



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

A Multimodal Omics Framework to Empower Target Discovery for Cardiovascular Regeneration

Citation for published version:

Li, Z, Brittan, M & Mills, NL 2023, 'A Multimodal Omics Framework to Empower Target Discovery for Cardiovascular Regeneration', *Cardiovascular Drugs and Therapy*. <https://doi.org/10.1007/s10557-023-07484-7>

Digital Object Identifier (DOI):

[10.1007/s10557-023-07484-7](https://doi.org/10.1007/s10557-023-07484-7)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

Cardiovascular Drugs and Therapy

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.





A Multimodal Omics Framework to Empower Target Discovery for Cardiovascular Regeneration

Ziwen Li¹ · Mairi Brittan¹ · Nicholas L. Mills¹

Accepted: 19 June 2023
© The Author(s) 2023

Abstract

Ischaemic heart disease is a global healthcare challenge with high morbidity and mortality. Early revascularisation in acute myocardial infarction has improved survival; however, limited regenerative capacity and microvascular dysfunction often lead to impaired function and the development of heart failure. New mechanistic insights are required to identify robust targets for the development of novel strategies to promote regeneration. Single-cell RNA sequencing (scRNA-seq) has enabled profiling and analysis of the transcriptomes of individual cells at high resolution. Applications of scRNA-seq have generated single-cell atlases for multiple species, revealed distinct cellular compositions for different regions of the heart, and defined multiple mechanisms involved in myocardial injury-induced regeneration. In this review, we summarise findings from studies of healthy and injured hearts in multiple species and spanning different developmental stages. Based on this transformative technology, we propose a multi-species, multi-omics, meta-analysis framework to drive the discovery of new targets to promote cardiovascular regeneration.

Keywords Single-cell RNA sequencing · Myocardial infarction · Heart failure · Target discovery · Cardiovascular regeneration

Introduction

Heart failure due to ischemic heart disease is associated with a high morbidity and mortality despite treatment [1]. Early restoration of blood flow by percutaneous coronary intervention has improved survival in acute myocardial infarction. However, myocardial hypoperfusion or “no-flow” phenomenon [2] caused by microvascular dysfunction is common and can impact the effectiveness of revascularisation, limiting tissue regeneration and resulting in irreversible injury and heart failure [3, 4]. With the ageing of the population and increasing availability of early treatment for acute myocardial infarction, the prevalence of heart failure

is projected to rise worldwide [5]. Numerous approaches, including pluripotent/stem cell transplantation, growth factor administration, and direct cell reprogramming, have been developed and tested in preclinical studies and clinical trials to promote tissue regeneration following ischemic injury, but the resulting benefits are transient and minimal [6]. New insights are required to broaden our understanding of the mechanisms underpinning endogenous cardiac regeneration and to develop more effective therapeutic strategies.

Human hearts, like most mammalian hearts, rapidly lose the ability to regenerate in the first week after birth [7–9]. This has presented a major challenge to regenerative medicine and has prompted researchers to either venture beyond mammals, studying species with superior regenerative capacity [10–12], or to turn back the clock, focusing on regenerative programs possessed by prenatal and early postnatal mammalian hearts. Furthermore, myocardial repair after infarction is a multi-stage process, involving inflammation, necrosis, angiogenesis, fibrosis, long-term remodeling, and maturation [13, 14], demanding a comprehensive regenerative strategy suitable for different stages of recovery.

Previous attempts to understand the mechanisms of myocardial repair and regeneration using traditional methods,

✉ Ziwen Li
Ziwen.Cass.Li@ed.ac.uk

Mairi Brittan
mbrittan@exseed.ed.ac.uk

Nicholas L. Mills
Nick.Mills@ed.ac.uk

¹ BHF Centre for Cardiovascular Science, The Queen's Medical Research Institute, University of Edinburgh, Edinburgh, UK

such as quantitative PCR, bulk-RNA sequencing, flow-cytometry, and immunofluorescent staining, have generated invaluable insights into the transcriptional and translational change in major cell populations as a whole, but our understanding of heterogeneity and plasticity of individual cells and cell subpopulations has remained limited. The breakthrough in whole transcriptome analysis in single cells since 2009 [15] has ushered in a new era of precision in our understanding of cell biology and disease [16]. Building on the unprecedented resolution offered by single-cell technologies, the rise of spatial transcriptomics has further empowered the interrogation of gene expression by individual cells through the provision of their “geographical coordinates” within the tissue of interest [17]. Recent applications of these technologies have identified the dynamic transcriptomic signatures of major cell populations and novel subpopulations involved in cardiovascular development and in adult heart homeostasis and the pathogenesis of ischaemic heart disease [18, 19]. This presents an exciting opportunity for researchers to thoroughly map the microenvironment composition and biochemical cues and discover novel therapeutic targets that may elicit regenerative responses in the diseased heart [20].

This review will discuss the latest findings using these new sequencing technologies and propose a multi-species, multi-omics, meta-analysis framework to drive the discovery of new targets to promote cardiovascular regeneration.

Understanding Cardiovascular Regeneration from a Single-Cell Perspective

Single-cell RNA sequencing (scRNA-seq) has been widely adopted to detect and quantify RNA molecules in individual cells, surpassing the traditional bulk-RNA-seq technology where a large number of cells are pooled and heterogeneity assessment is lacking [21]. The way we investigate organisms, organs, and tissues is transformed. Researchers have started building large-scale multi-organ single-cell transcriptomic atlases for different species such as *Tabula Sapiens* [22] and *Tabula Muris* [23] at an unprecedented speed. Cross-organ studies assessing major cell types, such as endothelial cells [24], smooth muscle cells [25], and fibroblasts [26], have generated valuable insights into how cells adapt to functional requirements at various anatomical locations. Facing ongoing healthcare challenges like cardiovascular disease and cancer [27–29] and public health crisis like the COVID-19 pandemic [30, 31], scientists have been able to quickly pinpoint the impacted cell types and transcriptional perturbations using scRNA-seq. Its wide application has led to the discovery of new lineage trajectories, cell populations and states, cell-cell interaction patterns, disease-specific transcriptional signatures, and potential

therapeutic targets in a series of state-of-the-art studies in the cardiovascular field.

The heart is a complex and dynamic organ. Profiling such an organ requires a careful review of the factors that shape the research questions, including (i) the anatomical region of interest, such as different heart chambers, layers of the heart wall, or specialised subregions like valvular apparatus; (ii) the major cell types to focus on, cardiomyocytes, endothelial cells, fibroblasts, etc.; (iii) the age of the heart, developmental, adult, or ageing; (iv) sex of the donor; and (v) status of the heart, whether it is healthy or diseased, such as myocardial infarction, cardiomyopathy, or heart failure. The studies included in this review have investigated some or a combination of these factors. We will review these studies and assimilate the findings in a holistic manner using the factors as a guide (Fig. 1).

Surveying Cardiac Anatomy at Single-Cell Level

The first endeavour to build a comprehensive human heart atlas is marked by the study carried out by Litviňuková and colleagues, in which they identified eleven major cell types including atrial cardiomyocytes, ventricular cardiomyocytes, fibroblasts, endothelial cells, pericytes, smooth muscle cells, immune cells, adipocytes, mesothelial cells, and neuronal cells, from 45,870 cells, 78,023 CD45+ enrich cells, and 363,213 nuclei [32]. A clear atrial-ventricular difference was found in the cell composition: atrial tissues contained 30.1% cardiomyocytes, 24.3% fibroblasts, and 12.2% endothelial cells, while ventricular tissues had more cardiomyocytes (49.2%), less fibroblasts (15.5%), and less endothelial cells (7.8%) [32]. The proportions of ventricular cardiomyocytes and fibroblasts were negatively correlated [32]. Tucker et al. showed that cardiomyocytes were the most diverse cell type between atria and ventricles, and the left-right difference was less pronounced [33]. In contrast, the endothelial cells were more distinct by sides (left versus right), with 220 differentially expressed genes for the left versus right comparison and only 43 differentially expressed genes for atrium versus ventricle [33]. Like the four chambers, the four cardiac valves were also shown to have distinct cellular composition: migratory valvular interstitial cells (VICs), mast cells, myofibroblasts, and shear stress-reactive valvular endothelial cells (VECs) were found in the aortic valve; shear stress-reactive VECs and mast cells were in the pulmonary valve; nitric oxide (NO)-associated VICs, protective VECs, and migratory anti-inflammatory T cells were found in tricuspid valve; and mast cells and myofibroblasts were in mitral valve [34]. Other specialised subregions such as sinoatrial node (SAN) have also been studied at single-cell level: a study of 718 SAN cells from 21 adult mice has identified *Vsn11* as a potential sinoatrial node marker and defined four distinct transcriptional clusters [35].

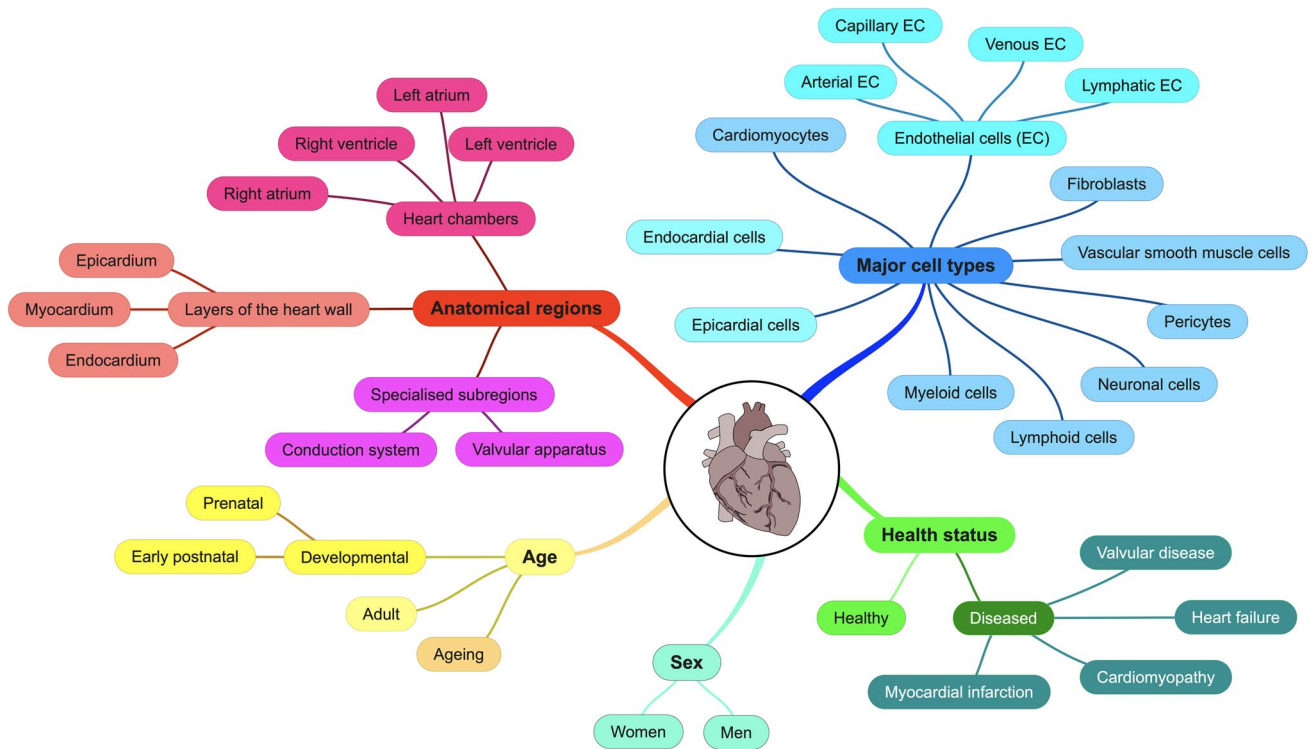


Fig. 1 Map of factors to consider for single-cell studies in the cardiovascular field. The heart is a complex organ, and single-cell profiling studies focusing on specific questions often only provide limited information. This road map highlights some of the primary systems and factors that are now important to consider in a holistic manner to improve our understanding of heart function, repair, and regeneration. These systems and factors include different anatomical regions in the heart (the four chambers, the three layers of the heart wall, and

specialised subregions like the conduction system and valvular apparatus), major cardiac cell types (endocardial cells, epicardial cells, cardiomyocytes, endothelial cells, fibroblasts, vascular smooth muscle cells, pericytes, neuronal cells, lymphoid cells, and myeloid cells), the health status of the heart (healthy or diseased such as myocardial infarction, cardiomyopathy, heart failure, valvular disease), sex (women and men), and age (prenatal, postnatal, adult, ageing)

The chamber wall has three layers. The inner endocardium consists of endothelial cells and connective tissue, lining the four chambers; the middle myocardium mainly contains cardiomyocytes, responsible for heart contraction; the outer epicardium covers the myocardium and is largely mesothelium [36]. The endocardium and epicardium are quiescent in the adult heart but are actively involved in cardiogenesis [37, 38]. A study of embryonic day (E) 11.5 mouse hearts showed that the endocardial-specific deletion of *Hand2* disrupted multiple endocardial signalling pathways and marker gene expression and caused congenital defects, including tricuspid atresia, double inlet left ventricle, and interventricular septal defects, suggesting that endocardial *Hand2* is required for cardiac morphogenesis [39]. In a study of E12.5 and E16.5 mouse hearts, researchers found that a group of *Slit2*⁺ epicardium-derived cells (EPDCs) emerged after epithelial-to-mesenchymal transition (EMT), functioning as guidepost cells for migrating endothelial cells. An epicardium-specific deletion of myocardium-related transcription factors *Mrtf-a* and *Mrtf-b* disrupted the expression of guidance molecules and led to the ectopic accumulation of

endothelial cells on the epicardial surface, indicating that paracrine signals from EPDCs are required for endothelial pathfinding and linking EMT to endothelial cell positioning [40].

Resolving Cellular Heterogeneity in the Infarct Heart

Cardiomyocytes, capillaries, and extracellular matrix (ECM) are the main components of the myocardium enabling muscle contraction. Capillaries guard the supply of oxygen and nutrients, and fibroblasts produce ECM and maintain the structural properties. In acute myocardial infarction, fibroblasts are activated by the invading immune cells and expand to stabilise the infarct region, while endothelial cells proliferate to support new vessel growth in the border zone to restore the blood flow and limit the infarct size. Angiogenesis and fibrosis are key processes in healing after ischaemic injury, which has prompted researchers to profile endothelial cells and fibroblasts in the healthy and injured hearts. Despite being transcriptionally and functionally distinct,

endothelial cells (ECs) and fibroblasts (FBs) show an interesting degree of commonality.

Firstly, both ECs and FBs have high heterogeneity, which changes with age. Litviňuková and colleagues identified ten EC populations in the adult human heart, including three capillary clusters (expressing *RGCC* and *CA4*, representing 57.4% of all ECs), one immune regulation-associated cluster (expressing *CX3CL1* and *IL6*), one arterial cluster (expressing *SEMA3G* and *EFNB2*), one venous cluster (expressing *NR2F2* and *ACKR1*), one atrial cluster (expressing *SMOC1* and *NPR3*, likely being endocardial cells), one lymphatic cluster (expressing *PROX1* and *PDPN*), and two clusters with fibroblast- and cardiomyocyte-like features respectively [32]. They also identified seven populations of FBs with an enriched expression of *DCN*, *GSN*, and *PDGFRA*, among which FB3 cluster expressed a high level of cytokine receptors, FB4 cluster expressed *POSTN* and *TNC* responsive to TGF β signalling, FB5 cluster expressed genes related to ECM modulation, FB6 and FB7 showed endothelial- and cardiomyocyte-like features respectively [32]. By comparing the hearts from young adult (3 months) and aged mice (24 months), researchers discovered parallel alterations in age-affected circulating factors and endothelial expressed receptors and identified an apelin receptor-enriched endothelial subpopulation that was reduced with age [41]. Another study comparing the adult (3 months old) and ageing (18 months old) hearts further showed that ageing predominantly affected fibroblasts among other cell types, and aged fibroblasts had a higher transcriptional heterogeneity, with increased signatures in angiogenesis, inflammatory responses, and osteoblast differentiation, compared to the young fibroblasts [42]. To further assess whether the paracrine activity also altered with age in the fibroblasts, the study compared conditioned medium derived from aged and young mouse cardiac fibroblasts in vitro and found that the conditioned medium from the aged heart did not affect proliferation but impaired tube formation, enhanced pro-inflammatory activities, and reduced autophagy in endothelial cells [42].

Secondly, ECs and FBs both have a complex developmental origin, which has little impact on injury response. A study of the developing mouse and human hearts revealed that the transcriptional heterogeneity of coronary endothelial cells was initially influenced by lineage, originated from endocardium or sinus venous and then by anatomical location, until it declined in the homeostatic adult heart, and endothelial cells in the adult heart had similar responses to injuries regardless of their origins [43]. Similarly, most fibroblasts in the adult heart were derived from the epicardium, with two small fractions from endothelial cells and Pax3-expressing cells respectively, but the expansion of fibroblasts after injury was consistent regardless of their developmental lineage [44].

Thirdly, both ECs and FBs have small subpopulations in the steady state that can expand rapidly upon activation after ischaemic injury. A previous study of the healthy and 7-day post-myocardial infarction mouse heart demonstrated that clonal proliferation of resident endothelial cells sustained the endothelial structure after infarction and established the first endothelial atlas in the infarcted adult heart [45]. We defined ten endothelial clusters, including injury-specific clusters 6, 7, and 8 expressing high levels of *Plvap*. Immunofluorescence of Plvap/PLVAP in injured mouse and human heart tissues showed consistent upregulation at the protein level. The siRNA knockdown of *Plvap* significantly reduced endothelial proliferation in vitro, identifying Plvap as a potential angiogenic target [45]. Ruiz-Villalba and colleagues presented the first scRNA-seq study focusing on cardiac fibroblasts and showed that a subpopulation of the activated fibroblasts, expressing *Cthrc1*, expanded rapidly in response to myocardial infarction, from 2.3% at the steady state to 12% at 7-day post-injury and to 34% at 14 days, and these cells were exclusively located in the infarct and border zones, presenting a strong pro-fibrotic signature [46]. *Cthrc1* knockout mice showed a dramatically poorer survival rate after myocardial infarction (80% wildtype versus 30% knockout) due to ventricular rupture as a result of decreased collagen deposition [46]. The study also identified a population of fibroblasts with a similar transcriptomic profile in a swine myocardial infarction model and in the heart tissue from patients with myocardial infarction and dilated cardiomyopathy [46].

Finally, both ECs and FBs have subpopulations that can potentially take on a cardiomyocyte fate via transdifferentiation/reprogramming. It is worth noticing that in the human heart, both ECs and FBs have small subpopulations expressing cardiomyocyte markers [32]. We also identified a small disease-specific endothelial cluster that co-expressed cardiomyocyte marker *Tnnt2* in the injured mouse heart at 7-day post-myocardial infarction and confirmed its existence in the injured heart tissue using immunofluorescence [45], indicating a fate shift towards cardiomyocyte in response to injuries. Using an inducible fibroblast-specific *Tcf21-iCre*, Tani et al. overexpressed cardiac transcription factors *Mef2c/Gata4/Tbx5/Hand2* in fibroblasts, which directly reprogrammed around 2% of cardiac fibroblasts into induced cardiomyocytes [47]. This was achieved by suppressing the expression of fibroblastic genes and converting pro-fibrotic cardiac fibroblasts to a quiescent anti-fibrotic state. In vivo reprogramming of cardiac fibroblasts using this technique has shown to improve cardiac function and reverse fibrosis in acute and chronic infarcted hearts [47].

Endothelial cells and fibroblasts play pivotal roles in angiogenesis and ECM modulation in the healthy and diseased hearts; however, the communication between the two has not been thoroughly investigated. Future studies on the

in situ cross-talk between endothelial cells and fibroblasts will provide invaluable insights into their synergy in tissue repair and regeneration and inspire new therapeutic strategies that can simultaneously target both populations.

Deciphering the Transcriptomic Complexity in the Failing Heart

The acute loss of myocardium and the resultant load increase following myocardial infarction triggers a cascade of biochemical signalling processes, which modulate post-injury repair and remodelling, involving dilation, hypertrophy, and scar formation [48]. Hypertrophy is an adaptive response to the increased load and progressive dilation during ventricular remodelling. In preclinical research, pressure overload hypertrophy is induced by transaortic banding (TAC) in mice. Ren et al. profiled the major cell types during the progression of pressure-induced hypertrophy at 0-, 2-, 5-, 8-, and 11-week post-TAC and described a stagewise change of cellular cross-talk [49]. They discovered that macrophage activation and subtype switching occurred at the middle stage (2–5 weeks), and fibroblasts and endothelial subtypes were correlated with disease progression especially at the late stage (5–11 weeks) [49]. Wang et al. studied 21,422 cells from the normal, failed, and partially recovered (left ventricular assist device treatment) adult human hearts and found that the contractility and metabolism of cardiomyocytes were prominently correlated with the changes in heart function [50]. They discovered that non-myocyte cells actively regulated myocyte behaviours and identified a subpopulation of *ACKR1*+ endothelial cells that could preserve cardiac function after injury [50].

Heart failure is a complex clinical syndrome, encompassing a diverse group of clinical features converging on impaired muscle function. Recent studies have mapped the transcriptional and cellular diversity of human hearts with different types of cardiomyopathy, including hypertrophic cardiomyopathy (HCM) [51], dilated cardiomyopathy (DCM) [51–54], ischaemic cardiomyopathy (ICM) [52], and arrhythmogenic cardiomyopathy (ACM) [54]. Rao et al. profiled non-cardiomyocyte cells from late-stage DCM and ICM hearts and identified AEBP1 as a crucial regulator of cardiac fibrosis in ACTA2+ myofibroblasts [52]. They also discovered a *CXCL8*^{hi}*CCR2*+*HLA-Dr*^{hi} macrophage subpopulation enriched in a severely fibrotic area, which interacted with endothelial cells and potentially facilitated leukocyte recruitment [52]. Koenig et al. defined the transcriptional landscape of 25 non-diseased and 13 non-ischaemic DCM left ventricles and investigated the impact of age, sex, and disease state on the transcriptomes [53]. They found that cell composition did change in DCM individuals but was not associated with the severity of heart failure, and no clear differences by sex were detected in the non-diseased

or DCM group. They observed a converging trend in cardiomyocytes toward common disease-associated states, while fibroblasts and myeloid cells were increasingly diversified. Endothelial cells and pericytes showed global transcriptional shifts but little to no change in cell composition [53]. Chaffin and colleagues profiled the left ventricles from 11 DCM, 15 HCM, and 16 non-failing hearts and identified a large transcriptional change in DCM or HCM left ventricles compared to the non-failing ones, though few differences were found between DCM and HCM despite their distinct genetic bases [51]. They compared HCM samples with preserved left ventricular ejection fraction (LVEF) to HCM samples with reduced LVEF and found no convincing transcriptional differences in any cell types, indicating a convergence to a common transcriptional profile in advanced cardiomyopathy [51]. Reichart and colleagues sequenced the samples from the left ventricle free wall, apex, septum, and right ventricle free wall from non-ischaemic failing hearts with pathogenic variants in DCM and ACM and performed a genotype-stratified analysis of cell lineages and transcriptional states [54]. The cell atlas from DCM and ACM ventricles showed different left and right ventricular responses and demonstrated genotype-specific signalling pathways, cell-cell communications, and differentially expressed genes at single-cell levels. They further applied a machine learning approach and successfully developed a graph attention network classifier that was trained on the snRNA-seq data from four major left ventricular cell types (cardiomyocytes, fibroblasts, endothelial cells, and myeloid cells) and was able to predict the genotypes with high accuracy [54]. Simonson et al. performed snRNA-seq on non-infarcted samples from the left ventricles of patients with long-term ICM and observed a significant decrease in cardiomyocytes and an endothelial subpopulation expressing high *ARHGAP28* and *FBNI* [55]. Genes associated with proliferation and angiogenesis were enriched in endothelial cells, and lymphatic endothelial cells increased in all ICM samples [55]. By combining the HCM and DCM data from Chaffin et al. [51], Simonson and colleagues also identified the genes that were up- and down-regulated in ICM but not in HCM or DCM [55]. Furthermore, they compared the genes with consistent change across all three cardiomyopathies to the druggable genome [56], resulting in 39 druggable genes (10 up-regulated and 29 down-regulated) including *CYD2J2*, *MYH6*, *COL14A1*, and *SPARCL1* [55].

While these individual studies only investigated one or two types of cardiomyopathy, they have collectively encompassed the full spectrum of cardiomyopathy (Table 1). Four [51–54] out of the five studies have primarily analysed a total of 81 DCM samples, meriting further investigation into the shared characteristics. All four studies have profiled major cardiac cell types including cardiomyocytes (though not in Rao et al.), fibroblasts, endothelial cells, and

Table 1 Overview of sample numbers and cell types in the cardiomyopathy studies [51–55]

Studies	HCM	DCM	ICM	ACM	Non-diseased heart	Cell types profiled
Rao et al. [52]	–	<i>N</i> = 5	<i>N</i> = 4	–	<i>N</i> = 2	Non-cardiomyocyte cells
Koenig et al. [53]	–	<i>N</i> = 13*	–	–	<i>N</i> = 25*	All cell types
Chaffin et al. [51]	<i>N</i> = 15	<i>N</i> = 11	–	–	<i>N</i> = 16	All cell types
Reichart et al. [54]	–	<i>N</i> = 52	–	<i>N</i> = 8	<i>N</i> = 18	All cell types
Simonson et al. [55]			<i>N</i> = 7		<i>N</i> = 8	All cell types
Total	<i>N</i> = 15	<i>N</i> = 81	<i>N</i> = 11	<i>N</i> = 8	<i>N</i> = 69	–

HCM hypertrophic cardiomyopathy, *DCM* dilated cardiomyopathy, *ICM* ischaemic cardiomyopathy, *ACM* arrhythmogenic cardiomyopathy, *N* number

*The referenced study reported that 27 healthy donors and 18 individuals with dilated cardiomyopathy were included. However, entries of only 38 patients were found in the supplementary patient metadata. The discrepancies are likely due to double counts of patients whose samples were processed using both snRNA-seq and scRNA-seq

macrophages. A decrease in cardiomyocytes was observed across all studies and in all genotypes except in the *LMNA* pathogenic variant (PV) carriers. An increase in fibroblasts was shown in all studies except Reichart et al. despite clear histopathological findings of fibrosis. Angiogenic signatures were up-regulated in endothelial cells across all studies, and a reduction in proliferating macrophages was observed in Koenig et al. and Chaffin et al. Both Chaffin et al. and Reichart et al. investigated the impact of *TTN* mutations. The former detected minimal transcriptional differences between the *TTN* loss-of-function mutation carriers and other DCM patients, while the latter observed both cell compositional and transcriptional changes that were specific to the *TTN* PV carriers. A high degree of agreement in the findings is readily appreciated in these studies; however, discrepancies and unanswered questions warrant follow-up investigation. Furthermore, subgroup analysis, such as a sex-specific profiling for the *TTN* PV carriers in Reichart et al., was not feasible due to limited sample size and would certainly benefit from incorporating external datasets, like the data from the *TTN* mutation carriers in Chaffin et al. A meta-analysis approach to integrate datasets and assimilate findings from individual studies would help ensure validity and stimulate new analyses as shown briefly above. In the process of curating the studies, knowledge/data gap could also be identified—more data from HCM, ICM, and ACM are clearly required here to obtain a complete map of the disease-specific transcriptome and to enable subgroup analyses stratified by age, sex, and genotype (Table 1).

Harnessing Endogenous Regenerative Capacity in Non-mammalian Species

Unlike humans, salamanders and zebrafish [57] can regenerate their hearts throughout life, making excellent models for cardiovascular regeneration. After injuries, axolotl hearts can undergo macrophage-modulated cardiomyocyte

proliferation to regenerate, and macrophage depletion leads to progressive scarring and impaired ventricular function [58]. An axolotl single-cell study revealed that axolotl hearts contained *Gata4+* endothelial cells, supporting both angiogenesis and cardiomyocyte regeneration [59]. Using scRNA-seq, three epicardial subpopulations and seven cell clusters have been identified in the developing [60] and adult zebrafish hearts [61] respectively. Transcription factor *prrx1b*, expressed in the epicardial- and epicardium-derived cells, was shown required for scar-free zebrafish heart regeneration. Loss of *prrx1b* caused excessive fibrosis and impaired cardiomyocyte proliferation by tuning *Nrg1* expression [62]. The transient activated epicardial progenitor cells (aEPCs), expressing *ptx3a* and *coll2a1b*, underwent epithelial-to-mesenchymal transition (EMT) to differentiate into mural cells and mesenchymal cells. A conditional ablation of these aEPCs was shown to block heart regeneration through reduced *Nrg1* expression and decreased mesenchymal population [63]. Heterogeneous *pdgfrb+* epicardium-derived cells were also found essential for heart regeneration [64]. In the endocardium, absence of *Runx1*-positive cells promoted myocardial proliferation and protection [65]. A transient subpopulation of collagen-12-expressing fibroblasts, which originated from both the epicardium and endocardium, has proved proregenerative [65]. A cardiomyocyte-focused study in zebrafish has revealed that border zone cardiomyocytes underwent metabolic reprogramming induced by *Nrg1*/*ErbB2* signalling after injuries, making them resemble embryonic cardiomyocytes [66]. Taken together, restoration of zebrafish and salamander hearts is accomplished through the efforts of multiple cell types and subpopulations via specific signalling pathways such as *Nrg1*. Further studies of the epicardium are required to gain a comprehensive understanding of its role in adult heart regeneration.

Studying the Developmental Heart to Inform Adult Cardiovascular Regeneration Pathways

Adult mammalian hearts regenerate poorly, but their prenatal and early postnatal counterparts restore fully with minimal fibrosis and function normally [7–9]. These observations have led to the development of two reprogramming strategies: to reactivate the foetal regenerative programs and to deactivate the senescence programs in the adult heart. An analysis of the E9.5, E12.5, and postnatal (P)1 mouse cardiomyocytes has revealed a “progenitor-like” subpopulation that was abundant in the early embryonic mouse hearts but decreased to minimal postnatally [67]. The same study also found that epicardial progenitors were unlikely to be the source of cardiomyocytes using an epicardial lineage tracing model, as the labelled clones were primarily fibroblasts and smooth muscle cells [67]. The combined application of the Fucci system and scRNA-seq has enabled direct comparisons between cycling and non-cycling cardiomyocytes at P0 (regenerative) and P7 (non-regenerative) and demonstrated that cycling cardiomyocytes had the highest mitotic rate at birth, which gradually declined during neonatal period [68]. The differentially up-regulated genes in cycling P0 cardiomyocytes included members of the Hippo/YAP and Nrg1/ErbB2 signalling pathways and genes required for cardiomyocyte cytokinesis like *Ect2* and *RhoA* [68]. Another study comparing the regenerative (P1) and non-regenerative (P8) neonatal hearts showed that many epicardial-enriched and injury-induced genes were specific to the regenerative hearts, among which RSPO1 was a key ligand mediating crosstalk between epicardial and endothelial cells and enhancing the angiogenic ability of endothelial cells [69]. In contrast to the regenerative hearts, epicardial cells from the non-regenerative hearts secreted pro-fibrotic factors post-injuries [69]. Epicardial-derived cells (EPDCs) are essential for cardiac regeneration during mammalian development [67, 69] and in lower vertebrates with superior regenerative capacity [60–64]. However, it is unclear how many of these cells exist and to what extent they can differentiate into other cardiac cell types to support regeneration in the adult human heart. Knight-Schrijver and colleagues have recently explored this uncharted space in a scRNA-seq study comparing the cellular composition and transcriptional states of foetal and adult human epicardium [70]. They found that the adult epicardium had a minimal population of mesenchymal EPDCs and reduced paracrine communication with other cardiac cell types and was lacking angiogenic epicardial programs including WNT signalling compared to their foetal counterpart [70]. They also revealed an age-associated transition of human epicardium from an angiogenic-responsive state toward an immune-responsive state, likely a previously overlooked element in understanding the role of immune response in cardiac regeneration.

A recent study on aortic endothelial cells in 4-, 26-, and 86-week-old mice has shed light on senescence heterogeneity among different endothelial subpopulations, showing upregulation of senescence-associated genes *P53* and *P21* and downregulation of proliferation-associated genes *Ki67* in the Gpihbp1+ EC1 subpopulation, compared to the Plvap+ EC2 [71]. Taking a more direct approach, Chen et al. overexpressed the Yamanaka factors [72] Oct4, Sox2, c-Myc, and Klf4 in cardiomyocytes for 6 days before or during injuries and achieved the repair of amputated apex with minimal scarring at P7, a time point when cardiomyocytes normally stopped proliferating and the regenerative capacity of the heart is completely lost [73]. They successfully demonstrated that the transient dedifferentiation and reprogramming of adult cardiomyocytes could improve left ventricular function, provided it was initiated as early as possible post-myocardial infarction, and they also cautioned that the process carried risks of cellular transformation and neoplasm formation if not tightly controlled [73]. In summary, steady progresses have been made towards rejuvenating adult heart tissue to promote injury-induced repair, and more in-depth studies are required to develop and validate these approaches and explore their safety and efficacy.

Multimodal Profiling of the Single-Cell Multiome

Turning back ageing clocks in the heart requires a deep understanding of epigenetic control of cardiovascular development, homeostasis, and diseases. Epigenetic changes accumulate during ageing [74], leading to disrupted gene expression, reduced tissue function, and diminished regenerative capacity. With technologies like scATAC-seq (assay for transposase-accessible chromatin using sequencing) and scMethyl-seq (DNA methylation sequencing), we can map the epigenetic landscape at single-cell levels and gain further insights into the transcriptional regulatory networks. Wang et al. compared the post-MI P1 and P8 mouse hearts using scATAC-seq and found that fibroblasts had the greatest number of both myocardial infarction-induced and infarction-repressed regions among the major cell types, and more myocardial infarction-responsive regions were found in the P8 hearts than in the P1 hearts, indicating a prominent open chromatin dynamics in fibroblasts concurrent with cardiac remodelling [69]. To assess the chromatin level change in the cardiomyocyte-like cells (iCMs) induced by reprogramming factors *Mef2c/Gata4/Tbx5* (MGT), Wang and colleagues performed scATAC-seq on cardiac fibroblasts at 3 days post-MGT treatment and revealed several transcription factor regulatory networks associated with the conversion, including FOS-AP1, SMAD, and TEAD [75]. They showed that accessible regions with the motifs of FOS-AP1 were rapidly closed upon MGT induction, and SMAD family members exhibited diverse modulation roles on iCM

reprogramming at different stages post-induction [75]. Single-cell profiling at the protein level has recently come into play. Schoof and colleagues presented the first large-scale single-cell mass spectrometry–based proteomics with comprehensive experiment and computational analysis workflow [76]. The novel approach allowed consistent quantification of 1000 proteins per cell across thousands of cells, and the companion software Sceptre could seamlessly normalise the data, integrate FACS data, and assist downstream analysis [76]. With the development and maturation of these new technologies, future studies can routinely perform simultaneous profiling of genome, epigenome, transcriptome, and proteome, covering the full genetic information flow from DNA to RNA to protein, providing comprehensive multi-omics map for healthy and diseased hearts at the single-cell level.

Triangulating Injury Responses Using Spatial Coordinates

Single-cell sequencing technology relies on the successful isolation of individual cells [77], during which the spatial information of each cell is irreversibly lost. By combining single-cell and spatial transcriptomics, researchers have gained insights into the spatial distribution of different cell types and subtypes and how individual cells with specific transcriptional profiles function collectively with the surrounding cells in the microenvironment. Asp et al. constructed the first spatial transcriptomics map for the embryonic human hearts spanning three developmental stages, enabling studies of cardiac morphogenesis [78]. Misra and colleagues analysed P2 and P10 heart tissues post-apical resection with spatial transcriptomics, which defined the scar tissue and its transitioning from a proliferative to secretory phenotype as the heart lost regenerative capacity. They also found a regenerative border zone defined by immature cardiomyocyte markers such as *Ankrd1* and strong expression of cardioprotection factor *Sprr1a*. Kuppe et al. constructed the first spatial multi-omics map of human myocardial infarction and characterised tissues from the ischaemic zone, boarder zone, remote zone, and fibrotic zone using cell-type deconvolution [20]. Boileau et al. developed Single-cell Nanopore Spatial Transcriptomics (SCNASt) software, based on long-read sequencing data barcode assignment, to generate a full-length near-single-cell transcriptional landscape of the tissue microenvironment [79]. Using this method, they profiled four mouse heart sections at 3 days post-myocardial infarction and discovered previously uncharacterised modes of transcription, such as intron retention, which was associated with the infarct area. They also identified 109 significant regional isoform switching genes across all comparisons between the remote, border, and infarct areas [79].

Difficulties in Connecting a Million Dots

Researchers have gained valuable insights into the many faceted nature of cardiac regeneration through numerous studies in a range of organism models over the years. However, this has not resolved into a clear picture of the mechanistic network governing regeneration in the adult human heart. It is largely due to failure in using small animal models to faithfully reproduce human disease and recapitulate human responses to interventions. After all, cardiac cellular composition varies among different species [32, 80, 81], and each species has unique regulatory programs in heart homeostasis and disease [82, 83]. The difference could sometimes preclude further development of drugs and therapies for reasons like the lack of human orthologous of a promising target gene discovered in mice [84]. Furthermore, the central dogma, from DNA to RNA and from RNA to protein, is at the centre of all biological processes, and each step also has many pre- and post-regulatory mechanisms. While DNA and RNA can cause structural and functional changes in protein, leading to phenotypic perturbations, disease-associated phenotypes are more often caused by altered regulatory mechanisms. Subsequently, inconsistency is observed between different layers of omics. Pei and colleagues profiled 13 HCM (with truncating *MYBPC3* mutations) and 10 control hearts using chromatin immunoprecipitation sequencing (ChIP-seq), RNA-seq, and proteomics [85]. They identified HCM-related changes in histone acetylation (4226 regions with higher acetylation and 5084 with lower acetylation), transcription (936 up-regulated genes and 1097 down-regulated genes), and translation (216 up-regulated proteins and 225 down-regulated proteins) [85]. Despite the rich information from each omics layer, only 5 genes including *ADPN*, *FMOD*, *MCAM*, *NPPA*, and *AASS* showed a consistent change in the same direction in the HCM versus control hearts at DNA, RNA, and protein levels [85], reflecting the complex relationship between different omics layers and cautioning us against drawing conclusions based on a single layer.

To better understand the mechanism and pinpoint the omics layer(s) to target for intervention at the single-cell level, researchers have started combining multiple single-cell omics technologies. Alexanian and colleagues performed scRNA-seq and scATAC-seq on the healthy and pressure overload mouse hearts and found that the fibroblasts had increased chromatin accessibility post-TAC [86]. They identified transcription factor (TF) MEOX1 as a central regulator of fibroblast activation and showed that *Meox1* knockdown reduced the formation of stress fibres and expression of *Acta2* in TGF β -treated fibroblasts [86]. Ameen and colleagues also integrated scATAC-seq

and scRNA-seq data from human foetal heart samples and mapped the landscapes of cis-regulatory elements (CREs) and genes that define major cell types during cardiogenesis [87]. They successfully trained deep learning models to interpret the cell-type-specific sequence of active TF binding sites, predict the impact of de novo non-coding mutations on chromatin accessibility, and infer the active TF binding sites disrupted by high-impact mutations [87]. They validated an endothelial-specific enhancer that harboured a predicted high-impact mutation related to *JARID2* using CRISPR/Cas9 deletion and showed reduced *JARID2* expression and depletion of tubes using a tube formation assay in iPSC-derived endothelial cells [87]. This analysis framework created a new way to study putative causal, de novo non-coding mutations associated with congenital heart diseases. Hocker and colleagues profiled ~80,000 cardiac cells using snATAC-seq on adult human heart tissues and created a comprehensive atlas of cardiac cis-regulatory elements (cCREs) [88]. By integrating snRNA-seq data from the matched samples, they identified differentially accessible cCREs underlying chamber- and cell-type-specific gene expression patterns and potential causal candidate cCREs involved in heart failure and atrial fibrillation (AF). They validated the role of one cCRE that contained AF risk variants in regulating the expression of potassium channel gene *KCNH2* and action potential repolarisation in hPSC cardiomyocytes, using CRISPR/Cas9 genome deletion [88].

On the one hand, it is increasingly difficult to effectively synthesise the knowledge from previous studies in the field; on the other hand, new technologies are generating more data and creating new avenues at unprecedented paces. Both require the creation of a framework to systematically and continuously assimilate findings from research studies to improve our understanding of the complex mechanism in cardiovascular diseases.

A Multi-species Multi-omics Meta-analysis Framework

Breaking through the barrier to heart regeneration in humans, we need to use a comprehensive approach and test all possibilities. We have previously carried out a meta-analysis of integrated scRNA-seq data from coronary endothelial cells extracted from injured and healthy developing adult mouse and human heart, which enabled us to gain new insights into the conserved pathways between the injured human and mouse hearts, unique regenerative programs used by the developing hearts, and transcriptional shifts in heart failure [89]. We identified VEGFC, KLF4, EGR1, and ZFP36 as potential new targets for vascular regeneration and generated a coronary endothelial cell meta-atlas

CrescENDO Shiny application to empower future research [89]. By constantly integrating data from species with higher regenerative capacity like axolotls and zebrafish from other omics technology such as scATAC-seq and spatial transcriptomics, and disease data spanning various timepoints after myocardial infarction and different stages of heart failure, we will be able to construct an exhaustive database that allows fast knowledge base building and rapid target discovery, prioritisation, and validation. Therefore, we propose a multi-species, multi-omics, meta-analysis framework (Fig. 2), to assimilate the latest scientific discoveries, inform new mechanisms, and generate novel targets for successful cardiovascular regeneration. There are four main stages in this framework: data generation and collation, target discovery and prioritisation, target testing and validation, and target recommendation. In stage 1, previously published and newly generated bulk, single-cell, and spatial datasets will be collated for each species; in stage 2, targets will be pulled out from individual studies with their regulatory network mapped out with the multi-omics data, and these targets will subsequently be prioritised based on the expression level, relevance to patients, and novelty; in stage 3, top ranked targets are tested in vitro (cultured primary cells and organoids), ex vivo (cultured tissue slices and biopsies), and in vivo (small and large animal models); in stage 4, the performance of the targets are assessed, and the ones with the greatest safety and efficacy will be recommended for further preclinical or even clinical investigation. As the framework grows, it will embrace new technologies/omics data and knowledge from other species that are clinically relevant; allow modular investigation and selective integration tunable to different therapeutic strategies, such as reactivating the developmental programs, targeting specific phases of the regeneration process, and utilising conserved regenerative pathways uniquely activated in non-mammalian species; and explore new testing platforms as they become available.

Establishing and maintaining such a framework require the wider research community to work together and act collectively, coordinating studies, providing datasets, creating new pipelines for multimodal data, automating target prioritisation, optimising downstream testing pathways, fine-tuning recommendation algorithms, and building and updating knowledge base. This in turn will empower broader and more impactful international research, encouraging scientists from different disciplines to join forces to tackle the challenges in cardiovascular regeneration and accelerate therapeutic innovation. The Human Cell Atlas Consortium [90] and CZ CELLxGENE Discover [91] have demonstrated the feasibility of large-scale single-cell data curation, integration, annotation, and sharing. Learning from their success will be instrumental in the early stage of building this framework. Multimodal approach has been adopted by artificial intelligence (AI)-augmented drug discovery companies

A multi-species, multi-omic, meta-analysis framework powering target discovery in cardiovascular regeneration

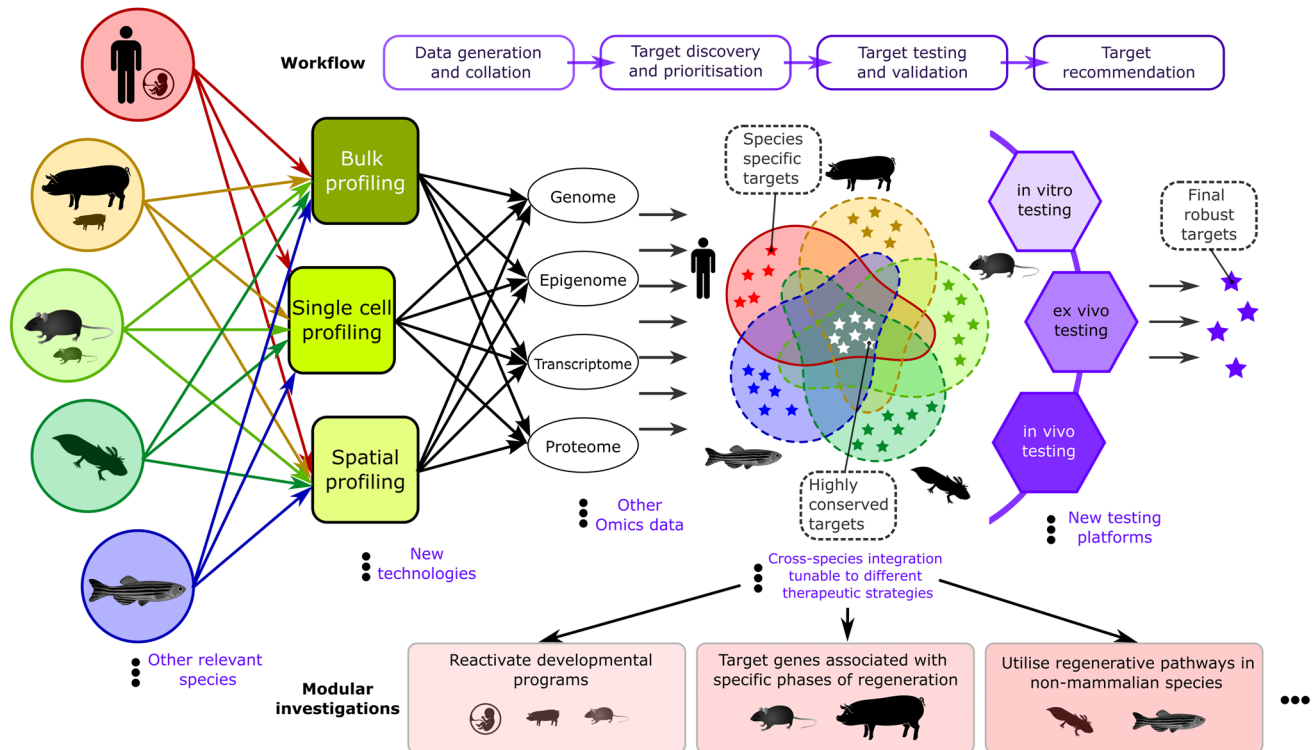


Fig. 2 Schematic of a multi-species, multi-omics, meta-analysis framework for target discovery in cardiovascular regeneration. There are four main stages in the framework: stage 1, bulk/single-cell/spatial profiling of genome, epigenome, transcriptome, and proteome data from human, large animal models like porcine, small animal models like mice and rats, highly regenerative non-mammal models like axolotl and zebrafish, will be collated and inventorised based on species and three main technologies, including bulk, single-cell, and spatial profiling; stage 2, multi-omics data (including genome, epigenome, transcriptome, and proteome profiling) are integrated vertically within each species and technology, potential targets and their regulatory networks are identified, and targets from individual integra-

tions are pulled out and prioritised based on the factors like expression level, relevance to human pathophysiology, and novelty; stage 3, top ranked targets are tested in in vitro, ex vivo, and finally in vivo experiments; stage 4, targets with the greatest safety and efficacy recommended for further preclinical or clinical investigations. The framework will embrace new technologies/omics data and knowledge from other species; allow modular investigation and selective integration tunable to specific therapeutic strategies, such as reactivating the developmental programs, targeting specific phases of the regeneration process, and utilising conserved regenerative pathways uniquely activated in non-mammalian species; and explore new testing platforms

such as BenevolentAI to help understand disease biology, build data foundation, and inform new targets [92]. Collaborating with leading AI drug discovery companies will significantly accelerate the establishment and refinement of the framework.

Understanding the Challenge Ahead

A fully functioning and well-maintained framework would serve as a powerful resource for the research community, but setting it up is not easy. A major challenge is multi-omics data integration, for unmatched (of the same modality but from different cells) and matched (from the same cells but of different modalities) datasets [93]. The former, also called horizontal integration, uses genes as the anchor for

integration, while the latter, vertical integration, uses cells. Horizontal integration has been used for integrating large-scale scRNA-seq datasets generated across multiple batches and technologies. The tools for horizontal integration, such as Seurat v3 [94], Harmony [95], and scVI [96], aim to remove batch and other technical variances while preserving the biological differences. Vertical integration can be divided further into local integration, focusing on defining putative interaction across omics layers, and global integration, with an aim to understand broader cellular states like the cell cycle. Tools like MOFA [97] and Seurat v4 [98] are used for global integration and LMM [99] for local integration. To integrate datasets of different modalities and from different sets of cells, we need to perform diagonal integration, which is much more challenging because of no anchors, and the findings are often difficult to interpret. Diagonal

integration is based on a critical assumption that similarities are preserved between different omics layers. One of the strategies for such integrations is to group cells into different cell types within each layer first, but this has its own assumption that the cell-type signatures are the same across omics layers, which is not always true. A recent integration tool GLUE (graph-linked unified embedding) explicitly modelled the regulatory interactions between omics layers and has shown to be more accurate and robust in challenging tasks like triple-omics integration [100].

The widely used specialised packages such as Seurat can handle certain integration tasks, but more machine/deep learning-based tools are required to resolve the interaction across omics layers to build a gene regulatory network. Meanwhile, we need to create gold-standard multi-omics datasets specifically for the heart and cardiovascular diseases, which will provide the ground truth for building and evaluating new multi-omics methods. Generating such datasets requires robust study design, careful assessment of statistical power and effect size, identification of confounders, biases, and sources of variation, and quality control. A good example of such datasets is The Cancer Genome Atlas Program (TCGA) dataset [101], which has curated 2.5 petabytes of genomic, epigenomic, transcriptomic, and proteomic data from over 20,000 primary cancer and matched normal samples spanning 33 cancer types since 2006. Many multi-omics integration tools are developed using the TCGA dataset, and further investigation is required to test whether these tools can be generalised beyond the cancer research setting and readily adopted in the cardiovascular field. CRAN also recommended a series of tools for multi-omics data integration and analysis in “CRAN Task View: Genomics, Proteomics, Metabolomics, Transcriptomics, and other Omics (Version 2023-04-03)” [102]. Time-series and spatial mapping studies add further complexity to the already challenging integration task. Luckily, solutions to predict how DNA, RNA, and protein co-vary in single cells using multi-omics time-series data have been demonstrated in the prediction competition “Open Problems—Multimodal Single-Cell Integration” hosted by Kaggle last year [103]. More multi-omics integration tools with the ability to handle time/spatial dimensions will become available in the foreseeable future.

As more datasets are curated, the computational burden and storage requirement will continue to increase and become impractical for operating in an on-site environment. Fortunately, cloud-based services are becoming increasingly popular and have started alleviating these burdens. However, this approach may not be widely adopted by the research community due to the lack of funding support, training, and/or other logistic concerns. Scaling will also be challenging. The multi-omics datasets are intrinsically large, especially the single-cell omics data, which significantly increase model complexity and computing time.

Application of dimensionality reduction and representation learning methods such as autoencoders can help mitigate the issue but at the risk of losing information during the process. Finally, findings from *in silico* analyses should be validated in biological settings, which may be overlooked when using a comprehensive and powerful framework.

The Era of Big Data and Precision Medicine

As high-throughput multi-omics data and healthcare data become increasingly accessible and artificial intelligence constantly makes breakthroughs, a new era of more personalised and effective medicines is likely to emerge [104]. Machine learning has been used to develop diagnostic, prognostic, and predictive tools from healthcare data [105, 106], and linking in multi-omics data will further drive the discovery of novel biomarkers, new drugs, and personalised preventative measures and therapeutic regimens. Analysing multi-omics data using machine learning is still at an early stage even though some of the techniques have already been explored for single omics data analysis [107, 108]. Mamoshina and colleagues applied multiple machine learning methods including support vector machines, random forest, and neural networks on muscle transcriptomics data and trained age predictors to predict human skeletal muscle age. They identified key genes and pathways including PPAR signalling and neurotransmitter recycling to be associated with skeletal muscle ageing via feature importance ranking [107]. Yang et al. followed the BERT (bidirectional encoder representations from transformers) approach in the natural language processing field; developed a pretrained deep neural network-based model, single-cell BERT, extensively bench-marked; and demonstrated to be superior at cell-type annotation, novel cell type discovery, and overcome batch effects in single-cell transcriptomic data analysis [108]. The nature of multi-omics data, including high sparsity, more features than observations, and the need for integration, currently poses unique challenges to machine learning applications [109]. However, as machine learning algorithms keep evolving, combining multi-omics and clinical data to inform healthcare decisions will become “the new normal”.

Conclusion

Single-cell profiling has revolutionised biomedical research and fuelled many new discoveries in the cardiovascular field. Multiple single-cell transcriptomic atlases have been established in different species, providing valuable high-quality references for future studies. Major cell subtypes in the healthy and diseased hearts have been identified and their interactions charted. Studies of injury models across species

have provided insights into shared mechanisms of myocardial regeneration, including the cell populations involved and the relevant signalling pathways. Comparative investigations have identified regenerative programs unique to the developing hearts that are not detected in the adult heart. Therefore, to now assimilate these insights efficiently and effectively from individual studies to inform new therapeutic strategies, we propose a multi-species, multi-omics, meta-analysis framework to constantly generate, evaluate, and prioritise new targets for cardiovascular regeneration. With the growing application of machine learning and artificial intelligence, we envision that linking and analysing health-care and multi-omics data using machine learning will lead to a new era for cardiovascular regeneration.

Author Contribution ZL conceived the work. ZL, MB, and NLM wrote the manuscript, and Ziwen Li compiled the figures. All authors revised the manuscript critically for important intellectual content and provided their final approval of the version to be published. All authors are accountable for the work.

Funding MB and ZL are supported by the British Heart Foundation (Grant number FS/16/4/31831). NLM is supported by the British Heart Foundation through a Chair Award (CH/F/21/90010), Programme Grant (RG/20/10/34966), and a Research Excellent Award (RE/18/5/34216).

Data Availability Not applicable.

Code Availability Not applicable.

Declarations

Ethics Approval Not applicable.

Consent to Participate Not applicable.

Consent for Publication Not applicable.

Competing Interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Lond Engl*. 2018;392:1789–858.
2. Jaffe R, Charron T, Puley G, Dick A, Strauss BH. Microvascular obstruction and the no-reflow phenomenon after percutaneous coronary intervention. *Circulation*. 2008;117:3152–6.
3. Van Herck PL, et al. Coronary microvascular dysfunction after myocardial infarction: increased coronary zero flow pressure both in the infarcted and in the remote myocardium is mainly related to left ventricular filling pressure. *Heart*. 2007;93:1231–7.
4. Lindner JR. Microvascular impairment after myocardial infarction. *Circ Cardiovasc Imaging*. 2020;13:e011083.
5. Savarese G, Lund LH. Global public health burden of heart failure. *Card Fail Rev*. 2017;3:7–11.
6. Hashimoto H, Olson EN, Bassel-Duby R. Therapeutic approaches for cardiac regeneration and repair. *Nat Rev Cardiol*. 2018;15:585–600.
7. Porrello ER, et al. Transient regenerative potential of the neonatal mouse heart. *Science*. 2011;331:1078–80.
8. Haubner BJ, et al. Complete cardiac regeneration in a mouse model of myocardial infarction. *Aging*. 2012;4:966–77.
9. Ye L, et al. Early regenerative capacity in the porcine heart. *Circulation*. 2018;138:2798–808.
10. Poss KD, Wilson LG, Keating MT. Heart regeneration in zebrafish. *Science*. 2002;298:2188–90.
11. Pedersen K, Rasmussen RK, Dittrich A, Lauridsen H. Cardiac regeneration in the axolotl is unaffected by alterations in leukocyte numbers induced by lipopolysaccharide and prednisolone. *BMC Res Notes*. 2021;14:157.
12. Wu H-Y, et al. Fos11 is vital to heart regeneration upon apex resection in adult *Xenopus tropicalis*. *Npj Regen Med*. 2021;6:1–16.
13. Richardson W, Clarke S, Quinn T, Holmes J. Physiological implications of myocardial scar structure. *Compr Physiol*. 2015;5:1877–909.
14. Prabhu SD, Frangogiannis NG. The biological basis for cardiac repair after myocardial infarction. *Circ Res*. 2016;119:91–112.
15. Tang F, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods*. 2009;6:377–82.
16. Aldridge S, Teichmann SA. Single-cell transcriptomics comes of age. *Nat Commun*. 2020;11:4307.
17. Marx V. Method of the year: spatially resolved transcriptomics. *Nat Methods*. 2021;18:9–14.
18. Paik DT, Cho S, Tian L, Chang HY, Wu JC. Single-cell RNA sequencing in cardiovascular development, disease and medicine. *Nat Rev Cardiol*. 2020;17:457–73.
19. Tan WLW, et al. Current and future perspectives of single-cell multi-omics technologies in cardiovascular research. *Nat Cardiovasc Res*. 2023;2:20–34.
20. Kuppe C, et al. Spatial multi-omic map of human myocardial infarction. *Nature*. 2022;608:766–77.
21. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*. 2009;10:57–63.
22. Consortium*, T. T. S, et al. The Tabula Sapiens: a multiple-organ, single-cell transcriptomic atlas of humans. *Science*. 2022;376:eabl4896.
23. Schaum N, et al. Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*. 2018;562:367–72.
24. Kalucka J, et al. Single-cell transcriptome atlas of murine endothelial cells. *Cell*. 2020;180:764–779.e20.
25. Vaja R, Atanur S, Mitchell JA, Ahmetaj-Shala B. A human atlas of smooth muscle cell gene expression: insights from the FANTOM consortium CAGE dataset. *Eur Heart J*. 2022;43:ehac544.3022.

26. Buechler MB, et al. Cross-tissue organization of the fibroblast lineage. *Nature*. 2021;593:575–9.
27. Salcher S, et al. High-resolution single-cell atlas reveals diversity and plasticity of tissue-resident neutrophils in non-small-cell lung cancer. *Cancer Cell*. 2022;40:1503–1520.e8.
28. Savas P, et al. Single-cell profiling of breast cancer T cells reveals a tissue-resident memory subset associated with improved prognosis. *Nat Med*. 2018;24:986–93.
29. Song H, et al. Single-cell analysis of human primary prostate cancer reveals the heterogeneity of tumor-associated epithelial cell states. *Nat Commun*. 2022;13:141.
30. Stephenson E, et al. Single-cell multi-omics analysis of the immune response in COVID-19. *Nat Med*. 2021;27:904–16.
31. Ren X, et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell*. 2021;184:1895–1913.e19.
32. Litviňuková M, et al. Cells of the adult human heart. *Nature*. 2020;588:466–72.
33. Tucker NR, et al. Transcriptional and cellular diversity of the human heart. *Circulation*. 2020;142:466–82.
34. Shu S, et al. Cellular landscapes of nondiseased human cardiac valves from end-stage heart failure–explanted heart. *Arterioscler Thromb Vasc Biol*. 2022;42:1429–46.
35. Liang D, et al. Cellular and molecular landscape of mammalian sinoatrial node revealed by single-cell RNA sequencing. *Nat Commun*. 2021;12:287.
36. Arackal, A. & Alsayouri, K. Histology, heart. StatPearls [Internet] (StatPearls Publishing, 2022).
37. Tang J, et al. Genetic fate mapping defines the vascular potential of endocardial cells in the adult heart. *Circ Res*. 2018;122:984–93.
38. Zhou B, et al. Adult mouse epicardium modulates myocardial injury by secreting paracrine factors. *J Clin Invest*. 2011;121:1894–904.
39. George RM, et al. Single-cell evaluation of endocardial Hand2 gene regulatory networks reveals HAND2-dependent pathways that impact cardiac morphogenesis. *Development*. 2023;150:dev201341.
40. Quijada P, et al. Coordination of endothelial cell positioning and fate specification by the epicardium. *Nat Commun*. 2021;12:4155.
41. Emechebe U, et al. Age-dependent transcriptional alterations in cardiac endothelial cells. *Physiol Genomics*. 2021;53:295–308.
42. Vidal R, et al. Transcriptional heterogeneity of fibroblasts is a hallmark of the aging heart. *JCI Insight*. 2019;4:e131092.
43. Phansalkar R, et al. Coronary blood vessels from distinct origins converge to equivalent states during mouse and human development. *eLife*. 2021;10:e70246.
44. Ali SR, et al. Developmental heterogeneity of cardiac fibroblasts does not predict pathological proliferation and activation. *Circ Res*. 2014;115:625–35.
45. Li Z, et al. Single-cell transcriptome analyses reveal novel targets modulating cardiac neovascularization by resident endothelial cells following myocardial infarction. *Eur Heart J*. 2019;40:2507–20.
46. Ruiz-Villalba A, et al. Single-cell RNA-seq analysis reveals a crucial role for collagen triple helix repeat containing 1 (CTHRC1) cardiac fibroblasts after myocardial infarction. *Circulation*. 2020;142:1831–47.
47. Tani H, et al. Direct reprogramming improves cardiac function and reverses fibrosis in chronic myocardial infarction. *Circulation*. 2023;147:223–38.
48. Sutton MG, St J, Sharpe N. Left ventricular remodeling after myocardial infarction. *Circulation*. 2000;101:2981–8.
49. Ren Z, et al. Single-cell reconstruction of progression trajectory reveals intervention principles in pathological cardiac hypertrophy. *Circulation*. 2020;141:1704–19.
50. Wang L, et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat Cell Biol*. 2020;22:108–19.
51. Chaffin M, et al. Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature*. 2022;608:174–80.
52. Rao M, et al. Resolving the intertwining of inflammation and fibrosis in human heart failure at the single-cell level. *Basic Res Cardiol*. 2021;116:55.
53. Koenig AL, et al. Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat Cardiovasc Res*. 2022;1:263–80.
54. Reichart D, et al. Pathogenic variants damage cell composition and single-cell transcription in cardiomyopathies. *Science*. 2022;377:eabo1984.
55. Simonson B, et al. Single-nucleus RNA sequencing in ischemic cardiomyopathy reveals common transcriptional profile underlying end-stage heart failure. *Cell Rep*. 2023;42:112086.
56. Finan C, et al. The druggable genome and support for target identification and validation in drug development. *Sci Transl Med*. 2017;9:eaag1166.
57. Ross Stewart KM, Walker SL, Baker AH, Riley PR, Brittan M. Hooked on heart regeneration: the zebrafish guide to recovery. *Cardiovasc Res*. 2022;118:1667–79.
58. Godwin JW, Debuque R, Salimova E, Rosenthal NA. Heart regeneration in the salamander relies on macrophage-mediated control of fibroblast activation and the extracellular landscape. *Npj Regen Med*. 2017;2:1–11.
59. Ye F, et al. Construction of the axolotl cell landscape using combinatorial hybridization sequencing at single-cell resolution. *Nat Commun*. 2022;13:4228.
60. Weinberger M, Simões FC, Patient R, Sauka-Spengler T, Riley PR. Functional heterogeneity within the developing zebrafish epicardium. *Dev Cell*. 2020;52:574–590.e6.
61. Sun J, et al. hapln1 defines an epicardial cell subpopulation required for cardiomyocyte expansion during heart morphogenesis and regeneration. *Circulation*. 2022;146:48–63.
62. de Bakker DEM, et al. Prrx1b restricts fibrosis and promotes Nrg1-dependent cardiomyocyte proliferation during zebrafish heart regeneration. *Development*. 2021;148:dev198937.
63. Xia Y, et al. Activation of a transient progenitor state in the epicardium is required for zebrafish heart regeneration. *Nat Commun*. 2022;13:7704.
64. Kapuria S, et al. Heterogeneous pdgfrb+ cells regulate coronary vessel development and revascularization during heart regeneration. *Dev Camb Engl*. 2022;149:dev199752.
65. Koth J, et al. Runx1 promotes scar deposition and inhibits myocardial proliferation and survival during zebrafish heart regeneration. *Development*. 2020;147:dev186569.
66. Honkoop H, et al. Single-cell analysis uncovers that metabolic reprogramming by ErbB2 signaling is essential for cardiomyocyte proliferation in the regenerating heart. *eLife*. 2019;8:e50163.
67. Sereti K-I, et al. Analysis of cardiomyocyte clonal expansion during mouse heart development and injury. *Nat Commun*. 2018;9:754.
68. Baniol M, et al. Identification and characterization of distinct cell cycle stages in cardiomyocytes using the Fucci transgenic system. *Exp Cell Res*. 2021;408:112880.
69. Wang Z, et al. Cell-type-specific gene regulatory networks underlying murine neonatal heart regeneration at single-cell resolution. *Cell Rep*. 2020;33:108472.

70. Knight-Schrijver VR, et al. A single-cell comparison of adult and fetal human epicardium defines the age-associated changes in epicardial activity. *Nat Cardiovasc Res.* 2022;1:1215–29.
71. Xie W, et al. Single-cell RNA sequencing and assay for transposase-accessible chromatin using sequencing reveals cellular and molecular dynamics of aortic aging in mice. *Arterioscler Thromb Vasc Biol.* 2022;42:156–71.
72. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell.* 2006;126:663–76.
73. Chen Y, et al. Reversible reprogramming of cardiomyocytes to a fetal state drives heart regeneration in mice. *Science.* 2021;373:1537–40.
74. Horvath S. DNA methylation age of human tissues and cell types. *Genome Biol.* 2013;14:3156.
75. Wang H, Yang Y, Qian Y, Liu J, Qian L. Delineating chromatin accessibility re-patterning at single cell level during early stage of direct cardiac reprogramming. *J Mol Cell Cardiol.* 2022;162:62–71.
76. Schoof EM, et al. Quantitative single-cell proteomics as a tool to characterize cellular hierarchies. *Nat Commun.* 2021;12:3341.
77. Li, Z., Ross Stewart, K. M., Bruton, F. A., Denvir, M. A. & Brittan, M. Isolation of cardiocardiac endothelial cells for transcriptomic analysis of the zebrafish *Danio rerio* (zebrafish) and mouse heart. in *Angiogenesis: Methods and Protocols* (ed. Benest, A. V.) 297–309 (Springer US, 2022). doi: https://doi.org/10.1007/978-1-0716-2059-5_23.
78. Asp M, et al. A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell.* 2019;179:1647–1660.e19.
79. Boileau E, et al. Full-length spatial transcriptomics reveals the unexplored isoform diversity of the myocardium post-MI. *Front Genet.* 2022;13:912572.
80. Patra C, et al. The zebrafish ventricle: a hub of cardiac endothelial cells for in vitro cell behavior studies. *Sci Rep.* 2017;7:2687.
81. Banerjee I, Fuseler JW, Price RL, Borg TK, Baudino TA. Determination of cell types and numbers during cardiac development in the neonatal and adult rat and mouse. *Am J Physiol Heart Circ Physiol.* 2007;293:H1883–91.
82. Schafer S, et al. IL-11 is a crucial determinant of cardiovascular fibrosis. *Nature.* 2017;552:110–5.
83. Allanki S, et al. Interleukin-11 signaling promotes cellular reprogramming and limits fibrotic scarring during tissue regeneration. *Sci Adv.* 2021;7:eabg6497.
84. Oh H, et al. Cardiac progenitor cells from adult myocardium: homing, differentiation, and fusion after infarction. *Proc Natl Acad Sci U S A.* 2003;100:12313–8.
85. Pei J, et al. Multi-omics integration identifies key upstream regulators of pathomechanisms in hypertrophic cardiomyopathy due to truncating MYBPC3 mutations. *Clin Epigenetics.* 2021;13:61.
86. Alexanian M, et al. A transcriptional switch governs fibroblast activation in heart disease. *Nature.* 2021;595:438–43.
87. Ameen M, et al. Integrative single-cell analysis of cardiogenesis identifies developmental trajectories and non-coding mutations in congenital heart disease. *Cell.* 2022;185:4937–4953.e23.
88. Hocker JD, et al. Cardiac cell type-specific gene regulatory programs and disease risk association. *Sci Adv.* 2021;7:eabf1444.
89. Li Z, et al. Multi-species meta-analysis identifies transcriptional signatures associated with cardiac endothelial responses in the ischaemic heart. *Cardiovasc Res.* 2022;cvac151 <https://doi.org/10.1093/cvr/cvac151>.
90. HCA. <https://www.humancellatlas.org/learn-more/>. Accessed 6 Mar 2023.
91. Chan Zuckerberg CELLxGENE Discover. Cellxgene Data Portal. <https://cellxgene.cziscience.com/>. Accessed 6 Mar 2023.
92. Machine learning and AI in biological science, drug discovery and medicine | Royal Society. <https://royalsociety.org/science-events-and-lectures/2023/03/AI-ML-in-biology-tof/>. Accessed 15 May 2023.
93. Argelaguet R, Cuomo ASE, Stegle O, Marioni JC. Computational principles and challenges in single-cell data integration. *Nat Biotechnol.* 2021;39:1202–15.
94. Stuart T, et al. Comprehensive integration of single-cell data. *Cell.* 2019;177:1888–1902.e21.
95. Korsunsky I, et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat Methods.* 2019;16:1289–96.
96. Lopez R, Regier J, Cole MB, Jordan MI, Yosef N. Deep generative modeling for single-cell transcriptomics. *Nat Methods.* 2018;15:1053–8.
97. Argelaguet R, et al. MOFA+: a statistical framework for comprehensive integration of multi-modal single-cell data. *Genome Biol.* 2020;21:111.
98. Hao Y, et al. Integrated analysis of multimodal single-cell data. *Cell.* 2021;184:3573–3587.e29.
99. Cuomo ASE, et al. Single-cell RNA-sequencing of differentiating iPS cells reveals dynamic genetic effects on gene expression. *Nat Commun.* 2020;11:810.
100. Cao Z-J, Gao G. Multi-omics single-cell data integration and regulatory inference with graph-linked embedding. *Nat Biotechnol.* 2022;40:1458–66.
101. The Cancer Genome Atlas Program (TCGA) - NCI. 2022. <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>. Accessed 15 May 2023.
102. Aubert, J., Hocking, T. D. & Vialaneix, N. CRAN Task View: genomics, proteomics, metabolomics, transcriptomics, and other omics. 2023. <https://CRAN.R-project.org/view=Omics>. Accessed 15 May 2023.
103. Open problems - multimodal single-cell integration. <https://kaggle.com/competitions/open-problems-multimodal>. Accessed 15 May 2023.
104. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. *Nat Med.* 2022;28:1773–84.
105. Lee KK, et al. Development and validation of a decision support tool for the diagnosis of acute heart failure: systematic review, meta-analysis, and modelling study. *BMJ.* 2022;377:e068424.
106. Doudesis D, et al. Validation of the myocardial-ischaemic-injury-index machine learning algorithm to guide the diagnosis of myocardial infarction in a heterogenous population: a prespecified exploratory analysis. *Lancet Digit Health.* 2022;4:e300–8.
107. Mamoshina P, et al. Machine learning on human muscle transcriptomic data for biomarker discovery and tissue-specific drug target identification. *Front Genet.* 2018;9:242.
108. Yang F, et al. scBERT as a large-scale pretrained deep language model for cell type annotation of single-cell RNA-seq data. *Nat Mach Intell.* 2022;4:852–66.
109. Reel PS, Reel S, Pearson E, Trucco E, Jefferson E. Using machine learning approaches for multi-omics data analysis: a review. *Biotechnol Adv.* 2021;49:107739.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.