

## Review

Illumination of neural development by *in vivo* clonal analysisMingrui Xu <sup>a, b, d</sup>, Jingjing Wang <sup>a, d</sup>, Xize Guo <sup>a, b, d</sup>, Tingting Li <sup>a, b, d</sup>, Xia Kuang <sup>a</sup>, Qing-Feng Wu <sup>a, b, c, \*</sup><sup>a</sup> Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China<sup>b</sup> University of Chinese Academy of Sciences, Beijing 100101, China<sup>c</sup> Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Beijing 100101, China

## ARTICLE INFO

## Article history:

Received 15 July 2018

Received in revised form

22 August 2018

Accepted 18 September 2018

Available online 12 October 2018

## Keywords:

*In vivo* clonal analysis

Lineage tracing

Neural stem cells

Neural development

Nuclear brain structure

## ABSTRACT

Single embryonic and adult neural stem cells (NSCs) are characterized by their self-renewal and differentiation potential. Lineage tracing via clonal analysis allows for specific labeling of a single NSC and tracking of its progeny throughout development. Over the past five decades, a plethora of clonal analysis methods have been developed in tandem with integration of chemical, genetic, imaging and sequencing techniques. Applications of these approaches have gained diverse insights into the heterogeneous behavior of NSCs, lineage relationships between cells, molecular regulation of fate specification and ontogeny of complex neural tissues. In this review, we summarize the history and methods of clonal analysis as well as highlight key findings revealed by single-cell lineage tracking of stem cells in developing and adult brains across different animal models.

© 2018 Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences. Production and hosting by Elsevier B.V. on behalf of KeAi. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. Introduction

The structural organization and function of the nervous system rely on the sequential generation of distinct neural cell types, such as neurons, astrocytes and oligodendrocytes.<sup>1,2</sup> Upon formation and closure of the neural tube, billions of neurons are generated from neural stem cells (NSCs) prenatally and trillions of neural connections are produced postnatally. In the developing brain, neuroepithelial cells (NEs), the earliest NSC type, first undergo symmetric proliferative cell division to expand the stem cell pool and then give rise to radial glial cells (RGCs) which undergo asymmetric neurogenic cell division.<sup>3,4</sup> The massive generation of neurons by asymmetrically dividing RGCs is followed by the production of glia at the postnatal stage. Thereafter, most of the RGCs lose their self-renewal and differentiation capacity, ending up with terminal differentiation into either glial cells or ependymal cells. At the same time, a subpopulation of RGCs in the subventricular zone of lateral

ventricle and subgranular zone of dentate gyrus retain their stem cell potential until adulthood.<sup>4</sup>

The heterogeneity and dynamic state of NSCs have long hindered efforts towards comprehensive characterization of these cells using population analysis methods.<sup>5,6</sup> To this end, quantitative single-cell approaches are required for detailed interpretations of NSC behavior and fate specification. Clonal analysis allows for the quantitative analysis of clones originated from a single stem cell, and provides unprecedented spatial and/or temporal resolution to investigate the lineage progression and fate specification of stem cells. Retrospectively, clonal analysis of stem cells can reveal the ontogenetic organization of diverse tissues during organogenesis and uncover cellular mechanisms underlying tissue homeostasis, thereby gaining key insights into cardinal properties of neural cells throughout the developmental continuum. Therefore, single-cell analysis is particularly critical for developmental and stem cell biological research.

Genetic-based single-cell lineage tracing, time-lapse live imaging and single-cell genomic/transcriptomic profiling represent the most cutting-edge methods of clonal analysis.<sup>7</sup> Although whole-organism lineage tracing has elucidated the lineage correlation and fate of all cells in the nematode *Caenorhabditis elegans* three decades ago,<sup>8</sup> the landscape view of heterogeneous NSCs in the insect, vertebrate and mammalian brain remains unclear. Here, we review the approaches and applications of clonal analysis in uncovering NSC behavior within biologically relevant contexts.

\* Corresponding author. Institute of Genetics and Developmental Biology, Chinese Academy of Sciences, Beijing 100101, China.

E-mail address: [wu\\_qingfeng@genetics.ac.cn](mailto:wu_qingfeng@genetics.ac.cn) (Q.-F. Wu).

Peer review under responsibility of Guangzhou Institutes of Biomedicine and Health, Chinese Academy of Sciences.

<sup>d</sup> These authors contributed equally to this review.

## 2. History and development of clonal analysis

There are multiple waves of technological innovation driving the development of clonal analysis in history (Fig. 1). The earliest instances of clonal analysis using light microscopy to illuminate cell cleavages in the invertebrate embryo date back to 1905<sup>9</sup> and provided a powerful tool to study stem cell behavior during embryogenesis. Time-lapse imaging at the single-cell resolution then allowed for four-dimensional (4D) reconstitution of an invertebrate embryo with an extremely simple cellular architecture, and a milestone was reached when a complete atlas and lineage tree of *Caenorhabditis elegans* was determined.<sup>8</sup> However, achieving 4D reconstitution of a vertebrate embryo by time-lapse imaging still proved to be challenging. Later on, various dyes were applied to label individual founder stem cells and perform clonal analysis in different stages of chick, zebrafish and mouse embryogenesis,<sup>10</sup> but the disadvantage of being diluted with successive rounds of cell division limits the application of these tracers.

With the advent of recombinant DNA technology in the 1980s, retrovirus encoding reporter genes such as  $\beta$ -galactosidase (lacZ) and a set of DNA barcode tags were used for stem cell labeling and lineage reconstruction.<sup>11–13</sup> The reconstitution of clonal relationships among lacZ-labeled cells depended on PCR amplification and sequencing of barcode tags, and artificial errors in lineage reconstruction were greatly minimized as the size of retroviral library expanded from hundreds to millions of barcode tags.<sup>13–16</sup>

Molecular engineering of green fluorescent protein (GFP) achieved a major breakthrough in 1995, markedly improving the spectral characteristics, fluorescent intensity and photostability of GFP.<sup>17</sup> Subsequently, GFP and its derivatives (YFP, RFP, CFP, etc.) were widely applied in retrovirus-based and genetic-based population lineage tracing.<sup>18–21</sup> The implementation of two recombinase systems (Cre-loxP and FLP-FRT) further enhanced the effectiveness of using fluorescent reporter genes in lineage tracing to achieve cell-type-specific or tissue-specific labeling.<sup>22–24</sup> The combination of multicolor reporter genes with recombinase systems in genetic animal models, such as Brainbow and Confetti mice, then permitted for clonal tracing of fluorescent cells *in vivo* and *ex vivo*,<sup>25,26</sup> whereas continuous observation of color-coded clones by time-lapse imaging served to inform on dynamic changes of clones arising from a single stem cell.<sup>27–31</sup> A caveat was that since such intrachromosomal recombination of multicolor reporter

genes yielded a widespread level of labeling, extensive video microscopy was needed to explicitly distinguish clones. Comparatively, genetic animals with interchromosomal recombination systems, like twin-spot generator (TSG) for *Drosophila* and mosaic analysis with double markers (MADM) for mice, can allow us to achieve extremely sparse labeling of stem cells for long-term clonal tracing.<sup>32–34</sup> In the hematologic system, transplantation of single cell bearing fluorescent proteins was recently performed for clonal analysis of hematopoietic stem cells.<sup>35</sup>

With the rise of the clustered regularly interspaced palindromic repeats (CRISPR)-Cas9 system for genomic editing of mammalian cells in 2012, an innovative technique termed genome editing of synthetic target arrays for lineage tracing (GESTALT) was subsequently used to track the fate of individual stem cells and reconstruct the lineage relationships in complex, multicellular systems.<sup>36,37</sup> Simultaneously, the development of single-cell genomic/transcriptomic profiling technologies now allows for the generation of a lineage tree that covers all identified cell types rooted to a single stem cell group in whole planarians, although only a minority of lineages are reconstituted from a large amount of cells in human beings.<sup>38–40</sup> In summation, these methodological advances in both prospective and retrospective clonal analysis complementarily offer valuable insights into nervous system development and neural lineages, particularly in understanding the intricate mechanisms governing mammalian brain patterning.

## 3. Advanced methods of clonal analysis

The classic clonal analysis method is to label a single founder cell and track its progeny over time by either snapshot-type endpoint analysis or real-time recording. Here we provide a broad overview of the genetic toolbox for clonal analysis applied to dissect the cellular architecture of the developing brain and discuss the advantages and potential caveats of each technique.

### 3.1. Sparse retroviral library labeling for clonal analysis

A barcoded retroviral library containing short DNA oligonucleotides with high complexity and fluorescent reporter gene can be injected into the brain at different embryonic stages for lineage tracing.<sup>16</sup> To achieve cell-type or tissue specificity, genetically-modified animals with fluorescent labeling of specific cell

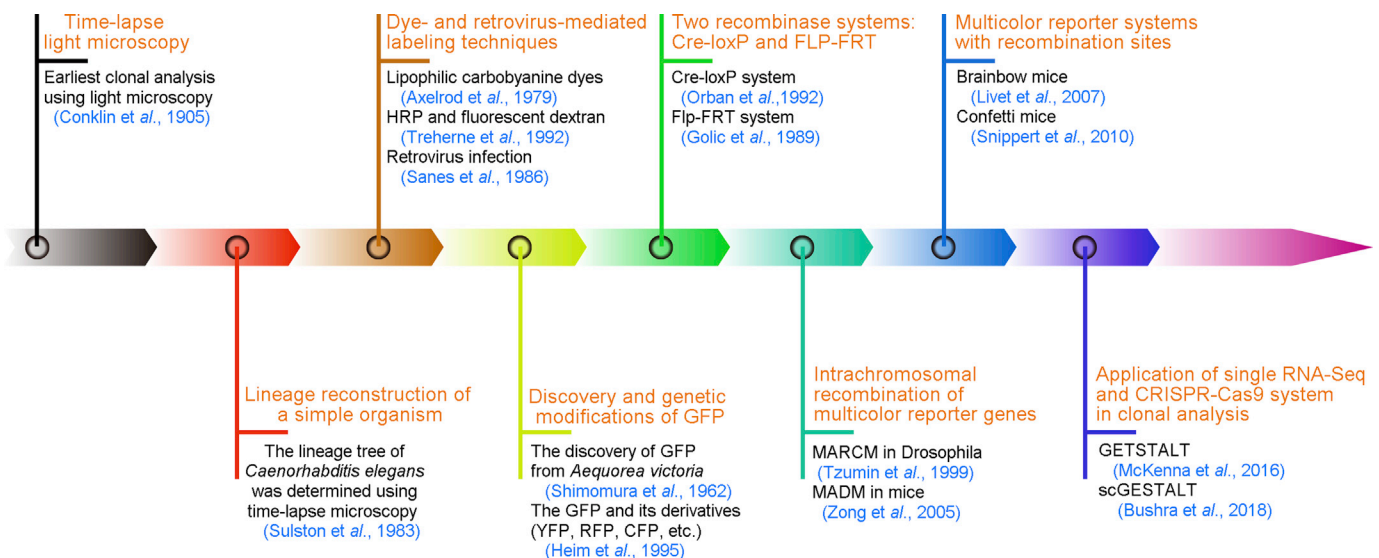
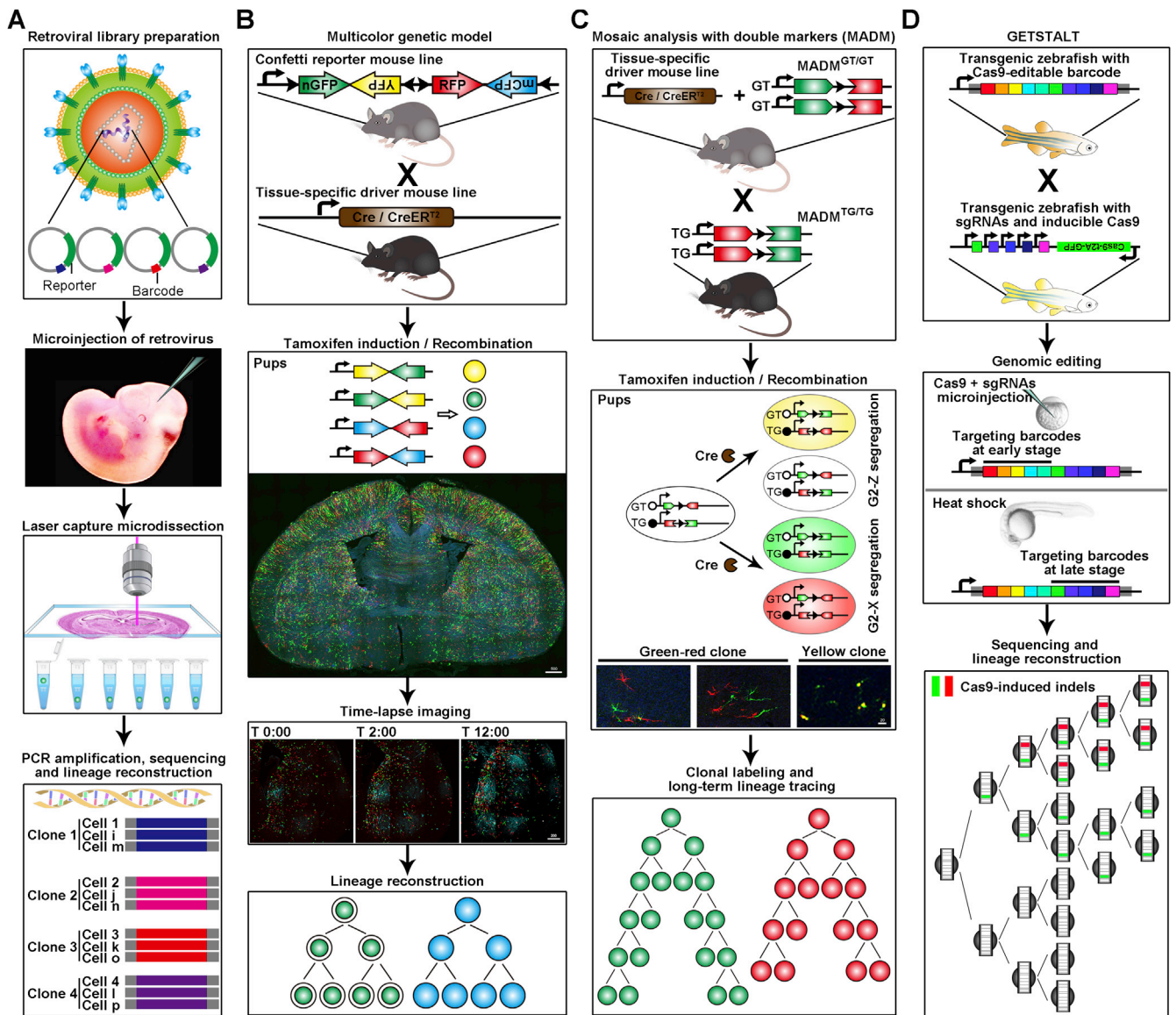


Fig. 1. History of clonal analysis.

population have been used as the host organisms for retroviral library infection.<sup>16</sup> The availability of different fluorescent colors in retroviruses and host animals allows us to precisely observe and dissect specific neural cell types by fluorescence-activated cell sorting (FACS) or laser capture microdissection (LCM). PCR amplification of integrated barcode tags can then reveal the clonal relationships among dissected cells, as sister neural cells originating from the same NSCs share the identical barcode tag (Fig. 2A). Notably, while LCM preserves the spatial information of dissected cells and gives us clues regarding the spatial distribution of clonally-related cells, single-cell RNAseq of FACS-sorted cells provides a large database pertaining to the heterogeneity of molecular profiles among sister cells.

Despite the power of retroviral library labeling in reconstructing cell lineages, there are a few limitations of this method that

introduce uncertainties to the interpretations of clonal analysis results. Firstly, the incorporation of retroviral libraries into the host cell genome may disrupt the function of endogenous genes and therefore interfere with the behavior of stem cells. Secondly, spontaneous silencing of reporter genes in retroviral vectors could result in a comparative underestimation of clonal size.<sup>37</sup> However, it has been shown that clones from the same founder cell seem to have a propensity to silence genes following an “all-or-nothing” pattern, which lowers the likelihood of skewed data interpretation.<sup>41,42</sup> Thirdly, recovering the barcode tag remains a significant challenge, with instances showing a recovery rate of only approximately 50% of barcode tags in labeled cells.<sup>16</sup> Lastly, some barcode tags appear to be overrepresented in the retroviral library, leading to cases whereby the progeny derived from more than a single NSC share the same barcode tag and are erroneously considered as



**Fig. 2. Multiplexed strategies for lineage tracing and clonal analysis.** (A) Sparse retroviral library labeling introduces a unique and heritable DNA barcode tag and a reporter gene into the genome of host cells. The spatial position and lineage relationships among labeled cells can be uncovered by laser capture microdissection (LCM) and PCR amplification. (B) Schematic of genetic multicolor labeling. In Confetti reporter system adapted from Brainbow, random Cre-induced recombination between different loxP sites allows for expression of different fluorescent proteins. (C) Scheme of mosaic analysis with double markers (MADM). MADM system consists of two reciprocally complementary chimeric marker genes hybridizing GFP and RFP in the same locus of homologous chromosomes. Intrachromosomal recombination during mitotic division induced by tamoxifen administration recovers functional fluorescent reporter genes. While G2-X segregation generates a double-labeled cells and a colorless cell, G2-Z segregation produces a pair of green and red cells. (D) Scheme of genome editing of synthetic target arrays for lineage tracing (GESTALT). GESTALT contains a designated barcode that progressively accumulates mutations (deletion and insertion) by CRISPR-Cas9 over cell division. The lineage tree is reconstructed by mapping mutations in the barcode among labeled cells.



sibling cells.<sup>14</sup> Even so, such “lumping” errors can be mitigated by vastly increasing the diversity of barcode tags. The advances in retroviral library generation, high-throughput deep sequencing and computational approaches have been partially circumventing these limitations.

### 3.2. Genetic multicolor labeling for clonal analysis and time-lapse imaging

The emergence of multicolor reporter animal models (including Brainbow, Confetti, MCFO, Flybow and ZebraBow) provides remarkable tools to decode the complexity of neural tissues.<sup>25–27,30,43–45</sup> By harnessing Cre/loxP and FLP/FRT recombinase systems, we can switch on the expression of more than three fluorescent proteins in a stochastic manner, yielding up to 90 distinguishable colors. Neural cells presenting a particular color are apt to share the same lineage, providing a readout to examine the clonal origin of cells (Fig. 2B). Nevertheless, it is still difficult to accurately identify clones from a multitude of labeled cells in the snapshot images, despite the progress of computational methods in defining color codes during multicolor fluorescence clonal tracing.<sup>46</sup> Recent advances in miniature two-photon microscope technology for *in vivo* live imaging will be useful to shed light on the features of clonal expansion and competition in the brain based on multicolor labeling.<sup>47</sup>

In combination with genetic labeling, time-lapse imaging provides a continuous observation of single stem cell and directly reveals its behavior within a niche, which can include the mode of cell division, cell cycle length and cell death. In nematode embryos, live imaging has revealed the genetic network involved in the lineage progression and fate specification of embryonic progenitors at the system level.<sup>48</sup> Due to the optical transparency of zebrafish, the ZebraBow genetic model has also been developed for *in vivo* multicolor imaging and long-term lineage tracing.<sup>30</sup> In addition, the deep penetration and sharp imaging quality of multi-photon imaging make it an ideal platform for monitoring stem cell activity in the vertebrate and mammalian brains,<sup>49,50</sup> whereas time-lapse imaging during mammalian embryogenesis remains as yet unfeasible.

### 3.3. Inducible genetic sparse labeling for clonal analysis

The development of mosaic analysis with a repressible cell marker (MARCM) allows for sparse labeling of cells in the *Drosophila* nervous system, which has many applications in studying lineage relationships. As such, the MARCM technique inspired the emergence of TSG and MADM reporter animal models, in which the two halves of two fluorescent reporter genes are hybridized, split by loxP or FRT sites, driven by a common promoter and placed at the same genomic loci on homologous chromosomes (Fig. 2C). Given that TSG and MADM reporter systems require Cre or FLP recombinase-mediated mitotic recombination between two homologous chromosomes, the efficiency of this genetic labeling is low. The fusion of recombinase with estrogen receptor (ER) under the control of a specific promoter can not only achieve cell-type or tissue specificity, but also enables the driver system to be temporally inducible. Here, an optimized dose titration of an inducer (such as tamoxifen) in the Cre-ER driver system can significantly promote the clonal specificity, allowing for a much higher resolution of lineage tracing at the single-clone level.<sup>51</sup> Hence, it would thereby be feasible to achieve genetic labeling of a single NSC throughout the whole brain at an early embryonic stage.<sup>33</sup>

Although inducible genetic sparse labeling achieves unprecedented spatiotemporal resolution, there remain several drawbacks. Similar to other snapshot-type endpoint analysis, this technique does not permit continuous observation of stem cells, and it is

difficult to reconstruct multiple cell lineages at one time by applying this approach if a single stem cell generates a number of progeny with widespread migration. Apart from being labor-intensive and time-consuming to perform this type of clonal analysis, the results generated may also biasedly label stem cells with specific properties.<sup>52</sup> The effects of these limitations can be reduced along with the development of research equipment with automatic serial sectioning and high-speed imaging.<sup>53,54</sup>

### 3.4. Genomic editing of synthetic target arrays for clonal analysis

Recently, GESTALT has been developed to reconstruct whole-organism lineages in zebrafish.<sup>36</sup> This approach relies on an inducible system that enables the progressive accumulation of unique, random mutations in custom-designed barcodes over multiple rounds of cell division (Fig. 2D). To achieve 11 barcode editing, a transgenic zebrafish line expressing fluorescent reporter gene with ten single guide RNA (sgRNA) targeting sites at its 3' untranslated region is generated, followed by the injection of Cas9 ribonucleoprotein containing sgRNAs that match each of the ten targets in the synthetic barcode.<sup>36</sup> Alternatively, induction of Cas9 by heat shock in zebrafish or Cre-ER driver system makes GESTALT temporally controllable.<sup>37</sup> The random insertions and deletions (indels) in the barcode can mark NSCs and their progeny with unique tags, while the clonal relationships among sampled cells will be reconstructed based on the pattern of shared indels. However, the potential deletion of multiple target sites in the barcode will hinder complete reconstruction of cell lineages in the brain. Other barcode recording methods used to reconstruct cell lineages include the *Polylox* barcoding system and MEMOIR system,<sup>55,56</sup> which provide insights into the cellular diversity of clonally-related cells via a combination of the barcode recording approach with high-throughput single-cell RNAseq.

While GESTALT is widely applicable in zebrafish and its strategy has been used to investigate cell lineages during mammalian embryogenesis,<sup>37,57</sup> several optimizations may help to further improve the accuracy of lineage reconstruction. Firstly, the large fragment deletion of tandem barcodes significantly impairs the barcode tag recovery and confounds the reconstruction of cellular pedigrees. Consequently, efforts to design barcodes that can avoid large fragment deletion may largely eliminate this potential limitation of GESTALT. Secondly, low expression level and inefficient capture of barcode transcripts would lead to low recovery rate of barcode tags. Finally, the results generated lack precise spatial information of analyzed cells; however, progress in the development of single-cell transcriptome atlases will be able to assign the analyzed cells to specific anatomical positions.<sup>58</sup>

Taken together, the integrated engineering of barcode tags, reporter genes, genetic recombination system, transposon system and Cas9 nuclease in the virus and host animals facilitates the development and innovation of diverse clonal analysis methods. Concomitantly, advances in imaging and sequencing technologies have also conferred greatly enhanced spatiotemporal resolution and molecular information as readouts of clonal lineage tracing.

## 4. Application of clonal analysis in the developing brain

### 4.1. *Drosophila*

As a powerful genetic animal model, *Drosophila* has been used to develop multiple elegant approaches of genetic labeling and clonal analysis. MARCM was originally used to label lineage-related cells in the fly brain for visualizing the cytoarchitecture and neural connections in neuroblast-derived clones.<sup>34,59</sup> With MARCM, it was disclosed that different neuronal subtypes within mushroom

bodies or antennal lobes are sequentially generated from corresponding distinct populations of neuroblasts.<sup>60,61</sup> Furthermore, the approach was used to uncover that type II neuroblasts exhibit a unique pattern of expansion and serve as the major source of glia at larval stage as compared with type I neuroblasts,<sup>62–68</sup> and also to show the production of two daughter cells with different fates from a single ganglion mother cell, which is mediated by Notch signaling.<sup>69,70</sup>

One notable extension of MARCM is to use genetic manipulation to switch cell fates at single-cell resolution.<sup>71–73</sup> As a second-generation system for clonal analysis in the fly brain, TSG was developed and used in neural regeneration studies.<sup>74</sup> For instance, it was reported that in the injured imaginal discs of larval brain, regenerative cells stop dividing as soon as the damaged tissue is restituted, but reactivate to proliferate when their cell fate is changed.<sup>75</sup> Another study using TS-MARCM revealed that projection neurons and local interneurons are generated in pairs via Notch-mediated binary fate specification of neural precursors in the lateral antennal lobe.<sup>76,77</sup> In particular, the clonal lineage tracing using Flybow showed that adult glia surrounding thoracic neuropil originate from ~40 astrocytes distributing around immature leg neuropil in third instar larval stage,<sup>45,78</sup> whereas all 12 embryonic peripheral glia persist to the end of larval stage with some being maintained until adulthood, and differentially contribute to three distinct glial layers surrounding peripheral nerves.<sup>79,80</sup> These studies are but several of many which highlight the importance of genetic tools developed for *Drosophila*, and have enabled cell lineage fate analyses at the fundamental level as well as facilitated studies of the mechanisms underlying fly brain development.

#### 4.2. Zebrafish

The optical transparency of zebrafish enables *in vivo* real-time imaging to easily observe cell expansion, migration and connection. Early studies using zebrafish as a model used dye lineage tracers to track clones in zebrafish central nervous system, and revealed that the division of clonally-related cells are synchronized and consistently oriented during cell movement to form clonal strings along the anterior–posterior axis.<sup>81,82</sup> Strikingly, the second wave of cell division in zebrafish was found to occur at the midline of nervous system by tracking the dye, converting the single clonal string into a bilateral pair of strings on each side of the neural tube.<sup>81,82</sup>

Together with studies in the central nervous system, the retina is also widely used to investigate the principles of neural development in zebrafish. In this regard, studies using a combination of *in vivo* mosaic labeling, live imaging and Zebrafish-based clonal analysis have revealed that retinal stem cells originate from bipotent progenitors in the medial epithelial layer of optic vesicles, and that these cells behave in a stochastic manner.<sup>30,83–85</sup> In recent years, the multi-pronged convergence of single-cell sequencing, CRISPR-Cas9 and recombinase systems has generated a variety of tools for lineage tracing such as GESTALT, ScarTrace and 2C-Cas9 technology to perform clonal analysis and interrogate gene function in zebrafish.<sup>36,37,86,87</sup> Future applications of these new methods are expected to uncover and characterize other fundamentally important rules involved in governing the formation of the vertebrate brain.

#### 4.3. Mouse

In the mammalian brain, clonal analysis has been applied to a wide spectrum of studies to reveal the behavior of neural progenitors, organization of sister neurons, origin of adult NSCs,

cytoarchitecture of neocortex, ontogeny of thalamic nuclei and molecular regulatory logic of NSC behavior.<sup>12,13,16,33,41,51,88–92</sup> The pioneering applications of barcoded retroviral libraries in the embryonic mouse brain suggested that clonally-related neurons in the cerebral cortex widely disperse across functional regions.<sup>12,13,93</sup> In contrast, later studies provided compelling evidences to support the radial unit hypothesis, which proposed that cerebral cortex is assembled by an array of radial columns and implied the generation of individual structural columns from single RGCs.<sup>92,94,95</sup> In corroboration with these findings, a recent study using MADM-based clonal analysis showed that lineage-related neurons are organized in the shape of columns or cones and span multiple layers.<sup>32</sup> Moreover, it was indicated that the production of excitatory projection neurons by single RGCs is unitary and intrinsically determined at the rate of ~8–9 neurons per asymmetric division,<sup>32</sup> and it has also been shown that excitatory sister neurons within the same clones preferentially develop electrical and chemical synapses.<sup>92,96</sup>

While a consensus that clonally-related excitatory projection neurons are organized into radial column has been reached, the spatial organization of sister inhibitory interneurons remains the subject of debate. In one study, sparse labeling of the neocortex was achieved by retrovirus injection into a knockin host mouse line expressing avian tumor virus receptor A (TVA), and the clonally-related interneurons exhibited vertically or horizontally aligned clusters.<sup>89</sup> Conversely, two studies using a combination of genetic animal and barcoded retroviral libraries revealed that sister interneurons display widespread distribution across functional regions and structural boundaries.<sup>42,88</sup> Therefore, further work regarding the dispersion patterns of clonally-related interneurons could be informative given that lineage alone may not clearly elucidate the distribution of sister interneurons.<sup>88,90</sup>

The organizational logic dictating the formation of nuclear brain structures (such as thalamus, hypothalamus and brainstem) is much less understood compared with the cytoarchitecture of laminated cerebral cortex. Recently, the spatiotemporal organization of thalamic nucleogenesis has been studied using MADM-based clonal analysis.<sup>33,91</sup> One study reported that anterior thalamic clones contribute to nuclei with cognitive functions while medial ventral posterior clones are integrated into nuclei with sensory/motor functions, and suggested that clones which are fate specified to first-order and higher-order sensory/motor nuclei are spatially segregated.<sup>91</sup> In comparison, another study showed that the ontogenetic relationship of distinct thalamic nuclei appeared to be determined by the spatial position of founder NEs/RGCs rather than the functional modality criteria, and further uncovered four fundamental features of thalamic nucleogenesis: 1) individual NEs/RGCs contribute to multiple thalamic nuclei; 2) thalamic nuclei are temporally generated within the lineage; 3) spatial coding of RGCs determines the fate specification of clones to specific sets of nuclei, and 4) the fate of intermediate progenitors (IPCs) is not specified to a single nucleus.<sup>33</sup> Although the aforementioned studies have provided a framework for studying cytoarchitecture and cellular connectivity in the mammalian thalamus, subsequent investigations into the mechanisms driving thalamic self-organization would provide a more comprehensive roadmap of nucleus formation.

In the germinal zone of the adult brain, *in vivo* clonal analysis studies have revealed different patterns of hippocampal NSC activation, including symmetric and asymmetric self-renewal.<sup>51</sup> Live imaging of adult neurogenesis in hippocampus suggests that NSCs firstly undergo symmetrically proliferative division, then transit to an asymmetrically neurogenic phase, and are lost after a burst of cell division.<sup>50</sup> Interestingly, individual adult NSCs were found to generate neurons and astrocytes but not oligodendrocytes.<sup>51</sup> Moreover, a study which carried out single-cell lineage tracing

using retroviral libraries with more than  $10^5$  barcoded tags indicated that adult NSCs in the subventricular zone of lateral ventricle originate from embryonic RGCs and are regionally specified as early as embryonic day 11.5 (E11.5).<sup>16</sup> This study also revealed that the prenatally-born neurons in the forebrain regions including cortex, striatum and septum are clonally related with postnatally-born neurons in the olfactory bulb by preserving the spatial coding information of embryonic progenitors in postnatal NSCs.<sup>16</sup> It is likely that integrated approaches utilizing clonal analysis and molecular profiling in the next decade will be essential to reveal the molecular regulatory logic of complex NSC behavior.

#### 4.4. Nonhuman primate

To date, there have been very few studies conducted on clonal analysis of NSCs in the nonhuman primate brain possibly due to intrinsic challenges of collecting data from primate. A recent study used long-term *ex vivo* live imaging to provide a continuous and clonal observation of neural progenitors in outer subventricular zone (OSVZ) of embryonic macaque brains, followed by unbiased sampling of cycling neural progenitors.<sup>97</sup> In this study, four distinct types of basal RGCs and one class of IPCs were identified and the database of clonal observation includes critical attributes of single progenitor behavior: cell cycle length, mode of division, direction of somal translocation, clonal size and cell fate transition. Specifically, it was highlighted that every type of neural progenitors self-renews and directly generates neurons, with variable rounds of cell division. It also appears that the cell cycle length is shortened at later developmental stage to expand the neural progenitor pool in OSVZ, while different types of neural progenitors display bidirectional transitions in fate.<sup>97</sup> Despite the general paucity of information in this area, these results unveil the hitherto unknown differences in the behavior of neural progenitors between rodent and nonhuman primate and pave the way for raising other important questions on species-specific characteristics of NSCs.

### 5. Conclusions and future directions

Clonal analysis has been developed and applied for a century since the first characterization of cell lineage in embryos. Today, the availability and integration of sophisticated toolkits incorporating genetic tools, genomic editing approaches, imaging and sequencing technologies have driven and opened up new avenues of research to developmental and stem cell biologists alike. These methods have been invaluable in elucidating the heterogeneity of NSCs in self-renewal and differentiation behavior, ontogenetic organization of neural tissues, homeostatic regeneration of neural progenitors after injury and molecular network controlling the behavior of NSCs. More importantly, the conceptual advances derived from clonal analysis in the brain are beneficial for the progress of regenerative medicine in context of the pathophysiology of brain development, wherein future functional studies will further establish the relationship between cell lineages and functional modality.

#### Conflict of interest

The authors declare that they have no conflict of interest.

#### Acknowledgements

We thank Samuel Z.H. Wong for comments. This work is supported by funding from the Hundred-Talent Program (Chinese Academy of Sciences), Strategic Priority Research Program of the

Chinese Academy of Sciences (XDA16020100) and National Natural Science Foundation of China (31771131) and to Q.F.W.

#### References

- Gage FH. Mammalian neural stem cells. *Science*. 2000;287(5457):1433–1438.
- Temple S. The development of neural stem cells. *Nature*. 2001;414(6859):112–117.
- Taverna E, Gotz M, Huttner WB. The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. *Annu Rev Cell Dev Biol*. 2014;30:465–502.
- Kriegstein A, Alvarez-Buylla A. The glial nature of embryonic and adult neural stem cells. *Annu Rev Neurosci*. 2009;32:149–184.
- Krieger T, Simons BD. Dynamic stem cell heterogeneity. *Development*. 2015;142(8):1396–1406.
- Hoppe PS, Coutu DL, Schroeder T. Single-cell technologies sharpen up mammalian stem cell research. *Nat Cell Biol*. 2014;16(10):919–927.
- Etzrodt M, Ende M, Schroeder T. Quantitative single-cell approaches to stem cell research. *Cell Stem Cell*. 2014;15(5):546–558.
- Sulston JE, Schierenberg E, White JG, Thomson JN. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev Biol*. 1983;100(1):64–119.
- Conklin EG. *The Organization and Cell-lineage of the Ascidian Egg*. 1905. Philadelphia.
- Clarke JD, Tickle C. Fate maps old and new. *Nat Cell Biol*. 1999;1(4):E103–E109.
- Turner DL, Cepko CL. A common progenitor for neurons and glia persists in rat retina late in development. *Nature*. 1987;328(6126):131–136.
- Walsh C, Cepko CL. Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. *Science*. 1992;255(5043):434–440.
- Walsh C, Cepko CL. Clonal dispersion in proliferative layers of developing cerebral cortex. *Nature*. 1993;362(6421):632–635.
- Reid CB, Liang I, Walsh C. Systematic widespread clonal organization in cerebral cortex. *Neuron*. 1995;15(2):299–310.
- Golden JA, Fields-Berry SC, Cepko CL. Construction and characterization of a highly complex retroviral library for lineage analysis. *Proc Natl Acad Sci U S A*. 1995;92(12):5704–5708.
- Fuentealba LC, Rompani SB, Parraguez JI, et al. Embryonic origin of postnatal neural stem cells. *Cell*. 2015;161(7):1644–1655.
- Heim R, Cubitt AB, Tsien RY. Improved green fluorescence. *Nature*. 1995;373(6516):663–664.
- Yuste R. Fluorescence microscopy today. *Nat Methods*. 2005;2(12):902–904.
- Woodworth MB, Girsakis KM, Walsh CA. Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat Rev Genet*. 2017;18(4):230–244.
- Venters SJ, Dias da Silva MR, Hyer J. Murine retroviruses re-engineered for lineage tracing and expression of toxic genes in the developing chick embryo. *Dev Dyn*. 2008;237(11):3260–3269.
- Papaioannou VE. Concepts of cell lineage in mammalian embryos. *Curr Top Dev Biol*. 2016;117:185–197.
- Kilby NJ, Snaith MR, Murray JA. Site-specific recombinases: tools for genome engineering. *Trends Genet*. 1993;9(12):413–421.
- Gates CA, Cox MM. FLP recombinase is an enzyme. *Proc Natl Acad Sci U S A*. 1988;85(13):4628–4632.
- Silver DP, Livingston DM. Self-excising retroviral vectors encoding the Cre recombinase overcome Cre-mediated cellular toxicity. *Mol Cell*. 2001;8(1):233–243.
- Snippert HJ, van der Flier LG, Sato T, et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell*. 2010;143(1):134–144.
- Livet J, Weissman TA, Kang H, et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature*. 2007;450(7166):56–62.
- Hampel S, Chung P, McKellar CE, Hall D, Looger LL, Simpson JH. Drosophila Brainbow: a recombinase-based fluorescence labeling technique to subdivide neural expression patterns. *Nat Methods*. 2011;8(3):253–259.
- Rinkevich Y, Lindau P, Ueno H, Longaker MT, Weissman IL. Germ-layer and lineage-restricted stem/progenitors regenerate the mouse digit tip. *Nature*. 2011;476(7361):409–413.
- Gupta V, Poss KD. Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature*. 2012;484(7395):479–484.
- Pan YA, Freundlich T, Weissman TA, et al. Zebrafish: multispectral cell labeling for cell tracing and lineage analysis in zebrafish. *Development*. 2013;140(13):2835–2846.
- Loulier K, Barry R, Mahou P, et al. Multiplex cell and lineage tracking with combinatorial labels. *Neuron*. 2014;81(3):505–520.
- Gao P, Postiglione MP, Krieger TG, et al. Deterministic progenitor behavior and unitary production of neurons in the neocortex. *Cell*. 2014;159(4):775–788.
- Wong SZH, Scott EP, Mu W, et al. In vivo clonal analysis reveals spatiotemporal regulation of thalamic neurogenesis. *PLoS Biol*. 2018;16(4):e2005211.
- Lee T, Luo L. Mosaic analysis with a repressible cell marker (MARCM) for Drosophila neural development. *Trends Neurosci*. 2001;24(5):251–254.
- Yamamoto R, Morita Y, Oeohara J, et al. Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell*. 2013;154(5):1112–1126.



36. McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. *Science*. 2016;353(6298):aa7907.
37. Raj B, Wagner DE, McKenna A, et al. Simultaneous single-cell profiling of lineages and cell types in the vertebrate brain. *Nat Biotechnol*. 2018;36(5):442–450.
38. Plass M, Solana J, Wolf FA, et al. Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science*. 2018;360(6391).
39. Shapiro E, Biezuner T, Linnarsson S. Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nat Rev Genet*. 2013;14(9):618–630.
40. Evrony GD, Lee E, Mehta BK, et al. Cell lineage analysis in human brain using endogenous retroelements. *Neuron*. 2015;85(1):49–59.
41. Mayer C, Jaglin XH, Cobbs LV, et al. Clonally related forebrain interneurons disperse broadly across both functional areas and structural boundaries. *Neuron*. 2015;87(5):989–998.
42. Harwell CC, Fuentealba LC, Gonzalez-Cerrillo A, et al. Wide dispersion and diversity of clonally related inhibitory interneurons. *Neuron*. 2015;87(5):999–1007.
43. del Valle Rodriguez A, Didiano D, Desplan C. Power tools for gene expression and clonal analysis in Drosophila. *Nat Methods*. 2011;9(1):47–55.
44. Nern A, Pfeiffer BD, Rubin GM. Optimized tools for multicolor stochastic labeling reveal diverse stereotyped cell arrangements in the fly visual system. *Proc Natl Acad Sci U S A*. 2015;112(22):E2967–E2976.
45. Hadjiconomou D, Rotkopf S, Alexandre C, Bell DM, Dickson BJ, Salecker I. Flybow: genetic multicolor cell labeling for neural circuit analysis in Drosophila melanogaster. *Nat Methods*. 2011;8(3):260–266.
46. Wu JW, Turcotte R, Alt C, Runnels JM, Tsao H, Lin CP. Defining clonal color in fluorescent multi-clonal tracking. *Sci Rep*. 2016;6:24303.
47. Zong W, Wu R, Li M, et al. Fast high-resolution miniature two-photon microscopy for brain imaging in freely behaving mice. *Nat Methods*. 2017;14(7):713–719.
48. Du Z, Santella A, He F, Tionson M, Bao Z. De novo inference of systems-level mechanistic models of development from live-imaging-based phenotype analysis. *Cell*. 2014;156(1–2):359–372.
49. Barbosa JS, Sanchez-Gonzalez R, Di Giaimo R, et al. Neurodevelopment. Live imaging of adult neural stem cell behavior in the intact and injured zebrafish brain. *Science*. 2015;348(6236):789–793.
50. Pilz GA, Bottes S, Betizeau M, et al. Live imaging of neurogenesis in the adult mouse hippocampus. *Science*. 2018;359(6376):658–662.
51. Bonaguidi MA, Wheeler MA, Shapiro JS, et al. In vivo clonal analysis reveals self-renewing and multipotent adult neural stem cell characteristics. *Cell*. 2011;145(7):1142–1155.
52. Hope K, Bhatia M. Clonal interrogation of stem cells. *Nat Methods*. 2011;8(4 Suppl):S36–S40.
53. Gong H, Xu D, Yuan J, et al. High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. *Nat Commun*. 2016;7:12142.
54. Economou MN, Clack NG, Lavis LD, et al. A platform for brain-wide imaging and reconstruction of individual neurons. *Elife*. 2016;5:e10566.
55. Frieda KL, Linton JM, Hormoz S, et al. Synthetic recording and in situ readout of lineage information in single cells. *Nature*. 2017;541(7635):107–111.
56. Pei W, Feyerabend TB, Rossler J, et al. Polylox barcoding reveals haematopoietic stem cell fates realized in vivo. *Nature*. 2017;548(7668):456–460.
57. Chan Michelle, Smith ZD, Grosswendt Stefanie, et al. Molecular recording of mammalian embryogenesis. *bioRxiv*. 2018.
58. Shah S, Lubeck E, Zhou W, Cai L. In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron*. 2016;92(2):342–357.
59. Lai SL, Lee T. Genetic mosaic with dual binary transcriptional systems in Drosophila. *Nat Neurosci*. 2006;9(5):703–709.
60. Lee T, Lee A, Luo L. Development of the Drosophila mushroom bodies: sequential generation of three distinct types of neurons from a neuroblast. *Development*. 1999;126(18):4065–4076.
61. Jefferis GS, Marin EC, Stocker RF, Luo L. Target neuron prespecification in the olfactory map of Drosophila. *Nature*. 2001;414(6860):204.
62. Boone JQ, Doe CQ. Identification of Drosophila type II neuroblast lineages containing transit amplifying ganglion mother cells. *Dev Neurobiol*. 2008;68(9):1185–1195.
63. Bowman SK, Rolland V, Betschinger J, Kinsey KA, Emery G, Knoblich JA. The tumor suppressors Brat and Numb regulate transit-amplifying neuroblast lineages in Drosophila. *Dev Cell*. 2008;14(4):535–546.
64. Bello BC, Izergina N, Caussinus E, Reichert H. Amplification of neural stem cell proliferation by intermediate progenitor cells in Drosophila brain development. *Neural Dev*. 2008;3(1):1–18.
65. Jiang Y, Reichert H. Programmed cell death in type II neuroblast lineages is required for central complex development in the Drosophila brain. *Neural Dev*. 2012;7(1):3.
66. Awasaki T, Lai SL, Ito K, Lee T. Organization and postembryonic development of glial cells in the adult central brain of Drosophila. *J Neurosci*. 2008;28(51):13742–13753.
67. Izergina N, Balmer J, Bello B, Reichert H. Postembryonic development of transit amplifying neuroblast lineages in the Drosophila brain. *Neural Dev*. 2009;4(1):1–13.
68. Viktorin G, Riebli N, Popkova A, Giangrande A, Reichert H. Multipotent neural stem cells generate glial cells of the central complex through transit amplifying intermediate progenitors in Drosophila brain development. *Dev Biol*. 2011;356(2):553–565.
69. Lin S, Lai SL, Yu HH, Chihara T, Luo L, Lee T. Lineage-specific effects of Notch/ Numb signaling in post-embryonic development of the Drosophila brain. *Development*. 2010;137(1):43.
70. Truman JW, Moats W, Altman J, Marin EC, Williams DW. Role of Notch signaling in establishing the hemilineages of secondary neurons in Drosophila melanogaster. *Development*. 2010;137(1):53.
71. Bello BC, Hirth F, Gould AP. A pulse of the Drosophila Hox protein abdominal-A schedules the end of neural proliferation via neuroblast apoptosis. *Neuron*. 2003;37(2):209–219.
72. Komiyama T, Johnson WA, Luo L, Jefferis GSXE. From lineage to wiring specificity. POU domain transcription factors control precise connections of Drosophila olfactory projection neurons. *Cell*. 2003;112(2):157–167.
73. Liu Z, Steward R, Luo L. Drosophila Lis1 is required for neuroblast proliferation, dendritic elaboration and axonal transport. *Nat Cell Biol*. 2000;2(11):776–783.
74. Griffin R, Sustar A, Bonvin M, et al. The twin spot generator for differential Drosophila lineage analysis. *Nat Methods*. 2009;6(8):600–602.
75. Sustar A, Bonvin M, Schubiger M, Schubiger G. Drosophila twin spot clones reveal cell division dynamics in regenerating imaginal discs. *Dev Biol*. 2011;356(2):576–587.
76. Lin S, Chih-Fei K, Yu HH, Huang Y, Tzumin L. Lineage analysis of Drosophila lateral antennal lobe neurons reveals notch-dependent binary temporal fate decisions. *Plos Biol*. 2012;10(11):e1001425.
77. Yu HH, Chen CH, Shi L, Huang Y, Lee T. Twin-spot MARCM to reveal the developmental origin and identity of neurons. *Nat Neurosci*. 2009;12(7):947–953.
78. Enriquez J, Rio LQ, Blazeski R, et al. Differing strategies despite shared lineages of motor neurons and glia to achieve robust development of an adult neuropil in Drosophila. *Neuron*. 2018;97(3).
79. Hilchen CMV, Bustos AE, Giangrande A, Technau GM, Altenhein B. Pre-determined embryonic glial cells form the distinct glial sheaths of the Drosophila peripheral nervous system. *Development*. 2013;140(17):3657–3668.
80. Hilchen CV, Altenhein B. Tracing cells throughout development: insights into single glial cell differentiation. *Fly (Austin)*. 2014;8(2):86–90.
81. Papan C, Campos-Ortega JA. A clonal analysis of spinal cord development in the zebrafish. *Dev Genes Evol*. 1997;207(2):71–81.
82. Kimmel CB, Warga RM, Kane DA. Cell cycles and clonal strings during formation of the zebrafish central nervous system. *Development*. 1994;120(2):265–276.
83. He J, Zhang G, Almeida AD, Cayouette M, Simons BD, Harris WA. How variable clones build an invariant retina. *Neuron*. 2012;75(5):786–798.
84. Tang X, Gao J, Jia X, et al. Bipotent progenitors as embryonic origin of retinal stem cells. *J Cell Biol*. 2017;216(6):1833–1847.
85. Wan Y, Almeida AD, Rulands S, et al. The ciliary marginal zone of the zebrafish retina: clonal and time-lapse analysis of a continuously growing tissue. *Development*. 2016;143(7):1099–1107.
86. Alemay A, Florescu M, Baron CS, Peterson-Maduro J, van Oudenaarden A. Whole-organism clone tracing using single-cell sequencing. *Nature*. 2018;556(7699):108–112.
87. Di Donato V, De Santis F, Auer TO, et al. 2C-Cas9: a versatile tool for clonal analysis of gene function. *Genome Res*. 2016;26(5):681–692.
88. Sultan KT, Han Z, Zhang XJ, et al. Clonally related GABAergic interneurons do not randomly disperse but frequently form local clusters in the forebrain. *Neuron*. 2016;92(1):31–44.
89. Brown KN, Chen S, Han Z, et al. Clonal production and organization of inhibitory interneurons in the neocortex. *Science*. 2011;334(6055):480–486.
90. Mayer C, Bandler RC, Fishell G. Lineage is a poor predictor of interneuron positioning within the forebrain. *Neuron*. 2016;92(1):45–51.
91. Shi W, Xianyu A, Han Z, et al. Ontogenetic establishment of order-specific nuclear organization in the mammalian thalamus. *Nat Neurosci*. 2017;20(4):516–528.
92. Yu YC, Bultje RS, Wang X, Shi SH. Specific synapses develop preferentially among sister excitatory neurons in the neocortex. *Nature*. 2009;458(7237):501–504.
93. Walsh C, Cepko CL. Clonally related cortical cells show several migration patterns. *Science*. 1988;241(4871):1342–1345.
94. Rakic P. Radial unit hypothesis of neocortical expansion. *Novartis Found Symp*. 2000;228:30–42. discussion 42–52.
95. Magavi S, Friedmann D, Banks G, Stolfi A, Lois C. Coincident generation of pyramidal neurons and protoplasmic astrocytes in neocortical columns. *J Neurosci*. 2012;32(14):4762–4772.
96. Yu YC, He S, Chen S, et al. Preferential electrical coupling regulates neocortical lineage-dependent microcircuit assembly. *Nature*. 2012;486(7401):113–117.
97. Betizeau M, Cortay V, Patti D, et al. Precursor diversity and complexity of lineage relationships in the outer subventricular zone of the primate. *Neuron*. 2013;80(2):442–457.