

## REVIEW

# Advances in single-cell sequencing and its application to musculoskeletal system research

Yongxiang Zhang<sup>1,2</sup> | Jingkai Wang<sup>1,2</sup> | Chao Yu<sup>1,2</sup> | Kaishun Xia<sup>1,2</sup> | Biao Yang<sup>1,2</sup> |  
 Yuang Zhang<sup>1,2</sup> | Liwei Ying<sup>1,2</sup> | Chenggui Wang<sup>1,2</sup> | Xianpeng Huang<sup>1,2</sup> | Qixin Chen<sup>1,2</sup> |  
 Li Shen<sup>1,3,4</sup>  | Fangcai Li<sup>1,2</sup> | Chengzhen Liang<sup>1,2</sup> 

<sup>1</sup>Department of Orthopedics Surgery, The Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, Zhejiang, China

<sup>2</sup>Zhejiang Key Laboratory of Bone and Joint Precision and Department of Orthopedics, Research Institute of Zhejiang University, Hangzhou, Zhejiang, China

<sup>3</sup>The MOE Key Laboratory of Biosystems Homeostasis & Protection and Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Life Sciences Institute, Zhejiang University, Hangzhou, China

<sup>4</sup>Hangzhou Innovation Center, Zhejiang University, Hangzhou, China

## Correspondence

Li Shen, Fangcai Li, and Chengzhen Liang,  
 Department of Orthopedics Surgery,  
 The Second Affiliated Hospital, Zhejiang  
 University School of Medicine, Hangzhou,  
 Zhejiang, China.

Email: li\_shen@zju.edu.cn (L. S.);

lifangcai@zju.edu.cn (F. L.);

liangchengzhen@zju.edu.cn (C. L.)

## Funding information

Scientific Research Fund of Zhejiang  
 Provincial Education Department, Grant/  
 Award Number: Y201941476; Nature  
 Science Foundation of Zhejiang Province,  
 Grant/Award Number: LQ18H060003,  
 LR18E030002, LY18H060004,  
 LY19H060005 and Y20H060063; Medical  
 and Health Innovation Talent Support  
 Program of Zhejiang Province, Grant/Award  
 Number: 2020RC011; National Natural  
 Science Foundation of China, Grant/Award  
 Number: 21774109, 51522304, 51973188,  
 81772379, 81902238, 81972096 and  
 82072465; Zhejiang University Education  
 Foundation Global Partnership Fund,  
 a project supported by the Scientific  
 Research Fund of Zhejiang Provincial  
 Education Department, Grant/Award  
 Number: Y201941476 and Y201941491;  
 Health Foundation of Zhejiang Province,  
 Grant/Award Number: 2018KY092 and  
 WKJ-ZJ-1903; China Postdoctoral Science  
 Foundation, Grant/Award Number:  
 2017M612011

## Abstract

In recent years, single-cell sequencing (SCS) technologies have continued to advance with improved operating procedures and reduced cost, leading to increasing practical adoption among researchers. These emerging technologies have superior abilities to analyse cell heterogeneity at a single-cell level, which have elevated multi-omics research to a higher level. In some fields of research, application of SCS has enabled many valuable discoveries, and musculoskeletal system offers typical examples. This article reviews some major scientific issues and recent advances in musculoskeletal system. In addition, combined with SCS technologies, the research of cell or tissue heterogeneity in limb development and various musculoskeletal system clinical diseases also provides new possibilities for treatment strategies. Finally, this article discusses the challenges and future development potential of SCS and recommends the direction of future applications of SCS to musculoskeletal medicine.

Chengzhen Liang, Fangcai Li and Li Shen are co-communication authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Cell Proliferation* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

The basic unit of an organism is cell, and multicellular lives in nature begin with a single cell. Although preliminary estimates suggest that every person is composed of at least 37.2 trillion cells,<sup>1</sup> a deeper understanding is still very limited concerning the functions of those cells. Thus, the Human Genome Project (HGP) was officially launched in 1990 to unravel the genetic code of all of the approximately 25,000 genes in the human body and map the human genome atlas.<sup>2</sup> In 2001, the publication of a working draft of the human genome was considered a milestone in HGP.<sup>3</sup> The Human Cell Atlas (HCA) was launched in 2007 to further describe all human cells.<sup>4</sup> HCA integrated the information of cell types, number, location, relationship and molecular composition that facilitates to describe the cellular basis of health and disease. Gene sequencing is considered as a highly reliable method for analysing cell genetic attribution.<sup>5</sup> The earliest method of gene sequencing can be traced back to in 1977,<sup>6</sup> where 'Sanger sequencing' was a revolutionary technology, which ushered the era of gene sequencing research. Over the last several decades, due to its accuracy, 'Sanger sequencing' has been widely incorporated to scientific investigations such as HGP and diagnosis of clinical genetic diseases.<sup>7</sup> However, because of its complexity and high cost, investigators have been trying to develop more efficient sequencing methods.

Next-generation sequencing (NGS) is the second revolutionary innovation of traditional gene sequencing.<sup>8</sup> NGS has the characteristics of high throughput and low cost of per base, also known as massively parallel sequencing (MPS).<sup>9</sup> Further advances in computer science and technology have enabled development of third-generation gene sequencing, with the ability of high-throughput, single-molecule sequencing. It can produce genome assemblies of unprecedented quality.<sup>7</sup> In 2009, Tang first reported a method for investigating the mRNA transcriptome using high-throughput sequencing in a single cell. This achievement is considered the beginning of widespread application of SCS technologies in scientific research.<sup>10</sup> Thanks to the application of SCS, cell biology and molecular biology had made great discovery in recent years. In 2018, researchers used SCS to create dynamic maps of gene expression during early embryonic development of zebrafish and frogs. Through integrating data on time scales in minutes to hours, describing the cells one by one, and tracking the eventual formation of embryo, investigators were able to build a complete map that revealed the entire developmental process from a single cell to an entire organism.<sup>11,12</sup> SCS can deepen our understanding of various aspects of cell function, such as tumorigenesis,<sup>13</sup> nerve degeneration,<sup>14</sup> immunology,<sup>15</sup> cell differentiation<sup>16</sup> and gene expression.<sup>17</sup>

Limb development is a fundamental event in musculoskeletal system. Researches reveal that mesenchymal stem cells (MSCs) can act as bone progenitor cells in bone marrow. Skeleton development begin with the migration of mesenchymal cells derived from the embryonic lineage to the sites of future bone.<sup>18</sup> Further, researches reveal that transcription factor SRY (sex determining

region Y)-box 9 (Sox9) plays a critical role in inducing osteogenic differentiation of MSCs.<sup>19-21</sup> Moreover, many morphogenetic or growth factors, such as WNTs, Hedgehogs, Notch, VEGF, FGFs, IGF-1, TGF- $\beta$  and PTHRP, have been found to be involved in the regulation of endochondral bone formation.<sup>22</sup> For musculoskeletal disease research, trauma, pain and limb malformation are main issues. More specifically, bone fracture, joint injury, osteoarthritis and intervertebral disc degeneration are the common clinical diseases. The analysis of cell heterogeneity and activation of specific types of stem cells have become hot issues in this field.<sup>23</sup> Stem cells are found in the bone marrow and periosteum. The stem cell population is made up of heterogeneous cells. Previous studies have focused on histomorphology and gene expression regulation, which skipped the cellular level. Some important information is hidden in heterogeneous cell populations. By combining SCS and lineage tracing, a unique subpopulation of periosteum stem cell is identified to contribute to bone regeneration.<sup>24</sup> More importantly, the contribution of periosteal stem cells to bone regeneration is higher than that of bone marrow stem cells.

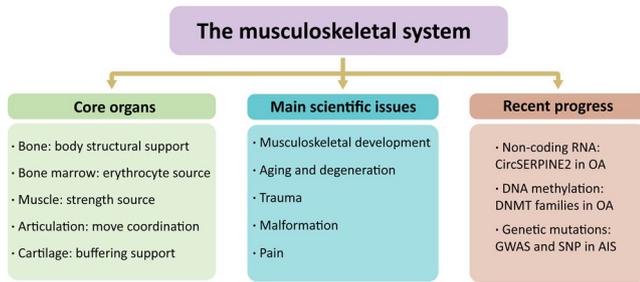
In this review, we first outline the major scientific issues and recent advances on the musculoskeletal system. Second, we discuss single-cell technologies in detail, including its technical features, superiority and its application in multi-omics studies. From our perspectives, cell heterogeneity has become the focus in limb development and clinical diseases of musculoskeletal system, and the application of SCS will provide us with unprecedented understanding of these issues. Moreover, the challenges and possible future directions of single-cell technologies are also discussed.

## 2 | THE MUSCULOSKELETAL SYSTEM

### 2.1 | Overview

Musculoskeletal system consists of bone, muscle, articulation, cartilage and other connective tissue that stabilize or connect bones (Figure 1).<sup>25</sup> In addition to supporting the body's weight, bone and muscle work together to keep the body in position or to produce controlled and precise movements.<sup>26</sup>

Bones provide structural support for the entire body and also protect the internal organs. Red bone marrow in long or flat bones lacunae produces blood cells. There are many attachments of muscle to the bones, which work as levers to change the magnitude and direction of the strength produced by the muscles.<sup>27</sup> Articulations play an important role in movement coordination. The stability and range of different articulations movement in the human body vary greatly, depending on the capacity of the associated muscles, tendons and ligaments. Cartilage is a kind of supportive buffering connective tissue.<sup>28</sup> Depending on its intercellular attribution, cartilage can be divided into three types: hyaline cartilage, elastic cartilage and fibrous cartilage. These three types of cartilage have different properties and functions in different parts of the body.



**FIGURE 1** An overview of the research issues of the musculoskeletal system

## 2.2 | Main scientific issues

The study of the musculoskeletal system is mainly divided into musculoskeletal development and clinical diseases. Chronic pain is the most common symptoms in orthopaedic clinics. In addition to injury-induced pain, degeneration and ageing of bones, cartilage and intervertebral discs (IVDs) are the main causes of pain. In the general population, bone ageing results in loss of volume and mass, usually manifested as osteoporosis and an increased risk of fracture.<sup>29,30</sup> Cartilage degeneration is common in joints. Degeneration of chondrocytes in the knee joint is a typical example. Ageing decreases cartilage thickness, which will lead to osteoarthritis (OA). Loss of IVD structural integrity can result in loss of IVD height, leading to collapse and compression of the spine and clinical symptoms, such as lower back pain.

Bone marrow MSCs can be stimulated to form osteoblasts, myocytes and fibroblasts. Change in the number or activity of these cells will affect a range of musculoskeletal tissues. Older MSCs exhibit a state of irreversible growth stagnation or senescence. Expression of p53, p21 and ageing-related  $\beta$ -galactosidase was increased compared to that in young MSCs, perhaps due to the down-regulation of age-related osteogenic genes, including Runx2 and osteocalcin.<sup>29</sup> Age-related changes in sex hormones can also affect normal bone biology. For example, postmenopausal women have gradually decreased oestrogen, while osteoclasts being released in response to oestrogen inhibition, increasing overall bone resorption.<sup>31</sup> During ageing, the molecular structure of articular cartilage changes in the same time. The most significant changes are reflected in decreased proteoglycans, aggrecan and water content of chondroitin sulphate (CS), crushed core protein well as increased keratan sulphate concentrations. Similarly, ageing IVDs show increased expression of matrix metalloproteinases (including MMP3 and MMP7) and disintegrin. These enzymes degrade collagen and prostaglandins, leading to the dehydration and disintegration of the extracellular matrix (ECM). Under these conditions, the ECM becomes granular, cracked or torn. The structural strength of the annulus fibrosus (AF) decrease makes it easy for the nucleus pulposus (NP) to excrete.<sup>32,33</sup>

In addition, malformation is a common concern of musculoskeletal system. The representative disease is scoliosis, among which adolescent idiopathic scoliosis (AIS) is the most common clinical type, occurring in 0.5%–3.0% of children.<sup>34</sup> It is currently believed that AIS

is a multi-factor disease with genetic predisposition.<sup>35</sup> However, the chromatin-level pathogenesis of scoliosis remains uncertain.

## 2.3 | Recent progress

High-throughput sequencing technologies have been widely used in musculoskeletal system. These applications have led to many discoveries in the regulation of gene diagnosis and non-coding genes in disease. For example, a variety of factors involve and regulate the process of OA. Recently, a novel CircRNA (CircSERPINE2) is identified in an RNA-seq comparing total RNA in several groups of clinical OA with that of normal control tissues. Transfection experiments suggest that CircSERPINE2 could reduce chondrocyte apoptosis and promote ECM synthesis. Further studies found that CircSERPINE2 plays a role in OA by targeting miR-1271 expression. The presumed target gene of miR-1271 is ERG [E26 transference specific (ETS)-related gene]. The CircSERPINE2-miR-1271-ERG axis is recognized as a new target for OA prevention and treatment.<sup>36</sup> Likewise, OA is a complete joint disease, and treatment strategies need to be able to target a wide range of relevant cell signalling pathways in cartilage synovium and bone. Studies over the past decade have revealed a large number of high-risk genetic loci for OA, which are predicted to increase disease risk by modulating the expression of target genes. Many risk loci are associated with epigenetic mediators.<sup>37</sup> Specifically, epigenetic factors can regulate gene expression by influencing chromatin status, and the regulation is accompanied by changes of multiple signal transduction pathways. For example, DNA methylation is catalysed by DNA methyltransferase families (DNMT1, DNMT3A and DNMT3B). A study found DNMT3A expression was up-regulated in a subset of OA patients ( $n = 71$ ) compared with the normal control group ( $n = 32$ ), and DNMT3A knockout was observed to reduce the catabolism of interleukin-1 $\beta$  (IL1- $\beta$ ) in the ECM.<sup>38</sup> However, DNMT3B had the opposite effect. It inhibits 4-aminobutyrate amino transferase (ABAT) degradation, while increased ABAT expression leads to enhance catabolic activity. Enzymes that undergo epigenetic changes can have a wide range of simultaneous effects in OA.<sup>39</sup> These effects cannot be achieved by targeting a single pathway.

In recent years, researchers have attempted to study the genetics of spinal malformations to explore their pathogenesis. Wu et al used comparative genomic hybridization to analyse congenital scoliosis in Han populations for the first time.<sup>40</sup> These investigators found large DNA deletions in the 16P11.2 region of the genome of patients with sporadic congenital scoliosis, and sequencing analysis confirmed that TBX6 gene in this deletion region was the source of this pathology. Moreover, the rare combination of null mutations and subtype alleles in TBX6 accounts for 11% of cases of congenital scoliosis in Han populations. This finding provides a theoretical basis for early diagnosis and genetic counselling of high-risk population. Moreover, Zhu et al used genome-wide association analysis (GWAS) to locate loci were 1p36.32, 2q36.1, 18q21.33 and 10q24.32 associated with AIS in Han Chinese girls.<sup>41</sup> The discovery that these genes

are closely related to bone growth and osteoblastic differentiation provides new insights into the genetic causes of AIS.

### 3 | SINGLE-CELL SEQUENCING TECHNOLOGIES

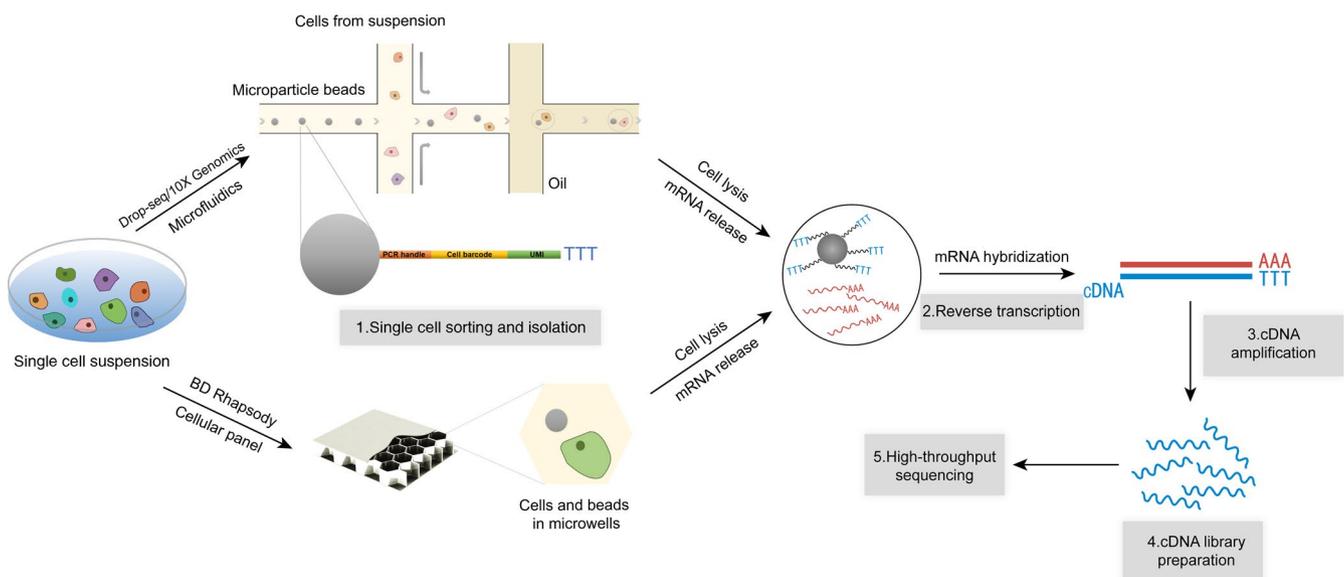
#### 3.1 | SCS can analyse cell heterogeneity, even in small samples

SCS technologies have obvious superiorities. Generally, most sequencing requires a sample size of at least one milligram, but the amount of DNA or RNA in a single cell is very small. For example, a typical cancer cell contains 6–12 pg of DNA and 10–50 pg of RNA (1%–5% mRNA).<sup>42</sup> In the past, genetic material from over ten thousand cells had to be extracted simultaneously to meet the sample size required for sequencing analysis. Yet, random expression of genes coding for proteins and metabolites causes cell heterogeneity, and genetic information in cells with the same phenotype may differ significantly. Simultaneous analysis of multiple cells will equalize cell signals and obscure cell heterogeneity.<sup>43</sup> In assessing the genetic clonal structure of tumours at the genome-wide level, the heterogeneity of tumour cells makes batch sequencing difficult and inaccurate. SCS is revolutionary because it can quantify all transcriptomes expressed in a single cell. High-resolution analysis can reveal intercellular differences that would normally be masked by batch sampling methods, providing a completely unbiased strategy for identifying and characterizing different cell populations. SCS also enables reliable transcriptome analysis at

the level of single cell. Some research focuses on limited number of cells, such as unicellular microorganisms or rare cells; in these cases, genetic information is especially difficult to obtain by conventional high-throughput sequencing. Yet, SCS can make these studies possible.<sup>44,45</sup>

#### 3.2 | Single-cell technologies and procedures

Single-cell transcriptome sequencing (scRNA-seq) is most widely used in the field of SCS technologies, generally including the following five steps: (1) Single-cell isolation, (2) Reverse transcription, (3) cDNA amplification, (4) Library preparation and (5) High-throughput sequencing.<sup>46</sup> A schematic diagram of the general process is shown in Figure 2. Isolation of a single cell is the most critical step in which the target cells must be accurately located and isolated from the samples. The initial two methods for single-cell separation, continuous dilution and micromanipulation, both had the disadvantages of low throughput and cell susceptibility to mechanical damage. More recent cell separation methods have included magnetic activated cell sorting (MACS),<sup>47</sup> fluorescence activated cell sorting (FACS)<sup>48</sup> and microfluidics.<sup>49</sup> Notably, microfluidics involves fewer step and has the advantages of high analytical sensitivity and specificity as well as high throughput. Drop-seq is a representative method to capture single cell by microfluidics.<sup>50</sup> This method uses barcode beads (Figure 2). The oligonucleotides on the beads include handle primer for PCR amplification, cellular barcode UMI (unique molecular identifier) that recognizes all oligonucleotides in a single cell and oligonucleotides (Oligo dT) that capture single-cell mRNA molecules.



**FIGURE 2** Basic procedures of single-cell transcriptome sequencing (scRNA-seq) and three common single-cell sorting platform. ScRNA-seq generally involves five basic procedures: single-cell isolation, reverse transcription, cDNA amplification, library preparation and high-throughput sequencing. Drop-seq, 10X Genomics and BD Rhapsody are the most widely used single-cell sorting platforms. They use microparticle beads with sequences linked by different cell barcodes, UMI and Oligo dT. These beads can capture mRNAs with poly A tails for reverse transcription and cDNA library preparation. Finally, all transcriptome information of a single cell is obtained by high-throughput sequencing

Based on Drop-seq, 10X Genomics was developed as a widely used SCS library preparation platform. 10X Genomics is similar to Drop-Seq method of capturing single cells, but it is integrated into an oil droplet-based microfluidic device that effectively captures a larger cell count with higher sensitivity.<sup>51</sup> In addition, BD Rhapsody is the another widely used SCS platform. Instead of microfluidic channel emits cells and collides with the outgoing magnetic beads, BD Rhapsody uses CytoSeq unique cellular panel.<sup>52</sup> First, the cell suspension is injected through the injection hole. It naturally subsides into the reaction hole. Then, beads are inserted through the injection hole in the same way to capture the cells in a single reaction hole. This method allows cells to be physically isolated from each other. Furthermore, in 2017, Chen et al. developed a new technology for single-cell space transcriptome sequencing: Geo-seq.<sup>53</sup> This method retains the original spatial information in single cell and elucidates the heterogeneity and spatial variation of cells. This approach has great potential in stem cell and embryonic development research. Recently, Georg Seelig's team developed a method called SPLiT-seq.<sup>54</sup> The cell acts as a reaction chamber, immobilizing the cell or nucleus, allowing efficient sample reuse. This is a low-cost method that does not require special customized instruments and enables the use of single-cell technology in a greater range of laboratory research.

### 3.3 | Single-cell multi-omics

The emergence of single-cell omics has provided unprecedented insight into embryonic development and the development of disease. Genomics, transcriptomics and proteomics are three important aspects of single-cell omics technologies, which allow us to understand the characteristic of the transmission and expression of genetic information with higher resolution and accuracy. Single-cell whole-genome sequencing is an effective method for screening single-cell SNP (single nucleotide polymorphism) and CNV (copy number variation). It is of great significance for the detection or prediction of tumour-driven mutations because it can evaluate the overall quality of DNA and annotate the variation.<sup>55</sup> For instance, hepatocellular carcinoma (HCC) is considered to be a cancer of genetic and phenotypic diversity. Duan et al profiled 96 tumour cells and 15 normal liver cells using single-cell whole-genome sequencing.<sup>56</sup> They found CNV occurred in the early stages of HCC and remained relatively stable during tumour progression. Single-cell transcriptome sequencing is a technology aiming to study of cell heterogeneity based on bulk transcriptome sequencing.<sup>57,58</sup> However, if the cell population is analysed only from a single-omics approach at a time, only local conditions in the gene regulatory network can be detected. The overall landscape cannot be accurately provided. Single-cell multi-omics is developed by combining various single-cell single-omics technologies. Genetic information, chromatin state, RNA expression and protein abundance can be analysed simultaneously to help us further understand the regulatory mechanism of gene expression.<sup>59,60</sup>

Researchers have developed a variety of single-cell multi-omics technologies. Specifically, DNA-mRNA sequencing (DR-seq) a sequencing method, lyses a single-cell and simultaneously amplifies the DNA and RNA.<sup>61</sup> The suspension is divided into two samples, one for RNA sequencing and the other for genomic DNA sequencing. This method keeps DNA and RNA amplifying together, meanwhile minimizing nucleic acid loss. Genome and transcriptome sequencing (G&T-seq) is another important method.<sup>62</sup> In G&T-seq, mRNA and DNA are physically separated from fully lysed cells using magnetic beads coated with short oligonucleotide sequences that bind mRNA. DNA and mRNA are then amplified and sequenced. This method keeps mRNA and DNA separate, allowing researchers to analyse each molecule using their chosen scheme. However, this approach can lead to loss of nucleic acids. DR-seq and G&T-seq are lysing whole cell, while simultaneous isolation of genomic DNA and total RNA (SIDR) was developed as a cell membrane lysis method.<sup>63</sup> The method captures and labels cells with magnetic beads that bind to specific antibodies. A hypotonic solution is used to destroy the plasma membrane and release total RNA, leaving the nuclear fibrous layer intact and genomic DNA (gDNA) remaining in the nucleus. The lysis suspension is then placed on a magnetic frame, where the supernatant is total RNA and the precipitate was gDNA. They can physically isolate of total RNA in gDNA, including mRNA and non-coding RNA (especially long non-coding RNA, lncRNA). In addition, SIDR has a more accurate comparison ratio than DR-Seq, and this method can be applied to more comprehensive studies of cell heterogeneity and complexity.

Notably, single-cell multi-omics has a significant application prospect. First, cell subtypes can be identified from heterogeneous cell populations. Parallel measurements of different omics can provide non-overlapping information about the identity of the cell, enabling more detailed and accurate classification of the cell population. Second, tracking the cell lineages throughout the development of multicellular animals is one of the most important aspects of developmental biology. Lineages can be traced through mutations or epigenetic modifications of DNA during cell division.<sup>64</sup> Matching single-cell transcriptome data can also reveal the changes in gene expression and the fate of cells during proliferation and differentiation.

### 3.4 | Single-cell ATAC-seq and epigenetics

Nucleosomes are the basic structural units of chromatin in eukaryotes. DNA binds to histones to form nucleosomes, which are further folded and compressed to form chromatin. Both replication and transcription of DNA require the opening of tightly packed chromatin structures to allow regulatory molecules to bind to DNA. The opened section of the chromatin is called accessible chromatin. The search for accessible chromatin regions is the most critical step in the study of transcriptional regulation. William et al used DNA transposase in combination with high-throughput sequencing to study chromatin accessibility, an approach known as ATAC-seq (Assay for Transposase Accessible Chromatin with high-throughput

sequencing).<sup>65</sup> ATAC-seq is a genome-wide measure of chromatin accessibility, which can obtain genome-wide protein binding site information. ATAC-seq can be used to screen for unknown specific transcription factors.

Single-cell ATAC-seq (scATAC-seq) is a single-cell resolution method for the study of chromatin accessibility. Specifically, sequences in open chromatin regions contain a large amount of motifs information about transcriptional initiation. Motif information can be found in the database for corresponding regulatory transcription factors, thus providing information for downstream experimental design. Frisé et al selected populations of resident non-nerve cells in the mice spinal cord for scRNA-seq and scATAC-seq.<sup>66</sup> They observed that the motifs of typical oligodendrocyte lineage transcription factors OLIG2 and SOX10 were highly available not only in oligodendrocyte progenitor cells (OPC) but also in ependymal cell populations.<sup>67</sup> However, OLIG2 and SOX10 are expressed in oligodendrocytes but not in ependymal cells. Their subsequent experiments confirmed that OLIG2 expression in ependymal cells activates a potential oligodendrocyte lineage program after spine cord injury (SCI). ScATAC-seq revealed that specific transcription factor OLIG2 was the core basis of Frisé's study. Because ATAC-seq is more repeatable and easier to operate than traditional methods, it has become the preferred method to study chromatin accessibility.

Environmental conditions affect the metabolism of cells and organisms. With changing conditions, the organism tries to adjust its metabolism in response to the environment. Gene mutations may also change biological differentiation pathways.<sup>68</sup> Epigenetic research explores the complex interaction between environmental conditions and genetic factors to illuminate the pathogenesis of disease.<sup>69</sup> Extensive evidence supports the view that epigenetic mechanisms (DNA methylation, post-translational modifications of histone tails and non-coding RNAs) are involved in the differentiation of bone cells.<sup>70</sup> Currently, epigenetic studies are mainly conducted at the bulk level. Different cell types with distinct epigenetic characteristics are often mixed together in tissues or organs, thus presenting a challenge for research.

NGS has greatly facilitated the study of epigenetics by enabling the study of DNA methylation at single-base resolution,<sup>71</sup> thus elucidating methylation patterns in the entire genome.<sup>72</sup> Wang et al developed CoBatch,<sup>73</sup> a single-cell ChIP-Seq technology that is universal, easy to operate and of high quality. The method used the fusion Protein PAT (Protein A-Tn5) to recognize and cut specific genomic regions to which antibodies bind, and combined with the barcode label to achieve high-throughput single-cell capture. Single-cell techniques can reveal the epigenetic heterogeneity of cells and describe the DNA methylation characteristic of cell subsets.<sup>74</sup> In one study, Angermueller et al observed methylation heterogeneity in two enhancers of mouse ESCs, *Esrrb1*, and found a negative correlation between methylation level and gene expression.<sup>75</sup> As mentioned above, of SCS technologies have a superior ability to analyse rare cells. One research focused targeted cells that develop in early mammalian embryos, including especially early human embryonic cells. The epigenome of these cells changes rapidly and significantly

during development after fertilization.<sup>76,77</sup> Therefore, it is important to elucidate epigenetic status by using epigenomic mapping techniques in single-cell level.

## 4 | CELL HETEROGENEITY IN THE MUSCULOSKELETAL SYSTEM

### 4.1 | Limb development

Bone and muscle are the most principal organs of musculoskeletal system. In the past, researches on skeleton biology had met with several obstacles. The chemical methods used for treatment of highly cellular tissues such as brain, liver and kidney, may not be applicable to skeleton tissue,<sup>78</sup> because some solutions cannot penetrate the extracellular matrix to reach bone cells, especially in mineralized tissue.<sup>79</sup> This problem is recognized in skeleton biology research, and a variety of tissue-specific RNA separation methods have been developed for bone,<sup>80,81</sup> cartilage,<sup>82</sup> ligament and tendon.<sup>83</sup> Exploring the potential stem cell properties of the periosteum, Debnath et al found that a number of groups of mesenchymal cells are identified with expressed Ctsk (cathepsin K) in the periosteum of long bones.<sup>84</sup> ScRNA-seq analysis was performed to divide the cell pool into four groups, one of which expressed progenitor cell/stem cell markers such as *Sox9* and *Col2a1*. Their findings reveal a type of periosteum stem cell (PSC) exists in long bones and skulls of mice. PSC shows clone pluripotency and self-renewal abilities, and it locates at the top of the differentiation hierarchy. Further research found that the human periosteum also contains a cell population similar to PSC. Moreover, muscle stem cell (MuSC) is required for skeletal muscle development, growth and regeneration.<sup>85</sup> Stefania et al identified the discrete transcriptional programs of homeostatic MuSC, injured MuSC and primary myoblasts (PMs) by scRNA-seq. A pseudotime dynamic map from MuSC initiation to PMs was delineated.<sup>86</sup> Zi et al used scRNA-seq to study the dynamic transcriptional characteristic of cells at different time points (E10.5, E12.5 and E15.5) during the process from limb bud initiation to the basic formation of musculoskeletal system tissues in mouse.<sup>87</sup> Pseudotime analysis showed that *SCX*<sup>+</sup> *HOXD13*<sup>+</sup> musculoskeletal stem progenitor clusters exists in the early stage of limb development. These clusters form distinct branching pathways corresponding to hard and soft connective tissue differentiation during late development. Similarly, Chan et al combined scRNA-seq data of mouse skeletal stem cell (mSSC) and human skeletal stem cells (hSSC) to identify hSSC characteristic, meanwhile revealed species conservation in skeletogenesis.<sup>88</sup> Furthermore, He et al discovered a self-renew cell cluster, embryonic skeletal stem/progenitor cell (eSSPC). This cluster can generate osteochondral lineage cells.<sup>89</sup> In summary, the ability of scRNA-seq to analyse a single cell significantly promotes the exploration of bone development and reveals the complexity of bone cells.

Several additional studies have focused on whole limbs. Because tetrapods are good models for studying the genetic and molecular

basis of vertebrate pattern formation, limb development in tetrapods has received extensive research attention.<sup>90</sup> The genome determines development of the limb.<sup>91</sup> Feregrino et al used scRNA-seq to profile transcription patterns in developing chicken autopods.<sup>92</sup> 17,628 cells were sequenced from three key stages of development in autopods of the chicken, with 1000 genes detected in each cells. Subsequently, 23 cell populations are identified by different transcriptomics patterns. Similarly, Natalie et al study used scRNA-seq to describe the transcription changes of embryonic hind limb development in mice.<sup>93</sup> They detected cell heterogeneity, even at the earliest point of development when limb buds could be artificially isolated. The identified cell clusters included the known cell types involved in the development of the hind limb, bone, cartilage and muscle. In addition, to study the limb regeneration mechanisms in salamanders, Gerber et al performed scRNA-seq in connective tissue cells during bud formation, arm growth and embryonic limb development.<sup>94</sup> These researchers observed that heterogeneous groups of connective tissue cells converge into a uniform and transient bud-based progenitor cell state, which is similar to the embryonic limb budding process in later stage. In addition, Qin et al isolated blastema tissues of salamander forelimbs at 3 days, 7 days and 21 days after treated as well as normal forelimbs of salamanders, then conducted scRNA-seq on 938 extracted cells.<sup>95</sup> They found a cluster of regenerative cells that is characterized by a significantly higher number of mitochondria than normal limb tissue cells. A novel COL2-mito subcluster is further defined as COL2<sup>+</sup> cells, which perform energy metabolite-related functions, such as responded to oxygen levels and ATP metabolism.

Collectively, these findings demonstrate the ability of scRNA-seq to isolate populations of developing limb cells at the molecular level. The discoveries of the trajectory of limb development also provide new insights into tissue regeneration in the musculoskeletal system.

## 4.2 | Osteoarthritis

OA is a chronic degenerative disease closely related to ageing and progressive joint dysfunction.<sup>96</sup> The main pathological change of OA is a disorder of articular cartilage homeostasis, accompanied by inflammation and degradation.<sup>97</sup> Understanding the role and degeneration mechanism of chondrocytes will be helpful to development of an innovative treatment strategy for OA. Ji et al performed scRNA-seq on 1464 chondrocytes selected from 10 OA patients after knee arthroplasty.<sup>98</sup> They found that the transcription factors SOX4, TRPS1 and EGR2 are strongly expressed in the early stages of OA and that DNAJC2, GZF1 and ETS2 are strongly expressed in the late stages of OA. Thus, the entire transcription process of degeneration of OA cartilage is described at the single-cell level, highlighting the potential of cell transcriptomes to diagnose and predict the results of OA cartilage degeneration. Thanks to the unique advantages of SCS in analysing cell heterogeneity, Ji et al found seven potential chondrocyte clusters, including four classical cell clusters and three new clusters. Further research revealed that these chondrocyte

clusters played protective or exacerbation role in the OA progression respectively.<sup>99</sup> Wang et al used scRNA-seq to identify chondrocyte populations and genetic characteristics of Kashin-Beck disease, OA and healthy chondrocytes.<sup>100</sup> Ten cell clusters are annotated by cell markers. A new subcluster, MTCs (mitochondrial chondrocytes), is identified based on the expression of a new set of biomarkers. These findings deepen our knowledge of the functions of OA chondrocytes. More importantly, they also offer new possibilities for treatment strategies of OA.

## 4.3 | Intervertebral disc degeneration

Many people suffer from neck pain and lower back pain, which may lead to loss of normal work ability and greatly reduce the quality of life. The main cause of thoracic lumbar pain is intervertebral disc degeneration (IVDD). In recent years, IVDD has become one of the hottest diseases in cell regeneration medicine research.<sup>101</sup> However, due to the lack of understanding of the molecular characteristic of healthy mature IVD cells, cell regeneration therapies have encountered some obstacles. Li et al chose mature bovine tailbone IVD as the study model because it appeared to be similar in anatomy, histology and biochemistry to human lumbar IVD from healthy young adults aged 15–40 years.<sup>102,103</sup> The researchers used RISH (RNA in situ hybridization) in bovine IVD to analyse transcriptional target genes in heterogeneous cell populations. They found two new biomarkers for AF (Lam1 and Thy1) and eight new NP biomarkers (Gli1, Gli3, Noto, Scx, Ptprc, Sox2, Zscan10, Loc101904175).<sup>104</sup> Building on the above result, they used single-cell resolution FL-RISH (fluorescent RNA in situ hybridization) to further evaluate and quantify the new biomarkers identified by colour rendering AP-RISH and z-scale analysis. Fernandes et al isolated NP and AF cells from healthy human IVD.<sup>105</sup> They found that FOXM1 and KDM4E transcription factors are key regulators of AF and NP gene networks. Previous research in mouse models had shown that notochord cells contribute to NP. However, retention of these cells in mature humans or bovine NP differs from that in mice or other rodents.<sup>106–108</sup> The difference indicates interspecies heterogeneity of IVD cells. Recently, Gan et al performed scRNA-seq of 108,108 NP, AF and cartilage endplate (CEP) cells in IVD, providing the most comprehensive systematic single-cell map of human IVD.<sup>109</sup> Furthermore, nucleus pulposus progenitor cells (NPPCs) as a specific subset in NP were identified by in vitro functional experiments. These cells have the ability of colony formation and three-system differentiation. Obviously, the applications of single-cell technologies have provided a better understanding of the cellular heterogeneity of human IVDs.

## 4.4 | Rheumatoid arthritis

Rheumatoid arthritis (RA) is a systemic autoimmune disease mainly involving surrounding joints. RA is characterized by chronic synovial

inflammation of the joints, which can lead to erosion and destruction of articular cartilage and surrounding tissues, resulting in joint deformity, stiffness and dysfunction, and finally shortened life span.<sup>110</sup> Previous studies have demonstrated that the destructive inflammatory environment of the joint synovium results from the action of a variety of cell types, including synovial fibroblasts, macrophages, osteoclasts and vascular endothelial cells.<sup>111,112</sup> Stephenson et al developed a low-cost droplet microfluidics control instrument.<sup>113</sup> They extracted 20,387 cells from synovial tissues of 5 RA patients for single-cell droplet microfluidics. First, they recognized different lymphocyte populations (T, B and NK cells) in these samples. In the CD4<sup>+</sup> T-helper cell population, a unique subgroup of peripheral blood T-helper cells (TPH) was detected, marked by high levels of MAF, CXCL13 and PDCD1 (PD1). In addition, unbiased cluster analysis reveals three different fibroblast subsets. Two constituent fibroblasts show distinct bifurcation in marker expression (fibroblast 1 and fibroblast 2) and further subdivision of the latter (fibroblast 2a and fibroblast 2b). Fibroblast 1 cells are stained primarily in the synovial lining, while fibroblast 2 cells are marked in the underlying region. Hyaluronic acid synthetase 1 (HAS1) is highly expressed in the lining fibroblasts, suggesting that fibroblasts 1 are responsible for synovium production and turnover. Similarly, Zhang et al performed single-cell transcriptomics sequencing and flow cytometry analysis of T cells, B cells, monocytes and fibroblasts in synovial tissue from 36 RA patients and 15 OA patients.<sup>114</sup> The role of specific clusters of cells in RA and chronic inflammation was precisely defined by the integration of cross-data mode at the single-cell level. Monocyte samples RNA-seq data suggested that genes associated with SC-M1 (IL1B<sup>+</sup> proinflammatory monocytes) are significantly up-regulated in leukocyte-rich RA samples. In contrast, the marker genes associated with SC-M2 were down-regulated in OA. In other words, the leukocyte-rich RA synovium had greater numbers of IL1B<sup>+</sup> monocytes and IFN-activated monocytes but lower numbers NUPR1<sup>+</sup> monocytes than the OA synovium. These data suggest that cytokine activation drives the expansion of the unique mononuclear population in the synovial membrane of active RA.

## 5 | CHALLENGES AND FUTURE DIRECTIONS

SCS is an emerging technology with great promise, but also with challenges of its application to scientific research. First, the biggest issue is cost, which can easily cost more than \$1,000 per sample, including cell capture and library preparation. Second, extraction of single cells is the starting step of SCS. With stroma-rich bone tissue, it is a difficulty to isolate sufficient amounts and qualities of RNA,<sup>115</sup> particularly in small animals. Sometimes it is essential to isolate a sufficient number of living cells from the bone marrow or bone tissue, because these living cells accurately represent the cellular diversity of tissues in the body. For instance, multiple displacement amplification (MDA) is the preferred method for genomic DNA amplification, specifically

commonly used in low DNA quantities clinical samples.<sup>116</sup> However, MDA has the defects of amplification bias and unbalanced genome coverage.<sup>117</sup> In addition, the current data algorithms used in computer bioinformatics still need further development, and new methods are needed to standardize the identification of new cell types.<sup>118</sup> Professional analysis of the vast amount of sequence data is also a major challenge. Despite these challenges, SCS still has excellent potential. For example, single-cell technology can be used to perform multi-layer analyses of the tumour cell genome and transcriptome, which has the potential to revolutionize methods of understanding tumour growth, and this potential exists at every stage of disease development. As we have discussed in this article, SCS has unique superiority to analyse cell heterogeneity in many contexts.<sup>119</sup> From this perspective, SCS can be applied to musculoskeletal research to identify cell heterogeneity between stem cells populations,<sup>44</sup> which can help the design of methods to induce specific stem cells to play a role in tissue repair in musculoskeletal diseases.

## 6 | CONCLUDING REMARKS

In 2018, 'development cell by cell' was named 'the top one scientific breakthrough of the year' by *science*.<sup>120</sup> SCS is the core technology of the researches. Public recognition of this growing and innovative technology shows that SCS has stimulated great changes in scientific research. Over the past decade, investigators have improved or simplified several procedures resulting in drastic reduction in costs. Thanks to the unique advantages of SCS, it is proved to be a technology with broad applications, yielding valuable data on microorganisms, tumorigenesis, and brain or nervous system. As we have seen, this technology has also enabled great advances in the study of heterogeneity of musculoskeletal system (Table 1), and it is believed that the continued application of SCS technologies will achieve more breakthrough in this field. The application of single-cell technologies in OA has stimulated new potential treatment approaches to this common disease in clinical. Combined with description of developmental trajectory of embryonic limb, SCS technologies can facilitate new cell subtypes discovery and further illuminate the composition of the biological skeleton. A better understanding of bone fracture, ligament tear and post-traumatic osteoarthritis can lead to come up with innovative treatment strategies to repair bone and joint tissues. In the future, we expect to see more SCS application in limb regeneration, bone metabolism and tumorigenesis in musculoskeletal system.

## ACKNOWLEDGEMENTS

This study was supported by grants from the Nature Science Foundation of Zhejiang Province (Y20H060063, LY19H060005, LQ18H060003, LR18E030002, LY18H060004), the Medical and Health Innovation Talent Support Program of Zhejiang Province (2020RC011), the National Natural Science Foundation of China (NO. 82072465, NO. 81772379, NO. 81972096, NO. 81902238,

TABLE 1 Cell heterogeneity in the musculoskeletal system researches

Subject	Technology	Method	Result	Reference
Bone development	ScRNA-seq	CEL-Seq2	Identification of Ctsk <sup>+</sup> periosteal stem cell	Debnath et al. <sup>15</sup>
Muscle regeneration	ScRNA-seq	10X Genomics	Depiction of muscle stem cell homeostatic and regeneration map	Stefania et al. <sup>16</sup>
Limb development	ScRNA-seq	Fluidigm C1	Identification of SCX <sup>+</sup> HOXD13 <sup>+</sup> musculoskeletal stem/progenitor cell	Zi et al. <sup>17</sup>
Limb development	ScRNA-seq	Smart-seq2	Identification of the Human Skeletal Stem Cell	Chan et al. <sup>18</sup>
Limb regeneration	ScRNA-seq	10X Genomics	Depiction of chicken autopods developmental transcriptome atlas	Feregino et al. <sup>19</sup>
Limb development	ScRNA-seq	10X Genomics	Depiction of embryonic mice hind limb developmental transcriptome atlas	Natalie et al. <sup>20</sup>
Limb regeneration	ScRNA-seq	10X Genomics	Description of pluripotency in the connective tissue cell of salamander	Gerber et al. <sup>21</sup>
Limb development	ScRNA-seq	Fluidigm C1	Identification of energy-metabolite-related COL2-mito cells	Qin et al. <sup>22</sup>
Osteoarthritis	ScRNA-seq	STRT-Seq	Identification of three new clusters effector chondrocytes	Ji et al. <sup>23</sup>
Osteoarthritis	ScRNA-seq	10X Genomics	Identification of mitochondrial chondrocytes	Wang et al. <sup>24</sup>
Intervertebral disc degeneration	ScRNA-seq	Drop-seq	Description of interspecies heterogeneity of IVD cells	Fernandes et al. <sup>25</sup>
Intervertebral disc degeneration	ScRNA-seq	10X Genomics	Depiction of systematic single-cell map of human IVD	Gan et al. <sup>26</sup>
Rheumatoid arthritis	ScRNA-seq	Drop-seq (low-cost droplet microfluidic)	Identification of different lymphocyte populations and two constituent fibroblasts	Stephenson et al. <sup>27</sup>
Rheumatoid arthritis	ScRNA-seq	CEL-Seq2	Definition overabundant stromal and immune cell populations in RA	Zhang et al. <sup>28</sup>

NO. 21774109, NO. 51973188, NO. 51522304), the Health Foundation of Zhejiang Province (2018KY092, WKJ-ZJ-1903), the China Postdoctoral Science Foundation (2017M612011), the Zhejiang University Education Foundation Global Partnership Fund, a project supported by the Scientific Research Fund of Zhejiang Provincial Education Department (Y201941476 and Y201941491) and the Scientific Research Fund of Zhejiang Provincial Education Department (Y201941476).

#### CONFLICT OF INTEREST

The authors declare that they have no competing interests.

#### AUTHOR CONTRIBUTIONS

ZY wrote original manuscript. WJ, YC, XK, YB, ZY, YL, WC and HX reviewed and edited. CQ, SL, LF and LC involved in conceptualization, and reviewed and edited. All authors reviewed the final version of manuscript.

#### DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

#### ORCID

Li Shen  <https://orcid.org/0000-0002-5696-2191>

Chengzhen Liang  <https://orcid.org/0000-0002-2948-2141>

#### REFERENCES

- Bianconi E, Piovesan A, Facchin F, et al. An estimation of the number of cells in the human body. *Ann Hum Biol.* 2013;40(6):463-471.
- Engel LW. The human genome project - history, goals, and progress to date. *Arch Pathol Lab Med.* 1993;117(5):459-465.
- Dobson R. Working draft of the human genome completed. *Bull World Health Organ.* 2000;78(9):1168-1169.
- Regev A, Teichmann SA, Lander ES, et al. The human cell atlas. *eLife.* 2017;6:30.
- Wu X, Yang B, Udo-Inyang I, et al. Research techniques made simple: single-cell RNA sequencing and its applications in dermatology. *J Invest Dermatol.* 2018;138(5):1004-1009.
- Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA.* 1977;74(12):5463-5467.
- van Dijk EL, Jaszczyszyn Y, Naquin D, Thermes C. The third revolution in sequencing technology. *Trends Genet.* 2018;34(9):666-681.
- Gu W, Miller S, Chiu CY. Clinical metagenomic next-generation sequencing for pathogen detection. In: Abbas AK, Aster JC, Feany MB, eds. *Annual Review of Pathology: Mechanisms of Disease*, vol. 14. Annual Reviews; 2019:319-338.
- Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. *Cell.* 2013;155(1):27-38.
- Tang FC, Barbacioru C, Wang YZ, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods.* 2009; 6(5):377-U386.
- Farrell JA, Wang Y, Riesenfeld SJ, Shekhar K, Regev A, Schier AF. Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science.* 2018;360(6392):979.

12. Wagner DE, Weinreb C, Collins ZM, Briggs JA, Megason SG, Klein AM. Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science (New York, NY)*. 2018;360(6392):981-987.
13. Baslan T, Hicks J. Unravelling biology and shifting paradigms in cancer with single-cell sequencing. *Nat Rev Cancer*. 2017;17(9):557-569.
14. Ofengeim D, Giagtzoglou N, Huh D, Zou CY, Yuan JY. Single-cell RNA sequencing: unraveling the brain one cell at a time. *Trends Mol Med*. 2017;23(6):563-576.
15. Neu KE, Tang Q, Wilson PC, Khan AA. Single-cell genomics: approaches and utility in immunology. *Trends Immunol*. 2017;38(2):140-149.
16. Trapnell C, Cacchiarelli D, Grimsby J, et al. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014;32(4):381-251.
17. Potter SS. Single-cell RNA sequencing for the study of development, physiology and disease. *Nat Rev Nephrol*. 2018;14(8):479-492.
18. Hall BK, Miyake T. The membranous skeleton: the role of cell condensations in vertebrate skeletogenesis. *Anat Embryol*. 1992;186(2):107-124.
19. Mori-Akiyama Y, Akiyama H, Rowitch DH, de Crombrugge B. Sox9 is required for determination of the chondrogenic cell lineage in the cranial neural crest. *Proc Natl Acad Sci USA*. 2003;100(16):9360-9365.
20. Kawakami Y, Rodriguez-León J, Izpisua Belmonte JC. The role of TGFbetas and Sox9 during limb chondrogenesis. *Curr Opin Cell Biol*. 2006;18(6):723-729.
21. Akiyama H, Chaboissier MC, Martin JF, Schedl A, de Crombrugge B. The transcription factor Sox9 has essential roles in successive steps of the chondrocyte differentiation pathway and is required for expression of Sox5 and Sox6. *Genes Dev*. 2002;16(21):2813-2828.
22. Kronenberg HM. Developmental regulation of the growth plate. *Nature*. 2003;423(6937):332-336.
23. Loebel C, Burdick JA. Engineering stem and stromal cell therapies for musculoskeletal tissue repair. *Cell Stem Cell*. 2018;22(3):325-339.
24. Ortinau LC, Wang H, Lei K, et al. Identification of functionally distinct Mx1+ $\alpha$ SMA+ periosteal skeletal stem cells. *Cell Stem Cell*. 2019;25(6):784-796.e785.
25. Berta A, Sumich JL, Kovacs KM. Chapter 8 - musculoskeletal system and locomotion. In: Berta A, Sumich JL, Kovacs KM, eds. *Marine Mammals*, 3rd ed. Academic Press; 2015:211-268.
26. Sylos-Labini F, Zago M, Guertin PA, Lacquaniti F, Ivanenko YP. Muscle coordination and locomotion in humans. *Curr Pharm Des*. 2017;23(12):1821-1833.
27. Turner CH. Bone strength: current concepts. *Ann N Y Acad Sci*. 2006;1068:429-446.
28. Oryan A, Sahviah S. Effectiveness of chitosan scaffold in skin, bone and cartilage healing. *Int J Biol Macromol*. 2017;104(Pt A):1003-1011.
29. Roberts S, Colombier P, Sowman A, et al. Ageing in the musculoskeletal system. *Acta Orthop*. 2016;87(sup363):15-25.
30. Adams DJ, Rowe DW, Ackert-Bicknell CL. Genetics of aging bone. *Mamm Genome*. 2016;27(7-8):367-380.
31. Zou Z, Liu W, Cao L, et al. Advances in the occurrence and biotherapy of osteoporosis. *Biochem Soc Trans*. 2020;48(4):1623-1636.
32. Rustenburg CME, Emanuel KS, Peeters M, Lems WF, Vergroesen PA, Smit TH. Osteoarthritis and intervertebral disc degeneration: quite different, quite similar. *JOR Spine*. 2018;1(4):e1033.
33. Zhao CQ, Wang LM, Jiang LS, Dai LY. The cell biology of intervertebral disc aging and degeneration. *Ageing Res Rev*. 2007;6(3):247-261.
34. Altaf F, Gibson A, Dannawi Z, Noordeen H. Adolescent idiopathic scoliosis. *BMJ (Clin Res ed)*. 2013;346:f2508.
35. Newton Ede MM, Jones SW. Adolescent idiopathic scoliosis: evidence for intrinsic factors driving aetiology and progression. *Int Orthop*. 2016;40(10):2075-2080.
36. Shen S, Wu Y, Chen J, et al. CircSERPINE2 protects against osteoarthritis by targeting miR-1271 and ETS-related gene. *Ann Rheum Dis*. 2019;78(6):826-836.
37. Rice SJ, Beier F, Young DA, Loughlin J. Interplay between genetics and epigenetics in osteoarthritis. *Nat Rev Rheumatol*. 2020;16(5):268-281.
38. Rahmati M, Nalesso G, Mobasheri A, Mozafari M. Aging and osteoarthritis: central role of the extracellular matrix. *Ageing Res Rev*. 2017;40:20-30.
39. Grandi FC, Bhutani N. Epigenetic therapies for osteoarthritis. *Trends Pharmacol Sci*. 2020;41(8):557-569.
40. Wu N, Ming X, Xiao J, et al. TBX6 null variants and a common hypomorphic allele in congenital scoliosis. *N Engl J Med*. 2015;372(4):341-350.
41. Zhu Z, Tang NL, Xu L, et al. Genome-wide association study identifies new susceptibility loci for adolescent idiopathic scoliosis in Chinese girls. *Nat Commun*. 2015;6:8355.
42. Livesey FJ. Strategies for microarray analysis of limiting amounts of RNA. *Brief Funct Genom Proteom*. 2003;2(1):31-36.
43. Su D, Fang C. The application of single cell technologies in stem cell tissue repair and drug development. *Chin J Cell Biol*. 2019;41(01):121-131.
44. Kester L, van Oudenaarden A. Single-cell transcriptomics meets lineage tracing. *Cell Stem Cell*. 2018;23(2):166-179.
45. Woyke T, Doud DFR, Schulz F. The trajectory of microbial single-cell sequencing. *Nat Methods*. 2017;14(11):1045-1054.
46. Hedlund E, Deng Q. Single-cell RNA sequencing: technical advancements and biological applications. *Mol Aspects Med*. 2018;59:36-46.
47. Murakami A, Oshiro H, Kanzaki S, et al. A novel method for isolating podocytes using magnetic-activated cell sorting. *Nephrol Dial Transplant*. 2010;25(12):3884-3890.
48. Qin YL, Wu L, Wang JG, et al. A fluorescence-activated single-droplet dispenser for high accuracy single-droplet and single-cell sorting and dispensing. *Anal Chem*. 2019;91(10):6815-6819.
49. Martino C, deMello AJ. Droplet-based microfluidics for artificial cell generation: a brief review. *Interface Focus*. 2016;6(4):10.
50. Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell*. 2015;161(5):1202-1214.
51. Zhang X, Li T, Liu F, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-seq systems. *Mol Cell*. 2019;73(1):130-142.e135.
52. Fan HC, Fu GK, Fodor SP. Expression profiling. Combinatorial labeling of single cells for gene expression cytometry. *Science (New York, NY)*. 2015;347(6222):1258367.
53. Chen J, Suo S, Tam PPL, Han J-DJ, Peng G, Jing N. Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. *Nat Protoc*. 2017;12(3):566-580.
54. Rosenberg AB, Roco CM, Muscat RA, et al. Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*. 2018;360(6385):176-182.
55. Zhang L, Dong X, Lee M, Maslov AY, Wang T, Vijj J. Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. *Proc Natl Acad Sci USA*. 2019;116(18):9014-9019.
56. Duan M, Hao J, Cui S, et al. Diverse modes of clonal evolution in HBV-related hepatocellular carcinoma revealed by single-cell genome sequencing. *Cell Res*. 2018;28(3):359-373.
57. Papalexis E, Satija R. Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat Rev Immunol*. 2018;18(1):35-45.
58. Wen L, Tang F. Single-cell sequencing in stem cell biology. *Genome Biol*. 2016;17:71.

59. Peng A, Mao X, Zhong J, Fan S, Hu Y. Single-cell multi-omics and its prospective application in cancer biology. *Proteomics*. 2020:e1900271.
60. Bock C, Farlik M, Sheffield NC. Multi-omics of single cells: strategies and applications. *Trends Biotechnol*. 2016;34(8):605-608.
61. Dey SS, Kester L, Spanjaard B, Bienko M, van Oudenaarden A. Integrated genome and transcriptome sequencing of the same cell. *Nat Biotechnol*. 2015;33(3):285-289.
62. Macaulay IC, Haerty W, Kumar P, et al. G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods*. 2015;12(6):519-522.
63. Han KY, Kim KT, Joung JG, et al. SIDR: simultaneous isolation and parallel sequencing of genomic DNA and total RNA from single cells. *Genome Res*. 2018;28(1):75-87.
64. Hu Y, An Q, Sheu K, Trejo B, Fan S, Guo Y. Single cell multi-omics technology: methodology and application. *Front Cell Dev Biol*. 2018;6:28.
65. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013;10(12):1213-1218.
66. Llorens-Bobadilla E, Chell JM, Le Merre P, et al. A latent lineage potential in resident neural stem cells enables spinal cord repair. *Science*. 2020;370(6512).
67. Meletis K, Barnabé-Heider F, Carlén M, et al. Spinal cord injury reveals multilineage differentiation of ependymal cells. *PLoS Biol*. 2008;6(7):e182.
68. Kang H, Jha S, Ivovic A, et al. Somatic SMAD3-activating mutations cause melorheostosis by up-regulating the TGF- $\beta$ /SMAD pathway. *J Exp Med*. 2020;217(5).
69. Fathollahi A, Aslani S, Jamshidi A, Mahmoudi M. Epigenetics in osteoarthritis: novel spotlight. *J Cell Physiol*. 2019;234(8):12309-12324.
70. Del Real A, Riancho-Zarrabeitia L, López-Delgado L, Riancho JA. Epigenetics of skeletal diseases. *Curr Osteoporos Rep*. 2018;16(3):246-255.
71. Lister R, Pelizzola M, Dowen RH, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*. 2009;462(7271):315-322.
72. Sheng S, Yanchun QU, Zhang J. The application of next generation sequencing on epigenetic study. *Hereditas*. 2014;36(03):256-275.
73. Wang Q, Xiong H, Ai S, et al. CoBATCH for high-throughput single-cell epigenomic profiling. *Mol Cell*. 2019;76(1):206-216.e207.
74. Cheow LF, Courtois ET, Tan Y, et al. Single-cell multimodal profiling reveals cellular epigenetic heterogeneity. *Nat Methods*. 2016;13(10):833-836.
75. Angermueller C, Clark SJ, Lee HJ, et al. Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity. *Nat Methods*. 2016;13(3):229-232.
76. Guo F, Yan L, Guo H, et al. The transcriptome and DNA methylome landscapes of human primordial germ cells. *Cell*. 2015;161(6):1437-1452.
77. Zhu C, Gao Y, Guo H, et al. Single-cell 5-formylcytosine landscapes of mammalian early embryos and ESCs at single-base resolution. *Cell Stem Cell*. 2017;20(5):720-731.e5.
78. Choi S, Ray HE, Lai S-H, Alwood JS, Globus RK. Preservation of multiple mammalian tissues to maximize science return from ground based and spaceflight experiments. *PLoS One*. 2016;11(12).
79. Ayturk U. RNA-seq in skeletal biology. *Curr Osteoporos Rep*. 2019;17(4):178-185.
80. Carter LE, Kilroy G, Gimble JM, Floyd ZE. An improved method for isolation of RNA from bone. *BMC Biotechnol*. 2012;12:5.
81. Kelly NH, Schimenti JC, Ross FP, van der Meulen MCH. A method for isolating high quality RNA from mouse cortical and cancellous bone. *Bone*. 2014;68:1-5.
82. Le Bleu HK, Kamal FA, Kelly M, Ketz JP, Zuscik MJ, Elbarbary RA. Extraction of high-quality RNA from human articular cartilage. *Anal Biochem*. 2017;518:134-138.
83. Grinstein M, Dingwall HL, Shah RR, Capellini TD, Galloway JL. A robust method for RNA extraction and purification from a single adult mouse tendon. *PeerJ*. 2018;6:17.
84. Debnath S, Yallowitz AR, McCormick J, et al. Discovery of a periosteal stem cell mediating intramembranous bone formation. *Nature*. 2018;562(7725):133-139.
85. Feige P, Brun CE, Ritso M, Rudnicki MA. Orienting muscle stem cells for regeneration in homeostasis, aging, and disease. *Cell Stem Cell*. 2018;23(5):653-664.
86. Dell'Orso S, Juan AH, Ko KD, et al. Single cell analysis of adult mouse skeletal muscle stem cells in homeostatic and regenerative conditions. *Development (Cambridge, England)*. 2019;146(12).
87. Yin Z, Lin J, Yan R, et al. Atlas of musculoskeletal stem cells with the soft and hard tissue differentiation architecture. *Adv Sci*. 2020;7(23):2000938.
88. Chan CKF, Gulati GS, Sinha R, et al. Identification of the human skeletal stem cell. *Cell*. 2018;175(1):43-56.e21.
89. He J, Yan J, Wang J, et al. Dissecting human embryonic skeletal stem cell ontogeny by single-cell transcriptomic and functional analyses. *Cell Res*. 2021;31(7):742-757.
90. Petit F, Sears KE, Ahituv N. Limb development: a paradigm of gene regulation. *Nat Rev Genet*. 2017;18(4):245-258.
91. Stricker S, Mundlos S. Mechanisms of digit formation: human malformation syndromes tell the story. *Dev Dyn*. 2011;240(5):990-1004.
92. Feregrino C, Sacher F, Parnas O, Tschopp P. A single-cell transcriptomic atlas of the developing chicken limb. *BMC Genom*. 2019;20.
93. Kelly NH, Huynh NPT, Guilak F. Single cell RNA-sequencing reveals cellular heterogeneity and trajectories of lineage specification during murine embryonic limb development. *Matrix Biol*. 2020;89:1-10.
94. Gerber T, Murawala P, Knapp D, et al. Single-cell analysis uncovers convergence of cell identities during axolotl limb regeneration. *Science (New York, NY)*. 2018;362(6413):421.
95. Qin T, Fan CM, Wang TZ, et al. Single-cell RNA-seq reveals novel mitochondria-related musculoskeletal cell populations during adult axolotl limb regeneration process. *Cell Death Differ*. 2021;28(3):1110-1125.
96. Edwards JJ, Khanna M, Jordan KP, Jordan JL, Bedson J, Dziedzic KS. Quality indicators for the primary care of osteoarthritis: a systematic review. *Ann Rheum Dis*. 2015;74(3):490-498.
97. Jin XZ, Beguerie JR, Zhang WY, et al. Circulating C reactive protein in osteoarthritis: a systematic review and meta-analysis. *Ann Rheum Dis*. 2015;74(4):703-710.
98. Ji Q, Zheng Y, Zhang G, et al. Single-cell RNA-seq analysis reveals the progression of human osteoarthritis. *Ann Rheum Dis*. 2019;78(1):100-110.
99. Fisch KM, Gamini R, Alvarez-Garcia O, et al. Identification of transcription factors responsible for dysregulated networks in human osteoarthritis cartilage by global gene expression analysis. *Osteoarthr Cartil*. 2018;26(11):1531-1538.
100. Wang X, Ning Y, Zhang P, et al. Comparison of the major cell populations among osteoarthritis, Kashin-Beck disease and healthy chondrocytes by single-cell RNA-seq analysis. *Cell Death Dis*. 2021;12(6):551.
101. Kraus P, Lufkin T. Implications for a stem cell regenerative medicine based approach to human intervertebral disk degeneration. *Front Cell Dev Biol*. 2017;5:17.
102. Demers CN, Antoniou J, Mwale F. Value and limitations of using the bovine tail as a model for the human lumbar spine. *Spine*. 2004;29(24):2793-2799.
103. Kraus P, Lufkin T. Bovine annulus fibrosus cell lines isolated from intervertebral discs. *Genomics Data*. 2016;10:83-84.

104. Li K, Kapper D, Youngs B, et al. Potential biomarkers of the mature intervertebral disc identified at the single cell level. *J Anat.* 2019;234(1):16-32.
105. Fernandes LM, Khan NM, Trochez CM, et al. Single-cell RNA-seq identifies unique transcriptional landscapes of human nucleus pulposus and annulus fibrosus cells. *Sci Rep.* 2020;10(1):15263.
106. Risbud MV, Shapiro IM. Notochordal cells in the adult intervertebral disc: new perspective on an old question. *Crit Rev Eukaryot Gene Expr.* 2011;21(1):29-41.
107. Kraus P, Yerden R, Kocsis V, Lufkin T. RNA in situ hybridization characterization of non-enzymatic derived bovine intervertebral disc cell lineages suggests progenitor cell potential. *Acta Histochem.* 2017;119(2):150-160.
108. Choi KS, Harfe BD. Hedgehog signaling is required for formation of the notochord sheath and patterning of nuclei pulposi within the intervertebral discs. *Proc Natl Acad Sci USA.* 2011;108(23):9484-9489.
109. Gan Y, He J, Zhu J, et al. Spatially defined single-cell transcriptional profiling characterizes diverse chondrocyte subtypes and nucleus pulposus progenitors in human intervertebral discs. *Bone Research.* 2021;9(1):37.
110. Smolen JS, Aletaha D, McInnes IB. Rheumatoid arthritis. *Lancet (London, England).* 2016;388(10055):2023-2038.
111. Rao DA, Gurish MF, Marshall JL, et al. Pathologically expanded peripheral T helper cell subset drives B cells in rheumatoid arthritis. *Nature.* 2017;542(7639):110-114.
112. McInnes IB, Schett G. The pathogenesis of rheumatoid arthritis. *N Engl J Med.* 2011;365(23):2205-2219.
113. Stephenson W, Donlin LT, Butler A., et al. Single-cell RNA-seq of rheumatoid arthritis synovial tissue using low-cost microfluidic instrumentation. *Nat Commun.* 2018;9(1):791.
114. Zhang F, Wei K, Slowikowski K, et al. Defining inflammatory cell states in rheumatoid arthritis joint synovial tissues by integrating single-cell transcriptomics and mass cytometry. *Nat Immunol.* 2019;20(7):928-942.
115. Svensson V, Vento-Tormo R, Teichmann SA. Exponential scaling of single-cell RNA-seq in the past decade. *Nat Protoc.* 2018;13(4):599-604.
116. Dean FB, Hosono S, Fang L, et al. Comprehensive human genome amplification using multiple displacement amplification. *Proc Natl Acad Sci USA.* 2002;99(8):5261-5266.
117. Ballantyne KN, van Oorschot RAH, Muharam I, van Daal A, Mitchell RJ. Decreasing amplification bias associated with multiple displacement amplification and short tandem repeat genotyping. *Anal Biochem.* 2007;368(2):222-229.
118. Stegle O, Teichmann SA, Marioni JC. Computational and analytical challenges in single-cell transcriptomics. *Nat Rev Genet.* 2015;16(3):133-145.
119. Wu N, Sun H, Zhao X, et al. MAP3K2-regulated intestinal stromal cells define a distinct stem cell niche. *Nature.* 2021;592(7855):606-610.
120. Pennisi E. Development cell by cell. *Science (New York, NY).* 2018;362(6421):1344-1345.

**How to cite this article:** Zhang Y, Wang J, Yu C, et al. Advances in single-cell sequencing and its application to musculoskeletal system research. *Cell Prolif.* 2022;55:e13161. doi:[10.1111/cpr.13161](https://doi.org/10.1111/cpr.13161)