

In situ polyadenylation enables spatial mapping of the total transcriptome

David W. McKellar¹, Madhav Mantri², Meleana Hinchman³, John S.L. Parker³, Praveen Sethupathy⁴, Benjamin D. Cosgrove^{1,*}, Iwijn De Vlaminck^{1,*}

¹Meinig School of Biomedical Engineering, Cornell University, Ithaca, NY, USA

²Department of Computational Biology, Cornell University, Ithaca, NY, USA

³Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA

⁴Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA

* Co-corresponding authors: vflaminck@cornell.edu (I.D.V) and bdc68@cornell.edu (B.D.C.)

INTRODUCTION

Spatial transcriptomics reveals the spatial context of gene expression, but current methods are limited to assaying polyadenylated (A-tailed) RNA transcripts. Here we demonstrate that enzymatic *in situ* polyadenylation of RNA enables detection of the full spectrum of RNAs, expanding the scope of sequencing-based spatial transcriptomics to the total transcriptome. We apply this Spatial Total RNA-Sequencing (STRS) approach to study skeletal muscle regeneration and viral-induced myocarditis. Our analyses reveal the spatial patterns of noncoding RNA expression with near-cellular resolution, identify spatially defined expression of noncoding transcripts in skeletal muscle regeneration, and highlight host transcriptional responses associated with local viral RNA abundance. *In situ* polyadenylation requires the addition of only a single step to a widely used protocol for spatial RNA-sequencing, and thus could be broadly and quickly adopted. Spatial RNA-sequencing of the total transcriptome will enable new insights into spatial gene regulation and biology.

MAIN TEXT

Spatial transcriptomics provides insight into the spatial context of gene expression¹⁻⁵. Current methods are restricted to capturing polyadenylated transcripts and are not sensitive to many species of non-A-tailed RNAs, including microRNAs, newly transcribed RNAs, and many non-host RNAs. Extending the scope of spatial transcriptomics to the total transcriptome would enable observation of spatial distributions of regulatory RNAs and their targets, link non-host RNAs and host transcriptional responses, and deepen our understanding of spatial biology.

Here, we demonstrate Spatial Total RNA-Sequencing (STRS), a method that enables spatial profiling of both the A-tailed and non-A-tailed transcriptome. This is achieved with a simple modification of a commercially available protocol for spatial RNA-sequencing. STRS uses poly(A) polymerase to add poly(A) tails to RNAs *in situ*. STRS otherwise follows conventional protocols to capture, spatially barcode, and sequence RNAs. STRS is compatible with existing approaches for sequencing-based spatial transcriptomics, is straightforward to implement, and adds minimal cost and time to an already widely used commercially available workflow. STRS enables the capture of many RNAs that are missed by conventional workflows, including noncoding RNAs, newly transcribed RNAs, and viral RNAs. To demonstrate the versatility of the method, we applied STRS to study the regeneration of skeletal muscle after injury and the pathogenesis of viral-induced myocarditis.

RESULTS

In situ polyadenylation enables capture of coding and noncoding RNAs

STRS adds a single step to a commercially available method for spatial RNA-sequencing (Visium Spatial

49 Gene Expression, 10x Genomics) to capture the total transcriptome⁶. As in the Visium method, the
50 sample is first sectioned, fixed with methanol, and stained for histology. After imaging, the sample is
51 rehydrated and then incubated with yeast poly(A) polymerase for 25 minutes at 37°C. This enzyme adds
52 poly(A) tails to the 3' end of all RNAs so that endogenous poly(A) tails are extended and non-A-tailed
53 transcripts are polyadenylated. After *in situ* polyadenylation, STRS again follows the Visium protocol
54 without modification (**Fig 1a**). One important feature of the Visium method that we leverage in STRS, is
55 its use of a strand-aware library preparation. We found that strandedness is critical for the study of
56 noncoding and antisense RNAs (see below) and must be considered in bioinformatic analyses (**Fig S1**).

57
58 To test the performance and versatility of STRS, we applied it to two distinct mouse tissue types: injured
59 hindlimb muscle⁵ and virally infected heart tissue⁴. We quantified the percentage of unique molecules
60 (UMIs) as a function of RNA biotype (GENCODE M28 annotations; **Fig 1b**). Compared to the Visium
61 method, we found similar counts for protein coding and other endogenously polyadenylated transcripts
62 (**Fig S1-2**). STRS enabled robust detection of several types of noncoding RNAs which are poorly
63 recovered or not detected at all by the Visium method, including ribosomal RNAs (rRNAs; mean of 5.4%
64 and 2.6% of UMIs for STRS and Visium respectively; computed across all Visium and STRS samples
65 included in this study), microRNAs (miRNAs; 0.4% in STRS versus 0.004% in Visium), transfer RNAs
66 (tRNAs; 0.4% in STRS versus 0.02% in Visium), small nucleolar RNAs (snoRNAs; 0.2% in STRS versus
67 0.002% in Visium), and several other biotypes (**Fig 1b, Fig S3-4**). STRS libraries also had an increased
68 fraction of unspliced transcripts (2.7% in Visium versus 18.3% in STRS). Unspliced or nascent RNA
69 counts have been used to predict transcriptional trajectories for single cells. Improved detection of
70 nascent RNAs may enable more accurate trajectory imputation and reveal the dynamics of spatial gene
71 expression. Finally, STRS libraries had an increased fraction of reads which map to intergenic regions,
72 reflecting increased capture of unannotated transcriptional products (22.2% in STRS versus 9.5% in
73 Visium; **Fig S1b-c**). We found that STRS captured many RNAs which were not present in Visium libraries.
74 Many of these features map outside of or antisense to known annotations (**Fig 1c**). We also found that
75 STRS detected many noncoding transcripts which are intragenic to other genes (**Fig 1c**). Standard short-
76 read sequencing was sufficient to delineate these features from the surrounding host genes, as reflected
77 by the expression count matrices for STRS versus Visium data (**Fig 1d**). Most importantly, we were able
78 to spatially map each of these features and visualize spatial patterns of gene expression (**Fig 1e**). We
79 found that features which were incompletely annotated (*ENSMUSG00002075551*) showed sparse spatial
80 expression. Several highly abundant genes showed homogenous patterns of expression, reflecting
81 putative (*Gm42826*) or known (*7SK*) housekeeping roles⁷.

82
83 We also asked whether *in situ* polyadenylation enables capture of non-A-tailed viral RNA. To this end,
84 we assayed murine heart tissues infected with Type 1-Lang reovirus (REOV), a segmented double-
85 stranded RNA virus that expresses ten transcripts which are not polyadenylated. No reovirus transcripts
86 were detected with the Visium workflow, whereas STRS enabled detection of more than 200 UMIs
87 representing all ten reovirus gene segments (**Fig 1f**). To deeply profile viral RNAs, we performed targeted
88 enrichment of viral-derived cDNA from the final sequencing libraries and re-sequenced the products. This
89 enrichment led to a further ~26-fold increase of the mean viral UMIs captured per spot (minimum L1
90 segment with 262 UMIs, maximum S4 segment with 1095 UMIs). Taken together, these findings
91 demonstrate that STRS enables the study of many types of RNAs that are not detectable with existing
92 technologies.

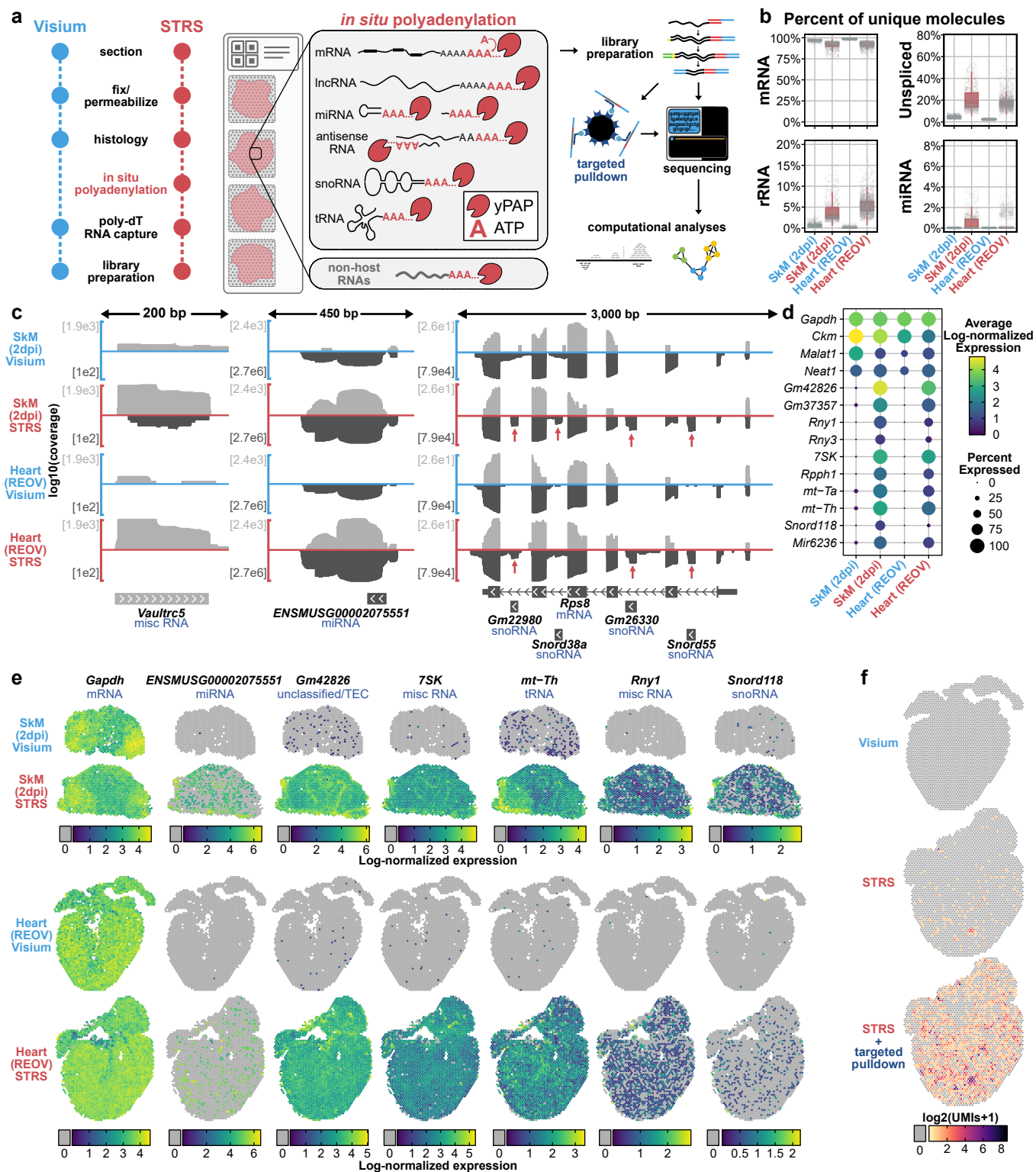


Figure 1. *In situ* polyadenylation enables spatial profiling of noncoding and non-host RNAs. **(a)** Workflow for Spatial Total RNA-Sequencing (STRS). **(b)** Comparison of select RNA biotypes between Visium and STRS datasets. Y-axis shows the percent of unique molecules (UMIs) for each spot. **(c)** Detection of coding and noncoding RNAs between Visium and STRS workflows. Color scale shows average log-normalized UMI counts. Dot size shows the percent of spots in which each RNA was detected. **(d)** Log₁₀-transformed coverage of deduplicated reads mapping to sense (light gray) and antisense (dark gray) strands at the *Vaultrc5*, *ENSMUSG00002075551*, and *Rps8* loci. Annotations shown are from GENCODE M28 and include one of the five isoforms for *Rps8* as well as the four intragenic features within introns of *Rps8*. **(e)** Spatial maps of coding and noncoding transcripts for Visium and STRS workflows. Spots in which the transcript was not detected are shown as gray. **(f)** Detection of reovirus transcripts using the standard workflow, STRS, and STRS with targeted pulldown enrichment. Spots in which the virus was not detected are shown as gray.

94 *Spatial total RNA-sequencing reveals spatial patterns of gene regulation in skeletal muscle regeneration*
95 Skeletal muscle regeneration is a coordinated system guided by complex gene regulatory networks^{5,8–12}.
96 We applied STRS to spatially map the coding and noncoding transcriptome in a mouse model of skeletal
97 muscle regeneration. We injured tibialis anterior muscles and then collected tissues at 2-, 5-, and 7-days
98 post-injury (dpi) in addition to an uninjured control (**Methods**). H&E imaging showed immune infiltration
99 centrally within tissue sections at 2 and 5dpi, which was mostly resolved by 7dpi (**Fig 2a**). Unsupervised
100 clustering identified spots in the injury loci, spots around the border of the injury loci, and spots under
101 intact myofibers (**Fig 2b, Methods**).

102
103 We performed differential gene expression analysis across the regional clusters to identify noncoding
104 RNAs specific to the injury locus (**Fig 2c, Methods**). We found several RNAs which were spatiotemporally
105 associated with injury locus, many of which are undetected or poorly detected by Visium (**Fig 2c-d**). *Meg3*
106 is an endogenously polyadenylated lncRNA which has been shown to regulate myoblast differentiation
107 in vitro. We found *Meg3* expression to be confined to the injury locus at 5dpi, when myoblast
108 differentiation and myocyte fusion occurs^{5,13}. *Gm10076*, a transcript with a biotype annotation conflict
109 (Ensembl: lncRNA; NCBI: pseudogene) and no known function, was highly and specifically expressed
110 within the injury locus 2dpi. *Gm10076* expression was reduced but still localized to the injury site by 5dpi
111 and returned to baseline levels by 7dpi. *Rpph1*, a ribozyme and component of the RNase P
112 ribonucleoprotein which has also been shown to play roles in tRNA and lncRNA biogenesis^{14,15}, showed
113 broad expression by 2dpi which peaked and localized to the injury site at 5 dpi. We also found that STRS
114 captured high levels of antisense transcripts for *Rpph1* which were not detected by the Visium chemistry.
115 This demonstrates that STRS can robustly profile both polyadenylated and non-polyadenylated RNAs
116 across heterogeneous tissues.

117
118 The role of miRNAs in skeletal muscle regeneration has been well-established^{11,16–18}. Mature miRNAs¹⁹
119 are ~22 nucleotides long, are not polyadenylated, and are not captured by the standard Visium workflow
120 (**Fig S5**). We asked if STRS was able to detect mature miRNAs. We first generated matched bulk small
121 RNAseq libraries from entire tibialis anterior muscles as a gold standard reference (n=2 per timepoint).
122 We used miRge3.0²⁰ to quantify mature miRNA abundance in the STRS and matched small RNA-
123 sequencing libraries (**Methods**). We found strong correlation in the abundance of the most highly
124 expressed miRNAs between STRS and small RNA-sequencing, with only minor drop-out of lowly
125 expressed miRNAs (**Fig 2e, Fig S5**). We identified many examples of mature miRNA expression in STRS
126 data, including expression of classic “myomiRs”, *miR-1a-3p*, *miR-133a/b-3p*, and *miR-206-3p* (**Fig 2f**)²¹.
127 Consistent with previous studies²², we detected static expression of *miR-1a-3p* across all four timepoints
128 (**Fig 2d**), whereas *miR-206-3p* was highly expressed within the injury locus five days post-injury, with
129 very low levels of expression detected at other timepoints.

130
131 *Spatial total transcriptomics spatially resolves viral infection of the murine heart*
132 We next explored the potential for STRS to profile host-virus interactions in a mouse model of viral-
133 induced myocarditis. We orally infected neonatal mice with type 1-Lang reovirus (REOV), a double-
134 stranded RNA virus with gene transcripts that are not polyadenylated. Within seven days of oral infection,
135 REOV spreads to the heart and causes myocarditis^{23–25}. We performed Visium and STRS on hearts
136 collected from REOV-infected and saline-injected control mice (**Fig 3a**). We found that reovirus
137 transcripts were only detected in the infected heart via STRS and that targeted enrichment of reovirus
138 transcripts enabled deeper profiling of viral infection (**Fig 1d, 3a; Methods**). Mapping these reads across
139 the tissue revealed pervasive infection across the heart (1,329/2,501 or 53% spots under the tissue; **Fig**
140 **1d**). Foci containing high viral UMI counts overlapped with myocarditic regions as identified by histology.

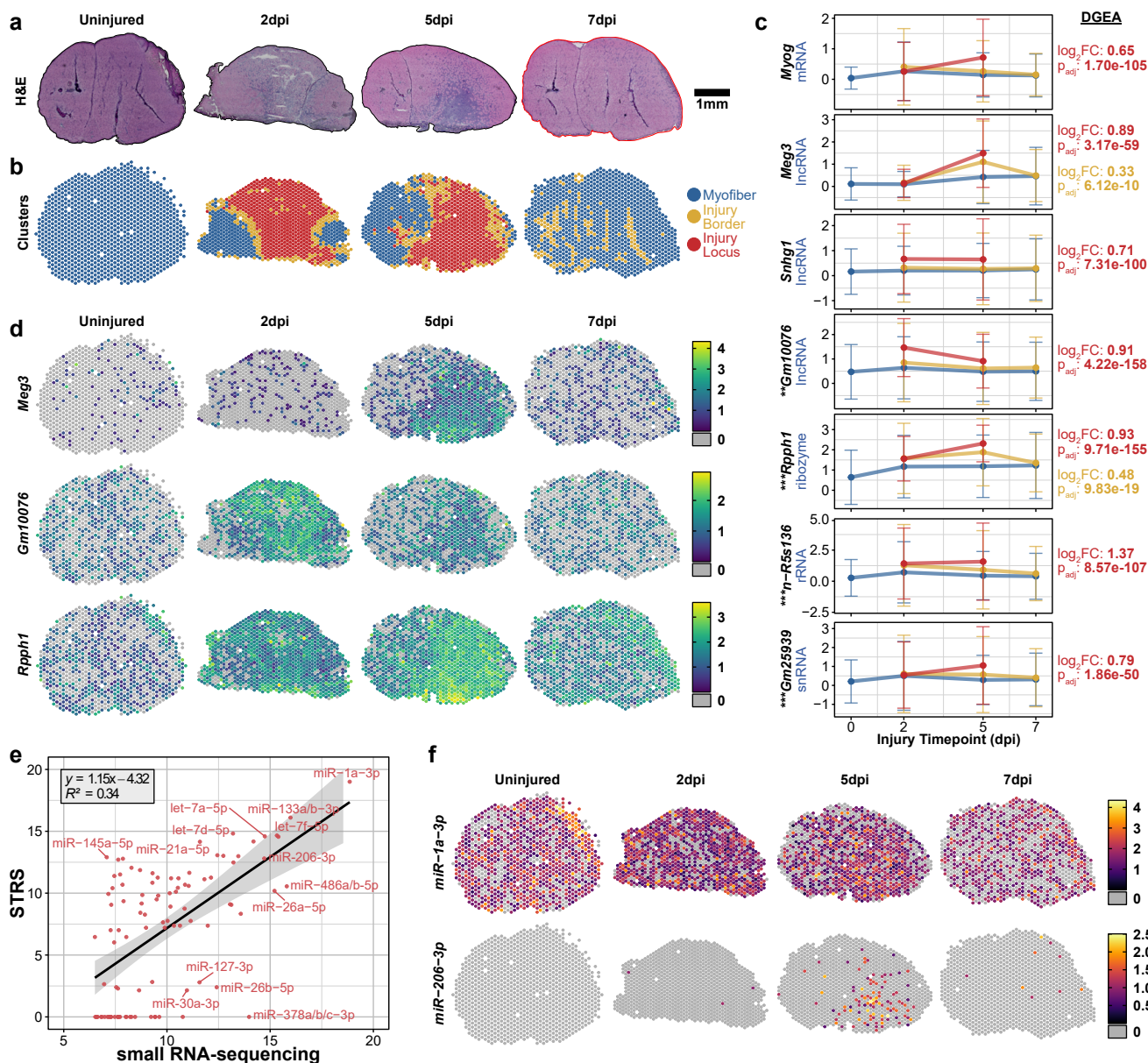


Figure 2. Spatial total RNA-sequencing of regenerating skeletal muscle **(a)** H&E histology of mouse tibialis anterior muscles collected 2-, 5-, and 7-days post-injury (dpi). **(b)** Clustering of spot transcriptomes based on total transcriptome repertoires (see **Methods**). **(c)** Differentially expressed RNAs across regional clusters. Y-axis shows log-normalized expression of each feature. Mean expression across each cluster is reported, colored according to the legend in (b). Error bars show standard deviation. Reported statistics to the right of plots reflect differential gene expression analysis performed across clusters on merged STRS samples (Wilcoxon, see **Methods**). Asterisks next to transcript names reflect differential expression analysis performed across skeletal muscle Visium and STRS samples (**p_{val_adj}<10⁻⁵⁰, ***p_{val_adj}<10⁻¹⁵⁰; Wilcoxon, see **Methods**) **(d)** Spatial maps for select features from (c). **(e)** Mature miRNA expression detected by STRS. Color scale shows log-normalized miRNA counts, quantified by miRge3.0 (**Methods**). **(f)** Average detection of miRNAs compared between small RNA-sequencing (n=8) and STRS (n=4). Axes show log₂ counts per million transcripts, normalized to the total number of transcripts which map to small RNA loci with miRge3.0. The top 100 most abundant miRNAs detected by small RNA-sequencing are shown.

141
 142 We next compared the read coverage profiles across the ten REOV gene segments for REOV-enriched
 143 libraries from Visium and STRS samples (**Fig 3d**). As expected, STRS libraries had a peak in coverage
 144 at the 3' end of viral gene segments. In contrast, the REOV-enriched Visium reads contained peaks in
 145 the middle of viral gene segments as expected for a chemistry that relies on the spurious capture of viral

146 RNA at poly(A) repeats within the transcripts²⁶. Interestingly, we found that STRS led to an
147 overrepresentation of reads from the 5' end of the sense [+] strand of all ten REOV segments. These
148 reads may represent incomplete transcripts generated by transcriptional pausing of the REOV RNA
149 polymerase or transcripts undergoing 3' exonucleolytic degradation. Finally, we detected the 3' end of
150 the antisense [-] strand for nine of the ten segments of the reovirus genome, suggesting that STRS
151 captures both strands of the dsRNA reovirus genome (**Fig 3d**). These antisense reads were present at
152 an average ratio of ~1:40 compared to the sense reads. The current model for synthesis of reovirus
153 dsRNA posits that dsRNA synthesis only occurs within a viral core particle after packaging of the ten viral
154 positive-sense RNAs. There are several possible explanations for our detection of the antisense strands.
155 One is that we are detecting negative-strand viral RNA that is part of dsRNA that has been released from
156 damaged viral particles either within the cytoplasm or within lysosomes. dsRNA released within
157 endolysosomes can be transported into the cytoplasm by RNA transmembrane receptors SIDT1 and
158 SIDT2^{27,28}. A second possibility is that antisense [-] viral RNA is synthesized prior to packaging of dsRNA
159 into viral particles.

160
161 Because STRS efficiently recovers viral RNA, we were able to directly correlate host transcriptomic
162 responses with viral transcript counts for spots in inflamed regions (generalized additive model,
163 **Methods**). We found inflammation-associated cytokine transcripts such as *Ccl2* and *Cxcl9*, and immune
164 cell markers such as *Gzma* and *Trbc2* to be upregulated in spots with high viral counts (**Fig 3e**). We
165 continued this analysis by performing unsupervised clustering (**Fig 3b**) and differential gene expression
166 analysis to identify transcripts associated with infection which are more readily detected by STRS (**Fig**
167 **3c**). *AW112010*, which has recently been shown to regulate inflammatory T cell states, was only found
168 in infected samples and was more abundant in the STRS data compared to Visium. We also found that
169 STRS led to increased detection of putative protein-coding genes, including *Ly6a2*, *Cxcl11*, and *Mx2*,
170 which were associated with infection. Interestingly, all three genes are annotated as pseudogenes in
171 GENCODE annotations but have biotype conflicts with other databases. The increased abundance as
172 measured by STRS could reflect differential mRNA polyadenylation for these transcripts. Overall, STRS
173 enabled more robust, spatially mapped analysis of the host response to infection by increasing the
174 breadth of captured transcript types and by providing direct comparison with viral transcript abundance.

175 **DISCUSSION**

177 Here, we demonstrate *in situ* polyadenylation of RNA in sectioned tissue to enable Spatial Total RNA-
178 Sequencing. Enzymatic polyadenylation is frequently implemented for bulk sequencing of total RNA and
179 was recently adopted for single-cell RNA-sequencing^{29,30}. STRS is the first implementation of *in situ* RNA-
180 labeling for spatial RNA-sequencing. STRS has several notable features. First, STRS is compatible with
181 a commercial workflow and requires the use of only one additional reagent. STRS can easily be adopted
182 by others as it requires minimal additional experimental time (~30 minutes) and cost (<\$90 per sample)
183 and does not require any specialized equipment. Second, because our RNA-labeling strategy is designed
184 to work with poly(dT) reverse transcription, STRS is compatible with other sequencing-based spatial
185 transcriptomics platforms. The resolution of our analyses is limited by the size and distribution of the
186 barcoded spots on the Visium slides. Future iterations of STRS which use higher resolution RNA-capture
187 platforms, including Slide-Seqv³¹, BGI Stereo-seq³², or new versions of Visium, promise substantial
188 improvements in spatial resolution. Because STRS is not targeted and does not require prior sequence
189 information, it is easily adapted to new biological systems and is well-suited for assaying unknown RNAs,
190 including novel RNAs or non-host transcripts.

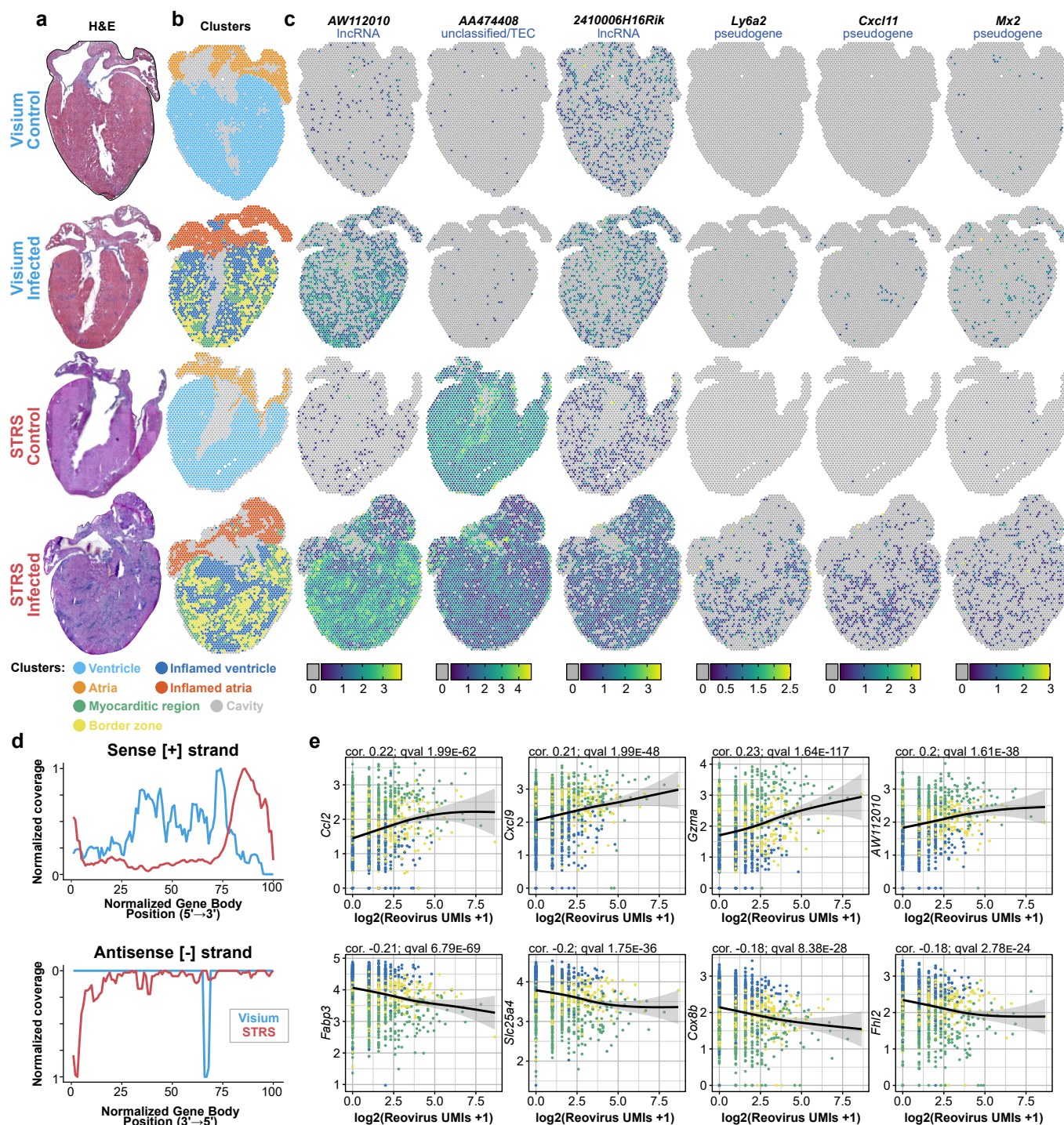


Figure 3. STRS enables simultaneous analysis of viral infection and host response. **(a)** H&E staining of control and reovirus-infected hearts, collected using the standard Visium workflow and STRS. **(b)** Tissue regions identified through unsupervised clustering of spot transcriptomes. **(c)** Log-normalized expression of noncoding and coding RNAs which are highly expressed in myocarditic regions. Spots in which transcripts were not detected are shown in gray. **(d)** Normalized coverage of deduplicated reads for the sense [+] and antisense [-] strands of all ten reovirus gene segments. X-axis shows the length-normalized position across the gene bodies of all ten reovirus segments. Note that the peak in antisense [-] coverage for the Visium sample (blue) corresponds to only 11 total reads. **(e)** Co-expression of pulldown-enriched reovirus UMIs versus infection-associated genes in spots underneath inflamed and myocarditic tissue. Spots are colored according to legend in (b). Correlation and q-value reported are from general additive model analysis (**Methods**).

193 We investigated the utility and versatility of STRS by applying it to two distinct models. First, we profiled
194 the noncoding RNA repertoires of infiltrating immune cells and regenerating myogenic cells at injury loci
195 in mouse muscle. Second, we analyzed the host transcriptome in response to mammalian orthoreovirus
196 infection. Members of the Reoviridae family of viruses synthesize non-polyadenylated viral mRNAs as do
197 Arenaviruses, and flaviviruses³³⁻³⁵. Because STRS can directly capture viral RNAs, we could directly
198 compare viral RNA abundance to gene expression changes in heart tissue. This enabled identification of
199 infection-related lncRNAs which were not detectable using standard techniques. Adding spatial context
200 has clarified the underlying biology of gene expression measurements. STRS improves on these facets
201 by extending the assayable transcriptome and enabling direct measurements of viral-derived RNA
202 transcripts.

203
204 With STRS, we demonstrated the first method to simultaneously map miRNAs and the mRNAs on which
205 they act. Because of their short length and known biases in adapter ligation, miRNAs are notoriously
206 difficult to assay^{36,37}. Furthermore, the Visium Gene Expression protocol uses a tagmentation-based
207 library preparation which depletes short molecules by either cutting the UMI/spot barcode or by producing
208 a read which is too short to confidently align to the genome. Despite these issues, we showed robust
209 detection for several known myomiRs and strong correlation with a gold-standard bulk method that does
210 not suffer from ligation or length biases. With future improvements to the library preparation strategy,
211 many of these hurdles can be further reduced.

212
213 This work highlights opportunities for improvements in current bioinformatic tools and resources for
214 single-cell and spatial transcriptomics. Current alignment and transcript counting tools are not optimized
215 for total RNA data and genome annotations are incomplete outside of protein coding genes. Furthermore,
216 new tools that go beyond UMI counts and better leverage the wealth of information in sequence read
217 alignment patterns are likely to be highly impactful.

218
219 In conclusion, we have demonstrated a versatile strategy for spatial mapping of the total transcriptome.
220 We think STRS will expand the scope of spatial transcriptomics and enable new types of analyses on
221 spatial gene regulation at tissue scale.

222

223 **METHODS**

224 *Mice*

225 The Cornell University Institutional Animal Care and Use Committee (IACUC) approved all animal
226 protocols, and experiments were performed in compliance with its institutional guidelines. For skeletal
227 muscle samples, adult female C57BL/6J mice were obtained from Jackson Laboratories (#000664; Bar
228 Harbor, ME) and were used at 6 months of age. For heart samples, confirmed pregnant female C57BL/6J
229 mice were ordered from Jackson Laboratories to be delivered at embryonic stage E14.5.

230

231 *Viral infection*

232 Litters weighing 3 gram/ pup were orally gavaged using intramedic tubing (Becton Dickinson Cat #427401
233 with 50 µl with 10⁷ PFU reovirus type 1-lang (T1L) strain in 1X phosphate buffered saline (PBS) containing
234 green food color (McCormick) via a 1ml tuberculin slip tip syringe (BD 309659) and 30G x 1/2 needle (BD
235 Cat #305106). Litters treated with 1X PBS containing green food color alone on the same day were used
236 as mock controls for the respective infection groups. The mock-infected and reovirus-infected pups were
237 monitored and weighed daily until the time points used in the study (7 days post infection). After
238 dissection, samples were embedded in O.C.T. Compound (Tissue-Tek) and frozen fresh in liquid
239 nitrogen.

240

241

242 *Muscle injury*

243 To induce muscle injury, both tibialis anterior muscles of old (20 months) C57BL/6J mice were injected
244 with 10 μ l of notexin (10 μ g/ml; Latoxan; France). Either before injury or 2-, 5-, or 7-days post-injury (dpi),
245 mice were sacrificed and tibialis anterior muscles were collected. After dissection, samples were
246 embedded in O.C.T. Compound (Tissue-Tek) and frozen fresh in liquid nitrogen.

247

248 *In situ polyadenylation and spatial total RNA-sequencing (STRS)*

249 Spatial total RNA-sequencing was performed using a modified version of the Visium protocol. 10 μ m thick
250 tissue sections were mounted onto the Visium Spatial Gene Expression v1 slides. For heart samples,
251 one tissue section was placed into each 6x6mm capture area. For skeletal muscle samples, two tibialis
252 anterior sections were placed into each capture area. After sectioning, tissue sections were fixed in
253 methanol for 20 minutes at -20°C. Next, H&E staining was performed according to the Visium protocol,
254 and tissue sections were imaged on a Zeiss Axio Observer Z1 Microscope using a Zeiss Axiocam 305
255 color camera. H&E images were shading corrected, stitched, rotated, thresholded, and exported as TIFF
256 files using Zen 3.1 software (Blue edition). After imaging, the slide was placed into the Visium Slide
257 Cassette. *In situ* polyadenylation was then performed using yeast Poly(A) Polymerase (yPAP; Thermo
258 Scientific, Cat #74225Z25KU). First, samples were equilibrated by adding 100 μ l of 1X wash buffer (20 μ l
259 5X yPAP Reaction Buffer, 2 μ l 40U/ μ l Protector RNase Inhibitor, 78 μ l nuclease-free H₂O) to each capture
260 area and incubating at room temperature for 30 seconds. The buffer was then removed. Next, 75 μ l of
261 yPAP enzyme mix (15 μ l 5X yPAP Reaction Buffer, 3 μ l of 600U/ μ l yPAP enzyme, 1.5 μ l 25mM ATP, 3 μ l
262 40U/ μ l Protector RNase Inhibitor, 52.5 μ l nuclease-free H₂O) was added to each reaction chamber. STRS
263 was also tested with 20U/ μ l of SUPERase-In RNase-Inhibitor, but we found that SUPERase was not able
264 to prevent degradation of longer transcripts during *in situ* polyadenylation (**Fig S6c-d**). The reaction
265 chambers were then sealed, and the slide cassette was incubated at 37°C for 25 minutes. The enzyme
266 mix was then removed. Prior to running STRS, optimal tissue permeabilization time for both heart and
267 skeletal muscle samples was determined to be 15 minutes using the Visium Tissue Optimization Kit from
268 10x Genomics. Following *in situ* polyadenylation, the standard Visium library preparation was followed to
269 generate cDNA and final sequencing libraries. The libraries were then pooled and sequenced according
270 to guidelines in the Visium Spatial Gene Expression protocol using either a NextSeq 500 or NextSeq
271 2000 (Illumina, San Diego, CA).

272

273 *Small RNA-sequencing*

274 For skeletal muscle samples, following the injury time course, tibialis anterior muscles were dissected
275 and snap frozen with liquid nitrogen. The Norgen Total RNA Purification Kit (Cat. 17200) was used to
276 extract RNA from 10 mg of tissue for each sample. For heart samples, following the infection time course,
277 hearts were dissected, embedded in OCT, and frozen in liquid nitrogen. RNA was extracted with Trizol
278 (Invitrogen, Cat. 15596026) and glycogen precipitation for a small fraction of each of the heart samples.
279 RNA quality was assessed via High Sensitivity RNA ScreenTape Analysis (Agilent, Cat. 5067-5579) and
280 all samples had RNA integrity numbers greater than or equal to 7.

281

282 Small RNA sequencing was performed at the Genome Sequencing Facility of Greehey Children's Cancer
283 Research Institute at the University of Texas Health Science Center at San Antonio. Libraries were
284 prepared using the TriLink CleanTag Small RNA Ligation kit (TriLink Biotechnologies, San Diego, CA).
285 Libraries were sequenced with single-end 50 \times using a HiSeq2500 (Illumina, San Diego, CA).

286

287 *Preprocessing and alignment of Spatial Total RNA-Sequencing data*

288 All code used to process and analyze these data can be found at <https://github.com/mckellardw/STRS>.
289 An outline of the pipelines used for preprocessing and alignment is shown in **Fig S1**.

290

291 Reads were first trimmed using cutadapt v3.4³⁸ to remove the following sequences: 1) poly(A) sequences
292 from the three prime ends of reads, 2) the template switch oligonucleotide sequence from the five prime
293 end of reads which are derived from the Visium Gene Expression kit (sequence:
294 CCCATGTA CTCTGCGTTGATACCACTGCTT), 3) poly(G) artifacts from the three prime ends of reads,
295 which are produced by the Illumina two-color sequencing chemistry when cDNA molecules are shorter
296 than the final read length, and 4) the reverse complement of the template switching oligonucleotide
297 sequence from the five prime ends of reads (sequence:
298 AAGCAGTGGTATCAACGCAGAGTACATGGG). Next, reads were aligned using either STAR v2.7.10a³⁹
299 or kallisto v0.48.0⁴⁰. Workflows were written using Snakemake v6.1.0⁴¹.

300
301 For STAR, the genomic reference was generated from the GRCm39 reference sequence using
302 GENCODE M28 annotations. For STAR alignment, the following parameters, based on work by Isakova
303 et al, were used: outFilterMismatchNoverLmax=0.05, outFilterMatchNmin=16,
304 outFilterScoreMinOverLread=0, outFilterMatchNminOverLread=0, outFilterMultimapNmax=50. Aligned
305 reads were deduplicated for visualization using umi-tools v1.1.2⁴².

306
307 For kallisto, a transcriptomic reference was also generated using the GRCm39 reference sequence and
308 GENCODE M28 annotations. The default k-mer length of 31 was used to generate the kallisto reference.
309 Reads were pseudoaligned using the `kallisto bus` command with the chemistry set to "VISIUM" and the
310 `fr-stranded` flag activated to enable strand-aware quantification. Pseudoaligned reads were then
311 quantified using bustools v0.41.0. First, spot barcodes were corrected with `bustools correct` using the
312 "Visium-v1" whitelist provided in the Space Ranger software from 10x Genomics. Next, the output bus
313 file was sorted and counted using `bustools sort` and `bustools count`, respectively. To estimate the
314 number of spliced and unspliced transcripts, reads pseudoaligned using kb-python v0.26.0, using the
315 "lemanno" workflow.

316
317 Spots were manually selected based on the H&E images using Loupe Browser from 10x Genomics.
318 Spatial locations for each spot were assigned using the Visium coordinates provided for each spot
319 barcode by 10x Genomics in the Space Ranger software ("Visium-v1_coordinates.txt"). Downstream
320 analyses with the output count matrices were then performed using Seurat v4.0.4^{43,44}. In addition to
321 manual selection, spots containing fewer than 500 detected features or fewer than 1000 unique
322 molecules were removed from the analysis. Counts from multimapping features were collapsed into a
323 single feature to simplify quantification.

324 325 *Mature microRNA quantification*

326 For STRS data: after trimming (see above), barcode correction with STAR v2.7.10a, and UMI-aware
327 deduplication with umi-tools v1.1.2, reads were split across all 4992 spot barcodes and analyzed using
328 miRge3.0 v0.0.9²⁰. Reads were aligned to the miRbase reference provided by the miRge3.0 authors.
329 MiRNA counts were log-normalized according to the total number of counts detected by kallisto and
330 scaled using a scaling factor of 1000. For small RNAseq data: Reads were first trimmed using trim_galore
331 v0.6.5. Reads were then aligned and counted using miRge3.0 v0.0.9.

332 333 *Unsupervised clustering and differential gene expression analysis of spot transcriptomes*

334 Spot UMI counts as generated by kallisto were used. First, counts were log-normalized and scaled using
335 default parameters with Seurat. Principal component analysis was then performed on the top 2000 most
336 variable features for each tissue slice individually. Finally, unsupervised clustering was performed using
337 the `FindClusters()` function from Seurat. The top principal components which accounted for 95% of
338 variance within the data were used for clustering. For skeletal muscle samples, a clustering resolution
339 was set to 0.8. For heart samples, clustering resolution was set to 1.0. Default options were used for all

340 other parameters. Finally, clusters were merged according to similar gene expression patterns and based
341 on histology of the tissue under each subcluster.

342
343 Differential gene expression analysis was performed using the `FindAllMarkers()` function from Seurat.
344 Default parameters were used, including the use of the Wilcoxon ranked sum test to identify differentially
345 expressed genes. To identify features enriched in the skeletal muscle STRS datasets, all Visium and
346 STRS were first merged and compared according to the method used (Visium vs. STRS). To identify
347 cluster-specific gene expression patterns, skeletal muscle samples were first clustered as described
348 above individually. STRS samples were then merged, and differential gene expression analysis was
349 performed across the three injury region groups.

350
351 *Targeted pulldown enrichment of viral fragments*
352 We performed hybridization-based enrichment of viral fragments on the Visium and STRS libraries for
353 reovirus-infected hearts using the xGen Hybridization and Wash Kit (IDT; 1080577)⁴. In this approach, a
354 panel of 5'-biotinylated oligonucleotides was used for capture and pulldown of target molecules of
355 interest, which were then PCR amplified and sequenced. We designed a panel of 202 biotinylated probes
356 tiled across the entire reovirus T1L genome to selectively sequence viral molecules from the sequencing
357 libraries (**Table S1**). After fragmentation and indexing of cDNA, 300ng of the final Visium or STRS
358 sequencing libraries from reovirus-infected hearts were used for xGen hybridization capture using the
359 xGen NGS Target Enrichment Kit protocol provided by the manufacturer. One round of hybridization
360 capture was performed for the STRS library followed by 14 cycles of PCR amplification. Because of the
361 reduced number of captured molecules, two rounds of hybridization were performed on the Visium
362 libraries. Enriched Visium libraries were PCR-amplified for 18 cycles after the first round of hybridization
363 and by 5 cycles after the second round of hybridization. Post-enrichment products were pooled and
364 sequenced on the Illumina NextSeq 500.

365
366 *Correlation analysis between reovirus counts and host gene expression*
367 We used a generative additive model (GAM) implemented in Monocle v2.18.0⁴⁵ to find genes that vary
368 with viral UMI count. A Seurat object for STRS data and viral UMI counts from the reovirus-infected heart
369 was converted to a CellDataSet object using the `as.CellDataSet()` command implemented in Seurat.
370 The expression family was set to "negative binomial" as suggested for UMI count data in the Monocle
371 documentation. The CellDataSet object was then preprocessed to estimate size factors and dispersion
372 for all genes. Genes expressed in fewer than 10 spots were removed. Within the remaining genes, we
373 then used the GAM implemented in the `differentialGeneTest()` command in Monocle to identify genes
374 that vary with log-transformed viral UMI counts. To find the direction in which these genes varied with
375 viral UMI counts, we calculated the Pearson correlation for all genes with log₂-transformed viral UMI
376 counts.

377
378 **DATA AND CODE AVAILABILITY**
379 Previously published spatial RNA-sequencing data were downloaded from Gene Expression Omnibus
380 (GEO) and are available under the following accession numbers; regenerating skeletal muscle⁵
381 GSE161318, infected heart tissue⁴ GSE189636. Spatial Total RNA-Sequencing data generated in this
382 study can be found on GEO under the accession number GSE200481. Small RNA-sequencing data are
383 available on GEO under the accession number GSE200480 A detailed protocol for performing STRS as
384 well as custom analysis scripts for aligning and processing STRS data can be found at
385 <https://github.com/mckellardw/STRS>.

386
387 **ACKNOWLEDGMENTS**
388 We thank Peter Schweitzer and colleagues in the Cornell Biotechnology Resource Center for their help

389 with sequencing the libraries. We thank the Cornell Center for Animal Resources and Education for
390 animal housing and care. We thank Ern Hwei Hannah Fong for helping with mouse procedures. We thank
391 Michael Shanahan and Zhao Lai for their help in generating the small RNA-sequencing data. We thank
392 Benjamin Grodner, Hao Shi, and other members of the Cosgrove and De Vlaminck labs for helpful
393 discussions and feedback. This work was supported by the US National Institutes of Health (NIH) grants
394 1DP2AI138242 to IDV, R21AI144557 to IDV and JSP, NIH grant R01AG058630 to BDC and IDV,
395 American Diabetes Association Pathway to Stop Diabetes Award 1-16-ACE-47 to PS, and T32EB023860
396 to DWM. The content is solely the responsibility of the authors and does not necessarily represent the
397 official views of the NIH.

398

399 AUTHOR CONTRIBUTIONS

400 DWM, IDV, and BDC designed the study. DWM, MM, and MH carried out the experiments. DWM and
401 MM analyzed the data. DWM, MM, IDV, and BDC wrote the manuscript. All authors provided feedback
402 and comments.

403

404 REFERENCES

- 405 1. Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial
406 transcriptomics. *Nature* **596**, 211–220 (2021).
- 407 2. Marx, V. Method of the Year: spatially resolved transcriptomics. *Nature Methods* **18**, 9–14
408 (2021).
- 409 3. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nature Methods* (2022)
410 doi:10.1038/s41592-022-01409-2.
- 411 4. Mantri, M. *et al.* Spatiotemporal transcriptomics reveals pathogenesis of viral myocarditis.
412 *bioRxiv* (2021) doi:10.1101/2021.12.07.471659.
- 413 5. McKellar, D. W. *et al.* Large-scale integration of single-cell transcriptomic data captures
414 transitional progenitor states in mouse skeletal muscle regeneration. *Communications Biology* **4**,
415 1280 (2021).
- 416 6. Ståhl, P. L. *et al.* Visualization and analysis of gene expression in tissue sections by spatial
417 transcriptomics. *Science* (1979) **353**, 78–82 (2016).
- 418 7. Egloff, S., Studniarek, C. & Kiss, T. 7SK small nuclear RNA, a multifunctional transcriptional
419 regulatory RNA with gene-specific features. *Transcription* **9**, 95–101 (2018).
- 420 8. Porpiglia, E. *et al.* High-resolution myogenic lineage mapping by single-cell mass cytometry.
421 *Nature Cell Biology* **19**, 558–567 (2017).
- 422 9. de Micheli, A. J. *et al.* Single-Cell Analysis of the Muscle Stem Cell Hierarchy Identifies
423 Heterotypic Communication Signals Involved in Skeletal Muscle Regeneration. *Cell Reports* **30**,
424 3583-3595.e5 (2020).
- 425 10. Wosczyzna, M. N. & Rando, T. A. A Muscle Stem Cell Support Group: Coordinated Cellular
426 Responses in Muscle Regeneration. *Developmental Cell* vol. 46 135–143 (2018).
- 427 11. Mccarthy, J. J. *The MyomiR Network in Skeletal Muscle Plasticity. Exerc. Sport Sci. Rev* vol. 39
428 www.acsm-essr.org (2011).
- 429 12. Bentzinger, C. F., Wang, Y. X., Dumont, N. A. & Rudnicki, M. A. Cellular dynamics in the muscle
430 satellite cell niche. *EMBO Rep* **14**, 1062–1072 (2013).
- 431 13. Dill, T. L., Carroll, A., Pinheiro, A., Gao, J. & Naya, F. J. The long noncoding RNA Meg3
432 regulates myoblast plasticity and muscle regeneration through epithelial-mesenchymal transition.
433 *Development (Cambridge)* **148**, (2021).
- 434 14. Zhang, Y. & Tang, L. Inhibition of breast cancer cell proliferation and tumorigenesis by long non-
435 coding RNA RPPH1 down-regulation of miR-122 expression. *Cancer Cell International* **17**,
436 (2017).

- 437 15. Zhang, P. *et al.* Long non-coding RNA Rpph1 promotes inflammation and proliferation of
438 mesangial cells in diabetic nephropathy via an interaction with Gal-3. *Cell Death and Disease* **10**,
439 (2019).
- 440 16. Alexander, M. S. & Kunkel, L. M. “Skeletal Muscle MicroRNAs: Their Diagnostic and Therapeutic
441 Potential in Human Muscle Diseases.” *Journal of Neuromuscular Diseases* **2**, 1–11 (2015).
- 442 17. Mok, G. F., Lozano-Velasco, E. & Münsterberg, A. microRNAs in skeletal muscle development.
443 *Seminars in Cell and Developmental Biology* vol. 72 67–76 (2017).
- 444 18. Giagnorio, E., Malacarne, C., Mantegazza, R., Bonanno, S. & Marcuzzo, S. MyomiRs and their
445 multifaceted regulatory roles in muscle homeostasis and amyotrophic lateral sclerosis. *Journal of*
446 *Cell Science* **134**, (2021).
- 447 19. Bartel, D. P. Metazoan MicroRNAs. *Cell* vol. 173 20–51 (2018).
- 448 20. Patil, A. H. & Halushka, M. K. miRge3.0: a comprehensive microRNA and tRF sequencing
449 analysis pipeline. *NAR Genomics and Bioinformatics* **3**, (2021).
- 450 21. Sempere, L. F. *et al.* Open Access Expression profiling of mammalian microRNAs uncovers a
451 subset of brain-expressed microRNAs with possible roles in murine and human neuronal
452 differentiation. *Sempere et al* vol. 5 <http://genomebiology.com/2004/5/3/R13> (2004).
- 453 22. Liu, N. *et al.* MicroRNA-206 promotes skeletal muscle regeneration and delays progression of
454 Duchenne muscular dystrophy in mice. *Journal of Clinical Investigation* **122**, 2054–2065 (2012).
- 455 23. Guo, Y. *et al.* The multi-functional reovirus $\sigma 3$ protein is a virulence factor that suppresses stress
456 granule formation and is associated with myocardial injury. *PLoS Pathogens* **17**, (2021).
- 457 24. Boehme, K. W., Lai, C. M. & Dermody, T. S. Mechanisms of Reovirus Bloodstream
458 Dissemination. in *Advances in Virus Research* vol. 87 (2013).
- 459 25. Sherry, B., Schoen, F. J., Wenske, E. & Fields, B. N. Derivation and characterization of an
460 efficiently myocarditic reovirus variant. *Journal of Virology* **63**, (1989).
- 461 26. Saikia, M. *et al.* Simultaneous multiplexed amplicon sequencing and transcriptome profiling in
462 single cells. *Nat Methods* **16**, 59–62 (2019).
- 463 27. Nguyen, T. A. *et al.* SIDT2 Transports Extracellular dsRNA into the Cytoplasm for Innate Immune
464 Recognition. *Immunity* **47**, 498-509.e6 (2017).
- 465 28. Nguyen, T. A. *et al.* SIDT1 Localizes to Endolysosomes and Mediates Double-Stranded RNA
466 Transport into the Cytoplasm. *The Journal of Immunology* **202**, 3483–3492 (2019).
- 467 29. Isakova, A., Neff, N. & Quake, S. R. Single-cell quantification of a broad RNA spectrum reveals
468 unique noncoding patterns associated with cell types and states. *Proceedings of the National*
469 *Academy of Sciences* **118**, e2113568118 (2021).
- 470 30. Salmen, F. *et al.* Droplet-based Single-cell Total RNA-seq Reveals Differential Non-Coding
471 Expression and Splicing Patterns during Mouse Development. *bioRxiv* 2021.09.15.460240
472 (2021) doi:10.1101/2021.09.15.460240.
- 473 31. Stickels, R. R. *et al.* Highly sensitive spatial transcriptomics at near-cellular resolution with Slide-
474 seqV2. *Nature Biotechnology* **39**, 313–319 (2021).
- 475 32. Chen, A. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball
476 patterned arrays. *bioRxiv* 2021.01.17.427004 (2021) doi:10.1101/2021.01.17.427004.
- 477 33. Uppendahl, L. D., Dahl, C. M., Miller, J. S., Felices, M. & Geller, M. A. Natural killer cell-based
478 immunotherapy in gynecologic malignancy: A review. *Frontiers in Immunology* vol. 8 (2018).
- 479 34. Polacek, C., Friebe, P. & Harris, E. Poly(A)-binding protein binds to the non-polyadenylated 3'
480 untranslated region of dengue virus and modulates translation efficiency. *Journal of General*
481 *Virology* **90**, (2009).
- 482 35. Hashizume, M., Takashima, A. & Iwasaki, M. A small stem-loop-forming region within the 3'-UTR
483 of a nonpolyadenylated LCMV mRNA promotes translation. *Journal of Biological Chemistry* **298**,
484 (2022).

- 485 36. Baran-Gale, J. *et al.* Addressing Bias in Small RNA Library Preparation for Sequencing: A New
486 Protocol Recovers MicroRNAs that Evade Capture by Current Methods. (2015)
487 doi:10.3389/fgene.2015.00352.
- 488 37. Jayaprakash, A. D., Jabado, O., Brown, B. D. & Sachidanandam, R. Identification and
489 remediation of biases in the activity of RNA ligases in small-RNA deep sequencing.
490 doi:10.1093/nar/gkr693.
- 491 38. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
492 *EMBnet J* **17**, (2011).
- 493 39. Kaminow, B., Yunusov, D. & Dobin, A. STARsolo: accurate, fast and versatile
494 mapping/quantification of single-cell and single-nucleus RNA-seq data. *bioRxiv*
495 2021.05.05.442755 (2021) doi:10.1101/2021.05.05.442755.
- 496 40. Melsted, P. *et al.* Modular, efficient and constant-memory single-cell RNA-seq preprocessing.
497 *Nature Biotechnology* 1–6 (2021) doi:10.1038/s41587-021-00870-2.
- 498 41. Mölder, F. *et al.* Sustainable data analysis with Snakemake. *F1000Res* **10**, 33 (2021).
- 499 42. Smith, T., Heger, A. & Sudbery, I. UMI-tools: Modeling sequencing errors in Unique Molecular
500 Identifiers to improve quantification accuracy. *Genome Research* **27**, (2017).
- 501 43. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).
- 502 44. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573-3587.e29 (2021).
- 503 45. Qiu, X. *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nature*
504 *Methods* **14**, 979–982 (2017).
505