



Computational Methods for Single-Cell Imaging and Omics Data Integration

Ebony Rose Watson, Atefeh Taherian Fard and Jessica Cara Mar**

Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, Brisbane, QLD, Australia

OPEN ACCESS

Edited by:

Reza M. Salek,
International Agency for Research on
Cancer (IARC), France

Reviewed by:

Oleg Mayboroda,
Leiden University Medical Center,
Netherlands
Gregorio Peron,
Ca' Foscari University of Venice, Italy

*Correspondence:

Atefeh Taherian Fard
a.taherianfard@uq.edu.au
Jessica Cara Mar
j.mar@uq.edu.au

Specialty section:

This article was submitted to
Molecular Diagnostics and
Therapeutics,
a section of the journal
Frontiers in Molecular Biosciences

Received: 14 September 2021

Accepted: 29 November 2021

Published: 17 January 2022

Citation:

Watson ER, Taherian Fard A and
Mar JC (2022) Computational
Methods for Single-Cell Imaging and
Omics Data Integration.
Front. Mol. Biosci. 8:768106.
doi: 10.3389/fmolb.2021.768106

Integrating single cell omics and single cell imaging allows for a more effective characterisation of the underlying mechanisms that drive a phenotype at the tissue level, creating a comprehensive profile at the cellular level. Although the use of imaging data is well established in biomedical research, its primary application has been to observe phenotypes at the tissue or organ level, often using medical imaging techniques such as MRI, CT, and PET. These imaging technologies complement omics-based data in biomedical research because they are helpful for identifying associations between genotype and phenotype, along with functional changes occurring at the tissue level. Single cell imaging can act as an intermediary between these levels. Meanwhile new technologies continue to arrive that can be used to interrogate the genome of single cells and its related omics datasets. As these two areas, single cell imaging and single cell omics, each advance independently with the development of novel techniques, the opportunity to integrate these data types becomes more and more attractive. This review outlines some of the technologies and methods currently available for generating, processing, and analysing single-cell omics- and imaging data, and how they could be integrated to further our understanding of complex biological phenomena like ageing. We include an emphasis on machine learning algorithms because of their ability to identify complex patterns in large multidimensional data.

Keywords: single cell imaging, single cell omics, data integration, machine learning, ageing

INTRODUCTION

Advances in high-throughput techniques have made it possible to collect largescale data from different types of regulatory information that controls a single cell. As a result, modelling approaches that combine multiple layers of cellular information deliver more informative insights than their single mode counterparts (Zhu et al., 2020). This review provides a comprehensive overview of the advanced technologies used for single cell imaging and omics sequencing, and the opportunities that exist to integrate these two types of data. We describe key advances in technologies and outline the major steps that are important for working with these two data types. Case studies are presented to illustrate some examples of integrating imaging and omics-level data. We emphasise the utility of this type of integration by focusing on studies that feature heterogeneous phenotypes in human health like ageing.

The substantial uptake of single cell-based technologies has been attractive in biomedical research because it is a known fact that human phenotypes are heterogeneous. Single cell omics methods like RNA-sequencing (scRNA-seq) have helped resolve this heterogeneity by providing a clearer

resolution of data so that pathways and master regulators can be identified with cell type-level specificity (Efremova and Teichmann, 2020). Single cell imaging methods like fluorescence microscopy have made it possible to acquire cellular features like morphology or cell area at such a high-throughput level that quantitative analyses can be done on populations of cells to investigate this heterogeneity (Marklein et al., 2018). Integrating these two types of technologies offers even more substantial benefits for inferring a more comprehensive model of cellular regulation. However, data integration brings additional challenges and single cell imaging and omics-level data have their own computational issues which is a focus of this review.

One of the major barriers to adopting high-throughput single-cell imaging lies with its computational requirements. For example, image processing, analysis and storage of the massive amount of data that is acquired during a simple imaging experiment are not trivial (Swedlow et al., 2021). For a typical research lab, this will usually require additional resources. These issues are compounded when integrating datasets collected from imaging and omics assays, which can drive the dimensions of the dataset into the hundreds of thousands, even whilst the number of biological samples remains small (Mirza et al., 2019). However, solutions are increasingly becoming more available and accessible through high-performance computing options on cloud platforms, along with high quality, open-source image processing and analysis software, and more efficient pipelines.

With the appropriate experimental assay and imaging technology, high-throughput cellular imaging can collect an impressive range of quantitative metrics that describe a single cell (Bray et al., 2016a). From quantifying basic morphological, intensity and textural features, to identifying the structure, number, and spatial distribution of sub-cellular elements such as organelles, proteins, and RNA sequences. Unlike omics technologies, in imaging many of these cellular features can even be measured in the same cell multiple times, giving insight into the spatiotemporal dynamics of single cells without having to destroy the cell (Nozaki et al., 2017). Single-cell imaging can also be used to explore the cause and effect relationship between specific genetic, chemical, and environmental perturbations and a variety of cellular phenotypes (Mattiuzzi Usaj et al., 2016). Consequently, microscopy remains the most informative tool for capturing associations and interactions between multiple molecular and cellular elements at high resolution.

Although the use of imaging data is well established in biomedical research, it has primarily been applied to observing phenotypes at the tissue or organ level, often using medical imaging techniques such as MRI, CT and PET (Shen et al., 2017). Such imaging has been complementary to omics-based data in biomedical research, where the goal is to identify associations between genotype and phenotype, along with functional changes at the tissue level (Antonelli et al., 2019). Now that omics and imaging techniques are becoming more accessible, it is feasible that single-cell imaging can act as an intermediary between these levels of information. As a result, integrating single-cell omics and single-cell imaging allows for a

more effective and comprehensive characterisation of the underlying mechanisms of a cellular phenotype.

All living organisms experience ageing, a phenomenon that is broadly defined as a gradual decline in physiological integrity, and consequently function, over the lifetime of an organism (López-Otín et al., 2013a). For humans, ageing can manifest through different symptoms, affecting a variety of organs and tissue types in a heterogeneous manner. Despite decades of research into practical and effective ageing interventions, advanced age remains the primary risk factor for many serious and chronic morbidities, including metabolic, cardiovascular, neoplastic, and neurodegenerative disorders (Niccoli and Partridge, 2012). From one individual to another, these age-associated pathologies vary in their severity and onset. Similarly, ageing within an individual is highly heterogeneous, with different tissues, cells and even cellular components that age according to different trajectories and rates.

Ageing is defined by a set of traits, termed the hallmarks of ageing (López-Otín et al., 2013b), which represent the key molecular and cellular components that are affected as organisms age. Once the level of damage within a cell reaches a certain threshold, it can initiate a cellular stress response known as senescence (Bhatia-Dey et al., 2016). Senescent cells secrete a variety of cytokines, chemokines, proteases, and other molecules that drive chronic inflammation in the tissue environment, leading to dysfunction and degradation that manifests as age-associated disease (Childs et al., 2015). Single-cell omics technologies have begun to provide insights into the mechanisms underlying senescence, sources of heterogeneity and the biological ageing process (Uyar et al., 2020). However, a complete picture cannot be formed without the addition of another technology: high-throughput cellular microscopy.

This review outlines the key methods currently available for the processing and analysis of single-cell omics and imaging data. We discuss how these data types can be used to further our understanding of biological processes, with a focus on applications in ageing. An emphasis has been included on machine learning algorithms, which can exceed human abilities in their capacity to identify extremely complex, subtle, and even sub-visual patterns in large multidimensional data. A range of post-hoc analysis methods can then be applied to extract meaningful biological information from these algorithms. We also explore how the integration of single-cell omics and single-cell imaging data using specific machine learning methods can exploit the distinct strengths of each technology to form a comprehensive understanding of ageing at the single-cell level.

AN OVERVIEW OF SINGLE-CELL OMICS TECHNOLOGIES

Historically, the sequencing methods that were used to capture genome-wide information required starting material that exceeded the amounts obtainable from a single cell. As a result, genomics and all of its related -omics technologies, have grown up in an era where information about the activity of genes and pathways has been obtained from mixtures of cells or

what is commonly referred to now as “bulk” samples. Measurements obtained from bulk samples result in the loss of cell-specific information because information from individual cells were averaged together to give a single, final data point. The transition from bulk to single cell-based approaches has had a major impact on genomics because it means that differences between cells can now be resolved rather than ignored as before. What used to be considered heterogeneity in data can now be clarified and sourced to differences in cell type or cell state because omics data can be captured for individual cells. The recent advances that have made single cell sequencing possible include improvements in single cell isolation, genome amplification, and barcoding which collectively have provided a platform to source information from different cellular and molecular levels without having to pool starting material. The current goal for genomics and its related technologies is to convert this information into actionable inferences that help describe the underlying biological mechanisms of different cells and tissue types.

Single cell-level omics data has also forced us to consider new implications, constraints and issues for the statistics that must be addressed for the analysis of this data. Because some of these statistical considerations are distinct from their counterparts for bulk data, it is necessary to adapt or invent new quantitative approaches that are appropriate for single cell data. For example, the most popular statistical approach for identifying differentially expressed genes for RNA-sequencing data is typically through an exact test for counts that have been fitted to a negative binomial distribution. Single cell RNA-sequencing data is more complex, with increased zeros and sometimes a multimodal distribution, and differential expression is typically assessed using a Wilcoxon signed-rank test (WSRT). The different statistic is necessary because of the increased heterogeneity in single cell data than in bulk data where the latter can be modelled more reliably with an approach that is based on parametric assumptions.

There are many other tasks in single cell data where the differences in the statistical approaches vary substantially from their bulk data counterparts. One prominent example is the data pre-processing and quality control pipelines where for single cell RNA-sequencing data, identifying low quality cells or detection of doublets are necessary for improving downstream data quality. Another example is clustering single cell RNA-sequencing data into groups of cells with similar expression profiles where the end goal is to identify cell types. This specific task has no direct parallel in bulk data because it is only at the level of single cells that information on cell types can be quantified.

Single-Cell Genomics

Single-cell genomics commonly refers to the capture of the DNA sequence of all genes in the genome of a single cell. In addition to identifying the genotype of a single cell's genes, this information enables the detection of rare and unique genomic alterations like single nucleotide polymorphisms (SNPs) and copy number variation (CNV). Understanding what genetic or genomic changes that occur in individual cells is instrumental to early detection of a disease such as in the case of an early-stage embryo with a genetic condition or identifying the spectrum of clonal

variation present in a tumour. For most genetic analysis studies, having access to an adequate quantity of high quality DNA is critical. There are various methods for amplifying the genome in preparation for single-cell whole genome sequencing (scWGS), for different applications including single SNP and CNV analysis. For example, multiple displacement amplification (MDA) method (Dean et al., 2002), can be carried out directly from biological samples and provides amplified DNA fragments that are uniformly represented across the genome. This method leverages the ϕ 29 DNA polymerase and random exonuclease-resistant primers in a simple isothermal reaction to amplify DNA strands with >10 kb in length. Similarly, MALBAC (Multiple Annealing and Looping Based Amplification Cycles) (Zong et al., 2012) provides amplified DNA through a series of temperature cycles, starting with melting genomic DNA into a single strand, random annealing of MALBAC primers to the DNA fragment, followed by extension to a semi- and then a full-amplicon. MALBAC claims to have a lower amplification bias as compared to methods with nonlinear amplification techniques like MDA and PCR-based methods. PicoPLEX (Rubicon Genomics PicoPLEX Kit) is a commercially-available whole genome amplification technology that performs DNA amplification of a single cell in a three-step single-tube reaction. Similar to the MALBAC method, the DNA template is denatured and pre-amplified using a quasi-random priming approach, creating a library of hairpin molecules that can be directly amplified into bulk quantities of DNA for further analysis (Table 1).

Single-cell genomics coupled with other single cell technologies can be used to construct information about the genome in a functional capacity and infer what molecular mechanisms underlie biological phenomena like cancer and development. For example, single-cell genomics has been extensively used in cancer to identify carcinogenic driver mutations, understand intratumor heterogeneity and its consequence on the transcriptome (Lim et al., 2020). In developmental biology, single-cell genome sequencing has been instrumental for reconstructing cellular ancestries in the form of a lineage tree (Hu et al., 2018a).

Numerous studies have used multi-omics sequencing to make the link between regulation of the genome with other omics at a single cell level (Lee et al., 2020). For example, Dey et al. (2015) used simultaneous sequencing of genomic DNA and mRNA from a single cell to investigate the correlation of CNVs to variability of the transcriptome in individual cells. They found that variations in CNV could potentially drive the gene expression variability observed in single cells.

Single-Cell Epigenomics

The epigenome is defined as the set of all changes occurring in a genome that does not involve alterations in DNA. Studying the epigenome therefore involves many different types of data depending on the specific epigenetic modification. For instance, a common type of epigenetic modification is DNA methylation where the addition of a methyl group to cytosine is a regulatory mechanism for controlling gene expression. Adaptations to bulk-level assays for capturing genome-wide

TABLE 1 | Single cell multi-omics technologies.

Single cell whole genome sequencing (scWGS)	
Multiple displacement amplification (MDA)	Dean et al. (2002)
Multiple Annealing and Looping Based Amplification Cycles (MALBAC)	Zong et al. (2012)
Rubicon Genomics PicoPLEX Kit	(Rubicon/Takara)
Single cell chromatin accessibility	
scDNase-seq (single-cell DNase sequencing)	Jin et al. (2015)
sci-ATAC-seq (single-cell combinatorial indexing assay for Transposase-accessible chromatin with sequencing)	Cusanovich et al. (2015)
scATAC-seq (single-cell assay for transposase-accessible chromatin using sequencing)	Buenrostro et al. (2015b)
NOMe-seq (nucleosome occupancy and methylation sequencing)	Kelly et al. (2012)
scMNase-seq (single-cell micrococcal nuclease sequencing)	Lai et al. (2018)
Single-cell epigenomics	
scRRBS (single-cell reduced representative bisulfite sequencing)	Guo et al. (2013)
scWGBS (single-cell whole-genome bisulfate sequencing)	Farlik et al. (2015)
snmC-seq (single-nucleus methylcytosine sequencing)	Luo et al. (2017)
sci-MET (single-cell combinatorial indexing for methylation)	Mulqueen et al. (2018)
Single-cell transcriptomic	
Drop-seq	Macosko et al. (2015)
Smart-seq/2 (switching mechanism at 5' end of the RNA transcript)	Goetz and Trimarchi, (2012); Picelli et al. (2014)
CEL-seq (Cell expression by linear amplification and sequencing)	Hashimshony et al. (2012)
MARS-seq	Jaitin et al. (2014)
STRT-seq	Islam et al. (2011)
Quartz-seq	Sasagawa et al. (2013)
SPLIT-seq	Rosenberg et al. (2018)
Single-cell proteomics	
MagRC (magnetic ranking cytometry)	Poudineh et al. (2017)
microengraving and droplet microfluidics	Haidas et al. (2019)
PLA (proximity ligation assay) and PEA (proximity extension assay)	Söderberg et al. (2006), Darmanis et al. (2016)
CytoF (cytometry by time of flight)	Bandura et al. (2009)
SCBCs (Single-cell barcode chips)	Labib and Kelley, (2020)
SCoPE-MS (Single Cell Proteomics by Mass Spectrometry)	Budnik et al. (2018); He et al. (2020)
Single-cell metabolomics	
ESI-MS (electrospray ionization mass spectrometry)	Huang et al., (2020); Li et al. (2020)
MALDI-MS (matrix-assisted laser desorption/ionization mass spectrometry)	Shanta et al., (2020)
SIMS (secondary ion mass spectrometry)	Leo et al., (2019)
Transcriptome with genome	
DR-seq (gDNA-mRNA sequencing)	Dey et al. (2015)
G&T-seq (genome and transcriptome sequencing)	Macaulay et al. (2015)
SIDR (simultaneous isolation of genomic DNA and total RNA)	Han et al. (2018)
TARGET-seq	Rodriguez-Meira et al. (2019)
scTrio-seq (single-cell triple omics sequencing)	Hou et al. (2016)
Transcriptome with epigenome	
sci-CAR (single-cell combinatorial indexing chromatin accessibility and mRNA)	Cao et al. (2018)
SNARE-seq (single-nucleus chromatin accessibility and mRNA expression sequencing)	Chen S. et al. (2019)
scNMT-seq (single-cell methylation and transcription sequencing)	Clark et al. (2018)
scM&T-seq (single-cell methylome and transcriptome sequencing)	Angermueller et al. (2016a)
scTrio-seq	Hou et al. (2016)
Transcriptome with proteome	
PEA/STA (proximity extension assay/specific RNA target amplification)	Genshaft et al. (2016)
PLAYR (proximity ligation assay for RNA)	Frei et al. (2016)
CITE-seq (cellular indexing of transcriptomes and epitopes by sequencing)	Stoeckius et al. (2017)
REAP-seq (RNA expression and protein sequencing assay)	Peterson et al. (2017)
RAID (single-cell RNA and immunodetection)	Gerlach et al. (2019)
ECCITE-seq (CRISPR-compatible cellular indexing of transcriptomes and epitopes by sequencing)	Mimitou et al. (2019)

DNA methylation events have resulted in the availability of single cell-level approaches to sequencing single cell DNA methylomes (Table 1).

Single-cell epigenomics sequencing provides insights into how the epigenome affects the transcriptome in a cell. There are

several single-cell epigenomics sequencing technologies that provide information on DNA modifications, DNA accessibility and chromosome conformation. For example, ATAC-seq (Assay of Transposase Accessible Chromatin sequencing) is an assay designed for detecting chromatin accessibility. In this method

hyperactive Tn5 transposases, loaded with sequencing adaptors, are probed in regions of open chromatin (*i.e.* nucleosome-free regions) and are used to generate sequencing libraries that can be amplified and sequenced (Buenrostro et al., 2015a). To capture *cis*-regulatory elements in individual cells, conventional ATAC-seq techniques have been implemented on droplet-based platforms for massively parallel sequencing and mapping transposase-accessible chromatin in tens of thousands of single cells (Yan et al., 2020). Other techniques include bisulfite sequencing that measures DNA methylation, Hi-C sequencing for measuring chromatin accessibility and chromosome conformation, and chromatin immune-precipitation that measures histone modifications and protein-DNA interaction (Lee et al., 2020).

Single-Cell Transcriptomics

The transcriptome is the set of all RNA transcripts, including coding (messenger RNA) and non-coding (such as microRNA and long non-coding RNA) which deliver information about protein-coding genes or RNA regulatory tasks, respectively. Single-cell transcriptomic technologies capture the gene expression levels of the transcriptome from thousands of single cells simultaneously (Hériché et al., 2019). The development of high-throughput protocols for single cell isolation and cell-specific barcoding technologies has enabled the generation of these datasets that allow cell-to-cell heterogeneity to be studied in a cellular population. Single-cell transcriptomic technologies have led to a host of new discoveries, including the detection of rare and new cell subtypes, the capture of cellular heterogeneity within a tissue, the identification of cellular states, and creating maps of developmental trajectories of specific cell types through pseudo temporal modelling and trajectory inference (Table 1).

Single cell methods differ in the strategies they adopt for individual protocol steps such as single cell isolation, library contraction and sequencing design as they are developed for different purposes. For example, Quartz-seq, MARS-seq and CEL-seq are UMI-based methods that measure transcripts at 3' end whereas Smart-seq and Smart-seq2 measure the full-length transcript (Ziegenhain et al., 2017; Lee et al., 2020). CEL-seq and Smart-seq use Fluidigm C1 (Wang and Navin, 2015) single cell isolation method while MARS-seq and Smart-seq2 use a FACS technique (Wang and Navin, 2015; Ziegenhain et al., 2017) (Table 1). Several largescale projects have been initiated to catalogue the comprehensive set of cell types in the human body (e.g. the Human Cell Atlas project) and to identify the spectrum of cell states at different stages of life (He et al., 2020; Lee et al., 2020).

Single-Cell Proteomics

Single-cell proteomics is one of the more recent areas of growth and new technologies to understand proteins at the single level and at scale are beginning to emerge. This is because unlike DNA and mRNA, proteins cannot be amplified. Nevertheless, there are several technologies that are mainly based on the applications of fluorescence-activated cell sorting (FACS), Western blotting, metal-tagged antibodies followed by mass cytometry to sort,

qualify phenotypes and high-multiplexing protein analysis (He et al., 2020). These methods are able to capture and analyse cell surface, cytoplasmic and secreted proteins (Labib and Kelley, 2020). For example, magnetic ranking cytometry (MagRC) (Poudineh et al., 2017) detects cell-surface proteins, while microengraving and droplet microfluidics (Haidas et al., 2019) detect the secreted protein. For cytoplasmic protein detection, methods include single-cell western blotting, proximity ligation assay (PLA) (Söderberg et al., 2006), proximity extension assay (PEA). Methods such as flow cytometry and single-cell barcode chips (SCBCs) are used for the analysis of proteins at all three cellular locations (Labib and Kelley, 2020). Although methods for single cell proteomics are mainly based on a limited number of proteins, the recently developed Single Cell Proteomics by Mass Spectrometry (SCoPE-MS) technique is able to detect more than 1000 proteins in a single cell (He et al., 2020; Budnik et al., 2018). It is worth highlighting that although innovations in mass spectrometry (MS) have improved the scope and scale of these technologies, as in the case of cytometry by time of flight (CyTOF (Bandura et al., 2009)), these methods are still not comparable to omics-level throughput.

Single-Cell Metabolomics

The aim of single-cell metabolite profiling is to study the effect of small molecules and metabolites in an epigenetic and transcriptomic profile in a single cell. Metabolites are arguably the end product of the basic central dogma process performed in the cell, providing a more immediate and holistic insight about the cellular phenotype. Metabolomics inform about the exact downstream effect and ultimate fates of the analytes, an information that other omics technologies fail to generate (Minakshi et al., 2019).

Screening single-cell metabolite profiles is challenging because these biomolecules have relatively short lifespans, are structurally diverse and chemically unstable *in vitro* (Minakshi et al., 2019; Zhu et al., 2021). However, refinements in the current single cell isolation techniques, mass spectrometry (MS) and high-throughput microfluidic-based methods have led to the detection of a limited number of metabolites present in the cell (Comi et al., 2017; Zhang and Vertes, 2018; Duncan et al., 2019; He et al., 2020). These methods include electrospray ionization mass spectrometry (ESI-MS) (Huang et al., 2020; Li et al., 2020), matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) (Shanta et al., 2020) and secondary ion mass spectrometry (SIMS) (Leo et al., 2019). Coupled with separation-based methods, MS is the most sensitive method for detecting a wide range of metabolites in a single cell. After the single cell is lysed, the complete cellular metabolome is separated by chromatography or electrophoresis on an automated platform such as a microfluidic device. The separated metabolites are then delivered to a MS platform for metabolite identification, quantitation, or downstream analysis (Minakshi et al., 2019). For a comprehensive review on single cell isolation strategies, sample preparation methods and single-cell metabolomics technologies refer to (Minakshi et al., 2019; Feng et al., 2020; Dueñas et al., 2021; Zhu et al., 2021).

INTEGRATIVE MULTI-OMICS SEQUENCING TECHNOLOGIES

Transcriptome With Genome

An important advantage of detecting multiple molecules from a single cell is that genotype-phenotype correlations can be accurately identified. These paired approaches can be used to link variation in genotype with their corresponding variation in transcriptional responses, and this information can be expanded to further applications like constructing lineage trees that map this variation. Detecting genomic mutations can also be performed with greater accuracy because they can be verified with corresponding mutations occurring in the RNA. Several methods have been developed for the simultaneous extraction and sequencing of the genome and transcriptome of a single cell (Table 1). These technologies differ in terms of how they capture cytoplasmic mRNA and nuclear DNA (genomic DNA). For example, scTrio-seq requires the cytoplasm and nucleus to be physically separated by centrifugation whereas G&T-seq separates poly-A-tailed mRNAs from gDNA using oligo-dT-coated magnetic beads. Next, the mRNA and gDNA will be independently amplified and sequenced using single-cell mono-omics sequencing technologies such as PicoPLEX (for gDNA) and Smart-seq2 (for mRNA). For further details, the characteristics of these technologies are summarised in (Hu et al., 2018b) and (Lee et al., 2020).

Transcriptome With Epigenome

Changes in DNA methylation and chromatin accessibility are directly linked to the regulation of gene expression. Advances in single-cell epigenomics and transcriptomics have now made it feasible to study how DNA methylation and histone modification vary with changes in transcription in a single cell (Clark et al., 2016). scM&T-seq (single-cell methylome and transcriptome sequencing) (Angermueller et al., 2016a) was the first method to be reported for combined DNA methylome and transcriptome analysis. Since then, other methods that combine the transcriptome with the epigenome have been developed, including scTrio-seq (Hou et al., 2016) which allows for the simultaneous profiling of DNA, methylome, genome and transcriptome within a single cell. A variety of methods exist where they differ in terms of the approaches that they adopt for isolating DNA and RNA and the subsequent mono-omics sequencing technology employed (Hu et al., 2018b; Lo and Zhou, 2018; Lee et al., 2020) (Table 1).

Transcriptome With Proteome

Methods that measure the transcriptome and proteome of a single-cell (Table 1) are designed for capturing proteins at different cellular locations and throughputs. For example, CITE-seq and REASP-seq can quantify cell-surface proteins with more than 80 antibodies and detect more than 20,000 genes in a single workflow (Hu et al., 2018b). RAID-seq on the other hand detects intracellular or phosphorylated proteins together with mRNA expression. ECCITE-seq is an extension of the CITE-seq method which provides a range of multi-modal

information including transcriptome, protein, clonotype, and CRISPR perturbation data at the single cell level (Mimitou et al., 2019; Lee et al., 2020). While the scale of single cell proteomics approaches is increasing with more modern innovations, it is important to recognize that the expectations for the throughput of these single-cell proteomics and integrated transcriptomic-proteomic approaches are not the same as for single-cell transcriptomic or epigenomic methods. At this stage, being able to capture single-cell level data for proteins is still only for smaller numbers of molecules at a time.

METHODS FOR MULTI-OMICS DATA ANALYSIS

All single-cell omics data are usually subjected to a variety of pre-processing steps that include alignment back to a reference, filtering to remove noise, and evaluation of quality control steps to assess overall reliability of the data. Subsequently, the data is subjected to a normalisation step which aims to reduce the amount of technical variation and thus increase the signal-to-noise ratio in the data. Other considerations for pre-processing of single cell data include detecting datapoints that may correspond to more than one cell, referred to as a doublet, and removing them from further analysis. Batch effects may induce patterns in the data that distract from studying genuine biological effects. The removal of these batch effects through different correction methods is therefore an important pre-processing step for this data type. Different statistical methods have been developed to address these pre-processing goals that are specific for their respective data type.

The applications of methods for sc-RNA data analysis have begun to evolve into a predictable workflow. These analysis steps include cell type identification from a heterogeneous cell population, regulatory-network based inference to identify regulatory relationship among marker genes, and cellular trajectory inference to study the temporal dynamics of the transcriptome during development or where cells may adopt one state along a continuum as they transition between states (Hwang et al., 2018; Lee et al., 2020).

Cell type identification from scRNA-seq data is mainly based on clustering methods (e.g. k-means, hierarchical, and graph-based) that operate off data that has been subjected to a dimensionality reduction (DR) technique. Principal component analysis (PCA) is a well-established unsupervised linear DR method. Other commonly used approaches are non-linear DR methods including t-distributed stochastic neighbour embedding (t-SNE) (van der Maaten and Hinton, 2008), locally linear embedding (LLE) (Roweis and Saul, 2000; Tenenbaum et al., 2000) and deep count autoencoder (DCA) (Eraslan et al., 2019). Among the frequently used packages for clustering and cell type annotation are Seurat (Stuart et al., 2019), SNN-cliq (Xu and Su, 2015), Garnett (Pliner et al., 2019) and SingleR (Aran et al., 2019). For an extensive review on cell type annotation and clustering methods refer to (Abdelaal et al., 2019; Wu and Zhang, 2020).

Cell trajectory inference involves ordering cells based on their transcription profile to identify continuous cell states and branch

points that represent key fate decisions along the trajectory. There is a plethora of trajectory inference packages with each relying on a different method and trajectory type (Saelens et al., 2019). For example, Monocle (Qiu et al., 2017) and SlingShot use a tree-based method (Street et al., 2018), PAGA (Wolf et al., 2019) uses a graph-based method, Wishbone (Setty et al., 2016) uses a bifurcation method whereas GPfates (Lönnerberg et al., 2017) is based on a multifurcation method.

Gene regulatory networks are important models for understanding the gene-gene and other types of interactions that control the transition from one cell type to another (Pratapa et al., 2020). Among the commonly used network-based inference methods that have been developed specifically for scRNA-seq data, some of the popular ones include the SCNS toolkit (Moignard et al., 2015), SCODE (Matsumoto et al., 2017) and SCENIC (Aibar et al., 2017).

Methods for single cell genomics and epigenomics analysis allow for the identification of genetic aberrations and epigenetic changes occurring at the single cell level (Gawad et al., 2016; Lee et al., 2020). Methods for identifying CNVs from scWGS data include Ginkgo (Garvin et al., 2015), baseqCNV (Fu et al., 2019), SCNV (Wang et al., 2018), SCCNV (Zhang et al., 2019), and SCOPE (Wang et al., 2019a). Moreover, several methods have been developed for the effective identification of SNVs from single cell whole genome sequencing data such as SCcaller (Dong et al., 2017), baseqSNV (Fu et al., 2019), MonoVar (Zafar et al., 2016), and SCAN-SNV (Luquette et al., 2019). Methods for identifying open chromatin sites and peak identification include chromVAR (Schep et al., 2017) and SCALE (Xiong et al., 2019), respectively. For an extensive review on these and other methods on multi-omics data analysis, we refer readers to (Hu et al., 2018b; Hwang et al., 2018; Chen H. et al., 2019; Saelens et al., 2019; Lee et al., 2020; Pratapa et al., 2020; Wu and Zhang, 2020).

SINGLE-CELL IMAGING

The growing need to visualise cellular elements at a molecular scale has driven rapid developments in all facets of microscopy imaging (Galler et al., 2014). Advances in single cell imaging have now gone beyond just visualising cells. Instead, identification and quantification of cellular and sub-cellular elements are routine. A variety of technologies have been developed or adapted to capture spatial, temporal, and morphological information at the single-cell sub-cellular level. For example, the spatial distribution of hundreds to thousands of unlabelled molecular species can be visualised at sub-cellular resolution with Imaging Mass Spectrometry (Buchberger et al., 2018). Cryo-electron microscopy has undergone a “resolution revolution,” where it is now capable of single-particle imaging at resolutions quickly approaching the sub-nanometre scale (Danev et al., 2019). Several imaging modes of atomic force microscopy have been developed to offer nanometre resolution imaging of structures in live cells, whilst simultaneously characterising mechanical, kinetic, thermodynamic and electrostatic properties (Dufrêne et al., 2017). Despite the rapid expansion of such sophisticated

instruments and technologies, optical microscopy has remained one of the foremost approaches in single-cell imaging, and as such will be the focus of this review.

Optical Microscopy

Optical microscopy has played a foundational role in the discovery and characterisation of biological structures, molecules, and processes since the 17th century. This type of technology remains popular due its simplicity, flexibility, and non-invasive nature (Masters, 2008). Although the core concept of utilising a light source and one or a series of lenses to generate magnified images remains, advances in optical and mechanical components have transformed the quality and functionality of optical microscopes considerably. Most notably, the automation of the sample preparation and image acquisition processes such as liquid handling, focusing, sample positioning and illumination and detection multiplexing, have transformed optical microscopes into sophisticated systems that are capable of imaging thousands to hundreds of thousands of samples at a single-cell resolution in a matter of hours (Lock and Strömblad, 2010; Mattiazzi Usaj et al., 2016; Mikami et al., 2018). These developments have also led to the incorporation of optical microscopes into other high-throughput single-cell technologies, as with imaging flow cytometers, enabling the collection of additional information on morphological, spatial, and textural features (Stavrakis et al., 2019). The quantity and diversity of cellular structures and biomolecules that can be specifically and sensitively identified within a single cell has also advanced significantly (Ozawa et al., 2013). These developments have enabled the systematic and quantitative investigation of single-cell biology with imaging data at similar scale and accessibility previously only seen in sequencing technologies, but with significant spatial and temporal information (Wollman and Stuurman, 2007).

These high-throughput microscopy systems and sophisticated labelling technologies can also be paired with large-scale systematic perturbations to provide insights into the influence of genetic or environmental factors on various cellular attributes (Boutros et al., 2015; Pegoraro and Misteli, 2017). Screening of comprehensive small molecule libraries is a common strategy for rapidly identifying and validating compounds in drug discovery and development (Bray et al., 2017; Boyd et al., 2020). Alternatively, chemical-genetic screens use libraries of characterised compounds, where the resulting phenotype (forward screening) or biological target (reverse screening) are known in advance (Choi et al., 2014). These screens facilitate the discovery of specific genes, proteins or pathways involved in cellular phenotypes of interest (Pegoraro and Misteli, 2017). Similarly, genetic screens utilise gene perturbation technologies such as RNAi (Schmidt et al., 2013) and CRISPR/Cas9 (Rauscher et al., 2017) to enable knockout, knockdown, or overexpression studies to target tens of thousands of genes at a time (Schuster et al., 2019).

Fluorescence Microscopy

In imaging, an investigation into complex aspects of cellular biology often starts with labelling for specific identification.

Depending on the study, a variety of biological attributes can be labelled, including certain cellular structures, organelles, macromolecules or even processes of interest. Fluorescence microscopy is an approach that offers excellent labelling specificity through the use of molecules called fluorophores, which have the capacity to absorb light of a specific wavelength and subsequently re-emit it at a longer wavelength. Paired with the properties of the fluorescence microscopes, high detection sensitivity can be achieved with minimal cell perturbation (Shashkova and Leake, 2017). There is also an increasing variety of fluorescence microscopy techniques available to suit a diverse range of applications where each come with their own trade-offs (Jensen, 2012; Combs and Shroff, 2017).

For example, confocal fluorescence (CFM) and light-sheet fluorescence (LSFM) microscopy are two techniques capable of producing high-resolution imaging of focal planes deep within samples, known as optical sectioning. This enables the reconstruction of three-dimensional cellular or subcellular structures in specimens, providing valuable spatial information (Long et al., 2012). Optical sectioning in CFM is achieved through the use of point-like illumination and detection pinholes that reject out-of-focus light. Whilst being highly cost-effective and accessible, CFM image-acquisition is slow, and produces moderate photo-bleaching and toxicity, as light must pass through the sample to reach the plane of interest (Jonkman and Brown, 2015). In comparison, LSFM performs high-speed optical sectioning by projecting a thin light sheet onto the sample from the side. This restricts illumination to the focal plane of interest, reducing photo-bleaching and toxicity significantly (Zagato et al., 2018). As a result, LSFM can perform high-resolution 3D imaging in live samples for long periods of time. Hof, Moreth (Hof et al., 2021) recently used LSFM to perform live imaging of the dynamic processes of organoid morphogenesis at the single-cell scale for up to 7 days. However, implementation of LSFM is substantially more challenging than CFM, including extensive and non-standard sample preparation (Zagato et al., 2018).

Super-resolution fluorescence microscopy (SRM), or nanoscopy are techniques that have the capacity to surpass the diffraction limit of optical resolution of approximately 200 nm are also available (Schermelleh et al., 2019). Several SRM techniques have the capacity to generate 2D and 3D images at a resolution of <50nm, with some reaching as high as <10 nm. Most SRM methods can also be successfully applied to live-cell imaging, with some approaches demonstrating a temporal resolution of only milliseconds (Balzarotti et al., 2017). SRM has already enabled the observation and quantification of *in situ* protein aggregation associated with various neuro-degenerative diseases, protein mobility within mitochondrial sub-compartments, and even the discovery of entirely new subcellular structures (Balzarotti et al., 2017). Several comprehensive reviews of SRM in cellular biology are available for further information (Sahl et al., 2017; Vangindertael et al., 2018; Schermelleh et al., 2019; Jacquemet et al., 2020).

Advanced Fluorescence Microscopy Techniques

The modification of fluorescence microscopy approaches has also created advanced techniques for the precise quantification of complex cellular dynamics in real time and at the nano-scale (De Los Santos et al., 2015). Data generated with these methods reveal insights into intra-cellular processes that are difficult to achieve with standard approaches. These techniques are highly tuned to specific applications through exploitation of specific fluorescence properties. For example, Fluorescence recovery after photobleaching (FRAP), Fluorescence Loss In Photobleaching (FLIP) and Fluorescence Localisation after Photobleaching (FLAP) all rely on the photobleaching of fluorophores that occurs due to the reactions between the fluorophore and the surrounding molecules during excitation (Ishikawa-Ankerhold et al., 2012). These techniques are commonly used to investigate molecular motility and diffusion, and explore the connections and molecular exchange happening between cellular compartments (Drummen, 2012).

Förster Resonance Energy Transfer (FRET) techniques are based on the distance-dependent transfer of excitation energy between a donor and an acceptor fluorophore, and can be adapted for an extensive variety of applications, including the motility, localisation, interactions and structural relationships of several molecular species (Algar et al., 2019). The application of this technique can facilitate the characterisation of complex processes such as signalling pathways or protein-folding dynamics (Kraimer et al., 2019). Fluorescence Lifetime Imaging Microscopy (FLIM) capitalises on the exponential decay in fluorescence emission after excitation, which is influenced by minute changes in the microenvironment such as pH, temperature or ion concentration (Datta et al., 2020). Many of these techniques provide complementary information, and as such are frequently applied in combination to yield comprehensive and rich imaging datasets of complex biological phenomena.

Labelling Strategies

Fluorophores commonly take the form of fluorescent proteins (Chudakov et al., 2010), synthetic organic molecules (Terai and Nagano, 2013), and fluorescent nanoparticles (Pratiwi et al., 2019), with assorted physiochemical properties to complement different labelling and microscopy techniques (Nienhaus and Nienhaus, 2017). An ongoing challenge of fluorescence microscopy is the limited capacity for *in situ* label multiplexing due to the broad excitation and emission spectra of many fluorophores, which results in bleed-through of signal between channels during imaging. As a result, only a small number of molecular targets can be imaged simultaneously in the same cell. To overcome this, the synthesis of new fluorescent labels with properties to extend the opportunities for effective multiplexing, such as increasingly narrow emission bands (Martino et al., 2019; Pandey and Bodas, 2020) or advanced optical encoding (Lin et al., 2018; Zhai et al., 2020) is a major area of focus, with fluorescent nanoparticles showing particular promise (Lee et al., 2018).

For both fixed and live cell imaging, preferential labelling can be employed to zoom in on certain cellular locations or types of molecules such as basic proteins, lipids, or nucleic acids. For example, the nucleus of live cells is commonly visualised using Hoechst 33342, a membrane-permeable dye which preferentially binds to AT-rich regions of double-stranded DNA (Chazotte, 2011). Fluorescent labelling of cellular components including membranes, organelles, cytoplasm, cytoskeleton, lysosomes, lipid droplets is similarly possible. Assays based on applying a combination of such stains, such as Cell Painting (Bray et al., 2016b), are popular for the generation of rich morphological profiles of single-cells at scale. Also available are fluorophores that report on particular chemical properties of the cellular environment, such as metal ions (Domaille et al., 2008), pH (Han and Burgess, 2010) or temperature (Okabe et al., 2018), often within specific compartments (Mizukami, 2017). Alternatively, when a certain molecule is of interest, fluorophores may be fused to a biomolecule, such as a protein, peptide, or nucleic acid, which acts as a specific probe for the target molecules. Common examples of this approach include immunofluorescence, Fluorescence *In Situ* Hybridization (FISH) and Genetically-Encoded labelling.

Immunofluorescence labelling uses antibodies with high specificity for a single target, typically a protein, as a probe (Joshi and Yu, 2017). This labelling technique is highly versatile, with an extensive range of commercially available fluorophore-labelled antibodies, which can be applied in different combinations to enable the labelling of several targets in a single cell (Buchwalow et al., 2005). Larger scale label-multiplexing can be achieved via performing cyclic immunofluorescence, whereby multiple rounds of labelling and imaging are conducted through the removal or inactivation of the fluorophore after each round (Wählby et al., 2002; Buchwalow et al., 2005; Ko et al., 2020). However, the applications of immunofluorescence for live cell imaging are generally limited to cell-surface or extra-cellular targets, as cells must be fixed and permeabilised before larger molecules such as antibodies are able to enter (Griffiths and Lucocq, 2014).

FISH techniques use fluorophore-labelled short nucleic acid sequences as the targeted probes of complementary RNA or DNA sequences (Huber et al., 2018). They are commonly applied to study genetic aberrations such as duplications, deletions, insertions, and translocations from the single gene to whole chromosome scale. Single-molecule FISH (smFISH) is a significant variation of FISH that allows for the accurate targeting and detection of individual RNA molecules, providing quantitative information on sub-cellular abundance, localisation and co-localisation of specific RNA sequences (Femino et al., 1998; Raj et al., 2008). smFISH can also be applied to many types of RNA molecules, including messenger RNA (mRNA) (Femino et al., 1998), long non-coding RNAs (Cabili et al., 2015), and ribosomal RNA (Buxbaum et al., 2014).

The smFISH techniques have been expanded further to accommodate greater scale in the number of molecules that can be detected within a single cell. For example, one adaptation called SeqFISH+ was able to capture *in situ* imaging of mRNAs for 10,000 genes in individual cells at high

resolution (Eng et al., 2019). Whilst FISH has been traditionally performed in fixed cells, CRISPR live-cell fluorescent *in situ* hybridization (LiveFISH) has recently been developed, enabling real-time imaging of DNA and RNA dynamics in live cells (Wang et al., 2019b).

Genetic encoding of labels, typically through fusion with the gene of a target protein at the DNA level, is a popular technique that ensures excellent target specificity *in vitro* and *in vivo*. GE labels may be intrinsically fluorescent proteins (Thorn, 2017) or tags designed to bind exogenous fluorophores with high specificity (Elia, 2021). Genetically-encoded labelling may also be used to label secondary targets, such as nucleic acids via RNA- or DNA-binding protein domains, or targeted to organelles of interest using specific protein localisation signals (Chudakov et al., 2010). Genetically-encoded sensors are also available for the visualisation and measurement of intra- and extra-cellular physiological, chemical and mechanical properties *in vivo* (Germond et al., 2016; Cost et al., 2019).

Single-cell imaging can also be conducted without the use of fluorescent labelling, using transmitted- or reflected-light microscopes. Label-free microscopy is a valuable technique for the study of cellular biology, offering greater simplicity and lower perturbation than many label-based methods, including fluorescence microscopy (Kasproicz et al., 2017). Furthermore, label-free imaging techniques typically offer distinct but complementary information to fluorescence microscopy, and as such the two techniques are often applied together (Figure 1). Brightfield microscopy creates a dark image on a light background as light is differentially absorbed, reflected, or refracted by biological structures. Moreover, a variety of techniques, such as darkfield, phase-contrast, polarised light, and differential interference contrast microscopy, have been developed with the capacity to enhance contrast optically, without compromising resolution, and resulting in detailed imaging of subcellular structures (Murphy and Davidson, 2012). The information that can be extracted from label-free images generated with such techniques is also expanding with the development of powerful computational algorithms. For example, a number of *in silico* labelling methods have been developed in recent years, with the capacity to predict multiplexed fluorescent labels in novel, unlabelled images with high accuracy in live and fixed cells (Christiansen et al., 2018). A recent model from Cheng, Fu (Cheng et al., 2021) predicts labels corresponding to the sub-cellular structures DNA, actin, endosome and the Golgi apparatus, as well as labels informing of cellular events such as proliferation and apoptosis. Similar models have been developed with the capacity to predict fluorescence labelling of 3D images (Ounkomol et al., 2018; Guo et al., 2020).

METHODS FOR SINGLE-CELL IMAGING DATA ANALYSIS

The development of sophisticated and automated methods for the processing and analysis of imaging data, typically via machine learning (ML) and computer vision (CV), has contributed

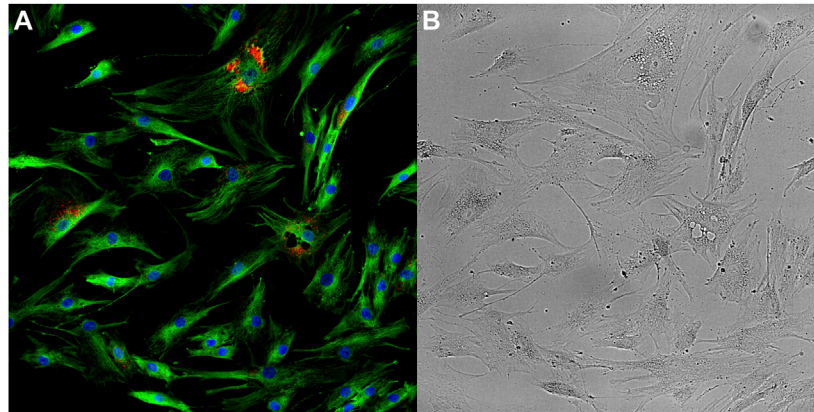


FIGURE 1 | Optical microscopy images taken of ageing mesenchymal stem cells. Fluorescence image **(A)** provides information on the abundance and distribution of DNA (blue), α -Tubulin (green) and Senescence-associated beta-galactosidase (red). Brightfield image **(B)** provides information on the cellular and sub-cellular morphology. Images have been enhanced for visualisation.

significantly to the increased popularity of biological image-based research (Danuser, 2011; Meijering, 2020). Despite this, the diversity of assays, techniques, and technologies available for generating cellular imaging data, along with the variability of experimental conditions and lack of standard imaging protocols, precludes the development of a ‘one-size-fits-all’ pipeline (Eliceiri et al., 2012). As such, only a broad overview of common approaches and generalised open source tools for the processing and analysis of fluorescence microscopy image data will be discussed in this review.

Some popular open source tools available for single-cell image data analysis include Icy (de Chaumont et al., 2012), ilastik (Berg et al., 2019), Microscopy Image Browser (Belevich et al., 2016), BioImageXD (Kankaanpää et al., 2012), Cytokit (Czech et al., 2019), KNIME (Fillbrunn et al., 2017), CellProfiler (McQuin et al., 2018) and FIJI (Schindelin et al., 2012) (ImageJ (Abràmoff et al., 2004)). The majority of these tools offer an implementation via a graphical or command line interface, and some in programming languages such as Python (Van Rossum, 2009) or R (R Core Team, 2020). Typically, these tools provide a variety of processing and analysis methods that can be “mixed and matched,” allowing the user to develop a customised pipeline to suit their specific needs. For example, CellProfiler includes over 70 independent modules designed for unique tasks, whilst there are several thousand modules available in the ImageJ ecosystem (McQuin et al., 2018). There are also a variety of powerful image processing libraries available in programming environments, including Scikit-image (van der Walt et al., 2014), Pillow (Clark, 2015) and OpenCV (Bradski, 2000) for Python, and EImage (Pau et al., 2010), imageHTS (Pau et al., 2020) in R. These, along with a variety of independently developed packages, can be applied in a similar manner for the development of a customised pipeline. Regardless of the nature of the interface, a conventional pipeline for single-cell imaging data consists of three main components: pre-processing for the correction of experimental or imaging artifacts, segmentation of the objects of interest, and an analysis of these objects.

Pre-Processing

The extent and specific methods applied for pre-processing of an image dataset will vary significantly depending on the type and quality of the images. Typically, all raw biological image data will require some form of denoising (Meinzel et al., 2018). A common source of systematic noise in microscopy imaging data is the presence of non-uniform illumination of the Field Of View (FOV), resulting from factors such as the light source, optical path, camera nonlinearity, or dust and staining artifacts. If left uncorrected, this non-uniformity can bias the measurements of properties of interest such as textural and intensity features, as well as interfere with the quality of processing steps downstream (Dey, 2019). The variety of illumination correction methods available is extensive (Singh et al., 2014; Smith et al., 2015; Peng et al., 2017; Nordenfelt et al., 2018), and are reviewed elsewhere for both general (Piccinini et al., 2012; Dey, 2019) and specific use cases (Liu et al., 2017). Other common pre-processing steps may include deconvolution to correct for signal blurring (Swedlow, 2013) and stitching and registration for samples split over multiple FOVs or imaged in multiple planes, wavelengths or modalities.

Object Segmentation

The accurate detection and segmentation of individual cells, or sub-cellular regions of interest, is an essential but challenging step in the quantitative analysis of cellular imaging data at the single-cell scale (Meijering, 2012). Traditional approaches to segmentation include thresholding (Otsu, 1979), feature detection (Kass et al., 1987) and watershed-based (Beucher and Meyer, 1993) methods. For particularly heterogeneous, noisy or complex datasets, machine learning models including U-Net (Falk et al., 2019), DeepCell (Van Valen et al., 2016), CDeep3M (Haberl et al., 2018), and CellPose (Stringer et al., 2021), are a popular choice. The segmentation of label-free images can be particularly challenging (Cameron et al., 2020; Liu et al., 2021a), and as such a number of methods have been developed specifically for this task (Vicar et al., 2019). A variety of segmentation methods designed for specific cell types (Li J et al.,

2019; Salvi et al., 2019) or datatypes, such as 3D images (Çiçek et al., 2016), are also available. The performance of segmentation methods have been reviewed and compared in detail elsewhere (Dima et al., 2011; Thomas and John, 2017; Caicedo et al., 2019; Cameron et al., 2020).

Analysis

Analysis of single-cell imaging data relies on extracting informative descriptors of phenotypic characteristics, or *features*, from the images. These features may be manually designed (handcrafted), and selected by the user, or automatically extracted from the data using machine learning algorithms, such as multi-layer artificial neural networks (ANNs). Using handcrafted features is often the more labour intensive approach, however they are also typically easier to interpret, and may even be defined in biologically meaningful terms such as cell membrane circularity or nuclei intensity (Caicedo et al., 2017). Automatically learned features usually take the form of abstract data representations, which are less intuitive, but may also more effectively capture the complexity of heterogenous and high dimensional datasets (Razavian et al., 2014).

These image-derived features describe phenotypic profiles of the system or condition under study (Caicedo et al., 2017; Grys et al., 2017), and are routinely utilised to group cells according to type (Zhang et al., 2017; Yao et al., 2019) or specific processes, such as phases of cell cycle (Eulenberg et al., 2017) via classification or clustering methods. Outlier detection methods can also be applied to identify rare or novel cell-types within heterogeneous populations (Mattiazzi Usaj et al., 2020). Phenotypic profiling of cellular responses to chemical (Kleinstreuer et al., 2014), environmental, and genetic (Rohban et al., 2017) perturbations is frequently applied for functional annotation and classification of the perturbants (Caicedo et al., 2016). Another common analysis is the quantification of the abundance and sub-cellular localisation of proteins (Pärnamaa and Parts, 2017) or RNA molecules (Samacoits et al., 2018) of interest. Other applications include lineage trajectory inference (Buggenthin et al., 2017), which commonly makes use of cell-tracking methods on live, long-term imaging data to accurately ascertain lineage progression (Piltti et al., 2018; Lugagne et al., 2020). Object tracking methods can be similarly applied to study subcellular dynamic processes, such as binding dynamics (Presman et al., 2017) or molecule trafficking (Chen et al., 2016), among others (Nketia et al., 2017; Brandão et al., 2021).

Multimodal Data Integration Techniques

Integrative approaches are commonly used for a range of different studies including classification (*e.g.* disease vs. normal), regression, annotation labelling (*e.g.* based on morphological or phenotypic descriptions), clustering, feature selection (biomarker discovery) and association studies. These studies share some strategies when categorising integrative approaches of multi-modal data. One strategy is to categorise the approaches into correlation analysis where the goal is to find correlations from the result obtained from the analysis of individual data types. Others include sequential analysis, where the analysis of one data type is followed by the integration of another data type), and integrative analysis where

integrative analysis of all data types are conducted to obtain an overall determination (**Figure 2**) (Lee et al., 2020).

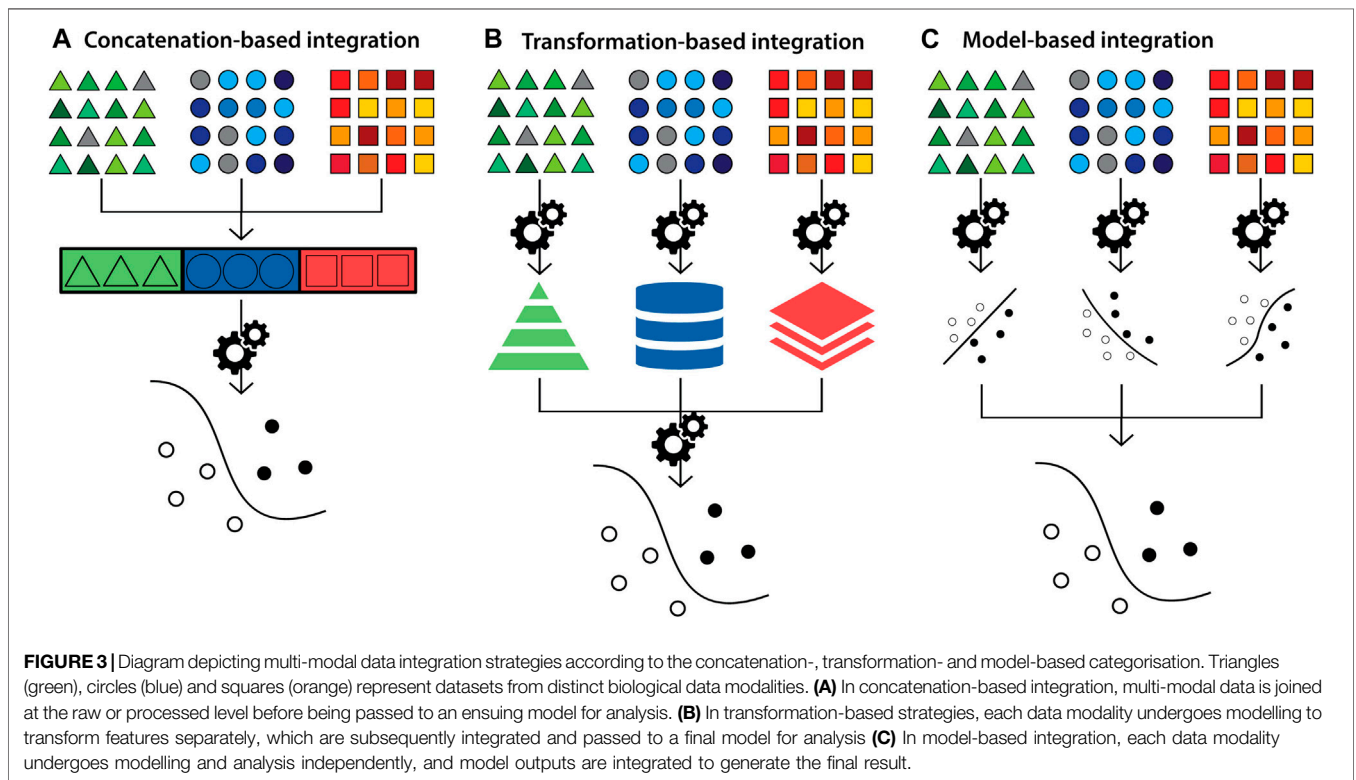
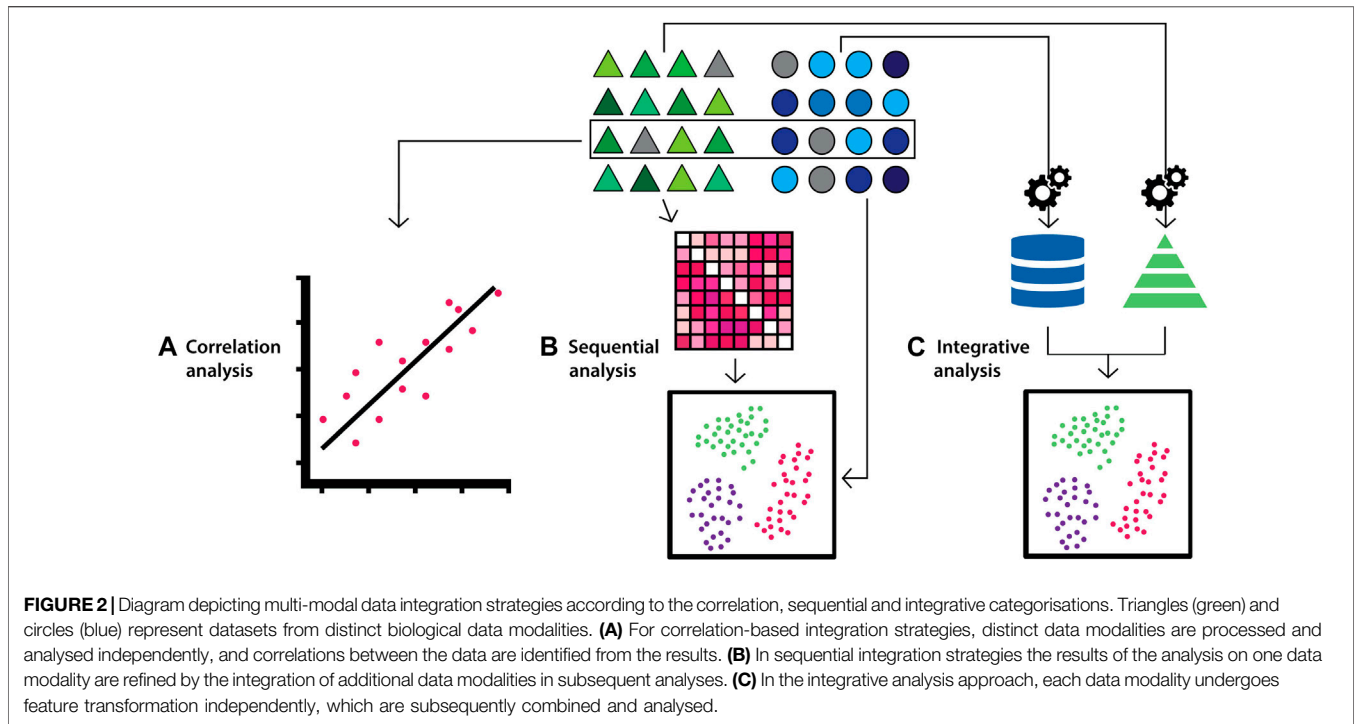
Another approach is to classify methods based on the strategies to build a multivariate final model. These methods are classified into concatenation-, transformation- and model-based integration (**Figure 3**) (Ritchie et al., 2015; Zitnik et al., 2019; Venugopalan et al., 2021). Concatenation-based classification involves combining datasets at the raw or processed level, followed by fitting into a supervised or unsupervised model and then analysis. Depending on the type of the data (*e.g.* images), the data is converted into a feature vector to be combined with other datasets. In transformation-based integration, the original data is transformed separately, and the modelling approach is applied at the level of the transformed matrices *i.e.* data types are integrated during the learning process. Model-based integration involves fitting separate models for individual data types and then combining their outputs to generate knowledge about the overall trait of interest (Ritchie et al., 2015; Venugopalan et al., 2021). The strengths and limitations of integration methods according to this classification strategy, and corresponding examples, are summarised in **Supplementary Table S1**. These methods are also referred to as early, intermediate and late integration, respectively (Li et al., 2016; Venugopalan et al., 2021).

Here we review the commonly-used methods for integrating images with omics data.

Annotation

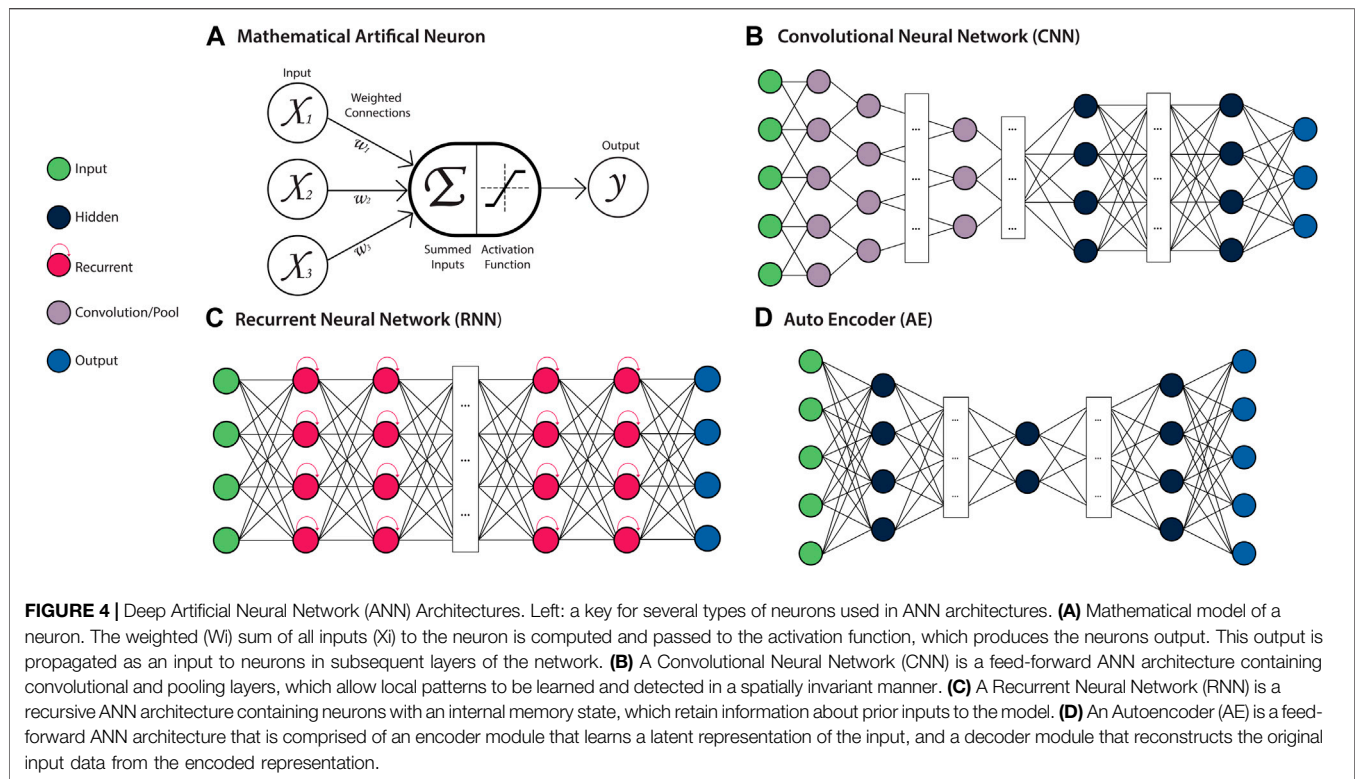
The first approach for integrating images with multi-omics data is to derive phenotypic information from imaging data, which is then utilised as annotations to aid in the interpretation of omics data. Thul, Åkesson (Thul et al., 2017) created an image-based map of the human subcellular proteome. They integrated transcriptomics data with high-resolution immunofluorescence microscopy images to determine the subcellular location of 12,003 proteins in various cell lines. Traditional image-derived annotations are usually manually curated in the form of morphological, biochemical, or physiological descriptions or measurements. Moreover, this information is also used for categorical classifications (*e.g.* presence or absence of a specific phenotype) (Hériché et al., 2019). For instance, in the context of cellular senescence, microscopy images of cells stained for senescence-specific markers such as senescence-associated beta-galactosidase (SA- β -gal) are often used to determine the presence of the senescent phenotype (Dimri et al., 1995). This determination can be further supported through quantifying the expression levels of proteins relevant to the senescent phenotype, such as cell-cycle arrest markers p21 and p16. Moreover, staining of the cellular membranes, cytoskeleton or cytoplasm provides morphological information. Cells present with a distinct morphology after transitioning into the senescent state, including enlarged and irregular cell shape, increased granularity and multinucleation (Biran et al., 2017).

Although traditional joint-analyses of multi-modal data are informative and relatively accurate, they are limited in their ability to identify patterns in complex biological data that often contain thousands of features. Therefore, features most relevant to each data type must first be identified and extracted



from the raw data before they can be integrated and analysed to draw biologically meaningful conclusions from them. However, as datasets increase in volume, dimensionality and heterogeneity, our ability to identify and extract meaningful features becomes

increasingly difficult and inefficient. This problem can be circumvented using more complex mathematical methods for multi-modal data representation and machine learning (ML) models to integrate multi-modal data.



Higher-Order Data Representation

The volume and complexity of data derived from images and multi-omics data brings the challenge of joining these data in an integrative framework (Hériché et al., 2019). Multi-modal integration methods look for patterns within and across data types, with or without prior knowledge (supervised or unsupervised) of the identity or label of the samples. Multiple high dimensional data can be incorporated and represented as higher order data structures or tensors (Chollet, 2018). Tensors then undergo dimensionality reduction to be integrated and jointly analysed (Li et al., 2016; Hériché et al., 2019). In the context of multimodal data integration, higher-order data representation and tensor factorisation methods have been used in the biological domain. For instance, Zhang, Liu (Zhang et al., 2012) used simultaneous non-negative matrix factorisation to integrate multi-omics cancer data. Argelaguet, Velten (Argelaguet et al., 2018) performed an integrative analysis of various biological data (drug response, mutation status, and transcriptome and DNA methylation profiles) using a joint matrix factorization approach formulated in a Bayesian framework. Last but not least, Acar, Papalexakis (Acar et al., 2014) performed a joint analysis of nuclear magnetic resonance and liquid chromatography–MS data using tensor factorisation.

Feature Vector Extraction

Both images and multi-omics data can be represented as numerical descriptors in the form of feature vectors (Hériché et al., 2019). Due to the high dimensional nature of the images and multi-omics data, it is often challenging to combine their

respective features in the original input space. Thus, new features from each data type can be extracted and then combined. Depending on the nature of the data, feature extraction methods such as matrix factorisation methods (e.g. PCA and NMF) or dimensionality reduction methods like autoencoders are applied (Li et al., 2016; Hériché et al., 2019). This is then followed by the classification or clustering on the combined features. The new features in the lower dimensional feature space are commonly numeric, providing a quantifiable measure of heterogeneity in each data mode and easy integration of their respective features. Feature vectors also provide a more efficient downstream analysis due to their reduced dimensionality (Li et al., 2016). Moreover, they can easily be incorporated into relational data (where the similarity between samples are known) by kernel feature extraction methods (Li and Ngom, 2014; Li et al., 2016).

Artificial Neural Networks

Artificial Neural Networks (ANN) are a class of ML algorithms that are based on many processing units (or “neurons”), typically organised into multiple layers which are inter-connected via edges to form a network (Figure 4A). These edges are assigned weights, which determine the strength of the connection between neurons and are adjusted throughout the network’s learning process to improve the model performance. The neurons of a network’s *input layer* contain the initialising data, which undergoes some transformation at the neurons of one or more *hidden layers*, followed by an *output layer* which produces the final result. The neurons contained within the

hidden layers compute the weighted sum of their inputs, apply an activation function, and produce the output (Angermueller et al., 2016b). The activation functions of neurons within hidden layers are typically non-linear, allowing inputs to be transformed in a manner that simultaneously increases the selectivity and invariance of the data representations (features) (LeCun et al., 2015). In “Deep” ANNs containing multiple hidden layers, the outputs from one layer act as input to the following layer. The compounding non-linearity allows for features of increasing complexity to be learned in a hierarchical manner as information progresses through the network. These features are optimised according to the specific task for which the model has been trained, typically classification, regression, or recognition (LeCun et al., 2015). As biological systems are inherently non-linear, this ability to generate intricate nonlinear input-output mappings is of great benefit for resolving the heterogeneity and complexity contained within biological data (Willy et al., 2003; Janson, 2012). The features learned by ANNs can also be extracted as feature vectors from the intermediate layers of the trained model, and subsequently combined for downstream integrative analyses (Chen et al., 2020). By utilising different layer types, neuron connections, activation functions, and learning rules, ANN architectures can be designed with a range of distinct behaviours and applications.

Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are a feed-forward ANN architecture designed to process input data in the form of multiple arrays (*i.e.* a tensor), making them particularly well-suited to processing raw image data, which usually takes the form of several two-dimensional arrays, representing each colour channel. They are also capable of processing sequence or signal data in the form of multiple one-dimensional arrays. CNNs are typically composed of multiple blocks of convolution and pooling layers which perform the feature learning task (Figure 4B). The convolutional layers contained within CNNs use arrays of weights (kernels) with a pre-defined shape to learn locally distinct patterns in the data through convolution operations. These patterns may represent edges or curves that form an object in an image, or a series of specific bases that form a transcription factor binding site in a genome sequence. The kernels are applied across the entirety of the data array, allowing these features to be detected in a spatially-invariant manner. Pooling layers perform down-sampling operations to merge semantically similar features, leading to robust feature detection and reduced model parameters (LeCun et al., 2015). The final layers of the CNN are fully-connected layers, where neurons are connected to every neuron in the previous layer, which map the learned features to the final output prediction. A more detailed explanation of CNNs can be found in several recent reviews (Gu et al., 2018; Khan et al., 2020).

Recurrent Neural Networks

Recurrent Neural Networks (RNNs) are a class of ANNs that are specialised for sequential data, such as DNA sequences or time series measurements. RNNs take a single element (*e.g.* an amino

acid in a protein sequence) as input at a time, allowing them to process sequences of variable length. The output generated by the neurons of the hidden layer for each element can then be passed as input to another neuron or looped directly back into that same neuron (Figure 4C). This cyclic processing allows the RNN to retain information pertaining to previous outputs in an internal ‘memory’, which is incorporated in the processing of the next element of the sequence. Accordingly, during each new cycle, the output of the hidden layer neurons is generated on the basis of both the new sequence element and the memory of previous sequence elements. As the memory capacity of the basic RNN architecture is relatively limited, a number of derivatives that have been developed to overcome this, including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures. These architectures are explored further in other reviews (Jozefowicz et al., 2015; Yu et al., 2019).

Autoencoder

One of the most popular ANN architectures applied for feature extraction is the Autoencoder (AE), which learns features in an unsupervised or ‘self-supervised’ manner. The task of an AE is to encode the input data into an internal representation through learning combinations of non-linear features, and then reconstruct the output from this encoding (Figure 4D). Through iteration, the AE model aims to find a codification of the data that enables the most accurate data reconstruction. This encoding can then be extracted to create a feature vector. AEs have been adapted to suit different data modalities through the incorporation of other ANN architectures, including convolutional AEs for multi-array data and LSTM autoencoders for sequence data (Charte et al., 2018). Various forms of regularisation can also be introduced to ensure the AE is learning a suitably meaningful encoding of the data, as is the case for sparse, denoising and contractive AEs (Zhai et al., 2018). Variational autoencoders (VAEs) are a class of AEs which aim to approximate the underlying distribution of the input data through implementing a variational Bayesian inference approach to encoding (Charte et al., 2018). The generative nature of VAEs make them particularly applicable to multi-modal data integration tasks (Simidjievski et al., 2019). AEs are covered in more detail in a number of recent reviews (Charte et al., 2018; Pulgar et al., 2020).

Transfer Learning

Transfer learning is the strategy of utilising knowledge learned by a previously trained ANN to enhance the performance of a new model with a different target domain or task. This approach is commonly applied for feature extraction, as ANNs trained on extremely large and diverse datasets tend to learn generic but high-quality features that are transferable across a variety of domain tasks (Pan and Yang, 2010). A number of high-performance models pre-trained on the ImageNet challenge dataset, consisting of 1.4 million images across 10,000 classes, have been utilised for feature extraction from biological imaging data with particular success (Russakovsky et al., 2015). For example, Khan *et al.* (Khan et al., 2019) extracted generic features from breast cytology images using three pre-trained

CNNs (GoogleNet (Szegedy et al., 2015), VGGNet (Simonyan and Zisserman, 2014), ResNet (He et al., 2016)), which enabled the detection and classification of malignant cells with an accuracy greater than 97% when combined.

These are but some of the ANN architectures most commonly utilised in multi-modal biological data integration studies. For an extensive review of ANNs and their biological applications, please refer to (LeCun et al., 2015; Angermueller et al., 2016b; Jones et al., 2017; Khamparia and Singh, 2019; Li Y et al., 2019; Tang et al., 2019; Emmert-Streib et al., 2020; Mahmud et al., 2021). Adaptations of many ANN architectures, including CNN, RNN and AE, that are designed to receive graph structured biological data such as gene regulatory networks as input are also available (Jin et al., 2021; Muzio et al., 2021).

CASE STUDIES

An attractive feature of AI is the ability to identify and extract informative patterns from complex, nonlinear data. Without the need for prior knowledge, AI unveils the mechanism underlying a complex biological process. Recently, ML and deep learning (DL) techniques have been developed and applied in many biomedical health and pharmaceutical-related fields (Gawehn et al., 2016; Mamoshina et al., 2016; Lenselink et al., 2017). These include, prediction of organic chemistry reactions (Wei et al., 2016), optimisation of chemical synthesis (Segler et al., 2018), prediction of pharmacological properties of drugs and drug repurposing (Aliper et al., 2016), modelling structural features of RNA-binding protein targets (Zhang et al., 2015), analysis of drug-induced liver injury (Xu et al., 2015), or the study of human long non-coding RNAs (Fan et al., 2015).

In the context of integrative analysis, depending on the nature of the task (classification, prediction, annotation, or marker discovery), the data types and the amount of data to handle, the constructed models from different ML algorithms are integrated into a single framework to capture the complex mechanism of biological systems. These frameworks are built based on different approaches and as such have different costs and benefits. Network-based fusion methods are able to infer direct or indirect associations in heterogeneous networks. Bayesian-based methods use prior information and model measurements in building the final model. Tree-based models make the final decision based on the trees constructed from individual or collective data types. Additionally, there is a range of deep ANNs that are used to integrate multi-modal data in a single framework (Bersanelli et al., 2016; Li et al., 2016). Here we discuss case-studies that have implemented commonly used frameworks for multi-modal data integration.

Kim et al. (Kim et al., 2013) used grammatical evolution neural network (GENN) to predict clinical outcomes for cancer patients by integrating gene copy number, DNA methylation, miRNA and gene expression data. Their computational platform ATHENA allows users to input multimodal omics data. In the first step, the noise variable from each genomic data is filtered out. Individual datasets then go through GENN modelling; the variables that best describe each genomic dataset are selected for the final GENN

modelling and integration. An advantage of this framework is its ability to model complex and non-linear relationships between variables, thus identifying interactions that influence variance in an outcome of interest. The final integrated model provides a global view of interaction within and between different levels of genomic data. They tested the final integrated framework on ovarian cancer data from the Cancer Genome Atlas and found that the identified interactions between multiple levels of genomics data are associated with an improved prognosis for ovarian cancer patients.

Chaudhary et al. (2018) used concatenation and DL to integrate mRNA expression, miRNA expression and DNA methylation data to improve clinical outcomes for patients with hepatocellular carcinoma. They implemented an AE model with three hidden layers. For each of the transformed features produced by the AE, they selected survival-associated features through a univariate Cox proportional hazards model. Next, they used these reduced new features to cluster the samples using the K-means clustering algorithm which led to the discovery of two subtypes with significant differences in survival. Furthermore, they validated these two subtypes in five independent cohorts which have an miRNA or mRNA or DNA methylation dataset.

In the context of single cell data integration, Tao et al. (2021) proposed a flexible framework, GLUER, for integrating single-cell omics and imaging data. After normalising the data for each modality, they employ a joint nonnegative matrix factorization (NMF) to identify common factor across data sets of different modalities while maintaining their biological differences. This is followed by using a mutual nearest neighbour (MNN) algorithm to map many-to-many relationships among cells across the data sets, generating factor loading matrices (dimensionality reduced matrices) for each data modality. One factor loading matrix is defined as a reference and the rest as query matrices. A distance between reference and query matrices is computed and used to determine the putative cell pairs between the two datasets. Finally they implement a CNN to learn nonlinear relationships between the factor loading matrices of reference and query datasets. The learnt functions are then used to co-embed the data by combining the reference factor loading matrix and query factor loading matrices.

Yang, Belyaeva (Yang et al., 2021) used AEs to integrate different single cell-sequencing modalities coupled with single cell-imaging data. Their study focused on identifying heterogenous cell states in human naïve CD4⁺ T-cells. In their framework, a different AE model is used to embed each of the data modalities into a shared latent space. The alignment and integration of each embedding within the latent space was performed using an adversarial training approach. Unlike other integration methods (Gundersen et al., 2020), this approach does not require paired data.

Stuart et al. (2019) used canonical correlation analysis (CCA) and MNN to develop a framework for reference assembly and transfer learning for transcriptomic, epigenomic, proteomic, and spatially-resolved single-cell data. First, they used CCA to jointly reduce the dimensionality of the reference and query datasets. These datasets originate from separate single cell experiments but

share cells from similar biological states. This is followed by identifying anchors (biologically-matched cells in a pair of datasets) using MNNs in the shared lower-dimensional space. Anchors encode the cellular relationships across datasets that will form the basis for all subsequent integration analyses. A score is assigned to each anchor pair based on the consistency of anchors across the neighbourhood structure of each dataset. Anchors and their score are then utilised to compute “correction” vectors for each query cell, transforming its expression so it can be jointly analysed as part of an integrated reference.

APPLICATIONS OF SINGLE-CELL AND INTEGRATIVE MULTI-MODAL DATA ANALYSIS IN AGEING STUDIES AND RELATED RESOURCES

Because ageing is a complex biological process, we have selected this particular area of biomedical and health, to showcase studies where multi-modal integration of data has had impact. The complexity of identifying regulators in ageing is due to the fact that ageing is influenced by genetic, epigenetic, transcription, metabolic and post-translation modifications. Systems-level multi-dimensional strategies are therefore required to capture the heterogeneity associated with the ageing phenotype. Due to their improved resolution and advancement, single-cell technologies allow for generation of largescale multi-modal data, which provides opportunities to integrate these datasets to inform our understanding of the mechanism of ageing and age-related disease.

Single cell multi-omics data have been used to discover novel cell types and cell state during ageing; detect cell population shifts and cell-state changes; identify tissue and cell-type specific genes and features; and identify ageing related genes in less abundant cell types (He et al., 2020). The results from these studies are applied in biomarker discovery, drug target identification, regenerative medicine, gene therapy, immune oncology and immunosenescence (Zhavoronkov et al., 2019). For example, Ma, Sun (Ma et al., 2020) created the first single-cell atlas of ageing and ageing interventions in rats that were subjected to a normal and caloric restriction (CR) diet. They studied the ageing-related changes in cell-type composition, gene expression and core transcription factors across tissues due to CR in young and aged rats. Zhang et al. (2019) utilised single-cell whole-genome sequencing to compare somatic mutations in human B lymphocytes in four age groups (newborn, adult, aged and centenarian). They found that somatic mutations increase from <500 per cell to >3,000 per cell across the human lifespan. For a comprehensive list of single cell omics studies in ageing refer to (He et al., 2020).

Single-cell imaging has been applied extensively to the discovery and characterisation of cell types and cell states associated with ageing and age-associated diseases. For example, Phillip et al. (2017) used a range of single-cell imaging technologies to quantify hundreds of biophysical and biomolecular properties of cells obtained from individuals

between 2 and 96 years of age. Based on these measurements they were able to identify key phenotypes associated with cellular ageing, such as reduced motility and increased cytoplasmic stiffness, which they used to develop a biological ageing clock. A number of models based on single-cell imaging data have also been developed for the identification of senescent cells (Oja et al., 2018; Kusumoto et al., 2021; Zhai et al., 2021). Wu et al. (2020) has also demonstrated that cellular morphology obtained from imaging data is predictive of the tumorigenic and metastatic potential of individual cells.

The functional annotation of genes linked to ageing and age-associated diseases has also been achieved via single-cell imaging. For instance, Jiao et al. (2019) performed an image-based genetic screen to construct morphological profiles of 125 genes from loci associated with Type-2 diabetes, adiposity, and insulin resistance. Clustering of these profiles revealed novel protein–protein and gene regulatory interactions relevant to Type-2 diabetes. High-throughput single-cell imaging is routinely applied for the discovery of therapeutic compounds to treat a range of age-associated diseases, including Alzheimer’s disease (Honarnejad et al., 2013), osteoarthritis (Nogueira-Recalde et al., 2019), Hutchinson–Gilford Progeria Syndrome (Kubben et al., 2016) and cancer (Caie et al., 2010; Moffat et al., 2014), as well as therapeutics for biological ageing as a whole (Sarkar et al., 2020), often through targeting cellular senescence (Fuhrmann-Stroissnigg et al., 2017).

Integrative multimodal analyses of biological imaging and omics data are popular in ageing-related research. Venugopalan et al. (2021) used deep AEs and CNNs to extract and integrate features from clinical, genomic and neurological imaging data to classify patients according to the severity of their Alzheimer’s disease stage. They also demonstrated that this multi-modal model outperformed single-modality models for the predictive task. Another common application of integrative analysis in age-associated disease is the integration of tissue-level imaging with genomic or transcriptomic data for the identification and classification of cancer sub-types (Shao et al., 2020; Liu et al., 2021b). Alternatively, Sailem and Bakal (2017) performed a correlation-based integrative analysis of single-cell morphology and bulk transcriptional data, finding that alterations in cell shape promoted breast cancer progression through the modulation of NF- κ B. Although these studies have typically been limited to the tissue level, the recent advances in single-cell technologies and computational methods discussed in this review hold great promise for enhancing our understanding of the molecular basis of the biological ageing process. For example, Meyer et al. (2020) recently developed a same-cell pharmacogenomics approach, *fate-seq*, which uses live imaging to predict the drug response of individual cells, that are subsequently isolated and profiled using single-cell RNA-seq. With this technique, they were able to identify the transcriptional profile responsible for modulating cancer-drug resistance.

There are many examples where AI has been successfully applied in longevity medicine research, including biomarker discovery (Putin et al., 2016; Moskalev et al., 2017; Zhavoronkov et al., 2021), using deep learning to predict

chronological age (Wang et al., 2017) and analysis of relationships between life-style traits (e.g. smoking) and accelerated ageing (Mamoshina et al., 2019). Please refer to the following reviews for a comprehensive overview on applications of AI in biomedicine (Fabris et al., 2017; Ching et al., 2018; Rifaioğlu et al., 2018; Tsigelny, 2018) and ageing research (Zhavoronkov et al., 2019; He et al., 2020).

The large amount of data generated in ageing research has been organised and disseminated in various databases. The publicly available databases consist of ageing phenotypes, longevity records, ageing- and senescence-related genes, and factors with lifespan-extending effects. These include Human Aging Genomic Resources (HAGR) containing GenAge, AnAge, GenDR, LongevityMap, DrugAge and CellAge (Tacutu et al., 2017). GenAge is a benchmark database for ageing- and longevity-associated genes. CellAge is a manually curated database of senescence-associated genes and DrugAge contains over 500 ageing-related drugs in model organisms. For more information about HAGR databases refer to (Tacutu et al., 2017). Other ageing-research related databases include Geroprotectors (Moskalev et al., 2015), AgeFactDB (Hühne et al., 2014), the Digital Ageing Atlas (Craig et al., 2015), AGEMAP (Zahn et al., 2007), SeneQuest (<https://senequest.net/>) by ICSA (International Cell Senescence Association). Last but not least, the Aging Atlas (Consortium, 2020) is a curated biomedical database which comprises of multi-omics datasets (sc-transcriptomics, epigenomics, proteomics and pharmacogenomics) and the tools to analyse and visualise the datasets.

CONCLUSION

In this review, we provide an overview of the current single-cell omics and imaging technologies, their respective methods for data analysis and common approaches for multi-modal data integration. While single-cell omics and imaging both represent two broad areas of interest, the intention of this review was not to provide an exhaustive treatment of these topics but instead offer a guide to help navigate the growing landscape of these two areas. We expect that the number of new techniques, data analysis approaches, and opportunities for integrating single-cell omics

REFERENCES

- Abdelaal, T., Michielsen, L., Cats, D., Hoogduin, D., Mei, H., Reinders, M. J. T., et al. (2019). A Comparison of Automatic Cell Identification Methods for Single-Cell RNA Sequencing Data. *Genome Biol.* 20 (1), 194. doi:10.1186/s13059-019-1795-z
- Abramoff, M. D., Magalhães, P. J., and Ram, S. J. (2004). Image Processing with ImageJ. *Biophotonics Int.* 11 (7), 36–42.
- Acar, E., Papalexakis, E. E., Gürdeniz, G., Rasmussen, M. A., Lawaetz, A. J., Nilsson, M., et al. (2014). Structure-revealing Data Fusion. *BMC Bioinformatics* 15 (1), 239. doi:10.1186/1471-2105-15-239
- Aibar, S., González-Blas, C. B., Moerman, T., Huynh-Thu, V. A., Imrichova, H., Hulselmans, G., et al. (2017). SCENIC: Single-Cell Regulatory Network Inference and Clustering. *Nat. Methods* 14 (11), 1083–1086. doi:10.1038/nmeth.4463

data with images will continue to grow and mature, and we hope that this review provides a reader, especially one who is a beginner to single cell biology, with enough content to learn about these areas more effectively and easily.

Single-cell omics technologies offer unprecedented opportunities to systematically explore cellular and molecular diversity at a single cell resolution. The data generated through these technologies have had a significant impact in understanding the heterogeneity in a cell population or tissue, leading to discovery of novel cell types, their function and their underlying genetic composition. Single cell imaging technologies capture morphological description of tissues and cells. Through the use of these technologies, we are also able to identify and quantify molecular profiles with single-molecule resolution. Advances in different single cell technologies that allow the capture of multiple features of a cell, in combination with the development of new multi-modal data integration approaches presented in this review are rapidly emerging and beginning to present promising results in different fields of biomedical research.

AUTHOR CONTRIBUTIONS

ERW and ATF researched the information that went into this manuscript. ERW, ATF, and JCM co-designed and co-wrote the manuscript.

FUNDING

JCM is funded by an Australian Research Council Future Fellowship (FT170100047).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmolb.2021.768106/full#supplementary-material>

- Algar, W. R., Hildebrandt, N., Vogel, S. S., and Medintz, I. L. (2019). FRET as a Biomolecular Research Tool - Understanding its Potential while Avoiding Pitfalls. *Nat. Methods* 16 (9), 815–829. doi:10.1038/s41592-019-0530-8
- Aliper, A., Plis, S., Artemov, A., Ulloa, A., Mamoshina, P., and Zhavoronkov, A. (2016). Deep Learning Applications for Predicting Pharmacological Properties of Drugs and Drug Repurposing Using Transcriptomic Data. *Mol. Pharmaceutics* 13 (7), 2524–2530. doi:10.1021/acs.molpharmaceut.6b00248
- Angermueller, C., Clark, S. J., Lee, H. J., Macaulay, I. C., Teng, M. J., Hu, T. X., et al. (2016). Parallel Single-Cell Sequencing Links Transcriptional and Epigenetic Heterogeneity. *Nat. Methods* 13 (3), 229–232. doi:10.1038/nmeth.3728
- Angermueller, C., Pärnamaa, T., Parts, L., and Stegle, O. (2016). Deep Learning for Computational Biology. *Mol. Syst. Biol.* 12 (7), 878. doi:10.15252/msb.20156651
- Antonelli, L., Guarracino, M. R., Maddalena, L., and Sangiovanni, M. (2019). Integrating Imaging and Omics Data: A Review. *Biomed. Signal Process. Control.* 52, 264–280. doi:10.1016/j.bspc.2019.04.032

- Aran, D., Looney, A. P., Liu, L., Wu, E., Fong, V., Hsu, A., et al. (2019). Reference-based Analysis of Lung Single-Cell Sequencing Reveals a Transitional Profibrotic Macrophage. *Nat. Immunol.* 20 (2), 163–172. doi:10.1038/s41590-018-0276-y
- Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J. C., et al. (2018). Multi-Omics Factor Analysis-A Framework for Unsupervised Integration of Multi-Omics Data Sets. *Mol. Syst. Biol.* 14 (6), e8124. doi:10.15252/msb.20178124
- Balzarotti, F., Eilers, Y., Gwosch, K. C., Gynnä, A. H., Westphal, V., Stefani, F. D., et al. (2017). Nanometer Resolution Imaging and Tracking of Fluorescent Molecules with Minimal Photon Fluxes. *Science* 355 (6325), 606–612. doi:10.1126/science.aak9913
- Bandura, D. R., Baranov, V. I., Ornatsky, O. I., Antonov, A., Kinach, R., Lou, X., et al. (2009). Mass Cytometry: Technique for Real Time Single Cell Multitarget Immunoassay Based on Inductively Coupled Plasma Time-Of-Flight Mass Spectrometry. *Anal. Chem.* 81 (16), 6813–6822. doi:10.1021/ac901049w
- Belevich, I., Joensuu, M., Kumar, D., Vihinen, H., and Jokitalo, E. (2016). Microscopy Image Browser: A Platform for Segmentation and Analysis of Multidimensional Datasets. *Plos Biol.* 14 (1), e1002340. doi:10.1371/journal.pbio.1002340
- Berg, S., Kutra, D., Kroeger, T., Straehle, C. N., Kausler, B. X., Haubold, C., et al. (2019). Ilastik: Interactive Machine Learning for (Bio)image Analysis. *Nat. Methods* 16 (12), 1226–1232. doi:10.1038/s41592-019-0582-9
- Bersanelli, M., Mosca, E., Remondini, D., Giampieri, E., Sala, C., Castellani, G., et al. (2016). Methods for the Integration of Multi-Omics Data: Mathematical Aspects. *BMC Bioinformatics* 17 Suppl 2 (2), 15. doi:10.1186/s12859-015-0857-9
- Beucher, S., and Meyer, F. (1993). The Morphological Approach to Segmentation: the Watershed Transformation. *Math. Morphol. image Process.* 34, 433–481.
- Bhatia-Dey, N., Kanherkar, R. R., Stair, S. E., Makarev, E. O., and Csoka, A. B. (2016). Cellular Senescence as the Causal Nexus of Aging. *Front. Genet.* 7, 13. doi:10.3389/fgene.2016.00013
- Biran, A., Zada, L., Abou Karam, P., Vadai, E., Roitman, L., Ovadya, Y., et al. (2017). Quantitative Identification of Senescent Cells in Aging and Disease. *Aging Cell* 16 (4), 661–671. doi:10.1111/acel.12592
- Boutou, M., Heigwer, F., and Laufer, C. (2015). Microscopy-Based High-Content Screening. *Cell* 163 (6), 1314–1325. doi:10.1016/j.cell.2015.11.007
- Boyd, J., Fennell, M., and Carpenter, A. (2020). Harnessing the Power of Microscopy Images to Accelerate Drug Discovery: what Are the Possibilities? *Expert Opin. Drug Discov.* 15 (6), 639–642. doi:10.1080/17460441.2020.1743675
- Bradski, G. (2000). *The OpenCV Library*. Dr. Dobb's Journal of Software Tools.
- Brandão, H. B., Gabriele, M., and Hansen, A. S. (2021). Tracking and Interpreting Long-Range Chromatin Interactions with Super-resolution Live-Cell Imaging. *Curr. Opin. Cell Biol.* 70, 18–26. doi:10.1016/j.celb.2020.11.002
- Bray, M.-A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., et al. (2016). Cell Painting, a High-Content Image-Based Assay for Morphological Profiling Using Multiplexed Fluorescent Dyes. *Nat. Protoc.* 11 (9), 1757–1774. doi:10.1038/nprot.2016.105
- Bray, M.-A., Singh, S., Han, H., Davis, C. T., Borgeson, B., Hartland, C., et al. (2016). Cell Painting, a High-Content Image-Based Assay for Morphological Profiling Using Multiplexed Fluorescent Dyes. *Nat. Protoc.* 11 (9), 1757–1774. doi:10.1038/nprot.2016.105
- Bray, M. A., Gustafsdottir, S. M., Rohban, M. H., Singh, S., Ljosa, V., Sokolnicki, K. L., et al. (2017). A Dataset of Images and Morphological Profiles of 30 000 Small-Molecule Treatments Using the Cell Painting Assay. *Gigascience* 6 (12), 1–5. doi:10.1093/gigascience/giw014
- Buchberger, A. R., DeLaney, K., Johnson, J., and Li, L. (2018). Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Anal. Chem.* 90 (1), 240–265. doi:10.1021/acs.analchem.7b04733
- Buchwalow, I. B., Minin, E. A., and Boecker, W. (2005). A Multicolor Fluorescence Immunostaining Technique for Simultaneous Antigen Targeting. *Acta Histochem.* 107 (2), 143–148. doi:10.1016/j.acthis.2005.01.003
- Budnik, B., Levy, E., Harmange, G., and Slavov, N. (2018). SCoPE-MS: Mass Spectrometry of Single Mammalian Cells Quantifies Proteomic Heterogeneity during Cell Differentiation. *Genome Biol.* 19 (1), 161. doi:10.1186/s13059-018-1547-5
- Buenrostro, J. D., Wu, B., Chang, H. Y., and Greenleaf, W. J. (2015a). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-wide. *Curr. Protoc. Mol. Biol.* 109 (1), 21–29. doi:10.1002/0471142727.mb2129s109
- Buenrostro, J. D., Wu, B., Litzenburger, U. M., Ruff, D., Gonzales, M. L., Snyder, M. P., et al. (2015b). Single-cell Chromatin Accessibility Reveals Principles of Regulatory Variation. *Nature* 523 (7561), 486–490. doi:10.1038/nature14590
- Buggenthin, F., Buettner, F., Hoppe, P. S., Endeke, M., Kroiss, M., Strasser, M., et al. (2017). Prospective Identification of Hematopoietic Lineage Choice by Deep Learning. *Nat. Methods* 14 (4), 403–406. doi:10.1038/nmeth.4182
- Buxbaum, A. R., Wu, B., and Singer, R. H. (2014). Single β -Actin mRNA Detection in Neurons Reveals a Mechanism for Regulating its Translatability. *Science* 343 (6169), 419–422. doi:10.1126/science.1242939
- Cabili, M. N., Dunagin, M. C., McClanahan, P. D., Biaisch, A., Padovan-Merhar, O., Regev, A., et al. (2015). Localization and Abundance Analysis of Human lncRNAs at Single-Cell and Single-Molecule Resolution. *Genome Biol.* 16 (1), 20. doi:10.1186/s13059-015-0586-4
- Caicedo, J. C., Cooper, S., Heigwer, F., Warchal, S., Qiu, P., Molnar, C., et al. (2017). Data-analysis Strategies for Image-Based Cell Profiling. *Nat. Methods* 14 (9), 849–863. doi:10.1038/nmeth.4397
- Caicedo, J. C., Roth, J., Goodman, A., Becker, T., Karhohs, K. W., Broisin, M., et al. (2019). Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry* 95 (9), 952–965. doi:10.1002/cyto.a.23863
- Caicedo, J. C., Singh, S., and Carpenter, A. E. (2016). Applications in Image-Based Profiling of Perturbations. *Curr. Opin. Biotechnol.* 39, 134–142. doi:10.1016/j.copbio.2016.04.003
- Caie, P. D., Walls, R. E., Ingleston-Orme, A., Daya, S., Houslay, T., Eagle, R., et al. (2010). High-content Phenotypic Profiling of Drug Response Signatures across Distinct Cancer Cells. *Mol. Cancer Ther.* 9 (6), 1913–1926. doi:10.1158/1535-7163.mct-09-1148
- Cameron, W. D., Bui, C. V., Bennett, A. M., Chang, H. H., and Rocheleau, J. V. (2020). Cell Segmentation Using Deep Learning: Comparing Label and Label-free Approaches Using Hyper-Labeled Image Stacks. *bioRxiv*. doi:10.1101/2020.01.09.900605v2
- Cao, J., Cusanovich, D. A., Ramani, V., Aghamirzaie, D., Pliner, H. A., Hill, A. J., et al. (2018). Joint Profiling of Chromatin Accessibility and Gene Expression in Thousands of Single Cells. *Science* 361 (6409), 1380–1385. doi:10.1126/science.aau0730
- Charte, D., Charte, F., García, S., del Jesus, M. J., and Herrera, F. (2018). A Practical Tutorial on Autoencoders for Nonlinear Feature Fusion: Taxonomy, Models, Software and Guidelines. *Inf. Fusion* 44, 78–96. doi:10.1016/j.inffus.2017.12.007
- Chaudhary, K., Poirion, O. B., Lu, L., and Garmire, L. X. (2018). Deep Learning-Based Multi-Omics Integration Robustly Predicts Survival in Liver Cancer. *Clin. Cancer Res.* 24 (6), 1248–1259. doi:10.1158/1078-0432.ccr-17-0853
- Chazotte, B. (2011). Labeling Nuclear DNA with Hoechst 33342. *Cold Spring Harb Protoc.* 2011 (1), pdb.prot5557. doi:10.1101/pdb.prot5557
- Chen, C., Zong, S., Wang, Z., Lu, J., Zhu, D., Zhang, Y., et al. (2016). Imaging and Intracellular Tracking of Cancer-Derived Exosomes Using Single-Molecule Localization-Based Super-resolution Microscope. *ACS Appl. Mater. Inter.* 8 (39), 25825–25833. doi:10.1021/acsami.6b09442
- Chen, H., Lareau, C., Andreani, T., Vinyard, M. E., Garcia, S. P., Clement, K., et al. (2019). Assessment of Computational Methods for the Analysis of Single-Cell ATAC-Seq Data. *Genome Biol.* 20 (1), 241. doi:10.1186/s13059-019-1854-5
- Chen, S., Lake, B. B., and Zhang, K. (2019). High-throughput Sequencing of the Transcriptome and Chromatin Accessibility in the Same Cell. *Nat. Biotechnol.* 37 (12), 1452–1457. doi:10.1038/s41587-019-0290-0
- Chen, R. J., Lu, M. Y., Wang, J., Williamson, D. F. K., Rodig, S. J., Lindeman, N. I., et al. (2020). Pathomic Fusion: An Integrated Framework for Fusing Histopathology and Genomic Features for Cancer Diagnosis and Prognosis. *IEEE Trans. Med. Imaging*.
- Cheng, S., Fu, S., Kim, Y. M., Song, W., Li, Y., Xue, Y., et al. (2021). Single-cell Cytometry via Multiplexed Fluorescence Prediction by Label-free Reflectance Microscopy. *Sci. Adv.* 7 (3), abe0431. doi:10.1126/sciadv.abe0431
- Childs, B. G., Durik, M., Baker, D. J., and van Deursen, J. M. (2015). Cellular Senescence in Aging and Age-Related Disease: from Mechanisms to Therapy. *Nat. Med.* 21 (12), 1424–1435. doi:10.1038/nm.4000
- Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., et al. (2018). Opportunities and Obstacles for Deep Learning in Biology and Medicine. *J. R. Soc. Interf.* 15 (141), 20170387. doi:10.1098/rsif.2017.0387

- Choi, H., Kim, J.-Y., Chang, Y. T., and Nam, H. G. (2014). Forward Chemical Genetic Screening. *Methods Mol. Biol.* 1062, 393–404. doi:10.1007/978-1-62703-580-4_21
- Chollet, F. (2018). *Deep Learning with Python. Vol. 361*. New York: Manning.
- Christiansen, E. M., Yang, S. J., Ando, D. M., Javaherian, A., Skibinski, G., Lipnick, S., et al. (2018). In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images. *Cell* 173 (3), 792–803. doi:10.1016/j.cell.2018.03.040
- Chudakov, D. M., Matz, M. V., Lukyanov, S., and Lukyanov, K. A. (2010). Fluorescent Proteins and Their Applications in Imaging Living Cells and Tissues. *Physiol. Rev.* 90 (3), 1103–1163. doi:10.1152/physrev.00038.2009
- Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, 424–432.
- Clark, A. (2015). *Pillow (PIL Fork) Documentation*. readthedocs.
- Clark, S. J., Lee, H. J., Smallwood, S. A., Kelsey, G., and Reik, W. (2016). Single-cell Epigenomics: Powerful New Methods for Understanding Gene Regulation and Cell Identity. *Genome Biol.* 17 (1), 72. doi:10.1186/s13059-016-0944-x
- Clark, S. J., Argelaguet, R., Kapourani, C.-A., Stubbs, T. M., Lee, H. J., Alda-Catalinas, C., et al. (2018). scNMT-seq Enables Joint Profiling of Chromatin Accessibility DNA Methylation and Transcription in Single Cells. *Nat. Commun.* 9 (1), 781. doi:10.1038/s41467-018-03149-4
- Combs, C. A., and Shroff, H. (2017). Fluorescence Microscopy: A Concise Guide to Current Imaging Methods. *Curr. Protoc. Neurosci.* 79, 2–25. doi:10.1002/cpns.29
- Comi, T. J., Do, T. D., Rubakhin, S. S., and Sweedler, J. V. (2017). Categorizing Cells on the Basis of Their Chemical Profiles: Progress in Single-Cell Mass Spectrometry. *J. Am. Chem. Soc.* 139 (11), 3920–3929. doi:10.1021/jacs.6b12822
- Consortium, A. A. (2020). Aging Atlas: a Multi-Omics Database for Aging Biology. *Nucleic Acids Res.* 49 (D1), D825–D830. doi:10.1093/nar/gkaa894
- Cost, A. L., Khalaji, S., and Grashoff, C. (2019). Genetically Encoded FRET-Based Tension Sensors. *Curr. Protoc. Cell Biol.* 83 (1), e85. doi:10.1002/cpcb.85
- Craig, T., Smelick, C., Tacutu, R., Wuttke, D., Wood, S. H., Stanley, H., et al. (2015). The Digital Ageing Atlas: Integrating the Diversity of Age-Related Changes into a Unified Resource. *Nucleic Acids Res.* 43, D873–D878. doi:10.1093/nar/gku843
- Cusanovich, D. A., Daza, R., Adey, A., Pliner, H. A., Christiansen, L., Gunderson, K. L., et al. (2015). Multiplex Single-Cell Profiling of Chromatin Accessibility by Combinatorial Cellular Indexing. *Science* 348 (6237), 910–914. doi:10.1126/science.aab1601
- Czech, E., Aksoy, B. A., Aksoy, P., and Hammerbacher, J. (2019). Cytokit: a Single-Cell Analysis Toolkit for High Dimensional Fluorescent Microscopy Imaging. *BMC Bioinformatics* 20 (1), 448. doi:10.1186/s12859-019-3055-3
- Danev, R., Yanagisawa, H., and Kikkawa, M. (2019). Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. *Trends Biochem. Sci.* 44 (10), 837–848. doi:10.1016/j.tibs.2019.04.008
- Danuser, G. (2011). Computer Vision in Cell Biology. *Cell* 147 (5), 973–978. doi:10.1016/j.cell.2011.11.001
- Darmanis, S., Gallant, C. J., Marinescu, V. D., Niklasson, M., Segerman, A., Flamourakis, G., et al. (2016). Simultaneous Multiplexed Measurement of RNA and Proteins in Single Cells. *Cel Rep.* 14 (2), 380–389. doi:10.1016/j.celrep.2015.12.021
- Datta, R., Heaster, T. M., Sharick, J. T., Gillette, A. A., and Skala, M. C. (2020). Fluorescence Lifetime Imaging Microscopy: Fundamentals and Advances in Instrumentation, Analysis, and Applications. *J. Biomed. Opt.* 25 (7), 1–43. doi:10.1117/1.JBO.25.7.071203
- de Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., et al. (2012). Icy: an Open Bioimage Informatics Platform for Extended Reproducible Research. *Nat. Methods* 9 (7), 690–696. doi:10.1038/nmeth.2075
- De Los Santos, C., Chang, C. W., Mycek, M. A., and Cardullo, R. A. (2015). FRET: FRAP, FLIM, and FRET: Detection and Analysis of Cellular Dynamics on a Molecular Scale Using Fluorescence Microscopy. *Mol. Reprod. Dev.* 82 (7–8), 587–604. doi:10.1002/mrd.22501
- Dean, F. B., Hosono, S., Fang, L., Wu, X., Faruqi, A. F., Bray-Ward, P., et al. (2002). Comprehensive Human Genome Amplification Using Multiple Displacement Amplification. *Proc. Natl. Acad. Sci.* 99 (8), 5261–5266. doi:10.1073/pnas.082089499
- Dey, N. (2019). Uneven Illumination Correction of Digital Images: A Survey of the State-Of-The-Art. *Optik* 183, 483–495. doi:10.1016/j.ijleo.2019.02.118
- Dey, S. S., Kester, L., Spanjaard, B., Bienko, M., and van Oudenaarden, A. (2015). Integrated Genome and Transcriptome Sequencing of the Same Cell. *Nat. Biotechnol.* 33 (3), 285–289. doi:10.1038/nbt.3129
- Dimas, A. A., Elliott, J. T., Filliben, J. J., Halter, M., Peskin, A., Bernal, J., et al. (2011). Comparison of Segmentation Algorithms for Fluorescence Microscopy Images of Cells. *Cytometry* 79A (7), 545–559. doi:10.1002/cyto.a.21079
- Dimri, G. P., Lee, X., Basile, G., Acosta, M., Scott, G., Roskelley, C., et al. (1995). A Biomarker that Identifies Senescent Human Cells in Culture and in Aging Skin *In Vivo*. *Proc. Natl. Acad. Sci.* 92 (20), 9363–9367. doi:10.1073/pnas.92.20.9363
- Domaille, D. W., Que, E. L., and Chang, C. J. (2008). Synthetic Fluorescent Sensors for Studying the Cell Biology of Metals. *Nat. Chem. Biol.* 4 (3), 168–175. doi:10.1038/nchembio.69
- Dong, X., Zhang, L., Milholland, B., Lee, M., Maslov, A. Y., Wang, T., et al. (2017). Accurate Identification of Single-Nucleotide Variants in Whole-Genome-Amplified Single Cells. *Nat. Methods* 14 (5), 491–493. doi:10.1038/nmeth.4227
- Drummen, G. (2012). Fluorescent Probes and Fluorescence (Microscopy) Techniques - Illuminating Biological and Biomedical Research. *Molecules* 17 (12), 14067–14090. doi:10.3390/molecules171214067
- Dueñas, M. E., and Lee, Y. J. (2021). “Single-Cell Metabolomics by Mass Spectrometry Imaging,” in *Cancer Metabolomics: Methods and Applications*. Editor S. Hu (Cham: Springer International Publishing), 69–82.
- Dufréne, Y. F., Ando, T., Garcia, R., Alsteens, D., Martinez-Martin, D., Engel, A., et al. (2017). Imaging Modes of Atomic Force Microscopy for Application in Molecular and Cell Biology. *Nat. Nanotech* 12 (4), 295–307. doi:10.1038/nnano.2017.45
- Duncan, K. D., Fyrestam, J., and Lanekoff, I. (2019). Advances in Mass Spectrometry Based Single-Cell Metabolomics. *Analyst* 144 (3), 782–793. doi:10.1039/c8an01581c
- Efremova, M., and Teichmann, S. A. (2020). Computational Methods for Single-Cell Omics across Modalities. *Nat. Methods* 17 (1), 14–17. doi:10.1038/s41592-019-0692-4
- Elia, N. (2021). Using Unnatural Amino Acids to Selectively Label Proteins for Cellular Imaging: a Cell Biologist Viewpoint. *Febs J.* 288 (4), 1107–1117. doi:10.1111/febs.15477
- Eliceiri, K. W., Berthold, M. R., Goldberg, I. G., Ibáñez, L., Manjunath, B. S., Martone, M. E., et al. (2012). Biological Imaging Software Tools. *Nat. Methods* 9 (7), 697–710. doi:10.1038/nmeth.2084
- Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., and Dehmer, M. (2020). An Introductory Review of Deep Learning for Prediction Models with Big Data. *Front. Artif. Intell.* 3, 4. doi:10.3389/fraci.2020.00004
- Eng, C.-H. L., Lawson, M., Zhu, Q., Dries, R., Kouloula, N., Takei, Y., et al. (2019). Transcriptome-scale Super-resolved Imaging in Tissues by RNA seqFISH+. *Nature* 568 (7751), 235–239. doi:10.1038/s41586-019-1049-y
- Eraslan, G., Simon, L. M., Mircea, M., Mueller, N. S., and Theis, F. J. (2019). Single-cell RNA-Seq Denoising Using a Deep Count Autoencoder. *Nat. Commun.* 10 (1), 390. doi:10.1038/s41467-018-07931-2
- Eulenbergh, P., Köhler, N., Blasi, T., Filby, A., Carpenter, A. E., Rees, P., et al. (2017). Reconstructing Cell Cycle and Disease Progression Using Deep Learning. *Nat. Commun.* 8, 463. doi:10.1038/s41467-017-00623-3
- Fabris, F., Magalhães, J. P. d., and Freitas, A. A. (2017). A Review of Supervised Machine Learning Applied to Ageing Research. *Biogerontology* 18 (2), 171–188. doi:10.1007/s10522-017-9683-y
- Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., et al. (2019). Author Correction: U-Net: Deep Learning for Cell Counting, Detection, and Morphometry. *Nat. Methods* 16 (4), 351. doi:10.1038/s41592-019-0356-4
- Fan, X.-N., and Zhang, S.-W. (2015). lncRNA-MFDL: Identification of Human Long Non-coding RNAs by Fusing Multiple Features and Using Deep Learning. *Mol. Biosyst.* 11 (3), 892–897. doi:10.1039/c4mb00650j
- Farlik, M., Sheffield, N. C., Nuzzo, A., Datlinger, P., Schönegger, A., Klughammer, J., et al. (2015). Single-cell DNA Methylome Sequencing and Bioinformatic Inference of Epigenomic Cell-State Dynamics. *Cel Rep.* 10 (8), 1386–1397. doi:10.1016/j.celrep.2015.02.001
- Femino, A. M., Fay, F. S., Fogarty, K., and Singer, R. H. (1998). Visualization of Single RNA Transcripts *In Situ*. *Science* 280 (5363), 585–590. doi:10.1126/science.280.5363.585
- Feng, D., Xu, T., Li, H., Shi, X., and Xu, G. (2020). Single-cell Metabolomics Analysis by Microfluidics and Mass Spectrometry: Recent New Advances. *J. Anal. Test.* 4 (3), 198–209. doi:10.1007/s41664-020-00138-9

- Fillbrunn, A., Dietz, C., Pfeuffer, J., Rahn, R., Landrum, G. A., and Berthold, M. R. (2017). KNIME for Reproducible Cross-Domain Analysis of Life Science Data. *J. Biotechnol.* 261, 149–156. doi:10.1016/j.jbiotec.2017.07.028
- Frei, A. P., Bava, F.-A., Zunder, E. R., Hsieh, E. W. Y., Chen, S.-Y., Nolan, G. P., et al. (2016). Highly Multiplexed Simultaneous Detection of RNAs and Proteins in Single Cells. *Nat. Methods* 13 (3), 269–275. doi:10.1038/nmeth.3742
- Fu, Y., Zhang, F., Zhang, X., Yin, J., Du, M., Jiang, M., et al. (2019). High-throughput Single-Cell Whole-Genome Amplification through Centrifugal Emulsification and eMDA. *Commun. Biol.* 2, 147. doi:10.1038/s42003-019-0401-y
- Fuhrmann-Stroissnigg, H., Ling, Y. Y., Zhao, J., McGowan, S. J., Zhu, Y., Brooks, R. W., et al. (2017). Identification of HSP90 Inhibitors as a Novel Class of Senolytics. *Nat. Commun.* 8, 422. doi:10.1038/s41467-017-00314-z
- Galler, K., Bräutigam, K., Große, C., Popp, J., and Neugebauer, U. (2014). Making a Big Thing of a Small Cell – Recent Advances in Single Cell Analysis. *Analyst* 139 (6), 1237–1273. doi:10.1039/c3an01939j
- Garvin, T., Aboukhalil, R., Kendall, J., Baslan, T., Atwal, G. S., Hicks, J., et al. (2015). Interactive Analysis and Assessment of Single-Cell Copy-Number Variations. *Nat. Methods* 12 (11), 1058–1060. doi:10.1038/nmeth.3578
- Gawad, C., Koh, W., and Quake, S. R. (2016). Single-cell Genome Sequencing: Current State of the Science. *Nat. Rev. Genet.* 17 (3), 175–188. doi:10.1038/nrg.2015.16
- Gawehn, E., Hiss, J. A., and Schneider, G. (2016). Deep Learning in Drug Discovery. *Mol. Inf.* 35 (1), 3–14. doi:10.1002/minf.201501008
- Genshaft, A. S., Li, S., Gallant, C. J., Darmanis, S., Prakadan, S. M., Ziegler, C. G. K., et al. (2016). Multiplexed, Targeted Profiling of Single-Cell Proteomes and Transcriptomes in a Single Reaction. *Genome Biol.* 17 (1), 188. doi:10.1186/s13059-016-1045-6
- Germond, A., Fujita, H., Ichimura, T., and Watanabe, T. M. (2016). Design and Development of Genetically Encoded Fluorescent Sensors to Monitor Intracellular Chemical and Physical Parameters. *Biophys. Rev.* 8 (2), 121–138. doi:10.1007/s12551-016-0195-9
- Gerlach, J. P., van Buggenum, J. A. G., Tanis, S. E. J., Hogeweg, M., Heuts, B. M. H., Muraro, M. J., et al. (2019). Combined Quantification of Intracellular (Phospho-)proteins and Transcriptomics from Fixed Single Cells. *Sci. Rep.* 9 (1), 1469. doi:10.1038/s41598-018-37977-7
- Goetz, J. J., and Trimarchi, J. M. (2012). Transcriptome Sequencing of Single Cells with Smart-Seq. *Nat. Biotechnol.* 30 (8), 763–765. doi:10.1038/nbt.2325
- Griffiths, G., and Lucocq, J. M. (2014). Antibodies for Immunolabeling by Light and Electron Microscopy: Not for the Faint Hearted. *Histochem. Cel Biol* 142 (4), 347–360. doi:10.1007/s00418-014-1263-5
- Grys, B. T., Lo, D. S., Sahin, N., Kraus, O. Z., Morris, Q., Boone, C., et al. (2017). Machine Learning and Computer Vision Approaches for Phenotypic Profiling. *J. Cel Biol* 216 (1), 65–71. doi:10.1083/jcb.201610026
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., et al. (2018). Recent Advances in Convolutional Neural Networks. *Pattern Recognition* 77, 354–377. doi:10.1016/j.patcog.2017.10.013
- Guo, H., Zhu, P., Wu, X., Li, X., Wen, L., and Tang, F. (2013). Single-cell Methylation Landscapes of Mouse Embryonic Stem Cells and Early Embryos Analyzed Using Reduced Representation Bisulfite Sequencing. *Genome Res.* 23 (12), 2126–2135. doi:10.1101/gr.161679.113
- Gundersen, G., Dumitrascu, B., Ash, J. T., and Engelhardt, B. E. (2020). “End-to-end Training of Deep Probabilistic CCA on Paired Biomedical Observations,” in *Proceedings of the 35th Uncertainty in Artificial Intelligence Conference. PMLR: Proceedings of Machine Learning Research*. Editors P. A. Ryan and G. Vibhav, 945–955.
- Guo, S. M., Yeh, L. H., Folkesson, J., Ivanov, I. E., Krishnan, A. P., Keefe, M. G., et al. (2020). Revealing Architectural Order with Quantitative Label-free Imaging and Deep Learning. *Elife* 9, e55502. doi:10.7554/eLife.55502
- Haberl, M. G., Churas, C., Tindall, L., Boassa, D., Phan, S., Bushong, E. A., et al. (2018). CDeep3M-Plug-and-Play Cloud-Based Deep Learning for Image Segmentation. *Nat. Methods* 15 (9), 677–680. doi:10.1038/s41592-018-0106-z
- Han, K. Y., Kim, K.-T., Joung, J.-G., Son, D.-S., Kim, Y. J., Jo, A., et al. (2018). SIDR: Simultaneous Isolation and Parallel Sequencing of Genomic DNA and Total RNA from Single Cells. *Genome Res.* 28 (1), 75–87. doi:10.1101/gr.223263.117
- Haidas, D., Bachler, S., Köhler, M., Blank, L. M., Zenobi, R., and Dittrich, P. S. (2019). Microfluidic Platform for Multimodal Analysis of Enzyme Secretion in Nanoliter Droplet Arrays. *Anal. Chem.* 91 (3), 2066–2073. doi:10.1021/acs.analchem.8b04506
- Han, J., and Burgess, K. (2010). Fluorescent Indicators for Intracellular pH. *Chem. Rev.* 110 (5), 2709–2728. doi:10.1021/cr900249z
- Hashimshony, T., Wagner, F., Sher, N., and Yanai, I. (2012). CEL-seq: Single-Cell RNA-Seq by Multiplexed Linear Amplification. *Cel Rep.* 2 (3), 666–673. doi:10.1016/j.celrep.2012.08.003
- He, K. M., Zhang, X., Ren, S., and Sun, J. (2016). “Deep Residual Learning for Image Recognition,” in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Las Vegas, NV, USA: IEEE), 770–778.
- He, X., Memczak, S., Qu, J., Belmonte, J. C. I., and Liu, G.-H. (2020). Single-cell Omics in Ageing: a Young and Growing Field. *Nat. Metab.* 2 (4), 293–302. doi:10.1038/s42255-020-0196-7
- Hériché, J.-K., Alexander, S., and Ellenberg, J. (2019). Integrating Imaging and Omics: Computational Methods and Challenges. *Annu. Rev. Biomed. Data Sci.* 2 (1), 175–197.
- Hof, L., Moreth, T., Koch, M., Liebisch, T., Kurtz, M., Tarnick, J., et al. (2021). Long-term Live Imaging and Multiscale Analysis Identify Heterogeneity and Core Principles of Epithelial Organoid Morphogenesis. *BMC Biol.* 19 (1). doi:10.1186/s12915-021-00958-w
- Honarnejad, K., Kirsch, A. K., Daschner, A., Szybinska, A., Kuznicki, J., and Herms, J. (2013). FRET-based Calcium Imaging. *J. Biomol. Screen.* 18 (10), 1309–1320. doi:10.1177/1087057113502672
- Hou, Y., Guo, H., Cao, C., Li, X., Hu, B., Zhu, P., et al. (2016). Single-cell Triple Omics Sequencing Reveals Genetic, Epigenetic, and Transcriptomic Heterogeneity in Hepatocellular Carcinomas. *Cell Res* 26 (3), 304–319. doi:10.1038/cr.2016.23
- Hu, Y., An, Q., Sheu, K., Trejo, B., Fan, S., and Guo, Y. (2018). Single Cell Multi-Omics Technology: Methodology and Application. *Front Cel Dev Biol* 6 (28), 28. doi:10.3389/fcell.2018.00028
- Hu, Y., An, Q., Sheu, K., Trejo, B., Fan, S., and Guo, Y. (2018). Single Cell Multi-Omics Technology: Methodology and Application. *Front. Cel Dev. Biol.* 6, 28. doi:10.3389/fcell.2018.00028
- Huang, Q., Mao, S., Khan, M., Li, W., Zhang, Q., and Lin, J.-M. (2020). Single-cell Identification by Microfluidic-Based *In Situ* Extracting and Online Mass Spectrometric Analysis of Phospholipids Expression. *Chem. Sci.* 11 (1), 253–256. doi:10.1039/c9sc05143k
- Huber, D., Voith von Voithenberg, L., and Kaigala, G. V. (2018). Fluorescence *In Situ* Hybridization (FISH): History, Limitations and what to Expect from Micro-scale FISH? *Micro Nano Eng.* 1, 15–24. doi:10.1016/j.mne.2018.10.006
- Hühne, R., Thalheim, T., and Sühnel, J. (2014). AgeFactDB—the JenAge Ageing Factor Database—Towards Data Integration in Ageing Research. *Nucleic Acids Res.* 42, D892–D896.
- Hwang, B., Lee, J. H., and Bang, D. (2018). Single-cell RNA Sequencing Technologies and Bioinformatics Pipelines. *Exp. Mol. Med.* 50 (8), 96. doi:10.1038/s12276-018-0071-8
- Ishikawa-Ankerhold, H. C., Ankerhold, R., and Drummen, G. P. C. (2012). Advanced Fluorescence Microscopy Techniques—FRAP, FLIP, FLAP, FRET and FLIM. *Molecules* 17 (4), 4047–4132. doi:10.3390/molecules17044047
- Islam, S., Kjällquist, U., Moliner, A., Zajac, P., Fan, J.-B., Lönnnerberg, P., et al. (2011). Characterization of the Single-Cell Transcriptional Landscape by Highly Multiplex RNA-Seq. *Genome Res.* 21 (7), 1160–1167. doi:10.1101/gr.110882.110
- Jacquemet, G., Carisey, A. F., Hamidi, H., Henriques, R., and Letierrier, C. (2020). The Cell Biologist’s Guide to Super-resolution Microscopy. *J. Cel Sci* 133 (11). doi:10.1242/jcs.240713
- Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., et al. (2014). Massively Parallel Single-Cell RNA-Seq for Marker-free Decomposition of Tissues into Cell Types. *Science* 343 (6172), 776–779. doi:10.1126/science.1247651
- Janson, N. B. (2012). Non-linear Dynamics of Biological Systems. *Contemp. Phys.* 53 (2), 137–168. doi:10.1080/00107514.2011.644441
- Jensen, E. C. (2012). Types of Imaging, Part 2: an Overview of Fluorescence Microscopy. *Anat. Rec.* 295 (10), 1621–1627. doi:10.1002/ar.22548
- Jiao, Y., Ahmed, U., Sim, M. F. M., Bejar, A., Zhang, X., Talukder, M. M. U., et al. (2019). Discovering Metabolic Disease Gene Interactions by Correlated Effects on Cellular Morphology. *Mol. Metab.* 24, 108–119. doi:10.1016/j.molmet.2019.03.001

- Jin, W., Tang, Q., Wan, M., Cui, K., Zhang, Y., Ren, G., et al. (2015). Genome-wide Detection of DNase I Hypersensitive Sites in Single Cells and FFPE Tissue Samples. *Nature* 528 (7580), 142–146. doi:10.1038/nature15740
- Jin, S., Zeng, X., Xia, F., Huang, W., and Liu, X. (2021). Application of Deep Learning Methods in Biological Networks. *Brief Bioinform* 22 (2), 1902–1917. doi:10.1093/bib/bbaa043
- Jones, W., Alasoo, K., Fishman, D., and Parts, L. (2017). Computational Biology: Deep Learning. *Emerg. Top. Life Sci.* 1 (3), 257–274. doi:10.1042/etsl20160025
- Jonkman, J., and Brown, C. M. (2015). Any Way You Slice It-A Comparison of Confocal Microscopy Techniques. *J. Biomol. Tech.* 26 (2), 54–65. doi:10.7171/jbt.15-2602-003
- Joshi, S., and Yu, D. (2017). “Immunofluorescence,” in *Basic Science Methods for Clinical Researchers*, 135–150. doi:10.1016/b978-0-12-803077-6.00008-4
- Jozefowicz, R., Zaremba, W., and Sutskever, I. (2015). “An Empirical Exploration of Recurrent Network Architectures,” in *International Conference on Machine Learning* (Lille, France: PMLR).
- Kankaanpää, P., Paavolainen, L., Tiitta, S., Karjalainen, M., Päivärinne, J., Nieminen, J., et al. (2012). BioImageXD: an Open, General-Purpose and High-Throughput Image-Processing Platform. *Nat. Methods* 9 (7), 683–689. doi:10.1038/nmeth.2047
- Kaspruwicz, R., Suman, R., and O’Toole, P. (2017). Characterising Live Cell Behaviour: Traditional Label-free and Quantitative Phase Imaging Approaches. *Int. J. Biochem. Cel Biol.* 84, 89–95. doi:10.1016/j.biocel.2017.01.004
- Kass, M., Witkin, A., and Terzopoulos, D. (1987). Snakes - Active Contour Models. *Int. J. Comput. Vis.* 1 (4), 321–331.
- Kelly, T. K., Liu, Y., Lay, F. D., Liang, G., Berman, B. P., and Jones, P. A. (2012). Genome-wide Mapping of Nucleosome Positioning and DNA Methylation within Individual DNA Molecules. *Genome Res.* 22 (12), 2497–2506. doi:10.1101/gr.143008.112
- Khamparia, A., and Singh, K. M. (2019). A Systematic Review on Deep Learning Architectures and Applications. *Expert Syst.* 36 (3), 12400. doi:10.1111/essy.12400
- Khan, A., Sohail, A., Zahoor, U., and Qureshi, A. S. (2020). A Survey of the Recent Architectures of Deep Convolutional Neural Networks. *Artif. Intell. Rev.* 53 (8), 5455–5516. doi:10.1007/s10462-020-09825-6
- Khan, S., Islam, N., Jan, Z., Ud Din, I., and Rodrigues, J. J. P. C. (2019). A Novel Deep Learning Based Framework for the Detection and Classification of Breast Cancer Using Transfer Learning. *Pattern Recognition Lett.* 125, 1–6. doi:10.1016/j.patrec.2019.03.022
- Kim, D., Li, R., Dudek, S. M., and Ritchie, M. D. (2013). ATHENA: Identifying Interactions between Different Levels of Genomic Data Associated with Cancer Clinical Outcomes Using Grammatical Evolution Neural Network. *BioData Mining* 6 (1), 23. doi:10.1186/1756-0381-6-23
- Kleinstreuer, N. C., Yang, J., Berg, E. L., Knudsen, T. B., Richard, A. M., Martin, M. T., et al. (2014). Phenotypic Screening of the ToxCast Chemical Library to Classify Toxic and Therapeutic Mechanisms. *Nat. Biotechnol.* 32 (6), 583–591. doi:10.1038/nbt.2914
- Ko, J., Oh, J., Ahmed, M. S., Carlson, J. C. T., and Weissleder, R. (2020). Ultra-fast Cycling for Multiplexed Cellular Fluorescence Imaging. *Angew. Chem. Int. Ed.* 59 (17), 6839–6846. doi:10.1002/anie.201915153
- Kraimer, G., Keller, S., and Schlierf, M. (2019). Structural Dynamics of Membrane-Protein Folding from Single-Molecule FRET. *Curr. Opin. Struct. Biol.* 58, 124–137. doi:10.1016/j.sbi.2019.05.025
- Kubben, N., Brimacombe, K. R., Donegan, M., Li, Z., and Misteli, T. (2016). A High-Content Imaging-Based Screening Pipeline for the Systematic Identification of Anti-progeroid Compounds. *Methods* 96, 46–58. doi:10.1016/j.ymeth.2015.08.024
- Kusumoto, D., Seki, T., Sawada, H., Kunitomi, A., Katsuki, T., Kimura, M., et al. (2021). Anti-senescent Drug Screening by Deep Learning-Based Morphology Senescence Scoring. *Nat. Commun.* 12 (1), 257. doi:10.1038/s41467-020-20213-0
- Labib, M., and Kelley, S. O. (2020). Single-cell Analysis Targeting the Proteome. *Nat. Rev. Chem.* 4 (3), 143–158. doi:10.1038/s41570-020-0162-7
- Lai, B., Gao, W., Cui, K., Xie, W., Tang, Q., Jin, W., et al. (2018). Principles of Nucleosome Organization Revealed by Single-Cell Micrococcal Nuclease Sequencing. *Nature* 562 (7726), 281–285. doi:10.1038/s41586-018-0567-3
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature* 521 (7553), 436–444. doi:10.1038/nature14539
- Lee, H., Gao, X., and Kim, Y.-P. (2018). Immuno-Nanoparticles for Multiplex Protein Imaging in Cells and Tissues. *Biochip J.* 12 (2), 83–92. doi:10.1007/s13206-018-2201-8
- Lee, J., Hyeon, D. Y., and Hwang, D. (2020). Single-cell Multiomics: Technologies and Data Analysis Methods. *Exp. Mol. Med.* 52 (9), 1428–1442. doi:10.1038/s12276-020-0420-2
- Lenselink, E. B., Ten Dijke, N., Bongers, B., Papadatos, G., van Vlijmen, H. W. T., Kowalczyk, W., et al. (2017). Beyond the Hype: Deep Neural Networks Outperform Established Methods Using a ChEMBL Bioactivity Benchmark Set. *J. Cheminform* 9 (1), 45. doi:10.1186/s13321-017-0232-0
- Leo, B. F., Fearn, S., Gonzalez-Cater, D., Theodorou, I., Ruenraroengsak, P., Goode, A. E., et al. (2019). Label-Free Time-Of-Flight Secondary Ion Mass Spectrometry Imaging of Sulfur-Producing Enzymes inside Microglia Cells Following Exposure to Silver Nanowires. *Anal. Chem.* 91 (17), 11098–11107. doi:10.1021/acs.analchem.9b01704
- Li, J., Wang, Y., and Zhang, Q. (2019). BEM-RCNN Segmentation Based on the Inadequately Labeled Moving Mesenchymal Stem Cells. *Pt Ii* 11663, 383–391. doi:10.1007/978-3-030-27272-2_34
- Li, Y., Wu, F. X., and Ngom, A. (2016). A Review on Machine Learning Principles for Multi-View Biological Data Integration. *Brief Bioinform* 19 (2), 325–340. doi:10.1093/bib/bbw113
- Li, Y., Huang, C., Ding, L., Li, Z., Pan, Y., and Gao, X. (2019). Deep Learning in Bioinformatics: Introduction, Application, and Perspective in the Big Data Era. *Methods* 166, 4–21. doi:10.1016/j.ymeth.2019.04.008
- Li, Y., and Ngom, A. (2014). Versatile Sparse Matrix Factorization: Theory and Applications. *Neurocomputing* 145, 23–29. doi:10.1016/j.neucom.2014.05.076
- Li, Z., Wang, Z., Pan, J., Ma, X., Zhang, W., and Ouyang, Z. (2020). Single-Cell Mass Spectrometry Analysis of Metabolites Facilitated by Cell Electromigration and Electroporation. *Anal. Chem.* 92 (14), 10138–10144. doi:10.1021/acs.analchem.0c02147
- Lim, B., Lin, Y., and Navin, N. (2020). Advancing Cancer Research and Medicine with Single-Cell Genomics. *Cancer Cell* 37 (4), 456–470. doi:10.1016/j.ccell.2020.03.008
- Lin, G., Baker, M. A. B., Hong, M., and Jin, D. (2018). The Quest for Optical Multiplexing in Bio-Discoveries. *Chem* 4 (5), 997–1021. doi:10.1016/j.chempr.2018.01.009
- Liu, L., Kan, A., Leckie, C., and Hodgkin, P. D. (2017). Comparative Evaluation of Performance Measures for Shading Correction in Time-Lapse Fluorescence Microscopy. *J. Microsc.* 266 (1), 15–27. doi:10.1111/jmi.12512
- Liu, T., Huang, J., Liaob, T., Pub, R., Liub, S., and Penga, Y. (2021). A Hybrid Deep Learning Model for Predicting Molecular Subtypes of Human Breast Cancer Using Multimodal Data. *Irbm*. doi:10.1016/j.irbm.2020.12.002
- Liu, X. B., Song, L., Liu, S., and Zhang, Y. (2021). A Review of Deep-Learning-Based Medical Image Segmentation Methods. *Sustainability* 13 (3), 1224. doi:10.3390/su13031224
- Lo, P. K., and Zhou, Q. (2018). Emerging Techniques in Single-Cell Epigenomics and Their Applications to Cancer Research. *J. Clin. Genom* 1 (1), 103. doi:10.4172/JCG.1000103
- Lock, J. G., and Strömblad, S. (2010). Systems Microscopy: an Emerging Strategy for the Life Sciences. *Exp. Cel Res.* 316 (8), 1438–1444. doi:10.1016/j.yexcr.2010.04.001
- Long, F., Zhou, J., and Peng, H. (2012). Visualization and Analysis of 3D Microscopic Images. *Plos Comput. Biol.* 8 (6), e1002519. doi:10.1371/journal.pcbi.1002519
- Lönnberg, T., Svensson, V., James, K. R., Fernandez-Ruiz, D., Sebina, I., Montandon, R., et al. (2017). Single-cell RNA-Seq and Computational Analysis Using Temporal Mixture Modelling Resolves Th1/Tfh Fate Bifurcation in Malaria. *Sci. Immunol.* 2 (9), eaal2192. doi:10.1126/sciimmunol.aal2192
- López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The Hallmarks of Aging. *Cell* 153 (6), 1194–1217. doi:10.1016/j.cell.2013.05.039
- López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M., and Kroemer, G. (2013). The Hallmarks of Aging. *Cell* 153 (6), 1194–1217. doi:10.1016/j.cell.2013.05.039
- Lugagne, J. B., Lin, H., and Dunlop, M. J. (2020). DeLTA: Automated Cell Segmentation, Tracking, and Lineage Reconstruction Using Deep Learning. *Plos Comput. Biol.* 16 (4), e1007673. doi:10.1371/journal.pcbi.1007673

- Luo, C., Keown, C. L., Kurihara, L., Zhou, J., He, Y., Li, J., et al. (2017). Single-cell Methyomes Identify Neuronal Subtypes and Regulatory Elements in Mammalian Cortex. *Science* 357 (6351), 600–604. doi:10.1126/science.aan3351
- Luquette, L. J., Bohrson, C. L., Sherman, M. A., and Park, P. J. (2019). Identification of Somatic Mutations in Single Cell DNA-Seq Using a Spatial Model of Allelic Imbalance. *Nat. Commun.* 10 (1), 3908. doi:10.1038/s41467-019-11857-8
- Ma, S., Sun, S., Geng, L., Song, M., Wang, W., Ye, Y., et al. (2020). Caloric Restriction Reprograms the Single-Cell Transcriptional Landscape of Rattus Norvegicus Aging. *Cell* 180 (5), 984–1001. e22. doi:10.1016/j.cell.2020.02.008
- Macauley, I. C., Haerty, W., Kumar, P., Li, Y. I., Hu, T. X., Teng, M. J., et al. (2015). G&T-seq: Parallel Sequencing of Single-Cell Genomes and Transcriptomes. *Nat. Methods* 12 (6), 519–522. doi:10.1038/nmeth.3370
- Macosko, E. Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* 161 (5), 1202–1214. doi:10.1016/j.cell.2015.05.002
- Mahmud, M., Kaiser, M. S., McGinnity, T. M., and Hussain, A. (2021). Deep Learning in Mining Biological Data. *Cogn. Comput.* 13, 1–33. doi:10.1007/s12559-020-09773-x
- Mamoshina, P., Kochetov, K., Cortese, F., Kovalchuk, A., Aliper, A., Putin, E., et al. (2019). Blood Biochemistry Analysis to Detect Smoking Status and Quantify Accelerated Aging in Smokers. *Sci. Rep.* 9 (1), 142. doi:10.1038/s41598-018-35704-w
- Mamoshina, P., Vieira, A., Putin, E., and Zhavoronkov, A. (2016). Applications of Deep Learning in Biomedicine. *Mol. Pharmaceutics* 13 (5), 1445–1454. doi:10.1021/acs.molpharmaceut.5b00982
- Marklein, R. A., Lam, J., Guvendiren, M., Sung, K. E., and Bauer, S. R. (2018). Functionally-Relevant Morphological Profiling: A Tool to Assess Cellular Heterogeneity. *Trends Biotechnol.* 36 (1), 105–118. doi:10.1016/j.tibtech.2017.10.007
- Martino, N., Kwok, S. J. J., Liapis, A. C., Forward, S., Jang, H., Kim, H.-M., et al. (2019). Wavelength-encoded Laser Particles for Massively Multiplexed Cell Tagging. *Nat. Photon.* 13 (10), 720–727. doi:10.1038/s41566-019-0489-0
- Masters, B. R. (2008). History of the Optical Microscope in Cell Biology and Medicine. *eLS*. doi:10.1002/9780470015902.a0003082
- Matsumoto, H., Kiryu, H., Furusawa, C., Ko, M. S. H., Ko, S. B. H., Gouda, N., et al. (2017). SCODE: an Efficient Regulatory Network Inference Algorithm from Single-Cell RNA-Seq during Differentiation. *Bioinformatics (Oxford, England)* 33 (15), 2314–2321. doi:10.1093/bioinformatics/btx194
- Mattiazzi Usaj, M., Sahin, N., Friesen, H., Pons, C., Usaj, M., Masinas, M. P. D., et al. (2020). Systematic Genetics and Single-Cell Imaging Reveal Widespread Morphological Pleiotropy and Cell-To-Cell Variability. *Mol. Syst. Biol.* 16 (2), e9243. doi:10.15252/msb.20199243
- Mattiazzi Usaj, M., Styles, E. B., Verster, A. J., Friesen, H., Boone, C., and Andrews, B. J. (2016). High-Content Screening for Quantitative Cell Biology. *Trends Cell Biol.* 26 (8), 598–611. doi:10.1016/j.tcb.2016.03.008
- McQuin, C., Goodman, A., Chernyshev, V., Kametsky, L., Cimini, B. A., Karhohs, K. W., et al. (2018). CellProfiler 3.0: Next-Generation Image Processing for Biology. *Plos Biol.* 16 (7), e2005970. doi:10.1371/journal.pbio.2005970
- Meijering, E. (2020). A Bird's-Eye View of Deep Learning in Bioimage Analysis. *Comput. Struct. Biotechnol. J.* 18, 2312–2325. doi:10.1016/j.csbj.2020.08.003
- Meijering, E. (2012). Cell Segmentation: 50 Years Down the Road [Life Sciences]. *IEEE Signal. Process. Mag.* 29 (5), 140–145. doi:10.1109/msp.2012.2204190
- Meinzel, W., Olivo-Marin, J.-C., and Angelini, E. D. (2018). Denoising of Microscopy Images: A Review of the State-Of-The-Art, and a New Sparsity-Based Method. *IEEE Trans. Image Process.* 27 (8), 3842–3856. doi:10.1109/tip.2018.2819821
- Meyer, M., Paquet, A., Arguel, M.-J., Peyre, L., Gomes-Pereira, L. C., Lebrigand, K., et al. (2020). Profiling the Non-genetic Origins of Cancer Drug Resistance with a Single-Cell Functional Genomics Approach Using Predictive Cell Dynamics. *Cel Syst.* 11 (4), 367–374. doi:10.1016/j.cels.2020.08.019
- Mikami, H., Lei, C., Nitta, N., Sugimura, T., Ito, T., Ozeki, Y., et al. (2018). High-Speed Imaging Meets Single-Cell Analysis. *Chem* 4 (10), 2278–2300. doi:10.1016/j.chempr.2018.06.011
- Mimitou, E. P., Cheng, A., Montalbano, A., Hao, S., Stoeckius, M., Legut, M., et al. (2019). Multiplexed Detection of Proteins, Transcriptomes, Clonotypes and CRISPR Perturbations in Single Cells. *Nat. Methods* 16 (5), 409–412. doi:10.1038/s41592-019-0392-0
- Minakshi, P., Ghosh, M., Kumar, R., Patki, H. S., Saini, H. M., Ranjan, K., et al. (2019). Single-Cell Metabolomics: Technology and Applications. *Single-Cell Omics*, 319–353. doi:10.1016/b978-0-12-814919-5.00015-4
- Mirza, B., Wang, W., Wang, J., Choi, H., Chung, N. C., and Ping, P. (2019). Machine Learning and Integrative Analysis of Biomedical Big Data. *Genes (Basel)* 10 (2), 87. doi:10.3390/genes10020087
- Mizukami, S. (2017). Targetable Fluorescent Sensors for Advanced Cell Function Analysis. *J. Photochem. Photobiol. C: Photochem. Rev.* 30, 24–35. doi:10.1016/j.jphotochemrev.2017.01.003
- Moffat, J. G., Rudolph, J., and Bailey, D. (2014). Phenotypic Screening in Cancer Drug Discovery - Past, Present and Future. *Nat. Rev. Drug Discov.* 13 (8), 588–602. doi:10.1038/nrd4366
- Moignard, V., Woodhouse, S., Haghverdi, L., Lilly, A. J., Tanaka, Y., Wilkinson, A. C., et al. (2015). Decoding the Regulatory Network of Early Blood Development from Single-Cell Gene Expression Measurements. *Nat. Biotechnol.* 33 (3), 269–276. doi:10.1038/nbt.3154
- Moskalev, A., Anisimov, V., Aliper, A., Artemov, A., Asadullah, K., Belsky, D., et al. (2017). A Review of the Biomedical Innovations for Healthy Longevity. *Aging* 9 (1), 7–25. doi:10.18632/aging.101163
- Moskalev, A., Chernyagina, E., de Magalhães, J. P., Barardo, D., Thoppil, H., Shaposhnikov, M., et al. (2015). Geroprotectors.org: a New, Structured and Curated Database of Current Therapeutic Interventions in Aging and Age-Related Disease. *Aging* 7 (9), 616–628. doi:10.18632/aging.100799
- Mulqueen, R. M., Pokholok, D., Norberg, S. J., Tornczyk, K. A., Fields, A. J., Sun, D., et al. (2018). Highly Scalable Generation of DNA Methylation Profiles in Single Cells. *Nat. Biotechnol.* 36 (5), 428–431. doi:10.1038/nbt.4112
- Murphy, D. B., and Davidson, M. W. (2012). *Fundamentals of Light Microscopy and Electronic Imaging*.
- Muzio, G., O'Bray, L., and Borgwardt, K. (2021). Biological Network Analysis with Deep Learning. *Brief. Bioinform.* 22 (2), 1515–1530. doi:10.1093/bib/bbaa257
- Niccoli, T., and Partridge, L. (2012). Ageing as a Risk Factor for Disease. *Curr. Biol.* 22 (17), R741–R752. doi:10.1016/j.cub.2012.07.024
- Nienhaus, G. U., and Nienhaus, K. (2017). “Fluorescence Labeling,” in *Fluorescence Microscopy*, 133–164. doi:10.1002/9783527687732.ch4
- Nketia, T. A., Sailem, H., Rohde, G., Machiraju, R., and Rittscher, J. (2017). Analysis of Live Cell Images: Methods, Tools and Opportunities. *Methods* 115, 65–79. doi:10.1016/j.jymeth.2017.02.007
- Nogueira-Recalde, U., Lorenzo-Gómez, I., Blanco, F. J., Loza, M. I., Grassi, D., Shirinsky, V., et al. (2019). Fibrates as Drugs with Senolytic and Autophagic Activity for Osteoarthritis Therapy. *Ebiomedicine* 45, 588–605. doi:10.1016/j.ebiom.2019.06.049
- Nordenfelt, P., Cooper, J. M., and Hochstetter, A. (2018). Matrix-masking to Balance Nonuniform Illumination in Microscopy. *Opt. Express* 26 (13), 17279–17288. doi:10.1364/oe.26.017279
- Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., et al. (2017). Dynamic Organization of Chromatin Domains Revealed by Super-resolution Live-Cell Imaging. *Mol. Cell* 67 (2), 282–293. e7. doi:10.1016/j.molcel.2017.06.018
- Oja, S., Komulainen, P., Penttilä, A., Nystedt, J., and Korhonen, M. (2018). Automated Image Analysis Detects Aging in Clinical-Grade Mesenchymal Stromal Cell Cultures. *Stem Cell Res Ther* 9 (1), 6. doi:10.1186/s13287-017-0740-x
- Okabe, K., Sakaguchi, R., Shi, B., and Kiyonaka, S. (2018). Intracellular Thermometry with Fluorescent Sensors for thermal Biology. *Pflugers Arch. - Eur. J. Physiol.* 470 (5), 717–731. doi:10.1007/s00424-018-2113-4
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man, Cybern.* 9 (1), 62–66. doi:10.1109/tsmc.1979.4310076
- Ounkomol, C., Seshamani, S., Maleckar, M. M., Collman, F., and Johnson, G. R. (2018). Label-free Prediction of Three-Dimensional Fluorescence Images from Transmitted-Light Microscopy. *Nat. Methods* 15 (11), 917–920. doi:10.1038/s41592-018-0111-2
- Ozawa, T., Yoshimura, H., and Kim, S. B. (2013). Advances in Fluorescence and Bioluminescence Imaging. *Anal. Chem.* 85 (2), 590–609. doi:10.1021/ac3031724
- Pan, S. J., and Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* 22 (10), 1345–1359. doi:10.1109/tkde.2009.191

- Pandey, S., and Bodas, D. (2020). High-quality Quantum Dots for Multiplexed Bioimaging: A Critical Review. *Adv. Colloid Interf. Sci.* 278, 102137. doi:10.1016/j.cis.2020.102137
- Pärnamaa, T., and Parts, L. (2017). Accurate Classification of Protein Subcellular Localization from High-Throughput Microscopy Images Using Deep Learning. *G3-Genes Genomes Genet.* 7 (5), 1385–1392. doi:10.1534/g3.116.033654
- Pau, G., Fuchs, F., Sklyar, O., Boutros, M., and Huber, W. (2010). EBIImage—an R Package for Image Processing with Applications to Cellular Phenotypes. *Bioinformatics* 26 (7), 979–981. doi:10.1093/bioinformatics/btq046
- Pau, G. Z., Xian, Z., Boutros, M., and Huber, W. (2020). *imageHTS: Analysis of High-Throughput Microscopy-Based Screens*.
- Pegoraro, G., and Misteli, T. (2017). High-Throughput Imaging for the Discovery of Cellular Mechanisms of Disease. *Trends Genet.* 33 (9), 604–615. doi:10.1016/j.tig.2017.06.005
- Peng, T., Thorn, K., Schroeder, T., Wang, L., Theis, F. J., Marr, C., et al. (2017). A BaSiC Tool for Background and Shading Correction of Optical Microscopy Images. *Nat. Commun.* 8, 14836. doi:10.1038/ncomms14836
- Peterson, V. M., Zhang, K. X., Kumar, N., Wong, J., Li, L., Wilson, D. C., et al. (2017). Multiplexed Quantification of Proteins and Transcripts in Single Cells. *Nat. Biotechnol.* 35 (10), 936–939. doi:10.1038/nbt.3973
- Phillip, J. M., Wu, P. H., Gilkes, D. M., Williams, W., McGovern, S., Daya, J., et al. (2017). Biophysical and Biomolecular Determination of Cellular Age in Humans. *Nat. Biomed. Eng.* 1 (7), 93. doi:10.1038/s41551-017-0093
- Piccinini, F., Lucarelli, E., Gherardi, A., and Bevilacqua, A. (2012). Multi-image Based Method to Correct Vignetting Effect in Light Microscopy Images. *J. Microsc.* 248 (1), 6–22. doi:10.1111/j.1365-2818.2012.03645.x
- Picelli, S., Faridani, O. R., Björklund, Å. K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-Seq from Single Cells Using Smart-Seq2. *Nat. Protoc.* 9 (1), 171–181. doi:10.1038/nprot.2014.006
- Piltti, K. M., Cummings, B. J., Carta, K., Manughian-Peter, A., Worne, C. L., Singh, K., et al. (2018). Live-cell Time-Lapse Imaging and Single-Cell Tracking of *In Vitro* Cultured Neural Stem Cells - Tools for Analyzing Dynamics of Cell Cycle, Migration, and Lineage Selection. *Methods* 133, 81–90. doi:10.1016/j.jmeth.2017.10.003
- Pliner, H. A., Shendure, J., and Trapnell, C. (2019). Supervised Classification Enables Rapid Annotation of Cell Atlases. *Nat. Methods* 16 (10), 983–986. doi:10.1038/s41592-019-0535-3
- Poudineh, M., Aldridge, P. M., Ahmed, S., Green, B. J., Kermanshah, L., Nguyen, V., et al. (2017). Tracking the Dynamics of Circulating Tumour Cell Phenotypes Using Nanoparticle-Mediated Magnetic Ranking. *Nat. Nanotech* 12 (3), 274–281. doi:10.1038/nnano.2016.239
- Pratapa, A., Jalihal, A. P., Law, J. N., Bharadwaj, A., and Murali, T. M. (2020). Benchmarking Algorithms for Gene Regulatory Network Inference from Single-Cell Transcriptomic Data. *Nat. Methods* 17 (2), 147–154. doi:10.1038/s41592-019-0690-6
- Pratiwi, F. W., Kuo, C. W., Chen, B.-C., and Chen, P. (2019). Recent Advances in the Use of Fluorescent Nanoparticles for Bioimaging. *Nanomedicine* 14 (13), 1759–1769. doi:10.2217/nnm-2019-0105
- Presman, D. M., Ball, D. A., Paakinaho, V., Grimm, J. B., Lavis, L. D., Karpova, T. S., et al. (2017). Quantifying Transcription Factor Binding Dynamics at the Single-Molecule Level in Live Cells. *Methods* 123, 76–88. doi:10.1016/j.jmeth.2017.03.014
- Pulgar, F. J., Charte, F., Rivera, A. J., and del Jesus, M. J. (2020). Choosing the Proper Autoencoder for Feature Fusion Based on Data Complexity and Classifiers: Analysis, Tips and Guidelines. *Inf. Fusion* 54, 44–60. doi:10.1016/j.inffus.2019.07.004
- Putin, E., Mamoshina, P., Aliper, A., Korzinkin, M., Moskalev, A., Kolosov, A., et al. (2016). Deep Biomarkers of Human Aging: Application of Deep Neural Networks to Biomarker Development. *Aging* 8 (5), 1021–1033. doi:10.18632/aging.100968
- Peng, T., Chen, G. M., and Tan, K. (2021). GLUER: Integrative Analysis of Single-Cell Omics and Imaging Data by Deep Neural Network. *bioRxiv* 2001, 2025427845. doi:10.1101/2021.01.25.427845
- Picard, M., Scott-Boyer, M.-P., Bodein, A., Périn, O., and Droit, A. (2021). Integration Strategies of Multi-Omics Data for Machine Learning Analysis. *Comput. Struct. Biotechnol. J.* 19, 3735–3746. doi:10.1016/j.csbj.2021.06.030
- Qiu, X., Mao, Q., Tang, Y., Wang, L., Chawla, R., Pliner, H. A., et al. (2017). Reversed Graph Embedding Resolves Complex Single-Cell Trajectories. *Nat. Methods* 14 (10), 979–982. doi:10.1038/nmeth.4402
- R Core Team, R. (2020). *A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging Individual mRNA Molecules Using Multiple Singly Labeled Probes. *Nat. Methods* 5 (10), 877–879. doi:10.1038/nmeth.1253
- Rauscher, B., Heigwer, F., Breinig, M., Winter, J., and Boutros, M. (2017). GenomeCRISPR - a Database for High-Throughput CRISPR/Cas9 Screens. *Nucleic Acids Res.* 45 (D1), D679–D686. doi:10.1093/nar/gkw997
- Razavian, A. S., Azizpour, H., Sullivan, J., and Carlsson, S. (2014). “CNN Features Off-The-Shelf: an Astounding Baseline for Recognition,” in 2014 Ieee Conference on Computer Vision and Pattern Recognition Workshops (Cvprw), 512–519. doi:10.1109/cvprw.2014.131
- Rifaioğlu, A. S., Atas, H., Martin, M. J., Cetin-Atalay, R., Atalay, V., and Doğan, T. (2018). Recent Applications of Deep Learning and Machine Intelligence on In Silico Drug Discovery: Methods, Tools and Databases. *Brief. Bioinform.* 20 (5), 1878–1912. doi:10.1093/bib/bby061
- Ritchie, M. D., Holzinger, E. R., Li, R., Pendergrass, S. A., and Kim, D. (2015). Methods of Integrating Data to Uncover Genotype-Phenotype Interactions. *Nat. Rev. Genet.* 16 (2), 85–97. doi:10.1038/nrg3868
- Rodriguez-Meira, A., Buck, G., Clark, S.-A., Povinelli, B. J., Alcolea, V., Louka, E., et al. (2019). Unravelling Intratumoral Heterogeneity through High-Sensitivity Single-Cell Mutational Analysis and Parallel RNA Sequencing. *Mol. Cell* 73 (6), 1292–1305. e1298. doi:10.1016/j.molcel.2019.01.009
- Rohban, M. H., Singh, S., Wu, X., Berthet, J. B., Bray, M.-A., Shrestha, Y., et al. (2017). Systematic Morphological Profiling of Human Gene and Allele Function via Cell Painting. *Elife* 6, e24060. doi:10.7554/elifesciences.24060
- Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., et al. (2018). Single-cell Profiling of the Developing Mouse Brain and Spinal Cord with Split-Pool Barcoding. *Science* 360 (6385), 176–182. doi:10.1126/science.aam8999
- Roweis, S. T., and Saul, L. K. (2000). Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science* 290 (5500), 2323–2326. doi:10.1126/science.290.5500.2323
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., et al. (2015). ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* 115 (3), 211–252. doi:10.1007/s11263-015-0816-y
- Saelens, W., Cannoodt, R., Todorov, H., and Saeys, Y. (2019). A Comparison of Single-Cell Trajectory Inference Methods. *Nat. Biotechnol.* 37 (5), 547–554. doi:10.1038/s41587-019-0071-9
- Sahl, S. J., Hell, S. W., and Jakobs, S. (2017). Fluorescence Nanoscopy in Cell Biology. *Nat. Rev. Mol. Cell Biol* 18 (11), 685–701. doi:10.1038/nrm.2017.71
- Sailem, H. Z., and Bakal, C. (2017). Identification of Clinically Predictive Metagenes that Encode Components of a Network Coupling Cell Shape to Transcription by Image-Omics. *Genome Res.* 27 (2), 196–207. doi:10.1101/gr.202028.115
- Salvi, M., Cerrato, V., Buffo, A., and Molinari, F. (2019). Automated Segmentation of Brain Cells for Clonal Analyses in Fluorescence Microscopy Images. *J. Neurosci. Methods* 325, 108348. doi:10.1016/j.jneumeth.2019.108348
- Samacoits, A., Chouaib, R., Safieddine, A., Traboulsi, A.-M., Ouyang, W., Zimmer, C., et al. (2018). A Computational Framework to Study Sub-cellular RNA Localization. *Nat. Commun.* 9 (1), 4584. doi:10.1038/s41467-018-06868-w
- Sarkar, T. J., Quarta, M., Mukherjee, S., Colville, A., Paine, P., Doan, L., et al. (2020). Transient Non-integrative Expression of Nuclear Reprogramming Factors Promotes Multifaceted Amelioration of Aging in Human Cells. *Nat. Commun.* 11 (1), 1545. doi:10.1038/s41467-020-15174-3
- Sasagawa, Y., Nikaido, I., Hayashi, T., Danno, H., Uno, K. D., Imai, T., et al. (2013). Quartz-Seq: a Highly Reproducible and Sensitive Single-Cell RNA Sequencing Method, Reveals Non-genetic Gene-Expression Heterogeneity. *Genome Biol.* 14 (4), 3097. doi:10.1186/gb-2013-14-4-r31
- Schep, A. N., Wu, B., Buenrostro, J. D., and Greenleaf, W. J. (2017). chromVAR: Inferring Transcription-Factor-Associated Accessibility from Single-Cell Epigenomic Data. *Nat. Methods* 14 (10), 975–978. doi:10.1038/nmeth.4401
- Schermelleh, L., Ferrand, A., Huser, T., Eggeling, C., Sauer, M., Biehlmair, O., et al. (2019). Super-resolution Microscopy Demystified. *Nat. Cell Biol* 21 (1), 72–84. doi:10.1038/s41556-018-0251-8

- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an Open-Source Platform for Biological-Image Analysis. *Nat. Methods* 9 (7), 676–682. doi:10.1038/nmeth.2019
- Schmidt, E. E., Pelz, O., Buhlmann, S., Kerr, G., Horn, T., and Boutros, M. (2013). GenomeRNAi: a Database for Cell-Based and *In Vivo* RNAi Phenotypes, 2013 Update. *Nucleic Acids Res.* 41, D1021–D1026. doi:10.1093/nar/gks1170
- Schuster, A., Erasmus, H., Fritah, S., Nazarov, P. V., van Dyck, E., Niclou, S. P., et al. (2019). RNAi/CRISPR Screens: from a Pool to a Valid Hit. *Trends Biotechnol.* 37 (1), 38–55. doi:10.1016/j.tibtech.2018.08.002
- Segler, M. H. S., Preuss, M., and Waller, M. P. (2018). Planning Chemical Syntheses with Deep Neural Networks and Symbolic AI. *Nature* 555 (7698), 604–610. doi:10.1038/nature25978
- Setty, M., Tadmor, M. D., Reich-Zeliger, S., Angel, O., Salame, T. M., Kathail, P., et al. (2016). Wishbone Identifies Bifurcating Developmental Trajectories from Single-Cell Data. *Nat. Biotechnol.* 34 (6), 637–645. doi:10.1038/nbt.3569
- Shanta, P. V., Li, B., Stuart, D. D., and Cheng, Q. (2020). Plasmonic Gold Templates Enhancing Single Cell Lipidomic Analysis of Microorganisms. *Anal. Chem.* 92 (9), 6213–6217. doi:10.1021/acs.analchem.9b05285
- Shao, W., Wang, T., Sun, L., Dong, T., Han, Z., Huang, Z., et al. (2020). Multi-task Multi-Modal Learning for Joint Diagnosis and Prognosis of Human Cancers. *Med. Image Anal.* 65, 101795. doi:10.1016/j.media.2020.101795
- Shashkova, S., and Leake, M. C. (2017). Single-molecule Fluorescence Microscopy Review: Shedding New Light on Old Problems. *Biosci. Rep.* 37 (4), 31. doi:10.1042/BSR20170031
- Shen, D., Wu, G., and Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. *Annu. Rev. Biomed. Eng.* 19 (1), 221–248. doi:10.1146/annurev-bioeng-071516-044442
- Simidjievski, N., Bodnar, C., Tariq, I., Scherer, P., Andres Terre, H., Shams, Z., et al. (2019). Variational Autoencoders for Cancer Data Integration: Design Principles and Computational Practice. *Front. Genet.* 10, 1205. doi:10.3389/fgene.2019.01205
- Simonyan, K., and Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv preprint arXiv:1409.1556*.
- Singh, S., Bray, M. A., Jones, T. R., and Carpenter, A. E. (2014). Pipeline for Illumination Correction of Images for High-throughput Microscopy. *J. Microsc.* 256 (3), 231–236. doi:10.1111/jmi.12178
- Smith, K., Li, Y., Piccinini, F., Csucs, G., Balazs, C., Bevilacqua, A., et al. (2015). CIDRE: an Illumination-Correction Method for Optical Microscopy. *Nat. Methods* 12 (5), 404–406. doi:10.1038/nmeth.3323
- Söderberg, O., Gullberg, M., Jarvius, M., Ridderstråle, K., Leuchowius, K.-J., Jarvius, J., et al. (2006). Direct Observation of Individual Endogenous Protein Complexes *In Situ* by Proximity Ligation. *Nat. Methods* 3 (12), 995–1000. doi:10.1038/nmeth947
- Stavrakis, S., Holzner, G., Choo, J., and deMello, A. (2019). High-throughput Microfluidic Imaging Flow Cytometry. *Curr. Opin. Biotechnol.* 55, 36–43. doi:10.1016/j.copbio.2018.08.002
- Stoeckius, M., Hafemeister, C., Stephenson, W., Houck-Loomis, B., Chattopadhyay, P. K., Swerdlow, H., et al. (2017). Simultaneous Epitope and Transcriptome Measurement in Single Cells. *Nat. Methods* 14 (9), 865–868. doi:10.1038/nmeth.4380
- Street, K., Risso, D., Fletcher, R. B., Das, D., Ngai, J., Yosef, N., et al. (2018). Slingshot: Cell Lineage and Pseudotime Inference for Single-Cell Transcriptomics. *BMC Genomics* 19 (1), 477. doi:10.1186/s12864-018-4772-0
- Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a Generalist Algorithm for Cellular Segmentation. *Nat. Methods* 18 (1), 100–106. doi:10.1038/s41592-020-01018-x
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., et al. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177 (7), 1888–1902. e21. doi:10.1016/j.cell.2019.05.031
- Swedlow, J. R., Kankaanpää, P., Sarkans, U., Goscinski, W., Galloway, G., Malacrida, L., et al. (2021). A Global View of Standards for Open Image Data Formats and Repositories. *Nat. Methods* 18, 1440. doi:10.1038/s41592-021-01113-7
- Swedlow, J. R. (2013). Quantitative Fluorescence Microscopy and Image Deconvolution. *Methods Cel Biol* 114, 407–426. doi:10.1016/b978-0-12-407761-4.00017-8
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). “Going Deeper with Convolutions,” in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Boston, MA: IEEE), 1–9.
- Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., et al. (2017). Human Ageing Genomic Resources: New and Updated Databases. *Nucleic Acids Res.* 46 (D1), D1083–D1090. doi:10.1093/nar/gkx1042
- Tang, B., Pan, Z., Yin, K., and Khateeb, A. (2019). Recent Advances of Deep Learning in Bioinformatics and Computational Biology. *Front. Genet.* 10, 214. doi:10.3389/fgene.2019.00214
- Tenenbaum, J. B., Silva, V. d., and Langford, J. C. (2000). A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science* 290 (5500), 2319–2323. doi:10.1126/science.290.5500.2319
- Terai, T., and Nagano, T. (2013). Small-molecule Fluorophores and Fluorescent Probes for Bioimaging. *Pflugers Arch. - Eur. J. Physiol.* 465 (3), 347–359. doi:10.1007/s00424-013-1234-z
- Thomas, R. M., and John, J. (2017). “A Review on Cell Detection and Segmentation in Microscopic Images,” in 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT) (Kollam, India: IEEE). doi:10.1109/iccpct.2017.8074189
- Thorn, K. (2017). Genetically Encoded Fluorescent Tags. *MBoC* 28 (7), 848–857. doi:10.1091/mbc.e16-07-0504
- Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., Geladaki, A., Ait Blal, H., et al. (2017). A Subcellular Map of the Human Proteome. *Science* 356 (6340), eaal3321. doi:10.1126/science.aal3321
- Tsigelny, I. F. (2018). Artificial Intelligence in Drug Combination Therapy. *Brief. Bioinform.* 20 (4), 1434–1448. doi:10.1093/bib/bby004
- Uyar, B., Palmer, D., Kowald, A., Murua Escobar, H., Barrantes, I., Möller, S., et al. (2020). Single-cell Analyses of Aging, Inflammation and Senescence. *Ageing Res. Rev.* 64, 101156. doi:10.1016/j.arr.2020.101156
- van der Maaten, L., and Hinton, G. (2008). Visualizing Data Using T-SNE. *J. Machine Learn. Res.* 9, 2579–2605.
- van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). Scikit-Image: Image Processing in Python. *PeerJ* 2, e453. doi:10.7717/peerj.453
- Van Rossum, G. D. (2009). *F. L., Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace.
- Van Valen, D. A., Kudo, T., Lane, K. M., Macklin, D. N., Quach, N. T., DeFelice, M. M., et al. (2016). Deep Learning Automates the Quantitative Analysis of Individual Cells in Live-Cell Imaging Experiments. *Plos Comput. Biol.* 12 (11), e1005177. doi:10.1371/journal.pcbi.1005177
- Vangindertael, J., Camacho, R., Sempels, W., Mizuno, H., Dedecker, P., and Janssen, K. P. F. (2018). An Introduction to Optical Super-resolution Microscopy for the Adventurous Biologist. *Methods Appl. Fluoresc.* 6 (2), 022003. doi:10.1088/2050-6120/aaae0c
- Venugopalan, J., Tong, L., Hassanzadeh, H. R., and Wang, M. D. (2021). Multimodal Deep Learning Models for Early Detection of Alzheimer’s Disease Stage. *Sci. Rep.* 11 (1), 3254. doi:10.1038/s41598-020-74399-w
- Vicar, T., Balvan, J., Jaros, J., Jug, F., Kolar, R., Masarik, M., et al. (2019). Cell Segmentation Methods for Label-free Contrast Microscopy: Review and Comprehensive Comparison. *BMC Bioinformatics* 20 (1), 360. doi:10.1186/s12859-019-2880-8
- Wählby, C., Erlandsson, F., Bengtsson, E., and Zetterberg, A. (2002). Sequential Immunofluorescence Staining and Image Analysis for Detection of Large Numbers of Antigens in Individual Cell Nuclei. *Cytometry* 47 (1), 32–41. doi:10.1002/cyto.10026
- Wang, H., Nakamura, M., Abbott, T. R., Zhao, D., Luo, K., Yu, C., et al. (2019). CRISPR-mediated Live Imaging of Genome Editing and Transcription. *Science* 365 (6459), 1301–1305. doi:10.1126/science.aax7852
- Wang, R., Lin, D.-Y., and Jiang, Y. (2019). SCOPE: a Normalization and Copy Number Estimation Method for Single-Cell DNA Sequencing. *bioRxiv*, 594267.
- Wang, X., Chen, H., and Zhang, N. R. (2018). DNA Copy Number Profiling Using Single-Cell Sequencing. *Brief Bioinform* 19 (5), 731–736. doi:10.1093/bib/bbx004
- Wang, Y., and Navin, N. E. (2015). Advances and Applications of Single-Cell Sequencing Technologies. *Mol. Cel* 58 (4), 598–609. doi:10.1016/j.molcel.2015.05.005
- Wang, Z., Li, L., Glicksberg, B. S., Israel, A., Dudley, J. T., and Ma’ayan, A. (2017). Predicting Age by Mining Electronic Medical Records with Deep Learning

- Characterizes Differences between Chronological and Physiological Age. *J. Biomed. Inform.* 76, 59–68. doi:10.1016/j.jbi.2017.11.003
- Wei, J. N., Duvenaud, D., and Aspuru-Guzik, A. (2016). Neural Networks for the Prediction of Organic Chemistry Reactions. *ACS Cent. Sci.* 2 (10), 725–732. doi:10.1021/acscentsci.6b00219
- Willy, C., Neugebauer, E. A. M., and Gerngro, H. (2003). The Concept of Nonlinearity in Complex Systems. *Eur. J. Trauma* 29 (1), 11–22. doi:10.1007/s00068-003-1248-x
- Wolf, F. A., Hamey, F. K., Plass, M., Solana, J., Dahlin, J. S., Göttgens, B., et al. (2019). PAGA: Graph Abstraction Reconciles Clustering with Trajectory Inference through a Topology Preserving Map of Single Cells. *Genome Biol.* 20 (1), 59. doi:10.1186/s13059-019-1663-x
- Wollman, R., and Stuurman, N. (2007). High Throughput Microscopy: from Raw Images to Discoveries. *J. Cel Sci* 120 (Pt 21), 3715–3722. doi:10.1242/jcs.013623
- Wu, P. H., Gilkes, D. M., Phillip, J. M., Narkar, A., Cheng, T. W., Marchand, J., et al. (2020). Single-cell Morphology Encodes Metastatic Potential. *Sci. Adv.* 6 (4), eaaw6938. doi:10.1126/sciadv.aaw6938
- Wu, Y., and Zhang, K. (2020). Tools for the Analysis of High-Dimensional Single-Cell RNA Sequencing Data. *Nat. Rev. Nephrol.* 16 (7), 408–421. doi:10.1038/s41581-020-0262-0
- Wang, T., Shao, W., Huang, Z., Tang, H., Zhang, J., Ding, Z., et al. (2021). MOGONET Integrates Multi-Omics Data Using Graph Convolutional Networks Allowing Patient Classification and Biomarker Identification. *Nat. Commun.* 12 (1), 3445. doi:10.1038/s41467-021-23774-w
- Xiong, L., Xu, K., Tian, K., Shao, Y., Tang, L., Gao, G., et al. (2019). SCALE Method for Single-Cell ATAC-Seq Analysis via Latent Feature Extraction. *Nat. Commun.* 10 (1), 4576. doi:10.1038/s41467-019-12630-7
- Xu, C., and Su, Z. (2015). Identification of Cell Types from Single-Cell Transcriptomes Using a Novel Clustering Method. *Bioinformatics* 31 (12), 1974–1980. doi:10.1093/bioinformatics/btv088
- Xu, Y., Dai, Z., Chen, F., Gao, S., Pei, J., and Lai, L. (2015). Deep Learning for Drug-Induced Liver Injury. *J. Chem. Inf. Model.* 55 (10), 2085–2093. doi:10.1021/acs.jcim.5b00238
- Yan, F., Powell, D. R., Curtis, D. J., and Wong, N. C. (2020). From Reads to Insight: a Hitchhiker's Guide to ATAC-Seq Data Analysis. *Genome Biol.* 21 (1), 22. doi:10.1186/s13059-020-1929-3
- Yang, K. D., Belyaeva, A., Venkatachalapathy, S., Damodaran, K., Katcoff, A., Radhakrishnan, A., et al. (2021). Multi-domain Translation between Single-Cell Imaging and Sequencing Data Using Autoencoders. *Nat. Commun.* 12 (1), 31. doi:10.1038/s41467-020-20249-2
- Yao, K., Rochman, N. D., and Sun, S. X. (2019). Cell Type Classification and Unsupervised Morphological Phenotyping from Low-Resolution Images Using Deep Learning. *Sci. Rep.* 9 (1), 13467. doi:10.1038/s41598-019-50010-9
- Yu, Y., Si, X., Hu, C., and Zhang, J. (2019). A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* 31 (7), 1235–1270. doi:10.1162/neco_a_01199
- Zafar, H., Wang, Y., Nakhleh, L., Navin, N., and Chen, K. (2016). Monovar: Single-Nucleotide Variant Detection in Single Cells. *Nat. Methods* 13 (6), 505–507. doi:10.1038/nmeth.3835
- Zagato, E., Toon, B., De Smedt, S. C., Katrien, R., Kristiaan, N., Kevin, B., et al. (2018). Technical Implementations of Light Sheet Microscopy. *Microsc. Res. Tech.* 81 (9), 941–958. doi:10.1002/jemt.22981
- Zahn, J. M., Poosala, S., Owen, A. B., Ingram, D. K., Lustig, A., Carter, A., et al. (2007). AGEMAP: a Gene Expression Database for Aging in Mice. *Plos Genet.* 3 (11), e201. doi:10.1371/journal.pgen.0030201
- Zhai, J., Zhang, S., Chen, J., and He, Q. (2018). “Autoencoder and its Various Variants,” in 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 415–419. doi:10.1109/smc.2018.00080
- Zhai, T., Li, Q., Shen, J., Li, J., and Fan, C. (2020). DNA Nanostructure-encoded Fluorescent Barcodes. *Aggregate* 1 (1), 107–116. doi:10.1002/agt2.8
- Zhai, W., Tan, J., Russell, T., Chen, S., McGonagle, D., Win Naing, M., et al. (2021). Multi-pronged Approach to Human Mesenchymal Stromal Cells Senescence Quantification with a Focus on Label-free Methods. *Sci. Rep.* 11 (1), 1054. doi:10.1038/s41598-020-79831-9
- Zhang, L., Dong, X., Lee, M., Maslov, A. Y., Wang, T., and Vijg, J. (2019). Single-cell Whole-Genome Sequencing Reveals the Functional Landscape of Somatic Mutations in B Lymphocytes across the Human Lifespan. *Proc. Natl. Acad. Sci. USA* 116 (18), 9014–9019. doi:10.1073/pnas.1902510116
- Zhang, L., Le Lu, fnm., Nogues, I., Summers, R. M., Liu, S., and Yao, J. (2017). DeepPap: Deep Convolutional Networks for Cervical Cell Classification. *IEEE J. Biomed. Health Inform.* 21 (6), 1633–1643. doi:10.1109/jbhi.2017.2705583
- Zhang, L., and Vertes, A. (2018). Single-Cell Mass Spectrometry Approaches to Explore Cellular Heterogeneity. *Angew. Chem. Int. Ed.* 57 (17), 4466–4477. doi:10.1002/anie.201709719
- Zhang, S., Liu, C.-C., Li, W., Shen, H., Laird, P. W., and Zhou, X. J. (2012). Discovery of Multi-Dimensional Modules by Integrative Analysis of Cancer Genomic Data. *Nucleic Acids Res.* 40 (19), 9379–9391. doi:10.1093/nar/gks725
- Zhang, S., Zhou, J., Hu, H., Gong, H., Chen, L., Cheng, C., et al. (2015). A Deep Learning Framework for Modeling Structural Features of RNA-Binding Protein Targets. *Nucleic Acids Res.* 44 (4), e32. doi:10.1093/nar/gkv1025
- Zhavoronkov, A., Bischof, E., and Lee, K.-F. (2021). Artificial Intelligence in Longevity Medicine. *Nat. Aging* 1 (1), 5–7. doi:10.1038/s43587-020-00020-4
- Zhavoronkov, A., Mamoshina, P., Vanhaelen, Q., Scheibye-Knudsen, M., Moskalev, A., and Aliper, A. (2019). Artificial Intelligence for Aging and Longevity Research: Recent Advances and Perspectives. *Ageing Res. Rev.* 49, 49–66. doi:10.1016/j.arr.2018.11.003
- Zhu, C., Preissl, S., and Ren, B. (2020). Single-cell Multimodal Omics: the Power of many. *Nat. Methods* 17 (1), 11–14. doi:10.1038/s41592-019-0691-5
- Zhu, G., Shao, Y., Liua, Y., Peia, T., Lia, L., Zhanga, D., et al. (2021). Single-cell Metabolite Analysis by Electrospray Ionization Mass Spectrometry. *Trac Trends Anal. Chem.* 143, 116351. doi:10.1016/j.trac.2021.116351
- Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., et al. (2017). Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol. Cel* 65 (4), 631–643. e4. doi:10.1016/j.molcel.2017.01.023
- Zitnik, M., Nguyen, F., Wang, B., Leskovec, J., Goldenberg, A., and Hoffman, M. M. (2019). Machine Learning for Integrating Data in Biology and Medicine: Principles, Practice, and Opportunities. *Inf. Fusion* 50, 71–91. doi:10.1016/j.inffus.2018.09.012
- Zong, C., Lu, S., Chapman, A. R., and Xie, X. S. (2012). Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. *Science* 338 (6114), 1622–1626. doi:10.1126/science.1229164

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Watson, Taherian Fard and Mar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.