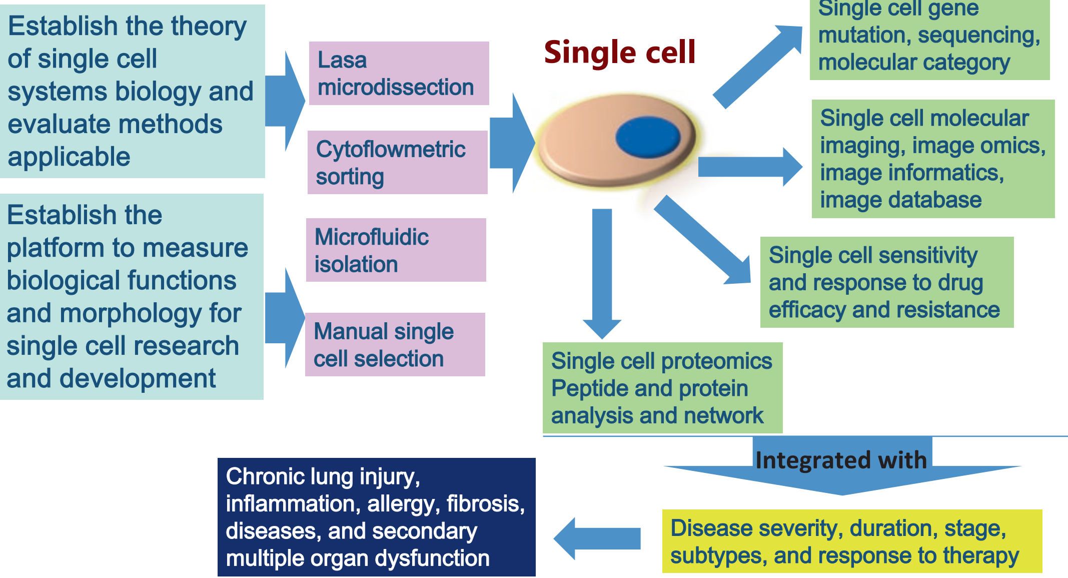


Fig. 15.1 The importance of single cell systems biology in clinical practice. The first step is to establish the theory of single cell systems biology and the platform to measure biological functions and morphology for single cell research and development. It is important to establish methods of single cell preparations applicable for clinical needs using different technologies. Isolated single cells can be used to measure single cell genomes, proteomics, image omics, and cell sensitivity to drugs. Values of single cell systems biology should be integrated with clinical phenotypes to investigate molecular mechanisms of lung diseases

gated. It will be more important if the single epithelial systems biology can be integrated with clinical questions, e.g., disease severity, duration, stage, subtypes, and response to therapy.

Figure 15.1 emphasizes the importance of single cell systems biology in clinical practice by establishing the theory of single cell systems biology and evaluate methods applicable for clinical needs and creating the new platforms to measure biological functions and morphology for single cell research and development. We evaluated a number of methods to isolate single cells and tried to find the applicability and efficiency of single cell isolation more suitable for clinical measurement [10–12]. There are a number of major obstacles to be overcome, e.g. the complexity of single cell isolation and purification processes, high cost of materials per sample, identification of targeted cells, and the repeatability of single cell analyses. Multi-directional analyses of single cell systems biology mainly contain four aspects: (1) single cell gene mutation, sequencing, and molecular category, (2) single cell molecular imaging, image omics, image informatics, and image database, (3) single cell sensitivity and response to drug efficacy and resistance, and (4) single cell proteomics, analysis and network of peptide and protein profiles. We specially emphasize the association of single cell systems biology with disease severity, duration, stage, subtypes, and response to therapy, and believe that single cell systems biology will bring more values in the application for Chronic lung injury, inflammation, allergy, fibrosis, diseases, and secondary multiple organ dysfunction (Fig. 15.1).

15.3 Comprehensive Pulmonary Single Cell Gene Expression and Sequencing

Single-cell gene expression and sequencing become more recognized parameters to describe intercellular heterogeneity, phenotypes, genealogies, somatic mosaicism, microbes, and disease

development, including single-cell DNA genome sequencing, DNA methylome sequencing, and RNA sequencing [13]. Of those, single-cell RNA sequencing (scRNA-seq) demonstrates transcriptomic cell-to-cell variation, new cell types, developmental processes, transcriptional stochasticity, transcriptome plasticity, and genome evolution [10]. Treutlein et al. primarily investigated the developmental and cellular hierarchy of the distal mouse lung epithelium using microfluidic scRNA-seq and classified cells into distinct groups using an unbiased genome-wide approach [17]. This is one of the pioneering studies to define the pulmonary epithelial differentiation and hierarchy by the principal component analysis of single cell transcriptomes and find five different cell populations and four different gene families. By comparing with epithelial cell known marker genes within the different clusters and correlating with previously reported epithelial cell type markers, Treutlein et al. suggested the intermediate population transitioning between the two alveolar lineages or a population of bipotential alveolar progenitors [17], although there is a need to be furthermore confirmed and validated in a large size of clinical samples.

Lung cancer heterogeneity is recently proposed as the critical clue to design and select the strategy of individualized therapies and re-define molecular category of tumor pathology for precision medicine therapies [18]. scRNA-seq is used for the detection of tumor heterogeneity Kim et al. applied scRNA-seq to measure the intratumoral genetic heterogeneity, single-cell heterogeneity of expressed single-nucleotide variants, and cell subgroups of lung adenocarcinoma in the model of patient-derived xenograft, which was established with harvested and seeded tumor cells from surgically resected patient tumor tissue [14]. This is one of the early outstanding studies on the commonest pathological subtype of non-small cell lung cancer. While, the major concern is whether the gene sequencing of lung cancer cells may be changed during the modeling where patient cancer cells grew within the animal system and during the cell culture of harvested cells from patient-derived xenograft. Dynamic gene heterogeneity is developing with inherent or environmental changes and involved in the evolutionary framework of tumor development with large-scale genomic alterations. The role of clonal heterogeneity in tumor evolution to drive tumor evolution and metastasis can be influenced by the immune microenvironment where the cell lives [15]. In clinical situations, the heterogeneity of cancer cell microenvironments between patients, primary and metastasis, locations, as well as therapies can also influence single-cell dynamic phenotypes of gene expression and sequencing. Targeted or non-targeted drugs per se can alter the microenvironment directly or indirectly to achieve pharmacological effects, leading to the occurrence of somatic gene mutations and changes of cell susceptibility to therapy. This is one of potential mechanisms by which the drug resistance develops during therapies.

In addition, scRNA-seq plays an important role in the detection of drug-associated mutations and specificity of targeted genes as well as drug-sensitive or drug-resistant heterogeneity between cells treated with targeting drugs. By comparing alterations of single-cell gene expression and sequencing, we can measure molecular mechanisms, drug efficacy and toxicity of identified core/driver genes and networks, to interpret the correlation of scRNA-seq findings with cell response, interaction, and phenotype. Suzuki et al. measured the transcriptome profiles and features of each cell responds to a molecularly targeted drug and the difference between parental cells and cells acquired drug resistance in lung adenocarcinoma-derived cell lines, by constructing and screening single-cell RNA-Seq libraries [19]. After then, the fusion gene transcript of major driver genes for carcinogenesis was further explored to identify a subclone of cells acquired resistance to vandetanib.

15.4 Single Cell Measurements in Lung Injury

Single cell analyses are applied for the heterogeneity between cells within the cancer tissue or in blood cancers, which contain multiple subclones or distinct clonal expansions, different from

known or detected categories [21]. Single cell analysis-detected intra- or inter-clonal heterogeneity not only occurs in cancer cells, but also in other diseases. Thus, Chu et al. called special attentions of clinicians and clinical researchers to the clinical application of single cell analysis [21]. For example, the process of epithelial cell repair and regeneration plays important roles in the development and recovery of acute and chronic lung injury by acting as the primary acceptor of the initial pathogens stimuli, while as the secondary activator to initiate the inflammatory reactions in/from other cells by producing the mediators and modulators [22–24]. Vaughan et al. applied single cell sequencing to define CC10− β4+ cells and rare Krt5-CreERT2-labeled cells, a subset of previously uncharacterized, rare lineage-negative epithelial stem/progenitor cells present within normal distal lung and found the enrichment of Myc, Klf4, pluripotency-associated transcription factors in the this population [24]. Of those, ΔNp63+ CC10− β4+ cells most closely related to the Krt5-traced cells also expressed cilia-associated genes and were activated through Notch signaling pathway, implying that distinct stem/progenitor cell pools may repopulate injured tissue and the outcomes of regeneration.

The airway remodeling, inflammation, alveolar destruction, and fibrosis are considered as one of critical characters of pulmonary fibrosis, which is one of chronic lung diseases with high mortality and morbidity. Xu et al. defined the distinct epithelial cell types harvested from animals with idiopathic pulmonary fibrosis by the difference of gene expression by scRNA-seq analysis and tracked the process of airway epithelial cell metaplasia and new differentiation/evolution during the repair of chronic lung injury [25]. Lung inflammation (e.g. leukocyte infiltration and over-production of inflammatory mediators), barrier dysfunction (e.g. edema, exclusion), and tissue injury (e.g. alveolar wall damage and repair) are the pathological features of acute and chronic lung injury. The epithelial repair is a consistent process in lung injury to heal the epithelial wound and recover the tissue damage, during which the epithelial cells can immigrate to the terminal and injured location as chemo-attracted, produce inflammatory mediators as activated, increase the number of cells as proliferated, and differentiate new types of epithelial cells as needed. Xu's study demonstrated scRNA-seq can be a critical tool to identify epithelial cell types and associated biological processes involved in the pathogenesis of pulmonary fibrosis [26]. It is no doubt that a number of subtype-specific and disease-specific biomarkers of epithelial cells can be identified by seRNAseq analyses, while it is a practical challenge to measure single epithelial cell profiles of patients with chronic lung diseases due to the limit of clinical biopsies and ethics as well as technical obstacles.

15.5 Pulmonary Single-Cell Multi-dimensions

In addition to scRNA-seq, a large number of cell function and morphology should be further explored at the single-cell level. For example, the surface damage of the cancer single-cell detected using the atomic force microscopy was used to show the drug toxicity [27]. It seems that the single-cell measurement can be used to clarify the individual cell response to drugs as a new angle of the efficacy and toxicity, although there is a question what the different responses of the cell are when it is a single, bulk solution, or intercellular communication within the tissue. In order to overcome the limit and inconsistent data of genotoxic properties, the alkaline single-cell microgel-electrophoresis also named comet assay has been used to measure the DNA-damaging ability of drugs decades ago [28]. The ethoxyresorufin-O-deethylase (EROD) activity of cytochrome P450 (CYP)1A1 was measured using the microspectrofluorometric technique in single living cell [29].

With an increasing of lung single cell measurements, a number of databases are established according to the categories of function and morphology. Du et al. developed a new web-based bioinformatics resource, named Lung Gene Expression iN Single-cell which is abbreviated as LungGENS [30], to investigate single-cell gene expression in the normal fetal mouse devel-

oping lung. This is an initial lung-specific single cell scRNA-seq database and one of a few organ-specific single cell databases opening for deep minding of lung single-cell gene expressions by searching a gene symbol, name, or lung cell type. It is no doubt that LungGENS will be an important tool to explore lung single cell function and differentiation and define lung cell types, gene signatures, gene-encoded proteins, and transcription factors. On basis of LungGENS, Du et al. furthermore developed Lung Gene Expression Analysis (LGEA) web portal as "an extended version of the LungGENS", to investigate and discover lung cell types and the dynamic changes in gene expression during lung development and function from the special aspect of scRNA-seq [31]. By integrating LungGENS and developed analytic pipeline 'SINCERA' [32], LGEA contains more comprehensive genomic and transcriptomic data from multiple platforms generated from RNA-seq from single cells, purified cell populations and whole tissue. It will be more powerful if lung single cell genomics, proteomics, metabolomics, and imaging can be integrated with the clinical phenotypes from the same individual patients.

15.6 Pulmonary Single Cell Gene Editing

There is an increasing evidence that the single-cell gene editing can be a powerful tool to define the perturbation and phenotype of a cell, of which the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas9) screen by integrating scRNA-seq with CROP-seq, CRISP-seq, Perturb-seq, with CRISPR/Cas9 can be one of the outstanding examples for gene editing. Although there are still a number of restrictions of CRISPR for clinical practice, CRISPR is considered as the powerful solution to open gene editing's "Gordian knot" [34]. The integration of those "seq" can measure the large-scale perturbations and phenotypes of a cell, as a single-cell CRISPR screen. Pooled CRISPR screening switches off multiple genes by introducing the Cas9 nuclease and single guide RNAs into cells, to define function- and target-oriented genes in the biological process, sensitivity and resistance against drugs, and screening of therapeutic strategies [16, 35].

There is a growing number of CRISPR-based studies on lung diseases to validate the value of identified targets and cancer cell sensitivity to drugs. Guernet et al. recently developed animal models of drug resistance in non-small cell lung cancer to epidermal growth factor receptor inhibitors by barcoding a specific sgRNA and a donor single-stranded DNA oligonucleotide with different genetic aberrations [36]. The high complex CRISPR-barcoding system per se is important and powerful to precisely target single specific cancer cells and identify even rare pre-existing resistant subclones. Such model can be applied for targeting drug screening and mechanism-based studies. Park et al. knocked down Gene 33 coding an adaptor protein with multiple cellular functions using CRISPR/cas9 and measured single lung epithelial cell RNA seq chronically exposed to a sublethal dose of hexavalent chromium [37]. This particular study demonstrated CRISPR-delated Gene 33 cells had the low capacity of proliferation and increased migration, and revealed a number of different genes between Gene 33-positive and negative cells. However, scRNA-seq and single cell biomedicine should be more important to describe the heterogeneity between cells, while single cell with gene editing will provide even more information on target-driven mechanism and heterogeneity.

15.7 How Far Single Cell Systems Biology Is from Clinical and Translational Medicine

It is expected that single cell systems biology should have more positive impact to the understanding of disease pathogenesis, early diagnosis, responses to therapy, and prognosis of patients with chronic lung diseases, although a large number of challenges are to be faced in future treatment regimens. Ellsworth et al. recently overviewed the latest development of methodologies and recent technological advances associ-

ated with single cell systems biology, and highlighted the knowledge of molecular heterogeneity at the genomic and transcriptomic levels of lung cancer clonal evolution and metastasis [20]. For example, gene editing by CRISPR-Cas9 using Non-homologous end joining repair and Homology directed repair can be a potential to repair the DNA damage and treat gene-dependent diseases [38], although there are a number of obstacles to be broken through.

Tang and Shrager proposed a new concept named "personalized molecular surgical therapy" that CRISPR/Cas-mediated genome editing can a new alternative to treat EGFR-mutant lung cancer [39]. The strategy of molecular surgery is designed on basis of the situation of target gene mutation, which is dependent upon the primary and secondary mutation, frequency, specificity, and length. "Molecular surgery" is proposed to edit target gene mutations in lung cancer cells using CRISPR/Cas9 through the design of single-guide RNA aiming the mutant sequences of target genes. Single cell CRISPR screening was proposed as one of the most powerful tools to evaluate the cell sensitivity and resistance to drugs [16], and single stem cell CRISPR can be a new alternative of therapies for genetic diseases [40].

In conclusion, pulmonary single cell systems biology is a new concept to understand the single cell from multi-directional aspects, including gene expression and sequencing, signal functioning, metabolism, proteomic profiling, imaging, and computational model of lung cells and lung diseases. Of those emerging area, the pulmonary single-cell dynamic phenotype is an important form to present time-dependent characteristics of pulmonary single cells, e.g., morphology, biological properties, bio-behaviors, genetic changes, and productions. Pulmonary scRNA-seq and target gene editing can be used for identification and validation of lung disease-specific biomarkers at a system-level. It is possible that "molecular surgery" by gene editing can be an alternative of future therapies for lung diseases, although lung diseases are more complex than mono-gene mutation.

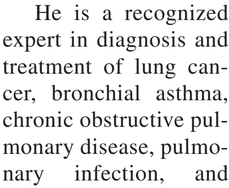
References

1. Wang X (2016) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361. https://doi.org/10.1007/s10565-016-9350-0
2. Liu X, Cho WC (2017) Precision medicine in immune checkpoint blockade therapy for non-small cell lung cancer. Clin Transl Med 6(1):7. https://doi.org/10.1186/s40169-017-0136-7
3. Reddy KP, Kong CY, Hyle EP, Baggett TP, Huang M, Parker RA et al (2017) Lung Cancer mortality associated with smoking and smoking cessation among people living with HIV in the United States. JAMA Intern Med 177(11):1613–1621. https://doi.org/10.1001/jamainternmed.2017.4349
4. Fang X, Netzer M, Baumgartner C, Bai C, Wang XD (2013) Genetic network and gene set enrichment analysis to identify biomarkers related to cigarette smoking and lung cancer. Cancer Treat Rev 39(1):77–88
5. Niu F, Wang DC, Lu J, Wu W, Wang X (2016) Potentials of single-cell biology in identification and validation of disease biomarkers. J Cell Mol Med 20(9):1789–1795. https://doi.org/10.1111/jcmm.12868
6. Wang W, Zhu B, Wang X (2017) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427. https://doi.org/10.1007/s10565-017-9400-2
7. Lawson MJ, Camsund D, Larsson J, Baltekin Ö, Fange D, Elf J (2017) In situ genotyping of a pooled strain library after characterizing complex phenotypes. Mol Syst Biol 13(10):947. https://doi.org/10.15252/msb.20177951
8. Liu H, Dong P, Ioannou MS, Li L, Shea J, Pasolli HA, et al (2017) Visualizing long-term single-molecule dynamics in vivo by stochastic protein labeling. Proc Natl Acad Sci U S A. 115(2):343–348. 201713895. https://doi.org/10.1073/pnas.1713895115
9. Zhao H, Adler KB, Bai C, Tang F, Wang X (2006) Epithelial proteomics in multiple organs and tissues: similarities and variations between cells, organs, and diseases. J Proteome Res 5(4):743–755
10. Wang XD (2015) In: Wang XD (ed) Single Cell Sequencing and Systems Immunology, Translational bioinformatics, vol 5. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9753-5
11. Zhang D, Wang X (2015) A simple protocol for single lung cancer cell isolation-making the single cell based lung cancer research feasible for individual investigator. In: Wang X (ed) Single cell sequencing and systems immunology, vol 5. Springer Netherlands, Dordrecht, pp 165–174
12. Wang J, Min Z, Jin M, Wang X (2015) Protocol for single cell isolation by flow cytometry. In: Wang X (ed) Single cell sequencing and systems immunology, vol 5. Springer Netherlands, Dordrecht, pp 155–163
13. Wang W, Gao D, Wang X (2017) Can single-cell RNA sequencing crack the mystery of cells? Cell

Biol Toxicol 34(1):1–6. https://doi.org/10.1007/s10565-017-9404-y
14. Kim KT, Lee HW, Lee HO, Kim SC, Seo YJ, et al (2015) Single-cell mRNA sequencing identifies subclonal heterogeneity in anti-cancer drug responses of lung adenocarcinoma cells. Genome Biol. 16(1):127. Published online 2015 19. https://doi.org/10.1186/s13059-015-0692-3
15. McGranahan N, Swanton C (2017) Cancer evolution constrained by the immune microenvironment. Cell 170(5):825–827. https://doi.org/10.1016/j.cell.2017.08.012
16. Wang W, Wang X (2017) Single-cell CRISPR screening in drug resistance. Cell Biol Toxicol 33(3):207–210. https://doi.org/10.1007/s10565-017-9396-7
17. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH et al (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509(7500):371–375. https://doi.org/10.1038/nature13173
18. Wang DC, Wang W, Zhu B, Wang X (2018) Lung Cancer heterogeneity and new strategies for drug therapy. Annu Rev Pharmacol Toxicol 58:531–546. https://doi.org/10.1146/annurev-pharmtox-010716-104523
19. Suzuki A, Matsushima K, Makinoshima H, Sugano S, Kohno T, Tsuchihara K, Suzuki Y (2015) Single-cell analysis of lung adenocarcinoma cell lines reveals diverse expression patterns of individual cells invoked by a molecular target drug treatment. Genome Biol 16:66. https://doi.org/10.1186/s13059-015-0636-y
20. Ellsworth DL, Blackburn HL, Shriver CD, Rabizadeh S, Soon-Shiong P, Ellsworth RE (2017) Single-cell sequencing and tumorigenesis: improved understanding of tumor evolution and metastasis. Clin Transl Med 6(1):15. https://doi.org/10.1186/s40169-017-0145-6
21. Chu MP, Kriangkum J, Venner CP, Sandhu I, Hewitt J, Belch AR, Pilarski LM. Addressing heterogeneity of individual blood cancers: the need for single cell analysis. Cell Biol Toxicol. 2017 Apr;33(2):83–97. doi: 10.1007/s10565-016-9367-4.
22. Shi L, Dong N, Ji D, Huang X, Ying Z, Wang X, Chen C (2017) Lipopolysaccharide-induced CCN1 production enhances interleukin-6 secretion in bronchial epithelial cells. Cell Biol Toxicol.:1–11. https://doi.org/10.1007/s10565-017-9401-1
23. Wang X, Adler KB, Erjefalt J, Bai CB (2007) Role of airway epithelial dysfunction in development of acute lung injury and acute respiratory distress syndrome. Expert Rev Respir Med 1(1):149–155. https://doi.org/10.1586/17476348.1.1.149
24. Shi L, Dong N, Fang XC, Wang XD (2016) Regulatory mechanisms of TGF-β1-induced fibrogenesis of human alveolar epithelial cells. J Cell Mol Med 20(11):2183–2193. https://doi.org/10.1111/jcmm.12918. Epub 2016 Jul 15
25. Vaughan AE, Brumwell AN, Xi Y, Gotts JE, Brownfield DG, Treutlein B et al (2015) Lineage-negative progenitors mobilize to regenerate lung epithelium after major injury. Nature 517(7536):621–625. https://doi.org/10.1038/nature14112
26. Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J et al (2016) Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. JCI Insight 1(20):e90558
27. Kao FS, Pan YR, Hsu RQ, Chen HM (2012) Efficacy verification and microscopic observations of an anticancer peptide, CB1a, on single lung cancer cell. Biochim Biophys Acta 1818(12):2927–2935. https://doi.org/10.1016/j.bbamem.2012.07.019
28. Villarini M1, Scassellati-Sforzolini G, Moretti M, Pasquini R (2000) In vitro genotoxicity of terbutryn evaluated by the alkaline single-cell microgel-electrophoresis "comet" assay. Cell Biol Toxicol 16(5):285–292
29. Taira Z, Yamase D, Ueda Y (2007) A new technique for assaying cytochrome P450 enzyme activity in a single cell. Cell Biol Toxicol 23(3):143–151
30. Du Y, Guo M (2015) Whitsett JA, et al. 'LungGENS': a web-based tool for mapping single-cell gene expression in the developing lung. Thorax 70:1092–1094
31. Du Y, Kitzmiller JA, Sridharan A, Ak P, Bridges JP, Misra PS et al (2017) Lung gene expression analysis (LGEA): an integrative web portal for comprehensive gene expression data analysis in lung development. Thorax 72(5):481–484. https://doi.org/10.1136/thoraxjnl-2016-209598
32. Guo M, Wang H, Potter SS et al (2015) SINCERA: a pipeline for single-cell RNA-Seq profiling analysis. PLoS Comput Biol 11:e1004575. https://doi.org/10.1371/journal.pcbi.1004575
33. Datlinger P, Rendeiro AF, Schmidl C, Krausgruber T, Traxler P, Klughammer J et al (2017) Pooled CRISPR screening with single-cell transcriptome readout. Nat Methods 14(3):297–301. https://doi.org/10.1038/nmeth.4177
34. Fang H, Wang W (2016) Could CRISPR be the solution for gene editing's Gordian knot? Cell Biol Toxicol 32(6):465–467
35. Sakuma T, Yamamoto T (2017) Magic wands of CRISPR-lots of choices for gene knock-in. Cell Biol Toxicol 33(6):501–505. https://doi.org/10.1007/s10565-017-9409-6
36. Guernet A, Mungamuri SK, Cartier D, Sachidanandam R, Jayaprakash A, Adriouch S et al (2016) CRISPR-barcoding for Intratumor genetic heterogeneity modeling and functional analysis of oncogenic driver mutations. Mol Cell 63(3):526–538. https://doi.org/10.1016/j.molcel.2016.06.017
37. Park S, Zhang X, Li C, Yin C, Li J, Fallon JT et al (2017) Single-cell RNA sequencing reveals an altered gene expression pattern as a result of CRISPR/cas9-mediated deletion of gene 33/Mig6 and chronic expo-

sure to hexavalent chromium in human lung epithelial cells. Toxicol Appl Pharmacol 330:30–39. https://doi.org/10.1016/j.taap.2017.07.003
38. Castillo A (2016) Gene editing using CRISPR-Cas9 for the treatment of lung cancer. Colomb Med (Cali) 47(4):178–180
39. Tang H, Shrager JB (2016) CRISPR/Cas-mediated genome editing to treat EGFR-mutant lung cancer: a personalized molecular surgical therapy. EMBO Mol Med 8(2):83–85. https://doi.org/10.15252/emmm.201506006
40. Paes BCMF, Moço PD, Pereira CG, Porto GS, de Sousa Russo EM et al (2017) Ten years of iPSC: clinical potential and advances in vitro hematopoietic differentiation. Cell Biol Toxicol 33(3):233–250. https://doi.org/10.1007/s10565-016-9377-2. Epub 2016 Dec 30

Yi-Ming Zeng, MD, is a Professor and President of the Second Affiliated Hospital of Fujian Medical University and member of State Council Expert for Special Allowance. He was honored as the Middle-aged Expert with Outstanding Contribution awarded by the National Health and Family Planning Commission of the People's Republic of China (NHFPC). His main *research* is focused on interventional pulmonology, sleep- breathing disorders, and noninvasive mechanical ventilation.



He is a recognized expert in diagnosis and treatment of lung cancer, bronchial asthma, chronic obstructive pulmonary disease, pulmonary infection, and interstitial lung disease. He also contributes to the establishment of medical standardizations, disease guidelines, and consensus development in China. Dr. Zeng serves as the Innovation Leading Talent in Fujian Province, Director of Sleep Medicine Key Laboratory of Fujian Province, Deputy Head of Interventional Group of the Chinese Medical Association Respiratory Branch, and Vice Chairman of Respiratory Endoscopy Special Committee of Chinese Anti-cancer Association. He undertakes a State Key Project of National Science and Technology in China, 10 projects from provincial or ministerial levels or the National Natural Science Foundation of China. He is the author of about 120 scientific publications.

Xiaoyang Chen, MD, PHD, is an associate chief physician of Department of Pulmonary and Critical Care Medicine, the Second Affiliated Hospital of Fujian Medical University. His main research is focused on sleep- breathing disorders, asthma and chronic obstructive pulmonary disease. Dr. Chen serves as vice director of Sleep Medicine Key Laboratory of Fujian Province. He is a member of the Youth Committee of Chinese Thoracic Society, and also a member of the Youth Committee of Chinese Society of Allergy. He undertakes a State Key Project of National Science and Technology in China, 2 projects from provincial or ministerial levels or the National Natural Science Foundation of China. He is the author of 12 scientific publications.

Xiangdong Wang, MD, PhD, is a Distinguished Professor of Medicine, Director of Shanghai Institute of Clinical Bioinformatics, Executive Director of Clinical Science Institute of Fudan University Zhongshan Hospital, Director of Fudan University Center of Clinical Bioinformatics, Deputy Director of Shanghai Respiratory Research Institute, and visiting professor of King's College of London. His main research is focused on clinical bioinformatics, disease-specific biomarkers, lung chronic diseases, cancer immunology, and molecular & cellular therapies. He is the author of more than 200 scientific publications with the impact factor about 900, citation number about 6918, h-index 46, i10-index 181, and cited journal impact factor about 7000.

16 The Significance of Single-Cell Biomedicine in Stem Cells

Weishan Zhuge, Furong Yan, Zhitu Zhu, and Xiangdong Wang

Abstract

Clinical application of stem cells (SCs) progresses significantly in the treatment of a large number of diseases, e.g. leukemia, respiratory diseases, kidney disease, cerebral palsy, autism, or autoimmune diseases. Of those, the population, biological phenotypes, and functions of individual SCs are mainly concerned, due to the lack of cell separation and purification processes. The single-cell technology, including microfluidic technology and single-cell genome amplification technology, is widely used to study SCs and gains some recognitions. The present review will address the importance of single-cell technologies in the recognition and heterogeneity of SCs and highlight the significance of current single-cell approaches in the understanding of SC phenotypes. We also discuss the values of single-cell studies to overcome the bottleneck in explore of biological mechanisms and reveal the therapeutic potentials of SCs in diseases, especially tumor-related diseases, as new diagnostic and therapeutic strategies

Keywords

Single-cell technology · Stem cell · Tumor · Heterogeneity · Microfluidics

16.1 Introduction

Stem cells (SCs) is a heterogeneous cell population with the capacity of self-renewal and differentiation and is suggested as an efficient alternative to treat a large number of diseases, especially human induced pluripotent stem cells (iPSC) [1–4]. SCs possess phenotypic and functional heterogeneity which are critical in the decision of SCs populations and are hardly identified, due to limitations of the separation and purification. There is growing evidence that the mechanisms of self-renewal and differentiation in stem-cell-dysfunction-associated diseases are explored at the single-cell level [2].

The single-cell technology rapidly develops to understand biology mechanisms and originations of SC population. For example, the microfluidic technology and single-cell genome amplification

Authors Weishan Zhuge and Furong Yan have equally contributed to this chapter as the first author.

W. Zhuge · X. Wang (✉)
Zhongshan Hospital Institute of Clinical Science, Fudan University Shanghai Medical School, Shanghai, China
e-mail: Xiangdong.wang@clintransmed.org

F. Yan
Department of Respiratory Pulmonary and Critical Care Medicine, The Second Hospital of Fujian Medical University, Quanzhou, Fujian Province, China

Z. Zhu (✉)
The First Hospital of Jinzhou Medical University, JinZhou, Liaoning Province, China

© Springer Nature Singapore Pte Ltd. 2018
J. Gu, X. Wang (eds.), *Single Cell Biomedicine*, Advances in Experimental Medicine and Biology 1068,
https://doi.org/10.1007/978-981-13-0502-3_16

are used to understand SC differentiation. The single-cell genome and transcriptome sequencing provides a valuable and practical insight into SCs-dominated gene screening and disease therapy. The present review will overview the application of single-cell technologies, discuss the challenges and opportunities in the application of single SCs, and highlight the significance of current single-cell approaches for SCs clinical application. We furthermore discuss the potentials to target the essential SCs for understanding of repair mechanisms in diseases, and the improvement and development of new therapies.

16.2 Needs of Single-Cell Technology

The single-cell technology becomes more important to explore SCs heterogeneity and ability to differentiate into other cells [5]. The degree of SCs diversity and heterogeneity is dependent upon the propensity of SCs, especially at single cell level, different from the community of seemingly "identical" cells. The bulk cell technology can hardly not capture dynamics or phenotypes of individual cells or cellular communities during the differentiation. Therefore, there is a great need of single-cell technology to distinguish and capture SCs origins and differentiations. Dynamic development of single-cell approaches will provide more opportunities to define SC evolution and remodeling.

16.3 Single-Cell Technology in Neural Stem Cells

One of vital challenges in single-cell gene expression surveys is to identify rare cell populations and lineage relationships, due to the need of more efficient technologies for cell capture and mRNA sequencing [6, 7]. Pollen et al. sampled single cells and analyzed the expression of stimulated genes in 301 cells from the developing human brain, to show the complexity of the human brain at single cell level [8]. In the "microfluidic" technology device, the individual cells were captured, flown into nano-scale chambers, and underwent efficient and accurate chemical transformation necessary for DNA sequencing. This particular study identified novel molecular features in diverse cell types in several brain disorders and the dysfunction of the brain cortex, such as autism and schizophrenia [8]. By integrating the single-cell RNA-seq technology with immunohistochemistry, Nowakowski et al. categorized single cells from developing cortex as astrocytes, radial glia, intermediate progenitor cells, and immature excitatory and inhibitory neurons and found that AXL receptor tyrosine kinase was highly enriched in the neural stem cells of developing fetal cerebral cortex and the retinal progenitor cells as the part of molecular mechanisms of Zika virus-associated fetal abnormalities and microcephaly [9]. The potential of neurotic SCs in differentiation and remodeling is summarized in Fig. 16.1.

Neural SCs are able to differentiate and renew itself to play the major role in repair and maintenance of the structure and function of the nerve systems. It is a challenge to study molecular properties of neural SCs, due to the relative unavailability, dynamic differentiation, and complex cellular environment of adult neural SCs. Luo et al. discovered a vital gene from ependymal $CD133^+$/GFAP-dormant cells in the neural proliferating zone of adult mice brain using single-cell transcriptome and weighted gene co-expression network analysis [10]. According to molecular properties of $CD133^+$/GFAP- E cells, a subset of dormant neural SCs was identified and isolated from regions of the brain like the lateral fourth ventricle without the neurogenic activity, as the source of $CD133^+$ ependymal neural SCs transformed into neural or glial cells by the stimulation of vascular endothelial growth factor/basic fibroblast growth factor [10]. One of important findings from this particular study is that the quiescent ependymal neural SCs normally existed in the ventricular areas could be activated and differentiated during injury. The single cell RNA sequencing (scRNAseq) can be a sensitive and accurate technique to analyze the expression of single cell microRNAs or RNA genome,

Fig. 16.1 The single-cell technology is applied for molecular properties of neural Stem cells (SCs). Neural single cells are harvested from developing cortex and are also differentiated into the cortex (e.g. radial glia, neurons, intermediate progenitor cells, or astrocytes) and neural retina (e.g. retinal progenitor cells). AXL receptor tyrosine kinase is highly enriched in the neural stem cells of developing fetal cerebral cortex and the retinal progenitor cells which are associated with Zika virus. CD133⁺/GFAP- E cells is identified and isolated from regions of the brain like the lateral fourth ventricle and can be transformed into neural or glial cells by the stimulation of vascular endothelial growth factor (VEGF)/basic fibroblast growth factor (bFGF). With single-cell technology to characterize the gene expression profiles of nestin+ tendon stem/progenitor cells from tendon tissue, nestin plays the crucial roles in progenitor cell fate decisions and phenotype maintenance

multiple transcription, and the corresponding resulting protein, to identify the new genes.

Nestin-rich tendon stem/progenitor cells were found in tendon-derived cells using single-cell gene analysis, showing the expression and function of nestin in repair and differentiation of tendon SCs after injury. Yin et al. isolated the single-cell transcript genes and analyzed their expression, to understand the function and expression of nestin during injury repair and differentiation [11]. This particular study isolated major tendon cells and subpopulation of nestin⁺ tendon stem/progenitor cells and demonstrated that the expression of nestin gene was necessary for differentiation and involvement of nestin⁺ tendon stem/progenitor cells in tissue repair (Fig. 16.1). Of those, TGF-β was proposed as a signal activating molecule to recruit nestin⁺ tendon stem/progenitor cells during tissue injury for repairing, remodeling, and regeneration of tendon. Such exploratory investigation offers new understanding the biological significance of tendons, as an alternative of new therapies.

16.4 Interaction and Development of Intestinal SCs

The tissue microenvironment contains a large number of vital extrinsic factors and signals govern SCs self-renewing and differentiation properties. Gracz et al. established an in vitro model to study the interactions between intestinal stem cells and surrounding cells and found that the direct contact between the intestinal SCs and PCs and contact-produced stemness are vital and necessary factors, reflected by the single cell gene expression with micro raft array [12]. The scRNA-seq technology allows us to observe dynamic changes of the gene expression over the course of enteroid development, which can be integrated with morphological differences in early asteroids. The development of stem cell niche or organoid can be detected using micro raft array. Before differentiating into matured cells, stem cells undergo the phase of transit-amplifying progenitors with the capacity of unidirectional differentiation to the cells with absorption or secretory function (Fig. 16.2), although the mechanism about the process of multilineage priming remains unclear. Kim et al. distinguish two discrete populations of $Lgr5^+$ intestinal SCs using microfluidic quantitative real time polymerase chain reaction, single-mRNA in situ hybridization, bDNA amplification, and single-cell gene expression [13]. $Lgr5^+$ cells may have potential progenitors' characteristics, to differ from and identify the priming of multi-lineage genes. Single cell technology is expected to explore SC biology and understand the genesis of the disease, cancers, and other disorders (Fig. 16.2).

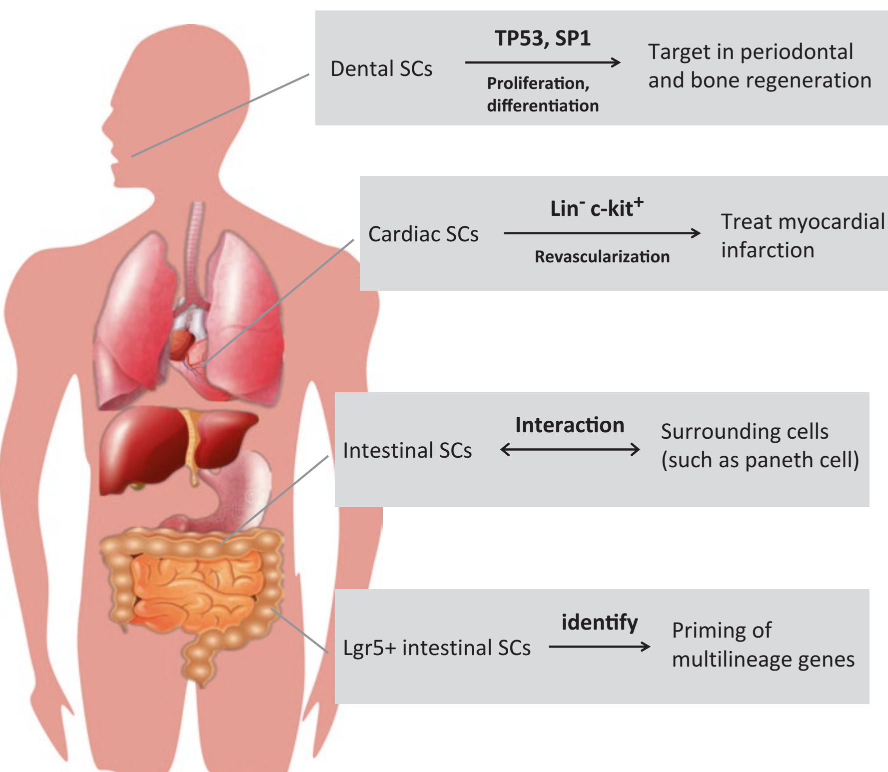


Fig. 16.2 The transcription factors TP53 and SP1 play an important role in the progress in periodontal and bone regeneration of dental stem cells (SCs), and may act as a promising potential target for the clinical applications of bioengineering in periodontal and bone regeneration. The injection of designated gene tagged Lin- $c\text{-}kit^+$ cells can treat myocardium infarction. The molecular mechanisms of multilineage gene priming is associated with the interaction between intestinal SCs and the surrounding cells

16.5 Single-Cell Technology in Cardiac SCs

C-kit is a stem cell factor receptor in multipotent adult cardiac SCs and the majority of the c-kit-positive cardiac SCs more than 90% are CD45⁺ CD31⁺ c-kit⁺ cells which have no SCs capacity, whiles the rest is the Lin⁻CD45⁻ c-kit⁺ cardiac cells with the properties of adult multipotent cardiac SCs, accounting for less than 10%. Vicinanza et al. separated Lin⁻ CD45⁻ c-kit⁺ cardiac cells from the heterogeneous c-kit-expressing cardiac cell population, via Lin and CD45-positive or -negative sorting, to distinguish between cardiac SCs and other c-kit⁺ cells [14]. The injection of designated gene tagged Lin⁻ c-kit⁺ cells into the infarcted myocardium could induce the significant regeneration of new arterioles and capillaries, but not CD45⁺ CD31⁺ c-kit⁺ cardiac cells [14]. A few of cardiac SCs with positive/negative markers appear in the regenerative area. Single cell biomedicine can provide more information and understanding of tissue SCs sources and identities for the application of SCs (Fig. 16.3).

16.6 Single-Cell Technology in Dental Stem Cells

The transcription factors TP53 and SP1 play the important roles in the progress in periodontal and bone regeneration of dental SCs. In the overex-

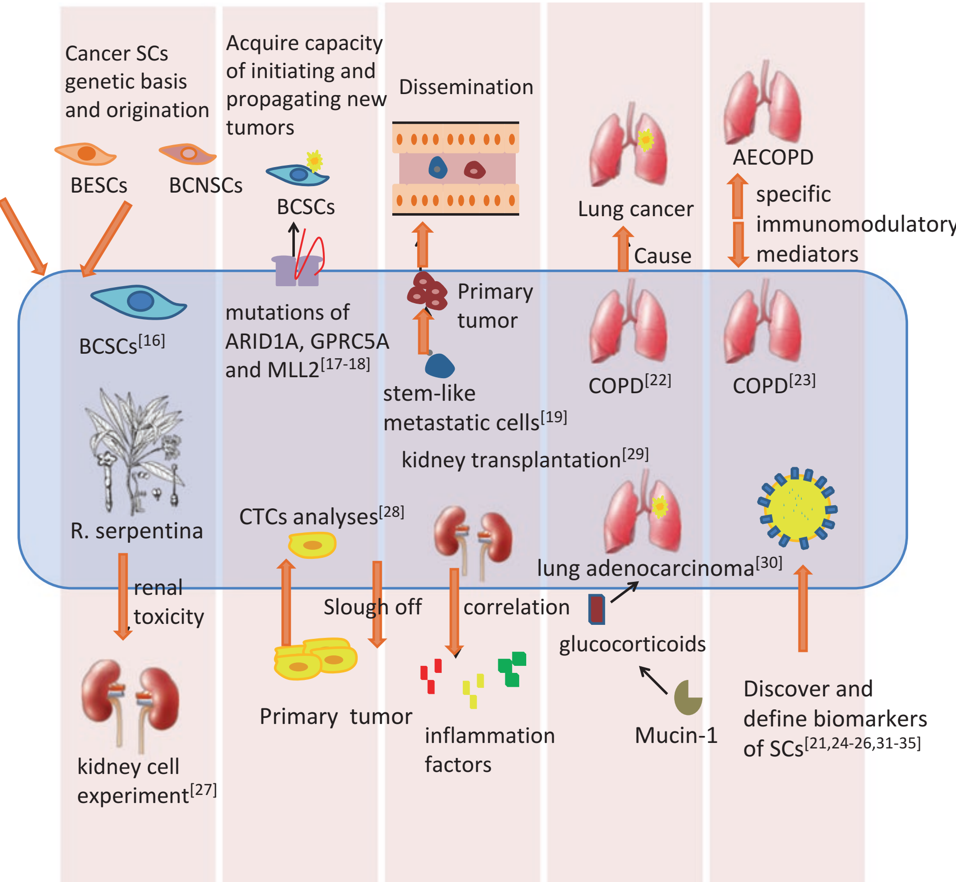


Fig. 16.3 The single-cell technology shows a powerful screening function to achieve the separation and identification stem cells (SCs) and biomarkers. *BCSCs* bladder cancer stem cells, *BCNSCs* bladder cancer non-stem cells, *BESCs* bladder epithelial stem cells, *ARID1A* AT-rich interaction domain 1A, *GPRC5A* G protein-coupled receptor class C group 5 member A, *MLL2* myeloid/lymphoid or mixed-lineage leukemia 2, *CTC* circulating tumor cells, *COPD* chronic pulmonary obstructive disease, and *AECOPD* acute exacerbation of COPD

pression of SP1 and TP53, SP1 could change cell proliferation and influence TP53 osteogenic differentiation of SCs in the dental follicle and pulp [15]. Of those, TP53 and SP1 could decide the expression of osteogenic marker genes, biological processes, and cell proliferation of SCs from human exfoliated deciduous teeth (Fig. 16.3). It lead to the new potential that targeted SCs act as a promising potential target for the clinical applications of bioengineering in periodontal and bone regeneration.

16.7 Single-Cell Technology in Tumor Stem Cells

Cancer SCs are considered as the significant source for tumor-formation, metastasis and drug-resistance, although the understanding of cancer SC genetic basis remains limited. By Using single-cell sequencing, Yang et al. found the homogeneity and heterogeneity of human bladder cancer SCs originated from bladder epithelial SCs or bladder cancer non-SCs [16]. Bladder cancer SCs have 21 main altered genes, of which ETS1, PITX2, PAWR, GPRC5A, MKL1 and RGS9BP are non-mention mutated genes in BC were identified in BCSCs. Of those genes, mutations of ARID1A, GPRC5A and MLL2 contribute to the self-renewal capability of bladder cancer SCs.

The metastases may arise from unique tumor cells with the capacity of initiating and propagating new tumors [17, 18]. Cancer SCs can be the major source of metastasis initiation and progression, as metastasis-initiating cells. There is one of the greatest challenges to discover and define tissue-specific, cell-specific, and disease-specific biomarkers of SCs, especially cancer SCs [19–33]. Biomarkers of initiating cells not only include 116 genes, but also biological behaviors, e.g. stemness, dormancy, cell cycle and proliferation gene, epithelial-to-mesenchymal transition, and mammary lineage specification, using microfluidics technology and single-cell multiplex gene sequencing analysis [34]. The metastatic cells from high-burden tissues resemble primary tumor cells, whereas low-burden tissues have tumor-initiating capacity, which are regulated by the activation of MYC gene and down-regulated by cyclin dependent kinase inhibitors [34]. The heterogeneity of those cancer SCs in the carcinogenesis and metastasis can be furthermore explored using single cell technologies, e.g. scRNA-seq (Fig. 16.3).

16.8 Conclusion

Single cell biomedicine is an emerging science of stem cells to understand the biological behaviors, e.g. engraftment, survival, migration, and differentiation, and heterogeneity of stem cells. Single cell technology becomes one of the powerful tools to explore the specific hematopoietic lineage of stem cells and potential applications, although there are a number of challenges to be faced [35]. The new application of single-cell methods can be an unexpected additional strategies to answer several longstanding questions of SC biology, as listed in Table 16.1. The strength and weakness of the single-cell technology as listed in Table 16.2 should be fully considered.

Table 16.1 Application of single-cell approaches in stem cells

Application	Significance
Neural stem cells	1. The penetration in biological mechanism of brain disorders
	2. Locate the vulnerable parts of ZIKV infection
	3. Activation signal in dormant neural
Tendon stem cells	Therapy of tendon injury.
Intestinal stem cells	1. Multilineage gene priming of stem cells
	2. Interaction between cells and stem cells
Cardiac stem cells	Separation and identification true cardiac stem cells
Dental stem cells	Bone regeneration-associated impact factors
Cancer stem cells,	
bladder cancer	1. Explore of genetic basis and origination
breast cancer	2. Prophylaxis and treatment metastatic disease

Table 16.2 Comparison the single-cell approaches used in stem cells

	Cellular culture and Isolation Techniques	Gene expression and sequencing technology
Approach	1. Cell isolated technique: Dilution separation technique; FACS; microscopy separation technique	1. (Multiplex) RT-PCR,Quantitative Single Cell Multiplex RT-PCR
	2. Clonogenic assay, spherogenesis assay	2. Analysis of gene expression profiles: whole genome sequencing; transcriptome sequencing,
	3. Transplantation technology	
Strength	Multiparameter flow cytometry based sorting; Amplification and cultivation of monoclonal cell group; high throughput screening	Genome-wide and transcriptome information; Effective identification of cell subsets
Weakness	Ex vivo purification may affect viability; insufficient prospective markers; in vitro culture may not reflect in vivo conditions	No combination of genome and transcriptome information; Cell activity disappeared after detection; Limits toward number of cells being probed;
Application	Blood, muscle, intestinal, skin, neurons, heart	Blood, brain, skin, intestinal, bone, muscle cancer, primordial germ cells

Single-cell technology provides new potentials to analyze the biological functions of SCs and the therapeutic roles of SCs in diseases. The single-cell technology overcomes the bottleneck to explore biological mechanisms of SC sand enriches our understanding on stem cell biology, therapy in diseases, and pathogenesis of cancer reoccurrence, metastasis, or drug resistance.

References

1. Barker RA, Parmar M, Studer L, Takahashi J (2017 Nov 2) Human trials of stem cell-derived dopamine neurons for Parkinson's disease: Dawn of a new era. Cell Stem Cell 21(5):569–573. https://doi.org/10.1016/j.stem.2017.09.014
2. Devine H, Patani R (2017) The translational potential of human induced pluripotent stem cells for clinical neurology: the translational potential of hiPSCs in neurology. Cell Biol Toxicol 33(2):129–144
3. Kumar D, Anand T, Kues WA (2017) Clinical potential of human-induced pluripotent stem cells: perspectives of induced pluripotent stem cells. Cell Biol Toxicol 33(2):99–112
4. Paes BCMF, Moço PD, Pereira CG, Porto GS, de Sousa Russo EM, Reis LCJ et al (2017) Ten years of iPSC: clinical potential and advances in vitro hematopoietic differentiation. Cell Biol Toxicol 33(3):233–250. https://doi.org/10.1007/s10565-016-9377-2
5. Ema H, Morita Y, Suda T (2014) Heterogeneity and hierarchy of hematopoietic stem cells. Exp Hematol 42(2):74–82.e2. https://doi.org/10.1016/j.exphem.2013.11.004
6. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R et al (2013) Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. Nature 498(7453):236–240. https://doi.org/10.1038/nature12172
7. Treutlein B, Brownfield DG, Wu AR, Neff NF, Mantalas GL, Espinoza FH et al (2014) Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq. Nature 509(7500):371–375. https://doi.org/10.1038/nature13173
8. Pollen AA, Nowakowski TJ, Shuga J, Wang X, Leyrat AA, Lui JH et al (2014) Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. Nat Biotechnol 32(10):1053–1058. https://doi.org/10.1038/nbt.2967
9. Nowakowski TJ, Pollen AA, Di Lullo E, Sandoval-Espinosa C, Bershteyn M, Kriegstein AR (2016) Expression analysis highlights AXL as a candidate Zika virus entry receptor in neural stem cells. Cell Stem Cell 18(5):591–596. https://doi.org/10.1016/j.stem.2016.03.012
10. Luo Y, Coskun V, Liang A, Yu J, Cheng L, Ge W et al (2015) Single-cell transcriptome analyses reveal signals to activate dormant neural stem cells. Cell 161(5):1175–1186. https://doi.org/10.1016/j.cell.2015.04.001
11. Yin Z, Hu JJ, Yang L, Zheng ZF, An CR, Wu BB et al (2016) Single-cell analysis reveals a nestin+ tendon stem/progenitor cell population with strong tenogenic potentiality. Sci Adv 2(11):e1600874. https://doi.org/10.1126/sciadv.1600874
12. Gracz AD, Williamson IA, Roche KC, Johnston MJ, Wang F, Wang Y et al (2015) A high-throughput platform for stem cell niche co-cultures and downstream gene expression analysis. Nat Cell Biol 17(3):340–349. https://doi.org/10.1038/ncb3104
13. Kim TH, Saadatpour A, Guo G, Saxena M, Cavazza A, Desai N et al (2016) Single-cell transcript profiles reveal multilineage priming in early progenitors derived from Lgr5(+) intestinal stem cells. Cell Rep 16(8):2053–2060. https://doi.org/10.1016/j.celrep.2016.07.056

14. Vicinanza C, Aquila I, Scalise M, Cristiano F, Marino F, Cianflone E et al (2017) Adult cardiac stem cells are multipotent and robustly myogenic: c-kit expression is necessary but not sufficient for their identification. Cell Death Differ 24(12):2101–2116. https://doi.org/10.1038/cdd.2017.130
15. Felthaus O, Viale-Bouroncle S, Driemel O, Reichert TE, Schmalz G, Morsczeck C (2012) Transcription factors TP53 and SP1 and the osteogenic differentiation of dental stem cells. Differentiation 83(1):10–16. https://doi.org/10.1016/j.diff.2011.08.008
16. Yang Z, Li C, Fan Z, Liu H, Zhang X, Cai Z et al (2017) Single-cell sequencing reveals variants in ARID1A, GPRC5A and MLL2 driving self-renewal of human bladder Cancer stem cells. Eur Urol 71(1):8–12. https://doi.org/10.1016/j.eururo.2016.06.025
17. Oskarsson T, Batlle E, Massagué J (2014) Metastatic stem cells: sources, niches, and vital pathways. Cell Stem Cell 14(3):306–321. https://doi.org/10.1016/j.stem.2014.02.002
18. Dieter SM, Ball CR, Hoffmann CM, Nowrouzi A, Herbst F, Zavidij O et al (2011) Distinct types of tumor-initiating cells form human colon cancer tumors and metastases. Cell Stem Cell 9(4):357–365. https://doi.org/10.1016/j.stem.2011.08.010
19. Frey O, Misun PM, Fluri DA, Hengstler JG, Hierlemann A (2014) Reconfigurable microfluidic hanging drop network for multi-tissue interaction and analysis. Nat Commun 5:4250. https://doi.org/10.1038/ncomms5250
20. Wang X (2016) New biomarkers and therapeutics can be discovered during COPD-lung cancer transition. Cell Biol Toxicol 32(5):359–361. https://doi.org/10.1007/s10565-016-9350-0
21. Shi L, Zhu B, Xu M, Wang X (2017) Selection of AECOPD-specific immunomodulatory biomarkers by integrating genomics and proteomics with clinical informatics. Cell Biol Toxicol 34:109. https://doi.org/10.1007/s10565-017-9405-x
22. Chen C, Shi L, Li Y, Wang X, Yang S (2016) Disease-specific dynamic biomarkers selected by integrating inflammatory mediators with clinical informatics in ARDS patients with severe pneumonia. Cell Biol Toxicol 32(3):169–184. https://doi.org/10.1007/s10565-016-9322-4
23. Gu J, Wang X (2016) New future of cell biology and toxicology: thinking deeper. Cell Biol Toxicol 32(1):1–3. https://doi.org/10.1007/s10565-016-9313-5
24. Wang W, Zhu B, Wang X (2017) Dynamic phenotypes: illustrating a single-cell odyssey. Cell Biol Toxicol 33(5):423–427. https://doi.org/10.1007/s10565-017-9400-2
25. Mossoba ME, Flynn TJ, Vohra S, Wiesenfeld PL, Sprando RL (2015) Human kidney proximal tubule cells are vulnerable to the effects of Rauwolfia serpentina. Cell Biol Toxicol 31(6):285–293. https://doi.org/10.1007/s10565-016-9311-7
26. Zhu Z, Qiu S, Shao K, Hou Y (2017) Progress and challenges of sequencing and analyzing circulating tumor cells. Cell Biol Toxicol. https://doi.org/10.1007/s10565-017-9418-5
27. Zhu D, Liu Z, Pan Z, Qian M, Wang L, Zhu T, Xue Y, Wu D (2016) A new method for classifying different phenotypes of kidney transplantation. Cell Biol Toxicol 32(4):323–332. https://doi.org/10.1007/s10565-016-9337-x
28. Xu M, Wang X (2017) Critical roles of mucin-1 in sensitivity of lung cancer cells to tumor necrosis factor-alpha and dexamethasone. Cell Biol Toxicol 33(4):361–371. https://doi.org/10.1007/s10565-017-9393-x
29. Wang X (2016) CBT profiles of cabozantinib approved for advanced renal cell carcinomas. Cell Biol Toxicol 32(4):259–261. https://doi.org/10.1007/s10565-016-9349-6
30. Bao L, Zhang Y, Wang J, Wang H, Dong N, Su X, Xu M, Wang X (2016) Variations of chromosome 2 gene expressions among patients with lung cancer or non-cancer. Cell Biol Toxicol 32(5):419–435. https://doi.org/10.1007/s10565-016-9343-z
31. Lin C, Chen S, Li Y (2017) T cell modulation in immunotherapy for hematological malignancies. Cell Biol Toxicol 33(4):323–327. https://doi.org/10.1007/s10565-017-9397-6
32. Shi L, Dong N, Ji D, Huang X, Ying Z, Wang X, Chen C (2017) Lipopolysaccharide-induced CCN1 production enhances interleukin-6 secretion in bronchial epithelial cells. Cell Biol Toxicol 34:39. https://doi.org/10.1007/s10565-017-9401-1
33. Arumugam P, Samson A, Ki J, Song JM (2017) Knockdown of clusterin alters mitochondrial dynamics, facilitates necrosis in camptothecin-induced cancer stem cells. Cell Biol Toxicol 33(3):307–321. https://doi.org/10.1007/s10565-016-9378-1
34. Lawson DA, Bhakta NR, Kessenbrock K, Prummel KD, Yu Y, Takai K et al (2015) Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. Nature 526(7571):131–135. https://doi.org/10.1038/nature15260
35. Devine H, Patani R (2017) The translational potential of human induced pluripotent stem cells for clinical neurology: the translational potential of hiPSCs in neurology. Cell Biol Toxicol 33(2):129–144. https://doi.org/10.1007/s10565-016-9372-7

Furong Yan, Master of Medicine, Researcher of Central Laboratory of the Second Affiliated Hospital of Fujian medical University. Her main research is focused on ion channel and signal transduction in cardiovascular diseases and lung diseases. She is the author of 3 scientific publications with the impact factor about 12.

Zhitu Zhu, MD, PhD, is a Professor and Chief Physician of Clinical Oncology, and a Vice-President of The First Hospital of Jinzhou Medical University, Jinzhou, China. He was honored with the Youth Talent of Liaoning Medical University His major research is focusing on gastric cancer and metabolomics, tissue biobank, carcinogenesis, and personalized diagnosis and treatment of the digestive tract tumor and clinical treatments.

He is a leading physician in chemotherapy, target therapy, and immune therapy for malignant tumor. He is the member of Chinese Medical Doctor Association Clinical Precision Medicine Committee and Society of Medical Biology Immune, and Chinese Pharmacological Society Chemotherapy Committee. He also serves as the Vice-Chairman of Liaoning Province Health Information Society.

Xiangdong Wang, MD, PhD, is a Distinguished Professor of Medicine, Director of Shanghai Institute of Clinical Bioinformatics, Executive Director of Clinical Science Institute of Fudan University Zhongshan Hospital, Director of Fudan University Center of Clinical Bioinformatics, Deputy Director of Shanghai Respiratory Research Institute, and visiting professor of King's College of London. His main research is focused on clinical bioinformatics, disease-specific biomarkers, lung chronic diseases, cancer immunology, and molecular & cellular therapies. He is the author of more than 200 scientific publications with the impact factor about 900, citation number about 6918, h-index 46, i10-index 181, and cited journal impact factor about 7000.