Dissecting the transcriptome in cardiovascular disease

Emma L. Robinson [ID] 1,2*, Andrew H. Baker [ID] 3, Mairi Brittan [ID] 3, Ian McCracken [ID] 3,
G. Condorelli [ID] 4, C. Emanueli [ID] 5, P.K. Srivastava5, C. Gaetano [ID] 6, T. Thum [ID] 7,
M. Vanhaverbeke8, C. Angione [ID] 9, S. Heymans1, Y. Devaux [ID] 10†, T. Pedrazzini11†, and
F. Martelli [ID] 12*†; on behalf of EU-CardioRNA COST Action CA17129

1Department of Cardiology, Cardiovascular Research Institute Maastricht (CARIM), Universiteitssingel 50, 6229 Maastricht University, Maastricht, The Netherlands; 2The Division of Cardiology, School of Medicine, University of Colorado Anschutz Medical Campus, Aurora, CO, USA; 3Centre for Cardiovascular Science, Queen's Medical Research Institute, University of Edinburgh, 47 Little France Crescent, Edinburgh, EH16 4TJ, UK; 4Humanitas Research Hospital, Humanitas University, Via Manzoni 113, Rozzano, MI 20089, Italy; 5Imperial College, National Heart and Lung Institute, Hammersmith campus, Du Cane Road, London W12 0NN, UK; 6Laboratorio di Epigenetica, Istituti Clinici Scientifici Maugeri IRCCS, Via Maugeri 4, Pavia 27100, Italy; 7Hannover Medical School, Institute of Molecular and Translational Therapeutic Strategies (IMTTS), Carl-Neuberg-Straße 1 30625 Hannover, Germany; 8UZ Gasthuisberg Campus, KU Leuven, Herestraat 49 3000 Leuven, Belgium; 9Department of Computer Science and Information Systems, Teesside University, Middlesbrough, TS4 3BX, UK; 10Cardiovascular Research Unit, Department of Population Health, Luxembourg Institute of Health, 1A-B, rue Thomas Edison, L-1445 Strassen, Luxembourg; 11Experimental Cardiology Unit, Division of Cardiology, Department of Cardiovascular Medicine, University of Lausanne Medical School, 1011 Lausanne, Switzerland; and 12Molecular Cardiology Laboratory, IRCCS-Policlinico San Donato, Piazza Edmondo Malan, 2, 20097 San Donato, Milan, Italy

Abstract	The human transcriptome comprises a complex network of coding and non-coding RNAs implicated in a myriad of biological functions. Non-coding RNAs exhibit highly organized spatial and temporal expression patterns and are emerging as critical regulators of differentiation, homeostasis, and pathological states, including in the cardiovascular system. This review defines the current knowledge gaps, unmet methodological needs, and describes the challenges in dissecting and understanding the role and regulation of the non-coding transcriptome in cardiovascular disease. These challenges include poor annotation of the non-coding genome, determination of the cellular distribution of transcripts, assessment of the role of RNA processing and identification of cell-type specific changes in cardiovascular physiology and disease. We highlight similarities and differences in the hurdles associated with the analysis of the non-coding and protein-coding transcriptomes. In addition, we discuss how the lack of consensus and absence of standardized methods affect reproducibility of data. These shortcomings should be defeated in order to make significant scientific progress and foster the development of clinically applicable non-coding RNA-based therapeutic strategies to lessen the burden of cardiovascular disease.
Keywords	Transcriptomics • Non-coding RNAs • Methodology standardisation • Translational cardiovascular research

1. Introduction

The completion of the human genome project in 2003—an international joint public and private effort—is a date that will go down in history.[1–3] Sequencing of all 46 chromosomes in the human genome revealed that we possess less than 25% of the predicted 100,000 protein-coding genes, a surprisingly similar number of protein-coding genes to that in the nematode worm.

We now know that the human transcriptome is much more complex than previously thought. The advent of deep sequencing tools of the last

---

15 years has revealed a complex network of at least 10,000 long non-coding RNAs, encoded in a myriad of genomic contexts.[4]

Non-coding RNAs are functional RNA molecules that lack protein-coding potential. Only a subset have so far been characterized, and many exhibit high spatial and temporal expression patterns and are emerging as key regulators of differentiation, development, homeostasis, and disease pathology, including in cardiovascular disease (CVD).

Delineating the transcriptome changes in cardiac cell types and defining their regulation in disease remodelling could not only enable us to further understand the cellular and molecular processes underlying pathology, but also provides the opportunity to select new, specific and more effective therapeutic targets to treat CVD.

This review describes the challenges in dissecting and understanding transcriptome remodelling in CVD, focusing on non-coding genes. We will then discuss strategies being implemented to overcome these confounding factors. Finally, we review hurdles yet to overcome to improve reproducibility of data, convincingly delineate the cellular and molecular processes underlying cardiovascular pathology and catalyse translational impact.

MicroRNAs (miRNAs) are 21- to 23-nt-long short, single-stranded non-coding RNAs involved in post-transcriptional regulation of gene expression through binding to complementary sites predominantly in the 3' untranslated regions (UTR) of pre-messenger RNA (mRNA, protein-coding).[5] This binding event represses translation of the mRNA, often targeting it for degradation. At least 60% of protein-coding genes are miRNA targets. Each miRNA can negatively regulate a number of different mRNAs and each of the miRNA-regulated protein-coding genes can be targeted by a number of miRNAs. This pleiotropic property of miRNAs has its advantages and disadvantages with regards to therapeutic potential. Complex diseases, including CVD, are most commonly multi-factorial and multi-genic, a consequence of multiple genetic and non-genetic factors. A miRNA often targets a number of mRNAs that converge on common biological pathways and processes. A large number of miRNAs are dysregulated in CVD, in the heart or blood, identified through pre-clinical and translational studies.[6]

MiRNA therapy has gained a lot of attention due to its specific targetability using chemically modified antisense technology, enabled by the inherent complementary base-pairing properties of nucleic acids. Whilst a handful of antisense therapies have been approved to date by the US Food and Drug Administration (FDA) targeting messenger RNAs, no miRNA drugs have thus far made it onto the market. Contributing to the low rate of translation to the clinic is a lack of consensus methodologies and experimental approaches. Pre-clinical work in large animal models has preceded clinical trials, which are currently ongoing.

Another facet of miRNA research in biomedical science is as circulating predictive, diagnostic, and prognostic biomarkers of disease. Blood, plasma, serum, and other biological fluids such as urine and tear drops have been found to be rich in RNA, either through release into the circulation from necrotic or apoptotic cells or having been actively secreted in microvesicles such as exosomes.[7]

Long non-coding RNAs (lncRNAs) are defined as transcribed non-protein-coding RNA molecules longer than 200 nt in length. LncRNA genes can overlap protein-coding genes, be located in an antisense orientation to protein-coding genes, intronic or intergenic (also known as long intervening noncoding RNAs (lincRNAs).[8] In general, lncRNAs are less well conserved than protein-coding genes and exhibit a greater degree of temporal expression patterns and cell-type specificity. Their biological functions are as heterogeneous and widespread as proteins.

LncRNAs are emerging as key mediators of differentiation, development, homeostasis, ageing, and disease.

Among lncRNAs, enhancer RNAs (eRNAs) describe a family of non-coding transcripts generated from enhancer regions of annotated genes that have a cis-regulatory role in transcription. Enhancer RNAs have been identified as regulators in cardiomyocyte(CM) differentiation development and homeostasis as well as being altered in response to pathological stimuli.[9–11]

A further recently discovered class of non-coding RNAs are the circular RNAs (circRNAs). CircRNAs are single-stranded and created by back splicing of 3'-5' linear coding or non-coding RNAs, forming covalently closed loops.[12] Originally thought to be a result of mis-splicing, circRNAs are now emerging as a major class of regulatory RNA molecules in disease and development. More than 30,000 circRNAs are now described in the human genome and diverse biological functions described and predicted including microRNA sponging and sequestration and protein binding affecting their stability.[13]

CircRNAs are being examined not only for their intracellular activity but also as stable and specific circulating biomarkers reflecting cardiac dysfunction and pathology. The stability of circRNAs in the circulation exceed that of linear RNAs and their diagnostic value for CVD has started to be evaluated.[14,15]

We are now aware that a significant number of lncRNAs and unannotated transcripts originally defined as purely non-coding have short, translated open reading frames (sORFs).[16] A number of strategies, including conservation analysis, have revealed micropeptides with important functional roles in biological processes. Accordingly, a recent sophisticated high-throughput examination of the 'translatome' in the healthy and diseased human heart by ribosome profiling revealed a vast number of potentially translated sORFs, including within 77 lincRNAs and 64 antisense transcripts.[17] Validation and biological relevance of these translatable regions remains to be elucidated.

2. RNA sources and RNA integrity

Whilst some aspects of CVD can be modelled in animals, due to our lack of understanding along with practicalities, biological differences, and multi-factorial aetiologies, human CVD can never be perfectly emulated in pre-clinical models. Thus, to confirm the translational potential of findings in cellular and animal studies, human specimen can be used to show conservation of transcriptome analyses.

Human specimens used in CVD research include, but are not limited to, liquid biopsies such blood, plasma, serum, buffy coat, isolated blood cells, and solid tissues, such as cardiac biopsies, explanted cardiac tissue from heart transplant patients and hearts extracted from cadavers of victims of sudden death (so-called 'donor' hearts). Guidelines for standardized sample collection and storage are not well established, which leads to variability in the quality of RNA available.

Isolation of good quality, intact RNA is required for techniques used for RNA and gene expression analysis including RTqPCR, RNA-sequencing (RNA-seq) and northern blotting. RNA integrity can be assessed qualitatively by running on a denaturing gel (e.g. with formaldehyde) and stained with a nucleic acid stain such as ethidium bromide. Intact RNA will give sharp, clean 28S and 18S ribosomal RNA (rRNA) bands, with the 28S approximately twice as intense as the 18S band. Degraded RNA will lack these sharp rRNA bands with a 2:1 ratio and have a smear-like appearance. A faint smear may appear between 1–2 kb, containing other

RNAs such as mRNAs and lncRNAs. With the knowledge that at least 80% of the total cellular RNA is rRNA, RNA integrity can additionally be assessed with the knowledge of the mass of total RNA loaded and the rRNA bands compared against an RNA ladder of known marker size and mass. Elementary variations of this method have been devised to avoid the use of hazardous denaturing substances and improve resolution.[18,19] Whilst gels offer a fast and simple means to assess RNA integrity with basic equipment used in molecular biology, even the latest nucleic acid gel imaging technologies and analysis tools are considered semi-quantitative at best.

The gold standard for assessing RNA integrity is the 2100 Bioanalyzer (Agilent). This device provides an automated electrophoresis strategy using microfluidics technology for RNA integrity analysis based on the relative rRNA species intensity. A range of microfluidics chips are available for different concentrations of RNA sample, with as little as 50 pg required by the Pico RNA analysis products. The 2100 Bioanalyzer calculates an RNA Integrity Number (RIN), which is an algorithm for robust and reliable assignment of quality values to RNA samples.[20] RIN calculation still includes 28S and 18S ratio as well as any RNA species detected between the 18S and smaller 5S band. An alternative to the 2100 Bioanalyzer, also by Agilent, is the TapeStation, which also uses capillary electrophoresis but requires slightly higher volumes and concentrations of RNA (at least 2 ul of $\geq$ 5 ng/ul).

Of note, these methods to assess RNA integrity are based on ribosomal RNAs. This is not suitable for assessing the quality of nuclear RNA or polyA RNA preparations, which are depleted of rRNA. Some semi-quantitative methodologies have been suggested for assessment of polyA RNA samples.[21]

There is a lack of consensus regarding the minimal RIN value that should be used for accurate execution of RNA analysis. Moreover, different techniques for RNA analysis may be able to withstand different RNA integrity thresholds to be accurate and reproducible. For example, cDNA library synthesis requires initial RNA degradation into fragments either through heating or enzymatic digestion. Many researchers and next-generation sequencing (NGS) technology companies and facilities arbitrarily state a minimum RIN value of 7 or 8 for deep RNA-sequencing. In the context of clinical samples in translational cardiovascular research, this minimum RIN is often unrealistic, with the example of fluid biopsies remaining on the bench at room temperature while the cardiologist tends to the patient.

In the context of transcriptome analysis, a number of confounding factors can affect the outcome of research findings. In particular, the quality and integrity of RNA, determined by a number of pre-analytical variables along with sampling and storage methods.[22] *Table 1* provides an overview of confounding factors and pre-analytical variables affecting RNA integrity and outcomes of gene expression analysis.

The post-mortem interval (PMI) is a primary confounding factor affecting quality and integrity of subsequently isolated biological material, such as DNA and RNA.

Firstly, at ambient temperatures, RNA is highly labile and RNase enzymes remain catalytically active, naturally causing a decrease in RNA integrity with increased PMI. Evidence also suggests that not all RNAs are created equal, some being more susceptible to degradation.[31]

Secondly, one could envisage stress-associated gene programmes, such as those induced by hypoxia, to be activated in the early PMI period. These two main influences can greatly affect the outcome of transcriptome analysis in CVD.

For extraction of tissues from sudden-death victims, the exact PMI is not always easy to determine and will inevitably not be consistent between samples. This is a common means by which tissues from 'healthy/control' individuals are collected. PMI in this case is usually determined by a pathologist based on physical observations of the corpse such as body temperature and *rigor mortis*. What's more, heart tissue collected during procedures involved in treatment and investigation of cardiovascular disease, such as endomyocardial biopsies, valve replacements or heart transplants, is handled very differently, generally placed into cardioplegic solution followed by appropriate and rapid storage.

Significantly, an analysis of GTEx (Genotype-Tissue Expression project) tissues and expression data of different human tissues with PMIs ranging from 1 to 27 hours and good quality RNA-seq data is available. A systematic analysis showed that cardiac and skeletal muscle displayed the greatest influence of PMI on the transcriptome, with over 850 genes significantly differentially expressed depending on PMI.[23]

Comparing the transcriptome profiles between diseased and healthy heart samples, the vastly different and heterogeneous PMIs and handling protocols will influence the findings. In addition, there are no published guidelines set by standardisation bodies or networks of experts

Table 1 Pre-analytical variables in translational cardiovascular research

Variable	Applicability to translational cardiovascular research	References
(1) Post-mortem interval	The time between death and sample collection and appropriate storage	[23–25]
(2) Anti-coagulants	The effects of anti-coagulants can present themselves in a number of ways in translational research. Many cardiovascular patients are administered drugs such as heparin or warfarin, which not only affects blood clotting but can also interfere with downstream molecular applications such as cDNA synthesis	[26,27]
	Anti-coagulant agents including EDTA and citrate are also used in the preparation of plasma, serum or white blood cells	
(3) Pre-processing interval	The time between sample extraction and processing (in the case of plasma, serum or white blood cell isolation) or storage (e.g. freezing cardiac samples)	[28]
(4) Processing method	The protocol used for extraction of plasma, serum or white blood cells isolation	
(5) Samples storage	Receptacles for liquid biopsies and solid tissue can affect preservation of RNA and DNA (e.g. cryo-tubes pre-coated with anti-coagulants such as EDTA or citrate or nuclease free tubes)	[29,30]

presenting consensus procedures in sample attainment, handling, or storage. A first step to improve reproducibility should be a detailed description of sample harvesting and handling in all scientific publications, information that is often missing.

The effect of the pre-processing time, as well as RNA integrity on the global depth of RNA-seq and on the outcome of gene expression analysis, was evaluated in only two studies of human peripheral blood mononuclear cells and human tissue (human placental samples).[32,33] RNA integrity notably affected the results of gene expression analysis.

The European Commission has acknowledged a need for international unity in governance and coordination of clinical biobanks.[34] Clinical biobanks need to adhere to minimum information requirements as set out in the Minimum Information About Biobank Data Sharing (MIABIS) tool developed by the EU BBMRI-ERIC research infrastructure.[35]

In conclusion, careful and correct collection, storage and sharing of information on pre-clinical variables as well as RNA integrity, evaluated by the gold standard RIN value, will enable consideration of these elements in translational cardiovascular research. These factors can then be incorporated as covariables in analysis platforms to identify independent gene expression changes.

3. Cell-type-specific vs. whole tissue transcriptomic changes

The heart is a highly heterogeneous organ, composed of at least six major cell types. The relative composition of the mammalian heart is a debated area, but it is generally accepted that CMs, cardiac fibroblasts (CFs), and endothelial cells (ECs) are the dominant cell types.[36,37] Other minor cell types include smooth muscle cells, pericytes, and macrophages. Relative cell type composition changes through development, ageing, and disease, with proliferation of most cell types occurring during mammalian development and growth. Conversely, ageing and disease are characterized by CM hypertrophy, attrition and apoptosis, by CF hyperproliferation and transformation into myofibroblasts, and by intramyocardial infiltration of immune cells. *Figure 1* depicts the composition of the adult heart and how different cell types are typically altered in response to chronic pathological stimuli.

Alike, the vasculature is composed of smooth muscle cells, ECs, pericytes as well as a significant layer of connective tissue and elastic fibres. Remodelling of the vessel bed composition, cellular biology, morphology, and structure occurs in injury and atherosclerosis.[38]

Considering all the aforementioned factors and substantial remodelling of cardiac cellular composition in disease, transcriptomic data from whole heart tissue are not representative for a true comparison of changes within a particular cell type and make translating bulk molecular changes to biological function difficult. The resulting data sets represent only composite changes across the constituent cell types for that sample, often diluting changes that are of relevance in cellular and functional remodelling.

A number of established protocols exist for isolation of CMs, CFs, ECs and immune cells from fresh heart tissue from animals and humans. These methods are centred around enzymatic digestion, often involving specialist equipment such as Langendorff apparatus for lengthy periods of aortic and coronary perfusion. Achieving high quality, high purity and consistent cell preparations proves variable between protocols, different
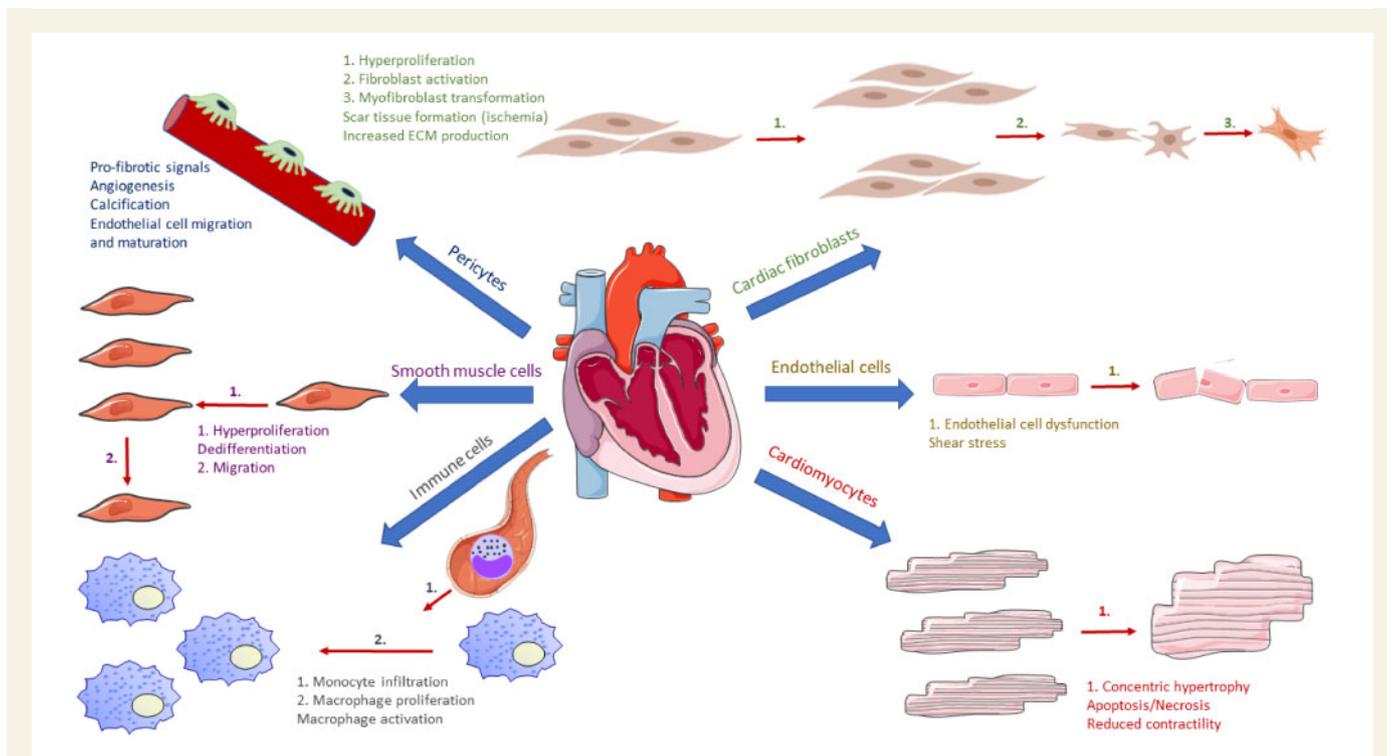
Figure 1 Depiction of the major cell types of the mature mammalian heart along with examples of ways cellular morphology and function can change in disease. Numbered arrows refer to the processes numbered in the text. Text and numbering is colour-coded according to cell type. From far-right clockwise: endothelial cells, cardiomyocytes, immune cells, smooth muscle cells, pericytes, and cardiac fibroblasts.

laboratories as well as between species, samples and experimental conditions.

One further significant caveat to isolation and separation of different cell types by enzymatic digestion is that fresh tissue is required, which needs to be maintained in cardioplegic solution since the point of extraction. Not only in some cases is access to fresh tissue impractical, such as in the case of biobanked human material, but the time of the preparations undertaken at ambient temperatures and conditions are conducive to DNA synthesis, transcription, RNA degradation as well as epigenetic modifier activity during the preparation time.[39]

Furthermore, as previously indicated, anti-coagulants such as heparin administered during extraction of the heart as well as anti-thrombotic medication in patients, have been shown to interfere with cDNA synthesis reactions used in RT-qPCR or RNA-seq library preparations.[26]

Variations on these methods for fresh CM and CF isolation have been developed with the aim of optimizing purity, cellular integrity, and minimizing the need for anti-coagulants and complex apparatus, including through direct needle perfusion of the extracted heart *ex vivo*.[40,41]

All confounding factors and circumstances considered, further approaches have been formulated to obtain cell type-specific gene expression or epigenomic information. Whilst intact living cells can only be isolated from fresh cardiac tissue, nuclei can be extracted from frozen tissue. Fluorescence immunolabelling of cell-type-specific nuclear antigens followed by flow cytometry enables collection of cell-type-specific DNA and nuclear RNA. Examples include the use of pericentriolar material-1 (PCM-1), nuclear-retained phospholamban, or troponin T for CM nuclear isolation.[42–45] A further benefit of this approach is that all steps can be performed at ice-cold temperatures and with addition of inhibitors of RNA Pol II and epigenetic mediators, such as sodium butyrate to suppress histone deacetylase activity, thereby reducing preparation-associated transcriptome and epigenetic remodelling.

It should be noted that nuclear transcriptome data give a different readout to cellular or bulk tissue RNA analysis. Firstly, 90% of the total RNA in the cell is in the cytoplasm as mature processed RNA species. The nucleus is home to pre-processed RNA such as pre-mRNA, pri- and pre-miRNAs, and nascent lncRNA transcripts. In addition, nuclear-localized RNA species, such as lncRNAs, will be overly represented in the nucleus. Nuclear content gives a readout of more of the nascent transcription rates whereas cellular RNA is reflective of additional post-transcriptional regulation such as miRNA mechanisms and mRNA stability.[46,47] Provided a pure nuclear population is isolated, transcripts from the 37 mitochondrial encoded genes will also be depleted. That said, direct comparison of results of global tissue or cellular vs nuclear RNA in a variety of models and species found strong correlations between the data sets from different origins.[48]

One further approach to deducing the relative contribution of cell types in big transcriptomic data from bulk tissue or a mixed cell population is through bioinformatic deconvolution using established knowledge drawn from gene expression profiles of the individual composing cell types. One such example of where this strategy is employed very effectively is with whole blood specimen, where gene expression profiles of blood cell types—predominantly white blood cells—and sub-types are well described.[49] An example of one such pipeline that has been developed for this purpose is Decon2.[50]

Major advantages of this strategy include minimization of confounding factors involved in tissue separation and specific cell or nuclear isolation and sequencing, with the resulting data representing the relative cell population and tissue transcriptome close to that in vivo, provided appropriate extraction and storage of the biospecimen.

That said, published examples of where bioinformatical deconvolution in this manner has been successfully employed in cardiovascular tissues are lacking. One probable reason for this is that transcriptional dysregulation is commonly seen in the context of disease, with cells of the cardiovascular system ectopically expressing genes atypical of their differentiated cell state. For example, CMs have been shown to express extracellular matrix protein in pro-fibrotic conditions and as fibroblasts are activated towards a myofibroblast state, express contractile fibres.[51,52] Leading up to the following section of this review article, with further single-cell RNA-sequencing (scRNA-seq) data being generated from cardiovascular tissues and shared through data archives, we can hope that bioinformatic deduction of the relative contribution of cell types and sub types to bulk tissue expression changes will soon evolve in the cardiovascular field.

4. Single-cell transcriptomic analysis

The advent of scRNA-seq technology provides an unparalleled opportunity for high-resolution profiling of the global transcriptional signature of individual cells in models of disease and human tissue. Since the seminal report of transcriptome sequencing at the single-cell level of a mouse blastomere a decade ago, there has been an early exponential rise in studies applying this technology in the cardiovascular system.[53] ScRNA-seq permits us to catalogue the cellular composition of a tissue and to classify cell and transcriptional heterogeneity and plasticity under different conditions in an unbiased fashion. Therefore, scRNA-seq holds great promise as an accessible technology that can provide pivotal in-depth insight into the molecular regulators associated with CVD and regeneration.

Analyses of bulk tissues produce signals that are an average across several cell-types in the tissue and so are unsuited to defining the role of individual cells. However, scRNA-seq allows the characterization of gene expression changes in individual cells and therefore an opportunity to deconvolve both cellular and disease-induced heterogeneities that serve biological functions.

In the heart, several papers reporting single-cell transcriptional profiling of cardiac lineages from adult mice have emerged over the past two years. In a landmark study, single-cell transcriptomic was applied to map changing cell subpopulations in the healthy and injured mouse heart 3 days after ischaemia-reperfusion.[54] This identified *cytoskeleton-associated protein 4* (*Ckap4*) as a specific marker of activated CFs following injury. In another study, Forte and colleagues used scRNA-seq to characterize the dynamic transcriptional changes in non-CM cells at multiple timepoints following MI, identifying the occurrence of an early myofibroblast response as a predictor of cardiac rupture.[55] A high-resolution single-cell gene expression atlas of isolated mouse cardiac ECs defined 10 transcriptionally-distinct, heterogeneous cardiac EC states resident in the adult mouse heart at a time of peak neovasculogenesis following ischemic injury, and provided a rich resource of targets with a potential role in vascular regeneration of the injured myocardium.[56] Sex-dependent transcriptional differences in cardiac cell subtypes were also shown, which may provide fundamental insight into the molecular cues that underpin sexually dimorphic responses to cardiac injury.[57] In the last year, studies have been published using single nuclear and cellular sequencing from cardiac tissue describing the transcriptional landscape and heterogeneity in different chambers of the heart and between sexes in the context of heath and heart failure (HF).[58,59] ScRNA-Seq has also revealed

the extent of immune system activation in a mouse model of pressure overload-induced HF.[60]

Outside the heart in the peripheral vascular setting, a number of studies have reported scRNA-seq of cells in blood vessels, for example in healthy and atherosclerotic mouse tissue.[38,61,62] This has characterized the transcriptional signatures of specific vascular lineages and subpopulations, including the molecular profiling of a potential endovascular progenitor cell, with a large cluster of ECs expressing clear mesenchyme-associated genes even in healthy aorta. Two independent relatively large-scale studies have used scRNA-seq to map the transcriptional landscape associated with pluripotent stem cell commitment to vascular lineages, with a focus on ECs.[63,64] McCracken sequenced over 100,000 single cells at several time-points during the differentiation protocol to identify many previously unknown transcription factors that are activated upon commitment of pluripotent cells to the endothelial lineage.[64] This knowledge may be key for future studies aiming to interrogate specific facets of vascular biology such as specification to distinct endothelial types, perhaps even organ-selective endothelial types.

See and colleagues applied single nuclear sequencing technology (snSeq) to identify both coding and non-coding RNA targets in CMs isolated from the left ventricles of healthy and failing hearts in mice and humans.[65] Core markers were defined, including non-coding RNAs, expressed by healthy CM subpopulations and common to both mice and humans. This study further characterized transcriptional heterogeneity in stressed CMs and identified lncRNA candidates, including *Gas5* and *Sghrt,* as regulators of gene expression networks associated with CM proliferation.

The application of single-cell technology has highlighted dynamic transcriptional plasticity and heterogeneity, including classification of previously unreported or rare cell subpopulations. However, discrepancies are already apparent. For example, Farbehi *et al.* found widespread expression of *Ckap4* across virtually all cardiac stromal populations in their analyses, dissimilar to the findings of Gladka and colleagues that *Ckap4* was specific to injured and activated CFs.[54,66] This can likely be explained by technical and methodological discrepancies such as differences in the number of cells sequenced, read depth per cell, thresholds applied to filter out background 'noise', and a lack of consensus over a standardized workflow and tools for analysis, as described in *Figure 2*. Great care must be taken when interpreting individual datasets following stand-alone analysis, and recent studies have emphasized the benefit of aligning multiple independent datasets using integrated bioinformatics pipelines, including from different species, as a means to identifying more robust and conserved targets. Archiving of raw sequencing datasets in open-access repositories may help to promote reproducibility and reduce potential issues caused by technical and biological discrepancies. An overview of the pros and cons of scRNA-seq technology is displayed in *Figure 3*.

Although several statistical frameworks exist for differential gene expression (DE) analysis for the bulk RNA-seq, methods for scRNA-seq are just emerging.[67,68] The real challenge in sn/scRNA-seq DE analysis is to take into account particular characteristics of scRNA-seq which include low library sizes, high noise levels, and a large fraction of dropout events.[69] In addition, the stochasticity of gene expression in cells results in multimodality of gene expression across different cells.[70] To address these challenges, several single-cell-specific methods have been developed. One example is that of model-based analysis of single-cell transcriptomics (MAST), which proposes to develop a generalized linear model to take into account bimodal distribution arisen as a result of either detectable transcripts or non-detectable genes that may have arisen because of stochastic dropout events.[71] Although MAST has been evaluated to be one of the best performing statistical frameworks for detecting DE genes for single-cell datasets, DE analysis of scRNA-seq data remains to be a challenge, as the tools developed for scRNA-seq focus to model either dropout events or multimodality but not both.[70]

The fast-paced evolution of snRNA-seq technology, and emerging applications such as snSeq of frozen tissues, will undoubtedly continue to provide critical new insights into cardiovascular homeostasis and disease processes that then inform critical proof-of-concept mechanistic studies. To complement the information obtainable by these techniques based on tissue disaggregation, another emerging application of next-generation sequencing is that of spatial transcriptomics. This method is an expansion on in situ hybridization into an unbiased high throughput method to analyse the presence and abundance of transcripts in a two- or three-dimensional context in complex tissues. One such example, Tomo-Seq (RNA Tomography for Spatially Resolved Transcriptomics), is able to obtain a genome-wide gene expression signature with high-spatial 2-D resolution in a mouse model of myocardial infarct (MI) spanning from the infarcted area to the remote area to identify new regulators of cardiac remodelling as well as in a zebrafish model of cryoinjury to identify mediators of cardiac regeneration.[72,73] Other newly established methods include solid-phase bar-coded RNA capture and so-called 3D-Cardiomics from which a transcriptome map has been constructed from microdissections of the healthy mouse heart.[74] One limitation to be noted is that, to eliminate genomic DNA incorporation into library construction, current methods for spatial transcriptomics rely on polyA capture and detect mature mRNA species only.

5. RNA splicing and circular RNAs

Most mRNAs and lncRNAs generated in mammalian cells are composed of exons separated by introns. Certain introns are spliced-out constitutively, but many splicing signals are not used by all cells nor in all circumstances, leading to alternative splicing (AS). AS allows expansion of the genome to generate a diverse array of functional RNA transcripts from a single gene. Being estimated to occur in about 90% of human genes, AS greatly increases the coding potential of the genome.[75] Of note, AS can generate not only different isoforms of the same protein, but also functional non-coding RNAs. Another example of splicing is that of the aforementioned 3'- to 5'-back-splicing-generating circRNA species.[12] The circRNA landscape expressed in the human and mouse heart as well as in differentiating CMs has revealed a large proportion of gene loci from stable detectable circRNAs.[76] The most abundant of these, *circSlc8a1-1*, was found to act as a *miR-133a* competing endogenous sponge, and its inhibition in pre-clinical pressure-overloaded models attenuated the pathological cardiac hypertrophic response.[77] Outside of the heart, circ_Lrp6, a circRNA enriched in vascular smooth muscle cells, regulates miRNA-145 function through acting as a sponge in vascular smooth muscle cells.[78]

It has been known for a long time that transcripts that are important for cardiac homeostasis and function can be aberrantly spliced in human heart disease, but the magnitude of mis-splicing events has only become evident in the past decade. Widespread use of microarrays and RNA-seq techniques has established a robust association of altered AS of pivotal cardiac genes, including many sarcomeric and ion channel genes, with cardiac hypertrophy, dilated cardiomyopathy (DCM), and HF.[79,80] An outline for the workflow for analysing splice variants from RNA-seq data with examples of computational tools available for each step is shown in *Figure 4*.

Figure 2 Schematic of a typical scRNA-seq workflow: single cells are separated and sequencing libraries generated using either droplet or plate-based scRNA-seq platforms. Raw sequencing data are processed and aligned to the reference genome to generate count matrices. Following quality control (QC) and normalisation, dimension reduction is performed to facilitate visualisation in low dimensional space. Subsequent clustering is performed to group cells by their similarity in transcriptional profile. Frequently used bioinformatic tools for each of these steps are indicated on the right side.

The first generation of transcriptomic analyses used exon-microarrays to identify splicing events, and, more recently, RNA-seq has revealed the pervasiveness of AS in an unbiased manner.[79,81] AS regulation is mediated by the interaction between *cis*-elements around the splicing site in both introns and exons, and *trans*-acting RNA-binding proteins (RBPs) that bind to specific *cis*-elements. RBPs are dynamically regulated with a variety of mechanisms and can promote or inhibit the splicing of a specific intron. A subset of RBPs regulating AS in the heart have been identified, including RBFOX-1 and -2, MBNL1 (muscleblind-like protein 1), CELF1 (CUGBP Elav-Like Family Member 1), and RBM-20 and -24 (RNA Binding Motif Protein 20 and 24).[75] Evidence of the importance of these RBPs for heart physio-pathology derives mostly from genetic engineering of mouse models and cell culture experiments. Indeed, so far, only mutations in RBM20 have been shown to be associated with cardiomyopathy.[82]

RBM20 mutations cause an early-onset and clinically aggressive form of familial DCM, associated with aberrant inclusion of specific exons in the Titin (TTN) transcript.[83] This, in turn, results in the expression of very large and compliant TTN isoforms affecting the physiological heart passive stiffness.[83]

TTN is a very large multi-exon gene encoding many transcript variants regulated by still poorly understood AS events. Transcriptomic analysis of cardiac-specific knockout mice indicates that TTN is also one of the direct splicing targets of murine Rbm24.[84] It remains an open question whether human RBM24 is a cardiomyopathy-associated gene, since no mutations in the three main isoforms of RBM24 have been identified in DCM patients.[85] Of note, specific circRNAs originating from TTN are also absent in RBM20 null mice.[86] Another RBP, Quaking, also regulates the formation of TTN-circRNAs, that are likely mediators of Quaking cardio-protective effects.[87]

RBPs can engage in antagonistic actions, as observed for CELF1 and MBNL1. While CELF1 decreases during heart development, MBNL1 increases.[88] Manipulation of these RBPs in the adult heart replicating their levels in the embryo results in reactivation of the embryonic splicing pattern.[88] Indeed, re-expression of a foetal splice variant programme is a phenomenon characterizing most cardiomyopathies.[81]

Figure 3 Advantages (Pros) and challenges (Cons) of single-cell sequencing technology.

In particular, in myotonic dystrophy type 1 (DM1), extensive AS changes result from the CUG expansions in the DMPK gene characterizing the disease.[89] These triplet-repeat RNAs accumulate in the nucleus, acting as molecular sponges and altering the function of MBNL, CELF1, and other splicing factors, leading to re-expression of foetal isoforms in adult tissues, including the myocardium. Accordingly, after respiratory failure, cardiac arrhythmias are the most common cause of death in DM1 patients.[90]

Another mechanism of AS regulation is made possible by splicing occurring co-transcriptionally for most genes, allowing the interaction between the splicing and the RNA-synthesis machinery.[91] Thus, AS can be affected by RNA transcription kinetics, as well as by chromatin modifications and structure.

Many computational methods have been developed to analyse RNA-seq datasets at exon level. All these methods rely on reads spanning each exonic junction, to estimate an abundance ratio between the two spliced isoforms (exclusion and inclusion isoforms), followed by a statistical test to assess the significance of the observed difference in splicing. Although implemented with a variety of statistical frameworks, all available methods suffer from high uncertainty for the AS events sampled with low sequencing coverage. High sequencing depth is necessary to obtain a representative picture of the gene expression pattern at isoform level. The situation may be even more critical for circRNAs: these RNA species are devoid of a 3'-end and published datasets deriving from poly-A positive RNA cannot be used for their detection.

A further, so-called third-generation sequencing technology, which may be beneficial in the identification of novel transcript isoforms, is that of long-read sequencing. Such long-read sequencing technologies have been developed by, for example, Pacific Biosciences PacBio or single-molecule, real-time (SMRT) sequencing and Oxford Nanopore Technologies (ONT) and are able to sequence full-length transcripts in excess of 10 kb in length. The advantages of long-read sequencing over classic short read include lower error rates in mapping, and reduce amplification bias, easier accurate assembly of regions of the genome comprising repetitive sequences, and *de novo* gene assembly for poorly annotated genomes and genes as well as valuation of copy number variants. This has been reviewed extensively previously.[92] Long-read sequencing has proven an effective strategy to identify novel transcripts and super repeat regions in very large muscle structural genes such as TTN and Nebulin (NEB), the biological relevance of which in cardiac and skeletal muscle is yet to be determined.[93]

Although there has been significant progress in the characterisation of isoform-level transcriptomics and the discovery of the regulatory aspects of AS networks, demonstration of the functional consequences of the expression of different isoforms of individual genes is still scarce. Therefore, understanding the physio-pathological relevance of AS is one of the biggest challenges of CVD functional genomics.

6. Subcellular compartmentalisation of RNA

Unequal cellular distribution is characteristic of the majority of RNA species. Concentrations of RNAs at specific locations are highly associated
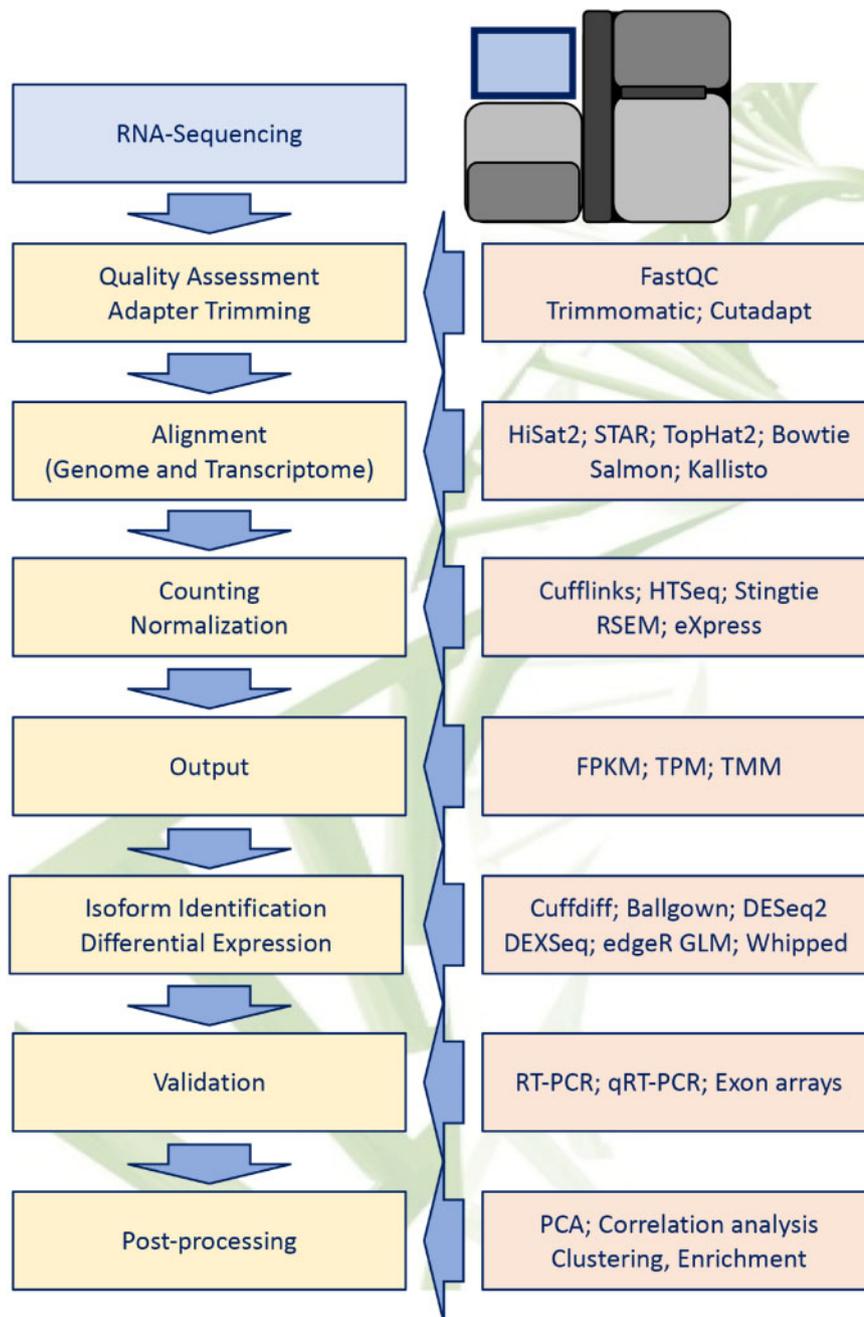
Figure 4 Schematic of typical alternative splicing (AS) analysis workflow using from RNA-seq data. Frequently used bioinformatic tools for each are indicated on the right side of each step.

with their functionality. A prime example is the transport of mRNAs from their site of production, the nucleus, to the cytoplasm, where they are actively translated.[94] Binding to ribosomes creates a microdomain enriched in mature mRNAs. In addition, subcellular concentrations of ribosome-bound mRNAs provide a means of controlling local protein production, allowing creation of functional protein gradients. In turn, localized protein translation regulates key cellular processes such as specification and differentiation. In the heart, Mbnl regulates mRNA localisation via binding to 3'UTRs of transcripts encoding membrane proteins, consistent with the Mbnl localisation at the periphery of the

cell.[88] In CMs, YB-1, a single-strand nucleic acid-binding protein that regulates expression of proteins associated with tissue repair, accumulates within the polysomes in the sarcoplasm, close to the intercalated disk, resulting in increased smooth muscle actin deposition in cardiac sarcomeres at this location. These examples suggest translation occurs at pre-specified sites to favour targeting of proteins at relevant locations.[95]

MiRNA processing and maturation is an example of regulated changes in localisation. MiRNA genes are genomically encoded individually or as clusters, intergenically or intronically to protein-coding genes.[5] Primary miRNA transcripts are synthesized in the nucleus, processed by Drosha,

exported to the cytoplasm and cleaved by Dicer to produce mature miRNAs. Then, miRNAs are loaded in a silencing complex (RISC) together with Argonaute proteins. This allows pairing to and targeting mRNAs to silence expression. Most mature miRNAs demonstrate therefore cytoplasmic enrichment. Moreover, processing bodies (P-bodies) are cytosolic organelles, which are the primary sites of miRNAs activity. Interestingly, the RISC is required for maintaining P-body integrity.[96] RISC is also associated with polysomes, suggesting that miRNA-mediated repression is initiated in the rough endoplasmic reticulum. Altogether, miRNAs are concentrated in the cytoplasm at specific *foci*, each representing a functional step in the silencing pathway leading to mRNA degradation. On the other hand, the miRNA-loaded RISC can shuttle back into the nucleus, where it can exert additional regulatory function on transcription and splicing in the nucleoplasm or even in the chromatin. Finally, mature miRNAs are also found in mitochondria.[97] They originate from the cytoplasm, but can also derive from the mitochondrial genome. Their actions involve fine-tuning of various mitochondrial functions via regulating the expression of respiratory chain subunits, and thereby mitochondrial electron transport and oxidative phosphorylation.

There is a further emerging role for Argonaute proteins in the formation of RNA-induced transcriptional activation (RITA) complexes in association with small activating RNAs (saRNAs). RITAs translocate into the nucleus to enhance gene expression. Description and understanding of endogenous saRNAs in cardiovascular disease are yet to be unveiled, and there may be a potential to harness saRNA biology as a therapeutic strategy.[98]

Molecular mechanisms controlling the correct targeting of lncRNAs to their ultimate site of action are intrinsic parts of their functionality. Nuclear-enriched lncRNAs are mainly involved in transcriptional and epigenetic regulation through cis- or trans-mechanisms, acting at their site of transcription or via travelling to remote locations in the genome respectively. A large number of lncRNAs are also exported to the cytoplasm. Visualisation of nuclear lncRNAs using RNA fluorescent *in situ* hybridisation identifies lncRNAs either in discrete nuclear *foci* or more homogeneously in the nucleoplasm. *Foci* of chromatin-bound lncRNAs reflect active lncRNAs transcription, lncRNAs that remain tethered to their own locus or accumulation at target loci, for instance through formation of RNA-DNA triplexes, a mechanism of lncRNA-mediated regulation of gene expression. Moreover, lncRNAs are observed in other compartments such as nuclear speckles and paraspeckles, organelles implicated in mRNA processing. In addition, lncRNAs can also be nuclear and cytoplasmic, or purely cytoplasmic.[99] In this vein, lncRNAs that act as decoy for miRNAs in competing endogenous network are usually found in the cytoplasm, the site of miRNA-mediated post-transcriptional regulation. This is also the case for lncRNAs directly interacting with mRNAs for controlling their stability and degradation.

Surprisingly, despite the fact that lncRNAs are not translated, many lncRNAs are found associated with ribosomes in polysome profiling experiments. As mentioned previously, this might reflect a role for lncRNAs in translation regulation or micropeptide production.[16,17] Last but not least, lncRNAs are found in exosomes suggesting a role for lncRNAs in cell-to-cell communication.[100]

7. Peripheral blood

Mechanistic studies support peripheral blood RNAs as regulators of CVD. Since peripheral blood is easily accessible, it is a promising source

of biomarkers for personalized healthcare in areas with room for improved diagnosis and treatment (e.g. HF, myocarditis, valvular disease progression, and cardiogenic shock). The first studies on circulating miRNAs in cardiovascular disease focused on their diagnostic potential. Plasma levels of miR-1a and miR-133a/b were among the first miRNAs identified to be upregulated in acute MI. Many subsequent studies postulated dynamic changes in peripheral blood miRNAs in different areas of CVD, systematically reviewed elsewhere.[101] However, a lack of protocol standardisation, small sample sizes without validation cohorts, and heterogeneous patient populations have sometimes led to inconsistent results.

When considering studies with miRNAs in larger populations with independent validation, the largest body of evidence is available for miR-1, miR-133a/b, miR-208, and miR-499. These muscle- and cardiac-enriched miRNAs increased significantly in patients undergoing septal ablation for hypertrophy, suggesting a troponin-like biomarker potential.[102,103] However, after adjusting for cardiac troponin T or clinical variables in independent studies, these miRNAs lost significant diagnostic or prognostic capabilities in patients with chest pain or acute coronary syndrome.[104,105] Other miRNAs, for example endothelium-enriched miR-126, have consistently been shown to be related to platelet function making them suitable as candidates for the prediction of atherothrombotic events in primary or secondary prevention.[106–108]

Besides miRNAs, the biomarker potential of mRNAs, lncRNAs, and circRNAs was investigated after a few landmark reports in stable coronary artery disease (CAD), acute MI, and HF.[14,109] Plasma levels of *LIPCAR* and whole blood levels of *QSOX1* and the circRNA *MICRA* have been identified and replicated as novel prognostic markers of subsequent left ventricular dysfunction in patients with MI, independent from classical prognostic markers including NT-proBNP.[15,110,111]

Nevertheless, none of these RNAs has as yet been assessed in a large interventional clinical study to tailor therapies or improve patient outcome. Several considerations may explain the challenge to translate experimental findings from transcriptomic studies back to the bedside (*Figure 5*). These challenges include the timing of blood sampling (e.g. lifestyle-related time-of-day fluctuations as well as stage of disease onset and progression), selecting the appropriate sample type, sample handling prior to storage, RNA isolation strategy, RNA quality and correct downstream RNA expression measurement with appropriate normalisation. RNA degradation is a serious limitation of transcriptomic studies in peripheral blood samples. In a clinical setting, liquid biospecimen cannot always be extracted and stored optimally to preserve the blood transcriptome. To overcome this challenge and eliminate other preanalytical variables, blood collection tubes containing RNA stabilizers have been developed. Even with these precautions, RNA integrity and potential genomic DNA contamination should be assessed before starting downstream applications. For cell-free RNA (e.g. in plasma or exosomes), no dedicated approach to minimize RNA degradation has been described. Heparin administration to the patient or in the sample may also result in spurious RNA expression results.[26] Finally, the quantity of extracted RNA may also constitute a hurdle for transcriptomic analyses. The abundance of RNA transcripts can be very low in plasma, particularly lncRNAs and circRNAs, which tend to be relatively weakly expressed in blood compartments such as exosomes. However, the continuous improvement of sequencing and PCR technologies is facilitating reliable transcriptomic studies from very low amounts of starting material. The ligation-free CATS technology has been developed to prepare RNA-seq libraries from low-input RNA down to 100 picograms.[112] Nonetheless, with these technologies, increased sensitivity

Figure 5 Challenges and opportunities in transcriptomic studies in peripheral blood.
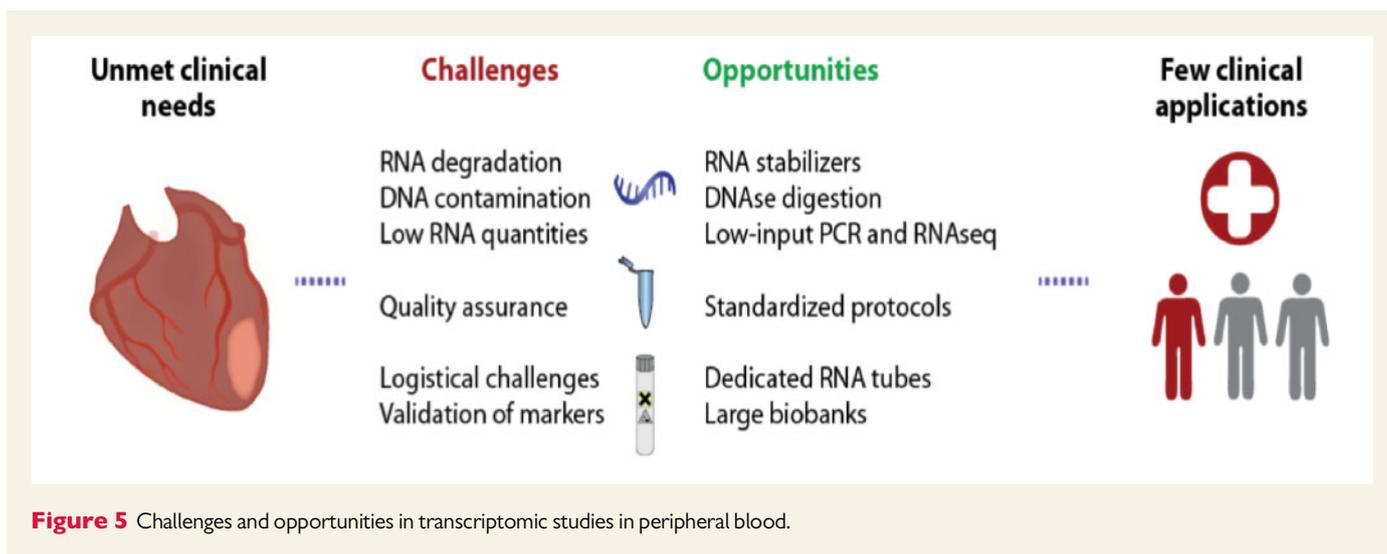
might also be synonymous of increased sensitivity to detection of non-RNA species (e.g. circulating genomic DNA).[113] A high-quality RNA is therefore necessary for transcriptomic studies from low-input RNA isolated from a small volume of plasma/serum or exosomes.

Only two RNA-based biomarkers in vitro diagnostic assays for cardiac conditions are commercially available. Corus® CAD (CardioDx, Palo Alto, CA, USA) uses whole blood RNA for the diagnosis of stable coronary artery disease in non-diabetic patients, and AlloMap® (CareDx, Brisbane, CA, USA) uses peripheral blood mononuclear cells for the detection of acute rejection in patients who underwent cardiac transplantation.[114,115] Both tests have a high negative predictive value, and the latter is the only RNA test in cardiovascular disease that demonstrated effectiveness in a clinical trial.[116]

8. Personalized medicine/tailored approaches

As previously discussed, the revolution in deeper and faster clinical genetic and transcriptomic phenotyping is enabling us to understand disease-associated changes in each patient. This allows the opportunity for personalizing therapeutic strategies.

Transcriptome changes in the blood (cells or plasma) may help to better diagnose or determine the prognosis of our patients. Indeed, many clinical guidelines emphasize the unmet need for personalized medicine approaches. For instance, biomarker-guided cardiovascular therapy represents an interesting approach to identify the right patients, make accurate diagnostic decisions, and monitor efficacy. NcRNA signatures provide valuable molecular insight of patient phenotypes and could add to traditional markers and established clinical variables.

Short circulating miRNAs show promise for making predictions of cardiovascular patients; plasma miRNA-1254 levels predicted changes in left ventricular (LV) volumes and LV ejection fraction at 6 months after ST elevation myocardial infarction (STEMI). The value of miR-1254 to inform tailored treatment selection and monitor ongoing efficacy deserves further investigation.[117] LncRNA biomarkers are also being investigated as novel predictive tools to monitor therapeutic effectiveness and to stratify patients. The lncRNA SENCR blood levels before the start of

pioglitazone therapy allowed distinction of responders from non-responders and were found to be a better predictor of the response to therapy than other traditional clinical variables.[118] Thus, assessment of lncRNA levels could constitute a major breakthrough in the development of personalized medicine to guide medical therapy, with important implications in terms of adverse effects and treatment costs. Another example demonstrating the potential of circulating miRNAs to predict the response to LVAD has been previously evaluated in patients with severe advanced HF. Morley-Smith *et al.* classified the clinical response to LVAD therapy based on NT-proBNP concentration at 3 months.[104] Selected patients with a NT-proBNP above the 50th percentile were identified as 'poor responders' whereas patients below the 50th percentile as 'good responders'. By screening >1000 miRNAs, miR-1202 was found to be a useful biomarker to predict the response to LVAD therapy. Baseline circulating miR-1202 levels stratified the cohort into poor and good responders with greater sensitivity and specificity than baseline NT-proBNP. Thus, miR-1202 emerges as an interesting biomarker to select LVAD therapy or alternative interventions, such as urgent cardiac transplantation.[119]

These examples illustrate that personalized medicine approaches show promise in becoming a reality in the near future. The rise of new technologies such as global and single-cell sequencing, the use of artificial intelligence, and machine-learning concepts will lead to tremendous new discoveries and clinically useful applications and finally also to better and safer therapies (*Figure 6*).

9. Brief conclusion

Recent breakthroughs in deep-sequencing technology have enabled us to delineate transcriptomes, genomes, and epigenomes from individual cells and cell types in complex heterogeneous organs and tissues of the CV system and unveil novel types, low or rarely expressed ncRNA transcripts.

With the rapid speed and reduced costs of DNA and RNA-sequencing, we are now on the brink of being in a position to fully characterize each individual patient at the molecular level. To fully realize personalized medicine, we need to bridge interdisciplinary gaps to address the

Figure 6 Potential clinical application of circulating non-coding RNAs in personalized cardiovascular medicine.
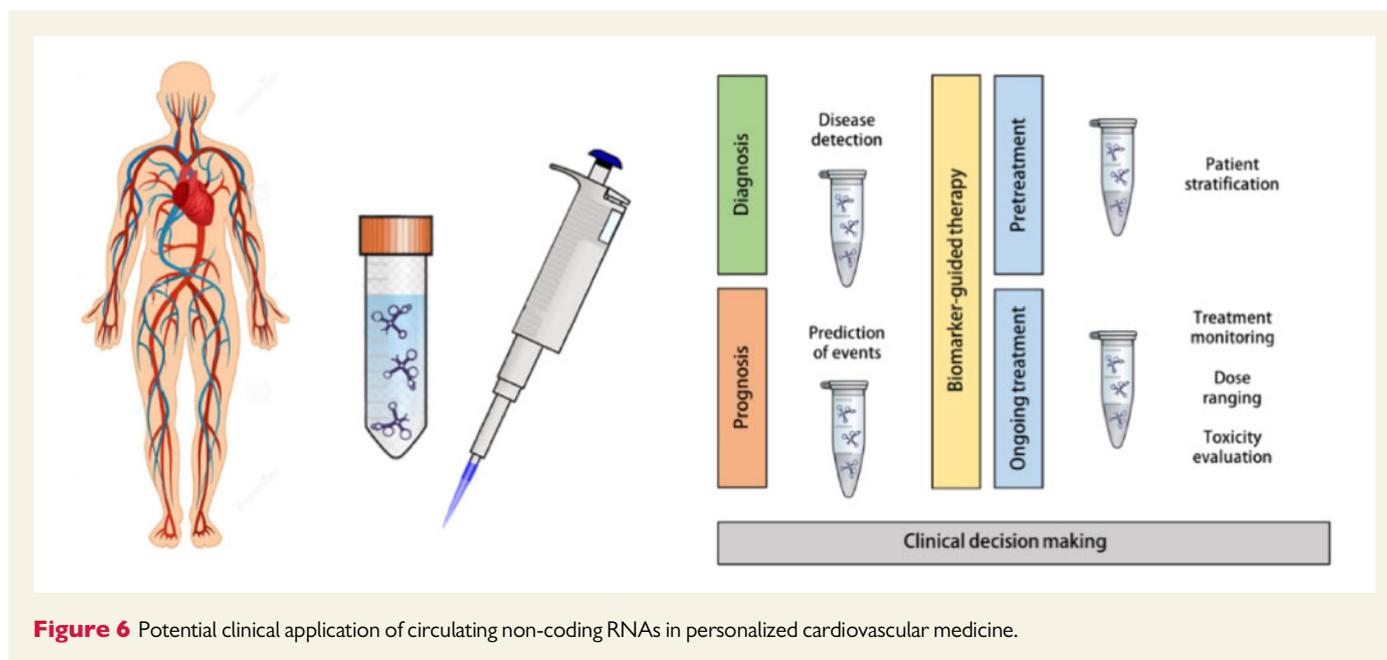
challenges posed by a deeper insight into the transcriptome regulation in the cardiovascular system, and to create unity and standardisation of methodologies: from sample preparation to bioinformatics analysis to interventional studies, for ultimately translating the findings into high-impact outcomes for public health.

Conflict of interest: none declared.

References

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann Y, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C, Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Raymond C, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, de la Bastide M, Dedhia N, Blöcker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglou S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kaspryzk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowki J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ,

Osoegawa K, Shizuya H, Choi S, Chen YJ, Szustakowki J, International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* 2001;**409**:860–921.

2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. *Science* 2001; **291**:1304–1351.

3. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature* 2004;**431**:931–945.

4. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, Oyama R, Ravasi T, Lenhard B, Wells C, Kodzius R, Shimokawa K, Bajic VB, Brenner SE, Batalov S, Forrest ARR, Zavolan M, Davis MJ, Wilming LG, Aidinis V, Allen JE, Ambesi-Impiombato A, Apweiler R, Aturaliya RN, Bailey TL, Bansal M, Baxter L, Beisel KW, Bersano T, Bono H, Chalk AM, Chiu KP, Choudhary V, Christoffels A, Clutterbuck DR, Crowe ML, Dalla E, Dalrymple BP, de Bono B, Della Gatta G, di Bernardo D, Down T, Engstrom P, Fagiolini M, Faulkner G, Fletcher CF, Fukushima T, Furuno M, Futaki S, Gariboldi M, Georgii-Hemming P, Gingeras TR, Gojobori T, Green RE, Gustincich S, Harbers M, Hayashi Y, Hensch TK, Hirokawa N, Hill D, Huminiecki L, Iacono M, Ikeo K, Iwama A, Ishikawa T, Jakt M, Kanapin A, Katoh M, Kawasawa Y, Kelso J, Kitamura H, Kitano H, Kollias G, Krishnan SPT, Kruger A, Kummerfeld SK, Kurochkin IV, Lareau LF, Lazarevic D, Lipovich L, Liu J, Liuni S, McWilliam S, Madan Babu M, Madera M, Marchionni L, Matsuda H, Matsuzawa S, Miki H, Mignone F, Miyake S, Morris K, Mottagui-Tabar S, Mulder N, Nakano N, Nakauchi H, Ng P, Nilsson D, Nishiguchi S, Nishikawa S, Nori F, Ohara O, Okazaki Y, Orlando V, Pang KC, Pavan WJ, Pavesi G, Pesole G, Petrovsky N, Piazza S, Reed J, Reid JF, Ring BZ, Ringwald M, Rost B, Ruan Y, Salzberg SL, Sandelin A, Schneider C, Schönbach C, Sekiguchi K, Semple C. A M, Seno S, Sessa L, Sheng Y, Shibata Y, Shimada H, Shimada K, Silva D, Sinclair B, Sperling S, Stupka E, Sugiura K, Sultana R, Takenaka Y, Taki K, Tammoja K, Tan SL, Tang S, Taylor MS, Tegner J, Teichmann SA, Ueda HR, van Nimwegen E, Verardo R, Wei CL, Yagi K, Yamanishi H, Zabarovsky E, Zhu S, Zimmer A, Hide W, Bult C, Grimmond SM, Teasdale RD, Liu ET, Brusic V, Quackenbush J, Wahlestedt C, Mattick JS, Hume DA, Kai C, Sasaki D, Tomaru Y, Fukuda S, Kanamori-Katayama M, Suzuki M, Aoki J, Arakawa T, Iida J, Imamura K, Itoh M, Kato T, Kawaji H, Kawagashira N, Kawashima T, Kojima M, Kondo S, Konno H, Nakano K, Ninomiya N, Nishio T, Okada M, Plessy C, Shibata K, Shiraki T, Suzuki S, Tagami M, Waki K, Watahiki A, Okamura-Oho Y, Suzuki H, Kawai J, Hayashizaki Y, FANTOM Consortium, RIKEN Genome Exploration Research Group and Genome Science Group (Genome Network Project Core Group). The transcriptional landscape of the mammalian genome. *Science* 2005;**309**:1559–1563.

5. Treiber T, Treiber N, Meister G. Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. *Nat Rev Mol Cell Biol* 2019;**20**:5–20.

6. Colpaert RMW, Calore M. MicroRNAs in cardiac diseases. *Cells* 2019;**8**:737.

7. Weber JA, Baxter DH, Zhang S, Huang DY, How Huang K, Jen Lee M, Galas DJ, Wang K. The MicroRNA spectrum in 12 body fluids. *Clin Chem* 2010;**56**:1733–1741.

8. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N,

Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* 2009;**458**:223–227.

9. Ounzain S, Micheletti R, Beckmann T, Schroen B, Alexanian M, Pezzuto I, Crippa S, Nemir M, Sarre A, Johnson R, Dauvillier J, Burdet F, Ibberson M, Guigó R, Xenarios I, Heymans S, Pedrazzini T. Genome-wide profiling of the cardiac transcriptome after myocardial infarction identifies novel heart-specific long non-coding RNAs. *Eur Heart J* 2015;**36**:353–368.

10. Alexanian M, Maric D, Jenkinson SP, Mina M, Friedman CE, Ting C-C, Micheletti R, Plaisance I, Nemir M, Maison D, Kernen J, Pezzuto I, Villeneuve D, Burdet F, Ibberson M, Leib SL, Palpant NJ, Hernandez N, Ounzain S, Pedrazzini T. A transcribed enhancer dictates mesendoderm specification in pluripotency. *Nat Commun* 2017;**8**:1806.

11. Salamon I, Serio S, Bianco S, Pagiatakis C, Crasto S, Chiariello AM, Conte M, Cattaneo P, Fiorillo L, Felicetta A, di Pasquale E, Kunderfranco P, Nicodemi M, Papait R, Condorelli G. Condorelli divergent transcription of the Nkx2-5 locus generates two enhancer RNAs with opposing functions. *iScience* 2020;**23**:101539.

12. Carrara M, Fuschi P, Ivan C, Martelli F. Circular RNAs: methodological challenges and perspectives in cardiovascular diseases. *J Cell Mol Med* 2018;**22**:5176–5187.

13. Wang Y, Lu T, Wang Q, Liu J, Jiao W. Circular RNAs: crucial regulators in the human body. *Oncol Rep* 2018;**40**:3119–3135.

14. Vausort M, Wagner DR, Devaux Y. Long noncoding RNAs in patients with acute myocardial infarction. *Circ Res* 2014;**115**:668–677.

15. Vausort M, Salgado-Somoza A, Zhang L, Leszek P, Scholz M, Teren A, Burkhardt R, Thiery J, Wagner DR, Devaux Y. Myocardial infarction-associated circular RNA predicting left ventricular dysfunction. *J Am Coll Cardiol* 2016;**68**:1247–1248.

16. Yeasmin F, Yada T, Akimitsu N. Micropeptides encoded in transcripts previously identified as long noncoding RNAs: a new chapter in transcriptomics and proteomics. *Front Genet* 2018;**9**:144.

17. van Heesch S, Witte F, Schneider-Lunitz V, Schulz JF, Adami E, Faber AB, Kirchner M, Maatz H, Blachut S, Sandmann C-L, Kanda M, Worth CL, Schafer S, Calviello L, Merriott R, Patone G, Hummel O, Wyler E, Obermayer B, Mücke MB, Lindberg EL, Trnka F, Memczak S, Schilling M, Felkin LE, Barton PJR, Quaife NM, Vanezis K, Diecke S, Mukai M, Mah N, Oh S-J, Kurtz A, Schramm C, Schwinge D, Sebode M, Harakalova M, Asselbergs FW, Vink A, de Weger RA, Viswanathan S, Widjaja AA, Gärtner-Rommel A, Milting H, Dos Remedios C, Knosalla C, Mertins P, Landthaler M, Vingron M, Linke WA, Seidman JG, Seidman CE, Rajewsky N, Ohler U, Cook SA, Hubner N. The translational landscape of the human heart. *Cell* 2019;**178**: 242–260.

18. Aranda PS, LaJoie DM, Jorcyk CL. Bleach gel: a simple agarose gel for analyzing RNA quality. *Electrophoresis* 2012;**33**:366–369.

19. Masek T, Vopalensky V, Suchomelova P, Pospisek M. Denaturing RNA electrophoresis in TAE agarose gels. *Anal Biochem* 2005;**336**:46–50.

20. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 2006;**7**:3.

21. Swift GH, Peyton MJ, MacDonald RJ. Assessment of RNA quality by semi-quantitative RT-PCR of multiple regions of a long ubiquitous mRNA. *Biotechniques* 2000;**28**: 524–531.

22. Sidova M, Tomankova S, Abaffy P, Kubista M, Sindelka R. Effects of post-mortem and physical degradation on RNA integrity and quality. *Biomol Detect Quantif* 2015;**5**: 3–9.

23. Zhu Y, Wang L, Yin Y, Yang E. Systematic analysis of gene expression patterns associated with postmortem interval in human tissues. *Sci Rep* 2017;**7**:5435.

24. White K, Yang P, Li L, Farshori A, Medina AE, Zielke HR. Effect of postmortem interval and years in storage on RNA quality of tissue at a repository of the NIH NeuroBioBank. *Biopreserv Biobank* 2018;**16**:148–157.

25. Elghamry HA, Hassan FM, Mohamed MI, Abdelfattah DS, Abdelaal AG. Estimation of the postmortem interval using GAPDH mRNA in skin and heart tissues of albino rats at different environmental conditions. *Egypt J Forensic Sci* 2018;**8**:69.

26. Boeckel J-N, Thomé CE, Leistner D, Zeiher AM, Fichtlscherer S, Dimmeler S. Heparin selectively affects the quantification of microRNAs in human blood samples. *Clin Chem* 2013;**59**:1125–1127.

27. Ling L, Camilleri ET, Helledie T, Samsonraj RM, Titmarsh DM, Chua RJ, Dreesen O, Dombrowski C, Rider DA, Galindo M, Lee I, Hong W, Hui JH, Nurcombe V, van Wijnen AJ, Cool SM. Effect of heparin on the biological properties and molecular signature of human mesenchymal stem cells. *Gene* 2016;**576**:292–303.

28. Lehmann S, Guadagni F, Moore H, Ashton G, Barnes M, Benson E, Clements J, Koppandi I, Coppola D, Demiroglu SY, DeSouza Y, De Wilde A, Duker J, Eliason J, Glazer B, Harding K, Jeon JP, Kessler J, Kokkat T, Nanni U, Shea K, Skubitz A, Somiari S, Tybring G, Gunter E, Betsou F, International Society for Biological and Environmental Repositories (ISBER) Working Group on Biospecimen Science. Standard preanalytical coding for biospecimens: review and implementation of the Sample PREanalytical Code (SPREC). *Biopreserv Biobank* 2012;**10**:366–374.

29. Häntzsch M, Tolios A, Beutner F, Nagel D, Thiery J, Teupser D, Holdt LM. Comparison of whole blood RNA preservation tubes and novel generation RNA extraction kits for analysis of mRNA and MiRNA profiles. *PLoS One* 2014;**9**: e113298.

30. Gautam A, Donohue D, Hoke A, Miller SA, Srinivasan S, Sowe B, Detwiler L, Lynch J, Levangie M, Hammamieh R, Jett M. Investigating gene expression profiles of whole blood and peripheral blood mononuclear cells using multiple collection and processing methods. *PLoS One* 2019;**14**:e0225137.

31. Tu C, Du T, Shao C, Liu Z, Li L, Shen Y. Evaluating the potential of housekeeping genes, rRNAs, snRNAs, microRNAs and circRNAs as reference genes for the estimation of PMI. *Forensic Sci Med Pathol* 2018;**14**:194–201.

32. Reiman M, Laan M, Rull K, Sõber S. Effects of RNA integrity on transcript quantification by total RNA sequencing of clinically collected human placental samples. *FASEB J* 2017;**31**:3298–3308.

33. Gallego Romero I, Pai AA, Tung J, Gilad Y. RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biol* 2014;**12**:42.

34. Union, P. O. of the E. Biobanks for Europe: a challenge for governance. 2012 http://op.europa.eu/en/publication-detail/-/publication/629eae10-53fc-4a52-adc2-210d4fcad8f2.

35. Merino-Martinez R, Norlin L, van Enckevort D, Anton G, Schuffenhauer S, Silander K, Mook L, Holub P, Bild R, Swertz M, Litton J-E. Toward Global Biobank Integration by Implementation of the Minimum Information About Biobank Data Sharing (MIABIS 2.0 Core). *Biopreserv Biobank* 2016;**14**:298–306.

36. Nag AC. Study of non-muscle cells of the adult mammalian heart: a fine structural analysis and distribution. *Cytobios* 1980;**28**:41–61.

37. Pinto AR, Ilinykh A, Ivey MJ, Kuwabara JT, D'Antoni ML, Debuque R, Chandran A, Wang L, Arora K, Rosenthal N, Tallquist MD. Revisiting Cardiac Cellular Composition. *Circ Res* 2016;**118**:400–409.

38. Dobnikar L, Taylor AL, Chappell J, Oldach P, Harman JL, Oerton E, Dzierzak E, Bennett MR, Spivakov M, Jørgensen HF. Disease-relevant transcriptional signatures identified in individual smooth muscle cells from healthy mouse vessels. *Nat Commun* 2018;**9**:4567.

39. O'Flanagan CH, Campbell KR, Zhang AW, Kabeer F, Lim JLP, Biele J, Eirew P, Lai D, McPherson A, Kong E, Bates C, Borkowski K, Wiens M, Hewitson B, Hopkins J, Pham J, Ceglia N, Moore R, Mungall AJ, McAlpine JN, Shah SP, Aparicio S, CRUK IMAXT Grand Challenge Team. Dissociation of solid tumor tissues with cold active protease for single-cell RNA-seq minimizes conserved collagenase-associated stress responses. *Genome Biol* 2019;**20**:210.

40. Ackers-Johnson M, Foo RS. Langendorff-free isolation and propagation of adult mouse cardiomyocytes. *Methods Mol Biol* 2019;**1940**:193–204.

41. Ackers-Johnson M, Li PY, Holmes AP, O'Brien S-M, Pavlovic D, Foo RS. A simplified, Langendorff-free method for concomitant isolation of viable cardiac myocytes and nonmyocytes from the adult mouse heart: novelty and significance. *Circ Res* 2016;**119**:909–920.

42. Bergmann O, Jovinge S. Isolation of cardiomyocyte nuclei from post-mortem tissue. *J Vis Exp* 2012; **65**: 4205.

43. Gilsbach R, Schwaderer M, Preissl S, Grüning BA, Kranzhöfer D, Schneider P, Nührenberg TG, Mulero-Navarro S, Weichenhan D, Braun C, Dreßen M, Jacobs AR, Lahm H, Doenst T, Backofen R, Krane M, Gelb BD, Hein L. Distinct epigenetic programs regulate cardiac myocyte development and disease in the human heart in vivo. *Nat Commun* 2018;**9**:391.

44. Bergmann O, Bhardwaj RD, Bernard S, Zdunek S, Barnabé-Heider F, Walsh S, Zupicich J, Alkass K, Buchholz BA, Druid H, Jovinge S, Frisén J. Evidence for cardiomyocyte renewal in humans. *Science* 2009;**324**:98–102.

45. Wu AZ, Xu D, Yang N, Lin S-F, Chen P-S, Cala SE, Chen Z. Phospholamban is concentrated in the nuclear envelope of cardiomyocytes and involved in perinuclear/nuclear calcium handling. *J Mol Cell Cardiol* 2016;**100**:1–8.

46. Barthelson RA, Lambert GM, Vanier C, Lynch RM, Galbraith DW. Comparison of the contributions of the nuclear and cytoplasmic compartments to global gene expression in human cells. *BMC Genomics* 2007;**8**:340.

47. Solnestam BW, Stranneheim H, Hällman J, Käller M, Lundberg E, Lundeberg J, Akan P. Comparison of total and cytoplasmic mRNA reveals global regulation by nuclear retention and miRNAs. *BMC Genomics* 2012;**13**:574.

48. Selewa A, Dohn R, Eckart H, Lozano S, Xie B, Gauchat E, Elorbany R, Rhodes K, Burnett J, Gilad Y, Pott S, Basu A. Systematic comparison of high-throughput single-cell and single-nucleus transcriptomes during cardiomyocyte differentiation. *Sci Rep* 2020;**10**:1535.

49. Monaco G, Lee B, Xu W, Mustafah S, Hwang YY, Carré C, Burdin N, Visan L, Ceccarelli M, Poidinger M, Zippelius A, Pedro de Magalhães J, Larbi A. RNA-seq signatures normalized by mRNA abundance allow absolute deconvolution of human immune cell types. *Cell Rep* 2019;**26**:1627–1640.

50. Aguirre-Gamboa R, de Klein N, di Tommaso J, Claringbould A, van der Wijst MG, de Vries D, Brugge H, Oelen R, Võsa U, Zorro MM, Chu X, Bakker OB, Borek Z, Ricaño-Ponce I, Deelen P, Xu C-J, Swertz M, Jonkers I, Withoff S, Joosten I, Sanna S, Kumar V, Koenen HJPM, Joosten LAB, Netea MG, Wijmenga C, Franke L, Li Y, BIOS Consortium. Deconvolution of bulk blood eQTL effects into immune cell subpopulations. *BMC Bioinformatics* 2020;**21**:243.

51. Heras-Bautista CO, Mikhael N, Lam J, Shinde V, Katsen-Globa A, Dieluweit S, Molcanyi M, Uvarov V, Jütten P, Sahito RGA, Mederos-Henry F, Piechot A, Brockmeier K, Hescheler J, Sachinidis A, Pfannkuche K. Cardiomyocytes facing fibrotic conditions re-express extracellular matrix transcripts. *Acta Biomater* 2019;**89**:180–192.

52. Wang J, Chen H, Seth A, McCulloch CA. Mechanical force regulation of myofibroblast differentiation in cardiac fibroblasts. *Am J Physiol Heart Circ Physiol* 2003;**285**:H1871–H1881.

53. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A, Lao K, Surani MA. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;**6**:377–382.

54. Gladka MM, Molenaar B, de Ruiter H, van der Elst S, Tsui H, Versteeg D, Lacraz GPA, Huibers MMH, van Oudenaarden A, van Rooij E. Single-cell sequencing of the healthy and diseased heart reveals cytoskeleton-associated protein 4 as a new modulator of fibroblasts activation. *Circ* 2018;**138**:166–180.

55. Forte E, Skelly DA, Chen M, Daigle S, Morelli KA, Hon O, Philip VM, Costa MW, Rosenthal NA, Furtado MB. Dynamic interstitial cell response during myocardial infarction predicts resilience to rupture in genetically diverse mice. *Cell Rep* 2020;**30**:3149–3163.

56. Li Z, Solomonidis EG, Meloni M, Taylor RS, Duffin R, Dobie R, Magalhaes MS, Henderson BEP, Louwe PA, D'Amico G, Hodivala-Dilke KM, Shah AM, Mills NL, Simons BD, Gray GA, Henderson NC, Baker AH, Brittan M. Single-cell transcriptome analyses reveal novel targets modulating cardiac neovascularization by resident endothelial cells following myocardial infarction. *Eur Heart J* 2019;**40**:2507–2520.

57. Humphries KH, Izadnegahdar M, Sedlak T, Saw J, Johnston N, Schenck-Gustafsson K, Shah RU, Regitz-Zagrosek V, Grewal J, Vaccarino V, Wei J, Bairey Merz CN. Sex differences in cardiovascular disease - Impact on care and outcomes. *Front Neuroendocrinol* 2017;**46**:46–70.

58. Tucker NR, Chaffin M, Fleming SJ, Hall AW, Parsons VA, Bedi KC, Akkad A-D, Herndon CN, Arduini A, Papangeli I, Roselli C, Aguet F, Choi SH, Ardlie KG, Babadi M, Margulies KB, Stegmann CM, Ellinor PT. Transcriptional and cellular diversity of the human heart. *Circ* 2020;**142**:466–482.

59. Wang L, Yu P, Zhou B, Song J, Li Z, Zhang M, Guo G, Wang Y, Chen X, Han L, Hu S. 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**:108–119.

60. Elisa M, Kunderfranco P, Peano C, Carullo P, Cremonesi M, Schorn T, Carriero R, Termanini A, Simone Colombo F, Jachetti E, Panico C, Faggian G, Fumero A, Torracca L, Molgora M, Cibella J, Pagiatakis C, Brummelman J, Alvisi G, Mazza EMC, Colombo MP, Lugli E, Condorelli G, Kallikourdis M. Single-cell sequencing of mouse heart immune infiltrate in pressure overload–driven heart failure reveals extent of immune activation. *Circ* 2019; **140**:2089–2107.

61. Cochain C, Vafadarnejad E, Arampatzi P, Pelisek J, Winkels H, Ley K, Wolf D, Saliba A-E, Zernecke A. Single-cell RNA-Seq reveals the transcriptional landscape and heterogeneity of aortic macrophages in murine atherosclerosis. *Circ Res* 2018;**122**:1661–1674.

62. Lukowski SW, Patel J, Andersen SB, Sim S-L, Wong HY, Tay J, Winkler I, Powell JE, Khosrotehrani K. Single-cell transcriptional profiling of aortic endothelium identifies a hierarchy from endovascular progenitors to differentiated cells. *Cell Rep* 2019;**27**:2748–2758.

63. Paik DT, Tian L, Lee J, Sayed N, Chen IY, Rhee S, Rhee J-W, Kim Y, Wirka RC, Buikema JW, Wu SM, Red-Horse K, Quertermous T, Wu JC. Large-scale single-cell RNA-seq reveals molecular signatures of heterogeneous populations of human induced pluripotent stem cell-derived endothelial cells. *Circ Res* 2018;**123**:443–450.

64. McCracken IR, Taylor RS, Kok FO, de la Cuesta F, Dobie R, Henderson BEP, Mountford JC, Caudrillier A, Henderson NC, Ponting CP, Baker AH. Transcriptional dynamics of pluripotent stem cell-derived endothelial cell differentiation revealed by single-cell RNA sequencing. *Eur Heart J* 2020;**41**:1024–1036.

65. See K, Tan WLW, Lim EH, Tiang Z, Lee LT, Li PYQ, Luu TDA, Ackers-Johnson M, Foo RS. Single cardiomyocyte nuclear transcriptomes reveal a lincRNA-regulated de-differentiation and cell cycle stress-response in vivo. *Nat Commun* 2017;**8**:225.

66. Farbehi N, Patrick R, Dorison A, Xaymardan M, Janbandhu V, Wystub-Lis K, Ho JW, Nordon RE, Harvey RP. Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury. *Elife* 2019;**8**:e43882.

67. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol* 2010;**11**:R106.

68. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010;**26**:139–140.

69. Soneson C, Robinson MD. Bias, robustness and scalability in single-cell differential expression analysis. *Nat Methods* 2018;**15**:255–261.

70. Wang T, Li B, Nelson CE, Nabavi S. Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data. *BMC Bioinformatics* 2019;**20**:40.

71. Finak G, McDavid A, Yajima M, Deng J, Gersuk V, Shalek AK, Slichter CK, Miller HW, McElrath MJ, Prlic M, Linsley PS, Gottardo R. MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol* 2015;**16**:278.

72. Wu C-C, Kruse F, Vasudevarao MD, Junker JP, Zebrowski DC, Fischer K, Noël ES, Grün D, Berezikov E, Engel FB, van Oudenaarden A, Weidinger G, Bakkers J. Spatially resolved genome-wide transcriptional profiling identifies BMP signaling as essential regulator of zebrafish cardiomyocyte regeneration. *Dev Cell* 2016;**36**:36–49.

73. Lacraz GPA, Junker JP, Gladka MM, Molenaar B, Scholman KT, Vigil-Garcia M, Versteeg D, de Ruiter H, Vermunt MW, Creyghton MP, Huibers MMH, de Jonge N, van Oudenaarden A, van Rooij E. Tomo-Seq identifies SOX9 as a key regulator of cardiac fibrosis during ischemic injury. *Circ* 2017;**136**:1396–1409.

74. Mohenska M, Tan NM, Tokolyi A, Furtado MB, Costa MW, Perry AJ, Hatwell-Humble J, K. van, Nim, H.T D, Nilsson SK, Powell DR, Rosenthal NA, Rossello FJ, Ramialison M, Polo JM. 3D-Cardiomics: a spatial transcriptional atlas of the mammalian heart. *bioRxiv* 2019;792002.

75. Beqqali A. Alternative splicing in cardiomyopathy. *Biophys Rev* 2018;**10**:1061–1071.

76. Tan WLW, Lim BTS, Anene-Nzelu CGO, Ackers-Johnson M, Dashi A, See K, Tiang Z, Lee DP, Chua WW, Luu TDA, Li PYQ, Richards AM, Foo RSY. A landscape of circular RNA expression in the human heart. *Cardiovasc Res* 2017;**113**:298–309.

77. Lim TB, Aliwarga E, Luu TDA, Li YP, Ng SL, Annadoray L, Sian S, Ackers-Johnson MA, Foo RS-Y. Targeting the highly abundant circular RNA circSlc8a1 in cardiomyocytes attenuates pressure overload induced hypertrophy. *Cardiovasc Res* 2019;**115**:1998–2007.

78. Hall IF, Climent M, Quintavalle M, Farina FM, Schorn T, Zani S, Carullo P, Kunderfranco P, Civilini E, Condorelli G, Elia L. Circ_Lrp6, a circular RNA enriched in vascular smooth muscle cells, acts as a sponge regulating miRNA-145 function. *Circ Res* 2019;**124**:498–510.

79. Kong SW, Hu YW, Ho JWK, Ikeda S, Polster S, John R, Hall JL, Bisping E, Pieske B, dos Remedios CG, Pu WT. Heart failure-associated changes in RNA splicing of sarcomere genes. *Circ Cardiovasc Genet* 2010;**3**:138–146.

80. Heinig M, Adriaens ME, Schafer S, van Deutekom HWM, Lodder EM, Ware JS, Schneider V, Felkin LE, Creemers EE, Meder B, Katus HA, Rühle F, Stoll M, Cambien F, Villard E, Charron P, Varro A, Bishopric NH, George AL, dos Remedios C, Moreno-Moral A, Pesce F, Bauerfeind A, Rüschendorf F, Rintisch C, Petretto E, Barton PJ, Cook SA, Pinto YM, Bezzina CR, Hubner N. Natural genetic variation of the cardiac transcriptome in non-diseased donors and patients with dilated cardiomyopathy. *Genome Biol* 2017;**18**:170.

81. Ames E, Lawson M, Mackey A, Holmes J. Sequencing of mRNA identifies re-expression of fetal splice variants in cardiac hypertrophy. *J Mol Cell Cardiol* 2013;**62**:99–107.

82. Beqqali A, Bollen IAE, Rasmussen TB, van den Hoogenhof MM, van Deutekom HWM, Schafer S, Haas J, Meder B, Sørensen KE, van Oort RJ, Mogensen J, Hubner N, Creemers EE, van der Velden J, Pinto YM. A mutation in the glutamate-rich region of RNA-binding motif protein 20 causes dilated cardiomyopathy through missplicing of titin and impaired Frank-Starling mechanism. *Cardiovasc Res* 2016;**112**:452–463.

83. Guo W, Schafer S, Greaser ML, Radke MH, Liss M, Govindarajan T, Maatz H, Schulz H, Li S, Parrish AM, Dauksaite V, Vakeel P, Klaassen S, Gerull B, Thierfelder L, Regitz-Zagrosek V, Hacker TA, Saupe KW, Dec GW, Ellinor PT, MacRae CA, Spallek B, Fischer R, Perrot A, Özcelik C, Saar K, Hubner N, Gotthardt M. RBM20, a gene for hereditary cardiomyopathy, regulates titin splicing. *Nat Med* 2012;**18**:766–773.

84. Liu J, Kong X, Zhang M, Yang X, Xu X. RNA binding protein 24 deletion disrupts global alternative splicing and causes dilated cardiomyopathy. *Protein Cell* 2019;**10**:405–416.

85. Gaertner A, Brodehl A, Milting H. Screening for mutations in human cardiomyopathy- is RBM24 a new but rare disease gene? *Protein Cell* 2019;**10**:393–394.

86. Khan MAF, Reckman YJ, Aufiero S, van den Hoogenhof MMG, van der Made I, Beqqali A, Koolbergen DR, Rasmussen TB, van der Velden J, Creemers EE, Pinto YM. RBM20 regulates circular RNA production from the titin gene. *Circ Res* 2016;**119**:996–1003.

87. Gupta SK, Garg A, Bär C, Chatterjee S, Foinquinos A, Milting H, Streckfuß-Bömeke K, Fiedler J, Thum T. Quaking inhibits doxorubicin-mediated cardiotoxicity through regulation of cardiac circular RNA expression. *Circ Res* 2018;**122**:246–254.

88. Kalsotra A, Xiao X, Ward AJ, Castle JC, Johnson JM, Burge CB, Cooper TA. A postnatal switch of CELF and MBNL proteins reprograms alternative splicing in the developing heart. *Proc Natl Acad Sci USA* 2008;**105**:20333–20338.

89. Falcone G, Perfetti A, Cardinali B, Martelli F. Noncoding RNAs: emerging players in muscular dystrophies. *Biomed Res Int* 2014;**2014**:503634.

90. Arbustini E, Di Toro A, Giuliani L, Favalli V, Narula N, Grasso M. Cardiac phenotypes in hereditary muscle disorders: JACC state-of-the-art review. *J Am Coll Cardiol* 2018;**72**:2485–2506.

91. Dvinge H. Regulation of alternative mRNA splicing: old players and new perspectives. *FEBS Lett* 2018;**592**:2987–3006.

92. Lukowski SW, Patel J, Andersen SB, Sim S-L, Wong HY, Tay J, Winkler I, Powell JE, Khosrotehrani K. Long-read human genome sequencing and its applications. *Nat Rev Genet* 2020;**21**:597–614.

93. Uapinyoying P, Goecks J, Knoblach SM, Panchapakesan K, Bonnemann CG, Partridge TA, Jaiswal JK, Hoffman EP. A long-read RNA-seq approach to identify novel transcripts of very large genes. *Genome Res* 2020;**30**:885–897.

94. Bahar Halpern K, Caspi I, Lemze D, Levy M, Landen S, Elinav E, Ulitsky I, Itzkovitz S. Nuclear retention of mRNA in mammalian tissues. *Cell Rep* 2015;**13**:2653–2662.

95. David JJ, Subramanian SV, Zhang A, Willis WL, Kelm RJ, Leier CV, Strauch AR. Y-box binding protein-1 implicated in translational control of fetal myocardial gene expression after cardiac transplant. *Exp Biol Med (Maywood)* 2012;**237**:593–607.

96. Eulalio A, Behm-Ansmant I, Schweizer D, Izaurralde E. P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol Cell Biol* 2007;**27**:3970–3981.

97. Duarte FV, Palmeira CM, Rolo AP. The role of microRNAs in mitochondria: small players acting wide. *Genes (Basel)* 2014;**5**:865–886.

98. Kwok A, Raulf N, Habib N. Developing small activating RNA as a therapeutic: current challenges and promises. *Ther Deliv* 2019;**10**:151–164.

99. Cabili MN, Dunagin MC, McClanahan PD, Biaesch A, Padovan-Merhar O, Regev A, Rinn JL, Raj A. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol* 2015;**16**:20.

100. Dragomir M, Chen B, Calin GA. Exosomal lncRNAs as new players in cell-to-cell communication. *Transl Cancer Res* 2018;**7**:S243–S252.

101. Navickas R, Gal D, Laucevičius A, Taparauskaitè A, Zdanytè M, Holvoet P. Identifying circulating microRNAs as biomarkers of cardiovascular disease: a systematic review. *Cardiovasc Res* 2016;**111**:322–337.

102. Liebetrau C, Möllmann H, Dörr O, Szardien S, Troidl C, Willmer M, Voss S, Gaede L, Rixe J, Rolf A, Hamm C, Nef H. Release kinetics of circulating muscle-enriched microRNAs in patients undergoing transcoronary ablation of septal hypertrophy. *J Am Coll Cardiol* 2013;**62**:992–998.

103. Schulte C, Barwari T, Joshi A, Theofilatos K, Zampetaki A, Barallobre-Barreiro J, Singh B, Sörensen NA, Neumann JT, Zeller T, Westermann D, Blankenberg S, Marber M, Liebetrau C, Mayr M. Comparative analysis of circulating noncoding RNAs versus protein biomarkers in the detection of myocardial injury. *Circ Res* 2019;**125**:328–340.

104. Widera C, Gupta SK, Lorenzen JM, Bang C, Bauersachs J, Bethmann K, Kempf T, Wollert KC, Thum T. Diagnostic and prognostic impact of six circulating microRNAs in acute coronary syndrome. *J Mol Cell Cardiol* 2011;**51**:872–875.

105. Devaux Y, Mueller M, Haaf P, Goretti E, Twerenbold R, Zangrando J, Vausort M, Reichlin T, Wildi K, Moehring B, Wagner DR, Mueller C. Diagnostic and prognostic value of circulating microRNAs in patients with acute chest pain. *J Intern Med* 2015;**277**:260–271.

106. Kaudewitz D, Skroblin P, Bender LH, Barwari T, Willeit P, Pechlaner R, Sunderland NP, Willeit K, Morton AC, Armstrong PC, Chan MV, Lu R, Yin X, Gracio F, Dudek K, Langley SR, Zampetaki A, de Rinaldis E, Ye S, Warner TD, Saxena A, Kiechl S, Storey RF, Mayr M. Association of microRNAs and YRNAs with platelet function. *Circ Res* 2016;**118**:420–432.

107. Zampetaki A, Willeit P, Tilling L, Drozdov I, Prokopi M, Renard J-M, Mayr A, Weger S, Schett G, Shah A, Boulanger CM, Willeit J, Chowienczyk PJ, Kiechl S, Mayr M. Prospective study on circulating MicroRNAs and risk of myocardial infarction. *J Am Coll Cardiol* 2012;**60**:290–299.

108. Jakob P, Kacprowski T, Briand-Schumacher S, Heg D, Klingenberg R, Stähli BE, Jaguszewski M, Rodondi N, Nanchen D, Räber L, Vogt P, Mach F, Windecker S, Völker U, Matter CM, Lüscher TF, Landmesser U. Profiling and validation of circulating microRNAs for cardiovascular events in patients presenting with ST-segment elevation myocardial infarction. *Eur Heart J* 2017;**38**:511–515.

109. Wingrove JA, Daniels SE, Sehnert AJ, Tingley W, Elashoff MR, Rosenberg S, Buellesfeld L, Grube E, Newby LK, Ginsburg GS, Kraus WE. Correlation of peripheral-blood gene expression with the extent of coronary artery stenosis. *Circ Cardiovasc Genet* 2008;**1**:31–38.

110. Vanhaverbeke M, Vausort M, Veltman D, Zhang L, Wu M, Laenen G, Gilljins H, Moreau Y, Bartunek J, Van de Werf F, Devaux Y, Janssens S, Sinnaeve PR, EU-CardioRNA COST Action CA17129. Peripheral blood RNA levels of QSOX1 and PLBD1 are new independent predictors of left ventricular dysfunction after acute myocardial infarction. *Circ Genom Precis Med* 2019;**12**:e002656.

111. Kumarswamy R, Bauters C, Volkmann I, Maury F, Fetisch J, Holzmann A, Lemesle G, de Groote P, Pinet F, Thum T. Circulating long noncoding RNA, LIPCAR, predicts survival in patients with heart failure. *Circ Res* 2014;**114**:1569–1575.

112. Mäki A, Tiirola M. Directional high-throughput sequencing of RNAs without gene-specific primers. *Biotechniques* 2018;**65**:219–223.

113. Verwilt J, Trypsteen W, Van Paemel R, De Preter K, Giraldez MD, Mestdagh P, Vandesompele J. When DNA gets in the way: a cautionary note for DNA contamination in extracellular RNA-seq studies. *Proc Natl Acad Sci USA* 2020;**117**:18934–18936.

114. Deng MC, Eisen HJ, Mehra MR, Billingham M, Marboe CC, Berry G, Kobashigawa J, Johnson FL, Starling RC, Murali S, Pauly DF, Baron H, Wohlgemuth JG, Woodward RN, Klingler TM, Walther D, Lal PG, Rosenberg S, Hunt S, CARGO Investigators. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am J Transplant* 2006;**6**:150–160.

115. Elashoff MR, Wingrove JA, Beineke P, Daniels SE, Tingley WG, Rosenberg S, Voros S, Kraus WE, Ginsburg GS, Schwartz RS, Ellis SG, Tahirkheli N, Waksman R, McPherson J, Lansky AJ, Topol EJ. Development of a blood-based gene expression

algorithm for assessment of obstructive coronary artery disease in non-diabetic patients. *BMC Med Genomics* 2011;**4**:26.

116. Pham MX, Teuteberg JJ, Kfoury AG, Starling RC, Deng MC, Cappola TP, Kao A, Anderson AS, Cotts WG, Ewald GA, Baran DA, Bogaev RC, Elashoff B, Baron H, Yee J, Valantine HA. Gene-expression profiling for rejection surveillance after cardiac transplantation. *N Engl J Med* 2010;**362**:1890–1190.

117. de Gonzalo-Calvo D, Cediel G, Bär C, Núñez J, Revuelta-Lopez E, Gavara J, Ríos-Navarro C, Llorente-Cortes V, Bodí V, Thum T, Bayes-Genis A. Circulating miR-1254 predicts ventricular remodeling in patients with ST-Segment-Elevation Myocardial Infarction: a cardiovascular magnetic resonance study. *Sci Rep* 2018;**8**: 15115.

118. de Gonzalo-Calvo, D., Kenneweg, F., Bang, C., Toro, R., van der Meer, R.W., Rijzewijk, L.J., Smit, J.W., Lamb, H.J., Llorente-Cortes, V., Thum, T. Circulating Long Noncoding RNAs in Personalized Medicine: Response to Pioglitazone Therapy in Type 2 Diabetes. J Am Coll Cardiol. 2016; 68(25):2914-2916.

119. Morley-Smith AC, Mills A, Jacobs S, Meyns B, Rega F, Simon AR, Pepper JR, Lyon AR, Thum T. Circulating microRNAs for predicting and monitoring response to mechanical circulatory support from a left ventricular assist device. *Eur J Heart Fail* 2014;**16**:871–879.