

# SUBCELLULAR TRANSCRIPTOMICS & PROTEOMICS: A COMPARATIVE METHODS REVIEW

Josie A. Christopher<sup>1,2</sup>, Aikaterini Geladaki<sup>1,3</sup>, Charlotte S. Dawson<sup>1,2</sup>, Owen L. Vennard<sup>1,2</sup>, Kathryn S. Lilley<sup>1,2</sup>

1. Cambridge Centre for Proteomics, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, CB2 1GA, UK
2. Milner Therapeutics Institute, Jeffrey Cheah Biomedical Centre, Puddicombe Way, Cambridge, CB2 0AW, UK
3. Department of Genetics, University of Cambridge, 20 Downing Place, Cambridge, CB2 3EJ, UK

## Abstract

The internal environment of cells is molecularly crowded, which requires spatial organization via subcellular compartmentalization. These compartments harbor specific conditions for molecules to perform their biological functions, such as coordination of the cell cycle, cell survival, and growth. This compartmentalization is also not static, with molecules trafficking between these subcellular neighborhoods to carry out their functions. For example, some biomolecules are multifunctional, requiring an environment with differing conditions or interacting partners, and others traffic to export such molecules. Aberrant localization of proteins or RNA species have been linked to many pathological conditions, such as neurological, cancer and pulmonary diseases. Differential expression studies in transcriptomics and proteomics are relatively common, but the majority have overlooked the importance of subcellular information. Additionally, subcellular transcriptomics and proteomics data do not always co-locate due to the biochemical processes that occur during and after translation, highlighting the complementary nature of these fields. In this review, we discuss and directly compare the current methods in spatial proteomics and transcriptomics, which include sequencing- and imaging-based strategies, to give the reader an overview of the current tools available. We also discuss current limitations of these strategies, as well as future developments in the field of spatial -omics.

## Introduction

Molecular biology is the study of cellular functions via processes such as molecular synthesis, modification, and interactions. RNA and proteins can have multiple roles and interacting partners that require close physical proximity to each other within the cell to function. Therefore, precise control of localization or co-localization by selective congregation and isolation of biochemical processes is integral and intrinsically linked to cellular functions. For instance, in context of transcription and

translation, mRNA is shuttled out of the nucleus, where it docks at ribosomes within the cytosol, at the endoplasmic reticulum (ER) or near the mitochondria, dependent on the coded protein and cellular conditions <sup>1-3</sup>. Translation of mRNA at the coded protein's functional site, rather than at a singular canonical and/or punctate location, is clearly demonstrated within polarized cells, such as neurons or intestinal epithelial cells <sup>4,5</sup>. Hence, studying subcellular localization not only gives insights into the organization of cellular compartments but how cells function, so techniques that provide spatial context are important tools in molecular biology.

The relationship between DNA, RNA and proteins does not represent a linear dogma. Interactions, or "interactomes", between nucleic acids and proteins are fundamental for cellular function. RNA-binding proteins (RBPs), originally thought to exclusively function in gene regulation via ribonucleoprotein (RNP) complex formation, have now been shown to have more extensive interplay between protein and RNA interactomes <sup>6</sup>. A prime example of RNA- and RBP-mediated regulation via subcellular re-localization is the short non-coding RNA transcript Y3 RNA, which orchestrates translocation of the RBP Rho 60-kDa autoantigen between the cytosol and nucleus as part of a UV-induced survival mechanism <sup>7,8</sup>. A more classic example of sub cellular control is during the cell cycle, where cyclins and cyclin-dependent kinases (CDKs) traffic between nuclei and cytosol <sup>9</sup>. An in-depth immunofluorescence study has recently captured single-cell variability of subcellular composition during the cell cycle <sup>10</sup>.

Aberrant trafficking of RNA and protein have been implicated in several pathological conditions, including amyotrophic lateral sclerosis (ALS) and pulmonary arterial hypertension (PAH), respectively <sup>11,12</sup>. A well-documented example of mis-localization causing severe disease is the most common mutation in cystic fibrosis, F508del. Immunofluorescence and subcellular fractionation strategies have shown this mutation causes the cystic fibrosis transmembrane regulator (CFTR) ion channel to misfold and accumulate at the ER, preventing CFTR expression at the plasma membrane (PM) and, consequently, impairing mucus clearance in the lungs <sup>13-15</sup>. This has aided the design of pharmacological intervention to correct this misfolding and subsequent mis-localization <sup>16</sup>. In many of these cases, early stages of disease can be identified by translocation events, which can precede or be independent to detectable changes in gene expression and, therefore, can only be studied at the subcellular level <sup>17</sup>. Despite this, temporal or differential expression is more commonly studied because it is more straightforward, though novel tools to study the spatial dimension on an -omics scale is opening new opportunities for a better understanding of cellular function.

Spatial proteomics and transcriptomics have often been reviewed independently with technical details covered in previous articles <sup>18-20</sup>. Here, we outline and directly compare methods (summarized in Table 1) that interrogate the spatial transcriptomic and proteomic within subcellular compartments of cells, rather than spatial information at the tissue-level, and suggest gaps in technology in need of further advancement. We aim to provide a resource for newcomers to spatial omics who wish to unpick the busy yet spatially organized environment within cells.

Method	Principle	Examples of biological insights	Live, fixed, or lysed samples?	<i>In situ</i> ?	Targeted?
<i>Imaging</i>					
<b>Affinity reagents</b>	Exogenous dyes or probes (e.g. antibodies or oligonucleotides) designed to target specific molecules of interest (MOI).	The largest database of human protein subcellular localizations using stringently validated antibodies, giving insights into cell variability and mapping subcellular localization of COVID-19 interactors <sup>21,22</sup> . smFISH aided the understanding of how liquid-liquid phase separation aids formation of rotavirus replication factories (considered virus-made membraneless organelles) <sup>23</sup> .	Primarily fixed samples (exception of live FISH <sup>24</sup> ).	✓	Targeted, label MOI
<b>Fluorescently tagged proteins</b>	Fluorescent proteins (typically) genetically fused to MOI and, therefore, co-expressed with the MOI.	Genetically fused fluorescent proteins were used to gain insight into the pH- and receptor-dependent endocytic entry of SARS virus into the host cell <sup>24</sup> .	Live/fixed.	✓	Targeted, label MOI
<b>Imaging flow cytometry (IFC)</b>	A combination of flow cytometry and microscopy to capture spatial information using fluorescent probes.	An IFC method was developed to provide a more informative diagnostic tool for types of acute leukemia <sup>25</sup> .	Live/fixed.	✓	Targeted, label MOI
<b>Imaging mass cytometry (IMC)</b>	Uses heavy-metal probes conjugated to antibodies which ablated pixel-by-pixel and measured using mass spectrometry. This improved multiplexing of probes due to the reduced spectral overlap compared to fluorescent strategies.	Used for cellular phenotyping of breast cancer, and lesions in multiple sclerosis and lymphoid organs <sup>26-28</sup> . Primarily used for tissue-level insights, rather than subcellular, though some subcellular information is achievable with the method.	Fixed.	✓	Targeted, label MOI
<b>Mass spectrometry imaging (MSI)</b>	Similar to IMC, but ablation leads to ionization of all molecules within the pixel, producing a separate spectra per pixel of the sample.	Primarily still tissue-level resolution, rather than subcellular resolution. Has been used for intraoperative imaging of pituitary adenomas for biomarkers that are usually difficult to detect efficiently <sup>29</sup> .	Typically fixed.	✓	Untargeted, cell-wide
<i>Biochemical separation</i>					
<b>Basic centrifugation/detergent-based</b>	Uses targeted centrifugation or detergent step(s) to achieve enrichment of a specific cellular component or organelle of interest.	Used in the study of mitochondrial transport in <i>Trypanosoma brucei</i> to aid understanding of parasitic physiology <sup>30</sup> and gain insights into proinflammatory gene regulation in context of subcellular dynamics of macrophages from mice <sup>31</sup> .	Lysed, <i>in vitro</i> .		Untargeted, enrich organelle(s) of interest
<b>Correlation profiling</b>	Uses multiple centrifugation or detergent steps of increasing spin speed/time or	Used to track the subcellular proteome of host cells over the course of HCMV infection in a spatial and	Lysed, <i>in vitro</i> .		Untargeted, cell-wide

	solubility, respectively, to collect an abundance profile of one or multiple subcellular components. Can be used for cell-wide analysis of molecules.	temporal context <sup>32</sup> . Also, used to identify that lysosomal trapping is important for the efficacy of drugs that aid antigen presentation <sup>33</sup> .		
<b>Electrophoresis-based</b>	Separates subcellular components via their charge state using modified electrophoresis techniques.	Used to assess the protein composition of the secretory pathway in plants that are otherwise difficult to resolve due to their similar density <sup>34</sup> .	Lysed, <i>in vitro</i> .	Untargeted, cell-wide
<i>Proximity labeling</i>				
<b>BioID and APEX</b>	Fusion of bait protein(s) to either a biotin ligase (e.g. BioID) or peroxidase (e.g. APEX) that covalently labels molecules in immediate proximity of the bait with a small, exogenous substrate. The substrate can then be purified along with the labeled molecules.	BioID has revealed novel organellar components of the <i>Trypanosoma brucei</i> , flies and worms <sup>35-38</sup> and identifying novel proteins involved in hyperpolarization that are linked to neurodegenerative diseases <sup>39,40</sup> APEX-seq identified stress-type-dependent RNA interactions with stress granules <sup>18,41</sup> .	Lysed, <i>in vivo</i> labeling.	Untargeted, label organelle(s) of interest

MOI: Molecule of interest

**Table 1. Summary of each method covered within this review.** The table includes a short description about the principle of each method, examples of their biological insights or applications and basic comparisons of the characteristics of the methods.

# 1. Imaging the spatial transcriptome and proteome

## 1.1. Microscopy-based imaging

Microscopy is the most well-established and largest branch of imaging with a variety of labelling strategies for targeting proteins and transcripts, often at a single-cell level. Conducting microscopy studies on a global spatial scale can be challenging and laborious, due to costly generation of antibodies or recombinant organisms, and limited multiplexing capacity. Additionally, sample preparation is rarely a one-size-fits-all process. For example, fixing is usually dependent on the subcellular compartment of interest and phototoxicity is a limiting factor in live-cell imaging. Fixing cells can disrupt molecular- and macro-organization and structures, causing artificial localization of molecules<sup>42</sup> but does not suffer from issues of phototoxicity, and can capture snapshots of transcripts and proteins, which rapidly fluctuate or have low copy number. Some applications are limited to fixed samples only, such as immunofluorescence or fluorescent *in situ* hybridization (FISH), whereas others have the capacity for live-cell imaging, such as genetically fusing fluorescent tags. Recent emergence of high-throughput and super-resolution microscopy has allowed mid- to large-scale spatial studies of transcripts and proteins, permitting quantitative measurements alongside the “seeing is believing” aspect at which imaging excels. Furthermore, while simultaneous genome-wide live-cell imaging is not yet possible, recent advancements in the field of high-content imaging are enabling faster image acquisition at higher resolution, though often with a trade-off between the two<sup>43</sup>. Both the technological advancements of the instrumentation and bioimaging informatics have been extensively reviewed<sup>44–48</sup>.

Here we briefly discuss the main labelling options and some alternative imaging approaches, whilst outlining the advantages and disadvantages, and giving representative examples of their use in subcellular research, specifically in the context of large-scale spatial studies. The following labelling strategies are not necessarily exclusive to each other, and combinational labelling protocols have been documented<sup>49–53</sup>.

### 1.1.1. Visualization of using affinity reagents

Antibodies & organelle-specific dyes

In the case of proteins, the use of antibodies against specific endogenous proteins of interest is often known as immunofluorescence or immunocytochemistry. Immunofluorescence can be highly sensitive when using signal amplifying reagents, such as secondary antibodies conjugated to various fluorophores. Readily available commercial antibodies make comparative studies of protein localization in different cell or tissue samples easy and fast, particularly in commonly used model organisms, such as humans and mice. Finding commercial antibodies for some less well-studied species and proteins can be more difficult. This can be overcome by genetically fusing an epitope, such as FLAG<sup>®</sup>, to the protein of interest and then using an antibody against this epitope to indirectly label the protein. However, in this case a fluorescent protein, such as GFP, genetically fused to the protein is often favored as it negates the need for the antibody labelling step. Chemical, organelle-specific dyes, such as DAPI for nuclei staining, can also be used alongside antibodies. Reviews are available detailing such dyes<sup>54,55</sup>. It should be noted that antibodies are prone to batch-to-batch variability and poor specificity that can yield false results from non-specific and variable binding. These drawbacks have caused major reproducibility crises amongst the scientific community<sup>56</sup>. Though, in recent years,

there has been a huge drive to address this key issue with commercial suppliers providing extensive validation and moving toward recombinant products with less batch variability. Additionally, with the increasing accessibility of CRISPR technology, validating specificity of antibodies using CRISPR knockouts is becoming common practice. Immunofluorescence-based methods are also restricted to static endpoint measurements since such experiments require cell fixation and permeabilization prior to intracellular staining (Figure 1a). Sample preparation can be very context specific and inappropriate selection of fixation and permeabilization approaches can affect protein localization by introducing artefacts or causing loss of soluble proteins<sup>20,57</sup>. However, standardization of sample preparation and developments in automation has allowed multiplexing of off-the-shelf antibodies to improve throughput<sup>42,58</sup>.

Limited global spatial proteomics experiments have been conducted, because of the aforementioned restrictions. The largest immunofluorescence-based subcellular proteomics study performed to date is the work of the Cell Atlas database. This work is part of the wider Human Protein Atlas (HPA) initiative, aiming to document the entirety of the human subcellular proteome in different human cell and tissue types to elucidate protein function and create a comprehensive biological resource for human proteins in health and disease<sup>59-61</sup>. HPA have collaborated with other international-scale projects, such as UniProt, NextProt, GO, ELIXIR, to provide publicly available databases of subcellular information for the wider scientific community<sup>62-65</sup>. During the past two decades, a near proteome-wide collection of antibodies has been created and validated for the purpose of this initiative<sup>66-70</sup>. This work used 14,000 antibodies to systematically map the spatial distribution of 12,003 proteins at single-cell resolution to one or more of 30 different subcellular niches. Of those proteins, 5,662 lacked subcellular localization information in the literature prior to this study. This classification was performed using a combination of manual and computational image analysis approaches<sup>60,71</sup>. Notably, the images were obtained using high resolution confocal microscopy, enabling assignment of proteins to fine, less-well characterized cellular structures, such as microtubule ends, cytokinetic bridge sub-compartments, and the nucleolar fibrillar center, as well as to functionally uncharacterized subcellular niches, such as rods and rings. Moreover, this work showed that approximately half of all human proteins (6,163 out of 12,003 proteins in this dataset) localize to multiple (two or more) subcellular niches. This dataset also revealed that more than one sixth of the human proteome displays variability in terms of expression levels or subcellular distribution at the level of single cells<sup>60</sup>. During the COVID-19 pandemic, with collaborators, HPA turned to mapping the distribution of the virus' key host interactor, ACE2, across >150 human tissues, as well as the human interactome of COVID-19 with the aim to determine whether readily available drugs can be repurposed in the fight against the virus<sup>21,22</sup>.

In immunofluorescence, the number of proteins of interest that can be probed in one sample is largely limited to the number of fluorochromes that can be used without causing signal interference by spectral overlap or fluorescent bleed-through into other channels. Therefore, traditionally only around 4-6 fluorochromes could be used at a time, where each primary antibody is labelled with its unique fluorochrome (e.g. conjugated to a secondary antibody). Recent developments in using either cyclical probing with antibodies, such as CycIF, or using them in combination with other types of probes, such as oligonucleotides as molecular barcodes in CODEX, has allowed for improved multiplexing<sup>58,72-74</sup>. Another example of overcoming fluorescent signal overlap was the use of unique and identifiable DNA origami structures with the blinking kinetics of DNA-PAINT that allowed multiplexing antibody probes in a single channel with super-resolution microscopy<sup>75,76</sup>.

The targeting of RNA transcripts directly via antibodies is far less prevalent than proteins. Antibodies against RNA antigens do exist, however, these are limited to global RNA applications. For example, antibodies against a subtype of RNA, such as ribosomal RNA (rRNA), or for epigenetic applications,

such as certain global modifications of RNA (e.g. methylation or acetylation groups). This restricts their primary use to immunoblotting or immunopurifications and is not typically applicable to imaging <sup>77,78</sup>.

### In situ hybridization (ISH)

*In situ* hybridization (ISH) was first discovered as a useful nucleic acid labelling tool in 1969 using radioactive tritium-labelled antisense sequences to image the nuclei of frog eggs <sup>79</sup>. Fluorescent ISH (FISH) was soon adopted as a safer, more stable alternative <sup>80</sup>. The oligonucleotides used in FISH are designed to hybridize on the RNA target by sequence complementarity. These oligonucleotides are labelled either directly or indirectly, via a secondary probe (such as an antibody) conjugated to a fluorophore (Figure 1d). Single-molecule resolution was enabled by “tiling” multiple antisense probes along a sequence of interest to boost signal and has been a powerful tool in understanding the role of RNA localization in biology, such as in meiosis and neuromuscular junctions <sup>81–83</sup>. The main restriction of FISH has been its low-throughput and need to be performed in fixed cells to prevent RNase and DNase degradation of nucleic acids, limiting its use for temporal applications. However, FISH imaging in live cells, known as “live FISH”, has been achieved with the caveats of using toxic permeabilization techniques and rapid sequestering of the molecular beacons in the nucleus <sup>84,85</sup>. It has only been with the recent developments within CRISPR/Cas9 technology, that live FISH has been possible without such drawbacks <sup>86</sup>. Despite live-imaging alternatives, such as aptamers, RNA FISH is still a gold-standard technique for RNA localization, and recent advancements, such as CRISPR/Cas9, has kept it current and pervasive.

FISH has a multitude of available signal-amplifying probes to choose from, which are particularly useful for overcoming hurdles commonly found in difficult targets and samples, such as short non-coding RNA and tissues <sup>87</sup>. Generally, these probes have branched structures that increase the molecular surface area for multiple fluorophores to bind to the molecule of interest, which form the basis for single molecule inexpensive FISH (smiFISH), FISH with Sequential Tethered and Intertwined ODN Complexes (FISH-STIC), branched DNA (bDNA) FISH and hybridization chain reaction (HCR) FISH <sup>87–91</sup>. For targets that require particularly high specificity, such as short non-coding RNAs, padlock probes can covalently “lock” and amplify the signal using a rolling circle mechanism (RCA) <sup>92,93</sup>. bDNA probes were used for a large-scale imaging study, which targeted 928 genes involved in cancer, endocytosis, and metabolism at a single-cell level <sup>94</sup>. The use of enzyme-amplification of ISH probes and 96-well plates enabled mapping of mRNA dynamics in embryogenesis of *Drosophila*, achieving analysis of 3,370 transcripts and demonstrated a correlation between mRNA localization and subsequent protein localization and function <sup>95</sup>. Additionally, super-resolution FISH was used alongside RNA-seq methods to track the dynamics of proteins in dendritic cells <sup>4</sup>.

Multiplexing is also a powerful feature of FISH with easy to perform probe generation and sequential rehybridization of the sample, allowing for multiple rounds of re-probing and fluorescent barcoding of thousands of molecules, with minimal loss of signal <sup>96</sup>. Novel methods, such as multiplexed error-robust FISH (MERFISH) and sequential barcoding FISH (seqFISH) methods, exploit such characteristics and, in theory, have the capability of generating spatial information of the entire known transcriptome in just eight rounds of hybridization and four dyes ( $4^8 = 65,536$ ) <sup>96–98</sup>. Realistically, this level of coverage is not achievable with the exponential increase in error-rates per round of hybridization. MERFISH employs an error-detection barcoding scheme to account for a proportion of this error and when used in conjunction with bDNA probes to amplify the signal across ~10,000 transcripts by 10.5-fold <sup>98</sup>. Optical overcrowding of transcripts is also a limiting factor for such techniques. seqFISH+ was developed to circumvent this optical overcrowding by expanding the fluorophore palette from 4-5 colors to 60 “pseudocolors” using molecular barcoding, allowing analysis of 24,000 genes in four rounds with one round of error correction <sup>99</sup>. MERFISH and seqFISH have provided insight into the spatial organization of the cell cycle, mouse hippocampus, and tissue development and homeostasis, as well as capturing nascent transcription active sites of genes <sup>100–103</sup>. Amplification is very powerful,

though it only provides a global increase of intensity across targets, which cannot distinguish real RNA spots/signals from non-specifically bound probes, which affects the resolution. To overcome this, experimentation of different split-probes was conducted to achieve impressively punctate transcript spots, that can only fluoresce when two probes dock within immediate proximity on a highly specific, shared bridge sequence<sup>104</sup>. An untargeted alternative to the above, fluorescent *in situ* sequencing (FISSEQ), used crosslinking and reverse transcription of RNA *in situ* to perform RNA-seq with cyclic fluorescent probe ligations directly on the sample, which was measured via confocal microscopy<sup>105</sup>. This method was demonstrated in a variety of sample types, such as primary fibroblasts, tissues and whole embryos, and could be powerful in applications such as cellular phenotyping and gene regulation. The original FISSEQ publication, uncovered that lncRNAs preferentially localize in the nucleus. The premise of this method is powerful but FISSEQ struggles to attain read counts comparable with standard scRNA-seq, is difficult to perform in tissues and is limited to short-reads<sup>105</sup>. A new variation of FISSEQ, known as INSTA-seq, has recently been developed for longer reads<sup>106</sup>. To determine the precise subcellular localization of transcripts, it is recommended that organelle-specific dyes or immunofluorescence or organellar proteins are used as counterstains in these approaches<sup>105</sup>. The versatility of FISH shows that it still has untapped potential in the transcriptomics world and some of the newer methods have been recently reviewed<sup>107</sup>.

### 1.1.2. Visualization using fluorescently tagged proteins

Fluorescent proteins (FPs)

Genetically fused fluorescent proteins (FPs) are the next most prolific method of fluorescently labelling molecules, with the work which allowed scientists to harness FPs for research winning the Nobel Prize in chemistry in 2008<sup>108–110</sup>. Since the discovery and enhanced engineering of FPs, their use has provided immense biological insights into multiple processes, including demonstrating pH- and receptor-dependent endocytic viral entry during SARS infection<sup>24</sup>. This strategy involves fusing a reporter protein gene, usually a fluorescent protein or a sequence that can be fluorescently labelled downstream, to a protein of interest using transfection. When the protein of interest is expressed, so is the fused reporter protein or sequence, which can then either be directly excited at the appropriate wavelength or labelled with a fluorophore (e.g. a fluorescent antibody) (Figure 1b). In contrast to strategies with affinity reagents, FPs allow for live cell imaging, capturing temporal protein dynamics. An innovative, multicolored system called Fluorescent Ubiquitination-based Cell Cycle Indicator (FUCCI) utilizes fused FP monomers to two proteins, Cdt1 and Geminin, that are specifically degraded in different parts of the cell cycle, at S/G2 and M/G1 phases, respectively. This strategy allows for cell-cycle-dependent multicolored labelling of the nuclei<sup>111</sup>. The strategy has allowed for deconvolution of cell cycle states and cellular processes that are otherwise difficult to distinguish. For example, it has been used to determine the relationship between the progression of double-stranded break repair and cell cycle status in living cells with the aim to help development and assessment of cancer therapies<sup>112</sup>. However, sensitivity can be an issue, as it has been shown that only a third of the most abundant proteins in mammalian cells can be detected using the most widely used FP, green fluorescent protein (GFP), although this can be mitigated via using more photostable or/and brighter tags<sup>113</sup>. Furthermore, it has been shown that in certain cases tagging endogenous proteins can interfere with specific properties of native molecules, including its subcellular localization. For instance, FPs have been found to erroneously locate at the endomembrane system of mammalian cells<sup>114,115</sup>. This localization artefact can be influenced by where the FP has been genetically encoded on the target protein (e.g. on the N- or C- terminus). This effect was extensively examined in budding yeast<sup>116</sup>. As well as this, protein fusion can also impair the normal expression, function, or degradation

patterns of the native protein. Therefore, verification is required to ensure that endogenous localization and expression of the target molecule is unaffected by genetic fusion.

*Saccharomyces cerevisiae* have highly efficient homologous recombination processes compared to mammalian cells, making it relatively easy to generate FP-fused libraries, while generally preserving the normal expression patterns of the endogenous genes. Therefore, the species was used to conduct the first genome-wide library of a eukaryote for live-cell imaging using GFP-tagging, achieving systematic localization of 75% of the yeast proteome to 22 distinct subcellular niches under normal culture conditions. This study provided novel localization information on 1,630 proteins<sup>117</sup>. Subsequent studies have used this yeast library under multiple conditions of environmental stress to uncover yeast protein localization dynamics, as well as providing a quantitative dimension<sup>118–122</sup> (reviewed in:<sup>123</sup>). Improved technology has led to further ease with creating genome-wide fluorescent fusion libraries. For example, the SWAp-Tag method, which allows efficient modification of a parental library and was employed for generating both an N- and C-terminally-tagged yeast proteomes<sup>124–126</sup>. Such extensive and numerous libraries enabled meta-analysis of protein localization dynamics in a quantitative manner with an unsupervised computational method<sup>127</sup>. Such approaches have been able to differentiate perturbation-specific re-localization events from more generalized stress responses, concluding that protein subcellular localization provides an important layer of cellular regulation, independent from modulation of protein expression levels<sup>17,127</sup>. Due to the efforts mentioned above, several databases containing imaging data on the spatial organization of the *S. cerevisiae* proteome are now publicly available<sup>128–133</sup>.

Similar efforts to systematically probe human protein subcellular localization using fluorescent reporter fusions have also been published, but so far have only covered a small proportion of the proteome. For example, a collection of N- and C-terminal GFP fusions to cDNA was generated to study protein localization in living human cells, resulting in localization assignment for 1,600 human proteins<sup>115</sup>. Similarly, an annotated reporter clone collection was built via exon tagging using retroviral particle-mediated delivery in 2006<sup>134</sup>. This collection has been used in combination with time-lapse fluorescence microscopy to track the abundance and localization dynamics of more than 1,000 endogenous proteins in living human cells under different conditions<sup>134–137</sup>. More recently, 1,311 proteins were fluorescently tagged using CRISPR-based fusion in multiple cell lines to achieve deep profiling of these proteins using 3D confocal microscopy, immunoprecipitation-MS and next-generation sequencing<sup>138</sup>.

The fluorescent tagging methods described above center on protein labelling, but variations of these approaches have also allowed probing of RNA localization. Typically, this has been possible by encoding RNA hairpins into the gene of interest, which when transcribed can then be targeted by a corresponding RNA-binding protein that is co-expressed and fused with fluorescent proteins<sup>139</sup>. The first and most used system of this kind is the MS2 system, which uses bacteriophage MS2 coat proteins (MCPs), which are RBPs, to target genetically inserted MS2 loops<sup>140</sup>. Similar systems exploiting FPs have been added to the RNA localization repertoire, such as the P77 bacteriophage coat protein (PCP) system<sup>141–146</sup>. tdTomato-labeled PCP (tdPCP) was used to successfully track individual mRNA molecules during translation at polysomes in different subcellular locations in dendrites<sup>147</sup>. This study also utilized SunTag molecules, which provide protein scaffolds for multimerization of fluorescent tags to boost poor signal and to study translation in real-time<sup>147,148</sup>. Several other methods have been developed for studying translation in both fixed and live cell applications, which are reviewed in detail in<sup>149</sup>. Single-molecule imaging of both translation and degradation in live cells can be achieved using the entertainingly named TRICK and TREAT methods, which both use dual-color MCP and PCP coat protein reporter systems. Translating RNA imaging by coat protein knockoff (TRICK) can distinguish

untranslated from translated transcripts by incorporating loops for MCP and PCP at different locations in sequence of the mRNA of interest<sup>150</sup>. During translation, the ribosome knocks off PCP in the coding region of the transcript leaving MCP behind<sup>151</sup>. 3(three)'-RNA end accumulation during turnover (TREAT) uses a similar concept, where PCP is used to label the 3' end of the transcript, which is lost during degradation<sup>152</sup>. TRICK and TREAT have both been used independently in HeLa cells under arsenite stress to show reporter mRNAs retained in P-bodies are suspended, neither being translated or degraded<sup>152–154</sup>.

MS2-based and MS2-like systems tended to suffer from low signal-to-noise ratios constitutive fusion of coat proteins to fluorescent proteins means that the fluorescence is independent of being bound to the sequence of interest. The signal-to-noise can be significantly improved by including a nuclear localization sequence, so unbound protein is sequestered in the nucleus to improve the background of cytoplasmic transcripts<sup>139,155</sup>. Also, much like FP-tagging, there is no clear rule as to where to genetically encode the RNA-stem loops within the endogenous transcript<sup>156</sup>. There has been evidence that introduction of MS2 coated stem loops in yeast causes inhibition of mRNA decay, leading to RNA fragments that can continue to fluoresce, leading to aberrant localization measurements<sup>157,158</sup>. However, there is debate whether this evidence was an artefact of gene expression and/or the methods used to assess this degradation<sup>159</sup>. To address these concerns, a modified coat-protein reporter system allowing for efficient RNA degradation was established in both yeast and mammalian cells<sup>160</sup>.

#### RNA aptamers

RNA aptamers have been used in both *in vitro* and *in vivo* imaging as affinity reagents and reporter tags, respectively<sup>161,162</sup>. They are short RNA oligo nucleotides that can be conjugated to fluorescent dyes or designed to bind and induce the fluorescence of exogenous small molecules such as 3,5-difluoro-4-hydroxybenzylideneimidazolidinone (DFHBI) which is structurally related to the GFP chromophore<sup>163</sup> (Figure 1c). DFHBI is structurally unstable, preventing its fluorescent activity until it is bound to the complementary active site of the fluorogenic RNA aptamer, bypassing the constitutive fluorescence that is caused by the persistent RBP-FP interaction in MS2-style systems.

The original DFHBI-binding RNA aptamer, Spinach, demonstrated excellent brightness with minimal background fluorescence and resistance to photobleaching. Typically, fluorogenic RNA aptamers are expressed fused to an RNA of interest (ROI) for subcellular RNA imaging in live cells [129, 131, 138]. Guet and colleagues used Spinach to show nuclear relocalization of *STL1* and *CTT1* transcripts in *Saccharomyces cerevisiae* upon osmotic stress<sup>164</sup>. Conversely, cyanine-conjugated RNA aptamers have been used as affinity reagents for live cell imaging of proteins including EGFR, human retinoblastoma protein and transferrin<sup>162,165–167</sup>

In comparison to antibodies, aptamers have improved versatility with flexible modifications, less batch-to-batch variability, less steric hindrance and are capable of labelling both nucleic acids and proteins<sup>162</sup>. However, Spinach, plus other RNA aptamers, have had issues with RNA degradation, intracellular folding, and thermal stability. Further aptamers, such as Spinach2 and Broccoli, have been designed to overcome these complications<sup>168,169</sup>. Additional fluorophores with corresponding fluorogenic aptamers have been designed to cover more of the visible and near-infrared spectra<sup>170–172</sup>. Indeed near-IR aptamers were the first to be adapted for live-cell super resolution RNA imaging, and have been used to detect subnuclear RNA structures in mammalian cells<sup>172,173</sup>. For further reviews of RNA aptamers, see<sup>174–176</sup>.

### 1.1.3 Imaging flow cytometry (IFC)

Imaging flow cytometry (IFC) could be considered an alternative microscopy-based technique and can achieve up to 20 nm resolution (Figure 2a)<sup>177</sup>. IFC combines the multiparameter capabilities of flow cytometry and the morphological and subcellular spatial capabilities of microscopy (including dark-field, light-field, and fluorescence). However, in IFC there tends to be a trade-off between throughput, sensitivity, and spatial resolution. To compensate for this, a technique to control the flow of cells in the microfluidics system was used to virtually “freeze” cells on the image sensor enabling longer exposure times in image acquisition<sup>177</sup>. This improved signal-to-noise, throughput, sensitivity, and resolution. Whilst IFC cannot perform super high-resolution imaging and capture more intricate subcellular features, its application has been particularly useful for rare cell events and in diagnostic contexts<sup>178</sup>. For example, it has been used as a diagnostic tool in acute leukaemia to assess PML protein bodies and the cytoplasmic versus nuclear localization of a characteristic antigen<sup>25</sup>. Another major consideration is that the approach requires cells to be in suspension and dissociation of adherent cells or tissues may cause aberrant localization of molecules. Whilst performed less frequently than protein analysis, RNA transcripts can be visualized using IFC<sup>179,180</sup>.

## 1.2. Non-microscopy-based imaging methods

Imaging techniques that do not rely on microscopy are also available to map subcellular localization. These typically consist of hybridizing flow cytometry and/or mass spectrometry to imaging. Whilst exciting, their use is still limited, therefore we only briefly provide an overview but direct to relevant sources of further reading.

### 1.2.1. Imaging mass cytometry

Imaging mass cytometry (IMC) uses a similar instrumental setup to mass cytometry, which hybridizes flow cytometry and mass spectrometry using a cyTOF (cytometry by time-of-flight). This technology does not suffer from the same degree of signal overlap compared to fluorescent tagging systems<sup>181,182</sup>. The mass spectrometry element allows discrimination between targets at an isotopic scale. This is achieved by coupling probes, commonly antibodies, to discrete heavy-metal isotope tags<sup>26,181</sup>. Currently, this tagging system allows around 40 targets of interest to be measured per single cell. In traditional mass cytometry, cells are passed through a microfluidics-style droplet system and through argon plasma at a high temperature when entering the instrument where covalent bonds within molecules are broken, releasing free, atomic-level ions. The ions enter a quadrupole where the heavy-metal isotope tags are selected. These tags go on to be separated by mass-to-charge in the TOF component of the instrument<sup>26,181</sup>. This is a destructive process, so cells cannot be sorted via this technique, unlike flow cytometry sorting methods (FACS), and spatial information is lost. Imaging mass cytometry (IMC) overcame this loss of spatial information by coupling laser ablation of tissue slide or cell culture a pixel at a time into a cyTOF (Figure 2c). In the first publication of this method, the ability to untangle the heterogeneity of breast cancer samples was demonstrated<sup>26</sup>. A similar study recently claimed to achieve subcellular resolution using IMC for 37 proteins in 483 breast cancer tumors to assess the phenogenomic correlation with protein expression<sup>183</sup>. Breast cancer samples were also used to simultaneously image 16 proteins and three mRNA targets using a combination of antibodies

and oligonucleotide probes, respectively<sup>184</sup>. Another variant of IMC was developed, which employed an ion beam to liberate metal ion reporters, known as multiplexed ion beam imaging (MIBI), which increased speed, sensitivity, and resolution, and has been reported to give “super-resolution” images of 5-30nm<sup>185,186</sup>. A variation of IMC was developed, which utilized an ion beam to liberate metal ion reporters, known as multiplexed ion beam imaging (MIBI), which increased speed, sensitivity, and resolution, and has been reported to give “super-resolution” images of 5-30nm<sup>185,186</sup>. Currently, IMC-like strategies have been used successfully for cellular phenotyping of lesions in multiple sclerosis and lymphoid organs, primarily at a tissue-level rather than subcellular level<sup>27,28</sup>. Yet their capabilities for providing such resolution are coming into fruition.

### 1.2.2. Mass spectrometry imaging (MSI)

Mass cytometry may be confused with MS-imaging (MSI), though MSI differs in instrumentation and does not require heavy isotope derivatized antibodies labelling. As with IMC, laser ablation is used to ionize individual “pixels” of a sample, with each pixel having a corresponding label-free spectrum, which allows deeper coverage of molecules than IMC (Figure 2b). However, the technique suffers from poor sensitivity and resolution (commercial instruments ranging 5-20  $\mu\text{m}$ ), so is predominantly only useful for macroscopic imaging where subcellular resolution is not in the scope of the experiment<sup>187-189</sup>. Though, hybrid MS setups have allowed this technology to improve its resolution. For example, researchers mixed and matched ion sources, such as atmospheric pressure and laser-induced post-ionization (MALDI-2) sources, coupled to orbitrap analyzers to achieve 1.4  $\mu\text{m}$  and <1  $\mu\text{m}$  resolution, respectively<sup>188,190</sup>. Due to MS vulnerability to contaminants, a lot of sample preparation methods, such as fixatives, are incompatible with this method and often flash-freezing is preferential, but further MS-friendly methods are under investigation<sup>187</sup>. Currently, MSI still suffers from shortfalls in achieving subcellular resolution so there is limited discussion in this review and more comprehensive details of MSI can be found in<sup>187-189</sup>. Arguably, MSI has yet to be fully integrated within subcellular -omics workflows due to limited resolution but advances in the technologies associated with the approach promise to improve the general utility of the MSI. Currently, MSI has been considered for tissue-level intraoperative imaging, particularly on difficult to image difficult to measure biomarkers in pancreatic adenoma<sup>29,191</sup>. What makes this approach particularly exciting is that it can be applied to any molecules that can be ionized, which include proteins, metabolites, or lipids.

## 2. Sequencing-based methods in spatial transcriptomics and proteomics

In the post-genomic era, advances in MS- and RNA-sequencing (RNA-seq)-based technology have allowed researchers to simultaneously quantify thousands of proteins and RNA species in whole cells and tissues. Along with the concurrent advancement of computational tools, powerful spatial -omics workflows can analyze the structure and molecular composition of specific or several subcellular compartments in one experiment. The methods in this section provide spatially-enriched samples of proteins or RNA on a subcellular-level that are measured downstream using MS or RNA-seq. Generally, these methods eliminate *in situ* spatial information during sample preparation and capture “bulk” information of all cells within a given sample. Therefore, achieving single-cell information using the following methods is still challenging, particularly in proteomics due to the inability to amplify proteins<sup>192-198</sup>. Details on the type of MS and RNA-seq approaches that could be coupled with the methods in this section are reviewed in<sup>199-201</sup>.

## 2.1 Biochemical separation

Established techniques that enrich or isolate cellular structures by their physicochemical properties have been in use for decades. Typically, subcellular distribution of molecules was assessed using target-specific enzymatic assays <sup>202</sup>, whereas modern techniques employ robust quantitative sequencing using RNA-seq and mass spectrometry (MS) <sup>199–201</sup>.

### 2.1.1. Basic centrifugation- and detergent-based fractionation

Centrifugation is one of the simplest methods to separate organelles based on their size, density, and shape. Organellar preparations using centrifugation date back to the late 1800s, initially to isolate nuclei <sup>203–205</sup>. Today, there are two generalized categories of centrifugal organellar fractionation, sedimentation, and equilibrium density centrifugation. These result in either an enriched pellet at the base of the tube or at the organelle's equivalent density within a sucrose (or equivalent) gradient, respectively. When coupled with current sequencing technologies, these enrichment strategies are powerful for exploring subcellular composition.

Early spatial proteomics studies focused on purification of a singular organelle of interest, giving insights into the molecular composition of many cellular compartments, such as the nucleolus, nucleus, nuclear pore, and mitochondria, across many cell/tissue types and models <sup>206–209</sup>. However, purifying subcellular compartments is challenging due to co-fractionation with other components of the cell, due to organelles having overlapping biochemical and biophysical properties and their constant interaction with one another. “Subtractive” or “differential” approaches account for this “contamination” or interactions. These methods involve quantitative comparisons of technically equivalent non-enriched fractions against organelle enriched fractions (Figure 3a). Proteins only detected or highly enriched in the organelle-enriched fractions are assigned to that organelle of interest. This strategy has provided valuable information on the subcellular proteomes of the human spliceosome <sup>210,211</sup>, rodent liver nuclear envelope <sup>212</sup>, rat lung endothelial cell plasma membrane and caveolae <sup>213,214</sup>, plus multiple subcellular niches in *S. cerevisiae* and other yeasts using diverse enrichment approaches <sup>215–222</sup>. However, despite accounting for contaminants, it is still difficult to confidently identify organellar proteins, as the composition of any co-fractionating organelle will be erroneously assigned to the organelle of interest. In addition, this technique is not always appropriate for multi-localized molecules or dynamic studies <sup>223</sup>. Coupling of subtractive proteomics with machine learning has improved classification of organellar proteomes <sup>224,225</sup>, which somewhat mitigated this issue by providing more robust statistical comparison between enriched and non-enriched fractions. These strategies have been used to establish biological functions and confident inventories of organellar proteomes, such as the mitochondria, peroxisome, lysosome, and autophagosome <sup>226–231</sup>. Such studies can be particularly useful for poorly characterized species, such as eukaryotic parasites, with the intention to aid biological understanding and pharmacological developments. For example, this strategy was used to assess the proteins involved in the mitochondrial “importome” of *Trypanosoma brucei* by coupling with RNA interference of a key translocase <sup>30</sup>.

Organelles can also be enriched using different detergent-containing buffers with increasing solubilization capacity to sequentially extract molecules from distinct parts of the cell <sup>232</sup>. For instance, the use of digitonin to permeabilize the plasma membrane or NP-40 to release contents of double-membrane organelles. The most popular workflow in proteomics achieves subcellular separation of

the cytosol, nucleus, cytoskeleton, and membranous compartments (such as those found in the secretory pathway) <sup>233</sup>. Modified protocols can further distinguish between DNA-associated and soluble nuclear proteins, or insoluble proteins in the cytosolic, nuclear, and membrane-bound components <sup>234</sup>. This approach was implemented in a phosphoproteomics study to resolve three crude subcellular compartments with a very limited amount of starting material <sup>235</sup>. Notably, detergent enrichment workflows have the advantage of preserving the cytoskeletal network, which is prone to fragmentation in centrifugal fractionation <sup>233</sup>. Differential detergent extraction is primarily reserved for proteomic studies. However, it has been used for studying polysomal RNA and in a two-step detergent protocol to investigate co-translational trafficking of mRNA from cytosolic polysomes to ER-bound polysomes <sup>236,237</sup>.

The development of equivalent biochemical fractionation methods to determine subcellular RNA localization is, in comparison, limited. Several studies use basic cell fractionation via centrifugation and detergent lysis followed by RNA-seq to infer transcript subcellular enrichment <sup>31,238,239</sup>. A sequential detergent strategy was employed to map spatial dynamics of RNA between the cytosol, nucleoplasm, and chromatin in inflammatory-stimulated macrophages by assessing the relative enrichment of transcripts in the different fractions to gain insights into proinflammatory gene regulation <sup>31</sup>. Similar protocols have been applied to obtain a static distribution of transcripts within these compartments in K562 and HEK293 cells <sup>238,239</sup>. Generally, subcellular enrichment can be assessed via western blot by using antibodies against corresponding organellar proteins <sup>240</sup>. However, this assumes that RNA species co-fractionate with their associated cellular compartment in the same way as proteins. Though, proteins and RNA behave differently under centrifugation <sup>241,242</sup>. This may be due to the variability of sedimentation rates and aggregation of RNA molecules which is dependent on the concentration of salts they are present in. This behavior of RNA is believed to have caused artefacts in early attempts at RNA purification and fractionation, leading to controversy over estimates of sedimentation coefficients for RNA molecules <sup>243–245</sup>. The cause of these artefacts may be explained by the size difference between proteins and mRNA, where the transcripts are orders of magnitude larger than proteins <sup>246</sup>. Cytosolic RNA may have similar buoyancy and sedimentation as some subcellular niches, particularly if they aggregate, leading to erroneous assignments of RNA species to various subcellular compartments. These protocols may be improved by including a similar quality control step involving qPCR for transcripts with known localizations.

Furthermore, the methods focus on enriching only a few compartments, which is also a problem shared with the subtractive proteomics methods mentioned previously <sup>31,238,239</sup>. The limited cellular compartment coverage for assessing RNA localization is partially addressed by the CeFra-seq method, which covered five fractions: nuclear, cytosolic, endomembrane, insoluble, and extracellular <sup>247</sup>. In this workflow, differential and density centrifugation were coupled with targeted detergent permeabilization <sup>241,247</sup>. The intention of this method is to enable measurement of global changes in RNA trafficking upon genetic or environmental stimuli. While these methods have provided important contributions to the understanding of components and functions of cellular architecture, they were too crude to resolve multiple organelles within the same experiment, particularly those compartments that are biophysically similar or highly interconnected, such as the secretory pathway <sup>235,248</sup>.

### 2.1.2. Protein/RNA correlation profiling

To address the limitations of more reductive centrifugation methods, correlation profiling was developed based on the principles of Christian de Duve, where localization of proteins can be determined without organellar purification <sup>202</sup>. Protein correlation profiling (PCP) involves organellar

enrichment using a density gradient alongside quantitative MS to measure abundance of peptides across the gradient. Localization of proteins is then inferred by comparing their gradient distribution patterns to those of known organelle marker proteins, usually performed using computational machine learning strategies, such as support vector machine (SVM) classification (Figure 3b)<sup>32,249–251</sup>. PCP was originally applied to single compartments of interest, such as the centrosome, peroxisome, lipid droplet, proteasome in various model organisms<sup>252–255</sup>. The technique was expanded for global organelle analyses in multiple mouse tissues<sup>256,257</sup>. Another density gradient centrifugation technique, Localization of Organelle Proteins by Isotope Tagging (LOPIT), employed isotope-coded affinity tagging (ICAT) to multiplex the gradient fractions and map the global subcellular proteome of the *Arabidopsis thaliana* root-derived callus material<sup>258,259</sup>. Since then, LOPIT has evolved alongside isobaric tagging technologies, allowing the study of the subcellular proteomes of diverse model systems, including human cell lines, chicken lymphocytes, and *D. melanogaster* embryos<sup>260–263</sup>. A further evolution of this protocol, hyperplexed LOPIT (hyperLOPIT), used a more complex density gradient to study pluripotent E14TG2a mouse embryonic stem cells and U-2 OS human bone osteosarcoma cells, which demonstrated highest subcellular resolution than any other MS-based spatial proteomics method available to date<sup>60,264</sup>. HyperLOPIT has also been employed to comprehensively map the subcellular organization of *S. cerevisiae*, cyanobacterium (*Synechocystis*), and *Toxoplasma gondii*<sup>265–267</sup>. The method has also been coupled to FFE (see section 1.1.3.) to analyze the protein composition of Golgi sub-compartments in *A. thaliana* cell-suspension cultures<sup>268</sup>. These comprehensive datasets are designed to provide holistic catalogues of system-wide proteomes to provide biological insight of organellar components, as well as the option to compare between systems and perturbations.

A similar complex density gradient was used to perform cell-wide temporal analysis of subcellular composition during human cytomegalovirus (HCMV) infection, capturing protein dynamics, providing unprecedented understanding of the organellar architecture of host cells during infection<sup>32</sup>. Differential centrifugation and/or detergent strategies can also be coupled to this correlation profiling approach, which the following methods utilize: Dynamic Organellar Maps (DOM), LOPIT-DC, and SubCellBarCode<sup>250,261,269</sup>. These methods vary in separation, labelling and analysis protocols that range in resolving power and come with their own advantages or limitations, depending on the study design and biological questions being addressed, offering flexible and customizable options for researchers. These methods can achieve high coverage, often >8,000 proteins, and some of these methods have achieved sub-organellar resolution, such as resolving the endoplasmic reticulum–Golgi intermediate compartment (ERGIC), ribosomal subunits, chromatin, and sub-nuclear compartments<sup>250,264,269</sup>. DOM has been used to further investigate protein trafficking after perturbation with compounds that enhance antigen import via lysosomal trapping<sup>33</sup>.

The big challenge of these correlation profiling experiments is the data analysis of their complex, multi-dimensional datasets. Answering apparently trivial questions can become challenging, specifically identifying protein translocation events and proteins localized in multiple cellular compartments. However, novel computational models, such as TRANSPIRE, BUNDLE, and MR scoring, have ventured to address questions on protein dynamics<sup>250,270,271</sup>. T-augmented Gaussian mixture model (TAGM) approaches have been developed to tackle questions on multiply localized proteins<sup>272,273</sup>. A recent evaluation of some of these approaches can be found in<sup>274</sup>.

In transcriptomics, sucrose gradients are frequently used to assess mRNA association with polysomes, which can be separated into cytosolic- and ER-polysome-bound transcripts<sup>237,275,276</sup>. It is thus reasonable to suggest that correlation profiling methods could be adapted for spatial transcriptomics assigning RNA subcellular localization based on distribution profile comparisons to curated 'RNA markers' of known localization. Current methods are typically protein-centric or are based on

separating RNA along with their protein interactors – e.g. within stress granules or ribosomes<sup>277–279</sup>. New approaches, such as ATLAS-seq, have been able to co-sediment different RNA using density gradients coupled to RNA-seq and then use hierarchical clustering to infer subcellular localization of transcripts<sup>242</sup>. When mapping the transcriptome of mouse liver, the authors found that transcripts that co-sediment tended to encode proteins that co-associate, including proteasomal subunit mRNAs. Additionally, alternatively spliced transcripts typically showed differential sedimentation patterns. However, transcripts with similar correlation profiles did not consistently co-localize when interrogated with an orthogonal method such as smiFISH, indicating need for further improvement of such approaches<sup>242</sup>. Thus, there is a current gap in technology and a requirement to develop a spatial transcriptomics technique that truly complements protein correlation profiling.

The key benefit of PCP-based methods are the ability to interrogate all cellular compartments at once and therefore be able to address dynamic and complex biological problems, which the majority of the other techniques discussed in this review struggle to do on the same scale. The primary drawback of all PCP methods is that they capture “average localization” of proteins within a given sample. Therefore, data from samples that are likely to contain a heterogeneous population of cells may be more complicated to interpret. Examples of such samples include tumor cell-lines in different cell cycle stages or cell types within a tissue. Techniques to resolve this, such as cell cycle synchronization or microdissection, reduce input material, which could complicate downstream sample preparation. Another consideration is that mechanical or chemical bursting of the outer membranes of cells is required for PCP, which may cause artefacts, such as lost interactions or leakage from organellar membranes.

### 2.1.3. Electrophoresis-based methods

Another biophysical property that can be exploited for organellar fractionation is charge states. Free-flow electrophoresis (FFE) relies on the same principles of electrophoresis, where particles in a biological sample are separated via their surface charge densities. It is a versatile technique that can separate a variety of charged analytes including low molecular mass organic compounds, proteins, peptides, macromolecular complexes, organelles, and whole cells under native or denaturing conditions in aqueous separation buffers. Notably, FFE-mediated cell fractionation experiments are characterized by fast separation and high sample recovery rates. Furthermore, FFE can be paired with various tools such as specific antibodies, lectins, chemical ligands, and proteases or other enzymes to optimize organelle separation, by introducing subtle changes in the surface charges of certain compartments with minimal disruption to their functional integrity<sup>233,280–283</sup>. FFE has been used in combination with centrifugal separation and MS analysis to resolve subpopulations of organellar networks that are otherwise difficult to capture, including the plasma membrane, components of the ER network, endosomes, lysosomes, phagosomes, peroxisomes, mitochondria, and plant tonoplasts (vacuole membranes)<sup>233,280–287</sup>. De Michele and colleagues were able to assign peripheral membrane proteins which are normally lost in traditional PM enrichment to the Arabidopsis PM, including the entire exocyst complex<sup>34</sup>. This method has been used to separate DNA in a size-dependent manner but has yet to be used to study RNAs, particularly in a subcellular context.

FIFFF (flow field-flow fractionation) is similar to FFE, but instead uses a “cross-flow” system that drives separation in a shape and size dependent manner, providing distinct elution patterns for different sample constituents<sup>281,288</sup>. FIFFF with MS has been used to analyze subcellular structures such as the mitochondria, extracellular vesicles, and lipoprotein particles<sup>288–290</sup>. The technique was used to separate and define a new subpopulation of small EVs termed exomeres, which are selectively

enriched in glycolytic and mTOR signalling proteins compared to larger EVs <sup>291</sup>. This method has also been used in a cell-wide context demonstrating simultaneous separation of multiple human subcellular compartments, albeit with lower resolution than centrifugation strategies <sup>292</sup>.

Much like standard single-cell RNA-seq (scRNA-seq), microfluidics-based electrophoresis has also been exploited for subcellular scRNA-seq. Single-cell integrated nucRNA and cytRNA-sequencing (SINC-seq) captured single-cells using a hydrodynamic trap, followed by selective electrolysis of the plasma membrane to attain intact nuclei <sup>293,294</sup>. The cytoplasmic RNA was then separated based on ionic mobility via electric field activation and used to construct individual RNA-seq libraries. SINC-seq in K652 cells showed that over pseudotime, the differentially expressed genes in cytRNA versus nucRNA showed less correlation in gene expression ~~captured the dynamics of cytRNA and nucRNA expression when~~ histone acetylation was perturbed using sodium butyrate. Impressively, this method resulted in only a 5.3% drop in reads when compared to standard scRNA-seq. In NanoSINC-seq, the microfluidic fractionation is coupled to Nanopore cDNA sequencing to compare isoform diversity between the cytoplasm and nucleus <sup>295</sup>. There is potential for the systems harnessed in FFE and FIFFF to be modified for single-cell sequencing of other organelles <sup>294</sup>.

## 2.2. Proximity labelling-based methods

Proximity labelling was originally developed for capturing the interactomes of specific proteins *in vivo* followed by downstream purification. Therefore, mapping interactomes of multiple bait proteins, can essentially capture the “local” spatial proteome or transcriptome of each bait <sup>296–298</sup>. Proximity labelling workflows consist of fusing a bait protein of interest to an enzyme, typically a biotin ligase or a peroxidase, which covalently labels proteins and RNA in the immediate vicinity of the bait with a small, exogenous substrate, typically biotin (Figure 3c). Due to the short half-life of the substrate, only molecules within a few tens to hundreds of nanometers of the bait are labelled. Therefore, there is no reliance on direct physical interactions <sup>296</sup>.

Proximity labelling overcomes shortcomings of traditional affinity purification protocols where interactions can be disrupted during sample preparation, with the caveat of contamination with promiscuous cellular components, such as diffusing components of the cytosol, and background biotinylation. However, these promiscuous molecules can be mostly accounted for with the appropriate controls and by referring to a contaminant repository for proximity labelling experiments, known as the CRAPome (<http://crapome.org>) <sup>299,300</sup>. Other considerations of proximity labelling include the variable elution of enriched molecules from affinity matrices, changes in expression, localization or function of the fused bait protein, and that amino acid residues targeted for biotinylation by the fused enzyme must be present on the surface of the proximal proteins. Therefore, proteins or RNA lacking these residues on the surface of their structure would be missed and it has been shown that proximity labelling favors intrinsically disordered regions where these residues are more likely to be exposed <sup>301</sup>. Additionally, compartments with highly dynamic and soluble molecules, such as the cytosol or nucleoplasm, are difficult to target using this strategy, because these subcellular niches do not offer a small, defined, membrane-enclosed space to which the bait protein can be specifically targeted, resulting in high rates of non-specific biotinylation. Multi-bait strategies for assessing subcellular localization across several organelles has started gaining popularity. However, there is no guarantee that the fused enzyme will have comparable activity in these different subcellular locations. This has implications on the quantitation of the data, where true proportions of the proteins or RNA across these compartments cannot be deduced and rather a qualitative list of species in those locations are produced. Additionally, this method is limited to biological systems that

can be genetically engineered. Despite this, proximity labelling excels at capturing membrane-bound organelles and proteins associated with insoluble cellular structures, such as various cytoskeletal components, which are challenging to isolate and reliably analyze with alternative methods<sup>302</sup>. Here, we cover the key enzymes used and how they are applied to both spatial transcriptomic and proteomic studies. The strategy has been invaluable for uncovering biological networks, albeit in a spatially restricted manner. Moreover, powerful, cell-wide proximity tagging studies have recently started to emerge, as well as transcript-capturing approaches, indicating what the future holds for this approach<sup>297</sup>.

### 2.2.1. BioID-based methods

BioID is a proximity labelling method that uses biotin ligases. Originally, wild-type BirA enzyme from *E.coli* was used, which catalyzes biotinylation in the presence of ATP (adenosine triphosphate) on molecules containing a biotin acceptor peptide (BAP) sequence<sup>303</sup>. This restricted targeting to proteins with BAP regions and exposed lysine residues. BioID overcame this limitation by engineering BirA to promiscuously biotinylate proteins<sup>304</sup>. This allowed for improved labelling efficiency, selectivity, faster incubation times, and higher signal-to-noise ratios. The newer generation of the BirA enzymes, such as TurboID or MiniTurbo, can achieve labelling within minutes rather than hours, enabling study of rapid and dynamic cellular processes<sup>35,305</sup>. These optimized enzymes have been recently reviewed in<sup>296</sup>.

BioID has been successfully applied to a range of interactions on a complex-, organellar- and dynamic-level in a variety of cell and tissue types, as well as entire organisms. In mammalian cells, BioID-based labelling strategies have provided insights into macromolecular complexes or subcellular niches, including the nuclear lamina, nuclear pore complex, nucleosome complexes, mediator transcription regulation complex, endoplasmic reticulum-peroxisome contacts, and focal adhesions<sup>299,304,306–310</sup>. In addition, BioID has been used in applications beyond immortalized mammalian cell lines. For example, the approach was used to study three unique subcellular niches in *Trypanosoma brucei*, the basal body, flagellum and bilobe, plus entire organisms, specifically flies and worms<sup>35–38</sup>, revealing novel organellar components of these organisms that are otherwise difficult to capture. Multi-bait studies were used to identify novel of the centrosome-cilium interface and ciliogenesis, to unpick the phosphorylation regulation in the Hippo signaling pathway, and the interactome of P-bodies and stress granules during normal and stressed conditions, using 58, 19, and 139 baits, respectively<sup>311–313</sup>. Impressively, examples of cell-wide BioID experiments have recently emerged that have captured 26,527 and 35,902 interactions located within 21 and 32 distinct cellular features, respectively, with the former identifying a further 9,390 interactions when coupled with affinity-purification MS (AP-MS)<sup>297,298</sup>. These studies demonstrate sub-organellar resolution that is difficult to achieve via correlation profiling.

Despite working well in proteomics, BioID has not yet been adapted to directly label RNA for spatial transcriptomics studies. Studies that do involve targeting BirA enzymes to RNA are typically focused on determining the protein interactors of specific transcripts, rather than biotinylation of RNA in certain organelles<sup>314–316</sup>. However, pairing BioID with ribosome profiling enabled identification of translated transcripts in a specific subcellular context. In this method, the biotin ligase was expressed as a fusion protein to localize it to the subcellular niche of interest, and ribosomes were expressed with a protease-cleavable biotin acceptor peptide (AviTag). With addition of biotin, only ribosomes in the vicinity of the ligase were biotinylated, after which they were isolated, and the translating RNA was sequenced. Proximity-specific ribosome profiling was used to profile the translome of the ER

and mitochondria in yeast and human cells<sup>317–319</sup>. Additionally, BioID was adapted for *in vivo* purposes to study the inhibitory neuronal network and synapse formation in live mice identifying both known and novel proteins involved in the hyperpolarization process with some linked to neurodegenerative diseases<sup>39,40</sup>.

### 2.2.2. APEX-based methods

APEX is also a proximity labelling technique that allows for the mapping of the proteome and transcriptome with spatial and temporal resolution. Instead of the BirA derivatives used for BioID, APEX uses a modified soybean-derived ascorbate peroxidase. The enzyme catalyses oxidation of a supplied biotin derivative (usually biotin-phenol) and generates spatially confined, short-lived biotin-phenoxy radicals that react with electron-rich side chains of amino acids (such as Tyr, Trp, His, and Cys), resulting in covalent biotin labelling of proteins and RNA in the near vicinity. The primary advantage of APEX over BioID is that the tagging reaction is faster, allowing for dynamic experiments capturing discrete time-points. Detailed comparisons of APEX and BioID, including their novel variants, can be found in recent reviews<sup>296,320</sup>.

The APEX-based system coupled to quantitative proteomics has been instrumental in exploring various subcellular compartments and networks. For example, its use in human cells allowed analysis of various mitochondrial sub-compartments, the ER membrane, and the endocytic system, as well as investigating interaction of the interactions of bioactive peptides (or “microproteins”)<sup>321–329</sup>. The advantages of APEX over BioID were epitomized in time-resolved studies. For example, APEX has been used to capture the transient interactome of G protein-coupled receptor (GPCR) signaling identifying novel regulators of the associated delta opioid receptor and  $\beta$ 2 adrenoceptor<sup>331,332</sup>. Another temporal study using APEX showed unique and aberrant stress granule dynamics in cells of amyloid lateral sclerosis (ALS), helping identify novel disease-relevant protein candidates<sup>330–332</sup>. An improved variant, APEX2, has allowed increased catalytic activity and labelling sensitivity compared to APEX. In contrast to the original monomeric APEX protein, APEX2 has the capacity of forming dimeric complexes, which has been shown to improve the activity as well as stability of ascorbate peroxidase enzymes<sup>323</sup>.

APEX technology has been recently adapted for spatial transcriptomics. There are two main applications of APEX for determining RNA subcellular localization: proximity labelling of protein crosslinked to RNA, and proximity labelling of the RNA nucleosides directly. The first method published was APEX-RIP, which combined APEX-dependent *in-situ* protein biotinylation with formaldehyde crosslinking of the labelled proteins to RNA in specific compartments or organelle interfaces. In a proof-of-concept study, APEX-RIP mapped transcripts associated with the nucleus, mitochondrial matrix, ER membrane, and cytosol in human cells<sup>333</sup>. A related method, Proximity-CLIP, achieved simultaneous profiling of both free and RNA-bound proteins at specific subcellular locations by combining APEX2-dependent protein tagging, an evolution of the original APEX enzyme, with UV-mediated protein-RNA crosslinking<sup>334</sup>. APEX-RIP has the benefit of not requiring RNA labelling whereas Proximity-CLIP uses 4-thiouridine labelling of RNA to enhance crosslinking efficiency<sup>334</sup>. However, APEX-RIP suffers from poor specificity in membrane-less regions of the cell due to the use of formaldehyde crosslinking<sup>333</sup>. These problems were addressed by an alternative implementation of APEX for RNA localization called APEX-seq<sup>3,41</sup>. This method took advantage of the discovery that the APEX2 enzyme directly labels nearby RNA, as well as proteins. The biotinylated RNA could then be affinity purified with streptavidin beads and sequenced. The APEX-Seq approach was used to generate an atlas of human RNA localization covering nine different subcellular niches and to probe the spatial organization of transcripts associated with translation initiation complexes, as well as repressive RNA

granules<sup>18,41</sup>. This study unveiled the dynamic, varied and stress-type-dependent nature of stress granules. It is important to remember that APEX-seq uses the same biotin-phenol substrate as standard protein proximity labelling, and subsequently proteins are also biotinylated. Recently, it has been shown that using different biotin substrates, such as biotin-aniline, improved labelling efficiency of nucleic acids versus proteins<sup>335</sup>. It is worth noting that biotin-phenol is toxic to cells at higher concentrations, which can limit its use in certain model systems, such as in tissue or whole organisms. BioID may be more appropriate in these experimental scenarios. Despite this, APEX-based labelling was adapted for application to *in vivo* model systems to map various subcellular compartments, such as nucleus-, mitochondrial matrix-, and Golgi apparatus-associated proximity networks, in live yeast, *Caenorhabditis elegans*, and *Drosophila*<sup>336–341</sup>.

### 3. Future Prospects

This review demonstrates there are several options available to researchers to address biological questions concerning the subcellular localization and trafficking of proteins and transcripts. However, the technical challenges can still be vast and differ between transcriptomics and proteomics, as well as the biological system and question in hand, which is the intrinsic reason why there is lack of a one-size-fits-all approach. Though there have been attempted hybrid methods where multiple probing across molecular species is possible, such as in IMC, these still often lack the same coverage or resolution as sequencing or microscopy<sup>184,187</sup>. Here we briefly discuss the methods that attempt to address fundamental limitations hindering the field and what the future may hold for subcellular omics.

Mass spectrometry-, RNA-seq-, and imaging-based methods are continuing to make great advances in improving coverage and resolution. Instrumentation advances have allowed researchers to push the boundaries in subcellular resolution and coverage<sup>200,201,342</sup>. Additionally, complex multiplexing strategies, such as DNA-PAINT, and super-resolution imaging are becoming standard practice in more research laboratories<sup>72,73,75,97,99</sup>. When combined, these techniques are not limited to studying localisation in membrane-bound organelles but can be extended to imaging biomolecular condensates<sup>23</sup>. Automation and artificial intelligence analyses are enabling deconvolution of vast quantities of imaging data to determine the extent of single-cell subcellular heterogeneity<sup>10,60,343</sup>, plus detection of translocations and multilocalized molecules<sup>270,271,273</sup>. Due to the impressive advances in methodologies and data analytics, scRNAseq is now relatively straightforward<sup>194,344</sup>. However, gaining subcellular scRNA-seq information is more challenging, with poor spatial resolution or poor read coverage, as seen from SINC-seq and FISSEQ, respectively<sup>294,345</sup>. Single-cell proteomics using MS is similarly in its infancy, hindered by the lack of a PCR-like amplification available for proteins. Despite emerging strategies for single-cell proteomics and sensitivity of MS instrumentation, there is still a long way to go to extract subcellular information on a single-cell level<sup>196,346</sup>. Additionally, many of these methods are enrichment-based and do not provide absolute quantitation, rather relative quantitation.

In addition to differential expression and subcellular distribution of RNA and protein, there are other dimensions of molecular biology that can influence physiological processes, subcellular localization and states of cells, such as molecular structure, stability, turnover, interactions, and modification (e.g. PTMs and splicing). Current approaches often do not capture this information, which is vital to understanding control of localization and molecular roles of proteins and RNAs. Capturing this information in addition to spatial data may deconvolute some of the ambiguity found between datasets as it is still unclear how these molecular characteristics, such as PTMs, influence protein and RNA distribution. This is becoming more achievable with improved MS technologies and protocols. For example, recent improvements in cross-linking mass spectrometry (XL-MS) technology allows inference of interactions and protein structures/folding, via forming identifiable chemical bridges between residues<sup>347–350</sup> and ion-mobility technology is improving throughput of “native, in-tact” proteins, as well as improved identification of multiply phosphorylated peptides<sup>351–353</sup>. Enrichment protocols and commercial kits are also aiding PTM analysis via MS with reduced starting material<sup>354,355</sup>. Nanopore sequencing is providing more straightforward and accurate RNA modification data, alongside other methods for RNA modification analysis, which include variants of next generation sequencing and LC-MS/MS analysis<sup>200,356,357</sup>. Additionally, straightforward and simultaneous enrichment of proteins, RNA and RBPs with crosslinking and phase separation, with reduced starting material compared to conventional methods, such as RNA interactome capture (RIC), is aiding the investigation of RNA-protein interaction across multiple cell types<sup>358,359</sup>. The efficacy of RIC opens the possibility of coupling it with different fractionation strategies to obtain functional maps of interacting molecules, bridging the RNA and protein fields. Harnessing these innovations in a spatial context would unearth new layers of cellular control.

## Conclusion

Coupling omics with localization studies is still largely in its infancy but is rapidly growing due to advancement of sample preparation strategies and equipment reaching a pinnacle with single-molecule tracking, sequencing, and current MS technology. Not only have subcellular -omics technologies aided our insight into global spatial organization (e.g. HPA Cell Atlas), biological processes (e.g. cell cycle and embryonic development) and pathologies (e.g. cancer biology), but are also emerging in diagnostic applications for patients<sup>10,25,26,60,95,343</sup>. The hope is that more cell biologists utilize these methods to enrich their own datasets, as well as contribute to growing repositories, such as UniProt and the HPA Cell Atlas, with the aim to unearth further understanding of the complex, multi-layered mechanisms of biological functions and disease<sup>60,62</sup>.

## Acknowledgements

The authors would like to thank Eneko Villanueva and Mariavittoria Pizzinga from the Medical Research Council (MRC) Toxicology Unit, Cambridge for their valuable suggestions for the manuscript. J.A.C. is funded through a BBSRC iCASE award with AstraZeneca (BB/R505304/1). A.G. was funded through the Alexander S. Onassis Public Benefit Foundation, the Foundation for Education and European Culture (IPEP), the A. G. Leventis Foundation and the Embiricos Trust Scholarship of Jesus College Cambridge. C.S.D. is funded by a Herchel Smith Research Studentship at the University of Cambridge, United Kingdom. O.L.V. is a BBSRC CASE student (BB/R505365/1). K.S.L. is funded by Wellcome Trust (110071/Z/15/Z) and EU Horizon 2020 program INFRAIA project Epic-XS (Project 823839).

## Contributions

Abstract (J.A.C.); Introduction (J.A.C., A.G.); Spatial proteomics methods (J.A.C., A.G.); Spatial transcriptomics methods (J.A.C., C.S.D.); Future prospects (J.A.C.); Conclusion (J.A.C.); Figures (O.L.V., J.A.C.); Overseeing and editing (J.A.C., K.S.L.).

## References

1. Kloc, M., Zearfoss, N. R. & Etkin, L. D. Mechanisms of subcellular mRNA localization. *Cell* vol. 108 533–544 (2002).
2. Dennerlein, S., Wang, C. & Rehling, P. Plasticity of Mitochondrial Translation. *Trends in Cell Biology* vol. 27 712–721 (2017).
3. Fazal, F. M. *et al.* Atlas of Subcellular RNA Localization Revealed by APEX-Seq. *Cell* **178**, 473–490.e26 (2019).
4. Cajigas, I. J. *et al.* The Local Transcriptome in the Synaptic Neuropil Revealed by Deep Sequencing and High-Resolution Imaging. *Neuron* **74**, 453–466 (2012).
5. Moor, A. E. *et al.* Global mRNA polarization regulates translation efficiency in the intestinal epithelium. *Science (80-. )*. **357**, 1299–1303 (2017).
6. Hentze, M. W., Castello, A., Schwarzl, T. & Preiss, T. A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology* vol. 19 327–341 (2018).
7. Sim, S. *et al.* The subcellular distribution of an RNA quality control protein, the ro autoantigen, is regulated by noncoding y RNA binding. *Mol. Biol. Cell* **20**, 1555–1564 (2009).
8. Sim, S. *et al.* The zipcode-binding protein ZBP1 influences the subcellular location of the Ro 60-kDa autoantigen and the noncoding Y3 RNA. *RNA* **18**, 100–110 (2012).
9. Yang, J. & Kornbluth, S. All aboard the cyclin train: Subcellular trafficking of cyclins and their CDK partners. *Trends in Cell Biology* vol. 9 207–210 (1999).
10. Mahdessian, D. *et al.* Spatiotemporal dissection of the cell cycle with single-cell proteogenomics. *Nature* **590**, 649–654 (2021).
11. Guo, L. & Shorter, J. Biology and pathobiology of TDP-43 and emergent therapeutic strategies. *Cold Spring Harb. Perspect. Med.* **7**, a024554 (2017).
12. Sehgal, P. B. & Lee, J. E. Protein trafficking dysfunctions: Role in the pathogenesis of pulmonary arterial hypertension. *Pulmonary Circulation* vol. 1 17–32 (2011).
13. Kopito, R. R. Biosynthesis and degradation of CFTR. *Physiol. Rev.* **79**, S167–S173 (1999).
14. Cheng, S. H. *et al.* Defective intracellular transport and processing of CFTR is the molecular basis of most cystic fibrosis. *Cell* **63**, 827–834 (1990).
15. Lukacs, G. L. *et al.* Conformational maturation of CFTR but not its mutant counterpart ( $\Delta$ F508) occurs in the endoplasmic reticulum and requires ATP. *EMBO J.* **13**, 6076–6086 (1994).
16. Ren, H. Y. *et al.* VX-809 corrects folding defects in cystic fibrosis transmembrane conductance regulator protein through action on membrane-spanning domain 1. *Mol. Biol. Cell* **24**, 3016–3024 (2013).
17. Lu, A. X. *et al.* Integrating images from multiple microscopy screens reveals diverse patterns of change in the subcellular localization of proteins. *Elife* **7**, 10.7554/eLife.31872 (2018).
18. Fazal, F. M. & Chang, H. Y. Subcellular Spatial Transcriptomes: Emerging Frontier for Understanding Gene Regulation. *Cold Spring Harb. Symp. Quant. Biol.* **84**, 31–45 (2020).
19. Christopher, J. A. *et al.* Subcellular proteomics. *Nat. Rev. Methods Prim.* **1**, 32 (2021).
20. Lundberg, E. & Borner, G. H. H. Spatial proteomics: a powerful discovery tool for cell biology.

- Nature Reviews Molecular Cell Biology* vol. 20 285–302 (2019).
21. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).
  22. Hikmet, F. *et al.* The protein expression profile of ACE2 in human tissues. *Mol. Syst. Biol.* **16**, e9610 (2020).
  23. Geiger, F. *et al.* Liquid–liquid phase separation underpins the formation of replication factories in rotaviruses. *EMBO J.* e107711 (2021) doi:10.15252/embj.2021107711.
  24. Wang, H. *et al.* SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Res.* **18**, 290–301 (2008).
  25. Grimwade, L. F., Fuller, K. A. & Erber, W. N. Applications of imaging flow cytometry in the diagnostic assessment of acute leukaemia. *Methods* vol. 112 39–45 (2017).
  26. Giesen, C. *et al.* Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
  27. Park, C. *et al.* The landscape of myeloid and astrocyte phenotypes in acute multiple sclerosis lesions. *Acta Neuropathol. Commun.* **7**, 130 (2019).
  28. Durand, M. *et al.* Human lymphoid organ cDC2 and macrophages play complementary roles in T follicular helper responses. *J. Exp. Med.* **216**, 1561–1581 (2019).
  29. Huang, K. T. *et al.* Rapid Mass Spectrometry Imaging to Assess the Biochemical Profile of Pituitary Tissue for Potential Intraoperative Usage. *Adv. Cancer Res.* **134**, 257–282 (2017).
  30. Peikert, C. D. *et al.* Charting organellar importomes by quantitative mass spectrometry. *Nat. Commun.* **8**, 1–14 (2017).
  31. Bhatt, D. M. *et al.* Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150**, 279–290 (2012).
  32. Beltran, P. M. J., Mathias, R. A., Cristea, I. M. & Beltran, J. A Portrait of the Human Organelle Proteome In Space and Time during Cytomegalovirus Infection. *Cell Syst.* **3**, 361–373 (2016).
  33. Kozik, P. *et al.* Small Molecule Enhancers of Endosome-to-Cytosol Import Augment Anti-tumor Immunity. *Cell Rep.* **32**, 107905 (2020).
  34. De Michele, R. *et al.* Free-Flow Electrophoresis of Plasma Membrane Vesicles Enriched by Two-Phase Partitioning Enhances the Quality of the Proteome from Arabidopsis Seedlings. *J. Proteome Res.* **15**, 900–913 (2016).
  35. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–898 (2018).
  36. Dang, H. Q. *et al.* Proximity interactions among basal body components in trypanosoma brucei identify novel regulators of basal body biogenesis and inheritance. *MBio* **8**, (2017).
  37. Hu, H., Zhou, Q. & Li, Z. SAS-4 protein in Trypanosoma brucei controls life cycle transitions by modulating the length of the flagellum attachment zone filament. *J. Biol. Chem.* **290**, 30453–30463 (2015).
  38. Morriswood, B. *et al.* Novel bilobe components in Trypanosoma brucei identified using proximity-dependent biotinylation. *Eukaryot. Cell* **12**, 356–367 (2013).
  39. Uezu, A. *et al.* Identification of an elaborate complex mediating postsynaptic inhibition.

- Science* (80- ). **353**, 1123–1129 (2016).
40. Spence, E. F. *et al.* In vivo proximity proteomics of nascent synapses reveals a novel regulator of cytoskeleton-mediated synaptic maturation. *Nat. Commun.* **10**, 1–16 (2019).
  41. Padrón, A., Iwasaki, S. & Ingolia, N. T. Proximity RNA Labeling by APEX-Seq Reveals the Organization of Translation Initiation Complexes and Repressive RNA Granules. *Mol. Cell* **75**, 875-887.e5 (2019).
  42. Stadler, C., Skogs, M., Brismar, H., Uhlén, M. & Lundberg, E. A single fixation protocol for proteome-wide immunofluorescence localization studies. *J. Proteomics* **73**, 1067–1078 (2010).
  43. Mattiuzzi Usaj, M. *et al.* High-Content Screening for Quantitative Cell Biology. *Trends in Cell Biology* vol. 26 598–611 (2016).
  44. Wollman, R. & Stuurman, N. High throughput microscopy: From raw images to discoveries. *J. Cell Sci.* **120**, 3715–3722 (2007).
  45. Eliceiri, K. W. *et al.* Biological imaging software tools. *Nat. Methods* **9**, 697–710 (2012).
  46. Aspelmeier, T., Egner, A. & Munk, A. Modern Statistical Challenges in High-Resolution Fluorescence Microscopy. *Annu. Rev. Stat. Its Appl.* **2**, 163–202 (2015).
  47. Khater, I. M., Nabi, I. R. & Hamarneh, G. A Review of Super-Resolution Single-Molecule Localization Microscopy Cluster Analysis and Quantification Methods. *Patterns* vol. 1 100038 (2020).
  48. Wu, Y. Le, Tschanz, A., Krupnik, L. & Ries, J. Quantitative Data Analysis in Single-Molecule Localization Microscopy. *Trends in Cell Biology* vol. 30 837–851 (2020).
  49. Pineau, I., Barrette, B., Vallières, N. & Lacroix, S. A novel method for multiple labeling combining in situ hybridization with immunofluorescence. *J. Histochem. Cytochem.* **54**, 1303–1313 (2006).
  50. Chaudhuri, A. D., Yelamanchili, S. V. & Fox, H. S. Combined fluorescent in situ hybridization for detection of microRNAs and immunofluorescent labeling for cell-type markers. *Front. Cell. Neurosci.* **7**, 160 (2013).
  51. VanZomeran-Dohm, A., Flannery, E. & Duman-Scheel, M. Whole-mount in situ hybridization detection of mRNA in GFP-marked drosophila imaginal disc mosaic clones. *Fly (Austin)*. **2**, 323–325 (2008).
  52. Zaglia, T. *et al.* Optimized protocol for immunostaining of experimental GFP-expressing and human hearts. *Histochem. Cell Biol.* **146**, 407–419 (2016).
  53. Oliva, J. & Swann, J. W. Fluorescence in situ hybridization method for co-localization of mRNA and GFP. *Biotechniques* **31**, 74–81 (2001).
  54. Kilgore, J. A., Dolman, N. J. & Davidson, M. W. A review of reagents for Fluorescence Microscopy of Cellular Compartments and Structures, part II: Reagents for non-vesicular organelles. *Curr. Protoc. Cytom.* **66**, 12.31.1-12.31.24 (2013).
  55. Zhu, H., Fan, J., Du, J. & Peng, X. Fluorescent Probes for Sensing and Imaging within Specific Cellular Organelles. *Acc. Chem. Res.* **49**, 2115–2126 (2016).
  56. Baker, M. Blame it on the antibodies. *Nature* **521**, 274–276 (2015).
  57. Schnell, U., Dijk, F., Sjollem, K. A. & Giepmans, B. N. G. Immunolabeling artifacts and the

- need for live-cell imaging. *Nature Methods* vol. 9 152–158 (2012).
58. Gut, G., Herrmann, M. D. & Pelkmans, L. Multiplexed protein maps link subcellular organization to cellular states. *Science* (80-. ). **361**, (2018).
  59. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* (80-. ). **347**, (2015).
  60. Thul, P. J. *et al.* A subcellular map of the human proteome. *Science* (80-. ). **356**, eaal3321 (2017).
  61. Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. *Science* (80-. ). **357**, (2017).
  62. Bateman, A. *et al.* UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).
  63. Zahn-Zabal, M. *et al.* The neXtProt knowledgebase in 2020: Data, tools and usability improvements. *Nucleic Acids Res.* **48**, D328–D334 (2020).
  64. Carbon, S. *et al.* The Gene Ontology resource: Enriching a GOLD mine. *Nucleic Acids Res.* **49**, D325–D334 (2021).
  65. Vizcaíno, J. A. *et al.* A community proposal to integrate proteomics activities in ELIXIR. *F1000Research* **6**, 875 (2017).
  66. Nilsson, P. *et al.* Towards a human proteome atlas: High-throughput generation of mono-specific antibodies for tissue profiling. *Proteomics* **5**, 4327–4337 (2005).
  67. Uhlén, M. *et al.* A human protein atlas for normal and cancer tissues based on antibody proteomics. *Mol. Cell. Proteomics* **4**, 1920–1932 (2005).
  68. Älgenäs, C. *et al.* Antibody performance in western blot applications is context-dependent. *Biotechnol. J.* **9**, 435–445 (2014).
  69. Skogs, M. *et al.* Antibody Validation in Bioimaging Applications Based on Endogenous Expression of Tagged Proteins. *J. Proteome Res.* **16**, 147–155 (2017).
  70. Stadler, C. *et al.* Systematic validation of antibody binding and protein subcellular localization using siRNA and confocal microscopy. *J. Proteomics* **75**, 2236–2251 (2012).
  71. Sullivan, D. P. *et al.* Deep learning is combined with massive-scale citizen science to improve large-scale image classification. *Nat. Biotechnol.* **36**, 820–832 (2018).
  72. Lin, J. R., Fallahi-Sichani, M. & Sorger, P. K. Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nat. Commun.* **6**, 1–7 (2015).
  73. Goltsev, Y. *et al.* Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968-981.e15 (2018).
  74. Lin, J. R. *et al.* Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *Elife* **7**, 10.7554/eLife.31657 (2018).
  75. Jungmann, R. *et al.* Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat. Methods* **11**, 313–318 (2014).
  76. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nat. Protoc.* **12**, 1198–1228 (2017).
  77. Doerr, A. RNA antibodies: Upping the ante. *Nat. Methods* **5**, 220 (2008).

78. Ye, J. D. *et al.* Synthetic antibodies for specific recognition and crystallization of structured RNA. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 82–87 (2008).
79. Gall, J. G. & Pardue, M. L. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proc. Natl. Acad. Sci. U. S. A.* **63**, 378–383 (1969).
80. Rudkin, G. T. & Stollar, B. D. High resolution detection of DNA-RNA hybrids in situ by indirect immunofluorescence [29]. *Nature* vol. 265 472–473 (1977).
81. Chen, J., McSwiggen, D. & Ünal, E. Single molecule fluorescence in situ hybridization (SmFISH) analysis in budding yeast vegetative growth and meiosis. *J. Vis. Exp.* **2018**, 10.3791/57774 (2018).
82. Ding, D. Q. *et al.* Chromosome-associated RNA–protein complexes promote pairing of homologous chromosomes during meiosis in *Schizosaccharomyces pombe*. *Nat. Commun.* **10**, 1–12 (2019).
83. Titlow, J. S., Yang, L., Parton, R. M., Palanca, A. & Davis, I. Super-resolution single molecule FISH at the drosophila neuromuscular junction. in *Methods in Molecular Biology* vol. 1649 163–175 (Humana Press Inc., 2018).
84. Simon, B., Sandhu, M. & Myhr, K. L. Live FISH: Imaging mRNA in living neurons. *J. Neurosci. Res.* **88**, 55–63 (2010).
85. Oomoto, I. *et al.* ECHO-liveFISH: In vivo RNA labeling reveals dynamic regulation of nuclear RNA foci in living tissues. *Nucleic Acids Res.* **43**, e126 (2015).
86. Wang, H. *et al.* CRISPR-mediated live imaging of genome editing and transcription. *Science (80-. )*. **365**, 1301–1305 (2019).
87. Pichon, X., Lagha, M., Mueller, F. & Bertrand, E. Molecular Cell Technology Review A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges. *Mol. Cell* **71**, 468–480 (2018).
88. Sinnamon, J. R. & Czaplinski, K. RNA detection in situ with FISH-STICs. *RNA* **20**, 260–266 (2014).
89. Wang, F. *et al.* RNAscope: A novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagnostics* **14**, 22–29 (2012).
90. Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat. Biotechnol.* **28**, 1208–1212 (2010).
91. Choi, H. M. T. *et al.* Third-generation in situ hybridization chain reaction: Multiplexed, quantitative, sensitive, versatile, robust. *Dev.* **145**, 10.1242/dev.165753 (2018).
92. Banér, J., Nilsson, M., Mendel-Hartvig, M. & Landegren, U. *Signal amplification of padlock probes by rolling circle replication*. *Nucleic Acids Research* vol. 26 <https://academic.oup.com/nar/article-abstract/26/22/5073/1021212> (1998).
93. Deng, R., Zhang, K., Sun, Y., Ren, X. & Li, J. Highly specific imaging of mRNA in single cells by target RNA-initiated rolling circle amplification. *Chem. Sci.* **8**, 3668–3675 (2017).
94. Battich, N., Stoeger, T. & Pelkmans, L. Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat. Methods* **10**, 1127–1136 (2013).
95. Lécuyer, E. *et al.* Global Analysis of mRNA Localization Reveals a Prominent Role in Organizing Cellular Architecture and Function. *Cell* **131**, 174–187 (2007).

96. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in situ RNA profiling by sequential hybridization. *Nature Methods* vol. 11 360–361 (2014).
97. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science (80-. )*. **348**, (2015).
98. Xia, C., Babcock, H. P., Moffitt, J. R. & Zhuang, X. Multiplexed detection of RNA using MERFISH and branched DNA amplification. *Sci. Rep.* **9**, 1–13 (2019).
99. Eng, C. H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**, 235–239 (2019).
100. Xia, C., Fan, J., Emanuel, G., Hao, J. & Zhuang, X. Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 19490–19499 (2019).
101. Shah, S., Lubeck, E., Zhou, W. & Cai, L. seqFISH Accurately Detects Transcripts in Single Cells and Reveals Robust Spatial Organization in the Hippocampus. *Neuron* **94**, 752-758.e1 (2017).
102. Mayr, U., Serra, D. & Liberali, P. Exploring single cells in space and time during tissue development, homeostasis and regeneration. *Development (Cambridge)* vol. 146 (2019).
103. Shah, S. *et al.* Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. *Cell* **174**, 363-376.e16 (2018).
104. Goh, J. J. L. *et al.* Highly specific multiplexed RNA imaging in tissues with split-FISH. *Nat. Methods* 1–5 (2020) doi:10.1038/s41592-020-0858-0.
105. Lee, J. H. *et al.* Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015).
106. Fürth, D., Hatini, V. & Lee, J. H. In Situ Transcriptome Accessibility Sequencing (INSTA-seq). *bioRxiv* 722819 (2019) doi:10.1101/722819.
107. Liao, J., Lu, X., Shao, X., Zhu, L. & Fan, X. Uncovering an Organ’s Molecular Architecture at Single-Cell Resolution by Spatially Resolved Transcriptomics. *Trends in Biotechnology* (2020) doi:10.1016/j.tibtech.2020.05.006.
108. Shimomura, O. Structure of the chromophore of Aequorea green fluorescent protein. *FEBS Lett.* **104**, 220–222 (1979).
109. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W. & Prasher, D. C. Green fluorescent protein as a marker for gene expression. *Science (80-. )*. **263**, 802–805 (1994).
110. Heim, R. & Tsien, R. Y. Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer. *Curr. Biol.* **6**, 178–182 (1996).
111. Sakaue-Sawano, A. *et al.* Visualizing Spatiotemporal Dynamics of Multicellular Cell-Cycle Progression. *Cell* **132**, 487–498 (2008).
112. Otsuka, K. & Tomita, M. Concurrent live imaging of DNA double-strand break repair and cell-cycle progression by CRISPR/Cas9-mediated knock-in of a tricistronic vector. *Sci. Rep.* **8**, 1–9 (2018).
113. Kamiyama, D. *et al.* Versatile protein tagging in cells with split fluorescent protein. *Nat. Commun.* **7**, 1–9 (2016).
114. Stadler, C. *et al.* Immunofluorescence and fluorescent-protein tagging show high correlation for protein localization in mammalian cells. *Nat. Methods* **10**, 315–323 (2013).

115. Simpson, J. C., Wellenreuther, R., Poustka, A., Pepperkok, R. & Wiemann, S. Systematic subcellular localization of novel proteins identified by large-scale cDNA sequencing. *EMBO Rep.* **1**, 287–292 (2000).
116. Weill, U. *et al.* Assessment of GFP Tag Position on Protein Localization and Growth Fitness in Yeast. *J. Mol. Biol.* **431**, 636–641 (2019).
117. Huh, W.-K. K. *et al.* Global analysis of protein localization in budding yeast. *Nature* **425**, 686–691 (2003).
118. Chong, Y. T. *et al.* Yeast proteome dynamics from single cell imaging and automated analysis. *Cell* **161**, 1413–1424 (2015).
119. Breker, M., Gymrek, M. & Schuldiner, M. A novel single-cell screening platform reveals proteome plasticity during yeast stress responses. *J. Cell Biol.* **200**, 839–850 (2013).
120. Déneraud, N. *et al.* A chemostat array enables the spatio-temporal analysis of the yeast proteome. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 15842–15847 (2013).
121. Tkach, J. M. *et al.* Dissecting DNA damage response pathways by analysing protein localization and abundance changes during DNA replication stress. *Nat. Cell Biol.* **14**, 966–976 (2012).
122. Newman, J. R. S. *et al.* Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. *Nature* **441**, 840–846 (2006).
123. Torres, N. P., Ho, B. & Brown, G. W. High-throughput fluorescence microscopic analysis of protein abundance and localization in budding yeast. *Critical Reviews in Biochemistry and Molecular Biology* vol. 51 110–119 (2016).
124. Yofe, I. *et al.* One library to make them all: Streamlining the creation of yeast libraries via a SWAp-Tag strategy. *Nat. Methods* **13**, 371–378 (2016).
125. Meurer, M. *et al.* Genome-wide C-SWAT library for high-throughput yeast genome tagging. *Nat. Methods* **15**, 598–600 (2018).
126. Weill, U. *et al.* Genome-wide SWAp-Tag yeast libraries for proteome exploration. *Nat. Methods* **15**, 617–622 (2018).
127. Lu, A. X. & Moses, A. M. An unsupervised knn method to systematically detect changes in protein localization in high-throughput microscopy images. *PLoS One* **11**, (2016).
128. Breker, M., Gymrek, M., Moldavski, O. & Schuldiner, M. LoQAtE-Localization and Quantitation ATlas of the yeast proteome. A new tool for multiparametric dissection of single-protein behavior in response to biological perturbations in yeast. doi:10.1093/nar/gkt933.
129. Riffle, M. & Davis, T. N. The yeast resource center public image repository: A large database of fluorescence microscopy images. *BMC Bioinformatics* **11**, 263 (2010).
130. Chuartzman, S. G. & Schuldiner, M. Database for High Throughput Screening Hits (dHITS): a simple tool to retrieve gene specific phenotypes from systematic screens done in yeast. *Yeast* **35**, 477–483 (2018).
131. Cherry, J. M. *et al.* Saccharomyces Genome Database: the genomics resource of budding yeast. doi:10.1093/nar/gkr1029.
132. Koh, J. L. Y. *et al.* CYCLOPs: A comprehensive database constructed from automated analysis of protein abundance and subcellular localization patterns in *Saccharomyces cerevisiae*. *G3*

- Genes, Genomes, Genet.* **5**, 1223–1232 (2015).
133. Dubreuil, B. *et al.* YeastRGB: comparing the abundance and localization of yeast proteins across cells and libraries. *Nucleic Acids Res.* **47**, 1245–1249 (2018).
  134. Sigal, A. *et al.* Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nat. Methods* **3**, 525–531 (2006).
  135. Sigal, A. *et al.* Generation of a fluorescently labeled endogenous protein library in living human cells. *Nat. Protoc.* **2**, 1515–1527 (2007).
  136. Cohen, A. A. *et al.* Dynamic proteomics of individual cancer cells in response to a drug. *Science (80-. )*. **322**, 1511–1516 (2008).
  137. Frenkel-Morgenstern, M. *et al.* Dynamic Proteomics: a database for dynamics and localizations of endogenous fluorescently-tagged proteins in living human cells. doi:10.1093/nar/gkp808.
  138. Cho, N. H. *et al.* OpenCell: proteome-scale endogenous tagging enables the cartography of human cellular organization. *bioRxiv* 2021.03.29.437450 (2021) doi:10.1101/2021.03.29.437450.
  139. Lampasona, A. A. & Czaplinski, K. RNA voyeurism: A coming of age story. *Methods* vol. 98 10–17 (2016).
  140. Bertrand, E. *et al.* Localization of ASH1 mRNA particles in living yeast. *Mol. Cell* **2**, 437–445 (1998).
  141. Daigle, N. & Ellenberg, J.  $\lambda$ N-GFP: An RNA reporter system for live-cell imaging. *Nat. Methods* **4**, 633–636 (2007).
  142. Chen, J. *et al.* High efficiency of HIV-1 genomic RNA packaging and heterozygote formation revealed by single virion analysis. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13535–13540 (2009).
  143. Yiu, H.-W., Demidov, V. V., Toran, P., Cantor, C. R. & Broude, N. E. RNA Detection in Live Bacterial Cells Using Fluorescent Protein Complementation Triggered by Interaction of Two RNA Aptamers with Two RNA-Binding Peptides. *Pharmaceuticals* **4**, 494–508 (2011).
  144. Yin, J. *et al.* Imaging of mRNA-protein interactions in live cells using novel mCherry trimolecular fluorescence complementation systems. *PLoS One* **8**, (2013).
  145. Valencia-Burton, M., McCullough, R. M., Cantor, C. R. & Broude, N. E. RNA visualization in live bacterial cells using fluorescent protein complementation. *Nat. Methods* **4**, 421–427 (2007).
  146. Wu, B., Chen, J. & Singer, R. H. Background free imaging of single mRNAs in live cells using split fluorescent proteins. *Sci. Rep.* **4**, (2014).
  147. Wang, C., Han, B., Zhou, R. & Zhuang, X. Real-Time Imaging of Translation on Single mRNA Transcripts in Live Cells. *Cell* **165**, 990–1001 (2016).
  148. Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S. & Vale, R. D. A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* **159**, 635–646 (2014).
  149. Biswas, J., Liu, Y., Singer, R. H. & Wu, B. Fluorescence imaging methods to investigate translation in single cells. *Cold Spring Harb. Perspect. Biol.* **11**, (2019).
  150. Russo, J. & Wilusz, J. Trick or TREAT: A Scary-Good New Approach for Single-Molecule mRNA Decay Analysis. *Molecular Cell* vol. 68 476–477 (2017).

151. Halstead, J. M. *et al.* TRICK: A Single-Molecule Method for Imaging the First Round of Translation in Living Cells and Animals. in *Methods in Enzymology* vol. 572 123–157 (Academic Press Inc., 2016).
152. Horvathova, I. *et al.* The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells Molecular Cell The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. *Mol. Cell* **68**, 615–625 (2017).
153. Halstead, J. M. *et al.* An RNA biosensor for imaging the first round of translation from single cells to living animals. *Science (80-. )*. **347**, 1367–1370 (2015).
154. Wilbertz, J. H. *et al.* Single-Molecule Imaging of mRNA Localization and Regulation during the Integrated Stress Response. *Mol. Cell* **73**, 946-958.e7 (2019).
155. Wu, B., Chao, J. A. & Singer, R. H. Fluorescence fluctuation spectroscopy enables quantitative imaging of single mRNAs in living cells. *Biophys. J.* **102**, 2936–2944 (2012).
156. Weil, T. T., Parton, R. M. & Davis, I. Making the message clear: Visualizing mRNA localization. *Trends in Cell Biology* vol. 20 380–390 (2010).
157. Garcia, J. F. & Parker, R. MS2 coat proteins bound to yeast mRNAs block 5' to 3' degradation and trap mRNA decay products: Implications for the localization of mRNAs by MS2-MCP system. *Rna* **21**, 1393–1395 (2015).
158. Heinrich, S., Sidler, C. L., Azzalin, C. M. & Weis, K. Stem-loop RNA labeling can affect nuclear and cytoplasmic mRNA processing. *Rna* **23**, 134–141 (2017).
159. Haimovich, G. *et al.* Use of the MS2 aptamer and coat protein for RNA localization in yeast: A response to 'MS2 coat proteins bound to yeast mRNAs block 5' to 3' degradation and trap mRNA decay products: Implications for the localization of mRNAs by MS2-MCP system'. *Rna* **22**, 660–666 (2016).
160. Tutucci, E. *et al.* An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat. Methods* **15**, 81–89 (2018).
161. Bai, J. *et al.* A protein-independent fluorescent RNA aptamer reporter system for plant genetic engineering. *Nat. Commun.* **11**, 1–14 (2020).
162. Yan, Q. *et al.* Using an RNA aptamer probe for super-resolution imaging of native EGFR. *Nanoscale Adv.* **1**, 291–298 (2019).
163. Paige, J. S., Wu, K. Y. & Jaffrey, S. R. RNA mimics of green fluorescent protein. *Science (80-. )*. **333**, 642–646 (2011).
164. Guet, D. *et al.* Combining Spinach-tagged RNA and gene localization to image gene expression in live yeast. *Nat. Commun.* **6**, 1–10 (2015).
165. Tan, X. *et al.* Fluoromodules Consisting of a Promiscuous RNA Aptamer and Red or Blue Fluorogenic Cyanine Dyes: Selection, Characterization, and Bioimaging. *J. Am. Chem. Soc.* **139**, 9001–9009 (2017).
166. Le, T. T. *et al.* A highly stable RNA aptamer probe for the retinoblastoma protein in live cells. *Chem. Sci.* **11**, 4467–4474 (2020).
167. Yoon, S. *et al.* Targeted Delivery of C/EBP $\alpha$ -saRNA by RNA Aptamers Shows Anti-tumor Effects in a Mouse Model of Advanced PDAC. *Mol. Ther. - Nucleic Acids* **18**, 142–154 (2019).
168. Strack, R. L., Disney, M. D. & Jaffrey, S. R. A superfolding Spinach2 reveals the dynamic nature

- of trinucleotide repeat-containing RNA. *Nat. Methods* **10**, 1219–1224 (2013).
169. Filonov, G. S., Moon, J. D., Svensen, N. & Jaffrey, S. R. Broccoli: Rapid selection of an RNA mimic of green fluorescent protein by fluorescence-based selection and directed evolution. *J. Am. Chem. Soc.* **136**, 16299–16308 (2014).
  170. Song, W. *et al.* Imaging RNA polymerase III transcription using a photostable RNA-fluorophore complex. *Nat. Chem. Biol.* **13**, 1187–1194 (2017).
  171. Filonov, G. S., Song, W. & Jaffrey, S. R. Spectral Tuning by a Single Nucleotide Controls the Fluorescence Properties of a Fluorogenic Aptamer. *Biochemistry* **58**, 1560–1564 (2019).
  172. Wirth, R., Gao, P., Nienhaus, G. U., Sunbul, M. & Jäschke, A. SiRA: A Silicon Rhodamine-Binding Aptamer for Live-Cell Super-Resolution RNA Imaging. *J. Am. Chem. Soc.* **141**, 7562–7571 (2019).
  173. Sunbul, M. *et al.* Super-resolution RNA imaging using a rhodamine-binding aptamer with fast exchange kinetics. *Nat. Biotechnol.* **39**, 686–690 (2021).
  174. Trachman, R. J., Truong, L. & Ferré-D'Amaré, A. R. Structural Principles of Fluorescent RNA Aptamers. *Trends in Pharmacological Sciences* vol. 38 928–939 (2017).
  175. Swetha, P., Fan, Z., Wang, F. & Jiang, J. H. Genetically encoded light-up RNA aptamers and their applications for imaging and biosensing. *Journal of Materials Chemistry B* vol. 8 3382–3392 (2020).
  176. Gao, T., Luo, Y., Li, W., Cao, Y. & Pei, R. Progress in the isolation of aptamers to light-up the dyes and the applications. *Analyst* vol. 145 701–718 (2020).
  177. Headland, S. E., Jones, H. R., D'Sa, A. S. V., Perretti, M. & Norling, L. V. Cutting-edge analysis of extracellular microparticles using imagestreamx imaging flow cytometry. *Sci. Rep.* **4**, 1–10 (2014).
  178. Doan, M. *et al.* Diagnostic Potential of Imaging Flow Cytometry. *Trends in Biotechnology* vol. 36 649–652 (2018).
  179. Pekle, E. *et al.* Application of Imaging Flow Cytometry for the Characterization of Intracellular Attributes in Chinese Hamster Ovary Cell Lines at the Single-Cell Level. *Biotechnol. J.* **14**, 1800675 (2019).
  180. Lalmansingh, A. S., Arora, K., DeMarco, R. A., Hager, G. L. & Nagaich, A. K. High-Throughput RNA FISH Analysis by Imaging Flow Cytometry Reveals That Pioneer Factor Foxa1 Reduces Transcriptional Stochasticity. *PLoS One* **8**, 76043 (2013).
  181. Spitzer, M. H. & Nolan, G. P. Mass Cytometry: Single Cells, Many Features. *Cell* vol. 165 780–791 (2016).
  182. Baharlou, H., Canete, N. P., Cunningham, A. L., Harman, A. N. & Patrick, E. Mass Cytometry Imaging for the Study of Human Diseases—Applications and Data Analysis Strategies. *Frontiers in Immunology* vol. 10 2657 (2019).
  183. Ali, H. R. *et al.* Imaging mass cytometry and multiplatform genomics define the phenogenomic landscape of breast cancer. *Nat. Cancer* **1**, 163–175 (2020).
  184. Schulz, D. *et al.* Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Syst.* **6**, 25-36.e5 (2018).
  185. Angelo, M. *et al.* Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–

- 442 (2014).
186. Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373-1387.e19 (2018).
  187. Buchberger, A. R., DeLaney, K., Johnson, J. & Li, L. Mass Spectrometry Imaging: A Review of Emerging Advancements and Future Insights. *Analytical Chemistry* vol. 90 240–265 (2018).
  188. Niehaus, M., Soltwisch, J., Belov, M. E. & Dreisewerd, K. Transmission-mode MALDI-2 mass spectrometry imaging of cells and tissues at subcellular resolution. *Nat. Methods* **16**, 925–931 (2019).
  189. Swales, J. G., Hamm, G., Clench, M. R. & Goodwin, R. J. A. Mass spectrometry imaging and its application in pharmaceutical research and development: A concise review. *Int. J. Mass Spectrom.* **437**, 99–112 (2019).
  190. Kompauer, M., Heiles, S. & Spengler, B. Atmospheric pressure MALDI mass spectrometry imaging of tissues and cells at 1.4- $\mu\text{m}$  lateral resolution. *Nat. Methods* **14**, 90–96 (2016).
  191. St John, E. R., Rossi, M., Pruski, P., Darzi, A. & Takats, Z. Intraoperative tissue identification by mass spectrometric technologies. *TrAC - Trends Anal. Chem.* **85**, 2–9 (2016).
  192. Macosko, E. Z. *et al.* Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
  193. Rich-Griffin, C. *et al.* Single-Cell Transcriptomics: A High-Resolution Avenue for Plant Functional Genomics. *Trends in Plant Science* vol. 25 186–197 (2020).
  194. Hwang, B., Lee, J. H. & Bang, D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Experimental and Molecular Medicine* vol. 50 (2018).
  195. Birnbaum, K. D. Power in Numbers: Single-Cell RNA-Seq Strategies to Dissect Complex Tissues. *Annu. Rev. Genet.* **52**, 203–221 (2018).
  196. Budnik, B., Levy, E., Harmange, G. & Slavov, N. SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* **19**, 161 (2018).
  197. Minakshi, P. *et al.* Single-cell proteomics: Technology and applications. in *Single-Cell Omics: Volume 1: Technological Advances and Applications* 283–318 (Elsevier, 2019). doi:10.1016/B978-0-12-814919-5.00014-2.
  198. Zhu, Y. *et al.* Single-cell proteomics reveals changes in expression during hair-cell development. *Elife* **8**, (2019).
  199. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M. C. & Yates, J. R. Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews* vol. 113 2343–2394 (2013).
  200. Stark, R., Grzelak, M. & Hadfield, J. RNA sequencing: the teenage years. *Nature Reviews Genetics* vol. 20 631–656 (2019).
  201. Zhang, F., Ge, W., Ruan, G., Cai, X. & Guo, T. Data-Independent Acquisition Mass Spectrometry-Based Proteomics and Software Tools: A Glimpse in 2020. *Proteomics* vol. 20 1900276 (2020).
  202. de Duve, C. Tissue fraction-past and present. *J. Cell Biol.* **50**, 20 (1971).
  203. Miescher, F. In Hoppe-Seyler's Medizinisch-chemische Untersuchungen. A. Hirschwald, Berlin. **4**, (1871).

204. BEHRENS, M. Untersuchungen an isolierten Zell- und Gewebestandteilen. I. Mitteilung: Isolierung von Zellkernen des Kalbsherzmuskels. *Z Physiol Chem* **209**, 59–7 (1932).
205. Bensley, RR; Hoerr, N. Studies on cell structure by the freezing-drying method VI. The preparation and properties of mitochondria. *Anat. Rec.* **60**, 449–55 (1934).
206. Andersen, J. S. *et al.* Directed proteomic analysis of the human nucleolus. *Curr. Biol.* **12**, 1–11 (2002).
207. De Castro Moreira Dos Santos, A. *et al.* Unveiling the trypanosoma Cruzi nuclear proteome. *PLoS One* **10**, (2015).
208. Cronshaw, J. M., Krutchinsky, A. N., Zhang, W., Chait, B. T. & Matunis, M. L. J. Proteomic analysis of the mammalian nuclear pore complex. *J. Cell Biol.* **158**, 915–927 (2002).
209. Taylor, S. W. *et al.* Characterization of the human heart mitochondrial proteome. *Nat. Biotechnol.* **21**, 281–286 (2003).
210. Zhou, Z., Licklider, L. J., Gygi, S. P. & Reed, R. Comprehensive proteomic analysis of the human spliceosome. *Nature* **419**, 182–185 (2002).
211. Neubauer, G. *et al.* Mass spectrometry and EST-database searching allows characterization of the multi-protein spliceosome complex. *Nat. Genet.* **20**, 46–50 (1998).
212. Schirmer, E. C., Florens, L., Guan, T., Yates, J. R. & Gerace, L. Nuclear membrane proteins with potential disease links found by subtractive proteomics. *Science (80-. )*. **301**, 1380–1382 (2003).
213. Oh, P. *et al.* Subtractive proteomic mapping of the endothelial surface in lung and solid tumours for tissue-specific therapy. *Nature* **429**, 629–635 (2004).
214. Yates, J. R., Gilchrist, A., Howell, K. E. & Bergeron, J. J. M. Proteomics of organelles and large cellular structures. *Nature Reviews Molecular Cell Biology* vol. 6 702–714 (2005).
215. Wiederhold, E. *et al.* The yeast vacuolar membrane proteome. *Mol. Cell. Proteomics* **8**, 380–392 (2009).
216. Valli, M. *et al.* A subcellular proteome atlas of the yeast *Komagataella phaffii*. *FEMS Yeast Res.* **20**, (2020).
217. Delom, F. *et al.* The plasma membrane proteome of *Saccharomyces cerevisiae* and its response to the antifungal calcofluor. *Proteomics* **6**, 3029–3039 (2006).
218. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* **4**, 484–494 (2009).
219. Wang, X., Li, S., Wang, H., Shui, W. & Hu, J. Quantitative proteomics reveal proteins enriched in tubular endoplasmic reticulum of *Saccharomyces cerevisiae*. *Elife* **6**, (2017).
220. Sickmann, A. *et al.* The proteome of *Saccharomyces cerevisiae* mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 13207–13212 (2003).
221. Vögtle, F. N. *et al.* Landscape of submitochondrial protein distribution. *Nat. Commun.* **8**, (2017).
222. Morgenstern, M. *et al.* Definition of a High-Confidence Mitochondrial Proteome at Quantitative Scale. *Cell Rep.* **19**, 2836–2852 (2017).
223. Gatto, L., Vizcaíno, J. A., Hermjakob, H., Huber, W. & Lilley, K. S. Organelle proteomics

- experimental designs and analysis. *Proteomics* **10**, 3957–3969 (2010).
224. Ohta, S. *et al.* The Protein Composition of Mitotic Chromosomes Determined Using Multiclassifier Combinatorial Proteomics. *Cell* **142**, 810–821 (2010).
  225. Kustatscher, G. *et al.* Proteomics of a fuzzy organelle: Interphase chromatin. *EMBO J.* **33**, 648–664 (2014).
  226. Calvo, S. E., Clauser, K. R. & Mootha, V. K. MitoCarta2.0: An updated inventory of mammalian mitochondrial proteins. *Nucleic Acids Res.* **44**, D1251–D1257 (2016).
  227. Güther, M. L. S., Urbaniak, M. D., Tavendale, A., Prescott, A. & Ferguson, M. A. J. High-confidence glycosome proteome for procyclic form trypanosoma brucei by epitope-tag organelle enrichment and SILAC proteomics. *J. Proteome Res.* **13**, 2796–2806 (2014).
  228. Islinger, M., Lüers, G. H., Ka, W. L., Loos, M. & Völkl, A. Rat liver peroxisomes after fibrate treatment: A survey using quantitative mass spectrometry. *J. Biol. Chem.* **282**, 23055–23069 (2007).
  229. Marelli, M. *et al.* Quantitative mass spectrometry reveals a role for the GTPase Rho 1 p in actin organization on the peroxisome membrane. *J. Cell Biol.* **167**, 1099–1112 (2004).
  230. Ray, G. J. *et al.* A PEROXO-Tag Enables Rapid Isolation of Peroxisomes from Human Cells. *iScience* **23**, (2020).
  231. Schmidtke, C. *et al.* Lysosomal proteome analysis reveals that CLN3-defective cells have multiple enzyme deficiencies associated with changes in intracellular trafficking. *J. Biol. Chem.* **294**, 9592–9604 (2019).
  232. McCarthy, F. M., Cooksey, A. M. & Burgess, S. C. Sequential detergent extraction prior to mass spectrometry analysis. *Methods Mol. Biol.* **528**, 110–118 (2009).
  233. Lee, Y. H., Tan, H. T. & Chung, M. C. M. Subcellular fractionation methods and strategies for proteomics. *Proteomics* vol. 10 3935–3956 (2010).
  234. Stasyk, T. & Huber, L. A. Zooming in: Fractionation strategies in proteomics. *Proteomics* vol. 4 3704–3716 (2004).
  235. Masuda, T., Sugiyama, N., Tomita, M., Ohtsuki, S. & Ishihama, Y. Mass Spectrometry-Compatible Subcellular Fractionation for Proteomics. *J. Proteome Res.* **19**, 75–84 (2020).
  236. Jagannathan, S., Nwosu, C. & Nicchitta, C. V. Analyzing mRNA localization to the endoplasmic reticulum via cell fractionation. *Methods Mol. Biol.* **714**, 301–321 (2011).
  237. Jagannathan, S., Reid, D. W., Cox, A. H. & Nicchitta, C. V. De novo translation initiation on membrane-bound ribosomes as a mechanism for localization of cytosolic protein mRNAs to the endoplasmic reticulum. *RNA* **20**, 1489–1498 (2014).
  238. Michael Werner, A. S., Ruthenburg Correspondence, A. J., Werner, M. S. & Ruthenburg, A. J. Nuclear Fractionation Reveals Thousands of Chromatin-Tethered Noncoding RNAs Adjacent to Active Genes. *CellReports* **12**, 1089–1098 (2015).
  239. Tilgner, H. *et al.* Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res.* **22**, 1616–1625 (2012).
  240. Mayer, A. & Churchman, L. S. A detailed protocol for subcellular RNA sequencing (subRNA-seq). *Curr. Protoc. Mol. Biol.* **2017**, 4.29.1-4.29.18 (2017).

241. Bouvrette, L. P. B. *et al.* CeFra-seq reveals broad asymmetric mRNA and noncoding RNA distribution profiles in *Drosophila* and human cells. *RNA* **24**, 98–113 (2018).
242. Adekunle, D. A. & Wang, E. T. Transcriptome-wide organization of subcellular microenvironments revealed by ATLAS-Seq. *Nucleic Acids Res.* **48**, 5859–5872 (2020).
243. Bramwell, M. E. & Harris, H. The origin of the polydispersity in sedimentation patterns of rapidly labelled nuclear ribonucleic acid. *Biochem. J.* **103**, 816–830 (1967).
244. Agrawal, H. O. & Bruening, G. Isolation of high-molecular-weight, P32-labeled influenza virus ribonucleic acid. *Proc. Natl. Acad. Sci. U. S. A.* **55**, 818–825 (1966).
245. Pons, M. W. Studies on influenza virus ribonucleic acid. *Virology* **31**, 523–531 (1967).
246. Milo, R., Jorgensen, P., Moran, U., Weber, G. & Springer, M. BioNumbers The database of key numbers in molecular and cell biology. *Nucleic Acids Res.* **38**, D750 (2009).
247. Lefebvre, F. A. *et al.* CeFra-seq: Systematic mapping of RNA subcellular distribution properties through cell fractionation coupled to deep-sequencing. *Methods* **126**, 138–148 (2017).
248. Martinez-Val, A. *et al.* Spatial-proteomics reveal in-vivo phospho-signaling dynamics at subcellular resolution. *bioRxiv* 2021.02.02.425898 (2021) doi:10.1101/2021.02.02.425898.
249. Gatto, L. *et al.* A foundation for reliable spatial proteomics data analysis. *Mol. Cell. Proteomics* **13**, 1937–1952 (2014).
250. Itzhak, D. N., Tyanova, S., Cox, J. & Borner, G. H. Global, quantitative and dynamic mapping of protein subcellular localization. *Elife* **5**, e16950 (2016).
251. Mulvey, C. M. *et al.* Using hyperLOPIT to perform high-resolution mapping of the spatial proteome. *Nat. Protoc.* **12**, 1110–1135 (2017).
252. Andersen, J. S. *et al.* Proteomic characterization of the human centrosome by protein correlation profiling. *Nature* **426**, 570–574 (2003).
253. Wiese, S. *et al.* Proteomics characterization of mouse kidney peroxisomes by tandem mass spectrometry and protein correlation profiling. *Mol. Cell. Proteomics* **6**, 2045–2057 (2007).
254. Krahmer, N. *et al.* Protein correlation profiles identify lipid droplet proteins with high confidence. *Mol. Cell. Proteomics* **12**, 1115–1126 (2013).
255. Sessler, N. *et al.* Analysis of the *Plasmodium falciparum* proteasome using Blue Native PAGE and label-free quantitative mass spectrometry. *Amino Acids* vol. 43 1119–1129 (Springer, 2012).
256. Foster, L. J. *et al.* A Mammalian Organelle Map by Protein Correlation Profiling. *Cell* **125**, 187–199 (2006).
257. Kislinger, T. *et al.* Global Survey of Organ and Organelle Protein Expression in Mouse: Combined Proteomic and Transcriptomic Profiling. *Cell* **125**, 173–186 (2006).
258. Dunkley, T. P. J., Watson, R., Griffin, J. L., Dupree, P. & Lilley, K. S. Localization of Organelle Proteins by Isotope Tagging (LOPIT). *Mol. Cell. Proteomics* **3**, 1128–1134 (2004).
259. Nikolovski, N. *et al.* Putative glycosyltransferases and other plant Golgi apparatus proteins are revealed by LOPIT proteomics. *Plant Physiol.* **160**, 1037–1051 (2012).
260. Shin, J. *et al.* Determining the content of vesicles captured by golgin tethers using LOPIT-DC. *bioRxiv* 841965 (2019) doi:10.1101/841965.

261. Geladaki, A. *et al.* Combining LOPIT with differential ultracentrifugation for high-resolution spatial proteomics. *Nat. Commun.* **10**, 331 (2019).
262. Hall, S. L., Hester, S., Griffin, J. L., Lilley, K. S. & Jackson, A. P. The organelle proteome of the DT40 lymphocyte cell line. *Mol. Cell. Proteomics* **8**, 1295–1305 (2009).
263. Tan, D. J. L. *et al.* Mapping organelle proteins and protein complexes in *Drosophila melanogaster*. *J. Proteome Res.* **8**, 2667–2678 (2009).
264. Christoforou, A. *et al.* A draft map of the mouse pluripotent stem cell spatial proteome. *Nat. Commun.* **7**, 9992 (2016).
265. Nightingale, D. J., Geladaki, A., Breckels, L. M., Oliver, S. G. & Lilley, K. S. The subcellular organisation of *Saccharomyces cerevisiae*. *Current Opinion in Chemical Biology* vol. 48 86–95 (2019).
266. Baers, L. L. *et al.* Proteome mapping of a cyanobacterium reveals distinct compartment organization and cell-dispersed metabolism. *Plant Physiol.* **181**, 1721–1738 (2019).
267. Barylyuk, K. *et al.* A subcellular atlas of *Toxoplasma* reveals the functional context of the proteome. *bioRxiv* 2020.04.23.057125 (2020) doi:10.1101/2020.04.23.057125.
268. Parsons, H. T. *et al.* Separating Golgi proteins from cis to trans reveals underlying properties of cisternal localization. *Plant Cell* **31**, 2010–2034 (2019).
269. Orre, L. M. *et al.* SubCellBarCode: Proteome-wide Mapping of Protein Localization and Relocalization. *Mol. Cell* **73**, 166-182.e7 (2019).
270. Kennedy, M. A., Hofstadter, W. A. & Cristea, I. M. TRANSPIRE: A Computational Pipeline to Elucidate Intracellular Protein Movements from Spatial Proteomics Data Sets. *J. Am. Soc. Mass Spectrom.* **31**, 1422–1439 (2020).
271. Crook, O. M., Davies, C. T. R., Gatto, L., Kirk, P. D. W. & Lilley, K. S. Inferring differential subcellular localisation in comparative spatial proteomics using BUNDLE. *bioRxiv* 2021.01.04.425239 (2021) doi:10.1101/2021.01.04.425239.
272. Crook, O. M., Mulvey, C. M., Kirk, P. D. W., Lilley, K. S. & Gatto, L. A Bayesian mixture modelling approach for spatial proteomics. *PLoS Comput. Biol.* **14**, e1006516 (2018).
273. Crook, O. M. *et al.* A semi-supervised Bayesian approach for simultaneous protein sub-cellular localisation assignment and novelty detection. *PLoS Comput. Biol.* **16**, e1008288 (2020).
274. Crook, O. M., Smith, T., Elzek, M. & Lilley, K. S. Moving Profiling Spatial Proteomics Beyond Discrete Classification. *Proteomics* 1900392 (2020) doi:10.1002/pmic.201900392.
275. Kleene, K. C., Bagarova, J., Hawthorne, S. K. & Catado, L. M. Quantitative analysis of mRNA translation in mammalian spermatogenic cells with sucrose and Nycodenz gradients. *Reprod. Biol. Endocrinol.* **8**, 155 (2010).
276. Aboulhoda, S., Di Santo, R., Therizols, G. & Weinberg, D. Accurate, Streamlined Analysis of mRNA Translation by Sucrose Gradient Fractionation. *BIO-PROTOCOL* **7**, (2017).
277. Jain, S. *et al.* ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **164**, 487–498 (2016).
278. Khong, A., Jain, S., Matheny, T., Wheeler, J. R. & Parker, R. Isolation of mammalian stress granule cores for RNA-Seq analysis. *Methods* **137**, 49–54 (2018).

279. Khong, A. *et al.* The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. *Mol. Cell* **68**, 808-820.e5 (2017).
280. Pasquali, C., Fialka, I. & Huber, L. A. Subcellular fractionation, electromigration analysis and mapping of organelles. *Journal of Chromatography B: Biomedical Sciences and Applications* vol. 722 89–102 (1999).
281. Satori, C. P., Kostal, V. & Arriaga, E. A. Review on recent advances in the analysis of isolated organelles. *Analytica Chimica Acta* vol. 753 8–18 (2012).
282. Drissi, R., Dubois, M. L. & Boisvert, F. M. Proteomics methods for subcellular proteome analysis. *FEBS J.* **280**, 5626–5634 (2013).
283. Gauthier, D. J. & Lazure, C. Complementary methods to assist subcellular fractionation in organellar proteomics. *Expert Review of Proteomics* vol. 5 603–617 (2008).
284. Parsons, H. T. & Lilley, K. S. Mass spectrometry approaches to study plant endomembrane trafficking. *Semin. Cell Dev. Biol.* **80**, 123–132 (2018).
285. Parsons, H. T. Preparation of highly enriched ER membranes using free-flow electrophoresis. in *Methods in Molecular Biology* vol. 1691 103–115 (Humana Press Inc., 2018).
286. Satori, C. P. *et al.* Bioanalysis of eukaryotic organelles. *Chemical Reviews* vol. 113 2733–2811 (2013).
287. Tharkeshwar, A. K., Gevaert, K. & Annaert, W. Organellar Omics—A Reviving Strategy to Untangle the Biomolecular Complexity of the Cell. *Proteomics* **18**, 1–12 (2018).
288. Moon, M. H. Flow field-flow fractionation: Recent applications for lipidomic and proteomic analysis. *TrAC - Trends in Analytical Chemistry* vol. 118 19–28 (2019).
289. Kang, D., Oh, S., Reschiglian, P. & Moon, M. H. Separation of mitochondria by flow field-flow fractionation for proteomic analysis. *Analyst* **133**, 505–515 (2008).
290. Oeyen, E. *et al.* Ultrafiltration and size exclusion chromatography combined with asymmetrical-flow field-flow fractionation for the isolation and characterisation of extracellular vesicles from urine. *J. Extracell. Vesicles* **7**, (2018).
291. Zhang, H. *et al.* Identification of distinct nanoparticles and subsets of extracellular vesicles by asymmetric flow field-flow fractionation. *Nat. Cell Biol.* **20**, 332–343 (2018).
292. Yang, J. S., Lee, J. Y. & Moon, M. H. High Speed Size Sorting of Subcellular Organelles by Flow Field-Flow Fractionation. *Anal. Chem.* **87**, 6342–6348 (2015).
293. George, L., Indig, F. E., Abdelmohsen, K. & Gorospe, M. Intracellular RNA-tracking methods. *Open Biology* vol. 8 (2018).
294. Abdelmoez, M. N. *et al.* SINC-seq: Correlation of transient gene expressions between nucleus and cytoplasm reflects single-cell physiology. *Genome Biol.* **19**, 66 (2018).
295. Oguchi, Y., Ozaki, Y., Abdelmoez, M. N. & Shintaku, H. NanoSINC-seq dissects the isoform diversity in subcellular compartments of single cells. *Sci. Adv.* **7**, (2021).
296. Samavarchi-Tehrani, P., Samson, R. & Gingras, A.-C. Proximity dependent biotinylation: key enzymes and adaptation to proteomics approaches. *Mol. Cell. Proteomics* mcp.R120.001941 (2020) doi:10.1074/mcp.r120.001941.
297. Go, C. D. *et al.* A proximity-dependent biotinylation map of a human cell. *Nature* **595**, 120–124 (2021).

298. Liu, X. *et al.* An AP-MS- and BioID-compatible MAC-tag enables comprehensive mapping of protein interactions and subcellular localizations. *Nat. Commun.* **9**, 1–16 (2018).
299. Lambert, J. P., Tucholska, M., Go, C., Knight, J. D. R. & Gingras, A. C. Proximity biotinylation and affinity purification are complementary approaches for the interactome mapping of chromatin-associated protein complexes. *J. Proteomics* **118**, 81–94 (2015).
300. Mellacheruvu, D. *et al.* The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10**, 730–736 (2013).
301. Minde, D. P., Ramakrishna, M. & Lilley, K. S. Biotin proximity tagging favours unfolded proteins and enables the study of intrinsically disordered regions. *Commun. Biol.* **3**, 1–13 (2020).
302. Rees, J. S., Li, X.-W., Perrett, S., Lilley, K. S. & Jackson, A. P. Protein Neighbors and Proximity Proteomics. *Mol. Cell. Proteomics* **14**, 2848–56 (2015).
303. Go, C. *et al.* A proximity biotinylation map of a human cell. *bioRxiv* 796391 (2019) doi:10.1101/796391.
304. Schatz, P. J. Use of peptide libraries to map the substrate specificity of a peptide-modifying enzyme: A 13 residue consensus peptide specifies biotinylation in *Escherichia coli*. *Bio/Technology* **11**, 1138–1143 (1993).
305. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).
306. May, D. G., Scott, K. L., Campos, A. R. & Roux, K. J. Comparative Application of BioID and TurboID for Protein-Proximity Biotinylation. *Cells* **9**, (2020).
307. Xie, W. *et al.* A-type Lamins Form Distinct Filamentous Networks with Differential Nuclear Pore Complex Associations. *Curr. Biol.* **26**, 2651–2658 (2016).
308. Kim, D. I. *et al.* Probing nuclear pore complex architecture with proximity-dependent biotinylation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E2453–E2461 (2014).
309. Remnant, L. *et al.* In vitro BioID: Mapping the CENP-A microenvironment with high temporal and spatial resolution. *Mol. Biol. Cell* **30**, 1314–1325 (2019).
310. Hua, R. *et al.* VAPs and ACBD5 tether peroxisomes to the ER for peroxisome maintenance and lipid homeostasis. *J. Cell Biol.* **216**, 367–377 (2017).
311. Dong, J. M. *et al.* Proximity biotinylation provides insight into the molecular composition of focal adhesions at the nanometer scale. *Sci. Signal.* **9**, rs4–rs4 (2016).
312. Gupta, G. D. *et al.* A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. *Cell* **163**, 1484–1499 (2015).
313. Couzens, A. L. *et al.* Protein interaction network of the mammalian hippo pathway reveals mechanisms of kinase-phosphatase interactions. *Sci. Signal.* **6**, (2013).
314. Youn, J.-Y. *et al.* High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. *Mol. Cell* **69**, 517–532.e11 (2018).
315. Ramanathan, M. *et al.* RN A-protein interaction detection in living cells. *Nat. Methods* **15**, 207–212 (2018).
316. Mukherjee, J. *et al.*  $\beta$ -Actin mRNA interactome mapping by proximity biotinylation. *Proc. Natl.*

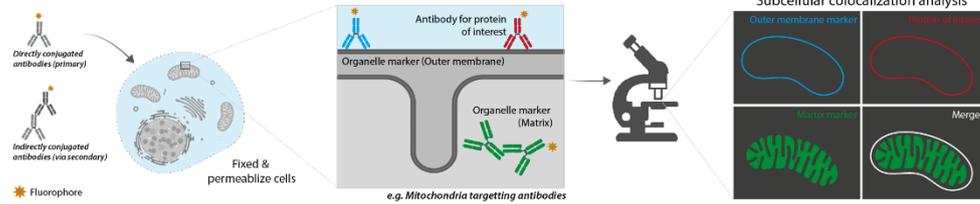
- Acad. Sci. U. S. A.* **116**, 12863–12872 (2019).
317. Lu, M. & Wei, W. Proximity labeling to detect RNA–protein interactions in live cells. *FEBS Open Bio* **9**, 1860–1868 (2019).
  318. Jan, C. H., Williams, C. C. & Weissman, J. S. Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science (80-. )*. **346**, (2014).
  319. Williams, C. C., Jan, C. H. & Weissman, J. S. Targeting and plasticity of mitochondrial proteins revealed by proximity-specific ribosome profiling. *Science (80-. )*. **346**, 748–751 (2014).
  320. Vardi-Oknin, D. & Arava, Y. Characterization of Factors Involved in Localized Translation Near Mitochondria by Ribosome-Proximity Labeling. *Front. Cell Dev. Biol.* **7**, 305 (2019).
  321. Trinkle-Mulcahy, L. Recent advances in proximity-based labeling methods for interactome mapping [version 1; referees: 2 approved]. *F1000Research* vol. 8 (2019).
  322. Rhee, H. W. *et al.* Proteomic mapping of mitochondria in living cells via spatially restricted enzymatic tagging. *Science (80-. )*. **339**, 1328–1331 (2013).
  323. Hung, V. *et al.* Proteomic Mapping of the Human Mitochondrial Intermembrane Space in Live Cells via Ratiometric APEX Tagging. *Mol. Cell* **55**, 332–341 (2014).
  324. Lam, S. S. *et al.* Directed evolution of APEX2 for electron microscopy and proximity labeling. *Nat. Methods* **12**, 51–54 (2014).
  325. Hung, V. *et al.* Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456–475 (2016).
  326. Hung, V. *et al.* Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *Elife* **6**, (2017).
  327. Han, S. *et al.* Proximity Biotinylation as a Method for Mapping Proteins Associated with mtDNA in Living Cells. *Cell Chem. Biol.* **24**, 404–414 (2017).
  328. Udeshi, N. D. *et al.* Antibodies to biotin enable large-scale detection of biotinylation sites on proteins. *Nat. Methods* **14**, 1167–1170 (2017).
  329. Del Olmo, T. *et al.* APEX2-mediated RAB proximity labeling identifies a role for RAB21 in clathrin-independent cargo sorting. *EMBO Rep.* **20**, (2019).
  330. Chu, Q. *et al.* Identification of Microprotein-Protein Interactions via APEX Tagging. *Biochemistry* **56**, 3299–3306 (2017).
  331. Paek, J. *et al.* Multidimensional Tracking of GPCR Signaling via Peroxidase-Catalyzed Proximity Labeling. *Cell* **169**, 338-349.e11 (2017).
  332. Lobingier, B. T. *et al.* An Approach to Spatiotemporally Resolve Protein Interaction Networks in Living Cells. *Cell* **169**, 350-360.e12 (2017).
  333. Markmiller, S. *et al.* Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. *Cell* **172**, 590-604.e13 (2018).
  334. Kaewsapsak, P., Shechner, D. M., Mallard, W., Rinn, J. L. & Ting, A. Y. Live-cell mapping of organelle-associated RNAs via proximity biotinylation combined with protein-RNA crosslinking. *Elife* **6**, (2017).
  335. Benhalevy, D., Anastasakis, D. G. & Hafner, M. Proximity-CLIP provides a snapshot of protein-occupied RNA elements in subcellular compartments. *Nat. Methods* **15**, 1074–1082 (2018).

336. Zhou, Y. *et al.* Expanding APEX2 Substrates for Proximity-Dependent Labeling of Nucleic Acids and Proteins in Living Cells. *Angew. Chemie Int. Ed.* **58**, 11763–11767 (2019).
337. Hwang, J. *et al.* A Golgi rhomboid protease Rbd2 recruits Cdc48 to cleave yeast SREBP. *EMBO J.* **35**, 2332–2349 (2016).
338. Hwang, J. & Espenshade, P. J. Proximity-dependent biotin labelling in yeast using the engineered ascorbate peroxidase APEX2. *Biochem. J.* **473**, 2463–2469 (2016).
339. Singer-Krüger, B. *et al.* APEX2-mediated proximity labeling resolves protein networks in *Saccharomyces cerevisiae* cells. *FEBS J.* **287**, 325–344 (2020).
340. Reinke, A. W., Mak, R., Troemel, E. R. & Ben, E. J. In vivo mapping of tissue- and subcellular-specific proteomes in *Caenorhabditis elegans*. *Sci. Adv.* **3**, e1602426 (2017).
341. Chen, C. L. *et al.* Proteomic mapping in live *Drosophila* tissues using an engineered ascorbate peroxidase. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 12093–12098 (2015).
342. Mannix, K. M., Starble, R. M., Kaufman, R. S. & Cooley, L. Proximity labeling reveals novel interactomes in live *Drosophila* tissue. *Development* **146**, (2019).
343. Sahl, S. J., Hell, S. W. & Jakobs, S. Fluorescence nanoscopy in cell biology. *Nature Reviews Molecular Cell Biology* vol. 18 685–701 (2017).
344. Mund, A. *et al.* AI-driven Deep Visual Proteomics defines cell identity and heterogeneity Proteomics Program, 2 Protein Signaling Program, and 3 Protein Imaging Platform. *bioRxiv* 2021.01.25.427969 (2021) doi:10.1101/2021.01.25.427969.
345. Haque, A., Engel, J., Teichmann, S. A. & Lönnerberg, T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine* vol. 9 1–12 (2017).
346. Lee, J. H. *et al.* Highly multiplexed subcellular RNA sequencing in situ. *Science (80-. )*. **343**, 1360–1363 (2014).
347. Labib, M. & Kelley, S. O. Single-cell analysis targeting the proteome. *Nature Reviews Chemistry* vol. 4 143–158 (2020).
348. Marx, V. Method of the Year: spatially resolved transcriptomics. *Nat. Methods* **18**, 9–14 (2021).
349. Leitner, A. Cross-linking and other structural proteomics techniques: How chemistry is enabling mass spectrometry applications in structural biology. *Chem. Sci.* **7**, 4792–4803 (2016).
350. Matzinger, M. & Mechtler, K. Cleavable Cross-Linkers and Mass Spectrometry for the Ultimate Task of Profiling Protein-Protein Interaction Networks in Vivo. *Journal of Proteome Research* vol. 20 78–93 (2021).
351. Hevler, J. F. *et al.* Selective cross-linking of coinciding protein assemblies by in-gel cross-linking mass spectrometry. *EMBO J.* **40**, e106174 (2021).
352. Steigenberger, B., Pieters, R. J., Heck, A. J. R. & Scheltema, R. A. PhoX: An IMAC-Enrichable Cross-Linking Reagent. *ACS Cent. Sci.* **5**, 1514–1522 (2019).
353. Fulcher, J. M. *et al.* Enhancing Top-Down Proteomics of Brain Tissue with FAIMS. *J. Proteome Res.* **20**, 2780–2795 (2021).
354. Zhong, Y., Hyung, S. J. & Ruotolo, B. T. Ion mobility-mass spectrometry for structural

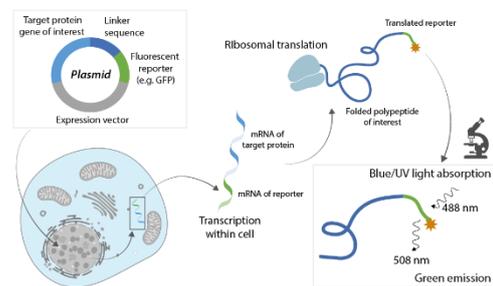
- proteomics. *Expert Review of Proteomics* vol. 9 47–58 (2012).
355. Zhao, H., Cunningham, D. L., Creese, A. J., Heath, J. K. & Cooper, H. J. FAIMS and Phosphoproteomics of Fibroblast Growth Factor Signaling: Enhanced Identification of Multiply Phosphorylated Peptides. *J. Proteome Res.* **14**, 5077–5087 (2015).
  356. Needham, E. J., Parker, B. L., Burykin, T., James, D. E. & Humphrey, S. J. Illuminating the dark phosphoproteome. *Science Signaling* vol. 12 8645 (2019).
  357. Humphrey, S. J., Karayel, O., James, D. E. & Mann, M. High-throughput and high-sensitivity phosphoproteomics with the EasyPhos platform. *Nat. Protoc.* **13**, 1897–1916 (2018).
  358. Chen, B., Yuan, B. F. & Feng, Y. Q. Analytical Methods for Deciphering RNA Modifications. *Anal. Chem.* **91**, 743–756 (2019).
  359. Motorin, Y. & Helm, M. Methods for RNA modification mapping using deep sequencing: Established and new emerging technologies. *Genes* vol. 10 35 (2019).
  360. Trendel, J. *et al.* The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell* **176**, 391-403.e19 (2019).
  361. Queiroz, R. M. L. *et al.* Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat. Biotechnol.* **37**, 169–178 (2019).

# Figures

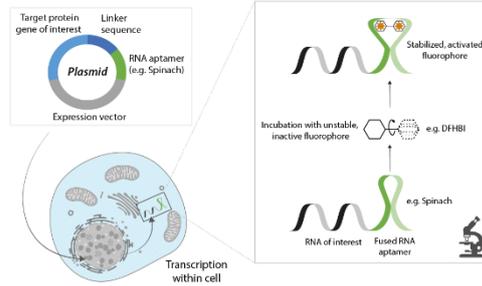
## a) Antibody staining (colocalization)



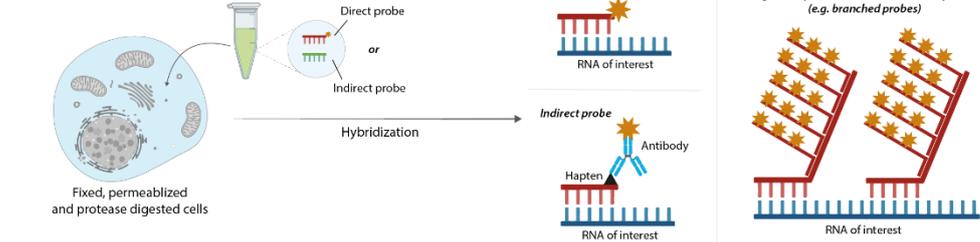
## b) Fluorescent protein reporters (e.g. GFP)



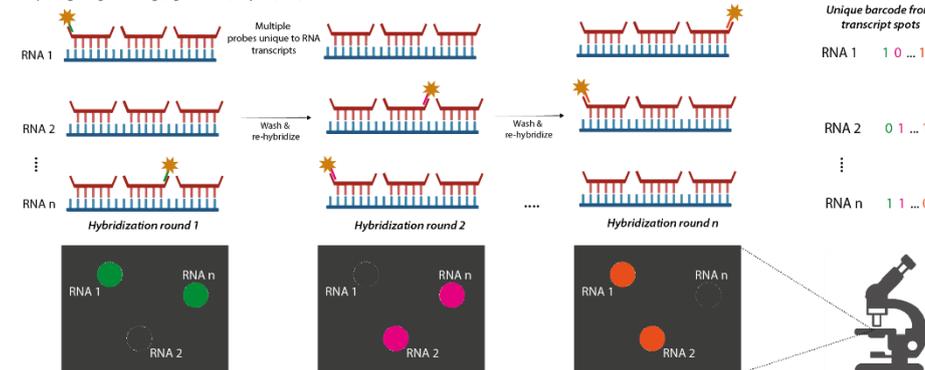
## c) RNA aptamers



## d) In-situ hybridization

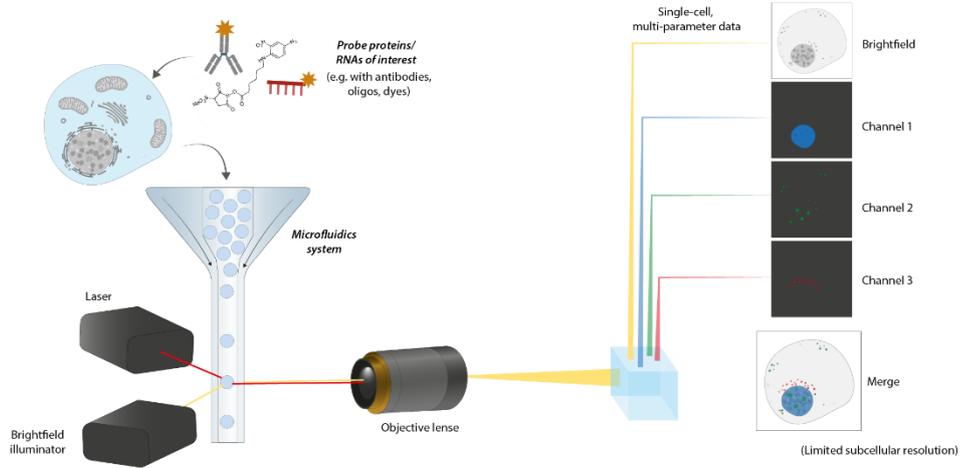


### Multiplexing using barcoding (e.g. MERFISH, seqFISH, etc.)

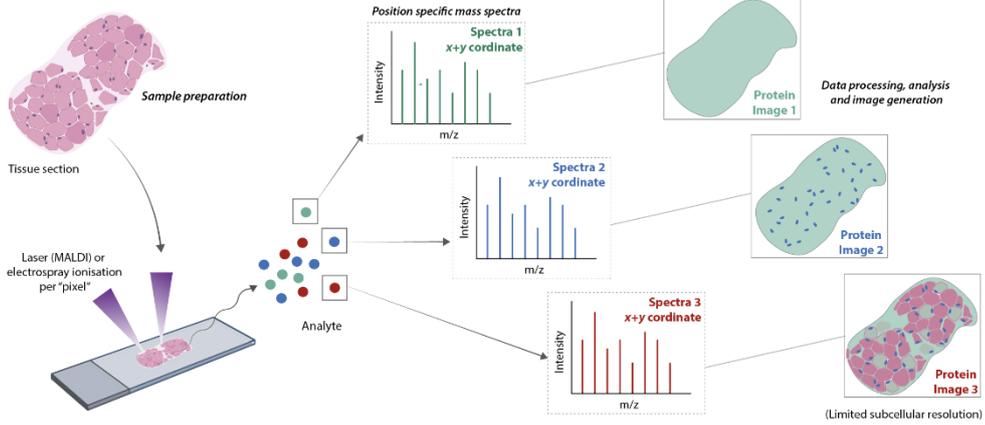


**Figure 1. Microscopy-based imaging approaches for subcellular proteomics or transcriptomics, focusing on the probing strategies.** A) Traditional antibody staining involves probing subcellular targets (such as the mitochondrial sub-structure) using monoclonal antibodies. These may be directly conjugated to a fluorescent label (direct immunofluorescence), or with a fluorescently labelled secondary antibody (indirect immunofluorescence). To determine subcellular location of proteins, an antibody against an organelle marker or a dye must be used alongside an antibody against the protein of interest. Then analysis can be performed to determine and quantify the co-localization of these antibodies/dyes. B) Fluorescent protein reporters, such as GFP, can be genetically engineered to be fused and expressed with a target gene/protein of interest. Therefore, allowing confocal imaging of molecules that have no antibody or require live-cell imaging. In MS2 labelling systems for RNA, fluorescent reporter proteins can be genetically fused to MCP. C) RNA aptamers are an alternative to MS2 systems for labelling RNA, which allow for fusion of an RNA structure that binds and stabilizes an exogenous fluorescent molecule (e.g. DFHBI). RNA aptamers can either be used as affinity reagents or as reporters. D) *In-situ* hybridization (ISH) employs a variety of antisense nucleic acid probes for the detection of RNA of interest in permeabilized and fixed cellular material. Recent ISH strategies have allowed for highly multiplexed experimental designs using molecular barcoding (e.g. seqFISH and MERFISH).

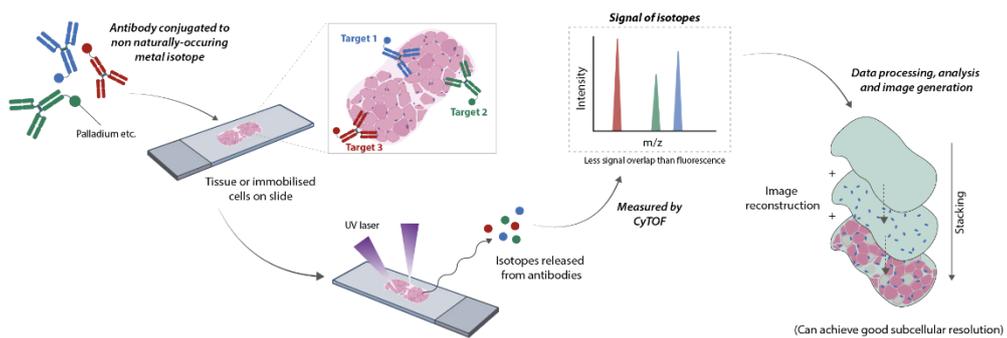
**a) Imaging flow cytometry (IFC)**



**b) Mass spectrometry imaging (MSI)**



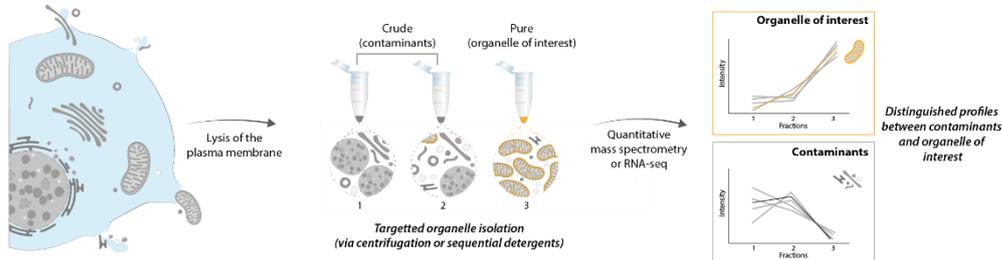
**c) Imaging mass cytometry (IMC)**



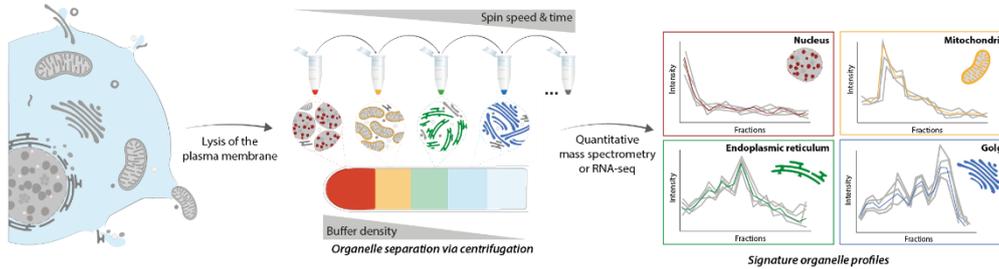
**Figure 2. Alternative imaging for subcellular proteomics and/or transcriptomics, which couple technologies in MS, microfluidics, and microdissection. A) Instrumentation coupling flow cytometry and microscopy allows for multiplexing of several protein/RNA targets using fluorescent labels,**

gaining both spatial and single-cell information. B) Micro-laser ablation and ionization of molecules, such as peptides, lipids, or metabolites, directly from tissue or cell culture sample enables label-free acquisition of mass spectra across each “pixel” of sample. Very rich datasets but still have poor resolution due to current technical limitations. C) Similar to MSI, micro-laser ablation allows for acquisition of spectra per “pixel” of a sample. Though, this method has improved subcellular resolution and uses labelling of antibodies conjugated to non-naturally occurring metal isotopes to quantify ~40 target proteins/RNAs of interest. The metal isotope signals have less signal overlap than fluorescent methods allowing improved multiplexing than traditional antibody probing.

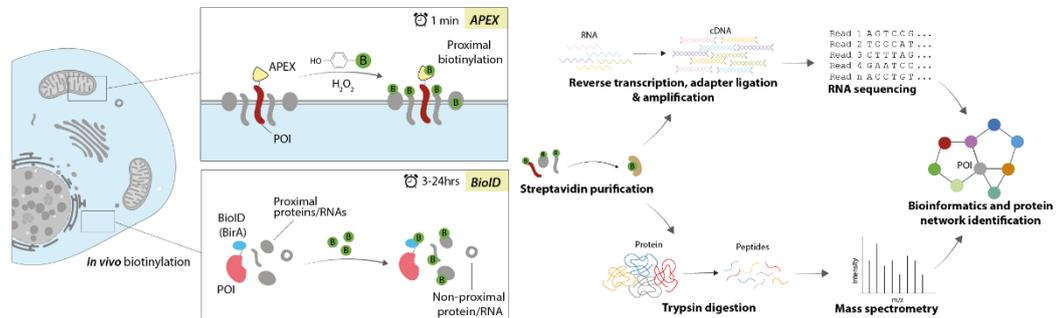
**a) Organelle-specific biochemical fractionation (e.g. subtractive)**



**b) Cell-wide biochemical fractionation (e.g. correlation profiling)**



**c) Proximity labelling**



**Figure 3. Sequencing-based approaches to subcellular proteomics and transcriptomics consist of biochemical organellar separation (A-B) or biotinylation of proximal molecules to a bait protein (C).** A) Quantifying proteins/RNAs in a targeted organelle-enrichment preparation (via centrifugation or detergents) against crude, contaminant samples can infer resident proteins/RNAs of the organelle of interest. Quantification of enriched samples can be performed using MS or RNA-seq. B) More extensive sequential centrifugation or detergent strategies can determine cell-wide residence of proteins/RNAs. The quantitative profiles of proteins/RNAs across the fractions aid identification of their localization by using organellar markers and machine learning techniques. C) A bait protein of interest (e.g. associated with a particular subcellular localization) is used to an enzyme which catalyzes the biotinylation of proximal proteins/RNAs in the cell once the substrate (e.g. biotin) is added to the cells *in vivo*. The biotinylated molecules can be purified and analyzed using either MS or RNA-seq.