Received: 7 December 2023 | Revised: 9 April 2024 | Accepted: 29 April 2024

DOI: 10.1111/mec.17382

OPINION

MOLECULAR ECOLOGY WILEY

From whole bodies to single cells: A guide to transcriptomic approaches for ecology and evolutionary biology

Katja M. Hoedjes[1] | Sonja Grath[2] | Nico Posnien[3] | Michael G. Ritchie[4] | Christian Schlötterer[5] | Jessica K. Abbott[6] | Isabel Almudi[7,8] | Marta Coronado-Zamora[9] | Esra Durmaz Mitchell[10,11] | Thomas Flatt[10] | Claudia Fricke[12] | Amanda Glaser-Schmitt[2] | Josefa González[9] | Luke Holman[13] | Maaria Kankare[14] | Benedict Lenhart[15] | Dorcas J. Orengo[7,8] | Rhonda R. Snook[16] | Vera M. Yılmaz[2] | Leeban Yusuf[4]

Correspondence
Katja M. Hoedjes, Amsterdam Institute for Life and Environment, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.
Email: k.m.hoedjes@vu.nl

Funding information
MCIN/AEI/10.13039/501100011033, Grant/Award Number: PID2020-115874GB-I00; Swiss National Science Foundation, Grant/Award Number: 310030_219283, 31003A-182262 and FZEB-0-214654; Austrian Science Funds, Grant/Award Number: P32935, P33734 and W1225; Natural Environment Research Council, UK, Grant/Award Number: NE/V001566/1; Deutsche Forschungsgemeinschaft, Grant/Award Number: GR 4495/4-1, PO 1648/4-1 and PO 1648/7-1; The Research Council of Finland, Grant/Award Number: 322980; Departament de Recerca i Universitats, Generalitat de Catalunya, Grant/Award Number: 2021 SGR 00417

Handling Editor: Sean Rogers

Abstract

RNA sequencing (RNAseq) methodology has experienced a burst of technological developments in the last decade, which has opened up opportunities for studying the mechanisms of adaptation to environmental factors at both the organismal and cellular level. Selecting the most suitable experimental approach for specific research questions and model systems can, however, be a challenge and researchers in ecology and evolution are commonly faced with the choice of whether to study gene expression variation in whole bodies, specific tissues, and/or single cells. A wide range of sometimes polarised opinions exists over which approach is best. Here, we highlight the advantages and disadvantages of each of these approaches to provide a guide to help researchers make informed decisions and maximise the power of their study. Using illustrative examples of various ecological and evolutionary research questions, we guide the readers through the different RNAseq approaches and help them identify the most suitable design for their own projects.

KEYWORDS
bulk RNAseq, cellular heterogeneity, deconvolution, gene expression, single-cell RNAseq, transcriptomics

1 | INTRODUCTION

Understanding phenotypic diversity and adaptation is a key goal of evolutionary and ecological research. Most phenotypes have a complex genetic basis that depends on the effects of a large number of genetic loci (sometimes thousands; Barton, 2022; Boyle et al., 2017), which are expressed in interaction with the environment. Gene transcription represents the first step towards translating genotype information into phenotypes, and its products interact in complex gene regulatory networks and signalling cascades. Moreover, the

The European Drosophila Population Genomics Consortium (DrosEU): Katja M. Hoedjes, Sonja Grath, Nico Posnien, Michael G. Ritchie, Christian Schlötterer, Jessica K Abbott, Marta Coronado-Zamora, Esra Durmaz Mitchell, Thomas Flatt, Claudia Fricke, Amanda Glaser-Schmitt, Josefa González, Luke Holman, Maaria Kankare, Benedict Lenhart, Dorcas J. Orengo, Rhonda R. Snook, Vera M. Yılmaz, Leeban Yusuf.

Katja M. Hoedjes, Sonja Grath, Nico Posnien, Michael G. Ritchie and Christian Schlötterer contributed equally to writing the manuscript.

For affiliations refer to page 7.

transcriptome integrates a range of genetic, environmental and physiological signals (e.g. Buchberger et al., 2019; Everett et al., 2020; Hill et al., 2021). Accordingly, evolutionary change in spatial and temporal patterns of gene expression is one of the main drivers of phenotypic differentiation (e.g. Barbosa-Morais et al., 2012; Brawand et al., 2011; Fukushima & Pollock, 2020; Gerhart & Kirschner, 2007; Mantica et al., 2024; Shapiro et al., 2004; Steiner et al., 2007). Gene expression analyses based on RNA sequencing (RNAseq) provide a powerful tool to study a wide range of ecological and evolutionary questions (reviewed e.g. in Oppenheim et al., 2015; Stark et al., 2019), in particular since they allow quantification of gene expression in organisms without high-quality reference genomes (Chalifa-Caspi, 2021; Cheng et al., 2018; Haas et al., 2013). Still, RNAseq analyses certainly profit from the availability of genome references. For instance, quantification of gene expression strongly depends on the quality of the available genome assembly and annotation (Torres-Oliva et al., 2016).

Isolation of RNA is the starting point for any RNAseq experiment. For many organisms, such as vertebrates and plants, RNA is usually isolated from specific organs, body parts or tissue samples for RNAseq because the whole body is too large. For smaller organisms, such as small arthropods or other invertebrates, one can also choose to conduct gene expression analysis on RNA extracted from the whole body. Although possible and also done, tissue-specific RNAseq is often more technically challenging for small organisms, and whole body RNAseq has therefore been common practice. In either case, one major complication of gene expression analyses in multicellular organisms lies in the large number of different cell types present within whole bodies, body parts and even specific tissues. For instance, the small nematode worm *Caenorhabditis elegans* is composed of 959 somatic cells, and the 302 cells of the nervous system fall into 128 different neuron types (Taylor et al., 2021). The human body consists of more than 3×10^{13} cells (Bianconi et al., 2013), and the first 400 cell types of 24 different tissues have only recently been characterised (Tabula Sapiens et al., 2022). Accordingly, RNAseq of whole bodies, body parts or tissues (i.e. bulk RNAseq) is based on a mixture of cells with specific expression patterns. This implies that bulk RNAseq analyses reflect the gene expression averaged across multiple cell types with distinct expression patterns. Recently developed single-cell RNAseq (scRNAseq) approaches allow quantifying expression at the cellular level to elucidate differences among multiple cell types, thereby presenting an alternative to bulk RNAseq approaches (e.g. Alfieri et al., 2022; Nguyen et al., 2018; Wang, Sun, et al., 2021).

It is often difficult to decide whether it is advantageous to study gene expression in specific tissues, body parts, a collection of organs, entire bodies or individual cells. In this opinion article, we highlight major technical and methodological advantages and limitations of gene expression studies based on whole bodies, organs and tissues (bulk RNAseq), as well as of recent scRNAseq methods. We propose guidelines for typical research questions in ecology and evolution, and we highlight how scRNAseq can enhance the merit of future and existing bulk RNAseq datasets, in particular when combined with whole body RNAseq.

2 | WHOLE BODY RNASEQ FOR A SYSTEMIC OVERVIEW OF GENE EXPRESSION

Bulk RNAseq methods depend on the principle that RNA is isolated from a heterogeneous set of cells, such as whole bodies or parts of it, and then sequenced in bulk. Whole body RNAseq has been especially popular in studies of small organisms (e.g. Bouvaine et al., 2012; Crawford et al., 2010; Teets et al., 2013; Winbush et al., 2012), as it is cost-effective and it can be challenging to obtain sufficient high-quality RNA from only a specific tissue of an individual. As whole body RNAseq provides an overview of the averaged gene expression patterns across all cells within an individual, the major advantage of this approach lies in the potential to provide a systemic overview of gene expression across the entire organism (Figure 1). Whole body RNAseq is therefore especially powerful when studying the evolutionary or ecological responses of phenotypes without a priori expectations about affected tissues, cell types, or specific physiological mechanisms. For example, studies of evolved and plastic responses underlying thermal adaptation have uncovered a diversity of physiological processes, which suggests broad systemic effects that depend on different tissues and cell types (e.g. Hsu et al., 2020; Kankare et al., 2016; Koniger & Grath, 2018; Mallard et al., 2020; Parker et al., 2021). In this case, focussing on specific tissues not only results in an incomplete analysis, but it could also potentially introduce a substantial bias if a non-representative tissue is selected, affecting subsequent interpretation. Nevertheless, whole body RNAseq also comes with its specific challenges and considerations, which we briefly discuss below.

2.1 | Genetic heterogeneity and pooling

Depending on the size of the organism (parts of) multiple individuals may be pooled prior to RNAseq library preparation. If all individuals have an identical genotype and are derived from the same environment, pooling clearly provides a benefit because stochastic inter-individual variation is reduced. Pooling individuals with different genotypes is frequently done when genetically identical individuals are not available. If many genetically distinct individuals are pooled, a reliable and representative average gene expression level of the population can be obtained. The drawback of this approach is that information about the heterogeneity in gene expression in the population, or differences in heterogeneity between populations, is lost. Reliable estimates of variance require very large population samples, however, which are often not possible due to limited access to material and restrictions in research budget. We assume that pooling individuals for RNAseq approaches comes with challenges that also apply for Pool-Seq of genomic DNA (e.g. Futschik & Schlötterer, 2010) and we propose that expression variation due to genetic heterogeneity warrants more attention in future research.

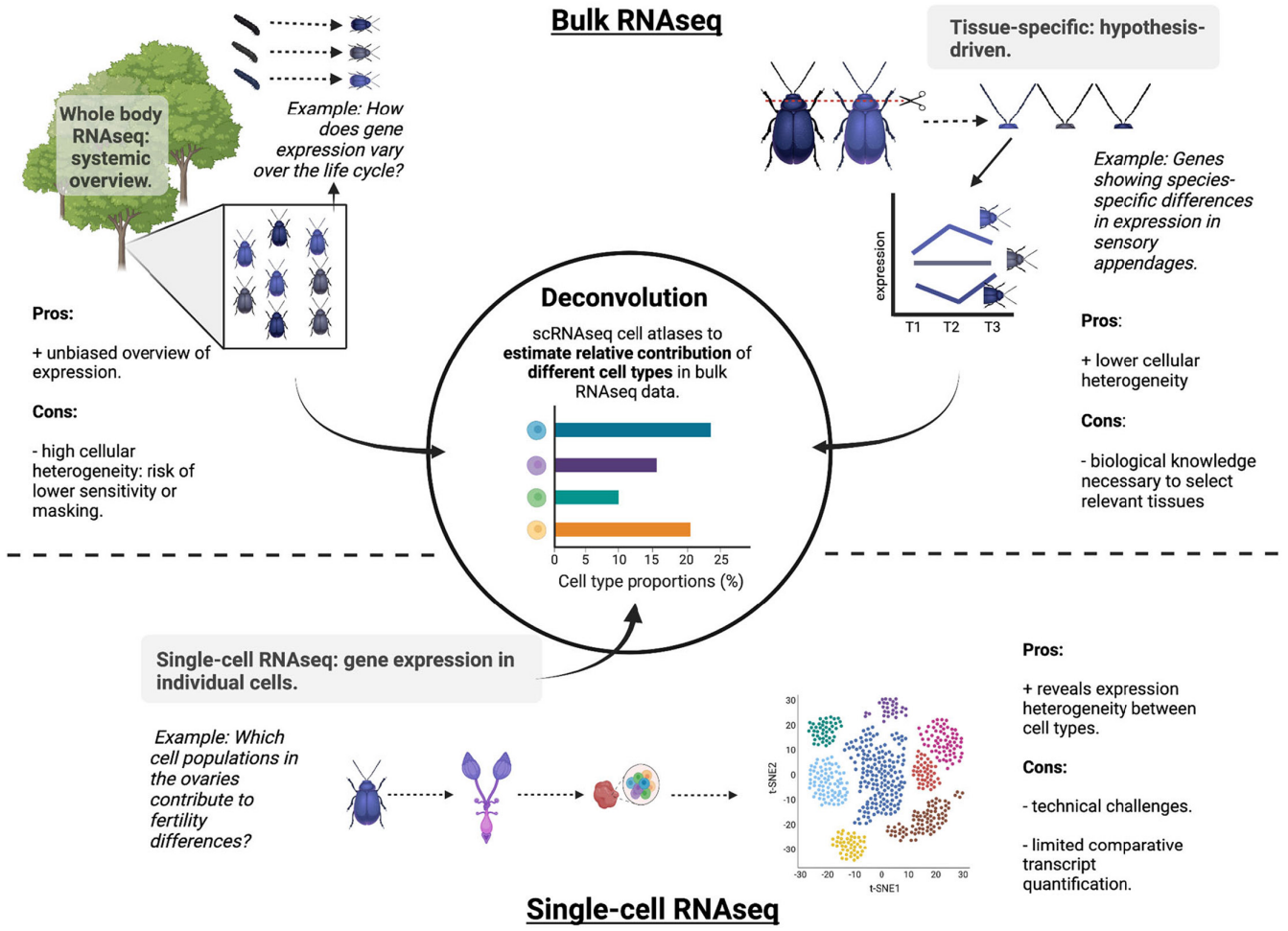

FIGURE 1 Advantages and disadvantages of RNAseq approaches, and opportunities for deconvolution. Each RNAseq approach, using either whole bodies, body parts or specific tissues (i.e. bulk RNAseq), or single cells (i.e. single-cell RNAseq), has unique advantages and disadvantages that make them more or less suitable to address particular questions in the fields of ecology and evolution. Integrating different approaches can provide additional advantages, by harnessing the strengths of each approach. Here, we highlight deconvolution, by combining bulk and single-cell RNAseq data, as a promising opportunity that has recently opened up for ecological and evolutionary research.

2.2 | Sensitivity

All bulk RNAseq approaches measure average mRNA abundance among all cells in a whole body or specific tissue. Thus, changes affecting only a small number of cells might be difficult to detect against the background of the rest of the body (or tissue) with no expression changes (Kuhn et al., 2012). For example, a comparative analysis of expression patterns of honey bee sting glands and digestive tract versus whole abdomen demonstrated that 81% and 69% of significantly differentially expressed genes in these two tissues, respectively, were not detected as differentially expressed in the composite tissue (Johnson et al., 2013). Likewise, in a study on diabetes in humans, 35% of eQTLs identified from pancreatic islet data could not be detected when analysing whole pancreas data (Vinuela et al., 2020). In both cases, a plausible explanation is that some expression differences are limited to particular cell types, which become more difficult to detect when other cell types are present in the RNA extract.

One possible solution to improve sensitivity is to increase either coverage or, preferentially, the number of biological replicates. For example, Jaksic et al. (2020) measured whole body gene expression in *Drosophila simulans* that had been experimentally evolved under either warm or cool conditions. Using a relatively large sample size of 20 populations (10 warm-evolved, 5 cool-evolved and 5 ancestral) from which they sequenced a pooled RNA sample from the whole body of 50 males each, they identified differences in gene expression in dopaminergic neurons in the brain. Although these neurons constitute only a small percentage of the cells of the entire body, significant differences in gene expression could be identified and functionally validated using RNA interference mediated knockdown and pharmacological intervention (Jaksic et al., 2020). Hence, given sufficient power to detect small changes, bulk RNAseq of whole bodies can be a promising discovery tool even if only a subset of the cell population is affected.

2.3 | Masking of signals in opposite directions

Many genes have pleiotropic functions, and are expressed in multiple cell types. For example, systematic studies of yeast found that more than 50% of tested mutations affected two or more organismal phenotypes (Dudley et al., 2005; Mattiazzi Usaj et al., 2020). Given widespread pleiotropy, in combination with transcriptional heterogeneity among cell types, it is conceivable that the experimental conditions studied might affect gene expression in multiple tissues or cell types, but in opposite directions. As bulk RNAseq only provides average expression levels of transcripts across all tissues and cells, such changes may partly or completely mask one another, essentially cancelling each other out so that no overall expression difference can be detected for a certain gene. Although this scenario is possible and warrants a careful interpretation of the outcomes of bulk RNAseq experiments, we expect that the exact cancellation of expression differences is unlikely in most cases. The previously discussed comparative study on honey bees indicated that a small proportion (<3%) of the differentially expressed genes exhibited an opposite differential expression pattern in the tissue-specific samples versus the whole abdomen, which could theoretically obscure interpretation of the results if the expression patterns cancel each other out completely (Johnson et al., 2013). However, if differences in gene expression are only partially masked by cell type heterogeneity in bulk RNAseq experiments, it will often still be possible to detect these differences when a sufficiently powerful experimental design (number of replicates and sequencing depth) is used.

2.4 | Allometric changes

Expression differences in whole body RNAseq samples could have two different causes, which are not mutually exclusive and could both represent adaptive responses: expression might differ within one or more cell types, or the relative proportion of different cell types in the whole body might be different. Relative size, in other words pleiotropy, allometry, of specific tissues (each differing in their gene expression profiles) can therefore be a key determinant of whole body RNAseq results, which is important to consider alongside transcriptional regulation. A frequently observed allometric change in evolutionary and ecological studies concerns, for example, the relative size of reproductive tissues. Here the allometric changes can be a mechanism by which an organism adapts to its environment; for example, males of polygamous species regularly increase testes size under increased sperm competition (Montgomery & Mank, 2016 and references therein). A point of concern is that allometric relationships could also lead to false negatives, when changes in expression occur but are offset by decreases in tissue size (Abell et al., 2022).

Hsu et al. (2020) proposed an approach to disentangle the effects of allometry versus transcriptional regulation on expression levels. They observed that ovaries of female flies that had evolved in a novel warm environment were enlarged, which made it difficult to draw conclusions on the evolution of transcriptional regulation. To overcome this limitation, Hsu et al. (2020) measured gene expression in gonads and carcasses of the same flies and determined a measure for allometric change, which was then used to correct the whole body expression data. This example shows that whole body RNAseq alone or in conjunction with tissue-specific RNAseq can be a powerful, unbiased approach to identify evolutionary or ecologically relevant expression changes, which tissue-specific RNAseq alone might fail to identify. Alternatively, tissue-specific RNAseq in conjunction with quantitative estimates of the relative contribution of that tissue between conditions (e.g. populations, environments and treatments) could also distinguish between regulatory evolution and allometric differences (Buono et al., 2021). This could be achieved with fluorescent imaging (e.g. immunohistology) or transgenic cell labelling methods in combination with fluorescence-activated cell sorting. In either case, it is important to recognise that both allometric and gene regulatory mechanisms can account for observed differences in whole body RNAseq data, but both mechanisms represent a genuine adaptive response that can be uncovered by whole body RNAseq.

3 | FROM SYSTEMIC TO HYPOTHESIS-DRIVEN RNASEQ

The evolution of complex body plans of multicellular organisms and their functions are intricately linked to tissue-specific transcriptomic differentiation (Barbosa-Morais et al., 2012; Brawand et al., 2011; Fukushima & Pollock, 2020; Mantica et al., 2024). For some types of research questions, a defined morphological structure or tissue may be of interest, in particular when specific target tissues responsible for the adaptive mechanism(s) have already been identified. In these cases, RNAseq of specific tissues and cell types may be the most logical step because the smaller number of different cell types increases the power to detect relevant gene expression differences (Figure 1). Also, complex phenotypic adaptations can depend on divergent transcriptional changes in different tissues, which might be missed when using a whole body RNAseq approach. For example, Salvador-Martinez et al. (2018) analysed spatial gene expression patterns from different *D. melanogaster* embryonic anatomical structures, which indicated that genes expressed in the digestive system and ectoderm-derived structures are under selective constraint, while genes expressed in the germ line showed high rates of adaptive substitution.

The key prerequisite for a successful tissue-specific RNAseq analysis is sufficient biological knowledge to select the relevant tissues for the question of interest. An example comes from analyses of post-mating gene expression changes in female fruit flies, triggered by the male ejaculate. Various studies have focussed on detecting differential expression in the female reproductive tract and the ovaries, which is extensive and can even be specific for distinct sub-tissues (e.g. Kapelnikov et al., 2008; Veltsos et al., 2022). At the same time, receptors for sex peptide, one of the most important components of the ejaculate triggering female post-mating

effects, have been found throughout the female body, including in the abdominal ganglia and the brain. This indicates that key targets of sex peptide are located in tissues outside the female reproductive system (e.g. Gioti et al., 2012; Pasquier & Robichon, 2022; Yapici et al., 2008) and would be missed by a tissue-specific RNAseq study. Hence these examples illustrate the power of a combined application of whole body and tissue-specific RNAseq. As with all bulk RNAseq approaches, tissue-specific RNAseq estimates gene expression of mixed cell types, and hence the technical limitations discussed above, also apply to tissue specific RNAseq, although to a lesser extent due to lower cellular heterogeneity within tissues as compared to whole bodies.

4 | SINGLE-CELL RNASEQ: CAPTURING GENE EXPRESSION DIFFERENCES AMONG CELLS

Recent single-cell sequencing methods (scRNAseq) can address some of the limitations of bulk RNAseq by measuring gene expression in large numbers of individual cells (e.g. Alfieri et al., 2022; Nguyen et al., 2018; Wang, He, et al., 2021). By revealing the heterogeneity in expression between cells, as well as variation in transcriptional profiles among species, populations and experimental conditions, scRNAseq has great potential to uncover mechanisms underpinning ecologically or evolutionarily salient phenotypic variation (Figure 1). For instance, scRNAseq was recently applied to fat bodies in fruit flies to uncover the cellular mechanisms underlying the trade-off between reproduction and immunity. It was demonstrated that limitations in the capacity of a specific subset of fat body cells to produce proteins constrains the immune response in reproducing females (Gupta et al., 2021). Despite its promises, scRNAseq comes with challenges that warrant careful consideration during the project planning phase.

4.1 | Sample availability and preparation

A typical scRNAseq workflow requires access to at least 50,000 live cells (Pollen et al., 2014) and the availability of fresh material is often limited in ecological or evolutionary studies that deal with field samples. Potential solutions could be the sequencing of single nuclei (snRNAseq), which can be isolated from flash frozen tissue samples to allow some level of tissue conservation (Denisenko et al., 2020; Wiegleb et al., 2022). Similarly, simultaneous cell dissociation and fixation has been applied to preserve tissue samples for scRNAseq (Garcia-Castro et al., 2021). However, obtaining sufficient material for small organisms or tiny tissue samples requires pooling of multiple individuals, which potentially introduces bias due to genetic heterogeneity (see Section 2.1 above). As the efficiency of capturing different cell types from a heterogeneous tissue varies considerably due to differences in cell size or cell shape, technical bias in cell type composition is expected (Darmanis et al., 2015; Yim et al., 2022). Thus, accurate and repeatable outcomes depend critically on the workflow used for sample preparation. Optimal parameters are highly tissue- and species-specific, which may necessitate laborious and often costly empirical optimisation, especially for non-model organisms (Svensson et al., 2018).

4.2 | RNA content and gene coverage

Each individual cell comprises very little RNA per gene resulting in high sampling variation and thus uncertainties for transcript quantification. This limitation means that, if no expression of a gene is detected in a certain cell, it is almost impossible to distinguish between a biological explanation (i.e. a gene is indeed not expressed) and a technical one (i.e. no reads due to inefficient sampling). The occurrence of such null data (also called dropout events) is much higher in scRNAseq compared to bulk RNAseq data (Bacher & Kendziorski, 2016). scRNAseq methods differ significantly in the number of genes that can unequivocally be detected per cell and the number of cells that can be analysed (Ziegenhain et al., 2017). For instance, the Smart-Seq2 method can detect many genes including low abundance transcripts in a few hundred cells in one run, while the 10× genomics method can process up to 10,000 cells per run, but with a higher noise level for low abundance transcripts (Wang, He, et al., 2021). While novel methods based on multiple rounds of in-cell barcoding (i.e. combinatorial indexing; e.g. split-pool ligation-based transcriptome sequencing (SPLiT-seq)) (Cao et al., 2019; Conte et al., 2023; Rosenberg et al., 2018) may mitigate this trade-off in the near future, scRNAseq applications are still limited for comparative transcript quantification. This is especially important to consider if the expected expression differences are rather small (Zhang et al., 2020), which is often the case in ecological or evolutionary research, for example, for studies examining environmental effects on individuals of the same species.

4.3 | Need for well-annotated reference genomes

The commonly used scRNAseq methods rely on sequencing the 3-prime ends of captured fragments (Ziegenhain et al., 2017), which can only be linked to particular genes if a well-annotated reference genome is available. Moreover, in snRNAseq nascent RNA is sequenced as well as mature mRNA, resulting in about one quarter of all reads originating from introns (Grindberg et al., 2013). Those reads can only be unequivocally assigned to a gene if the reads are mapped against a well-annotated genome, which is often unavailable for non-model systems used in ecological or evolutionary research. This is in contrast to bulk RNAseq data, which covers entire transcripts facilitating the de novo assembly of reference transcriptomes for subsequent transcript quantification (Grabherr et al., 2011; Haas et al., 2013; Raghavan et al., 2022), avoiding the need for a reference genome. Hence, differential gene expression analyses based on bulk

RNAseq are readily applicable for non-model organisms without the need for a reference genome.

In light of the outlined technical limitations and significantly higher costs, we think that scRNAseq has not yet reached the required maturity for a routine application in most research in ecology and evolution. In the following section, however, we discuss how single cell expression atlases obtained by shallow sequencing of a large number of cells can be used to enhance the interpretation of existing and future whole body or tissue-specific RNAseq data.

5 | DECONVOLUTION: MERGING TWO WORLDS

Biological interpretation of gene expression data from a heterogeneous cell population is the biggest challenge for bulk RNAseq. To overcome this problem, various approaches to assess whether differentially expressed genes are enriched for certain molecular or biological functions (i.e. gene ontology (GO) enrichment) (Gene Ontology et al., 2023; Thomas et al., 2022) or for specific signalling pathways (i.e. pathway enrichment) (Reimand et al., 2019) are routinely applied. Such functional enrichment analyses can help to extract biologically meaningful hypotheses or pinpoint candidate cell types or tissues from the often large numbers of differentially expressed genes (e.g. Stanford et al., 2020). These analyses can readily be applied in many model organisms for which databases with gene-to-function information are available. For instance, a study on the evolution of gene expression variance in flies used this approach to show that a small set of genes with a significant loss in expression variance was over-represented among genes with catabolic function and expression in the gut, suggesting that gene expression evolved in response to a less variable diet in the laboratory (Lai & Schlotterer, 2022). Similarly, Green et al. (2022) identified differentially expressed genes in response to copper exposure using whole body RNAseq of natural *Drosophila melanogaster* populations. A GO enrichment analysis identified the midgut as a candidate tissue and established a link between the preservation of gut acidity and tolerance to copper (Green et al., 2022). However, to leverage such functional databases for non-model organisms, a solid gene orthology assignment must be established prior to the enrichment analyses. While functional enrichment analyses are often the only way to extract biological information from gene lists, they depend on current knowledge, which is inherently incomplete and biased (Dessimoz & Skunca, 2017).

scRNAseq data provides exciting new opportunities to extract biological information from lists of differentially expressed genes obtained from bulk RNAseq approaches. The assessment of transcriptional profiles of distinct cell types has resulted in the development of single cell atlases across tissues, conditions, developmental stages (e.g. Allen et al., 2020; Brunet Avalos et al., 2019; Chen et al., 2021; Corrales et al., 2022; Hu, Comjean, et al., 2021; Hu, Tattikota, et al., 2021; Karaiskos et al., 2017; Li et al., 2022; Najle et al., 2023; Papatheodorou et al., 2020; Suo et al., 2022; Wang, Sun, et al., 2021; Xu et al., 2021) and even for entire organisms (Li et al., 2022). Such single cell atlases can directly be used to ask whether genes identified in bulk RNAseq approaches may be expressed in specific cell types. Moreover, deconvolution methods have been established, which estimate the relative contribution of different cell types in bulk RNAseq datasets (Mohammadi et al., 2017; Venet et al., 2001). When information about the gene expression profile of each cell constituting a bulk RNAseq sample is available, it is possible to decompose the average gene expression value into the expression of individual cell types that share an expression profile (Figure 1). Most expression deconvolution methods require prior information about cell-type specific marker genes (Zaitsev et al., 2019) or expression profiles encompassing multiple genes (Avila Cobos et al., 2018, 2020), whose expression level is directly correlated with the abundance of that cell type in a heterogeneous tissue. While such cell-type specific marker genes or expression profiles are typically identified based on extensive prior knowledge in model organisms (e.g. spatial expression data and functional assays), scRNAseq now allows the identification of such markers for many non-model systems too (e.g. Andrade Barbosa et al., 2021).

Expression deconvolution has successfully been applied to estimate the dynamics of cell type composition in evolving yeast populations exposed to environmental stress, the induction of extensive DNA damage and during sexual reproduction (i.e. sporulation) (Lu et al., 2003). In another eco-evolutionary study, a scRNAseq analysis of three populations of three-spined stickleback fish (*Gasterosteus aculeatus*) that exhibit natural variation in parasite resistance revealed differences in the composition and cell-type specific expression profiles of immune cells. Moreover, immune cell-type specific marker genes identified using the scRNAseq data were used to re-analyse previous tissue-specific RNAseq datasets of F_2 crosses to show that the response of gene expression in antigen-presenting cells to infection is most likely the result of regulatory variation and not due to an increase in the number of antigen-presenting cells (Fuess & Bolnick, 2023). Combining single cell expression atlases with bulk RNAseq datasets to estimate the expression profiles of individual cell types thus provides an important extra layer of biological insight that can be obtained from whole body or tissue-specific RNAseq datasets. We argue that the impact of deconvolution methods is largest for complex tissues or even whole bodies, and can mitigate some of the drawbacks associated with bulk RNAseq approaches, as discussed above. As such, the combination of whole body RNAseq with scRNAseq has the potential to bring together the best of both worlds by combining a systemic overview of gene expression with cell-type specific information.

6 | CONCLUSIONS AND RECOMMENDATIONS

In light of such exciting opportunities, but also the challenges of different RNAseq approaches, it is often not easy to decide which

strategy is best suited for a certain research question. The first important consideration is the phenotype that is being studied. For instance, if one wants to identify genes underlying the evolution of the morphology of a specific organ, such as the formation of pigment spots on insect wings (Hanly et al., 2019; Wee et al., 2023) or the size and shape of beetle horns (Emlen et al., 2007; Ohde et al., 2018) it is logical to restrict RNAseq analyses to that specific tissue. Similarly, the evolution of many behavioural traits is for an important part likely linked to variation in the composition and function of the central nervous system (CNS). Therefore, CNS-specific RNAseq experiments will be powerful to identify meaningful candidate genes (even though the CNS itself is a complex tissue with multiple cell types performing distinct roles). We caution, however, that often phenotypes cannot be restricted to single tissues. A nice example for this challenge comes from phenotypes related to the stress response (Horvath et al., 2023) or life history (Rodrigues et al., 2021) of an organism, where more systemic transcriptomic changes involving multiple organs and tissues are expected. In such cases, whole body RNAseq analyses could limit the bias due to missing biologically important signals.

A second important consideration is the level of prior knowledge, available tools and access to annotated genomes of the organism of study. While extensive mechanistic details are publicly available for many organs in model organisms, such data are often missing for non-model systems. If no hypotheses about specific tissues or cell types underlying a certain phenotypic trait are available, a whole body bulk RNAseq approach provides a systemic and comprehensive overview of potential gene expression changes. If putative target tissues have been identified based on prior whole body expression data or based on similar phenotypes in other models, tissue-specific RNAseq has the potential to assess variable expression of genes with minute differences in expression or lower overall expression levels, which may be missed in whole body RNAseq.

A third consideration is the opportunity to combine different RNAseq approaches to strengthen the biological interpretation of RNAseq data, while diminishing the inherent weaknesses of each individual approach. The combination of whole body RNAseq to establish first hypotheses followed by integration of tissue-specific expression information has been successfully applied in ecological and evolutionary research to identify mechanisms of phenotypic variation (Abbott et al., 2020; Lai & Schlotterer, 2022). Moreover, with increasing availability of single-cell expression data for model organisms and non-model organisms alike, existing and future bulk RNAseq data can now be (re-)interpreted to gain novel insights into tissue- and cell-type specific gene expression divergence. Most importantly, expression deconvolution methods have a great potential to distinguish between regulatory evolution and allometric differences, respectively, observed in bulk RNAseq data. We argue that future eco-evolution research would profit from a community driven development of tissue databases and single cell atlases for entire small organisms or specific tissues to facilitate the identification of cell-type specific or condition-dependent markers. Such reference databases and markers will allow a better interpretation and integration of future bulk RNAseq data, which is still the best accessible and often most informative technology for most questions in eco-evolution research.

In conclusion, we argue that both bulk RNAseq, whether whole body, body-part or tissue-specific, and scRNAseq have unique advantages and disadvantages. The best choice of RNAseq approach depends strongly on the model system, prior knowledge of the phenotype, and the biological level of interest (e.g. gene regulatory networks, physiology, allometry or phylogenies). Integrating different RNAseq methods allows to harness the strengths, versatility and opportunities of each approach to study research questions in ecology and evolution.

AUTHOR CONTRIBUTIONS

Conceptualisation: KMH, SG, NP, MGR, CS, EDM, JKA, TF, AG-S, JG, LH, MK, BL, VMY, LY; Visualisation: LY; Writing – Original Draft Preparation: KMH, SG, NP, MGR, CS; Writing – Review and Editing: KMH, SG, NP, MGR, CS, EDM, JKA, IA, MC-Z, TF, CF, AG-S, JG, LH, MK, BL, DJO, RRS, VMY, LY.

AFFILIATIONS

[1]Amsterdam Institute for Life and Environment, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[2]Division of Evolutionary Biology, LMU Munich, Planegg-Martinsried, Germany
[3]Department of Developmental Biology, Göttingen Center for Molecular Biosciences (GZMB), University of Göttingen, Göttingen, Germany
[4]Centre for Biological Diversity, University of St Andrews, St Andrews, UK
[5]Institut für Populationsgenetik, Vetmeduni Vienna, Vienna, Austria
[6]Biology Department, Lund University, Lund, Sweden
[7]Departament de Genètica, Microbiologia i Estadística, Universitat de Barcelona, Barcelona, Spain
[8]Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Barcelona, Spain
[9]Institute of Evolutionary Biology, CSIC, UPF, Barcelona, Spain
[10]Department of Biology, University of Fribourg, Fribourg, Switzerland
[11]Functional Genomics and Metabolism Research Unit, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark
[12]Institute for Zoology/Animal Ecology, Martin-Luther-University Halle-Wittenberg, Halle (Saale), Germany
[13]School of Applied Sciences, Edinburgh Napier University, Edinburgh, UK
[14]Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland
[15]Department of Biology, University of Virginia, Charlottesville, Virginia, USA
[16]Department of Zoology, Stockholm University, Stockholm, Sweden

ACKNOWLEDGEMENTS

The authors were assisted by a Special Topics Network (STN) grant from the European Society for Evolutionary Biology (ESEB) to facilitate networking of the European *Drosophila* Population Genomics consortium (DrosEU; https://droseu.net/). SG and NP are supported by a Deutsche Forschungsgemeinschaft grant (GR 4495/4-1, PO 1648/7-1, PO 1648/4-1). MGR is supported by Natural Environment Research Council, UK (NE/V001566/1). CS is supported by the Austrian Science Funds (P32935, P33734, W1225). TF is supported by the Swiss National Science Foundation (310030_219283,

FZEB-0-214654,31003A-182262). JG is supported by grant PID2020-115874GB-I00 funded by MCIN/AEI/10.13039/501100011033 and from grant 2021 SGR 00417 funded by Departament de Recerca i Universitats, Generalitat de Catalunya. MK was supported by the grant 322980 from The Research Council of Finland.

CONFLICT OF INTEREST STATEMENT

The authors declare that there is no conflict of interest.

DATA AVAILABILITY STATEMENT

Not applicable.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results.

ORCID

Katja M. Hoedjes https://orcid.org/0000-0001-6090-232X
Nico Posnien https://orcid.org/0000-0003-0700-5595
Christian Schlötterer https://orcid.org/0000-0003-4710-6526
Thomas Flatt https://orcid.org/0000-0002-5990-1503
Claudia Fricke https://orcid.org/0000-0002-0691-6779
Josefa González https://orcid.org/0000-0001-9824-027X
Rhonda R. Snook https://orcid.org/0000-0003-1852-1448
Leeban Yusuf https://orcid.org/0000-0001-6383-1176

REFERENCES

Abbott, J. K., Chippindale, A. K., & Morrow, E. H. (2020). The micro-evolutionary response to male-limited X-chromosome evolution in *Drosophila melanogaster* reflects macroevolutionary patterns. *Journal of Evolutionary Biology*, *33*, 738–750.

Abell, N. S., DeGorter, M. K., Gloudemans, M. J., Greenwald, E., Smith, K. S., He, Z., & Montgomery, S. B. (2022). Multiple causal variants underlie genetic associations in humans. *Science*, *375*, 1247–1254.

Alfieri, J. M., Wang, G., Jonika, M. M., Gill, C. A., Blackmon, H., & Athrey, G. N. (2022). A primer for single-cell sequencing in non-model organisms. *Genes (Basel)*, *13*, 380.

Allen, A. M., Neville, M. C., Birtles, S., Croset, V., Treiber, C. D., Waddell, S., & Goodwin, S. F. (2020). A single-cell transcriptomic atlas of the adult *Drosophila* ventral nerve cord. *eLife*, *9*, e54074.

Andrade Barbosa, B., van Asten, S. D., Oh, J. W., Farina-Sarasqueta, A., Verheij, J., Dijk, F., van Laarhoven, H. W. M., Ylstra, B., Garcia Vallejo, J. J., van de Wiel, M. A., & Kim, Y. (2021). Bayesian log-normal deconvolution for enhanced in silico microdissection of bulk gene expression data. *Nature Communications*, *12*, 6106.

Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P., & De Preter, K. (2020). Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nature Communications*, *11*, 5650.

Avila Cobos, F., Vandesompele, J., Mestdagh, P., & De Preter, K. (2018). Computational deconvolution of transcriptomics data from mixed cell populations. *Bioinformatics*, *34*, 1969–1979.

Bacher, R., & Kendziorski, C. (2016). Design and computational analysis of single-cell RNA-sequencing experiments. *Genome Biology*, *17*, 63.

Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., Slobodeniuc, V., Kutter, C., Watt, S., Colak, R., Kim, T., Misquitta-Ali, C. M., Wilson, M. D., Kim, P. M., Odom, D. T., Frey, B. J., & Blencowe, B. J. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science*, *338*, 1587–1593.

Barton, N. H. (2022). The "new synthesis". *Proceedings of the National Academy of Sciences of the United States of America*, *119*, e2122147119.

Bianconi, E., Piovesan, A., Facchin, F., Beraudi, A., Casadei, R., Frabetti, F., Vitale, L., Pelleri, M. C., Tassani, S., Piva, F., Perez-Amodio, S., Strippoli, P., & Canaider, S. (2013). An estimation of the number of cells in the human body. *Annals of Human Biology*, *40*, 463–471.

Bouvaine, S., Behmer, S. T., Lin, G. G., Faure, M. L., Grebenok, R. J., & Douglas, A. E. (2012). The physiology of sterol nutrition in the pea aphid *Acyrthosiphon pisum*. *Journal of Insect Physiology*, *58*, 1383–1389.

Boyle, E. A., Li, Y. I., & Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. *Cell*, *169*, 1177–1186.

Brawand, D., Soumillon, M., Necsulea, A., Julien, P., Csardi, G., Harrigan, P., Weier, M., Liechti, A., Aximu-Petri, A., Kircher, M., Albert, F. W., Zeller, U., Khaitovich, P., Grützner, F., Bergmann, S., Nielsen, R., Pääbo, S., & Kaessmann, H. (2011). The evolution of gene expression levels in mammalian organs. *Nature*, *478*, 343–348.

Brunet Avalos, C., Maier, G. L., Bruggmann, R., & Sprecher, S. G. (2019). Single cell transcriptome atlas of the *Drosophila* larval brain. *eLife*, *8*, e50354.

Buchberger, E., Reis, M., Lu, T. H., & Posnien, N. (2019). Cloudy with a chance of insights: Context dependent gene regulation and implications for evolutionary studies. *Genes (Basel)*, *10*, 492.

Buono, L., Corbacho, J., Naranjo, S., Almuedo-Castillo, M., Moreno-Marmol, T., de la Cerda, B., Sanabria-Reinoso, E., Polvillo, R., Diaz-Corrales, F. J., Bogdanovic, O., Bovolenta, P., & Martínez-Morales, J. R. (2021). Analysis of gene network bifurcation during optic cup morphogenesis in zebrafish. *Nature Communications*, *12*, 3866.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, *566*, 496–502.

Chalifa-Caspi, V. (2021). RNA-Seq in nonmodel organisms. *Methods in Molecular Biology*, *2243*, 143–167.

Chen, D., Sun, J., Zhu, J., Ding, X., Lan, T., Wang, X., Wu, W., Ou, Z., Zhu, L., Ding, P., Wang, H., Luo, L., Xiang, R., Wang, X., Qiu, J., Wang, S., Li, H., Chai, C., Liang, L., … Xu, X. (2021). Single cell atlas for 11 non-model mammals, reptiles and birds. *Nature Communications*, *12*, 7083.

Cheng, H., Wang, Y., & Sun, M. A. (2018). Comparison of gene expression profiles in nonmodel eukaryotic organisms with RNA-seq. *Methods in Molecular Biology*, *1751*, 3–16.

Conte, M. I., Fuentes-Trillo, A., & Dominguez Conde, C. (2023). Opportunities and tradeoffs in single-cell transcriptomic technologies. *Trends in Genetics*, *40*, 83–93.

Corrales, M., Cocanougher, B. T., Kohn, A. B., Wittenbach, J. D., Long, X. S., Lemire, A., Cardona, A., Singer, R. H., Moroz, L. L., & Zlatic, M. (2022). A single-cell transcriptomic atlas of complete insect nervous systems across multiple life stages. *Neural Development*, *17*, 8.

Crawford, J. E., Guelbeogo, W. M., Sanou, A., Traore, A., Vernick, K. D., Sagnon, N., & Lazzaro, B. P. (2010). De novo transcriptome sequencing in *Anopheles funestus* using Illumina RNA-seq technology. *PLoS One*, *5*, e14202.

Darmanis, S., Sloan, S. A., Zhang, Y., Enge, M., Caneda, C., Shuer, L. M., Hayden Gephart, M. G., Barres, B. A., & Quake, S. R. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 7285–7290.

Denisenko, E., Guo, B. B., Jones, M., Hou, R., de Kock, L., Lassmann, T., Poppe, D., Clement, O., Simmons, R. K., Lister, R., & Forrest, A. R. R. (2020). Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biology*, *21*, 130.

Dessimoz, C., & Skunca, N. (2017). *The gene ontology handbook*. Springer Nature.

Dudley, A. M., Janse, D. M., Tanay, A., Shamir, R., & Church, G. M. (2005). A global view of pleiotropy and phenotypically derived gene function in yeast. *Molecular Systems Biology*, *1*, 2005.0001.

Emlen, D. J., Corley Lavine, L., & Ewen-Campen, B. (2007). On the origin and evolutionary diversification of beetle horns. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(Suppl 1), 8661–8668.

Everett, L. J., Huang, W., Zhou, S., Carbone, M. A., Lyman, R. F., Arya, G. H., Geisz, M. S., Ma, J., Morgante, F., St Armour, G., Turlapati, L., Anholt, R. R. H., & Mackay, T. F. C. (2020). Gene expression networks in the *Drosophila* genetic reference panel. *Genome Research*, *30*, 485–496.

Fuess, L. E., & Bolnick, D. I. (2023). Single-cell RNA sequencing reveals microevolution of the stickleback immune system. *Genome Biology and Evolution*, *15*, evad053.

Fukushima, K., & Pollock, D. D. (2020). Amalgamated cross-species transcriptomes reveal organ-specific propensity in gene expression evolution. *Nature Communications*, *11*, 4459.

Futschik, A., & Schlötterer, C. (2010). The next generation of molecular markers from massively parallel sequencing of pooled DNA samples. *Genetics*, *186*, 207–218.

Garcia-Castro, H., Kenny, N. J., Iglesias, M., Alvarez-Campos, P., Mason, V., Elek, A., Schonauer, A., Sleight, V. A., Neiro, J., Aboobaker, A., Permanyer, J., Irimia, M., Sebé-Pedrós, A., & Solana, J. (2021). ACME dissociation: A versatile cell fixation-dissociation method for single-cell transcriptomics. *Genome Biology*, *22*, 89.

Gene Ontology Consortium, Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L., Hill, D. P., Lee, R., Mi, H., Moxon, S., Mungall, C. J., Muruganugan, A., Mushayahama, T., Sternberg, P. W., Thomas, P. D., … Westerfield, M. (2023). The gene ontology knowledgebase in 2023. *Genetics*, *224*, iyad031.

Gerhart, J., & Kirschner, M. (2007). The theory of facilitated variation. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(Suppl 1), 8582–8589.

Gioti, A., Wigby, S., Wertheim, B., Schuster, E., Martinez, P., Pennington, C. J., Partridge, L., & Chapman, T. (2012). Sex peptide of *Drosophila melanogaster* males is a global regulator of reproductive processes in females. *Proceedings of the Biological Sciences*, *279*, 4423–4432.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nature Biotechnology*, *29*, 644–652.

Green, L., Coronado-Zamora, M., Radio, S., Rech, G. E., Salces-Ortiz, J., & Gonzalez, J. (2022). The genomic basis of copper tolerance in *Drosophila* is shaped by a complex interplay of regulatory and environmental factors. *BMC Biology*, *20*, 275.

Grindberg, R. V., Yee-Greenbaum, J. L., McConnell, M. J., Novotny, M., O'Shaughnessy, A. L., Lambert, G. M., Arauzo-Bravo, M. J., Lee, J., Fishman, M., Robbins, G. E., Lin, X., Venepally, P., Badger, J. H., Galbraith, D. W., Gage, F. H., & Lasken, R. S. (2013). RNA-sequencing from single nuclei. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 19802–19807.

Gupta, V., Frank, A. M., Matolka, N., & Lazzaro, B. P. (2021). Reliance on polyfunctional tissue leads to a reproduction-immunity tradeoff due to inherent constraint. *bioRxiv*.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., … Regev, A. (2013). De novo transcript sequence reconstruction from RNA-seq using the trinity platform for reference generation and analysis. *Nature Protocols*, *8*, 1494–1512.

Hanly, J. J., Wallbank, R. W. R., McMillan, W. O., & Jiggins, C. D. (2019). Conservation and flexibility in the gene regulatory landscape of heliconiine butterfly wings. *EvoDevo*, *10*, 15.

Hill, M. S., Vande Zande, P., & Wittkopp, P. J. (2021). Molecular and evolutionary processes generating variation in gene expression. *Nature Reviews. Genetics*, *22*, 203–215.

Horvath, V., Guirao-Rico, S., Salces-Ortiz, J., Rech, G. E., Green, L., Aprea, E., Rodeghiero, M., Anfora, G., & Gonzalez, J. (2023). Gene expression differences consistent with water loss reduction underlie desiccation tolerance of natural *Drosophila* populations. *BMC Biology*, *21*, 35.

Hsu, S. K., Jaksic, A. M., Nolte, V., Lirakis, M., Kofler, R., Barghi, N., Versace, E., & Schlotterer, C. (2020). Rapid sex-specific adaptation to high temperature in *Drosophila*. *eLife*, *9*, e53237.

Hu, Y., Comjean, A., Rodiger, J., Liu, Y., Gao, Y., Chung, V., Zirin, J., Perrimon, N., & Mohr, S. E. (2021). FlyRNAi.Org-the database of the *Drosophila* RNAi screening center and transgenic RNAi project: 2021 update. *Nucleic Acids Research*, *49*, D908–D915.

Hu, Y., Tattikota, S. G., Liu, Y., Comjean, A., Gao, Y., Forman, C., Kim, G., Rodiger, J., Papatheodorou, I., Dos Santos, G., Mohr, S. E., & Perrimon, N. (2021). DRscDB: A single-cell RNA-seq resource for data mining and data comparison across species. *Computational and Structural Biotechnology Journal*, *19*, 2018–2026.

Jaksic, A. M., Karner, J., Nolte, V., Hsu, S. K., Barghi, N., Mallard, F., Otte, K. A., Svecnjak, L., Senti, K. A., & Schlotterer, C. (2020). Neuronal function and dopamine signaling evolve at high temperature in *Drosophila*. *Molecular Biology and Evolution*, *37*, 2630–2640.

Johnson, B. R., Atallah, J., & Plachetzki, D. C. (2013). The importance of tissue specificity for RNA-seq: Highlighting the errors of composite structure extractions. *BMC Genomics*, *14*, 586.

Kankare, M., Parker, D. J., Merisalo, M., Salminen, T. S., & Hoikkala, A. (2016). Transcriptional differences between diapausing and non-diapausing *D. montana* females reared under the same photoperiod and temperature. *PLoS One*, *11*, e0161852.

Kapelnikov, A., Rivlin, P. K., Hoy, R. R., & Heifetz, Y. (2008). Tissue remodeling: A mating-induced differentiation program for the *Drosophila* oviduct. *BMC Developmental Biology*, *8*, 114.

Karaiskos, N., Wahle, P., Alles, J., Boltengagen, A., Ayoub, S., Kipar, C., Kocks, C., Rajewsky, N., & Zinzen, R. P. (2017). The *Drosophila* embryo at single-cell transcriptome resolution. *Science*, *358*, 194–199.

Koniger, A., & Grath, S. (2018). Transcriptome analysis reveals candidate genes for cold tolerance in *Drosophila ananassae*. *Genes (Basel)*, *9*, 624.

Kuhn, A., Kumar, A., Beilina, A., Dillman, A., Cookson, M. R., & Singleton, A. B. (2012). Cell population-specific expression analysis of human cerebellum. *BMC Genomics*, *13*, 610.

Lai, W. Y., & Schlotterer, C. (2022). Evolution of phenotypic variance in response to a novel hot environment. *Molecular Ecology*, *31*, 934–945.

Li, H., Janssens, J., De Waegeneer, M., Kolluru, S. S., Davie, K., Gardeux, V., Saelens, W., David, F. P. A., Brbic, M., Spanier, K., Leskovec, J., McLaughlin, C. N., Xie, Q., Jones, R. C., Brueckner, K., Shim, J., Tattikota, S. G., Schnorrer, F., Rust, K., … Zinzen, R. P. (2022). Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science*, *375*, eabk2432.

Lu, P., Nakorchevskiy, A., & Marcotte, E. M. (2003). Expression deconvolution: A reinterpretation of DNA microarray data reveals dynamic changes in cell populations. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 10370–10375.

Mallard, F., Nolte, V., & Schlotterer, C. (2020). The evolution of phenotypic plasticity in response to temperature stress. *Genome Biology and Evolution*, *12*, 2429–2440.

Mantica, F., Iniguez, L. P., Marquez, Y., Permanyer, J., Torres-Mendez, A., Cruz, J., Franch-Marro, X., Tulenko, F., Burguera, D., Bertrand, S., Doyle, T., Nouzova, M., Currie, P. D., Noriega, F. G., Escriva, H., Arnone, M. I., Albertin, C. B., Wotton, K. R., Almudi, I., … Irimia, M. (2024). Evolution of tissue-specific expression of ancestral genes across vertebrates and insects. *Nature Ecology and Evolution*, https://doi.org/10.1038/s41559-024-02398-5

Mattiazzi Usaj, M., Sahin, N., Friesen, H., Pons, C., Usaj, M., Masinas, M. P. D., Shuteriqi, E., Shkurin, A., Aloy, P., Morris, Q., Boone, C., & Andrews, B. J. (2020). Systematic genetics and single-cell imaging reveal widespread morphological pleiotropy and cell-to-cell variability. *Molecular Systems Biology*, *16*, e9243.

Mohammadi, S., Zuckerman, N., Goldsmith, A., & Grama, A. (2017). A critical survey of deconvolution methods for separating cell types in complex tissues. *Proceedings of the IEEE*, *105*, 340–366.

Montgomery, S. H., & Mank, J. E. (2016). Inferring regulatory change from gene expression: The confounding effects of tissue scaling. *Molecular Ecology*, *25*, 5114–5128.

Najle, S. R., Grau-Bove, X., Elek, A., Navarrete, C., Cianferoni, D., Chiva, C., Canas-Armenteros, D., Mallabiabarrena, A., Kamm, K., Sabido, E., Gruber-Vodicka, H., Schierwater, B., Serrano, L., & Sebe-Pedros, A. (2023). Stepwise emergence of the neuronal gene expression program in early animal evolution. *Cell*, *186*, 4676–4693.e29.

Nguyen, Q. H., Pervolarakis, N., Nee, K., & Kessenbrock, K. (2018). Experimental considerations for single-cell RNA sequencing approaches. *Frontiers in Cell and Development Biology*, *6*, 108.

Ohde, T., Morita, S., Shigenobu, S., Morita, J., Mizutani, T., Gotoh, H., Zinna, R. A., Nakata, M., Ito, Y., Wada, K., Kitano, Y., Yuzaki, K., Toga, K., Mase, M., Kadota, K., Rushe, J., Lavine, L. C., Emlen, D. J., & Niimi, T. (2018). Rhinoceros beetle horn development reveals deep parallels with dung beetles. *PLoS Genetics*, *14*, e1007651.

Oppenheim, S. J., Baker, R. H., Simon, S., & DeSalle, R. (2015). We can't all be supermodels: The value of comparative transcriptomics to the study of non-model insects. *Insect Molecular Biology*, *24*, 139–154.

Papatheodorou, I., Moreno, P., Manning, J., Fuentes, A. M., George, N., Fexova, S., Fonseca, N. A., Fullgrabe, A., Green, M., Huang, N., Huerta, L., Iqbal, H., Jianu, M., Mohammed, S., Zhao, L., Jarnuczak, A. F., Jupp, S., Marioni, J., Meyer, K., … Brazma, A. (2020). Expression Atlas update: From tissues to single cells. *Nucleic Acids Research*, *48*, D77–D83.

Parker, D. J., Envall, T., Ritchie, M. G., & Kankare, M. (2021). Sex-specific responses to cold in a very cold-tolerant, northern *Drosophila* species. *Heredity (Edinb)*, *126*, 695–705.

Pasquier, C., & Robichon, A. (2022). Temporal and sequential order of nonoverlapping gene networks unraveled in mated female *Drosophila*. *Life Science Alliance*, *5*, e202101119.

Pollen, A. A., Nowakowski, T. J., Shuga, J., Wang, X., Leyrat, A. A., Lui, J. H., Li, N., Szpankowski, L., Fowler, B., Chen, P., Ramalingam, N., Sun, G., Thu, M., Norris, M., Lebofsky, R., Toppani, D., Kemp, D. W., II, Wong, M., Clerkson, B., … West, J. A. A. (2014). Low-coverage single-cell mRNA sequencing reveals cellular heterogeneity and activated signaling pathways in developing cerebral cortex. *Nature Biotechnology*, *32*, 1053–1058.

Raghavan, V., Kraft, L., Mesny, F., & Rigerte, L. (2022). A simple guide to de novo transcriptome assembly and annotation. *Briefings in Bioinformatics*, *23*, bbab563.

Reimand, J., Isserlin, R., Voisin, V., Kucera, M., Tannus-Lopes, C., Rostamianfar, A., Wadi, L., Meyer, M., Wong, J., Xu, C., Merico, D., & Bader, G. D. (2019). Pathway enrichment analysis and visualization of omics data using g:Profiler, GSEA, Cytoscape and EnrichmentMap. *Nature Protocols*, *14*, 482–517.

Rodrigues, M. A., Merckelbach, A., Durmaz, E., Kerdaffrec, E., & Flatt, T. (2021). Transcriptomic evidence for a trade-off between germline proliferation and immunity in *Drosophila*. *Evolution Letters*, *5*, 644–656.

Rosenberg, A. B., Roco, C. M., Muscat, R. A., Kuchina, A., Sample, P., Yao, Z., Graybuck, L. T., Peeler, D. J., Mukherjee, S., Chen, W., Pun, S. H., Sellers, D. L., Tasic, B., & Seelig, G. (2018). Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science*, *360*, 176–182.

Salvador-Martinez, I., Coronado-Zamora, M., Castellano, D., Barbadilla, A., & Salazar-Ciudad, I. (2018). Mapping selection within *Drosophila melanogaster* embryo's anatomy. *Molecular Biology and Evolution*, *35*, 66–79.

Shapiro, M. D., Marks, M. E., Peichel, C. L., Blackman, B. K., Nereng, K. S., Jonsson, B., Schluter, D., & Kingsley, D. M. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, *428*, 717–723.

Stanford, B. C. M., Clake, D. J., Morris, M. R. J., & Rogers, S. M. (2020). The power and limitations of gene expression pathway analyses toward predicting population response to environmental stressors. *Evolutionary Applications*, *13*, 1166–1182.

Stark, R., Grzelak, M., & Hadfield, J. (2019). RNA sequencing: The teenage years. *Nature Reviews. Genetics*, *20*, 631–656.

Steiner, C. C., Weber, J. N., & Hoekstra, H. E. (2007). Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biology*, *5*, e219.

Suo, C., Dann, E., Goh, I., Jardine, L., Kleshchevnikov, V., Park, J. E., Botting, R. A., Stephenson, E., Engelbert, J., Tuong, Z. K., Polanski, K., Yayon, N., Xu, C., Suchanek, O., Elmentaite, R., Domínguez Conde, C., He, P., Pritchard, S., Miah, M., … Teichmann, S. A. (2022). Mapping the developing human immune system across organs. *Science*, *376*, eabo0510.

Svensson, V., Vento-Tormo, R., & Teichmann, S. A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. *Nature Protocols*, *13*, 599–604.

Tabula Sapiens, C., Jones, R. C., Karkanias, J., Krasnow, M. A., Pisco, A. O., Quake, S. R., Salzman, J., Yosef, N., Bulthaup, B., Brown, P., Harper, W., Hemenez, M., Ponnusamy, R., Salehi, A., Sanagavarapu, B. A., Spallino, E., Aaron, K. A., Concepcion, W., Gardner, J. M., … Wyss-Coray, T. (2022). The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans. *Science*, *376*, eabl4896.

Taylor, S. R., Santpere, G., Weinreb, A., Barrett, A., Reilly, M. B., Xu, C., Varol, E., Oikonomou, P., Glenwinkel, L., McWhirter, R., Poff, A., Basavaraju, M., Rafi, I., Yemini, E., Cook, S. J., Abrams, A., Vidal, B., Cros, C., Tavazoie, S., … Miller, D. M., III. (2021). Molecular topography of an entire nervous system. *Cell*, *184*, 4329–4347.e23.

Teets, N. M., Kawarasaki, Y., Lee, R. E., Jr., & Denlinger, D. L. (2013). Expression of genes involved in energy mobilization and osmoprotectant synthesis during thermal and dehydration stress in the Antarctic midge, Belgica antarctica. *Journal of Comparative Physiology. B, Biochemical, Systemic, and Environmental Physiology*, *183*, 189–201.

Thomas, P. D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L. P., & Mi, H. (2022). PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Science*, *31*, 8–22.

Torres-Oliva, M., Almudi, I., McGregor, A. P., & Posnien, N. (2016). A robust (re-)annotation approach to generate unbiased mapping references for RNA-seq-based analyses of differential expression across closely related species. *BMC Genomics*, *17*, 392.

Veltsos, P., Porcelli, D., Fang, Y., Cossins, A. R., Ritchie, M. G., & Snook, R. R. (2022). Experimental sexual selection reveals rapid evolutionary divergence in sex-specific transcriptomes and their interactions following mating. *Molecular Ecology*, *31*, 3374–3388.

Venet, D., Pecasse, F., Maenhaut, C., & Bersini, H. (2001). Separation of samples into their constituents using gene expression data. *Bioinformatics*, *17*(Suppl 1), S279–S287.

Vinuela, A., Varshney, A., van de Bunt, M., Prasad, R. B., Asplund, O., Bennett, A., Boehnke, M., Brown, A. A., Erdos, M. R., Fadista, J., Hansson, O., Hatem, G., Howald, C., Iyengar, A. K., Johnson, P., Krus, U., MacDonald, P. E., Mahajan, A., Manning Fox, J. E., …

McCarthy, M. I. (2020). Genetic variant effects on gene expression in human pancreatic islets and their implications for T2D. *Nature Communications*, *11*, 4912.

Wang, J., Sun, H., Jiang, M., Li, J., Zhang, P., Chen, H., Mei, Y., Fei, L., Lai, S., Han, X., Song, X., Xu, S., Chen, M., Ouyang, H., Zhang, D., Yuan, G. C., & Guo, G. (2021). Tracing cell-type evolution by cross-species comparison of cell atlases. *Cell Reports*, *34*, 108803.

Wang, X., He, Y., Zhang, Q., Ren, X., & Zhang, Z. (2021). Direct comparative analyses of 10X genomics chromium and smart-seq2. *Genomics, Proteomics & Bioinformatics*, *19*, 253–266.

Wee, J. L. Q., Murugesan, S. N., Wheat, C. W., & Monteiro, A. (2023). The genetic basis of wing spots in *Pieris canidia* butterflies. *BMC Genomics*, *24*, 169.

Wiegleb, G., Reinhardt, S., Dahl, A., & Posnien, N. (2022). Tissue dissociation for single-cell and single-nuclei RNA sequencing for low amounts of input material. *Frontiers in Zoology*, *19*, 27.

Winbush, A., Reed, D., Chang, P. L., Nuzhdin, S. V., Lyons, L. C., & Arbeitman, M. N. (2012). Identification of gene expression changes associated with long-term memory of courtship rejection in Drosophila Males.. *G3 (Bethesda)*, *2*, 1437–1445.

Xu, J., Li, Y., Li, H., Tarashansky, A. J., Kalicki, C. H., Hung, R.-J., Hu, Y., Comjean, A., Kolluru, S. S., Wang, B., Quake, S. R., Luo, L., McMahon, A. P., Dow, J. A. T., & Perrimon, N. (2021). A cell atlas of the fly kidney. *bioRxiv*.

Yapici, N., Kim, Y. J., Ribeiro, C., & Dickson, B. J. (2008). A receptor that mediates the post-mating switch in *Drosophila* reproductive behaviour. *Nature*, *451*, 33–37.

Yim, A. K. Y., Wang, P. L., Bermingham, J. R., Jr., Hackett, A., Strickland, A., Miller, T. M., Ly, C., Mitra, R. D., & Milbrandt, J. (2022). Disentangling glial diversity in peripheral nerves at single-nuclei resolution. *Nature Neuroscience*, *25*, 238–251.

Zaitsev, K., Bambouskova, M., Swain, A., & Artyomov, M. N. (2019). Complete deconvolution of cellular mixtures based on linearity of transcriptional signatures. *Nature Communications*, *10*, 2209.

Zhang, M. J., Ntranos, V., & Tse, D. (2020). Determining sequencing depth in a single-cell RNA-seq experiment. *Nature Communications*, *11*, 774.

Ziegenhain, C., Vieth, B., Parekh, S., Reinius, B., Guillaumet-Adkins, A., Smets, M., Leonhardt, H., Heyn, H., Hellmann, I., & Enard, W. (2017). Comparative analysis of single-cell RNA sequencing methods. *Molecular Cell*, *65*, 631–643.e4.

How to cite this article: Hoedjes, K. M., Grath, S., Posnien, N., Ritchie, M. G., Schlötterer, C., Abbott, J. K., Almudi, I., Coronado-Zamora, M., Durmaz Mitchell, E., Flatt, T., Fricke, C., Glaser-Schmitt, A., González, J., Holman, L., Kankare, M., Lenhart, B., Orengo, D. J., Snook, R. R., Yılmaz, V. M., & Yusuf, L. (2024). From whole bodies to single cells: A guide to transcriptomic approaches for ecology and evolutionary biology. *Molecular Ecology*, *00*, e17382. https://doi.org/10.1111/mec.17382