

1 Visual proteomics using whole-lamella 2D template matching

2

3 This manuscript (permalink) was automatically generated from jojoelfe/deco\_lace\_template\_matching\_manuscript@0b9b63c  
4 on February 25, 2022.

5 **Authors**

- 6 • **Johannes Elferich**  0000-0002-9911-706X ·  jojoelfe RNA Therapeutic Institute, UMass Chan Medical  
7 School; HHMI
- 8 • **Nikolaus Grigorieff**  0000-0002-1506-909X ·  nikogrigorieff RNA Therapeutic Institute, UMass Chan  
9 Medical School; HHMI

10 **Abstract**

11 Localization of biomolecules inside a cell is an important goal of biological imaging. Fluorescence microscopy can  
12 localize biomolecules inside whole cells and tissues, but its ability to count biomolecules and accuracy of the spatial  
13 coordinates is limited by the wavelength of visible light. On the other hand, cryo-electron microscopy (cryo-EM)  
14 provides highly accurate position and orientation information of biomolecules but is often limited to small fields of  
15 view inside a cell, providing only limited biological context. In this study we use a new data-acquisition scheme  
16 called “Beam-Imageshift for Large Area Cryo-Electron microscopy” (BILACE) to collect high-resolution cryo-EM  
17 data over entire sections (100 – 200 nm thick lamellae) of neutrophil-like mouse cells, representing roughly 1% of the  
18 total cellular volume. We use 2D template matching to determine localization and orientation of the large ribosomal  
19 subunit in these sections. Furthermore, we use 2D template matching to test detected targets for the presence of  
20 the small ribosomal subunit and used the relative orientations of the ribosomes to assign them to polysomes. These  
21 results provide “maps” of translational activity across sections of mammalian cells. We envision that using this high-  
22 throughput cryo-EM data collection approach together with 2D template matching will advance visual proteomics  
23 to complement other single-cell “omics” techniques, such as flow-cytometry and single-cell sequencing.

<sup>24</sup> **Introduction**

<sup>25</sup> Understanding of cellular processes requires knowledge of the amounts, location, interactions, and conformations  
<sup>26</sup> of biomolecules inside the cell. Techniques that measure this can broadly be divided into label- and label-free  
<sup>27</sup> techniques. In label-techniques a probe is physically attached to a molecule of interest that is able to produce  
<sup>28</sup> a high signal-to-noise signal, such as a fluorescent molecule. In label-free techniques the physical properties of  
<sup>29</sup> molecules themselves are used for detection. An example for this is proteomics using mass-spectrometry [1]. The  
<sup>30</sup> advantage of label-free techniques is that they can provide information over thousands of molecules, while label-  
<sup>31</sup> techniques offer higher fidelity information for a few molecules. Especially spatial information can most of the time  
<sup>32</sup> only be achieved using label-techniques, such as fluorescence microscopy [2].

<sup>33</sup> Cryo-electron microscopy has the potential to directly measure the arrangement of atoms that compose biomole-  
<sup>34</sup> cules inside of cells, thereby allowing label-free detection with high spatial accuracy. This has been called “visual  
<sup>35</sup> proteomics” [3]. Since cryo-EM requires thin samples (<500nm), imaging of biomolecules inside cells is either re-  
<sup>36</sup> stricted to small organisms, thin regions of large cells, or requires thinning of the sample. This can be achieved  
<sup>37</sup> either by mechanical sectioning [4] or by milling using a focused ion beam (FIB) [5]. This complex workflow restricts  
<sup>38</sup> throughput of cryo-EM imaging of cells. This is exacerbated by the fact that at the required magnifications, typical  
<sup>39</sup> field of views (FOV) are very small compared to mammalian cells and the FOV achieved by label-techniques such  
<sup>40</sup> as fluorescence light microscopy. The predominant cryo-EM technique for detection of biomolecules according to  
<sup>41</sup> their shape in cells at the moment is cryo-electron tomography [6]. However, the requirement of physically tilt the  
<sup>42</sup> stage at every FOV, together with a more complex workflow that requires highly accurate alignment of various  
<sup>43</sup> projection, further restricts the throughput for molecular detection.

<sup>44</sup> An alternative approach is to identify molecules by their structural “fingerprint” in single projection using “2D  
<sup>45</sup> template-matching” [7,8,9]. In this method an experimentally obtained 3D model of a biomolecule is used to  
<sup>46</sup> calculate the expected electron density, which is called the template. The template is then projected on a fine  
<sup>47</sup> angular grid and the projections are used to find local cross-correlation peaks in a cryo-EM micrograph. Since  
<sup>48</sup> locations of the biomolecule in the Z-direction causes predictable aberrations to the projection image, this method  
<sup>49</sup> can be used to calculate 3D coordinates and orientations of a biomolecule in a cellular sample [8]

<sup>50</sup> Hematopoiesis is the process of generating the various cell types of the blood in the bone marrow. Disregulation  
<sup>51</sup> of the process results in diseases like leukemia. Understanding how hematopoietic stem and progenitor cells are  
<sup>52</sup> programmed to differentiate to the appropriate cell type would provide new insight how hematopoiesis can be  
<sup>53</sup> misregulated. Of special interest is the regulation of translation during hematopoiesis. This is exemplified by the  
<sup>54</sup> observation that genetic defects in the ribosome machinery often leads to hematopoietic disease[10]. As such direct  
<sup>55</sup> quantification of ribosome location, number and conformational states could lead to new insight into hematopoietic  
<sup>56</sup> disease [11].

57 Here we apply 2D-template matching of ribosomes to cryo-FIB milled neutrophil-like murine cells [12]. To in-  
58 crease the amount of collected data and to provide un-biased sampling of the whole lamella we devised a new  
59 data-acquisition scheme, Defocus-corrected large area cryo-electron microscopy (DeCo-LACE). We characterize  
60 aberration cause by the used large beam-image shifts and highly focused beams and find that they can be ade-  
61 quately correct to enable ribosome detection by 2D-template matching. The resulting data provides a description  
62 of ribosome distribution in the whole lamellae, which represent roughly 2% of the cellular volume. We find highly  
63 heterogeneous density of ribosome within the cell and can identify discrete clusters of presumably translationally  
64 active ribosomes, by testing for the presence of the small ribosomal subunit. The high accuracy of location and ori-  
65 entation of each detected ribosome also allows us to cluster ribosome molecules into potential polysomes. Analysis  
66 of the throughput in this method suggests that for the foreseeable future computation will be the bottleneck for  
67 visual proteomics.

## 68 Materials and Methods

### 69 Grid preparation

70 ER-HoxA9 cells were maintained in RPMI supplemented with 10% FBS, penicillin/streptomycin, SCF, and estrogen  
71 [12] at 37C and 5% CO2. 120h prior to grid freezing cells were washed twice in PBS and cultured in the same  
72 medium, except without estrogen. Differentiation was verified by staining with Hoechst-dye and inspection of  
73 nuclear morphology. Cells were then counted and diluted to  $1 \times 10^6$  cells/ml. Grids ( either 200 mesh copper grids,  
74 with a silicone-oxide and 2um holes with a 2um spacing or 200 mesh gold grids with a thin gold film and 2 um  
75 holes in 2um sapcing) were glow-discharged from both sides using a ... for ... . 3.5 ul of cells suspension was added to  
76 grids on the thin-film side and grids were automatically blotted from the back-side using a GP2 cryoplunger (Leica)  
77 for ... s and rapidly plunged into liquid ethane at -185C.

### 78 FIB-milling

79 Grids were loaded into a Acquilos 2 FIB/SEM microscope with a stage cooled to -190C. Grids were sputter-coated  
80 with platinum for 15s at 45 mA and then coated with a layer of platinum-precursor by opening the GIS-valve for  
81 45s. An overview of the grid was created by montaging SEM images and isolated cells at the center of gridsquares  
82 were selected for FB-milling. Lamella were generated automatically using the AutoTEM software, resulting in 6-8  
83 um wide lamella with 150-200 um thickness as determined by FIB-imaging of the lamella edge.

### 84 Data collection

85 Grids were loaded into a Krios Titam TEM operated at 300 keV. The microscope was setup with a cross-grating grid  
86 on the stageby setting the beam-diameter to 900 nm, resulting in the beam being completely visible in the camera.  
87 To establish fringe-free conditions, the “Fine eucentric” procedure of serialEM was used to move a square of the

88 cross-grating grid to the eucentric position of the microscope. The effective defocus was then set to 2 um, using the  
89 “autofocus” routine of serialEM. The objective focus of the microscope was changed until no fringes were visible.  
90 The stage was then moved in Z until images had a apparent defocus of 2 um. The differnce in stage Z-position  
91 between the eucentric and fringe-free conditions was calculate d and noted to move other areas into fringe-free  
92 condition.

93 Low magnification montages were used to find lamella and lamella that were sufficiently thin and free of contamina-  
94 tion were selected for automated data collection. The corners of the lamella were manually annotated in SerialEM  
95 and translated into Beam-Imageshift values using SerialEm calibration. A hexagonal patter of beam-imageshift  
96 positions was calculated that covered the area between he four corners in a serpentine way, with a  $\sqrt{3} * 400$  nm  
97 horizontal spacing and 800 nm vertical spacing. Exposures were then taking at each position with a 30 e/A total  
98 dose. After each exposure that defocus was estimated using the ctffind function of SerialEM and the focus for th  
99 next exposure was corrected by the difference between the estimated focus and the desired defocus of 800 um. Also  
100 after each exposure the deviation of the beam from the center of the camera was measured and corrected using the  
101 “CenterBeamFrom IMage” command of SerialEM.

102 After datacollection a 20s exposure at 2250x magnification of the lamella at 200um defocus was taken for visualiza-  
103 tion purposes.

#### 104 Data pre-processing

105 Movies were gain-corrected and motion-corrected using a custom version of unblur. To avoid influence of the beam-  
106 edge on motion-correction only a quarter of the movie in the center of the camera was considered for calculation of  
107 the estimated motion. After movie frames were aligned and summed a mask for the illuminated area was calculated  
108 by lowpass filtering the image at ... A, thresholding the image at 10% of the maximal value and then lowpass filtering  
109 the mask at ... A. This mask was then used to replace un-illuminated area with gaussian noise, with the same mean  
110 and standard deviation as the illuminated area. The contrast-transfer function (CTF) was estimated using ctffind,  
111 searching between 0.02 and 2 um defocus.

#### 112 Template matching

113 The search template was generated from the cryo-EM structure of the mouse large ribosomal subunit (PDB 6SWA).  
114 The ... subunit was deleted from the model and the simulate program of the cisTEM suite was used to calculate an  
115 density map from the atomic coordinates. The match\_template program was used to search for this template in  
116 the preprocessed images, using 1.5 deg angular step in out-of-plane angles and 1.0 deg in-plane. 21 defocus planes in  
117 200nm steps centered around the defocus estimates by ctffind were searches. Matches were defined as peaks above  
118 a threshold calulated according to ... .(7.75 for most images).

<sup>119</sup> **Data analysis**

<sup>120</sup> **Results**

<sup>121</sup> **2D-Template matching can be used to find ribosomal subunits in cryo-FIB thinned lamella of mam-**  
<sup>122</sup> **malian cells**

<sup>123</sup> To test whether we could detect individual ribosomes in mammalian cells we prepared cryo-lamella of mouse  
<sup>124</sup> neutrophil-like cells. Low-magnification images of these lamellas clearly shows cellular features consistent with a  
<sup>125</sup> neutrophil-like phenotype, mainly a segmented nucleus and a plethora of membrane-organelles, corresponding to  
<sup>126</sup> the granules and secretory vesicles of neutrophils. We then proceeded to acquire micrographs on this lamella with  
<sup>127</sup> a defocus of 0.5-1.0 um, 30 e/A<sup>2</sup>/s exposure and 1.5 Å pixelsize. We manually selected multiple locations in the  
<sup>128</sup> lamella and focused using standard low-dose techniques, i.e. by first ensuring correct focus by imaging a sacrificial  
<sup>129</sup> area. The resolute micrographs showed no signs of crystalline ice and had donut-rings to resolution, indicating  
<sup>130</sup> successful vitrification.

<sup>131</sup> We used an atomic model of the 60S mouse ribosomal subunit (6SWA) for 2D template matching. In a subset  
<sup>132</sup> of images the distribution of cross-correlation scores significantly exceeded the distribution expected from non-  
<sup>133</sup> significant matching (Figure 1B). In the resulting scaled maximum-intensity maps, clear peaks with SNR thresholds  
<sup>134</sup> up to 10 were apparent (Figure 1C). By using the criterion described by for thresholding potential matches we  
<sup>135</sup> found that in images of cytosolic compartments we found evidence of 10-500 ribosomes in the imaged areas. Notably  
<sup>136</sup> we found no matches in images that were taken in the nuclear compartment. In the cytosolic areas we found a  
<sup>137</sup> drastically different number of matches, in some areas we found only ~ 50 matches per image area, corresponding  
<sup>138</sup> to a concentration of ..., while in another area we found more than 500 matches, corresponding to a concentration  
<sup>139</sup> of ... .

<sup>140</sup> **cryo-EMILIA for 2D imaging of whole lamella**

<sup>141</sup> In order to obtain high-resolution data for complete lamella we used a new approach for data collection. This  
<sup>142</sup> approach uses three key strategies: (1) ensures that every electron that exposes the sample is collected on the  
<sup>143</sup> camera (2) uses beam-image shift to precisely and quickly raster the surface of the lamella and (3) uses a focusing  
<sup>144</sup> strategy that does not rely on a sacrificial area.

<sup>145</sup> To ensure that every electron exposing the sample was captured by the detector, we focused the electron beam so  
<sup>146</sup> that the entire beam was placed on the detector. During canonical low-dose imaging the microscope is configured  
<sup>147</sup> so that the focal plane is identical to the eucentric plane of the specimen stage. This leaves the C2 aperture out  
<sup>148</sup> of focus, resulting in ripples at the edge of the beam (Figure 2B). While these ripples are low-resolution features  
<sup>149</sup> that might not interfere with 2D template matching, which is designed to be robust to low-resolution noise, we also

150 tested collecting data under a condition where the C2 aperture is in focus (Figure 2C).

151 We then centered a lamella under the electron beam and used beam-image shift of the microscope to systematically  
152 raster the whole surface of the lamella in a hexagonal pattern. Instead of focusing in a sacrificial area, we determined  
153 the defocus after every exposure using a routine implemented in SerialEM modeled after CTFFind. The focus was  
154 then adjusted based on the difference between desired and measured defocus. Since we used a serpentine pattern  
155 for data collection every exposure is close to the previous exposure making drastic changes in the defocus unlikely.  
156 Furthermore we started our acquisition pattern on the platinum deposition edge, so the initial exposure where the  
157 defocus was not yet adjusted did not contain any biologically relevant information.

158 We used this strategy to collect data on 8 lamella, 4 using the eucentric focus condition and 4 using the fringe-free  
159 condition. We were able to highly consistently collect data with a defocus of 8 um (Figure 2D), both in the eucentric  
160 focus and fringe-free focus condition. Together with the nominal defocus of the microscope this data results in a  
161 topological map of the lamella. To ensure that data was collected consistently, we mapped defocus values as a  
162 function of the applied Beam-image shift (Figure 2E). This demonstrated that the defocus was consistent over the  
163 lamella, with outliers only at isolated images and in images containing contamination. We also plotted the measure  
164 objective astigmatism of the lamella and found that it varies with the applied Beam-image shift, become more  
165 astigmatic mostly due to beam-image shift in the X direction. While approaches exist to correct this during the  
166 data-collection, we opted to not use these mechanism for these early experiments and instead rely on computational  
167 correction of these aberrations in order to characterize them.

## 168 **2D-Template matching of cryo-EMILIA data reveals ribosome distribution**

169 We developed customized preprocessing protocol to images obtained by cryo-EMILIA to enable their use for 2D-  
170 template matching. First we restricted calculation of cross-correlation coefficients between individual movie frames  
171 to the central portion of the image to prevent artifacts from the beam edges on estimation of motion. Then we  
172 calculated a mask that defined the illuminated area of the micrographs and used it to fill non-illuminated areas with  
173 gaussian noise, matching mean and standard deviation to the illuminated signal (Figure 3A). The so processed  
174 images were suitable for 2D-template matching and we were able to obtain matches with the same model used  
175 for the data in Figure 1.

## 176 **Quantitative analysis of translation activity**

## 177 **Discussion**

- 178 • Elizabeth Wright and Grant Jensen Montage tomography papers  
179 • Waffle method for higher throughput, automation of fib-milling

- 180 • Throughput and bottlenecks  
 181 • Visual proteomics  
 182 • Granules containing ribosomes?  
 183 • Threshold implications (no matches on most images)

184 **Figures**

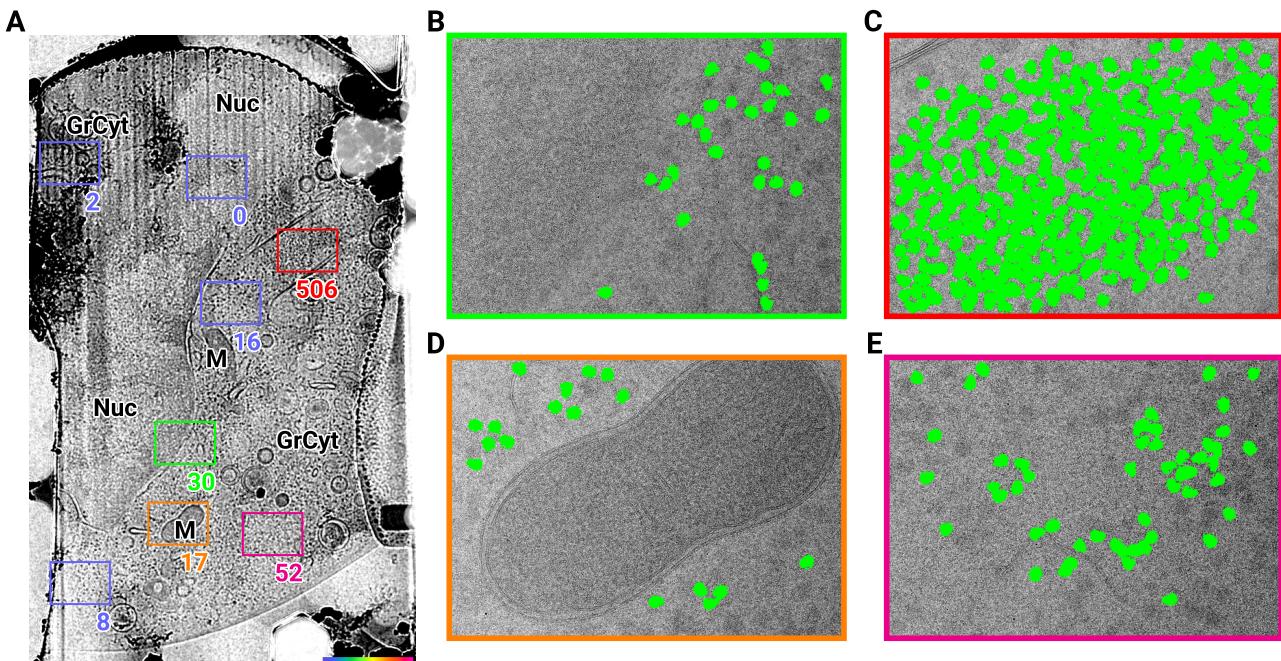


Figure 1: 2D template matching of the large subunit of the ribosome in fib-milled neutrophil-like cells (A) Overview image of the lamella. Major cellular regions are labeled, as Nucleus (Nuc), Mitochondria (M), and granular cytoplasm (GrCyt). FOVs where high-magnification images for template matching were acquired are indicated as boxes with the number of matches indicated on the bottom right. FOVs displayed in Panels B-E are color-coded. (B-E) FOVs with projection of ribosome LSU matches shown in green. (B) Perinuclear region, the only matches are in the cytoplasmic half. (C) Cytoplasmic region with high density of ribosomes (D) Mitochondrium, as expected there are only matches in the cytoplasmic region (E) Cytoplasm, with low density of ribosomes.

185 **References**

- 186 1. **Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis** Noelle M Griffin, Jingyi Yu, Fred Long, Phil Oh, Sabrina Shore, Yan Li, Jim A Koziol, Jan E Schnitzer  
*Nature Biotechnology* (2010-01) <https://doi.org/fshgnc> DOI: 10.1038/nbt.1592 · PMID: 20010810 · PMCID: PMC2805705

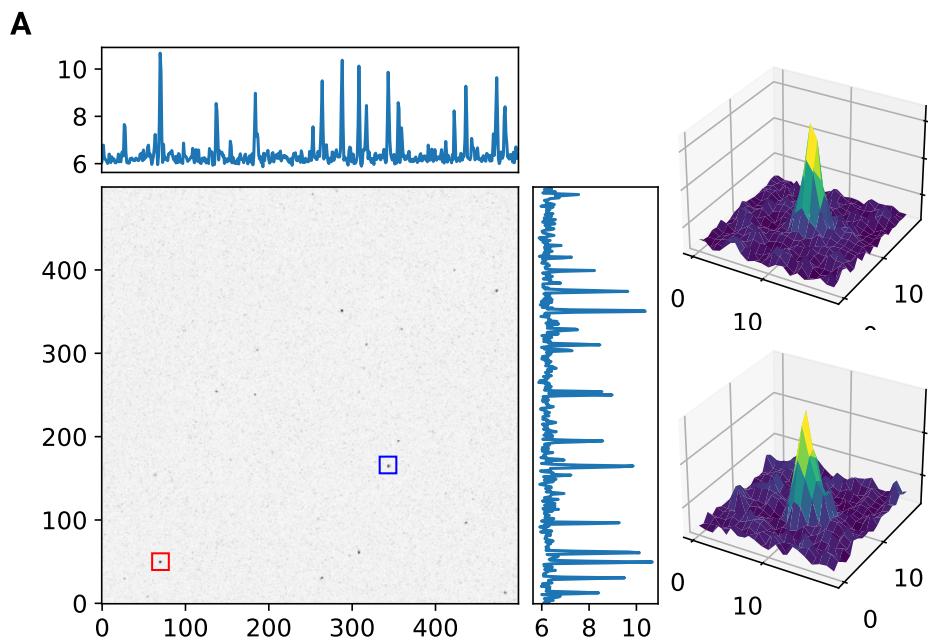


Figure S1: This is an example-figurern

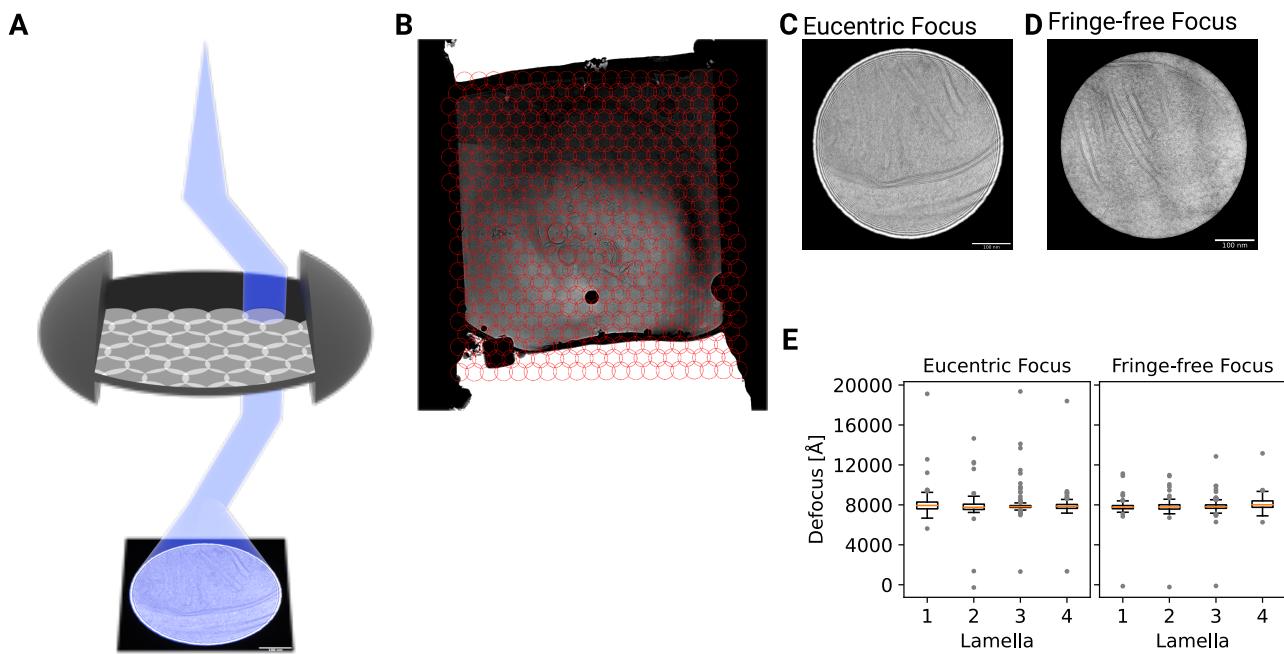


Figure 2: This is an example-figurern

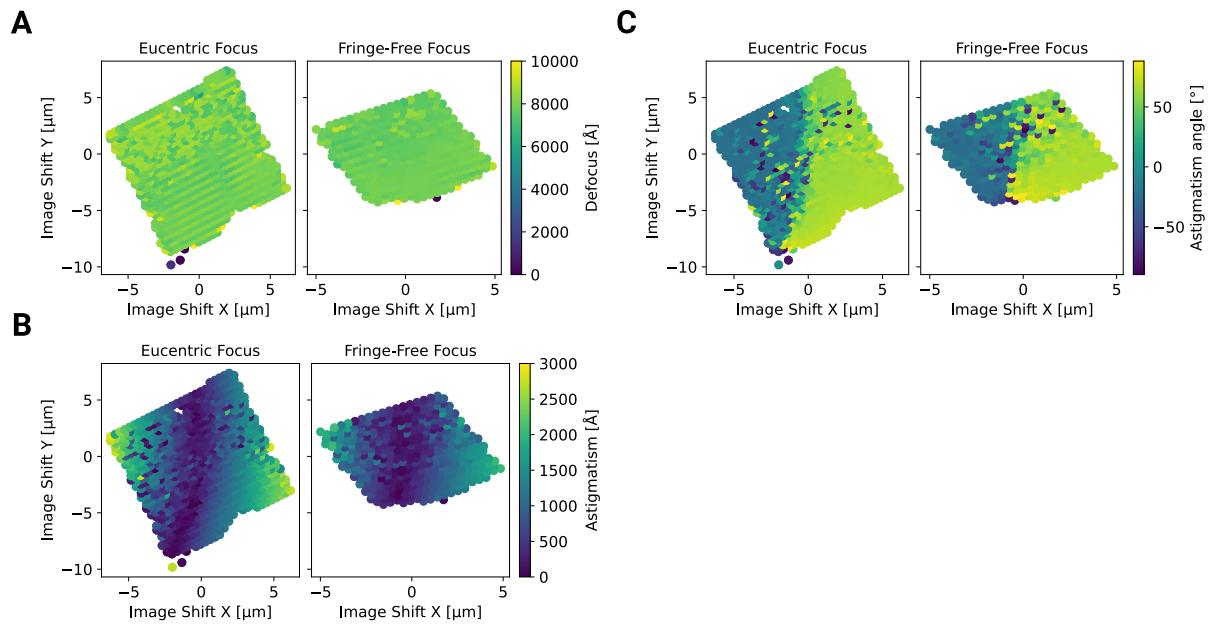


Figure 3: This is an example-figure

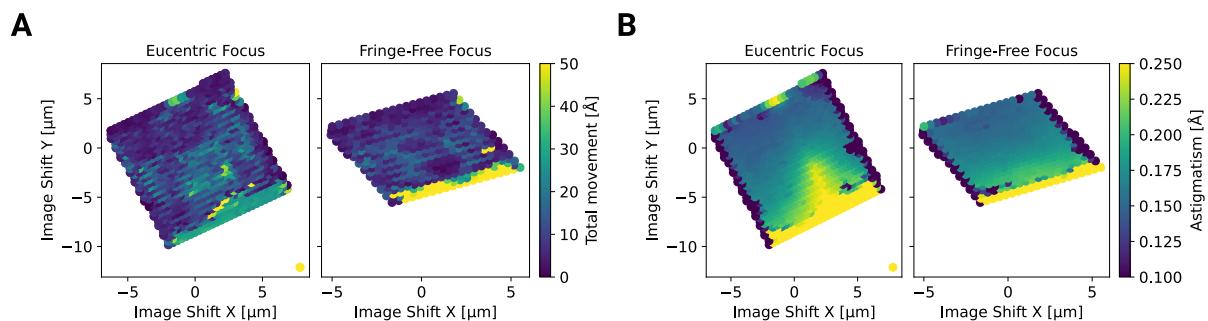


Figure 4: This is an example-figure

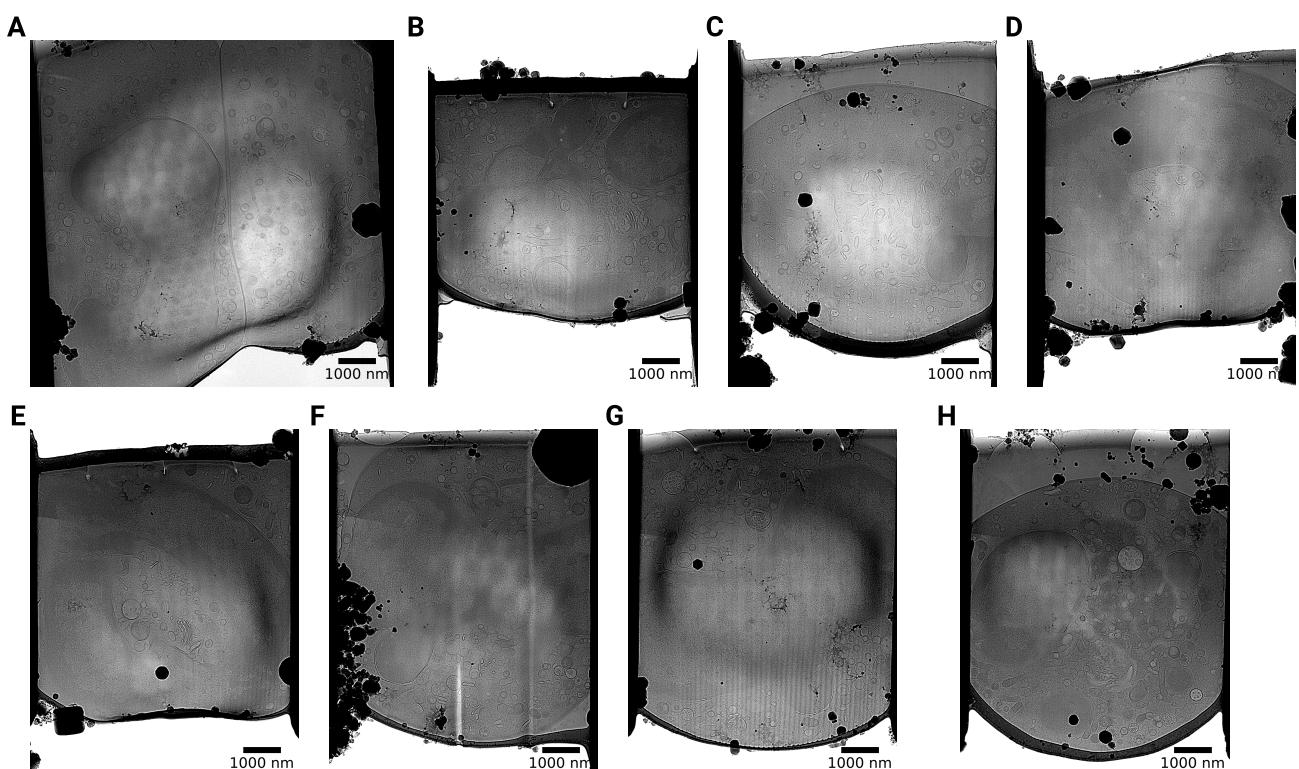


Figure 5: This is an example-figure



Figure 6: This is an example-figure

- 188 2. **Fluorescence microscopy** Jeff W Lichtman, José-Angel Conchello  
*Nature Methods* (2005-11-18) <https://doi.org/bbpg4n> DOI: 10.1038/nmeth817 · PMID: 16299476
- 189
- 190 3. **A visual approach to proteomics** Stephan Nickell, Christine Kofler, Andrew P Leis, Wolfgang Baumeister  
*Nature Reviews Molecular Cell Biology* (2006-02-15) <https://doi.org/d6d5mq> DOI: 10.1038/nrm1861 · PMID: 16482091
- 191
- 192 4. **Electron microscopy of frozen hydrated sections of vitreous ice and vitrified biological samples**  
AW McDowall, J-J Chang, R Freeman, J Lepault, CA Walter, J Dubochet  
*Journal of Microscopy* (1983-07) <https://doi.org/bdnzmv> DOI: 10.1111/j.1365-2818.1983.tb04225.x · PMID: 6350598
- 193
- 194 5. **Opening windows into the cell: focused-ion-beam milling for cryo-electron tomography**  
Elizabeth Villa, Miroslava Schaffer, Jürgen M Plitzko, Wolfgang Baumeister  
*Current Opinion in Structural Biology* (2013-10) <https://doi.org/f537jp> DOI: 10.1016/j.sbi.2013.08.006 · PMID: 24090931
- 195
- 196 6. **Electron tomography of cells** Lu Gan, Grant J Jensen  
*Quarterly Reviews of Biophysics* (2011-11-15) <https://doi.org/czj9hr> DOI: 10.1017/s0033583511000102 · PMID: 22082691
- 197
- 198 7. **Single-protein detection in crowded molecular environments in cryo-EM images** JPeter Rickgauer, Nikolaus Grigorieff, Winfried Denk  
*eLife* (2017-05-03) <https://doi.org/gnq4q4> DOI: 10.7554/elife.25648 · PMID: 28467302 · PMCID: PMC5453696
- 199
- 200 8. **Label-free single-instance protein detection in vitrified cells** JPeter Rickgauer, Heejun Choi, Jennifer Lippincott-Schwartz, Winfried Denk  
*Cold Spring Harbor Laboratory* (2020-04-24) <https://doi.org/gpbjfd> DOI: 10.1101/2020.04.22.053868

- 202 9. **Locating macromolecular assemblies in cells by 2D template matching with cisTEM** Bronwyn  
A Lucas, Benjamin A Himes, Liang Xue, Timothy Grant, Julia Mahamid, Nikolaus Grigorieff  
*eLife* (2021-06-11) <https://doi.org/gkkc49> DOI: 10.7554/elife.68946 · PMID: 34114559 · PMCID:  
PMC8219381
- 203
- 204 10. **Hallmarks of ribosomopathies** Kim R Kampen, Sergey O Sulima, Stijn Vereecke, Kim De Keersmaecker  
*Nucleic Acids Research* (2019-07-27) <https://doi.org/gpbjfm> DOI: 10.1093/nar/gkz637 · PMID: 31350888  
· PMCID: PMC7026650
- 205
- 206 11. **Diagnostic and prognostic implications of ribosomal protein transcript expression patterns in  
human cancers** James M Dolezal, Arie P Dash, Edward V Prochownik  
*BMC Cancer* (2018-03-12) <https://doi.org/gc87j9> DOI: 10.1186/s12885-018-4178-z · PMID: 29530001 ·  
PMCID: PMC5848553
- 207
- 208 12. **Inhibition of Dihydroorotate Dehydrogenase Overcomes Differentiation Blockade in Acute  
Myeloid Leukemia** David B Sykes, Youmna S Kfoury, François E Mercier, Mathias J Wawer, Jason M  
Law, Mark K Haynes, Timothy A Lewis, Amir Schajnovitz, Esha Jain, Dongjun Lee, ... David T Scadden  
*Cell* (2016-09) <https://doi.org/f3r5jr> DOI: 10.1016/j.cell.2016.08.057 · PMID: 27641501 · PMCID:  
PMC7360335

209