

A cell–cell atlas approach for understanding symbiotic interactions between microbes

Elisabet Alacid and Thomas A Richards



Natural environments are composed of a huge diversity of microorganisms interacting with each other to form complex functional networks. Our understanding of the operative nature of host–symbiont associations is limited because propagating such associations in a laboratory is challenging. The advent of single-cell technologies applied to, for example, animal cells and apicomplexan parasites has revolutionized our understanding of development and disease. Such cell atlas approaches generate maps of cell-specific processes and variations within cellular populations. These methods can now be combined with cellular-imaging so that interaction stage versus transcriptome state can be quantized for microbe–microbe interactions. We predict that the combination of these methods applied to the study of symbioses will transform our understanding of many ecological interactions, including those sampled directly from natural environments.

Address

Department of Zoology, University of Oxford, 11a Mansfield Road, Oxford OX1 3SZ, UK

Corresponding authors:

Alacid, Elisabet (elisabet.alacidfernandez@zoo.ox.ac.uk),
Richards, Thomas A (thomas.richards@zoo.ox.ac.uk)

Current Opinion in Microbiology 2021, **64**:47–59

This review comes from a themed issue on **Environmental Microbiology**

Edited by **Paul Hoskisson** and **John Bruce**

For complete overview of the section, please refer to the article collection, “**Environmental Microbiology**”

Available online 13th October 2021

<https://doi.org/10.1016/j.mib.2021.09.001>

1369-5274/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Introduction

Ecosystems are composed of a vast diversity of microbes performing critical functions [1,2]. Yet, microorganisms do not function in isolation and depend on interactions involving metabolite exchange and trophic cascades [3]. Such interactions are additive, resulting in complex networks. An elegant demonstration of the complexity of such systems can be seen in planktonic food webs [4]. Heterotrophic/mixotrophic microorganisms connect primary production with higher-trophic and lower-trophic levels through predation and symbioses, i) channeling nutrients into food webs, ii) altering host population

dynamics and iii) driving species succession [4,5]. Predation — or grazing — was considered the dominant heterotrophic interaction in aquatic food webs. Such a view has been amended due to a growing appreciation for a diversity of interactions (e.g. symbioses) [4,6,7]. Many of these interactions remain cryptic with culture-based experimentation falling-down in the attempt to recreate naturally occurring multi-faceted, temporally distinct and metabolically primed interactions. As a result, many species are difficult to culture because they depend on such interactions. The aim of this article is to identify new opportunities through integrating state-of-the-art technologies to the study of cellular interactions between microbes both in enigmatic and cultured symbioses and those sampled directly from natural environments.

Critical limitations in our understanding of microbial interactions

Symbiotic interactions encompass parasitism, mutualism and commensalism and often depend on cellular contact. Such interactions are dynamic in nature, with the status of the interaction determined by changing functions at the cellular level driven by transcriptional responses to the interaction. Our knowledge about microbial interactions is limited due to: i) a historical focus on specific multicellular organisms (e.g. humans, livestock, and agricultural plants), and ii) technical limitations that mean systematic sampling is difficult [8,9]. This difficulty arises because the interaction can be hidden inside a host and can also encompass a range of life-cycle stages which are difficult to recover or recapitulate.

The application of high-throughput diversity tag sequencing of environmental samples has revealed phylogenetic groups putatively assigned as symbionts are diverse [6,7] and present in many protistan taxonomic groups. Despite the potential relevance of such symbiotic relationships, most studies have not made progress beyond charting phylogenetic diversity. In the majority of cases the ecology of these sequence-identified ‘symbionts’ is inferred and therefore requires validation [4,10]. In an attempt to catalogue known associations, Bjorbaekmo *et al.* [11^{*}] have generated a Protist Interaction Database (PIDA) using bibliographic survey, inferring parasitism is underrepresented. Specifically, parasitism accounts for 18% of the database entries, while predation (39%) or alternative symbiotic relationships (43%) (mostly photosymbioses) dominate. Understanding commonality and variance between types of interaction,

as well as the mechanism and nature of metabolic interaction (e.g. which metabolites are taken from host, which metabolites are discarded) is lacking. Such knowledge is critical for understanding the role such interactions play in ecological processes [3].

In order to better understand the function and physiology of aquatic microorganisms, bulk community sequencing approaches have been applied (i.e. meta-transcriptomic [12[•]]). These approaches rely on genome database sampling to assign putative taxonomic and functional identity to each sequence. Such databases are biased towards free-living and photosynthetic organisms or terrestrial parasites of medical, veterinary and agricultural importance [13] with only a few genomes and transcriptomes from host-symbiont associations available [9,14–18]. Furthermore, analysis based on such data is biased towards the recognition of metabolic pathways and proteomic components of previously characterised systems. Such studies provide a start, allowing some taxonomic identification and functional annotation, but do not provide a systematic analysis of the basis of interactions. New solutions are required.

Another limitation of bulk approaches is that they encompass ‘ensemble’ samples including both colonized-host, uncolonized-host and host-free symbiont cells, so the sampling of interaction processes are blurred. This problem is compounded by the fact that most interactions are likely to be characterised by a diversity of developmental stage-specific functions. Attempts have been made to obtain lab-synchronized cell populations, and although a certain degree of ‘staging’ is achievable [14,19], a perfectly synchronized culture is intractable; cell heterogeneity and rare cell states are missed. These limitations are manifold when targeting interactions from natural samples, which are composed of millions of cells from different taxa displaying different developmental stage-specific phenotypes, and then transcriptome/function is averaged. Critical processes associated with interaction phenotypes will be diluted below the limit of detection by the wider signals within ensemble samples, making it impossible to identify transcriptional response associate with a specific phase of an interaction. Therefore, bulk measurements miss: i) the diversity of developmental processes intrinsic to each interaction stage, and ii) the phenotypic diversity within each interaction stage. If we are going to understand the molecular and cellular processes governing symbiotic-cell associations, determining the specific circumstance of each cellular entity within a population is vital. Again, new solutions are required.

A way forward using the cell atlas approach

It is now possible to capture, separate, and functionally profile single cells [20] moving away from ensemble

community analysis towards studies of the individual. This work is generating a shift in the field of genomics, developmental biology and associated disciplines. Although these methods were developed in medical-research-related sciences, Ku and Seb -Pedr s [21] have highlighted how such methods can be applied to microbial eukaryotes. Pioneering studies have revealed the existence of microbial interactions between single-celled organisms and provide some genomic insights into cryptic-groups [22,23]. In the past, such approaches were limited by technological constraints such as low input DNA and RNA material, restricted sample replication, and high rates of sequencing errors/noise (e.g. non-uniformity in coverage, sparse data, false-positive errors, and amplification biases). They also failed to assign a cell sample to a cellular image, preventing the identification of a developmental phase or a specific biological-moment in an interaction. In order to increase resolution to identify relevant cellular processes from intrinsic noise, technologies have moved towards an increase in sample events [24]. These technologies are now revealing the intricacies of cellular function at increasing temporal and spatial resolution, and capturing the heterogeneity present within populations [25].

Cell atlas initiatives have arisen from the potential of single-cell technologies [26] and are being applied to multicellular organisms [27] with the aim of describing the transcriptional characteristics of cell-type identities within a continuum of possible states encompassing both development process and tissue types (i.e. cellular map/atlas). Thus, cell atlases have revolutionized our understanding of tissue development and disease [28,29]. For parasitic protists, cell atlas approaches have been used for the Apicomplexa animal parasites *Plasmodium* [30,31^{••}] and *Toxoplasma* [32^{••}], allowing the identification of molecular phenotypes associated with life-cycle stages. In aquatic organisms, atlas approaches have been applied to deeply derived branches of the metazoan radiation, gaining insights into the evolutionary origins of animal cell types and genome-regulatory mechanisms [33–37]. To date, only a single beautiful study exists addressing a microbial aquatic interaction between a virus and its microalgal host [38^{••}].

Such methods have the potential to deconvolve host-symbiont interactions and quantitatively analyze the gene expression signatures of complex symbiont life-cycles. Currently, cell atlas approaches are only based on gene expression (the transcriptome). Yet, such approaches could be expanded if mRNA profiling can be combined with live-cell imaging, allowing a description of cell morphological state and interaction characteristics tied to a transcriptome, so that the progression of a cellular interaction versus transcriptome state can be ordered into a continuum and then quantized to

developmental checkpoints. As new technologies are developed such approaches can be further amended to include small-RNA, metabolomic and proteomic profiling [39]. Thus, the combination of several approaches to generate single-cell atlases of protist host–symbiont interactions will advance our understanding of the molecular and cellular biology of these understudied groups and will allow comparative biology of transcriptional programs of different types of symbiotic relationship. Such studies will allow us to reconstruct the evolution of symbiotic systems in protist lineages. In the remaining sections of this article we discuss the application of cell atlas technologies to study cell–cell host–symbiont interactions.

The potential of cell atlas technologies for the study of natural microbial interactions

Cells, the basic units of biological structure and function, vary broadly in type and state. This is also evident for symbiotic associations. We must therefore separate such cell states to capture differences in function associated to interaction state.

Single-cell isolation

The first step is to identify, capture and isolate cells from a heterogeneous population (e.g. asynchronous host–symbiont cultures or microbes from environmental samples). The sampling step needs to occur in a short time period in order to capture the native state of each interacting player. Changes in the host induced by symbiont and transitions between life-cycle stages during infection can happen quickly [19,30,31^{••}]. Microfluidic platforms (droplet-based [40[•]] or nano/microwell-based [41]) and fluorescence-activated cell sorting (FACS) [42] can be used to separate cells. These platforms can sort thousands of cells in a few hours. Such techniques are efficient when targeting abundant cells (i.e. from cultures or bloom-forming natural populations). However, without modification such approaches can be limited for sampling rare microbes. Another potential problem of single-cell isolation that should be considered is the differential sensitivity of organisms or interaction stages to the different cell-separation methods. For instance, some protists cells, especially those with delicate cellular appendages and which lack cell walls often cannot survive a FACS-sorting method, causing cell lysis before imaging. Furthermore, the cell isolation procedure itself can cause stress to sensitive cell types affecting their gene expression profiles in comparison to their native state. A solution to these problems is the fixation of cells before isolation, preventing early cell lysis and changes in gene transcription due to stress. A recent method based on acetic acid, methanol and glycerol (ACME) exposure has been developed to dissociate and fix animal cells from different lineages and different environments (including freshwater and marine habitats) for single-cell

transcriptome analysis [43]. This protocol produces fixed single cells in suspension, with high RNA integrity and preserves morphology allowing imaging of intact cells. This method is also compatible with cell staining and FACS, droplet-based approaches and combinatorial methods for single-cell transcriptomics (discussed below). Furthermore, it allows cryopreservation of samples, broadening the experimental setup that can be performed, including preservation of organisms collected during field sampling. We foresee value in the adaptation of this method for the study of cell–cell interactions in host–symbiont systems sensitive to live cell isolation or for those unculturable associations where the only possibility of capture is direct sampling from nature.

Microfluidic approaches

Droplet-based approaches, where single cells are encapsulated in *aqua* droplets within an oil emulsion, have been used for profiling thousands of *Plasmodium* cells [30]. Although droplet-based approaches are less costly, we foresee important limitations when applied to aquatic samples. First, this method does not allow the selection or discrimination of specific cells from heterogeneous populations. This method therefore has an inherent low detection/recovery efficiency. Second, it cannot distinguish between cells and non-cellular particles (which can be abundant in environmental samples). Third, it produces a high rate of multiple event recoveries, which in the case of cell–cell interactions can generate false positives for organismal interactions, invalidating an experiment. Although the number of doublets can be reduced by dilution, this process can increase the number of empty droplets. Very recently, ‘single-cell combinatorial fluidic indexing’ (scifi) has been developed to overcome this limitation. The method combines combinatorial pre-indexing of entire cellular transcriptomes with subsequent scRNA-seq using microfluidics, allowing resolved sampling of individual cellular transcriptomes from droplets containing multiple cells [40[•]]. Similarly, nano-based and microwell-based methods, where the fluid is driven through a flow cell by laminar flow leading to cells sinking into a microwell array by gravity, cannot select for cells based on phenotype information. However, microchips containing wells can be inspected and imaged to identify empty, singlet or multiplet wells, and also to discriminate living cells from dead or non-cellular particles [41].

Fluorescence-activated cell sorting (FACS) approaches

FACS methodologies can be used to sort cells based on cellular properties, such as presence/absence of pigments which, for example, allows the separation of phototrophs and heterotrophs [44]. FACS can also be used in combination with dyes to distinguish between cells versus non-living cells, senescent cells or particles. The use of some

dyes can also allow for targeting of specific cells or cell–cell interactions. Fluorescent labelling of intracellular acidic vacuoles has been used to enrich sampling for phagotrophs and therefore grazers [45], while tubulin stains have allowed for the enriched sampling of flagellated cells, a trait common in both grazers and parasites [46]. The use of FACS in combination with cellular dyes has allowed the identification of infected versus non-infected host cells and the infection stages of both *Plasmodium* and *Toxoplasma* [30,32^{••},47] (discussed further below). Another advantage of FACS over microfluidic approaches is the reduced rate of multi-sample recovery as it can be adapted to strictly select for single cells.

The main challenge when targeting host–symbiont interactions from nature is their low abundance in many environments. A recent study demonstrated that the single-cell recovery rate of targeted environmental microbial eukaryotes and the subsequent quality of the genome sequences recovered increased when a FACS enrichment step was performed before single-cell isolation [48]. The enrichment consists of a primary step where the target population is bulk sorted into sterile source media followed by a second round of FACS to capture single cells into microwell plates.

FACS-based approaches currently demonstrate a higher degree of accuracy for sampling targeted cell–cell interactions. Nevertheless, such approaches are continuously being improved, for example, with the development of image-based single-cell sorting automation using droplet microfluidics [49^{*}] or the combination of Printed Droplet Microfluidics (PDM) linked to optical and single-cell RNA sequencing (scRNA-seq) in nanoplates [50^{••}]. One additional exciting prospect, for both these approaches, is that they can be linked to high-throughput imaging (combining brightfield and fluorescence) close to the point of RNA fixation.

Single-cell imaging

Cell imaging just before scRNA-seq can be extremely valuable and is a key step for the study of microbial interactions. Characterising the morphology of the microbial host and its symbiont can allow assignment of a transcriptional profile to the progression of a symbiosis. Links between morphological/taxonomical information of single cells and their genomes have previously been attempted using light microscopy of manually isolated single-cells [51,52]. This approach has limited scalability and produces images of relatively low resolution. With the advent of high-throughput imaging (HTI) it is now possible to automatically obtain high-quality images of single cells and to extract numerical data for cellular features [53]. The challenge of such approaches is to balance the volume of image data (accuracy) and image-processing speed (response time).

When linking cell morphotypes to their expression profiles, we need technologies that will allow the simultaneous capture of intact single cells, high-throughput imaging and then RNA preservation. High-throughput ‘intelligent image-activated cell sorting’ (IACS) [54^{••}] allows the acquisition of multidimensional images, enabling discriminated sorting of cells or cell clusters using high-content images. The images obtained during the sorting can then be extracted and further processed for morphotype classification. This technique has been applied to the cell sorting of different microalgal species from cultures and ranging from 3 μm to 30 μm in diameter. Its performance for investigation of environmental samples or its efficiency at discrimination of finer-scale characteristics has yet to be validated but could prove to be a very useful approach.

An established HTI approach for multi-parametric data acquisition at the single-cell level is high-content screening (HCS). This technology consists of a set of analytical methods based on automated microscopy (including brightfield and fluorescence), image processing and visualization tools to extract quantitative data [55]. HCS can be combined with deep learning methods to automatically classify high-resolution images, allowing the sorting of single cells into categories (morphotypes) based on cellular features [56]. An additional advantage of this technology is it can be conducted post-sorting and close to RNA fixation. Recently, it has been demonstrated that it is possible to reconstruct fine-scale developmental trajectories from HCS images by combining deep-learning methods, which encompass feature extraction, and make use of a nonlinear dimensionality-reduction technique (uniform manifold approximation and projection, UMAP) for manifold learning [57^{••}]. In this study, Rappez *et al.* published a deep neural network that allowed the reconstruction of a continuous cell-cycle trajectory solely based on high-resolution images. Such strategies are likely to underpin further development of how we can interpret imaging for cross comparison with ‘omics data.

During the course of symbiotic interactions, the symbiont and the host experience cellular and physiological changes associated to the diversity of phenotypes produced by the developing interaction. In many cases these can be measured by changes in cellular morphology. Light and fluorescence microscopy have been used in several studies to follow the infection process in aquatic protistan cell–cell associations. Changes in host and symbiont developmental stages, produced by both intracellular and extracellular symbionts, can be tracked by changes in the intensity of autofluorescence and fluorescent dyes that stain cellular organelles, cell wall components or specific molecular targets [5,14,58–60]. Thus, the application of deep learning and UMAP from HCS images obtained from single-celled symbiotic associations can be used to reconstruct the interaction trajectory of

host–symbiont interactions from complex communities. In this review we do not cover the data-analysis strategies and existing methods for image-based cell phenotype/morphotype profiling. For a step-by-step computational workflow for image-based cell profiling, including different methods for each step, the pros and cons of different methodological approaches, and the applicability of these methods to different experimental types and scales we refer the reader to an article by Caicedo *et al.* [61] and the references therein. For a more detailed information on deep-learning methods applied to recognition, profiling, and prediction of image-based cell phenotyping we refer the readers to an article by Pratapa *et al.* [62].

Single-cell RNA sequencing

After single-cell isolation the stock approach to analysis is single-cell RNA sequencing. Diverse scRNA-seq protocols exist based on the transcript coverage (full-length or partial 5' or 3' ends), transcript quantification, and if the method allows for strand-specific sequencing [63]. When reference genomes or transcriptomes are available, or can be obtained from cultures, partial sequencing methods, such as MARS-seq [25] or Drop-seq [64] could be viable. These methods allow accurate gene expression estimation by using 'unique molecular identifiers' (UMIs) [65]. UMIs are random short sequences used to barcode the different individual transcripts allowing correction for amplification bias. UMIs provide absolute counts of each transcript template and allow comparisons between cells with minimized normalization. When no genomic information is available (for instance, when sampling uncultured host–symbiont associations from aquatic environments), full-length transcript sequencing may be preferable, such as the Smart-seq3 protocol [66], which also allows for the use of UMIs. These full-length mRNA sequencing methods provide information on the internal sequence and structure of the RNAs including, for example, alternative splicing and identification of the relative abundance of isoforms [63]. They do not, however, provide strand-specific information and give lower throughput.

A novel method called 'single-cell amplification and sequencing of full-length RNAs by Nanopore platform' (SCAN-seq) using the smartseq2 protocol has been tested in mouse embryonic cells and oocytes [67]. This work has demonstrated recovery of full-length transcripts, specifying the template strand for transcription of a specific RNA molecule, but did not allow for template quantification using UMIs. New protocols currently enabling the use of UMIs with long-read sequencing technologies, such as Oxford Nanopore and PacBio, are being developed [68]. Such technologies provide potential for the study of poorly sampled host–symbiont interactions in a quantitative and accurate manner, while also identifying alternative splicing, when genomic information is lacking. However, the key limitation of these approaches for single-cell

methods is the requirement for substantial quantities of starting materials with high-purity.

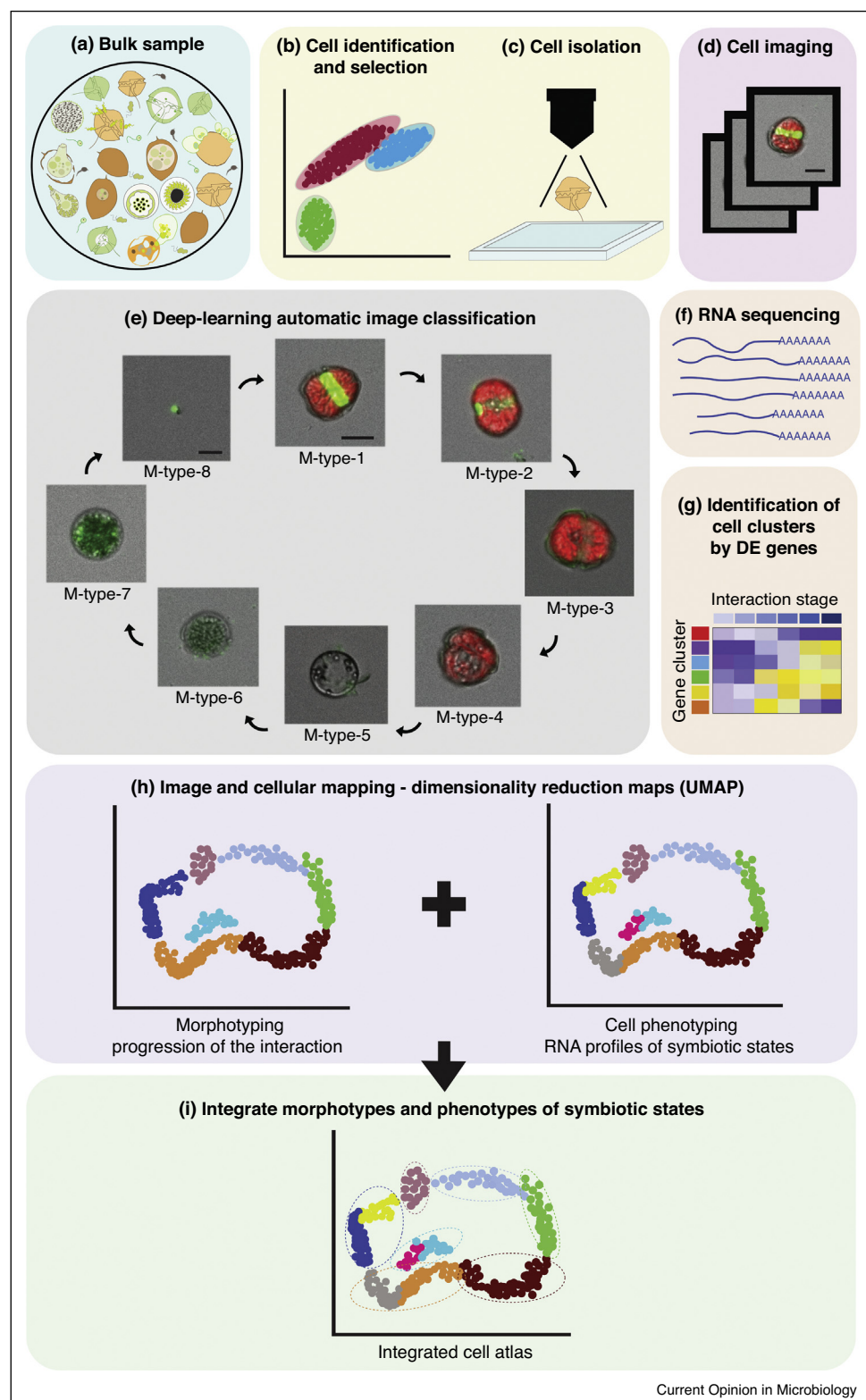
Examples of the cell atlas approach for the study of protist interactions

The combination of FACS and droplet-based approaches for single-cell isolation, combined with Smart-seq2 and Drop-seq protocols, have been applied to generate transcriptomes for numerous life-cycle stages of *Plasmodium* (i.e. the Malaria Cell Atlas [30]). Howick *et al.* profiled 1787 parasites using FACS in combination with Smart-seq2, generating full-length mRNA transcriptomes at 10 time points, covering all life-cycle stages across both infections of mosquitos and the mammalian host. Fine-scale transcriptional patterns of development, including marker genes associated with each parasite stage, cellular strategy (replicative, growth, and sexual phase), and host environment were identified. The authors also sequenced 15,858 cells of three species of Malaria parasite using droplet sequencing. The combination of the two methods allowed the developmental-infection trajectories across species during the pathogenic phase of the malarial life-cycle. A single-cell atlas of the most virulent agent of human malaria has also been published recently, revealing the gene usage across the transmission cycle through the mosquito and elucidating the developmental trajectories and transcriptional signatures of the transmissible stages [31**]. These studies provide useful information for drug and vaccine development identifying pathways which can be specifically targeted for processes involved in parasite transmission.

Another example of the single-cell atlas approach for a protist parasite is that of *Toxoplasma* during the course of *in vitro* asexual development [32**]. Xue *et al.* isolated >5400 *Toxoplasma* cells in both tachyzoite and bradyzoite stages using FACS, and performed Smart-seq2 sequencing to reconstruct transcriptional dynamics of asynchronous *Toxoplasma* from cultures. These data identified hidden cell states and rare parasite forms with highly unusual patterns of gene expression, which were shown to be associated to specific transcription factors. The fine-scale transcriptomic profiling of single cells during the cell-cycle and asexual development revealed that the antigenic repertoire of this parasite is more heterogeneous than previously appreciated, a key factor for vaccine and therapy design.

The use of FACS to sort cells from a virus-microalga system across the infection-cycle, coupled with MARS-seq2.0 (which included the use of UMIs), allowed the identification of the transcriptomic profile of 5000 cells during the infection process [38**]. The authors reconstructed the viral transcriptional trajectory, revealing viral-genetic-programs composed of genes with distinct promoter elements that orchestrate sequential expression. These 'viral-genetic-programs' were defined as

Figure 1



transcriptional kinetic classes specific to infection stage that each temporally regulate the progression of the viral infection. The work also demonstrates heterogeneity in viral transcript patterns among individual cells allowing a comprehensive understanding of the temporal dynamics of viral infection of an important aquatic microalga.

Conclusion: an integrated cell-cell atlas approach to study host-symbiont interactions

We argue that in order to benefit from the advantages of high-throughput single-cell technologies and accompanying computational methodologies, while addressing some of the challenges outlined above, an integrated cell-cell atlas approach represents a useful way forward. Figure 1 outlines a model of how this could work. This approach combines the latest advances in single-cell isolation, imaging and RNA sequencing, together with state-of-the-art deep learning methods for the purpose of classifying interaction-stage. The proposed cell atlas approach (Figure 1) will allow researchers to quantize the progression of a cellular interaction versus the transcriptome state. Many symbiotic interactions, especially parasitic outbreaks, occur together with the seasonal peaks of their hosts (i.e. blooms) [58,69]. Therefore, to study host-symbiont interactions it is important to collect environmental samples when the activity and abundance of these interactions are at a peak. We have listed some candidate microbial interactions for which model interaction-systems could be developed using available culture collections, or alternatively are accessible as

environmental samples from consistent natural bloom events (Table 1). Analyses of many of these interactions would establish the methodology, help to deconvolve cell-cell interactions occurring in highly complex systems (e.g. aquatic environments) and in many cases greatly aid both the analysis of environmental meta- 'omic data and parasite surveillance efforts. Ultimately, we predict the development of an integrated cell-cell atlas approach to microbial host-symbiont systems will transform the way we study interactions, how we approach comparative biological analyses of parasitic/mutualistic/commensal associations within and between systems, and help us understand the cellular systems that determine critical interactions shaping complex environments.

Conflict of interest statement

Nothing declared.

Acknowledgements

We would like to thank David Milner and Ben Jenkins for their help during image acquisition and Josephine Blagrove for her help in processing the high-resolution single-cell images. We are also very grateful to David Green and Cecilia Rad Menendez from the Culture Collection of Algae and Protozoa (CCAP) and Jamie McGowan (Earlham Institute) for their help in the identification of candidate host-symbiont model interaction-systems. This work is supported by The Royal Society (UK) under the Newton International Fellowship (grant number NF170346, call 2017) awarded to E. A. and T.A.R. T.A.R. is supported by a Royal Society University Research Fellowship (URF/R/191005) and European Research Council Consolidator Grant (CELL-in-CELL: 819507).

(Figure 1 Legend) An integrated cell-cell atlas approach to study host-symbiont microbial interactions.

Overview: **(a)** Environmental sample or asynchronized culture containing diverse stages of symbiosis and symbiont-free organisms. **(b)** Identification and selection of targeted symbionts using fluorescent cytometry using a combination of cell-trait-dyes (for examples, see Table 1). **(c)** Isolation of individual cells into plates using single-cell platforms (e.g. FACS), allowing identification, selection of low abundant cell types from mixed samples based on phenotype characteristics. Such approaches allow the collection of interacting cells in different stages from the environment or mixed cultures without the need for synchronization. **(d)** Flat-bottomed microwell plates containing sorted single-cells are imaged with a high-resolution system. The image shows a non-infected *Alexandrium minutum* (Dinoflagellata) cell taken with a High-Content Screening system (ImageXpress Pico Automated Cell Imaging System, molecular Devices S.L., US) using transmission light and fluorescent microscopy. Chlorophyll of the chloroplasts are visible in red colour due to autofluorescence of chlorophyll. Nucleus is visible in green color due to fluorescence (Sybr Green I – 5X concentration stain). Scale bar = 10 μ m. **(e)** High-resolution images are automatically classified into categories or morphotypes (M-types) using deep learning. Images used correspond to single cells isolated from an asynchronized co-culture in the laboratory during the infection cycle of the protist parasite *Parvilucifera sinerae* (Perkinsea) infecting the toxic bloom-forming dinoflagellate host *A. minutum*. Images were taken with the same instrument and same conditions and stained as explained in (d). The interaction cycle can be classified in different morphotypes: M-type-1, non-infected *A. minutum* (host) cell; M-type-2, a parasite zoospore (free-living swimming infective stage) penetrates the host cell and starts to consume the nucleus of the host causing a reduction of the green fluorescence; M-type-3, the parasite has completely consumed the host nucleus, which can be visualized by the lack of green fluorescence; M-type-4, the parasite grows inside the host while consuming the host contents including the chloroplasts, which translates in a reduction of the red fluorescence in the area where the parasite is located (round clear body); M-type-5, the parasite has completely consumed the host, visualized as no sign of red fluorescence; M-type-6, the parasite starts reproduction becoming a multinucleated stage (green fluorescence); M-type-7, the parasite fills its structure with zoospores, each with a nucleus (green fluorescence); M-type-8, the zoospores are released from the parasite structure and freely swim to find a new host to infect, restarting the cycle. They can be identified by their small size and the green nuclear fluorescence. Scale bar in M-type-1 is 10 μ m and applies for all images except M-type-8, which bar represents 5 μ m. The classification algorithm can provide a confidence score for each image classified, allowing only high-confidence classifications to be selected for downstream analyses. The number of cells for single-cell sequencing of each interaction state could be pre-selected, in order to equally cover the different phases of the interaction. **(f)** Once the microwell-plate has been imaged and the cells/wells have been fixed and selected, single-cell RNA sequencing is applied to obtain the transcriptomic profiles of individual cells. The sequencing protocol chosen would depend on the research question, biological hypothesis to be addressed and the genomic information available for the system. **(g)** Single-cell RNA data analyses for the identification of cells clusters based on differential expressed (DE) genes. **(h)** Image and cellular mapping using dimensionality reduction maps to track the physical progression and transcriptional dynamics across the interaction-cycle. **(i)** Integration of cellular maps allows for a quantized understanding of a cell-cell interaction. All of this progress is possible using environmental samples given efficient targeting without needing to raise a culture of a complex interaction.

Table 1

List of candidate microbial interactions for which model interaction-systems could be developed using available culture collections

Host-Symbiont phylum	Species model system (Host-Symbiont)	Type of symbiosis	Identification of interaction stages	Culture available (Host/Symbiont)	Genomic information available		Type of environment	Peak of interaction in nature	References
					Symbiont	Host			
Ciliophora — Chlorophyta	<i>Paramecium bursaria</i> — <i>Micractinium conductrix</i> / <i>Chlorella variabilis</i>	Facultative photosymbiosis	Non-digested <i>Chlorella</i> — red autofluorescence; Digested algae — reduction in red autofluorescence; Host digestive vacuoles (DV) stage-II (acidic) — LysoSensor yellow/blue DND-160 fluorescence/ Gomori staining (light microscopy); DV stage-III (non-acidic) — LysoSensor negative (lack of fluorescence), morphology (light microscopy); DV budding- Perialgal vacuole of SGC (single green <i>Chlorella</i>) — LysoSensor negative, Gomori negative, red autofluorescence; nuclei — SYBR green I	Yes/Yes; Culture collection of Algae and Protozoa, (CCAP), Scotland, United Kingdom	Genome (<i>C. variabilis</i> , https://phycocosm.jgi.doe.gov/ChlNC64A_1/ChlNC64A_1.home.html ; <i>M. conductrix</i> , https://www.ncbi.nlm.nih.gov/bioproject/PRJNA290385) and transcriptome	Genome (ParameciumDB, https://paramecium.i2bc.paris-saclay.fr/) and transcriptome	Freshwater	During periods of nutrient limitation and water-column stratification/ in nature host is not found without symbiont	[17,70]
Dinoflagellata — Ciliophora — Cryptophyta	<i>Dinophysis acuminata</i> — <i>Myrionecta rubra</i> — <i>Geminigera cryophila</i>	Kleptoplasty (acquired phototrophy-reduced symbionts)	Morphological changes — light microscopy; Kleptoplastids — red/orange autofluorescence; nucleus of host and symbiont — SYBR green I	Yes/Yes/Yes; <i>Dinophysis</i> (e.g. DAEP01, Woods Hole Oceanographic Institution, MA, USA); <i>G. cryophila</i> and <i>M. rubra</i> (National Center for Marine Microalgae and Microbiota (NCMA), Bigelow, ME, USA)	<i>G. cryophila</i> transcriptome and genome (JGI Project Id: 1201759); <i>M. rubra</i> transcriptome	<i>D. acuminata</i> transcriptome	Marine	Seasonal blooms of <i>M. rubrum</i> and <i>D. acuminata</i>	[9,16,18]

Table 1 (Continued)

Host–Symbiont phylum	Species model system (Host–Symbiont)	Type of symbiosis	Identification of interaction stages	Culture available (Host/Symbiont)	Genomic information available		Type of environment	Peak of interaction in nature	References
					Symbiont	Host			
Dinoflagellata — Dinoflagellata	<i>Alexandrium catenella</i> — <i>Amoebophrya ceratii</i>	Parasitism (Endobiotic)	Host chloroplasts — red autofluorescence; host theca (cellulose) — calcofluor white; nucleus — DAPI stain/SYBR green I. <i>Amoebophrya</i> early stages — green autofluorescence; late stages (vermiform) — DAPI/propidium iodide (nuclei)	Yes/Yes; Roscoff Culture Collection (RCC) Station Biologique de Roscoff, Roscoff, France	Genome and transcriptome	Transcriptome	Marine	Seasonal blooms	[9,15,19,58]
Dinoflagellata — Perkinsea	<i>Alexandrium minutum</i> — <i>Parvilucifera infectans</i>	Parasitism (Endobiotic)	Host chloroplasts — autofluorescence; host theca — calcofluor white; nucleus — DAPI/ SYBR green I; early stages of infection — reduction in red fluorescence inside the host and reduction of blue/green fluorescence (DAPI/ SYBR green I); late stages of infection (sporangium) — SYBR green I/DAPI/ propidium iodide	Yes/Yes; Roscoff Culture Collection (RCC), Station Biologique de Roscoff, Roscoff, France	No	Transcriptome	Marine	Seasonal blooms	[9,59,69]
Ochrophyta — Chytridiomycota	<i>Asterionella formosa</i> — <i>Zygorhizidium affluens</i>	Parasitism (Epibiotic)	Host chloroplasts — red autofluorescence; Infected host — reduction in red autofluorescence. Chytrid cysts — calcofluor white, Wheat Germ Agglutinin-Texas red/FITC; chytrid and host nuclei — DAPI/ SYBR green I	Yes/Yes; Culture collection of Algae and Protozoa, CCAP, Scotland, United Kingdom	No	Genome (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA247668)	Freshwater	Seasonal blooms	[60]
Xanthophyta — Aphelida	<i>Tribonema gayanum</i> — <i>Paraphelidium tribonemae</i>	Parasitism (Endobiotic)	<i>Tribonema</i> cell wall (cellulose) — calcofluor white; host chloroplasts — red autofluorescence; infected host — reduction in red autofluorescence. Parasite chitin-bearing infective cysts — Wheat Germ Agglutinin-Texas red/FITC; parasite and host nuclei — DAPI/ SYBR green I	Yes/Yes; Culture collection of parasitic protists at the Zoological Institute RAS (CCPP ZIN RAS), Laboratory of Protozoology, Zoological Institute of the Russian Academy of Sciences, St. Petersburg, Russia	Transcriptome	No (closest genome available <i>Tribonema minus</i> , GCA_017506865.1, submitted 2021/03/22, NCBI)	Freshwater	Unknown. Probably related to host abundance. Host is relatively common in freshwater environments (ditches, lakes, ponds, and bogs).	[14]

Table 1 (Continued)

Host–Symbiont phylum	Species model system (Host–Symbiont)	Type of symbiosis	Identification of interaction stages	Culture available (Host/Symbiont)	Genomic information available		Type of environment	Peak of interaction in nature	References
					Symbiont	Host			
Chytridiomycota — Rozellida	<i>Rhizoclostratium globosum</i> — <i>Rozella rhizoclostratii</i>	Parasitism (Endobiotic)	Parasite cysts with chitin wall — Wheat Germ Agglutinin-Texas red; Host dispersive cyst and vegetative stage — Wheat Germ Agglutinin-Texas red; Host and parasite nucleus — DAPI/SYBR green I	Yes/Yes; Collection of Zoosporic Eufungi (CZEUM), University of Michigan, MI, USA	No (closest genome available: <i>Rozella allomyces</i>)	Genome (JGI Project Id: 1060049) and transcriptome (JGI Project Id: 1060050)	Freshwater	Unknown. All established cultures in CZEUM were isolated in Spring and Summer.	[71,72]
Chlorophyta — Blastocladiomycota	<i>Haematococcus pluvialis</i> — <i>Paraphysoderma sedebokerense</i>	Facultative parasitism (Epibiotic)	Host chloroplasts in green cells — red autofluorescence; infected host — reduction of red autofluorescence and change of algal cells' colour from green or red to brown (light microscopy). Parasite ameboid zoospores with D-mannose and/or D-glucose surface — FITC-Concanavalin A; Parasite chitin-bearing cysts — Wheat Germ Agglutinin-Texas red/FITC; parasites early and late sporangium — Nile Red; Host and parasite nucleus (DAPI/SYBR green I)	Yes/Yes; Norwegian Institute for Water Research Culture Collection of Algae (NIVA-CCA), Oslo, Norway.	Genome (JGI Project Id: 1098865) and Transcriptome (JGI Project Id: 1099000)	Genome and transcriptome	Freshwater	Unknown. The parasite causes severe epidemics in microalgal mass cultures for commercial purposes	[73–76]

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Falkowski PG, Fenchel T, Delong EF: **The microbial engines that drive Earth's biogeochemical cycles.** *Science* 2008, **320**:1034-1039.
2. Caron DA, Alexander H, Allen AE, Archibald JM, Armbrust EV, Bachy C, Bell CJ, Bharti A, Dyhrman ST, Guida SM *et al.*: **Probing the evolution, ecology and physiology of marine protists using transcriptomics.** *Nat Rev Microbiol* 2017, **15**:6-20.
3. Morris JJ, Lenski RE, Zinser ER: **The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss.** *mBio* 2012, **3**:e00036-12.
4. Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ: **Rethinking the marine carbon cycle: Factoring in the multifarious lifestyles of microbes.** *Science* 2015, **347**:1257594.
5. Klawonn I, den Wyngaert SV, Parada AE, Arandia-Gorostidi N, Whitehouse MJ, Grossart H-P, Dekas AE: **Characterizing the "fungal shunt": Parasitic fungi on diatoms affect carbon flow and bacterial communities in aquatic microbial food webs.** *Proc Natl Acad Sci U S A* 2021, **118**:e2102225118.
6. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinosa JC, Roux S, Vincent F *et al.*: **Determinants of community structure in the global plankton interactome.** *Science* 2015, **348**:1262073.
7. deVargas C, Audic S, Henry N, Decelle J, Mahe F, Logares R, Lara E, Berner C, Bescot NL, Probert I *et al.*: **Eukaryotic plankton diversity in the sunlit ocean.** *Science* 2015, **348**:1261605.
8. del Campo J, Sieracki ME, Molestina R, Keeling P, Massana R, Ruiz-Trillo I: **The others: our biased perspective of eukaryotic genomes.** *Trends Ecol Evol* 2014, **29**:252-259.
9. Keeling PJ, Burki F, Wilcox HM, Allam B, Allen EE, Amaral-Zettler LA, Armbrust EV, Archibald JM, Bharti AK, Bell CJ *et al.*: **The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing.** *PLoS Biol* 2014, **12**:e1001889.
10. Röttjers L, Faust K: **From hairballs to hypotheses—biological insights from microbial networks.** *Fems Microbiol Rev* 2018, **42**:761-780.
11. Bjorbaekmo MFM, Evenstad A, Røsaeg LL, Krabberød AK, Logares R: **The planktonic protist interactome: where do we stand after a century of research?** *ISME J* 2020, **14**:544-559.
- Huge bibliographic survey identifies ~2500 documented protist interactions from 500 publications over 150 years. The work classifies published interactions by symbiotic status. Such resources will only gain power in light of unbiased large-scale sampling of interactions which addresses functional biology of symbiosis, as proposed here.
12. Coordinators TO, Carradec Q, Pelletier E, Silva CD, Alberti A, Seeluthner Y, Blanc-Mathieu R, Lima-Mendez G, Rocha F, Tirichine L *et al.*: **A global ocean atlas of eukaryotic genes.** *Nat Commun* 2018, **9**:373.
- Large-scale marine meta-transcriptomes identifies 116 million expressed eukaryotic genes, a vast diversity of which are genes of unknown function. This work illustrates how a multitude of sampled eukaryotic genes are orphans in terms of taxonomic assignment, functional annotation and understanding of expression pattern relative to life-cycle development. This paper highlights the scale of the 'orphan' problem, a problem which will be partially addressed by the processes set out in this article.
13. Sibbald SJ, Archibald JM: **More protist genomes needed.** *Nat Ecol Evol* 2017, **1**:0145.
14. Torruella G, Grau-Bové X, Moreira D, Karpov SA, Burns JA, Sebé-Pedrós A, Völcker E, López-García P: **Global transcriptome analysis of the aphelid *Paraphelidium tribonemae* supports the phagotrophic origin of fungi.** *Commun Biol* 2018, **1**:231.
15. John U, Lu Y, Wohlrab S, Groth M, Janouškovec J, Kohli GS, Mark FC, Bickmeyer U, Farhat S, Felder M *et al.*: **An aerobic eukaryotic parasite with functional mitochondria that likely lacks a mitochondrial genome.** *Sci Adv* 2019, **5**:eaav1110.
16. Johnson MD, Oldach D, Delwiche CF, Stoecker DK: **Retention of transcriptionally active cryptophyte nuclei by the ciliate *Myrionecta rubra*.** *Nature* 2007, **445**:426-428.
17. Kodama Y, Suzuki H, Dohra H, Sugii M, Kitazume T, Yamaguchi K, Shigenobu S, Fujishima M: **Comparison of gene expression of *Paramecium bursaria* with and without *Chlorella variabilis* symbionts.** *Bmc Genomics* 2014, **15**:183.
18. Wisecaver JH, Hackett JD: **Transcriptome analysis reveals nuclear-encoded proteins for the maintenance of temporary plastids in the dinoflagellate *Dinophysis acuminata*.** *Bmc Genomics* 2010, **11**:366.
19. Lu Y, Wohlrab S, Glöckner G, Guillou L, John U: **Genomic insights into processes driving the infection of *Alexandrium tamarense* by the parasitoid *Amoebophrya* sp.** *Eukaryot Cell* 2014, **13**:1439-1449.
20. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Tuch BB, Siddiqui A *et al.*: **mRNA-Seq whole-transcriptome analysis of a single cell.** *Nat Methods* 2009, **6**:377-382.
21. Ku C, Sebé-Pedrós A: **Using single-cell transcriptomics to understand functional states and interactions in microbial eukaryotes.** *Philos Trans R Soc B* 2019, **374**:20190098.
22. Kolisko M, Boscaro V, Burki F, Lynn DH, Keeling PJ: **Single-cell transcriptomics for microbial eukaryotes.** *Curr Biol* 2014, **24**:R1081-R1082.
23. Yoon HS, Price DC, Stepanauskas R, Rajah VD, Sieracki ME, Wilson WH, Yang EC, Duffy S, Bhattacharya D: **Single-cell genomics reveals organismal interactions in uncultivated marine protists.** *Science* 2011, **332**:714-717.
24. Svensson V, Vento-Tormo R, Teichmann SA: **Exponential scaling of single-cell RNA-seq in the past decade.** *Nat Protoc* 2018, **13**:599-604.
25. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, Mildner A, Cohen N, Jung S, Tanay A *et al.*: **Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types.** *Science* 2014, **343**:776-779.
26. Wilbrey-Clark A, Roberts K, Teichmann SA: **Cell Atlas technologies and insights into tissue architecture.** *Biochem J* 2020, **477**:1427-1442.
27. Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J *et al.*: **A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution.** *Science* 2019, **365**:eaax1971.
28. Stubbington MJT, Rozenblatt-Rosen O, Regev A, Teichmann SA: **Single-cell transcriptomics to explore the immune system in health and disease.** *Science* 2017, **358**:58-63.
29. Uhlen M, Zhang C, Lee S, Sjöstedt E, Fagerberg L, Bidkhori G, Benfeitas R, Arif M, Liu Z, Edfors F *et al.*: **A pathology atlas of the human cancer transcriptome.** *Science* 2017, **357**:eaan2507.
30. Howick VM, Russell AJC, Andrews T, Heaton H, Reid AJ, Natarajan K, Butungi H, Metcalf T, Verzier LH, Rayner JC *et al.*: **The malaria cell atlas: single parasite transcriptomes across the complete *Plasmodium* life cycle.** *Science* 2019, **365**:eaaw2619.
31. Real E, Howick VM, Dahalan FA, Witmer K, Cudini J, Andradi-Brown C, Blight J, Davidson MS, Dogga SK, Reid AJ *et al.*: **A single-cell atlas of *Plasmodium falciparum* transmission through the mosquito.** *Nat Commun* 2021, **12**:3196.
- Cell atlas approaches are used to deconvolve transcriptome diversity of a virulent malarial parasite during its infection cycle. Developmental trajectories are identified and compared between parasites of the insect midgut and the salivary glands. Conserved and variable transcriptional responses between rodent and human malaria are compared.
32. Xue Y, Theisen TC, Rastogi S, Ferrel A, Quake SR, Boothroyd JC: **A single-parasite transcriptional atlas of *Toxoplasma gondii***

reveals novel control of antigen expression. *eLife* 2020, **9**: e54129

Cell atlas approaches are used to understand developmental progression of this parasitic protist providing key understanding of the transcriptome factors that drive developmental progression. Identified transcription factors are also shown to be linked to switches in surface antigens. Comparison to the *Plasmodium berghei* atlas data identifies 'concerted' genetic programs in these parasites.

33. Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé M-P, Renno J, Loe-Mie Y, Lifshitz A, Mukamel Z, Schmutz S *et al.*: **Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-seq.** *Cell* 2018, **173**:1520-1534.e20.
 34. Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ *et al.*: **Comprehensive single-cell transcriptional profiling of a multicellular organism.** *Science* 2017, **357**:661-667.
 35. Sebé-Pedrós A, Chomsky E, Pang K, Lara-Astiaso D, Gaiti F, Mukamel Z, Amit I, Hejnal A, Degnan BM, Tanay A: **Early metazoan cell type diversity and the evolution of multicellular gene regulation.** *Nat Ecol Evol* 2018, **2**:1176-1188.
 36. Siebert S, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, Juliano CE: **Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution.** *Science* 2019, **365**: eaav9314.
 37. Musser JM, Schippers KJ, Nickel M, Mizzon G, Kohn AB, Pape C, Hammel JU, Wolf F, Liang C, Hernández-Plaza A *et al.*: **Profiling cellular diversity in sponges informs animal cell type and nervous system evolution.** *bioRxiv* 2019:758276 <http://dx.doi.org/10.1101/758276>.
 38. Ku C, Sheyn U, Sebé-Pedrós A, Ben-Dor S, Schatz D, Tanay A, Rosenwasser S, Vardi A: **A single-cell view on alga-virus interactions reveals sequential transcriptional programs and infection states.** *Sci Adv* 2020, **6**:eaba4137
- The viral transcriptional trajectories are identified for the alga *Emiliania huxleyi* during virus infection, demonstrating the effect of infection on host gene transcription. These data reveal how the virus acts to shut down the nuclear functions while mitochondrial and plastid functions of the host persist.
39. Ma A, McDermaid A, Xu J, Chang Y, Ma Q: **Integrative methods and practical challenges for single-cell multi-omics.** *Trends Biotechnol* 2020, **38**:1007-1022.
 40. Datlinger P, Rendeiro AF, Boenke T, Senekowitsch M, Krausgruber T, Barreca D, Bock C: **Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing.** *Nat Methods* 2021, **18**:635-642 <http://dx.doi.org/10.1038/s41592-021-01153-z>
- Development of a new microfluidic droplet method that allows combinatorial pre-indexing of entire cellular transcriptomes within droplets allowing subsequent targeted scRNA-seq using microfluidics. The authors foresee the method will be straightforward to adapt to other microfluidic droplet applications and to alternative methods for scRNA-seq.
41. Goldstein LD, Chen Y-JJ, Dunne J, Mir A, Hubschle H, Guillory J, Yuan W, Zhang J, Stinson J, Jaiswal B *et al.*: **Massively parallel nanowell-based single-cell gene expression profiling.** *BMC Genomics* 2017, **18**:519.
 42. Tung JW, Heydari K, Tirouvanziam R, Sahaf B, Parks DR, Herzenberg LA, Herzenberg LA: **Modern flow cytometry: a practical approach.** *Clin Lab Med* 2007, **27**:453-468.
 43. García-Castro H, Kenny NJ, Iglesias M, Álvarez-Campos P, Mason V, Elek A, Schönauer A, Sleight VA, Neiro J, Aboobaker A *et al.*: **ACME dissociation: a versatile cell fixation-dissociation method for single-cell transcriptomics.** *Genome Biol* 2021, **22**:89.
 44. Needham DM, Yoshizawa S, Hosaka T, Poirier C, Choi CJ, Hehenberger E, Irwin NAT, Wilken S, Yung C-M, Bachy C *et al.*: **A distinct lineage of giant viruses brings a rhodopsin photosystem to unicellular marine predators.** *Proc Natl Acad Sci U S A* 2019, **116**:20574-20583.
 45. Martínez-García M, Brazel D, Poulton NJ, Swan BK, Gomez ML, Masland D, Sieracki ME, Stepanauskas R: **Unveiling in situ interactions between marine protists and bacteria through single cell sequencing.** *ISME J* 2012, **6**:703-707.
 46. Wideman JG, Monier A, Rodríguez-Martínez R, Leonard G, Cook E, Poirier C, Maguire F, Milner DS, Irwin NAT, Moore K *et al.*: **Unexpected mitochondrial genome diversity revealed by targeted single-cell genomics of heterotrophic flagellated protists.** *Nat Microbiol* 2020, **5**:154-165.
 47. Reid AJ, Talman AM, Bennett HM, Gomes AR, Sanders MJ, Illingworth CJR, Billker O, Berriman M, Lawnczak MK: **Single-cell RNA-seq reveals hidden transcriptional variation in malaria parasites.** *eLife* 2018, **7**:e33105.
 48. Ciobanu D, Clum A, Ahrendt S, Andreopoulos WB, Salamov A, Chan S, Quandt CA, Foster B, Meier-Kolthoff JP, Tang YT *et al.*: **A single-cell genomics pipeline for environmental microbial eukaryotes.** *IScience* 2021, **24**:102290.
 49. Sesen M, Whyte G: **Image-based single cell sorting automation in droplet microfluidics.** *Sci Rep* 2020, **10**:8736
- Demonstrates that cell sorting is possible based on live imaging within a microfluidic platform. The work goes on to show sorting of red blood cells with high specificity.
50. Zhang JQ, Siltanen CA, Liu L, Chang K-C, Gartner ZJ, Abate AR: **Linked optical and gene expression profiling of single cells at high-throughput.** *Genome Biol* 2020, **21**:49
- Constructs a unique printed droplet microfluidic work flow for high-throughput optical-image analysis coupled to RNA sequencing for analysis of single mouse and human cells. Using a similar approach to a flow cytometer, cells are sorted into individual nanolitre wells for sequencing. Wells are indexed by addition of oligos allowing the sequence data to be coupled to image data.
51. Gawryluk RMR, del Campo J, Okamoto N, Strassert JFH, Lukeš J, Richards TA, Worden AZ, Santoro AE, Keeling PJ: **Morphological identification and single-cell genomics of marine diplomonads.** *Curr Biol* 2016, **26**:3053-3059.
 52. Strassert JFH, Karnkowska A, Hehenberger E, del Campo J, Kolisko M, Okamoto N, Burki F, Janoušková J, Poirier C, Leonard G *et al.*: **Single cell genomics of uncultured marine alveolates shows paraphyly of basal dinoflagellates.** *ISME J* 2018, **12**:304-308.
 53. Pegoraro G, Misteli T: **High-throughput imaging for the discovery of cellular mechanisms of disease.** *Trends Genet* 2017, **33**:604-615.
 54. Nitta N, Sugimura T, Isozaki A, Mikami H, Hiraki K, Sakuma S, Iino T, Arai F, Endo T, Fujiwaki Y *et al.*: **Intelligent image-activated cell sorting.** *Cell* 2018, **175**:266-276.e13
- Demonstrates the development of technology to achieve intelligent image-activated cell sorter (IACS) resulting in targeted sorting at high rate. 'In a sense, the intelligent IACS is an imaging version of FACS'. The high quality of the images indicates the system is capable of identifying features of various cell types ranging from 3 to 30 µm diameter. The work includes validation experiments performed using microalgal cells (using protein localisation data) and blood cells (based on interactions).
55. Usaj MM, Styles EB, Verster AJ, Friesen H, Boone C, Andrews BJ: **High-content screening for quantitative cell biology.** *Trends Cell Biol* 2016, **26**:598-611.
 56. Kraus OZ, Grys BT, Ba J, Chong Y, Frey BJ, Boone C, Andrews BJ: **Automated analysis of high-content microscopy data with deep learning.** *Mol Syst Biol* 2017, **13**:924.
 57. Rappez L, Rakhlin A, Rigopoulos A, Nikolenko S, Alexandrov T: **DeepCycle reconstructs a cyclic cell cycle trajectory from unsegmented cell images using convolutional neural networks.** *Mol Syst Biol* 2020, **16**:e9474 <http://dx.doi.org/10.15252/msb.20209474>
- Represents a big step forward by presenting a model able to reconstruct a continuous and closed cell cycle trajectory from unsegmented single-cell brightfield and fluorescent microscopy images. The approach combines deep-learning and dimensionality-reduction techniques. The work includes validation of the cell trajectory data by correlation analysis with real time measurements taken from live-cell imaging of cells throughout the cell cycle.
58. Chambouvet A, Morin P, Marie D, Guillou L: **Control of toxic marine dinoflagellate blooms by serial parasitic killers.** *Science* 2008, **322**:1254-1257.

59. Alacid E, Reñé A, Garcés E: **New insights into the parasitoid *Parvilucifera sinerae* life cycle: the development and kinetics of infection of a bloom-forming dinoflagellate host.** *Protist* 2015, **166**:677-699.
60. Rad-Menéndez C, Gerphagnon M, Garvetto A, Arce P, Badis Y, Sime-Ngando T, Gachon CMM: **Rediscovering *Zygorhizidium affluens* Canter - molecular taxonomy, infectious cycle, and cryopreservation of a chytrid infecting the bloom-forming diatom *Asterionella formosa*.** *Appl Environ Microb* 2018, **84**:AEM.01826-18.
61. Caicedo JC, Cooper S, Heigwer F, Warchal S, Qiu P, Molnar C, Vasilevich AS, Barry JD, Bansal HS, Kraus O *et al.*: **Data-analysis strategies for image-based cell profiling.** *Nat Methods* 2017, **14**:849-863.
62. Pratapa A, Doron M, Caicedo JC: **Image-based cell phenotyping with deep learning.** *Curr Opin Chem Biol* 2021, **65**:9-17.
63. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, Enard W: **Comparative analysis of single-cell RNA sequencing methods.** *Mol Cell* 2017, **65**:631-643.e4.
64. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM *et al.*: **Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets.** *Cell* 2015, **161**:1202-1214.
65. Islam S, Zeisel A, Joost S, Manno GL, Zajac P, Kasper M, Lönnerberg P, Linnarsson S: **Quantitative single-cell RNA-seq with unique molecular identifiers.** *Nat Methods* 2014, **11**:163-166.
66. Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks G-J, Larsson AJM, Faridani OR, Sandberg R: **Single-cell RNA counting at allele and isoform resolution using Smart-seq3.** *Nat Biotechnol* 2020, **38**:708-714.
67. Fan X, Tang D, Liao Y, Li P, Zhang Y, Wang M, Liang F, Wang X, Gao Y, Wen L *et al.*: **Single-cell RNA-seq analysis of mouse preimplantation embryos by third-generation sequencing.** *PLoS Biol* 2020, **18**:e3001017.
68. Karst SM, Ziels RM, Kirkegaard RH, Sørensen EA, McDonald D, Zhu Q, Knight R, Albertsen M: **High-accuracy long-read amplicon sequences using unique molecular identifiers with Nanopore or PacBio sequencing.** *Nat Methods* 2021, **18**:165-169.
69. Alacid E, Reñé A, Camp J, Garcés E: **In situ occurrence, prevalence and dynamics of *Parvilucifera* parasitoids during recurrent blooms of the toxic dinoflagellate *Alexandrium minutum*.** *Front Microbiol* 2017, **8**:1624.
70. Kodama Y, Fujishima M: **Synchronous induction of detachment and reattachment of symbiotic *Chlorella* spp. from the cell cortex of the host *Paramecium bursaria*.** *Protist* 2013, **164**:660-672.
71. Letcher PM, Longcore JE, Quandt CA, da SilvaLeite D, James TY, Powell MJ: **Morphological, molecular, and ultrastructural characterization of *Rozella rhizoclostratii*, a new species in Cryptomycota.** *Fungal Biol* 2017, **121**:1-10.
72. James TY, Pelin A, Bonen L, Ahrendt S, Sain D, Corradi N, Stajich JE: **Shared signatures of parasitism and phylogenomics unite Cryptomycota and microsporidia.** *Curr Biol* 2013, **23**:1548-1553.
73. Gutman J, Zarka A, Boussiba S: **Evidence for the involvement of surface carbohydrates in the recognition of *Haematococcus pluvialis* by the parasitic blastoclad *Paraphysoderma sedebokerensis*.** *Fungal Biol* 2011, **115**:803-811.
74. Asatryan A, Boussiba S, Zarka A: **Stimulation and isolation of *Paraphysoderma sedebokerense* (Blastocladiomycota) propagules and their infection capacity toward their host under different physiological and environmental conditions.** *Front Cell Infect Microbiol* 2019, **9**:72.
75. Luo Q, Chao B, Tao M, Huang Y, Zheng Y, Lv Y, Jia L, Wang C, You X, Jia B *et al.*: **Genome and transcriptome sequencing of the astaxanthin-producing green microalga, *Haematococcus pluvialis*.** *Genome Biol Evol* 2018, **11**:evy263.
76. Lin J, Yan H, Zhao L, Li Y, Nahidian B, Zhu M, Hu Q, Han D: **Interaction between the cell walls of microalgal host and fungal carbohydrate-activate enzymes is essential for the pathogenic parasitism process.** *Environ Microbiol* 2021 <http://dx.doi.org/10.1111/1462-2920.15465>.