

1 Establishing methods to monitor H5N1 influenza virus in dairy cattle milk

2

3 Elyse Stachler*, Andreas Gnirke, Kyle McMahon, Michael Gomez, Liam Stenson, Charelisse

4 Guevara-Reyes, Hannah Knoll, Toni Hill, Sellers Hill, Katelyn S. Messer, Jon Arizti-Sanz, Fatinah

5 Albeez, Elizabeth Curtis, Pedram Samani, Natalia Wewior, David H. O'Connor, William Vuyk,

6 Sophia Khoury, Matthew K. Schnizlein, Nicole C. Rockey, Zachariah Broemmel, Michael Mina,

7 Lawrence C. Madoff, Shirlee Wohl, Lorraine O'Connor, Catherine M. Brown, Al Ozonoff, Daniel

8 J. Park, Bronwyn L. MacInnis^{1*}, Pardis C. Sabeti^{1*}

9

10 Author affiliations: Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA (E.

11 Stachler, A. Gnirke, K. McMahon, M. Gomez, L. Stenson, C. Guevara-Reyes, H. Knoll, T. Hill, S.

12 Hill, K. S. Messer, J. Arizti-Sanz, F. Albeez, E. Curtis, P. Samani, N. Wewior, A. Ozonoff, D. J. Park,

13 B. L. MacInnis, P. C. Sabeti); University of Puerto Rico - Rio Piedras, San Juan, Puerto Rico, USA

14 (C. Guevara-Reyes); Harvard University, Cambridge, Massachusetts, USA (P. Samani, B.

15 MacInnis, P. C. Sabeti); University College London, London, United Kingdom (P. Samani);

16 University of Wisconsin-Madison, Madison, Wisconsin, USA (D. O'Connor, W. Vuyk); The

17 University of Texas at Austin, Austin, Texas, USA (S. Khoury); Michigan State University, East

18 Lansing, MI, USA (M. K. Schnizlein); Duke University, Durham, NC, USA (N. C. Rockey, Z.

19 Broemmel); Immune Observatory, Boston, Massachusetts, USA (M. Mina); Massachusetts

20 Department of Public Health, Boston, Massachusetts (L. C. Madoff, S. Wohl, C. M. Brown);

21 University of Massachusetts Chan Medical School, Worcester, Massachusetts (L. C. Madoff);

22 Brigham and Women's Hospital, Boston, Massachusetts, USA (S. Wohl); Massachusetts

23 Department of Agricultural Resources, Boston, Massachusetts (L. O'Connor); Boston Children's
24 Hospital, Boston, Massachusetts, USA (A. Ozonoff); Harvard Medical School, Boston,
25 Massachusetts, USA (A. Ozonoff); Howard Hughes Medical Institute, Chevy Chase, Maryland,
26 USA (P. C. Sabeti)

27 ¹These senior authors contributed equally to this article.

28 *Corresponding author

29

30

31 **Abstract**

32 Highly Pathogenic Avian Influenza strain H5N1 has caused a multi-state outbreak among US
33 dairy cattle, spreading across 15 states and infecting hundreds of herds since its onset. We
34 rapidly developed and optimized PCR-based detection assays and sequencing protocols to
35 support H5N1 molecular surveillance. Using 214 retail milk from 20 states for methods
36 development, we found that H5N1 concentrations by digital PCR strongly correlated with qPCR
37 cycle threshold (Ct) values, with dPCR exhibiting greater sensitivity. We also found that
38 metagenomic sequencing after hybrid selection was best for higher concentration samples
39 while amplicon sequencing performs best for lower concentrations. By establishing these
40 methods, we were able to support the creation of a statewide surveillance program to test bulk
41 milk samples monthly from all cattle dairy farms within Massachusetts, which remain negative
42 to date. The methods, workflow, and recommendations described here provide a framework
43 for others aiming to conduct H5N1 surveillance efforts.

44

45 Introduction

46 Highly Pathogenic Avian Influenza strain H5N1 has caused large scale outbreaks in wild
47 and domestic birds resulting in mass mortality, culling events, and economic losses (1). Viral
48 spillover to mammals has become more frequent recently, including outbreaks with mammal-
49 to-mammal transmission and sporadic human infections (2). In March 2024, H5N1 clade
50 2.3.4.4b was found in unpasteurized milk produced by infected dairy cattle in the US, the first
51 confirmation of an outbreak that has grown to span more than 695 herds in 15 states as of
52 October 24, 2024 (3,4). Phylogenetic analysis suggests the outbreak resulted from a single bird-
53 to-cattle spillover event (5). The outbreak has subsequently spread through interstate transport
54 of cattle, milking practices, and shared milking machinery and farm equipment (6,7). While
55 confirmed human cases have thus far been sporadic and primarily with mild symptoms, the
56 spread of H5N1 in cattle threatens the dairy industry and risks further adaptation to
57 mammalian hosts, including humans.

58 This outbreak has highlighted the need for rapidly deployable H5N1 molecular
59 surveillance capacity to detect infections, monitor viral spread and evolution, identify
60 transmission routes, target interventions to protect agricultural assets and food supply, and
61 prevent broader human transmission. Cow milk has emerged as an ideal sample source for
62 H5N1 detection and surveillance during this outbreak; the virus is shed in high concentrations in
63 milk, likely due to its affinity for infecting mammary gland epithelial cells (8). However, retail
64 milk undergoes intense processing steps, including ultra-pasteurization and homogenization,
65 with unknown effects on viral RNA quality.

66 While early reports detected the presence of H5N1 in pasteurized retail milk (9-12),
67 optimized, robust, and scalable protocols are needed to routinely and reliably detect and
68 characterize H5N1 for agricultural and public health surveillance. Using commercially available
69 milk as a sample source, we evaluated and optimized methods for nucleic acid extraction, H5N1
70 RNA detection by digital and quantitative PCR, and library construction and sequencing
71 approaches to produce near complete genomes. In partnership with the Massachusetts
72 Departments of Agricultural Resources and Public Health, we then implemented quantitative
73 detection protocols at scale to support routine statewide surveillance testing for H5N1 on all
74 cattle milk-producing farms. Here we describe these methods and applications, along with
75 considerations for implementation in other settings aiming to establish H5N1 detection and
76 surveillance capacity.

77

78 **Results**

79

80 *Validation, optimization, and performance of methods to detect H5N1 RNA in milk*

81 To validate and optimize performance for molecular detection, we tested performance
82 of an H5N1 assay using primers targeting the H5 subtype of the HA gene including sequences of
83 the current outbreak strain (13) (H5_Taq) by both quantitative (qPCR) and digital PCR (dPCR).
84 We optimized primer and probe concentrations using synthetic H5N1 RNA, selecting for
85 optimal linearity, sensitivity, accuracy, precision, and qPCR assay efficiency (Figures A1 and A2).

86 Overall, the H5N1 assay displayed robust performance on both platforms, with dPCR
87 outperforming qPCR in limit of detection (LOD) and precision. The LOD₉₀ was 5 copies/μL by

88 dPCR and 10 copies/ μ L by qPCR. Based on known concentrations of standard material, we
89 found dPCR concentrations correlated well with qPCR Ct values (Figure 1D). In addition, dPCR
90 exhibited lower coefficients of variations, ranging from 10.5-26.4% compared to 18.0-111.5%
91 for qPCR evaluated for 2.5-25 copies/ μ L (Figure 1C). Both assays maintained linearity over their
92 dynamic ranges (Figure 1A and 1B). Across all qPCR standard curves performed ($n = 8$),
93 efficiencies ranged from 87-111% (Figure A3), with the 10^2 copies standard concentration
94 detected 93% of the time (Figure A5). This detection rate was lower than expected based on
95 the LOD analysis, likely due to RNA standard degradation over time and highlighting the
96 importance of standard material integrity for qPCR.

97 As a positive internal control for nucleic acid extraction in cattle milk, we designed a PCR
98 assay targeting the bovine Ribonuclease P gene (both DNA and RNA; RP_Bov). By dPCR,
99 linearity was maintained across all dilutions tested (Figure 1A) and the LOD_{90} was 10 copies/ μ L.
100 Based on the superior performance of dPCR for the H5N1 target, the RP_Bov assay was not
101 evaluated as a qPCR assay. Overall, all PCR assays performed well with minimal optimization.

102 We next evaluated pre-processing and extraction protocols to optimize sample
103 preparation for subsequent H5N1 detection and sequencing. We tested two commercially
104 available extraction kits, MagMAX Prime Viral/Pathogen (Prime) and MagMAX CORE (CORE), by
105 spiking serial dilutions of synthetic H5N1 nucleic acid into milk. We tested milk with various fat
106 contents and examined the effect of pre-centrifugation (at either 1200xg or 12,000xg) on
107 outcomes. We also tested the MagMAX Wastewater kit (Wastewater) head-to-head with the
108 CORE kit on a subset of eight retail milk samples previously found to be positive with CORE kit
109 testing.

110 All three extraction kits demonstrated similar recovery of H5N1 from milk, with the
111 CORE kit exhibiting slightly better performance. The CORE (Figure 2) and Prime (Figure A6) kits
112 showed comparable results in terms of total recovery (down to $\sim 10^4$ H5N1 copies/mL milk) and
113 linearity ($R^2_{\text{Prime,DNA}}=0.93$, $R^2_{\text{CORE,DNA}}=0.96$, and $R^2_{\text{CORE,RNA}}=0.98$). Milk fat content and pre-
114 centrifugation exhibited no significant effect on target detection. As well, there was no
115 significant difference in detection of H5N1 ($p=0.20$) or RP_Bov ($p=0.17$) using the Wastewater
116 extraction kit on retail milk samples (Figure A7). We selected the CORE kit for ongoing testing
117 given its low detection limit and slightly better detection of RP_Bov, as well as practical
118 considerations including a manufacturer's protocol for processing milk and kit availability.

119

120 *Detection of H5N1 in retail milk samples*

121 To validate protocols on *in situ* H5N1 in milk, we sourced 214 retail milk cartons with
122 diverse characteristics, including fat content and pasteurization processes, from 61 processing
123 plants in 20 states (see Table, Figure 4A). Of these, 55 (26%) tested positive for presence of
124 H5N1 RNA by dPCR, while 48 (22%) tested positive by qPCR. Positive samples were from
125 processing plants in four states with reported H5N1 outbreaks (Colorado, Idaho, Michigan, and
126 Texas). We also detected one positive sample by both dPCR and qPCR originating from a
127 processing plant in Missouri, which has not reported H5N1 in cattle. Notably, the location of the
128 processing plant reported on milk containers may or may not correspond to the state(s) in
129 which the milk was initially collected, and this linkage is not publicly available.

130 H5N1 detection and concentrations were strongly correlated between dPCR and qPCR
131 platforms, with dPCR showing greater sensitivity. The platforms gave concordant

132 positive/negative results for 95% (n=203/214) of samples (Figure 3A, see methods for
133 thresholding details). Nine samples were positive only by dPCR, which may be due to the
134 slightly enhanced LOD of the dPCR assay. Conversely, two samples were positive only by qPCR,
135 possibly due to the more stringent thresholding criteria for dPCR. Further, H5N1 RNA dPCR
136 concentrations correlated strongly with qPCR Ct values ($R^2=0.81$, Figure 3B), suggesting the
137 assay is robust on either platform.

138 We used the RP_Bov assay as an internal sample process control to confirm sample
139 integrity and ensure proper collection and extraction, especially useful to interpret negative
140 H5N1 results. RP_Bov concentrations averaged 560 copies/ μ L extract (Figure 4B), with 98% of
141 samples falling within one standard deviation (we noted that four samples stored at 4°C for
142 longer prior to processing than other samples fell below this lower limit). Thus, detection of
143 RP_Bov below \sim 100 copies/ μ L could be effectively used as a measure of milk sample and
144 process integrity.

145

146 *Sequencing of H5N1 from retail milk*

147 We attempted to recover genomes from 23 H5N1 positive retail milk samples, across a
148 range of characteristics including virus concentration, milk type, and pasteurization process. To
149 obtain higher H5N1 concentrations for library preparation, we first extracted, pooled, and
150 concentrated ten samples from each milk container. Ultra-pasteurized samples exhibited
151 significantly lower concentration factors compared to pasteurized samples as measured by
152 H5N1 copy number (ultra-pasteurized_{AVG}=4.6, pasteurized_{AVG}=8.9, $p=0.015$, Figure A9). Despite
153 being highly concentrated, samples showed no evidence of PCR inhibition by dPCR ($p=0.89$,

154 Figure A10). The recovered RNA content and quality from these samples spanned a wide range
155 as determined by H5N1 copies ($14\text{-}9\times 10^4$ copies/ μL), total RNA concentration (0.1-67.4 ng/ μL),
156 H5N1 copies/ng RNA ($2\text{-}1.5\times 10^4$ copies/ng), and RIN (2-7.2) score (Table A7).

157 We evaluated three library construction methods to assess their efficacy in producing
158 genomes across the range of H5N1 concentrations and pasteurization processes: untargeted
159 metagenomic RNA sequencing (RNA-Seq), hybrid-selected RNA-Seq (hsRNA-Seq) enriched for
160 human respiratory viruses including influenza A (albeit not explicitly H5N1) (14), and amplicon
161 sequencing (Amp-Seq) of tiled 250-bp H5N1 PCR products (15). We produced near-complete
162 (>70% assembly) H5N1 genomes from all 23 samples: 12 by hsRNA-Seq ($\geq 80\%$) and 11 by Amp-
163 Seq ($\geq 74\%$). Hybrid selection greatly increased the chances of genome recovery for higher
164 concentration extracts (>500 copies/ μL), with hsRNA-Seq outperforming RNA-Seq for 11 of 12
165 samples. At lower concentrations, Amp-Seq resulted in the most complete genomes (Figure 5).
166 Notably, we modified the PCR cycling conditions of a previously reported H5N1 Amp-Seq
167 protocol (15), which resulted in improved amplicon generation and genome assemblies (Figure
168 A11; see Methods). However, there was considerable variability in PCR efficiency across
169 amplicons, with a small fraction of amplicons producing the vast majority of sequencing reads
170 (Figure A12), thus still requiring a large amount of sequences per sample for complete genome
171 assembly despite enrichment for specific PCR amplicons.

172 Phylogenetic analysis showed geographic clustering with other publicly available H5N1
173 genomes associated with the dairy cattle outbreak (Figure A13), suggesting the origin of the
174 viruses was consistent with the US state of the processing plant of the milk. Notably, the
175 positive originating from Missouri (which has no reports of H5N1 in cattle) clustered with

176 samples from Texas and Michigan, likely pointing to the farm location the milk originated from
177 despite being processed in a Missouri plant.

178

179 *Implementation at scale: statewide H5N1 surveillance of pre-pasteurized milk from*
180 *Massachusetts farms*

181 By establishing a robust workflow for detection and sequencing of H5N1 from milk, we
182 were positioned to partner with the Massachusetts Departments of Agricultural Resources and
183 of Public Health to support a mandatory surveillance program for H5N1 testing in milk from
184 cattle dairy farms across the state. This program, launched in August 2024, was implemented
185 preemptively in the absence of H5N1 detection in the state or surrounding region to confirm
186 the absence of H5N1, and to serve as an early warning system should a local outbreak occur.
187 State authorities worked with farms to collect samples from bulk milk tanks from all 95 cattle
188 dairy farms across Massachusetts, initially within a 3 week period, followed by a rotating
189 sampling schedule such that all farms were tested monthly. Based on our process validation
190 using retail milk samples, we extracted bulk milk samples using the CORE extraction kit and
191 performed dPCR for detection of H5_Taq and RP_Bov. While the surveillance program is
192 currently ongoing, to date we have completed four rounds of statewide testing; H5N1 has not
193 been detected in any sample from this surveillance program. The RP_Bov positive control has
194 been routinely detected at similar levels to commercially available milk, providing confidence in
195 the negative results obtained for H5N1 (Figure A14).

196

197

198 Discussion

199 This study contributes validated methods for the whole workflow from sample to
200 analyzed data for rapid deployment for potential future epidemiological studies and public
201 health surveillance. We used retail milk products to optimize and validate methods for sample
202 processing, detection, and sequencing of H5N1 in pasteurized milk, allowing us to implement
203 these workflows at scale to support ongoing surveillance for H5N1 in pre-pasteurized milk from
204 bulk tanks collected at the farm level. The workflow described here exhibits strong
205 performance characteristics, while requiring minimal sample processing. We found direct
206 nucleic acid extraction from milk regardless of fat content was efficient; with pre-centrifugation
207 offering no increase in viral RNA recovery, in accordance with previous findings (16,17).

208 The H5N1 PCR assay performed well by both qPCR and dPCR and can be easily deployed
209 on either platform. Both qPCR and dPCR have been widely adopted for viral surveillance, each
210 with unique strengths: qPCR has a broader dynamic range of detection, higher throughput
211 capabilities, lower cost per sample, and common instrument availability (18,19), while dPCR
212 has been shown to have superior performance in clinical studies (20, 21), enhanced sensitivity
213 (18, 19, 21-23), absolute quantification without reliance on a positive control standard (18, 20),
214 robustness to PCR inhibitors (19, 24), and lower inter-laboratory variation (25). In cases where
215 qPCR is used, dPCR can be useful for quantification of standards to ensure quality of materials
216 (20, 26) to mitigate the effect of qPCR standard degradation and inaccuracy. The choice of PCR
217 platform should weigh the features of each method, use case of the data, and the resources
218 available to the laboratory.

219 Despite intense milk pre-processing (such as ultra-pasteurization), near-complete H5N1
220 genomes were readily recovered. Phylogenetic value of the genomes are limited in this context,
221 as the original source farm(s) or state(s) are not provided on retail milk containers and the
222 provided location of the processing plant may or may not correlate to the originating farm
223 location. Using genome data, we verified that the positive sample obtained from a Missouri
224 processing plant likely originated from a Texas farm. Notably, all other samples clustered
225 geographically with publicly available H5N1 genomes originating from the state in which the
226 milk was processed.

227 Our protocol development work enabled a partnership with state officials to perform
228 preemptive mandatory H5N1 testing of raw milk from dairy farms across Massachusetts. To
229 date the virus has not been detected in New England, likely due to being geographically far
230 from where outbreaks have occurred thus far, the smaller size of the region's dairy industry,
231 and minimal interstate cattle transport compared to larger dairy industries (27). The prevalence
232 of viral RNA in milk offers a unique surveillance mechanism to easily monitor lactating herds by
233 testing pooled bulk milk tank samples, saving time and resources compared to individual cow
234 testing. The sensitivity of our workflow allows for preemptive surveillance of H5N1 for the
235 typical size of a Massachusetts cattle dairy farm (approximately 10,000 cows on 125 farms)
236 (28). Based on our LOD of 10^4 copies/mL milk, in order to detect 1 infected cow in a herd size of
237 either 100 or 1000 cows, the infected cow would have to be shedding 10^6 or 10^7 H5N1
238 copies/mL milk, respectively. This is within the concentration range of live virus shed by
239 infected cattle (10^4 - $10^{8.8}$ TCID₅₀/mL) (8).

240 Based on the methods testing and validation above, we offer guidance on establishing
241 efficient, robust, and scalable H5N1 surveillance from bulk milk for implementation in
242 molecular laboratory settings. For a detailed protocol and checklist, please refer to the guide
243 provided in the Appendix. In brief, to obtain positive control material for protocol validation,
244 look to source samples from outbreak-associated states by targeting milk cartons with USDA
245 codes originating from processing plants in those states. The latest outbreak information (4)
246 and USDA processing plant codes (29) can be easily accessed online. PCR assay performance
247 should include linearity >90%, qPCR efficiency between 90-110%, and an LOD ≤ 10 copies/ μ L
248 extract. Performance of the extraction method can be validated by spiking in synthetic nucleic
249 acid into a negative milk sample, confirming that linearity > 90% and the process LOD $\leq 10^4$
250 copies/mL milk. For sequencing of H5N1 in milk, we recommend metagenomic sequencing after
251 targeted hybrid selection for highly concentrated samples (>500 copies/ μ L) and sequencing of
252 PCR amplicons for less concentrated or unconcentrated samples. Where permitted, near-
253 complete genomes should be submitted to public databases. Finally, for unpasteurized milk
254 testing, milk can be pasteurized in the laboratory per the USDA protocol (30) to limit risk of
255 exposure to live virus while maintaining the molecular integrity of the sample for H5N1
256 detection and sequencing.

257 Enabling more labs to set up decentralized surveillance will allow us to stay ahead of
258 current and future outbreaks of public concern. We hope the guidelines provided here can
259 serve as a blueprint for rapid validation of new molecular detection methods and establishment
260 of surveillance systems for the current H5N1 outbreak and beyond.

261

262 **Methods**

263

264 *PCR assay design and characterization*

265 For detection of H5N1 RNA, a previously published assay targeting the HA fragment of
266 the H5N1 virus (denoted as H5_Taq) was used (13). Also, an internal extraction control assay
267 was designed to target the bovine Ribonuclease P gene (RP_Bov). Digital PCR (dPCR) reactions
268 were run on a Qiacuity One (5plex, Qiagen), while quantitative PCR (qPCR) assays were
269 conducted on a QuantStudio 6 Flex (ThermoFisher). Additional details including primer
270 sequences, kit information, primer and probe concentrations, and cycling conditions are
271 reported in the Appendix. For dPCR, a sample had to have \geq three positive partitions to be
272 defined positive (13, 31). For qPCR, a sample was considered positive if 2 of 3 replicates
273 amplified for at least one of the extraction replicates. The amplified Ct values were then
274 averaged for subsequent analysis. PCR assay performance was evaluated using synthetic RNA of
275 the H5N1 HA segment. The limit of detection (LOD₉₀) was defined as the lowest concentration
276 of target copies detected in at least 18/20 replicates. Linearity was evaluated from serial
277 dilutions over the dynamic range for both PCR platforms. Finally, the standard curves of all
278 qPCR runs were evaluated for overall assay efficiency.

279

280 *Extraction kit evaluation*

281 Three commercially available extraction kits were evaluated for their potential to
282 recover nucleic acid from a milk matrix. All kits chosen were bead-based, high-throughput kits
283 compatible with the KingFisher Flex instrument (ThermoFisher). The MagMAX Prime kit and

284 MagMAX CORE kit were tested by spiking serial dilutions of an 1800bp synthetic nucleic acid
285 fragment of the H5N1 HA sequence. Parameters evaluated were fat content (whole vs. 2%),
286 dilution with phosphate-buffered saline (PBS), and pre-centrifugation (either 12000xg for 10
287 minutes or 1200xg for 30 minutes). Finally, the MagMAX Wastewater kit was evaluated head-
288 to-head with the MagMAX CORE kit on a subset of retail milk samples previously identified as
289 positive through evaluation with the CORE kit. For this comparison, all samples were re-
290 extracted with the CORE kit as well as processed with the Wastewater kit on the same day. All
291 kits were used following the manufacturer's instructions, using 200 μ L of milk as input.

292

293 *Sourcing and nucleic acid isolation of retail milk*

294 Milk was purchased around greater Boston and obtained from states reported to be
295 impacted by the H5N1 outbreak. To source milk from across the country, milk was purchased
296 using mobile delivery apps by targeting local brands of milk and delivered to a collaborator. The
297 collaborator aliquoted the milk into two falcon tubes, sealed in plastic bags, packed on ice, and
298 shipped back to the Broad Institute overnight for processing. Upon receipt, one set of aliquots
299 was stored at -80°C for preservation. The other aliquot proceeded to nucleic acid extraction.

300 All milk samples were extracted in duplicate using the MagMAX CORE kit on a KingFisher
301 Flex following the manufacturer's "Simple Workflow" for 200 μ L of milk input. A subset of
302 positive samples were chosen to be evaluated by sequencing. To obtain enough RNA for
303 sequencing, these samples were re-extracted by the CORE kit 10 times and subsequently
304 concentrated using the RNA-Clean and Concentrator Kit (Zymo) using manufacturer protocols,
305 including on column DNase treatment and adjusting the amount of binding buffer and ethanol

306 to match the total elution volume and passing the entire volume in multiple loading steps
307 through the column. Effects of inhibition of concentrated samples were evaluated by the dPCR
308 H5_Taq assay spiking in two different volumes of template (1 μ L vs 2 μ L of input) and calculating
309 total concentration per μ L of extract.

310

311 *Surveillance of raw milk samples from Massachusetts farms*

312 Bulk tank samples from all Massachusetts cattle dairy farms (n = 95) were collected by
313 MDAR employees on a monthly basis starting on August 6, 2024 and delivered to the Broad
314 Institute for processing. Samples were collected in 2 ounce plastic containers, typical for bulk
315 tank sampling, and transported on ice to the lab for testing. Samples were immediately
316 pasteurized onsite in a heated water bath by ensuring the internal temperature of the
317 collection bottle reached 72°C for at least 15 seconds per USDA protocols (30). The samples
318 were immediately placed on ice to cool and then proceeded through the same workflow as
319 described above for nucleic acid extraction and subsequent dPCR analysis.

320

321 *Preparation of sequencing libraries*

322 Three methods were evaluated for library construction: untargeted metagenomic RNA
323 sequencing (RNA-Seq) using the xGen RNA library prep kit (IDT), hybrid-selected RNA-Seq
324 (hsRNA-Seq) of xGen RNA-Seq libraries enriched using the Respiratory Virus Research Capture
325 panel as bait (Twist Biosciences) with the Target Enrichment Standard Hybridization v2 kit
326 following the manufacturer's protocols, and amplicon sequencing (Amp-Seq) of tiled 250-bp
327 H5N1 PCR products as described by Vuyk et al (15). For Amp-Seq, we used a modified

328 thermoprofile from the original published protocol, namely shortening the annealing step and
329 adding an extension step (Table A6). Full methods are described in detail in the Appendix. RNA
330 samples for sequencing were run on an RNA 6000 Pico Bioanalyzer chip (Agilent) to determine
331 total RNA concentration, size distribution, and RNA Integrity (RIN) scores. Water and a water
332 extraction control served as blank negative controls for library construction. Pooled sequencing
333 libraries were sequenced with paired-end 151-base reads on 300-cycle NextSeq 2000 cartridges
334 (Illumina).

335

336 *Genomic analysis*

337 Detailed methods for genomic analysis are available in the Appendix. Briefly, NextSeq
338 sequencing runs were basecalled and demultiplexed using Picard using custom specified read
339 structures to accommodate for the xGen library protocol, resulting in hard-trimmed reads
340 containing only target sequences and obviating the need for post-alignment based trimming
341 during consensus sequence generation. For each sequencing library, consensus influenza
342 genomes were produced using a standard consensus generation pipeline utilized previously
343 (32-35). The H5N1 Bovine/texas/24-029328-01/2024 reference genome (PP599462.1 through
344 PP599469.1) was used as the reference for all assemblies. For each sample sequenced by
345 multiple methods, we utilized the hsRNA-Seq genome if it recovered >75% of the genome,
346 otherwise using the Amp-Seq genome. Phylogenetic analysis was performed by releasing
347 successful genomes on NCBI Genbank, allowing automatic incorporation into the Moncla
348 Lab/Nextstrain avian-flu builds for the cattle-associated outbreak. Sequence data is available at
349 NCBI/INSDC under BioProject PRJNA1134696.

350

351 *Statistical analyses*

352 All statistical analysis was completed in GraphPad Prism with statistical significance
353 defined as $p < 0.05$. Correlation between dPCR concentration and qPCR Ct value as well as tests
354 of linearity were fit by simple linear regression. One-way ANOVA was used to compare effects
355 of different conditions when multiple conditions were compared whereas a paired t-test was
356 used when two conditions were being compared.

357

358 **Funding Statement**

359 This work was supported by funding from the Howard Hughes Medical Institute (HHMI)
360 Investigator Program (to P.C.S.), the US Centers for Disease Control and Prevention (BAA
361 75D30122C15113 to P.C.S and PGCoe NU50CK000629 to S.W., L.M., P.C.S, and B.L.M.), and the
362 NIH National Institute of Allergy and Infectious Diseases (GCID U19AI110818 to P.C.S. and D.J.P.
363 and CREID U01AI151812 to P.C.S.). J.A.S. is supported in this project by philanthropic funding
364 from the TED Audacious Project. D.H.O.'s and W.V.'s work on this project was funded by Heart
365 of Racing and the UW Institute for Clinical and Translational Research's Pilot Award program.
366 This publication was supported by the Office of Advanced Molecular Detection, Centers for
367 Disease Control and Prevention through Cooperative Agreement Number CK22-2204. The
368 content is solely the responsibility of the authors and does not necessarily represent the official
369 views or policies of the Centers for Disease Control and Prevention or the U.S. government. This
370 study has been approved for public release; distribution is unlimited.

371

372

373 **Acknowledgements**

374 The authors thank the Boston Globe Media Partners, LLC staff for sourcing and donating
375 local New England milk samples and John Rinn, Noreen Beckie, Kristen Koneschik, and Michael
376 Butts for collecting US-based milk samples, and to other teammates/collaborators for helpful
377 discussions.

378

379 **Conflicts of Interest**

380 P.C.S. is co-founder and shareholder in Sherlock Biosciences and Delve Bio, and is a
381 board member and shareholder of Danaher Corporation. DHO is a co-founder and managing
382 member of Pathogenuity LLC.

383

384 **References**

- 385 1. Peacock T, Moncla L, Dudas G, VanInsberghe D, Sukhova K, Lloyd-Smith JO, et al. The
386 global H5N1 influenza panzootic in mammals. *Nature*. Forthcoming 2024.
387 <https://doi.org/10.1038/s41586-024-08054-z>.
- 388 2. Graziosi G, Lupini C, Catelli E, Carnaccini S. Highly Pathogenic Avian Influenza (HPAI) H5
389 Clade 2.3.4.4b Virus Infection in Birds and Mammals. *Animals (Basel)*. 2024 May
390 2;14(9):1372. doi: 10.3390/ani14091372. PMID: 38731377; PMCID: PMC11083745.
- 391 3. World Health Organization. Avian Influenza A(H5N1) – United States of America. 2024
392 April 9 [cited 2024 Oct 1]. <https://www.who.int/emergencies/disease-outbreak->

- 393 news/item/2024-DON512.
- 394 4. US Department of Agriculture Animal and Plant Health Inspection Service. HPAI
395 Confirmed Cases in Livestock. 2024 Jul 3 [cited 2024 Dec 3].
396 [https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-](https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/hpai-confirmed-cases-livestock)
397 [detections/hpai-confirmed-cases-livestock](https://www.aphis.usda.gov/livestock-poultry-disease/avian/avian-influenza/hpai-detections/hpai-confirmed-cases-livestock).
- 398 5. Nguyen T-Q, Hutter C, Markin A, Thomas M, Lantz K, Killian ML, et al. Emergence and
399 interstate spread of highly pathogenic avian influenza A(H5N1) in dairy cattle. unpub.
400 data, <https://doi.org/10.1101/2024.05.01.591751>.
- 401 6. Halwe NJ, Cool K, Breithaupt A, Schön J, Trujillo JD, Nooruzzaman M, et al. H5N1 clade
402 2.3.4.4b dynamics in experimentally infected calves and cows. *Nature*. Forthcoming
403 2024. <https://doi.org/10.1038/s41586-024-08063-y>.
- 404 7. Le Sage V, Campbell AJ, Reed DS, Duprex WP, Lakdawala SS. Persistence of Influenza
405 H5N1 and H1N1 Viruses in Unpasteurized Milk on Milking Unit Surfaces. *Emerg Infect*
406 *Dis*. 2024 Aug;30(8):1721-1723. doi: 10.3201/eid3008.240775. Epub 2024 Jun 24. PMID:
407 38914418; PMCID: PMC11286056.
- 408 8. Caserta LC, Frye EA, Butt SL, Laverack M, Nooruzzaman M, Covalada LM, et al. Spillover
409 of highly pathogenic avian influenza H5N1 virus to dairy cattle. *Nature*. 2024 Jul 25; 634,
410 669–676. <https://doi.org/10.1038/s41586-024-07849-4>.
- 411 9. Lauring AS, Martin ET, Eisenberg M, Fitzsimmons WJ, Salzman E, Leis AM, et al.
412 Surveillance of H5 HPAI in Michigan using retail milk. unpub. data,
413 <https://doi.org/10.1101/2024.07.04.602115>.
- 414 10. Spackman E, Jones DR, McCoig AM, Colonius TJ, Goraichuk IV, Suarez DL.

- 415 Characterization of highly pathogenic avian influenza virus in retail dairy products in the
416 US. *J Virol.* 2024 Jul 23;98(7):e0088124. doi: 10.1128/jvi.00881-24. Epub 2024 Jul 3.
417 PMID: 38958444; PMCID: PMC11264905.
- 418 11. Suarez DL, Goraichuk PV, Killmaster L, Spackman E, Clausen NJ, Colonius TJ, et al.
419 Testing of retail cheese, butter, ice cream and other dairy products for highly pathogenic
420 avian influenza in the US. unpub. data, doi:10.1101/2024.08.11.24311811.
- 421 12. Tarbuck N, Jones J, Franks J, Kandeil A, DeBeauchamp J, Miller L, et al. Detection of
422 A(H5N1) influenza virus nucleic acid in retail pasteurized milk. unpub. data,
423 doi:10.21203/rs.3.rs-4572362/v1.
- 424 13. Wolfe MK, Duong D, Shelden B, Chan EMG, Chan-Herur V, Hilton S, et al. Detection of
425 Hemagglutinin H5 Influenza A Virus Sequence in Municipal Wastewater Solids at
426 Wastewater Treatment Plants with Increases in Influenza A in Spring, 2024. *ES&T*
427 *Letters.* 2024 May 20;11(6), 526-532. doi: 10.1021/acs.estlett.4c00331.
- 428 14. Metsky HC, Siddle KJ, Gladden-Young A, Qu J, Yang DK, Brehio P, et al. Capturing
429 sequence diversity in metagenomes with comprehensive and scalable probe design.
430 *Nat. Biotechnol.* 2019 February 4; 37, 160–168. [https://doi.org/10.1038/s41587-018-](https://doi.org/10.1038/s41587-018-0006-x)
431 [0006-x](https://doi.org/10.1038/s41587-018-0006-x).
- 432 15. Vuyk WC, Lail A, Emmen I, Hassa N, Tiburcio PB, Newman C, et al. Whole genome
433 sequencing of H5N1 from dairy products with tiled 250bp amplicons v1. 2024 Jun 5.
434 [protocols.io. dx.doi.org/10.17504/protocols.io.kqdg322kpv25/v1](https://doi.org/10.17504/protocols.io.kqdg322kpv25/v1).
- 435 16. Minsky BB, Kadam A, Tanner NA, Cantor E, Patton GC, New England Biolabs. Facilitating
436 Purification and Detection of Viral Nucleic Acids from Milk. New England Biolabs. 2024

- 437 Jun. <https://www.neb.com/en-us/-/media/nebus/files/application->
438 [notes/appnote_facilitating_purification_and_detection_of_viral_nucleicacids_from_mil](https://www.neb.com/en-us/-/media/nebus/files/application-)
439 [k.pdf?rev=d5777c15b56b470fb97e69ae6d855cfd&sc_lang=en-us.](https://www.neb.com/en-us/-/media/nebus/files/application-)
- 440 17. Lail A, Vuyk WC, Emmen I, O'Connor D. RNA extraction from milk for HPAI surveillance
441 v1. 2024 May 10. [protocols.io](https://doi.org/10.17504/protocols.io.n2bvjn6obgk5/v1). doi:10.17504/protocols.io.n2bvjn6obgk5/v1.
- 442 18. Tiwari A, Ahmed W, Oikarinen S, Sherchan SP, Heikinheimo A, Jiang G, et al. Application
443 of digital PCR for public health-related water quality monitoring. *Sci Total Environ*. 2022
444 Sep 1;837:155663. doi: 10.1016/j.scitotenv.2022.155663. Epub 2022 May 4. PMID:
445 35523326.
- 446 19. Park S, Rana A, Sung W, Munir M. Competitiveness of Quantitative Polymerase Chain
447 Reaction (qPCR) and Droplet Digital Polymerase Chain Reaction (ddPCR) Technologies,
448 with a Particular Focus on Detection of Antibiotic Resistance Genes (ARGs). *Applied*
449 *Microbiology*. 2021 Oct 2; 1(3):426-444.
450 <https://doi.org/10.3390/applmicrobiol1030028>.
- 451 20. Kuypers J, Jerome KR. Applications of Digital PCR for Clinical Microbiology. *J Clin*
452 *Microbiol*. 2017 May 23;55(6):1621-1628. doi: 10.1128/JCM.00211-17. Epub 2017 Mar
453 15. PMID: 28298452; PMCID: PMC5442518.
- 454 21. Liu C, Shi Q, Peng M, Lu R, Li H, et al. Evaluation of droplet digital PCR for quantification
455 of SARS-CoV-2 Virus in discharged COVID-19 patients. *Aging (Albany NY)*. 2020 Nov
456 1;12(21):20997-21003. doi: 10.18632/aging.104020. Epub 2020 Nov 1. PMID: 33136068;
457 PMCID: PMC7695381.
- 458 22. Ahmed W, Simpson SL, Bertsch PM, Bibby K, Bivins A, Blackall LL, et al. Minimizing errors

- 459 in RT-PCR detection and quantification of SARS-CoV-2 RNA for wastewater surveillance.
460 Sci Total Environ. 2022 Jan 20;805:149877. doi: 10.1016/j.scitotenv.2021.149877. Epub
461 2021 Aug 25. PMID: 34818780; PMCID: PMC8386095.
- 462 23. Ahmed W, Smith WJM, Metcalfe S, Jackson G, Choi PM, Morrison M, et al. Comparison
463 of RT-qPCR and RT-dPCR Platforms for the Trace Detection of SARS-CoV-2 RNA in
464 Wastewater. ACS ES T Water. 2022 Nov 11;2(11):1871-1880. doi:
465 10.1021/acsestwater.1c00387. Epub 2022 Jan 28. PMID: 36380768; PMCID:
466 PMC8848507.
- 467 24. Hinkle A, Greenwald HD, Metzger M, Thornton M, Kennedy LC, Loomis K, et al.
468 Comparison of RT-qPCR and digital PCR methods for wastewater-based testing of SARS-
469 CoV-2. unpub. data, doi:10.1101/2022.06.15.22276459.
- 470 25. Boxman ILA, Molin R, Persson S, Juréus A, Jansen CCC, Sosef NP, et al. An international
471 inter-laboratory study to compare digital PCR with ISO standardized qPCR assays for the
472 detection of norovirus GI and GII in oyster tissue. Food Microbiol. 2024 Jun;120:104478.
473 doi: 10.1016/j.fm.2024.104478. Epub 2024 Jan 12. PMID: 38431324.
- 474 26. Park C, Lee J, Hassan ZU, Ku KB, Kim SJ, Kim HG, et al. Comparison of Digital PCR and
475 Quantitative PCR with Various SARS-CoV-2 Primer-Probe Sets. J Microbiol Biotechnol.
476 2021 Mar 28;31(3):358-367. doi: 10.4014/jmb.2009.09006. PMID: 33397829; PMCID:
477 PMC9705847.
- 478 27. Piore A. Traces of bird flu have made it into store-bought milk in New England, but at
479 very low levels. The Boston Globe (MA). 2024 May 17.
480 <https://www.bostonglobe.com/2024/05/17/metro/inactivated-bird-flu-virus-found-in->

- 481 milk-sold-in-massachusetts/.
- 482 28. USDA National Agricultural Statistics Service. 2023 Agricultural Statistics Annual Bulletin
483 New England. 2024 Sep.
484 [https://www.nass.usda.gov/Statistics_by_State/New_England_includes/Publications/An
485 nual_Statistical_Bulletin/2023/2023_NewEngland_Annual_Bulletin.pdf](https://www.nass.usda.gov/Statistics_by_State/New_England_includes/Publications/Annual_Statistical_Bulletin/2023/2023_NewEngland_Annual_Bulletin.pdf)
- 486 29. Where Is My Milk From?. 2023 May 1 [cited 2024 Jul 31].
487 <https://www.whereismilkfrom.com>.
- 488 30. Food and Drug Administration, Department of Health and Human Services. Code of
489 Federal Regulations: Part 131–Milk and Cream. 1977 Mar 15.
490 <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-B/part-131>.
- 491 31. Huisman JS, Scire J, Caduff L, Fernandez-Cassi X, Ganesanandamoorthy P, Kull A, et al.
492 Wastewater-Based Estimation of the Effective Reproductive Number of SARS-CoV-2.
493 Environ Health Perspect. 2022 May 26;130(5):57011. doi: 10.1289/EHP10050. Epub
494 2022 May 26. PMID: 35617001; PMCID: PMC9135136.
- 495 32. Holmes EC, Dudas G, Rambaut A, Andersen KG. The evolution of Ebola virus: Insights
496 from the 2013-2016 epidemic. Nature. 2016 Oct 13;538(7624):193-200. doi:
497 10.1038/nature19790. PMID: 27734858; PMCID: PMC5580494.
- 498 33. Lemieux JE, Siddle KJ, Shaw BM, Loreth C, Schaffner SF, Gladden-Young A, et al.
499 Phylogenetic analysis of SARS-CoV-2 in Boston highlights the impact of superspreading
500 events. Science. 2021 Feb 5;371(6529):eabe3261. doi: 10.1126/science.abe3261. Epub
501 2020 Dec 10. PMID: 33303686; PMCID: PMC7857412.
- 502 34. Wohl S, Metsky HC, Schaffner SF, Piantadosi A, Burns M, Lewnard JA, et al. Combining

503 genomics and epidemiology to track mumps virus transmission in the United States.
504 PLoS Biol. 2020 Feb 11;18(2):e3000611. doi: 10.1371/journal.pbio.3000611. PMID:
505 32045407; PMCID: PMC7012397.

506 35. Park D, Tomkins-Tinch C, Ye S, Jungreis I, Arenas FN, Shlyakhter I, et al.
507 broadinstitute/viral-pipelines: v2.3.6.0. Zenodo. 2024 Nov. doi:
508 10.5281/zenodo.4306357

509 36. Metsky HC, Matranga CB, Wohl S, Schaffner SF, Freije CA, Winnicki SM, et al. Zika virus
510 evolution and spread in the Americas. Nature. 2017 Jun 15;546(7658):411-415. doi:
511 10.1038/nature22402. Epub 2017 May 24. PMID: 28538734; PMCID: PMC5563848.

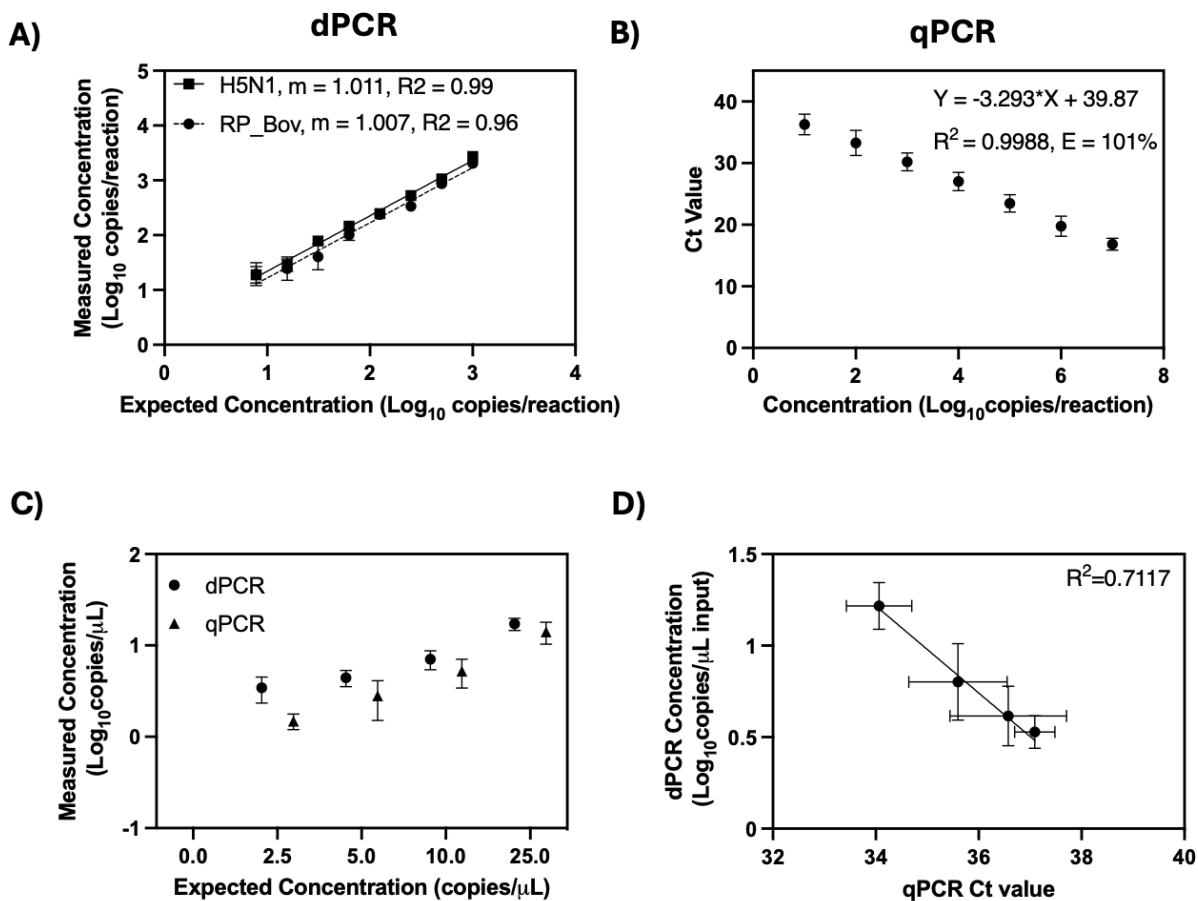
512 37. Bassano I, Ramachandran VK, Khalifa MS, Lilley CJ, Brown MR, van Aerle R, et al.
513 Evaluation of variant calling algorithms for wastewater-based epidemiology using mixed
514 populations of SARS-CoV-2 variants in synthetic and wastewater samples. Microb
515 Genom. 2023 Apr;9(4):mgen000933. doi: 10.1099/mgen.0.000933. PMID: 37074153;
516 PMCID: PMC10210938.

517
518
519
520
521
522
523

524 **Table: Breakdown of milk samples tested and their results by processing plant state. States**
525 **designated with an asterisk (*) indicate states that had reported cases of H5N1 in cattle at**
526 **the time of testing.**

Processing plant state	Number of cartons tested	Number of positives	Positivity rate (%)
AZ	1		
CA	10		
CO*	59	33	56
CT	4		
IA*	9		
ID*	12	5	42
KS*	2		
KY	1		
MA	18		
ME	2		
MI*	14	5	36
MN*	9		
MO	3	1	33
NC*	7		
NH	6		
NY	2		
OH*	3		
TX*	42	13	31
UT	7		
VA	3		
Total	214	57	27

527



528

529 **Figure 1: Validation and characterization of dPCR and qPCR H5N1 assays on synthetic spike-in**

530 **samples.** (A-B) Detection of (A) dPCR (H5_Taq and RP_Bov) and (B) qPCR (H5_Taq) assays using

531 serial dilutions of synthetic H5N1 RNA standard material. For qPCR data, we combined and

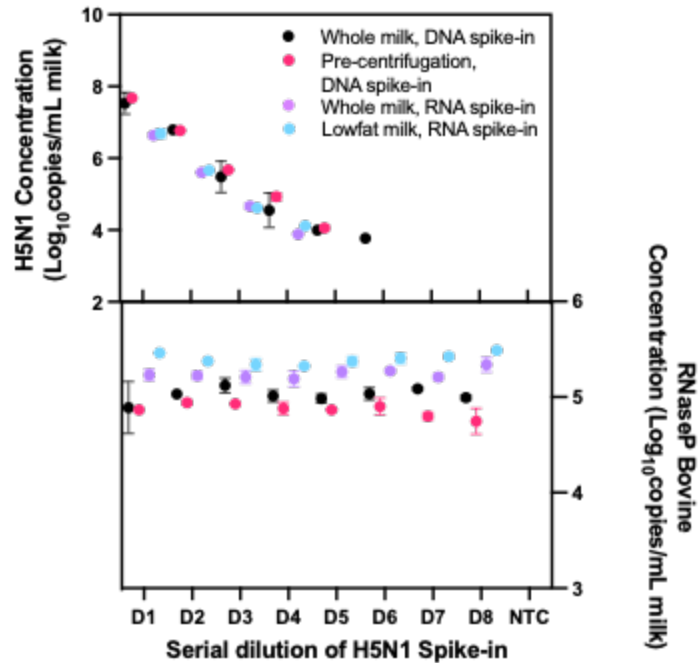
532 jointly analyzed all standard curve data from runs during retail milk testing. (C-D) Limit of

533 Detection (LOD) analysis for (C) measured concentrations compared to expected concentrations

534 for both qPCR and dPCR and (D) correlation of dPCR concentrations with qPCR Ct values. Fitted

535 lines in (A) and (D) represent simple linear regression lines of best fit.

536



537

538 **Figure 2: dPCR detection of H5N1 synthetic nucleic acid (top) and RNaseP Bovine (bottom) for**

539 **the MagMAX CORE extraction kit. For direct extraction, we extracted 200 μ L of milk spiked with**

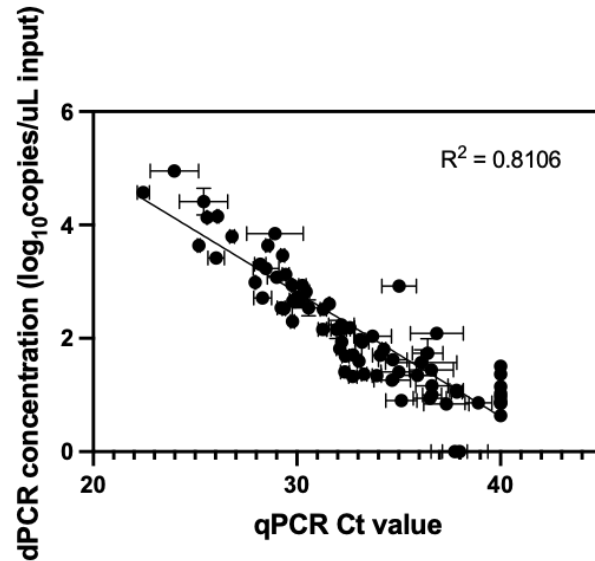
540 **serial dilutions of H5N1 synthetic fragments. For pre-centrifugation, we centrifuged samples for**

541 **12000xg for 10 minutes following spike-in, after which we extracted 200 μ L.**

A)

		qPCR	
		+	-
dPCR	+	46	9
	-	2	157

B)



542

543 **Figure 3: Comparison of dPCR and qPCR H5N1 testing on retail milk samples.** (A) Agreement of

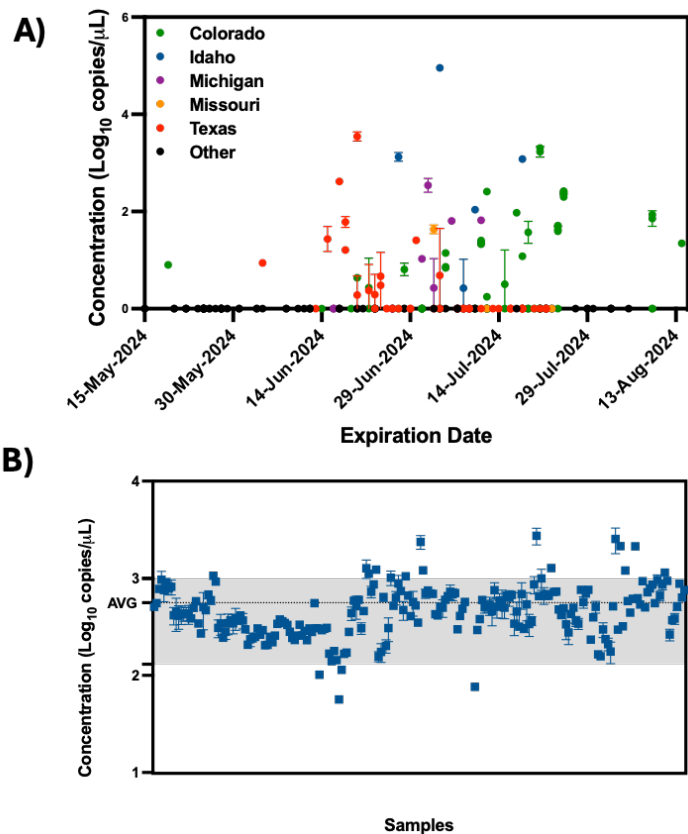
544 positive and negative calls of milk samples between the two platforms. (B) Correlation of H5N1

545 measured by dPCR concentration compared with qPCR Ct value. For plotting purposes, samples

546 not detected by dPCR were graphed with a dPCR concentration of 0 copies/ μ L while samples not

547 detected by qPCR were graphed with a Ct value of 40.

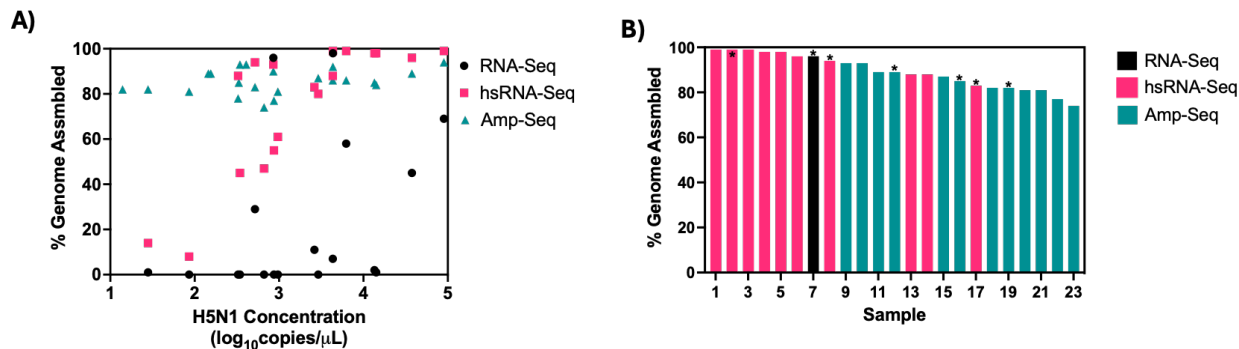
548



549
550 **Figure 4: H5N1 and Bovine Ribonuclease P (RP_Bov) for all retail milk samples as measured by**
551 **dPCR.** A) The concentration of H5N1 as a function of processing state and expiration date. B)
552 RP_Bov data for all samples. The gray-shaded region corresponds to the average RP_Bov
553 concentration of all data plus and minus one standard deviation.

554

555



556

557 **Figure 5:** (A) Completeness of H5N1 genome assemblies generated by unbiased metagenomics

558 (RNA-Seq), virus-enriched hybrid-selected metagenomics (hsRNA-Seq) and targeted H5N1

559 Amplicon Sequencing (Amp-Seq) as a function of H5N1 copies/μL RNA. (B): the most complete

560 H5N1 assembly produced for each sample sorted by length and the underlying sequencing

561 approach. Ultra-pasteurized samples are indicated by an (*) above the bar.