Establishing farm dust as a useful viral metagenomic surveillance matrix

Kirsty T. T. Kwok[a]*, Myrna M. T. de Rooij[b], Aniek B. Messink[b], Inge M. Wouters[b], Lidwien A. M. Smit[b], Dick J.J. Heederik[b], Marion P. G. Koopmans[a], My V. T. Phan[a]†#

[a] Department of Viroscience, Erasmus Medical Center, Rotterdam, The Netherlands

[b] Institute for Risk Assessment Sciences (IRAS), Utrecht University, Utrecht, The Netherlands

Running Head: Viromes from farm dust and farmed chicken feces

*Current affiliation: MRC-University of Glasgow Centre for Virus Research, Glasgow, United Kingdom

†Current affiliation: London School of Hygiene and Tropical Medicine, London, United Kingdom

Address correspondence to Kirsty T.T. Kwok (Kirsty.Kwok@glasgow.ac.uk) and My V.T. Phan (my.phan@lshtm.ac.uk)

Keywords: virome; dust; one health; airborne farm dust; viral metagenomics; picornavirus; astrovirus; calicivirus; poultry

Word counts for main text: 3439 words (max = 4500 words)

ABSTRACT (192 words, word limit = 200 words)

Farm animals may harbor many viral pathogens, some with zoonotic potential which can possibly cause severe clinical outcomes in animals and humans. Documenting viral content of dust may provide information of potential sources and movement of viruses. Here, we describe a dust sequencing strategy that provides detailed viral sequence characterization from farm dust samples and use this method to document the virus communities from chicken farm dust samples and paired feces collected at multiple time points during the production cycle in broiler farms in the Netherlands. From the sequencing data, *Parvoviridae* and *Picornaviridae* were the most prevalent families, detected in 85-100% of all feces and dust samples. Surprisingly large genomic diversity was identified in *Picornaviridae* and viruses from the *Caliciviridae* and *Astroviridae* were also obtained. This study provides a unique characterization of virus communities in farmed chickens and paired farm dust samples and our sequencing methodology enabled the recovery of viral genome sequences from farm dust, providing important tracking details for virus movement between livestock animals and their farm environment. This study serves as a proof of concept supporting that dust sampling could potentially be incorporated as part of viral metagenomic surveillance.

1. INTRODUCTION

Many emerging infectious diseases are of zoonotic origins [1]. Approximately 70 % of zoonoses are proposed to originate from wildlife and outbreaks of MERS-CoV [2], Nipah virus [3] and SARS-CoV-2 in minks [4] have shown that livestock and farmed animals can act as intermediate hosts and/or reservoirs which can sustain zoonotic transmission. In addition to passing viruses from wildlife to humans, livestock and farmed animals are also reservoir of zoonotic pathogens, such as bacteria *Coxiella burnetii* (*C. burnettii*) that causes Q-fever in goats, sheep and cattle [5], and avian influenza [6,7]. Additionally, epidemiological studies have found evidence for an increased risk of respiratory disease in persons exposed to large-scale animal farming [8,9]. In part, this can be attributed to farm dust exposure for which adverse health effects have been demonstrated [10–12], and the potential of zoonotic infections in these cases could be important yet under-reported. For instance, a Q-fever outbreak affecting approximately 4,000 people was reported in the Netherlands in 2007-2010 that was linked to circulation of *C. burnettii* in local dairy goat farms and dairy sheep farms [13]. Q-fever was not notifiable in small ruminant farms in the

30 Netherlands before the outbreak, making early detection of zoonotic transmission of *C. burnettii*
31 challenging [14]. Similarly, MERS-CoV had been circulating among dromedary camels long before it was
32 recognized as a cause of severe respiratory disease in humans [15]. These livestock-associated zoonotic
33 disease outbreaks emphasize that continued surveillance of health risks due to livestock exposure is
34 essential, especially when the intensification of livestock production is widely practiced to meet
35 increasing food demands of growing populations [16].
36 Advances in next-generation sequencing (NGS) technology can provide a detailed description of the
37 viral content of environmental samples offering great potential for viral surveillance. This metagenomic
38 NGS (mNGS) approach (using random priming rather than target specific primers) allows the detection
39 of viruses present in the samples without prior knowledge of its presence [17,18] as well as characterization
40 of complex viral communities [19,20]. Indeed, reports of virus diversity in animals including bats [21–23], pigs
41 [24–26], wild birds [27,28], and chickens [29] have been published since the development of NGS methods.
42 Chickens (*Gallus gallus domesticus*) are the most abundant domestic livestock in the world.
43 According to the Food and Agriculture Organization of the United Nations (FAO), the global chicken
44 population was >22 billion birds in 2017 [30]. Chicken farming ranges from backyard to large-scale
45 commercial farms that house ten-thousands of chickens. The Netherlands is the second largest
46 agricultural exporter in the world [31,32]. There are approximately 0.1 billion housed chickens and around
47 1,700 chicken farms in the Netherlands in 2020 [33]. Chicken can carry various pathogens such as
48 *Salmonella* [34], *Campylobacter* [35] and avian influenza viruses including subtypes that can cause severe
49 infection in humans like H5N1 [36] and H7N9 [6]; and it has been found that avian influenza virus can be
50 transmitted to humans as a result of close exposure to contaminated chicken [7]. Monitoring virus
51 diversity in chickens and other livestock may be a key component for identifying potential zoonotic
52 threats. Meanwhile, little is known about the virome composition of farm dust, although it has been
53 suggested that farm dust might play a role in infectious disease transmission [37,38]. Farm dust typically
54 consists of fragments of animal feed, bedding material, animal feces, animal dander, mites and
55 microorganisms [39], as well as fragments of poultry feathers in the case of poultry farm dust. The
56 complex composition of dust particles allows them to accumulate and serve as a vector of biologically
57 active material including antibiotic resistance genes, endotoxin, and, as shown here, complete viral
58 genomes. Farm dust exposure may cause adverse health effects [40,41]. Most studies on farm dust have

59 focused on measuring antimicrobial gene sequences [42] and endotoxins [10,12], with no published reports
60 documenting the viral sequence content of farm dust samples.
61 Here, we describe a metagenomic NGS characterization of virus diversity in a paired sample set of
62 farm dust and chicken feces collected at multiple time points in three commercial chicken farms across
63 the Netherlands. We develop and present a sequencing methodology that enables recovery of virus
64 genome sequences in farm dust samples. We show that the chicken fecal virome and farm dust virome
65 show similar patterns over time, host age and farms. This study provides useful insights in
66 understanding the virus communities in chicken and surrounding farm environments and provides an
67 important tool for monitoring viral movement by dust, a previously underappreciated source of viral
68 movement through our environment.
69

70 ## 2. RESULTS

71 *2.1 Comparative virus prevalence in chicken farm dust samples and paired feces*

72 A total of 46 individual farm dust samples and 56 pooled chicken feces were collected. Five
73 chicken fecal samples were excluded due to insufficient cDNA concentration for library preparation.
74 Individual farm dust samples collected at the same sampling moment from the same farm were pooled
75 together during sample processing for sequencing, making up a final set of 13 pooled farm dust
76 samples. Metagenomic deep sequencing of pooled chicken farm dust samples (N=13) and pooled
77 chicken feces (N=51) generated ca 1-3 million paired-end short reads per sample.
78 From the total of 64 chicken farm dust samples and chicken feces analyzed, viruses belonging to
79 the family *Parvoviridae* were consistently detected in all samples, followed by *Picornaviridae* with a 85%
80 and 100% prevalence in chicken dust samples and chicken feces respectively (**Table 1**). *Astroviridae*
81 were the third most prevalent virus family detected in 85% of farm dust samples and 82% chicken feces,
82 followed by *Caliciviridae* which were present in 69% (farm dust samples) and 78% (chicken feces)
83 prevalence. The overall pattern of virus detection observed in dust samples was similar to that in
84 chicken feces, some more variability was observed within farms over time (**Table 1** and **Figure S1**).
85 Although the detection frequency of *Parvoviridae* was slightly higher than that of *Picornaviridae*,
86 the number of *Picornaviridae* contigs detected was the highest (**Figure 1**). Within the *Picornaviridae*
87 family, *Sicinivirus* was consistently detected in 69% of farm dust samples and all chicken feces (**Table 2**),

88 followed by *Anativirus* (farm dust samples: 61%; chicken feces: 70%), and *Megrivirus* (farm dust
89 samples: 46%; chicken feces: 43%). Of note, the number of contigs detected in different picornavirus
90 genera corresponded to the prevalence pattern (**Figure 2 & Figure S2**). In particular, *Anatvirus*,
91 *Gallivirus*, *Megrivirus* and *Sicinivirus* were detected in all three farms, while *Orivirus* was only identified
92 in farm F_01 and F_03 and *Avisivirus* was only found in farm F_03 (**Figure S2**). When looking at virus
93 family level, we did not observe any substantial differences in overall viral contents (**Figure 1**) and virus-
94 specific prevalence (**Figure S1**) of samples from different farms and from different production cycles
95 (different chicken ages) within farms.
96
97 *2.2 Genomic diversity in viruses identified from chicken farm dust samples and paired feces*
98 2.2.1 *Picornaviridae*
99 Phylogenetic construction was performed to compare the reported *Picornaviridae* sequences
100 from 5 genera (*Avisvirus*, *Anativirus*, *Gallivirus*, *Megrivirus* and *Sicinivirus*) with global sequences. Forty-
101 six picornavirus sequences with ≥ 80% genome coverage (farm dust samples: 8; chicken feces: 38) were
102 included in the phylogenetic analysis of the polyprotein region. The maximum-likelihood (ML)
103 phylogenetic tree (**Figure 3**) showed that the sequences identified from farm F_03, albeit small number
104 (N=12), belonged to all 5 different genera *Avisvirus*, *Anativirus*, *Gallivirus*, *Megrivirus* and *Sicinivirus*,
105 suggesting a high viral diversity circulating among chicken in this farm. Interestingly, farm F_02 with the
106 highest number of picornavirus contigs identified (N=31) has a moderate level of viral diversity, with
107 these contigs belonging to 3 genera *Anativirus*, *Sicinivirus* and *Megrivirus*. Only 3 picornavirus genomic
108 sequences identified were from farm F_01, belonging to *Sicinivirus* (N=1) and *Anativirus* (N=2).
109 *Sicinivirus* and *Anativirus* are the most prevalent virus genera found with virus sequences detected in all
110 3 farms, while sequences of *Gallivirus* and *Avisivirus* were less prevalent, found in only 1 farm, farm
111 F_03.
112 Genome sequences detected from farm dust samples belonged to *Sicinivirus* and *Megrivirus*
113 from farm F_02 that clustered with sequences identified in chicken feces from the same farm (**Figure 3**).
114 Viruses in the *Sicinivirus* genus were detected in 69% of dust samples and all chicken feces samples with
115 the highest number of viral contigs identified in each sample as compared to other genera in the
116 *Picornaviridae* family. Slightly longer branches in the *Sicinivirus* sequences in the ML tree indicates

117 higher diversity within this *Sicinivirus* clade (**Figure 3)**. We further reconstructed a ML tree for all
118 identified *Sicinivirus* polyprotein nucleotide sequences (N=15; farm dust samples: 4; chicken feces: 11) in
119 this study and compared them with global reference sequences (**Figure 4**). Two sequences from chicken
120 feces (V_M_005_picorna_5 from farm F_01 and V_M_034_picorna_3 from farm F_03) were most
121 distinct from the rest of the identified sequences (sharing only 69.4-72.3% nt identity) and were most
122 closely related to strain JSY previously identified in China (78.9-79.0% nt identity). Other *Sicinivirus*
123 sequences from the same farm often formed a monophyletic lineage, suggesting within-farm *Sicinivirus*
124 sequences are highly genetically similar. Viruses identified from chicken feces and farm dust samples
125 consistently clustered with each other forming their own sub-clusters within major clades in the ML
126 trees, indicating high similarity between sequences from feces and dusts and that they are more closely
127 related as compared to global reference sequences.
128
129 2.2.2 *Astroviridae*
130 Sixteen astrovirus sequences with ≥ 80% genome coverage were identified in this study, among
131 which only 1 astrovirus sequence was found in farm dust (VE_7_astro_14; farm F_02). There was only 1
132 astrovirus genomic sequence from farm F_01. RNA-dependent RNA polymerase (RdRp) and capsid
133 regions of these sequences were extracted for phylogenetic analyses (**Figure 5A and 5B**, respectively).
134 ML trees of both regions showed that these 16 astrovirus sequences belonged to two distinct lineages,
135 A1 and A2 (**Figure 5**). Sequences from samples from farm F_02 and F_03 were found in both lineages A1
136 and A2, suggesting co-circulation of different astrovirus strains in both farms. The 12 sequences from
137 lineage A1 were most closely related to avian nephritis virus strains previously reported in China and
138 Brazil (GenBank accession number: HM029238, MN732558 and MH028405), while the 4 sequences
139 from lineage A2 were more closely related to each other and clustered with chicken astroviruses
140 previously identified in Malaysia, Canada and the USA (GenBank accession number: MT491731-2,
141 MT789782, MT789784 and JF414802). The 12 sequences from lineage A1 shared 96-100% amino acid
142 (aa) identity in the RdRp region while 4 sequences from lineage A2 shared a higher 99-100% aa identity.
143 For the less conserved capsid region, three small sub-lineages can be observed within lineage A1 (**Figure**
144 **5B**). Notably, the only astrovirus sequence found in dust (strain VE_7_astro_14) was identical to
145 astrovirus recovered from its paired chicken feces sample (strain V_M_038_astro_4) at the same farm.

146

147 2.2.3 *Caliciviridae*

148 Similar to astrovirus distribution in feces versus dust, 13 out of 14 calicivirus sequences identified
149 (with ≥ 80% genome coverage) were recovered from fecal samples with only 1 calicivirus sequence
150 identified from a dust sample. Partial ORF1 polyprotein region of these sequences was extracted to
151 construct ML tree comparing the local calicivirus with global sequences. The ML tree showed that the
152 reported calicivirus sequences were from two distinct lineages (C1 and C2) (**Figure 6**). All calicivirus
153 sequences from farm F_02 and F_03 belonged to lineage C1, while sequences from F_01 belonged to
154 both lineages C1 and C2. The 11 sequences in lineage C1 shared 98.7-100% aa identity despite being
155 collected from different farms and they shared 97.2-99.1% aa similarity with two partial ORF1
156 sequences identified in Germany in 2010 (GenBank accession number: JQ347523 and JQ347527). BLAST
157 searches and multiple sequence alignment indicated that these 11 local calicivirus sequences are the
158 first report of complete ORF1 sequence of that particular chicken calicivirus lineage. The only calicivirus
159 sequence from dust (VE_8_calici_66) clustered in lineage C1. The 3 sequences in lineage C2 are highly
160 similar, sharing 99.9-100% aa identity and were all collected from farm F_01 at two time points. These 3
161 sequences shared 93.8-96.8% aa similarity with reference sequences from Germany, USA, Korea and
162 Brazil (*Bavaria virus*, *Bavovirus* genus, GenBank accession number: HQ010042, MN810875, KM254170-
163 1, MG846433-4).

164

165 2.2.4 Other virus families

166 Three nearly-complete coronavirus genome sequences were identified in 3 chicken feces (V_M_001,
167 V_M_002 and V_M_003) from farm F_01. BLAST analyses showed that these sequences are infectious
168 bronchitis virus (IBV) strains and share 99.92-99.95% identity at the nucleotide level with IBV vaccine
169 strain 4/91 (GenBank accession number: KF377577), suggesting these IBV may possibly be viral shedding
170 following vaccination or environmental contamination as vaccines are generally provided in drinking
171 water.

172

173 *2.3 MASH analysis*

174 The phylogenetic comparison of large contig sequences presented above support the findings about
175 the similarities between sequences from dust samples and possible farm animal sources. However,
176 additional information can be obtained with global analyses that would compare all available sequence
177 data from the dust material. A method that allows a comparison of the large amount of unclassified
178 sequences that are commonly observed in NGS data would provide additional information. We
179 therefore used a kmer comparison tool MASH [43,44] that prepares a hash description of all kmers of a
180 given length and then allows rapid quantitative comparison of similarity of the kmers sets generated
181 from different sequenced samples. We used the Mash distance function to calculate a Jaccard distance
182 value for all pairs of dust sequence samples.
183 We compared all assembled viral contigs of the dust samples. Figure 7 panel A shows the Jaccard
184 distance between all 13 dust samples from 3 farms, colored by farm source. In all comparisons, pairs of
185 dust samples from the same farm have lower Jaccard distance values (darker blue in the figure) than
186 inter-farm samples. It is also apparent that farm F_01 and farm F_02 were slightly more related to each
187 other than either were to farm F_03. Similar patterns were obtained using all quality-controlled short
188 read data from each dust sample (Figure 7 Panel B).
189

190 ## 3. DISCUSSION

191 Animals including livestock kept in close proximity to humans can act as intermediate hosts and/or
192 sources for zoonotic disease transmission [45]. The current COVID-19 pandemic has raised awareness of
193 zoonotic risks and has prompted the need for more proactive surveillance strategies involving animal
194 surveillance, which could potentially guide zoonotic outbreak preparedness. Chickens are the most
195 abundant livestock in the world [30] and have been identified as reservoirs for various zoonotic pathogens
196 [35,36]. In this study, we employed viral metagenomic deep sequencing to document the viruses present in
197 farm dust samples and paired chicken feces collected within the same farms. These data revealed
198 potential virus flow between chickens and their farm environment.
199 In this study, we described a random-primed viral metagenomic deep sequencing strategy to viral
200 sequences from dust samples. We were able to obtain long viral sequences providing ≥ 80% viral
201 genome coverage from farm dust samples. Farm dust is known to be associated with adverse
202 respiratory effects observed in farm workers with prolonged exposure [40]; however, whether viruses play

203 a role is thus far unknown due to a knowledge gap in virus detection and virus diversity in farm dust.
204 Characterizing farm dust viromes could aid in investigating possible health effects of occupational and
205 environmental exposure to virus-containing farm dust. The method we have developed could provide a
206 platform for future surveillance in farm animals and environmental samples.

207 Our phylogenetic analyses indicated that for all three farms, viruses identified from farm dust
208 samples were genetically closer to viruses identified from chicken feces collected from the same farm
209 than from the other two farms in the study. Furthermore, the Jaccard similarity analysis further showed
210 that for both known and unknown sequences the dust sequences were closer to other samples from the
211 same farm, indicating some source specificity in the sampling method. The combined data support the
212 idea that farm dust could potentially be a good proxy for the animals in that farm. This observation is in
213 agreement with a previous study that reported the correlation found in the bacterial antimicrobial
214 resistomes between animal feces and farm dust [46]. More importantly, the detection of likely chicken-
215 derived virus sequences in dust samples suggested that dust-borne transmission of viruses is a real
216 possibility for chicken and/or zoonotic pathogenic viruses [47]. Although infectivity of the viruses
217 identified was not determined in our study, the presence of nearly complete viral genome sequences
218 and the complex protein/lipid/nucleic acid composition of dust makes it likely that dust indeed contains
219 infectious virus particles. Future investigation on the infectivity of viruses in farm dust will provide
220 useful insights in mapping virus flows between chickens and their surrounding environment, which
221 collectively would guide future outbreak control strategies within and between farms.

222 Viruses identified in chicken feces from different farms were phylogenetically closely related to each
223 other and formed sub-lineages or monophyletic lineages when comparing to global related sequences.
224 This could possibly be explained by the pyramid structure of the broiler production system, as chicks
225 found at each farm were all bought from a few suppliers; in other words, chicks could have been
226 infected from the supplier and then brought the virus to different broiler farms with further spread
227 within the farms, leading to similar viruses circulating in different broiler farms. Of note, the pyramid
228 production structure was also suggested as a possible transmission routes for ESBL/pAmpC producing
229 bacteria between broiler farms [48]. Future characterization of virus diversity of chickens from the top part
230 of the pyramid (suppliers) is needed to confirm the speculation.

231 Viruses belonging to five genera of the *Picornaviridae* family were identified in our study, suggesting
232 a great diversity of picornaviruses circulating in chickens, which is consistent with previous studies that
233 detected various picornavirus genera in chicken respiratory samples [49,50]. Astroviruses identified in this
234 study clustered in two distinct phylogenetic lineages; these astroviruses could present possible new
235 genotypes. For astrovirus, all chicken astroviruses are grouped to one single species according to the
236 current taxonomy assignment from International Committee on Taxonomy of Viruses [51]. Current
237 classification of avastrovirus is based on phylogenetic analysis of capsid (ORF2) amino acid sequences [52].
238 More comprehensive assignment of genotypes is hindered by the limited number of complete capsid
239 sequences. A similar observation was seen for *Caliciviridae*, where we identified two types of
240 caliciviruses forming two distinct lineages in the ML tree. Lineage C2 is closely related to the *Bavovirus*
241 genus, while the sequences from C1 are most related to an unclassified chicken calicivirus strain
242 F10026n [53] that only possessed a partial ORF1 sequence. Our data have provided 16 astrovirus and 14
243 calicivirus sequences, which helps expand the current knowledge of chicken astrovirus and calicivirus
244 diversity and could potentially contribute to improved taxonomic assignment. Identification of novel
245 calicivirus and astrovirus genome sequences indicates that continued effort is needed for increasing
246 understanding of virus diversity in chicken feces.
247 We observed that viruses in the *Parvoviridae* and *Picornaviridae* families were consistently detected
248 in chicken feces and farm dust samples at multiple time points in all three farms. These virus families
249 could potentially be used as signature markers when monitoring chicken and farm dust exposure.
250 Additionally, these viruses could also be used as markers to assess the efficiency of disinfection
251 procedures in farms and their impact on surrounding environments although stabilities of different
252 viruses may differ due to different structures and we caution that viral activity assays have not yet been
253 performed on this material. Further and perhaps larger scale of longitudinal surveillance with the
254 inclusion of samples from layer hens and more environmental samples from surrounding areas are
255 certainly necessary to confirm our observation and the applicability of using these identified viruses as
256 signature virus fingerprint.
257 One limitation that should be considered is that fecal and dust samples were pooled and dust by its
258 very nature is a pooled sample (i.e. possibly containing sequences from multiple individual infections).
259 The possibility that the resulting assembled genomes are chimeras assembled from multiple individual

260 infections should be considered in interpreting the results. Truly striking and virologically important
261 variants should be examined using alternate methods such as direct isolation from individually infected
262 animals. Although we collected chicken feces and farm dust samples longitudinally, limited number of
263 samples per each time point especially for farm dust samples has hindered in-depth statistical
264 comparisons between samples collected from different time points.
265 In conclusion, we provide a dataset of virus sequences generated from farm dust samples and feces
266 and show a correlation between sequences from dust and feces samples obtained within the same
267 farm. These results support the idea that dust sampling can provide an accurate description of the farm
268 virome and could potentially be incorporated as part of viral metagenomic surveillance matrix. In the
269 long term, understanding virus flow between animals and humans is vital for identifying potential
270 zoonotic threats and zoonotic outbreak control and preparedness.
271

272 ## 4. MATERIALS AND METHODS

273 *4.1 Sample collection*

274 Pooled chicken feces and farm dust samples were collected in three commercial broiler farms in the
275 Netherlands at 4-5 time points from May to August 2019. The three commercial broiler farms are
276 located in three different regions in the Netherlands (Western part: Noord/Zuid-Holland region;
277 Northern part: Friesland/Groningen region; Eastern part: Gelderland region). Detailed sampling strategy
278 and sample metadata is described in **Table 3**. Each pooled poultry fecal sample contains fresh fecal
279 material from 3-4 chicks. Farm dust samples were collected using a passive air sampling approach using
280 electrostatic dustfall collectors (EDCs) [54,55]. Electrostatic cloths were sterilized through incubation at
281 200°C for 4 hours. Sterilized electrostatic cloths were then fixed to a pre-cleaned plastic frame. EDCs
282 were exposed for 7 days at 1 meter above the floor with the electrostatic cloths facing up in broiler
283 farms to enable sampling of settling airborne dust instead of resuspended dust from the floor. EDCs
284 were contained in a sterile plastic bag before and after sampling. All samples were transported under
285 cold chain management and stored at -20°C/-80°C before processing.
286

287 *4.2 Sample processing and metagenomic deep sequencing*

288 Chicken fecal samples were processed as previously described [56]. Briefly, chicken fecal suspension
289 was prepared in Phosphate-buffered saline (PBS) and treated with TURBO DNAse (Thermo Fisher, USA)
290 at 37 °C for 30 minutes, and then subjected to total nucleic acid extraction using QIAamp viral RNA mini
291 kit (Qiagen, Germany) according to manufacturer's instruction without addition of carrier RNA. For dust
292 samples, electrostatic cloths were incubated in 3% beef extract buffer for 1 hour on rolling as previously
293 published [57]. After incubation, the suspension was collected and centrifugated at 4,000 g at 4°C for 4
294 minutes to pellet any large particles or debris. Total viruses in dust suspension were concentrated using
295 polyethylene glycol (PEG) similar to virus concentration in sewage published previously [58,59]. Briefly, PEG
296 6000 was added to each dust suspension to make up a final 10% PEG 6000 (Sigma-Aldrich, USA)
297 concentration, followed by pH adjustment to pH 4 and overnight incubation at 4°C with shaking. After
298 overnight incubation, sample was centrifuged at 13,500 g at 4°C for 90 minutes. Supernatant was
299 removed and the remaining pellet was resuspended in 500 μL of pre-warmed glycine buffer, and then
300 subjected to 5-minute centrifugation at 13,000 g at 4°C. Supernatant was collected, and supernatants
301 from EDC samples that were collected in the same farm at the same time point (N=1-4) were pooled
302 together for further processing. Viral-enriched dust suspension samples were then treated with TURBO
303 DNase as previously described [56] to remove non-encapsulated DNA, followed by total nucleic acid
304 extraction using MagMAX™ viral RNA isolation kit (Thermo Fisher, USA) according to manufacturer'
305 instructions but without the use of carrier RNA. Reverse transcription and second strand cDNA synthesis
306 of chicken fecal samples and dust samples was performed as previously described [56], using SuperScript
307 III reverse transcriptase and non-ribosomal hexamers [60] (reverse transcription) and Klenow fragment 3'-
308 5' exo- (New England BioLabs, USA) (2nd strand cDNA synthesis), respectively.
309 Standard Illumina libraries were prepared using Nextera XT DNA library kit following the
310 manufacturer's instruction. Final libraries were sequenced on the Illumina MiSeq platform (600 cycles;
311 paired-end). Chicken fecal samples with a multiplex range of 16-19 samples. Dust samples and negative
312 blank controls were sequenced with a multiplex of 12 samples per run.
313
314 *4.3 Sequencing data analysis*
315 Raw reads were subjected to adapter removal using Trim Galore/default Illumina software, followed
316 by quality trimming using QUASR [61] with a threshold of minimum length of 125 nt and median Phred

317 score ≤ 30. The resulting quality controlled reads were *de novo* assembled using metaSPAdes v3.12.0 [62].
318 *De novo* assembled contigs were classified using UBLAST [63] against eukaryotic virus family protein
319 databases as previously described [23,26]. We set a detection threshold of contig with minimum amino acid
320 identity of 70%, minimum length of 300 nt and an e-value threshold of 1×10^{-10} when interpreting our
321 contig classification results. Contig classification results were analyzed and visualized using R packages
322 including dplyr, reshaped2 and ComplexHeatmap [64–66].
323
324 *4.4 Phylogenetic analysis*
325 Sequences with ≥ 80% of genome coverage from the *Picornaviridae*, *Astroviridae* and *Caliciviridae*
326 families were used for phylogenetic analyses, comparing them with global sequences retrieved from
327 GenBank. Multiple sequence alignment was performed using MAFFT [67] and manually checked in
328 Geneious v2021.0.3 (Biomatters, New Zealand). Best-fit models of evolution for phylogenetic tree
329 reconstruction were estimated using ModelFinder module in IQ-TREE v1.6.11 [68], determined by Akaike
330 Information Criterion (AIC). Maximum-likelihood trees were constructed using RAxML-NG v1.0.1 [69] with
331 100 pseudo-replicates. Resulting trees were visualized and annotated using R package ggtree [70].
332
333 *4.5 Kmer analysis with MASH*
334 To compare Jaccard distances between dust samples, all quality controlled short reads, or resulting
335 assembled viral contigs were analyzed using the Triangle function from MASH v2.3 [43,44] with a kmer size
336 of 32 (-k 32) and a sketch size of 10,000 (-s 10000). The resulting distances were visualized in a heatmap
337 using R package ggplot2 [71].
338
339 *4.6 Data availability*
340 The raw reads are available in the SRA under the BioProject accession number PRJNA670873 (chicken
341 feces) and PRJNA701384 (farm dust samples) (**Table S1**). All sequences in phylogenetic analysis have
342 been deposited in GenBank under the accession numbers MW684778 to MW684847 (**Table S2**).
343
344 **ACKNOWLEDGMENTS**

We thank all the participating farms for their cooperation, Pella van der Wal and Gaby van Dijk from Erasmus MC for providing sequencing service. We thank Prof. Matt Cotten for his advice on Mash analysis and critical reading and comments on the manuscript. This virome study was funded by ZonMW TOP project 91217040 (K.T.T.K., M.M.T.dR., L.A.M.S., D.J.J.H. and M.P.G.K.) and a Marie Sklodowska-Curie Individual Fellowship, supported by Horizon 2020 (grant no. 799417; M.V.T.P.). M.V.T.P was additionally supported by the Wellcome Trust and FCDO (grant agreement number 220977/Z/20/Z). Sample collection was supported with internal funds from the Utrecht University (M.M.T.dR., A.B.M., and I.M.W.).

AUTHOR CONTRIBUTION

K.T.T.K., M.P.G.K. and M.V.T.P. conceived the study. K.T.T.K. and M.V.T.P. performed experiments. M.M.T.dR and I.M.W. designed sampling strategy. M.M.T.dR. and A.B.M. conducted fieldwork, handled sample collection and collected associated metadata. K.T.T.K. and M.V.T.P. analyzed data. K.T.T.K. drafted the manuscript. M.V.T.P., M.P.G.K., L.A.M.S., I.M.W., M.M.T.dR and A.B.M. edited the manuscript. M.V.T.P., M.P.G.K. and I.M.W. supervised the study. M.P.G.K., D.J.J.H. and L.A.M.S. acquired the funding.

REFERENCES

1. Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nature* (2008). doi:10.1038/nature06536
2. Azhar, E. I. *et al.* Evidence for camel-to-human transmission of MERS coronavirus. *N. Engl. J. Med.* (2014). doi:10.1056/NEJMoa1401505
3. Hsu, V. P. *et al.* Nipah virus encephalitis reemergence, Bangladesh. *Emerg. Infect. Dis.* (2004). doi:10.3201/eid1012.040701
4. Oude Munnink, B. B. *et al.* Transmission of SARS-CoV-2 on mink farms between humans and mink and back to humans. *Science (80-.).* **371**, 172–177 (2020).
5. Ganter, M. Zoonotic risks from small ruminants. *Vet. Microbiol.* (2015). doi:10.1016/j.vetmic.2015.07.015
6. Chen, Y. *et al.* Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: Clinical analysis and characterisation of viral genome. *Lancet* (2013). doi:10.1016/S0140-6736(13)60903-4
7. Koopmans, M. *et al.* Transmission of H7N7 avian influenza A virus to human beings during a large outbreak in commercial poultry farms in the Netherlands. *Lancet* (2004). doi:10.1016/S0140-6736(04)15589-X
8. Kalkowska, D. A. *et al.* Associations between pneumonia and residential distance to livestock farms over a five-year period in a large population-based study. *PLoS One* (2018). doi:10.1371/journal.pone.0200813
9. Smit, L. A. M. *et al.* Increased risk of pneumonia in residents living near poultry farms: does the upper respiratory tract microbiota play a role? *Pneumonia* (2017). doi:10.1186/s41479-017-0027-0
10. Kirychuk, S. P. *et al.* Endotoxin and dust at respirable and nonrespirable particle sizes are not consistent between cage- and floor-housed poultry operations. *Ann. Occup. Hyg.* (2010). doi:10.1093/annhyg/meq047
11. Rimac, D. *et al.* Exposure to poultry dust and health effects in poultry workers: Impact of mould and mite allergens. *Int. Arch. Occup. Environ. Health* (2010). doi:10.1007/s00420-009-0487-5
12. Dosman, J. A. *et al.* Respiratory response to endotoxin and dust predicts evidence of

inflammatory response in volunteers in a swine barn. *Am. J. Ind. Med.* (2006). doi:10.1002/ajim.20339

13. Schimmer, B. *et al.* Sustained intensive transmission of Q fever in the south of the Netherlands, 2009. *Eurosurveillance* **14**, (2009).
14. Van Der Hoek, W. *et al.* Epidemic Q fever in humans in the Netherlands. *Adv. Exp. Med. Biol.* (2012). doi:10.1007/978-94-007-4315-1_17
15. Corman, V. M. *et al.* Antibodies against MERS coronavirus in dromedary camels, Kenya, 1992-2013. *Emerg. Infect. Dis.* (2014). doi:10.3201/eid2008.140596
16. Jones, B. A. *et al.* Zoonosis emergence linked to agricultural intensification and environmental change. *Proc. Natl. Acad. Sci. U. S. A.* (2013). doi:10.1073/pnas.1208059110
17. Barzon, L., Lavezzo, E., Militello, V., Toppo, S. & Palù, G. Applications of next-generation sequencing technologies to diagnostic virology. *International Journal of Molecular Sciences* (2011). doi:10.3390/ijms12117861
18. Bexfield, N. & Kellam, P. Metagenomics and the molecular identification of novel viruses. *Veterinary Journal* (2011). doi:10.1016/j.tvjl.2010.10.014
19. Dávila-Ramos, S. *et al.* A review on viral metagenomics in extreme environments. *Front. Microbiol.* (2019). doi:10.3389/fmicb.2019.02403
20. Kristensen, D. M., Mushegian, A. R., Dolja, V. V. & Koonin, E. V. New dimensions of the virus world discovered through metagenomics. *Trends Microbiol.* (2010). doi:10.1016/j.tim.2009.11.003
21. Zheng, X. yan *et al.* Viral metagenomics of six bat species in close contact with humans in southern China. *Arch. Virol.* (2018). doi:10.1007/s00705-017-3570-3
22. Dacheux, L. *et al.* A preliminary study of viral metagenomics of french bat species in contact with humans: Identification of new mammalian viruses. *PLoS One* (2014). doi:10.1371/journal.pone.0087194
23. Phan, M. V. T. *et al.* Identification and characterization of Coronaviridae genomes from Vietnamese bats and rats based on conserved protein domains. *Virus Evol.* (2018). doi:10.1093/ve/vey035
24. Chen, Q. *et al.* Metagenomic analysis of the RNA fraction of the fecal virome indicates high diversity in pigs infected by porcine endemic diarrhea virus in the United States. *Virol J* **15**, (2018).

25. Shan, T. *et al.* The Fecal Virome of Pigs on a High-Density Farm. *J. Virol.* (2011). doi:10.1128/jvi.05217-11

26. Phan, M. V. T. *et al.* Unbiased whole-genome deep sequencing of human and porcine stool samples reveals circulation ofmultiple groups of rotaviruses and a putative zoonotic infection. *Virus Evol.* (2016). doi:10.1093/ve/vew027

27. Vibin, J. *et al.* Metagenomics detection and characterisation of viruses in faecal samples from Australian wild birds. *Sci. Rep.* (2018). doi:10.1038/s41598-018-26851-1

28. Vibin, J., Chamings, A., Klaassen, M., Bhatta, T. R. & Alexandersen, S. Metagenomic characterisation of avian parvoviruses and picornaviruses from Australian wild ducks. *Sci. Rep.* (2020). doi:10.1038/s41598-020-69557-z

29. Lima, D. A. *et al.* Faecal virome of healthy chickens reveals a large diversity of the eukaryote viral community, including novel circular ssDNA viruses. *J. Gen. Virol.* (2017). doi:10.1099/jgv.0.000711

30. FAO. Gateway to poultry production and products. (2021). Available at: http://www.fao.org/poultry-production-products/production/poultry-species/chickens/en/. (Accessed: 23rd January 2021)

31. Netherlandsandyou.nl. Agriculture and food: the Netherlands and Finland. Available at: https://www.netherlandsandyou.nl/your-country-and-the-netherlands/finland/and-the-netherlands/agriculture-and-food. (Accessed: 23rd January 2021)

32. CBS. *CBS Internationalisation Monitor 2016-II Agribusiness*. (2016).

33. Centraal Bureau voor de Statistiek. Agriculture; crops, livestock and land use by general farm type, region. (2021). Available at: https://www.cbs.nl/en-gb/figures/detail/80783eng?q=chicken. (Accessed: 23rd January 2021)

34. Finstad, S., O'Bryan, C. A., Marcy, J. A., Crandall, P. G. & Ricke, S. C. Salmonella and broiler processing in the United States: Relationship to foodborne salmonellosis. *Food Res. Int.* (2012). doi:10.1016/j.foodres.2011.03.057

35. Silva, J. *et al.* Campylobacter spp. As a foodborne pathogen: A review. *Frontiers in Microbiology* (2011). doi:10.3389/fmicb.2011.00200

36. Sims, L. D. *et al.* Origin and evolution of highly pathogenic H5N1 avian influenza in Asia. *Vet. Rec.* (2005). doi:10.1136/vr.157.6.159

37. De Rooij, M. M. T. *et al.* Detection of Coxiella Burnetii in ambient air after a large Q fever outbreak. *PLoS One* (2016). doi:10.1371/journal.pone.0151281

38. Zhao, Y. *et al.* Airborne transmission may have played a role in the spread of 2015 highly pathogenic avian influenza outbreaks in the United States. *Sci. Rep.* (2019). doi:10.1038/s41598-019-47788-z

39. Cambra-López, M., Torres, A. G., Aarnink, A. J. A. & Ogink, N. W. M. Source analysis of fine and coarse particulate matter from livestock houses. *Atmos. Environ.* (2011). doi:10.1016/j.atmosenv.2010.10.018

40. Viegas, S. *et al.* Occupational exposure to poultry dust and effects on the respiratory system in workers. *J. Toxicol. Environ. Heal. - Part A Curr. Issues* (2013). doi:10.1080/15287394.2013.757199

41. Danuser, B., Weber, C., Künzli, N., Schindler, C. & Nowak, D. Respiratory symptoms in Swiss farmers: An epidemiological study of risk factors. *Am. J. Ind. Med.* (2001). doi:10.1002/ajim.1032

42. Munk, P. *et al.* Abundance and diversity of the faecal resistome in slaughter pigs and broilers in nine European countries. *Nat. Microbiol.* (2018). doi:10.1038/s41564-018-0192-9

43. Ondov, B. D. *et al.* Mash: Fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **17**, 1–14 (2016).

44. Ondov, B. D. *et al.* Mash Screen: High-throughput sequence containment estimation for genome discovery. *Genome Biol.* **20**, 1–13 (2019).

45. Gardy, J. L. & Loman, N. J. Towards a genomics-informed, real-time, global pathogen surveillance system. *Nature Reviews Genetics* (2018). doi:10.1038/nrg.2017.88

46. Luiken, R. E. C. *et al.* Farm dust resistomes and bacterial microbiomes in European poultry and pig farms. *Environ. Int.* (2020). doi:10.1016/j.envint.2020.105971

47. Scoizec, A. *et al.* Airborne detection of H5N8 highly pathogenic avian influenza virus genome in poultry farms, France. *Front. Vet. Sci.* (2018). doi:10.3389/fvets.2018.00015

48. Dame-Korevaar, A., Fischer, E. A. J., van der Goot, J., Stegeman, A. & Mevius, D. Transmission routes of ESBL/pAmpC producing bacteria in the broiler production pyramid, a literature review. *Preventive Veterinary Medicine* (2019). doi:10.1016/j.prevetmed.2018.12.002

49. Lau, S. K. P. *et al.* Chickens host diverse picornaviruses originated from potential interspecies transmission with recombination. *J. Gen. Virol.* (2014). doi:10.1099/vir.0.066597-0

50. Boros, Á. *et al.* A diarrheic chicken simultaneously co-infected with multiple picornaviruses: Complete genome analysis of avian picornaviruses representing up to six genera. *Virology* (2016). doi:10.1016/j.virol.2015.12.002

51. Lefkowitz, E. J. *et al.* Virus taxonomy: The database of the International Committee on Taxonomy of Viruses (ICTV). *Nucleic Acids Res.* (2018). doi:10.1093/nar/gkx932

52. Donato, C. & Vijaykrishna, D. The broad host range and genetic diversity of mammalian and avian astroviruses. *Viruses* (2017). doi:10.3390/v9050102

53. Wolf, S. *et al.* Discovery and genetic characterization of novel caliciviruses in German and Dutch poultry. *Arch. Virol.* (2012). doi:10.1007/s00705-012-1326-7

54. Noss, I. *et al.* Passive airborne dust sampling with the electrostatic dustfall collector: Optimization of storage and extraction procedures for endotoxin and glucan measurement. *Ann. Occup. Hyg.* (2010). doi:10.1093/annhyg/meq026

55. Liebers, V. *et al.* Assessment of airborne exposure to endotoxin and pyrogenic active dust using Electrostatic Dustfall Collectors (EDCs). *J. Toxicol. Environ. Heal. - Part A Curr. Issues* (2012). doi:10.1080/15287394.2012.674919

56. Kwok, K. T. T. *et al.* Genome Sequences of Seven Megrivirus Strains from Chickens in The Netherlands . *Microbiol. Resour. Announc.* (2020). doi:10.1128/mra.01207-20

57. Prussin, A. J. *et al.* Seasonal dynamics of DNA and RNA viral bioaerosol communities in a daycare center. *Microbiome* (2019). doi:10.1186/s40168-019-0672-z

58. Strubbia, S. *et al.* Metavirome sequencing to evaluate norovirus diversity in sewage and related bioaccumulated oysters. *Front. Microbiol.* (2019). doi:10.3389/fmicb.2019.02394

59. Nieuwenhuijse, D. F. *et al.* Setting a baseline for global urban virome surveillance in sewage. *Sci. Rep.* (2020). doi:10.1038/s41598-020-69869-0

60. Endoh, D. *et al.* Species-independent detection of RNA virus by representational difference analysis using non-ribosomal hexanucleotides for reverse transcription. *Nucleic Acids Res.* **33**, e65 (2005).

61. Watson, S. J. *et al.* Viral population analysis and minority-variant detection using short read next-generation sequencing. *Philos. Trans. R. Soc. B Biol. Sci.* (2013). doi:10.1098/rstb.2012.0205

62. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. MetaSPAdes: A new versatile

metagenomic assembler. *Genome Res.* (2017). doi:10.1101/gr.213959.116

63. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* (2010). doi:10.1093/bioinformatics/btq461

64. Wickham, H., François, R., Henry, L. & Müller, K. dplyr: A Grammar of Data Manipulation. (2019). Available at: https://cran.r-project.org/package=dplyr.

65. Wickham, H. Reshaping data with the reshape package. *J. Stat. Softw.* (2007). doi:10.18637/jss.v021.i12

66. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* (2016). doi:10.1093/bioinformatics/btw313

67. Katoh, K., Misawa, K., Kuma, K. I. & Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* (2002). doi:10.1093/nar/gkf436

68. Nguyen, L. T., Schmidt, H. A., Von Haeseler, A. & Minh, B. Q. IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* (2015). doi:10.1093/molbev/msu300

69. Kozlov, A. M., Darriba, D., Flouri, T., Morel, B. & Stamatakis, A. RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* (2019). doi:10.1093/bioinformatics/btz305

70. Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T. Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* (2017). doi:10.1111/2041-210X.12628

71. Wickham, H. ggplot2. *ggplot2* (2009). doi:10.1007/978-0-387-98141-3

Figure legends

Figure 1. An overview of virus contigs detected in chicken feces and chicken farm dust samples. Each column represents one sample. Associated metadata is shown at the top panel. The color intensity of the heatmap (bottom panel) is determined by number of contigs with minimum length of 300 nt/bp, at least 70% identity and an e-value threshold of $1x10^{-10}$ when comparing with closest reference in our database. Sample order is assorted by farm ID, sampling time point. Time point is an arbitrary number. Samples were collected at the same time and place when both Farm ID and Time point match.

Figure 2. An overview of picornavirus contigs detected in chicken feces and chicken farm dust samples. Each column represents one sample. Associated metadata is shown at the top panel. The color intensity of the heatmap (bottom panel) is determined by number of contigs with minimum length of 300 nt, at least 70% identity and an e-value threshold of $1x10^{-10}$ when comparing with closest reference in our database. Sample order is assorted by farm ID, sampling time point. Time point is an arbitrary number. Samples were collected at the same time and place when both Farm ID and Time point match.

Figure 3. Maximum-likelihood phylogenetic tree of 76 (sequences from this study: 46; reference sequences: 30) picornavirus partial polyprotein amino acid sequences (3786 sites) reconstructed using a best-fit evolutionary model LG+FC+R10. For clarity, only bootstrap supports >= 70 are shown as a navy-blue node point. Tip labels of our samples are color-coded by corresponding farm ID. Reference sequences are in grey. Tree was mid-point rooted for clarity.

Figure 4. Maximum-likelihood phylogenetic tree of 30 sicinivirus partial polyprotein nucleotide sequences (sequences from this study: 15; reference sequences: 15; 8672 sites) reconstructed using a best-fit evolutionary model GTR+F+R10. For clarity, only bootstrap supports >= 70 are shown as a navy-blue node point. Tip labels of our samples are color-coded by corresponding farm ID. Reference sequences are in grey and tree was mid-point rooted for clarity.

Figure 5. A. Maximum-likelihood phylogenetic tree of 33 astrovirus partial RdRp amino acid sequences (sequences from this study: 16; reference sequences: 17; 548 sites) reconstructed using a best-fit evolutionary model LG+FC+I+G4m. B. Maximum-likelihood phylogenetic tree of 33 astrovirus partial capsid amino acid sequences (sequences from this study: 16; reference sequences: 17; 908 sites) reconstructed using a best-fit evolutionary model LG+FC+I+G4m. For clarity, only bootstrap supports >= 70 are shown as a navy-blue node point. Tip labels of our samples are color-coded by corresponding farm ID. Reference sequences are in grey and tree was mid-point rooted for clarity.

Figure 6. Maximum-likelihood phylogenetic tree of 30 calicivirus partial polyprotein amino acid sequences (sequences from this study: 14; reference sequences: 16; 2794 sites) reconstructed using a best-fit evolutionary model LG+FC+R10. For clarity, only bootstrap supports >= 70 are shown as a navy-blue node point. Tip labels of our samples are color-coded by corresponding farm ID. Reference sequences are in grey and tree was mid-point rooted for clarity.

Figure 7. A. Jaccard distance analysis of assembled viral contigs of 13 pooled farm dust samples. B. Jaccard distance analysis of quality-controlled short reads of 13 pooled farm dust samples. The color intensity of the heatmaps is determined by MASH distance. Corresponding MASH distances are also shown in the heatmap. Sample labels are color-coded by farm ID and farm ID was added as prefix for clarity.

Table 1. Prevalence of different eukaryotic virus families in chicken feces and farm dust samples

Sample types		Chicken feces (N=51)	Farm dust (N=13)
Virus family	***Astroviridae***	42 (82%)	11 (85%)
	Caliciviridae	40 (78%)	9 (69%)
	Coronaviridae	19 (37%)	3 (23%)
	Parvoviridae	51 (100%)	13 (100%)
	Picobirnaviridae	8 (16%)	2 (15%)
	Picornaviridae	51 (100%)	11 (85%)
	Reoviridae	7 (14%)	3 (23%)

Table 2. Prevalence of picornavirus genera in chicken feces and farm dust samples

Sample types		Chicken feces (N=51)	Farm dust (N=13)
Picornavirus genera	***Anativirus***	36 (70%)	8 (61%)
	Avisivirus	3 (5%)	0 (0%)
	Gallivirus	16 (31%)	4 (30%)
	Megrivirus	22 (43%)	6 (46%)
	Orivirus	8 (15%)	0 (0%)
	Sicinivirus	51 (100%)	9 (69%)

Table 3. Detailed sampling strategy

Farm #	Sampling moment identifier	Time point	No. of pooled fecal sample(s) collected	No. of pooled EDC sample(s) collected (no. of individual EDC sample[s] collected)	Week production cycle
F_01	F_01_1	1	4	1 (4)	3
	F_01_2	2	4	1 (4)	4
	F_01_3	3	4	1 (4)	5
	F_01_4	4	4	1 (1)	4
	F_01_5	5	4	0 (0)	5
F_02	F_02_1	1	4	1 (2)	4
	F_02_2	2	4	1 (4)	5
	F_02_3	3	4	1 (4)	2
	F_02_4	4	4	1 (4)	4
	F_02_5	5	4	1 (4)	5
F_03	F_03_1	1	4	1 (3)	4
	F_03_2	2	4	1 (4)	5
	F_03_3	3	4	1 (4)	4
	F_03_4	4	4	1 (4)	5

Supplementary materials legends.

Figure S1. An overview of virus family prevalence in chicken feces and farm dust samples. Associated metadata is shown at the top panel. The color intensity of the heatmap (bottom panel) is determined by number of contigs with minimum length of 300 nt, at least 70% identity and an e-value threshold of $1x10^{-10}$ when comparing with closest reference in our database. Sample order is assorted by farm ID, sampling time point. Time point is an arbitrary number. Samples were collected at the same time and place when both Farm ID and Time point match.

Figure S2. An overview of picornavirus genera prevalence in chicken feces and farm dust samples. Associated metadata is shown at the top panel. The color intensity of the heatmap (bottom panel) is determined by number of contigs with minimum length of 300 nt, at least 70% identity and an e-value threshold of $1x10^{-10}$ when comparing with closest reference in our database. Sample order is assorted by farm ID, sampling time point. Time point is an arbitrary number. Samples were collected at the same time and place when both Farm ID and Time point match.

Table S1. A list of corresponding BioProject, BioSample and SRA accession numbers for all samples.

Table S2. A list of corresponding GenBank accession numbers for sequences used in phylogenetic analyses.

Figure 1

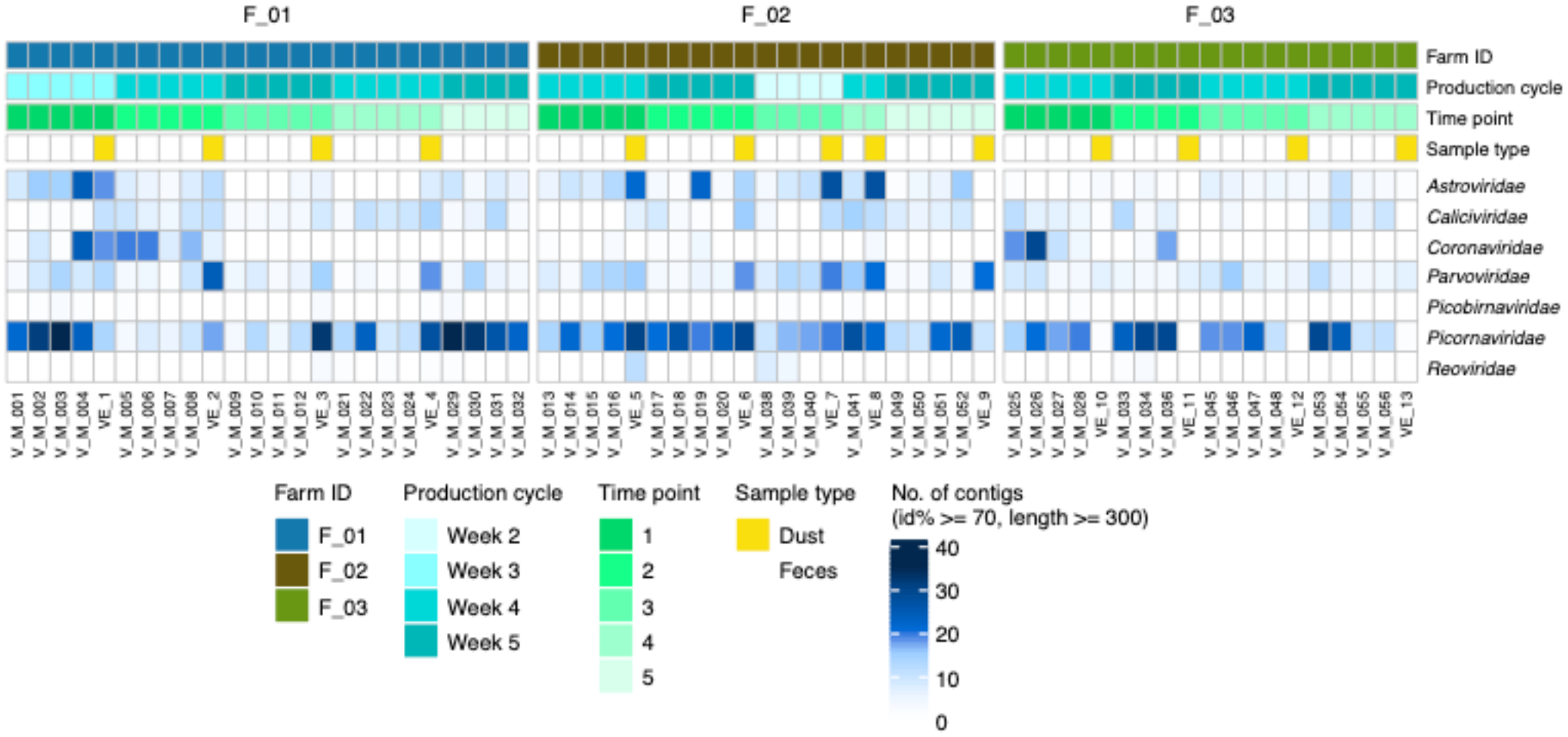

F_01
F_02
F_03
Farm ID
Production cycle
Time point
Sample type
Astroviridae
Caliciviridae
Coronaviridae
Parvoviridae
Picobirnaviridae
Picornaviridae
Reoviridae
V_M_001
V_M_002
V_M_003
V_M_004
VE_1
V_M_005
V_M_006
V_M_007
V_M_008
VE_2
V_M_009
V_M_010
V_M_011
V_M_012
VE_3
V_M_021
V_M_022
V_M_023
V_M_024
VE_4
V_M_029
V_M_030
V_M_031
V_M_032
V_M_013
V_M_014
V_M_015
V_M_016
VE_5
V_M_017
V_M_018
V_M_019
V_M_020
VE_6
V_M_038
V_M_039
V_M_040
VE_7
V_M_041
VE_8
V_M_049
V_M_050
V_M_051
V_M_052
VE_9
V_M_025
V_M_026
V_M_027
V_M_028
VE_10
V_M_033
V_M_034
V_M_036
VE_11
V_M_045
V_M_046
V_M_047
V_M_048
VE_12
V_M_053
V_M_054
V_M_055
V_M_056
VE_13
Farm ID
F_01
F_02
F_03
Production cycle
Week 2
Week 3
Week 4
Week 5
Time point
1
2
3
4
5
Sample type
Dust
Feces
No. of contigs
(id% >= 70, length >= 300)
40
30
20
10
0

Figure 2

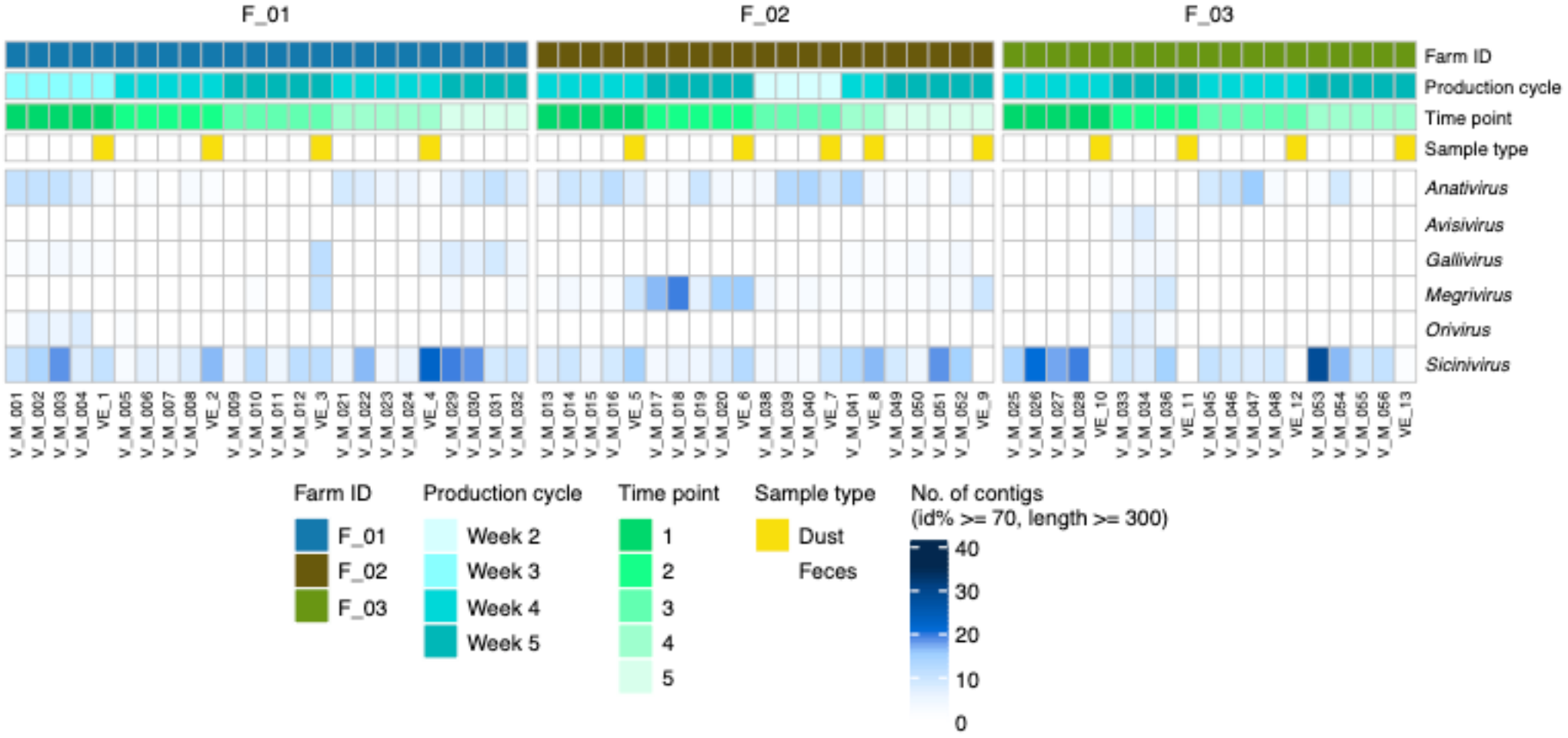

F_01
F_02
F_03
Farm ID
Production cycle
Time point
Sample type
Anativirus
Avisivirus
Gallivirus
Megrivirus
Orivirus
Sicinivirus
V_M_001
V_M_002
V_M_003
V_M_004
VE_1
V_M_005
V_M_006
V_M_007
V_M_008
VE_2
V_M_009
V_M_010
V_M_011
V_M_012
VE_3
V_M_021
V_M_022
V_M_023
V_M_024
VE_4
V_M_029
V_M_030
V_M_031
V_M_032
V_M_013
V_M_014
V_M_015
V_M_016
VE_5
V_M_017
V_M_018
V_M_019
V_M_020
VE_6
V_M_038
V_M_039
V_M_040
VE_7
V_M_041
VE_8
V_M_049
V_M_050
V_M_051
V_M_052
VE_9
V_M_025
V_M_026
V_M_027
V_M_028
VE_10
V_M_033
V_M_034
V_M_036
VE_11
V_M_045
V_M_046
V_M_047
V_M_048
VE_12
V_M_053
V_M_054
V_M_055
V_M_056
VE_13
Farm ID
F_01
F_02
F_03
Production cycle
Week 2
Week 3
Week 4
Week 5
Time point
1
2
3
4
5
Sample type
Dust
Feces
No. of contigs
(id% >= 70, length >= 300)
40
30
20
10
0

Figure 3

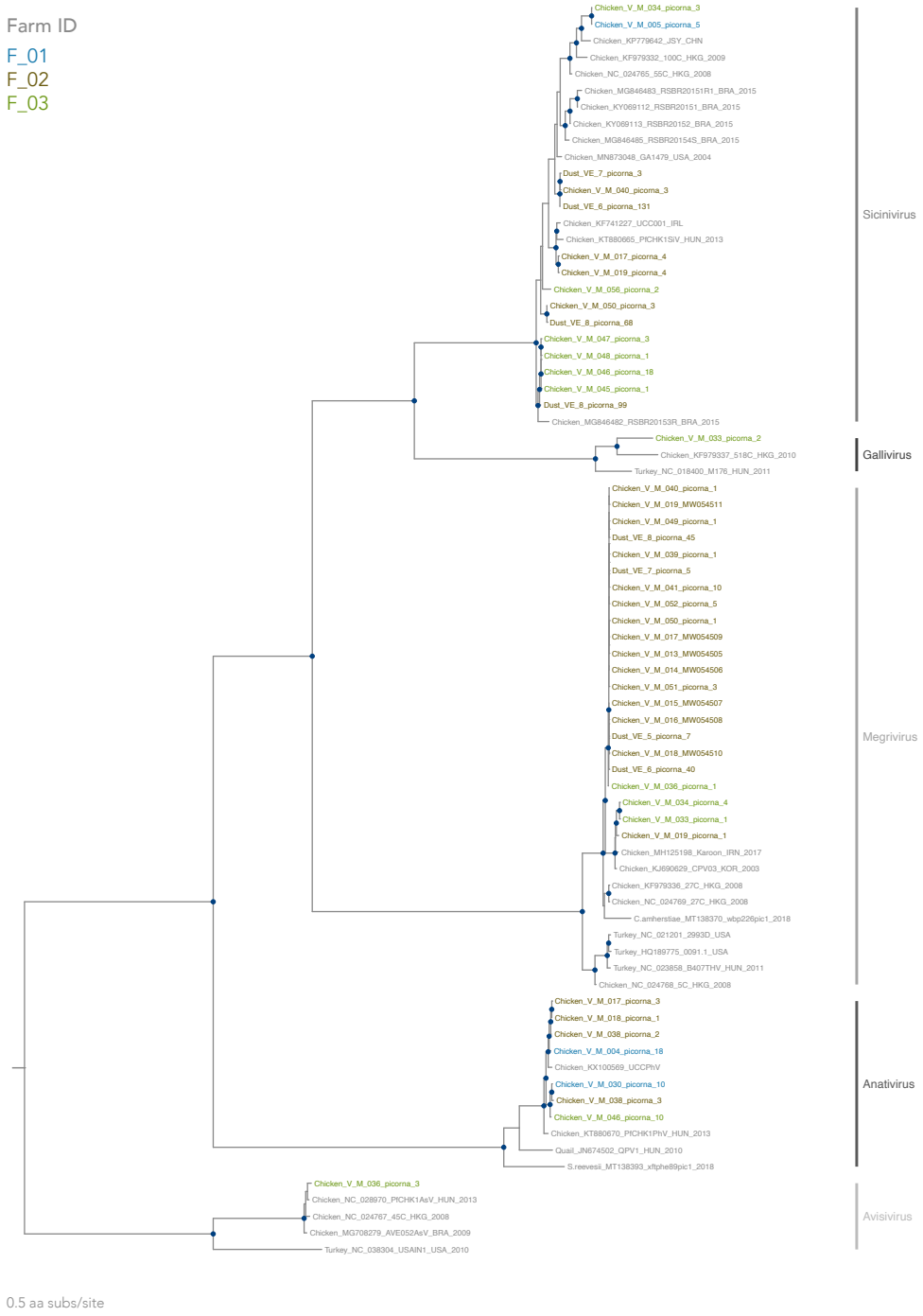

Farm ID
F_01
F_02
F_03
Chicken_V_M_034_picorna_3
Chicken_V_M_005_picorna_5
Chicken_KP779642_JSY_CHN
Chicken_KF979332_100C_HKG_2009
Chicken_NC_024765_55C_HKG_2008
Chicken_MG846483_RSBR20151R1_BRA_2015
Chicken_KY069112_RSBR20151_BRA_2015
Chicken_KY069113_RSBR20152_BRA_2015
Chicken_MG846485_RSBR20154S_BRA_2015
Chicken_MN873048_GA1479_USA_2004
Dust_VE_7_picorna_3
Chicken_V_M_040_picorna_3
Dust_VE_6_picorna_131
Chicken_KF741227_UCC001_IRL
Chicken_KT880665_PtCHK1SIV_HUN_2013
Chicken_V_M_017_picorna_4
Chicken_V_M_019_picorna_4
Chicken_V_M_056_picorna_2
Chicken_V_M_050_picorna_3
Dust_VE_8_picorna_68
Chicken_V_M_047_picorna_3
Chicken_V_M_048_picorna_1
Chicken_V_M_046_picorna_18
Chicken_V_M_045_picorna_1
Dust_VE_8_picorna_99
Chicken_MG846482_RSBR20153R_BRA_2015
Sicinivirus
Chicken_V_M_033_picorna_2
Chicken_KF979337_518C_HKG_2010
Turkey_NC_018400_M176_HUN_2011
Gallivirus
Chicken_V_M_040_picorna_1
Chicken_V_M_019_MW054511
Chicken_V_M_049_picorna_1
Dust_VE_8_picorna_45
Chicken_V_M_039_picorna_1
Dust_VE_7_picorna_5
Chicken_V_M_041_picorna_10
Chicken_V_M_052_picorna_5
Chicken_V_M_050_picorna_1
Chicken_V_M_017_MW054509
Chicken_V_M_013_MW054505
Chicken_V_M_014_MW054506
Chicken_V_M_051_picorna_3
Chicken_V_M_015_MW054507
Chicken_V_M_016_MW054508
Dust_VE_5_picorna_7
Chicken_V_M_018_MW054510
Dust_VE_6_picorna_40
Chicken_V_M_036_picorna_1
Chicken_V_M_034_picorna_4
Chicken_V_M_033_picorna_1
Chicken_V_M_019_picorna_1
Chicken_MH125198_Karoon_IRN_2017
Chicken_KJ690629_CPV03_KOR_2003
Chicken_KF979336_27C_HKG_2008
Chicken_NC_024769_27C_HKG_2008
C.amherstiae_MT138370_wbp226pic1_2018
Turkey_NC_021201_2993D_USA
Turkey_HQ189775_0091.1_USA
Turkey_NC_023858_B407THV_HUN_2011
Chicken_NC_024768_5C_HKG_2008
Megrivirus
Chicken_V_M_017_picorna_3
Chicken_V_M_018_picorna_1
Chicken_V_M_038_picorna_2
Chicken_V_M_004_picorna_18
Chicken_KX100569_UCCPhV
Chicken_V_M_030_picorna_10
Chicken_V_M_038_picorna_3
Chicken_V_M_046_picorna_10
Chicken_KT880670_PtCHK1PhV_HUN_2013
Quail_JN674502_QPV1_HUN_2010
S.reevesii_MT138393_xltphe89pic1_2018
Anativirus
Chicken_V_M_036_picorna_3
Chicken_NC_028970_PtCHK1AsV_HUN_2013
Chicken_NC_024767_45C_HKG_2008
Chicken_MG708279_AVE052AsV_BRA_2009
Turkey_NC_038304_USAIN1_USA_2010
Avisivirus
0.5 aa subs/site

Figure 4

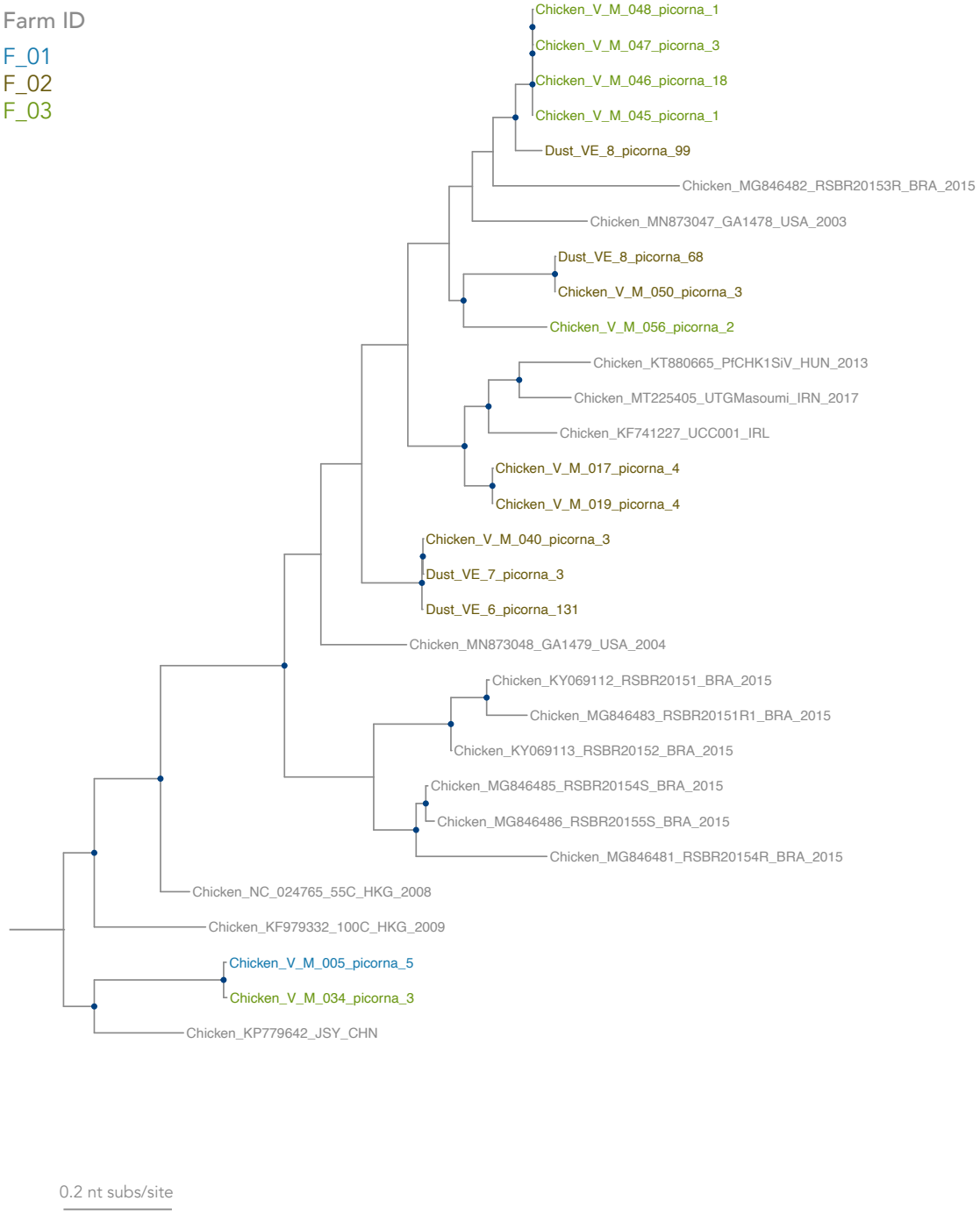

Farm ID
F_01
F_02
F_03
Chicken_V_M_048_picorna_1
Chicken_V_M_047_picorna_3
Chicken_V_M_046_picorna_18
Chicken_V_M_045_picorna_1
Dust_VE_8_picorna_99
Chicken_MG846482_RSBR20153R_BRA_2015
Chicken_MN873047_GA1478_USA_2003
Dust_VE_8_picorna_68
Chicken_V_M_050_picorna_3
Chicken_V_M_056_picorna_2
Chicken_KT880665_PfCHK1SiV_HUN_2013
Chicken_MT225405_UTGMasoumi_IRN_2017
Chicken_KF741227_UCC001_IRL
Chicken_V_M_017_picorna_4
Chicken_V_M_019_picorna_4
Chicken_V_M_040_picorna_3
Dust_VE_7_picorna_3
Dust_VE_6_picorna_131
Chicken_MN873048_GA1479_USA_2004
Chicken_KY069112_RSBR20151_BRA_2015
Chicken_MG846483_RSBR20151R1_BRA_2015
Chicken_KY069113_RSBR20152_BRA_2015
Chicken_MG846485_RSBR20154S_BRA_2015
Chicken_MG846486_RSBR20155S_BRA_2015
Chicken_MG846481_RSBR20154R_BRA_2015
Chicken_NC_024765_55C_HKG_2008
Chicken_KF979332_100C_HKG_2009
Chicken_V_M_005_picorna_5
Chicken_V_M_034_picorna_3
Chicken_KP779642_JSY_CHN
0.2 nt subs/site

Figure 5

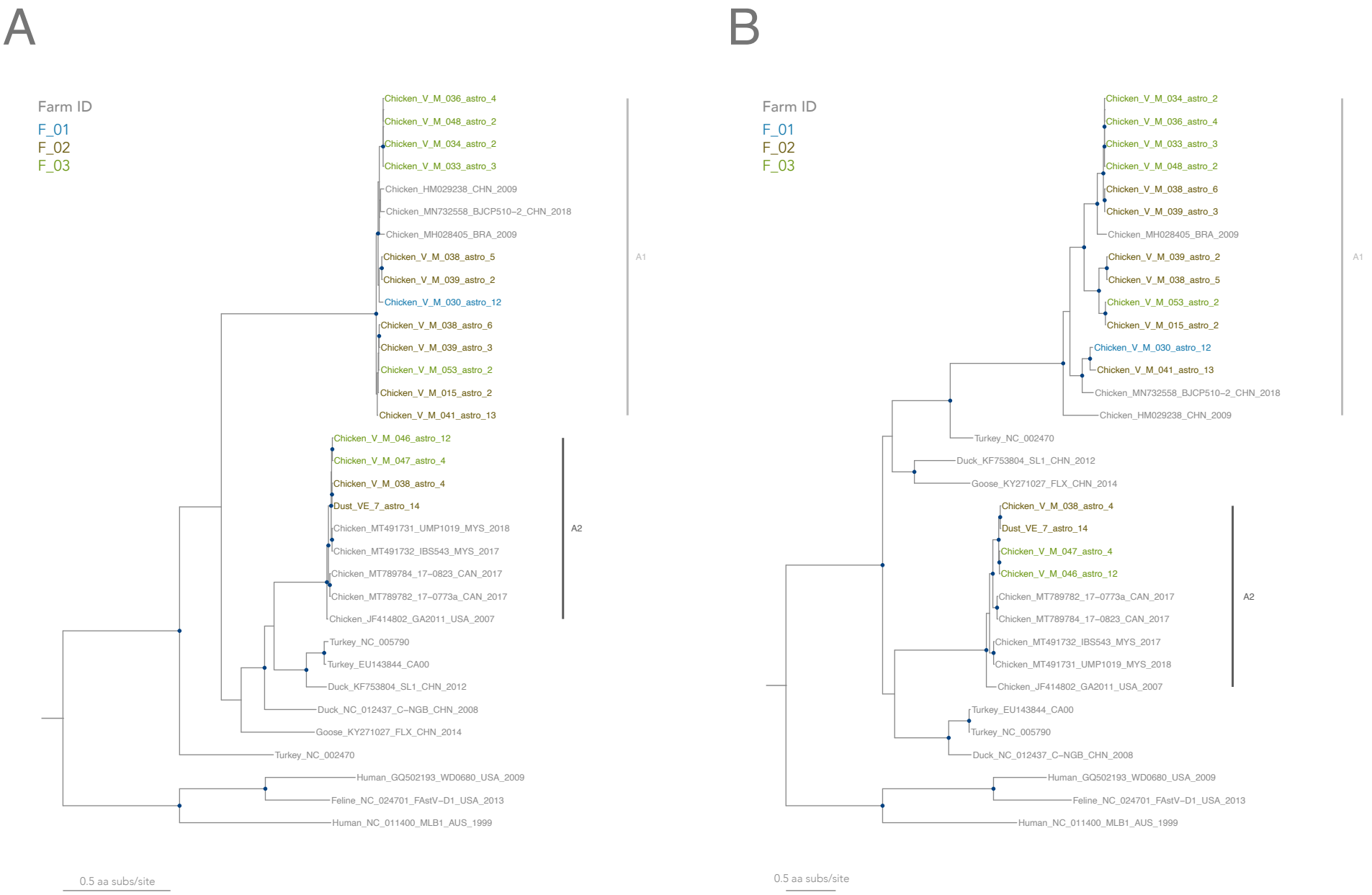

A
Farm ID
F_01
F_02
F_03
Chicken_V_M_036_astro_4
Chicken_V_M_048_astro_2
Chicken_V_M_034_astro_2
Chicken_V_M_033_astro_3
Chicken_HM029238_CHN_2009
Chicken_MN732558_BJCP510–2_CHN_2018
Chicken_MH028405_BRA_2009
Chicken_V_M_038_astro_5
Chicken_V_M_039_astro_2
Chicken_V_M_030_astro_12
Chicken_V_M_038_astro_6
Chicken_V_M_039_astro_3
Chicken_V_M_053_astro_2
Chicken_V_M_015_astro_2
Chicken_V_M_041_astro_13
A1
Chicken_V_M_046_astro_12
Chicken_V_M_047_astro_4
Chicken_V_M_038_astro_4
Dust_VE_7_astro_14
Chicken_MT491731_UMP1019_MYS_2018
Chicken_MT491732_IBS543_MYS_2017
Chicken_MT789784_17–0823_CAN_2017
Chicken_MT789782_17–0773a_CAN_2017
Chicken_JF414802_GA2011_USA_2007
A2
Turkey_NC_005790
Turkey_EU143844_CA00
Duck_KF753804_SL1_CHN_2012
Duck_NC_012437_C–NGB_CHN_2008
Goose_KY271027_FLX_CHN_2014
Turkey_NC_002470
Human_GQ502193_WD0680_USA_2009
Feline_NC_024701_FAstV–D1_USA_2013
Human_NC_011400_MLB1_AUS_1999
0.5 aa subs/site
B
Farm ID
F_01
F_02
F_03
Chicken_V_M_034_astro_2
Chicken_V_M_036_astro_4
Chicken_V_M_033_astro_3
Chicken_V_M_048_astro_2
Chicken_V_M_038_astro_6
Chicken_V_M_039_astro_3
Chicken_MH028405_BRA_2009
Chicken_V_M_039_astro_2
Chicken_V_M_038_astro_5
Chicken_V_M_053_astro_2
Chicken_V_M_015_astro_2
Chicken_V_M_030_astro_12
Chicken_V_M_041_astro_13
Chicken_MN732558_BJCP510–2_CHN_2018
Chicken_HM029238_CHN_2009
A1
Turkey_NC_002470
Duck_KF753804_SL1_CHN_2012
Goose_KY271027_FLX_CHN_2014
Chicken_V_M_038_astro_4
Dust_VE_7_astro_14
Chicken_V_M_047_astro_4
Chicken_V_M_046_astro_12
Chicken_MT789782_17–0773a_CAN_2017
Chicken_MT789784_17–0823_CAN_2017
Chicken_MT491732_IBS543_MYS_2017
Chicken_MT491731_UMP1019_MYS_2018
Chicken_JF414802_GA2011_USA_2007
A2
Turkey_EU143844_CA00
Turkey_NC_005790
Duck_NC_012437_C–NGB_CHN_2008
Human_GQ502193_WD0680_USA_2009
Feline_NC_024701_FAstV–D1_USA_2013
Human_NC_011400_MLB1_AUS_1999
0.5 aa subs/site

Figure 6

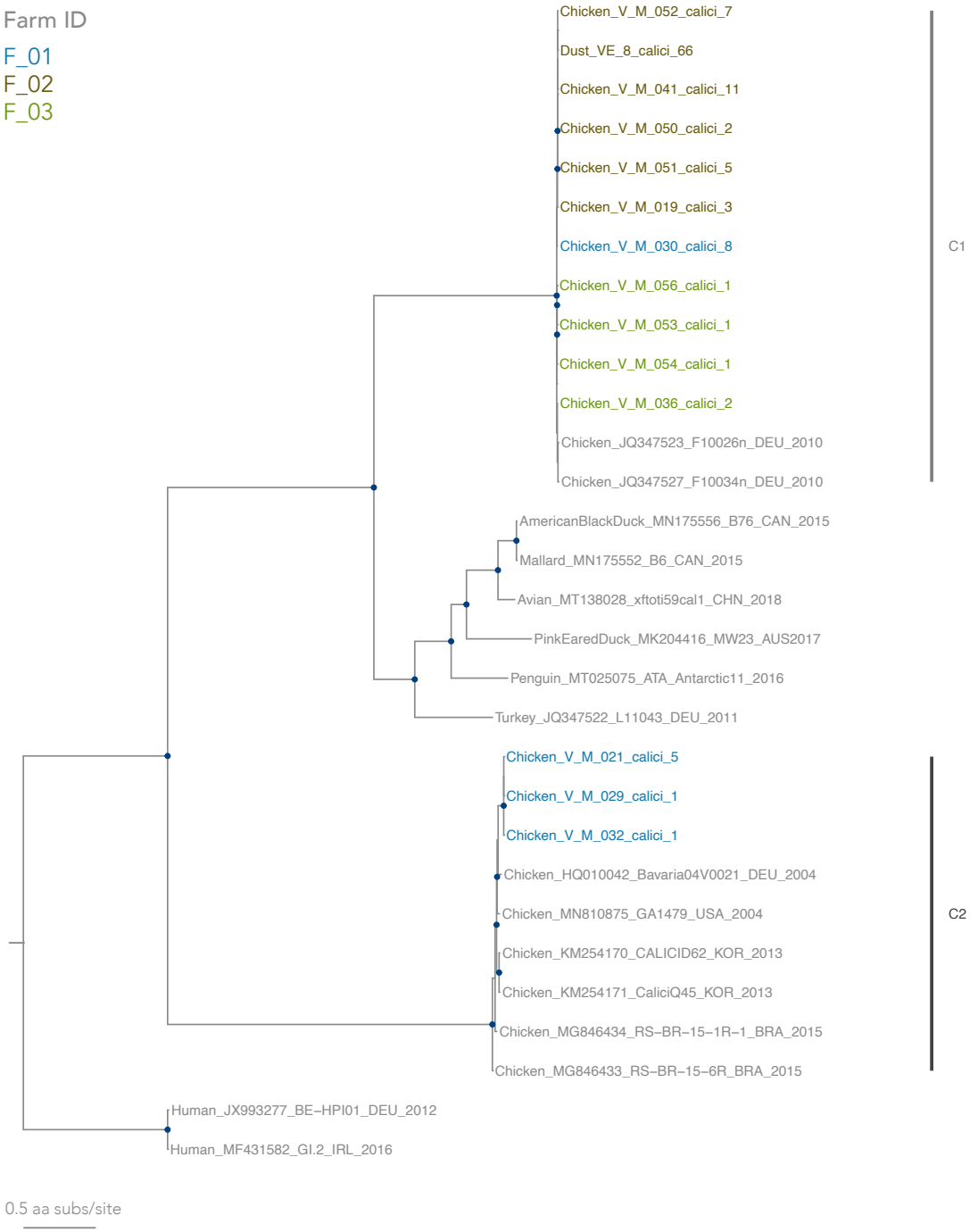
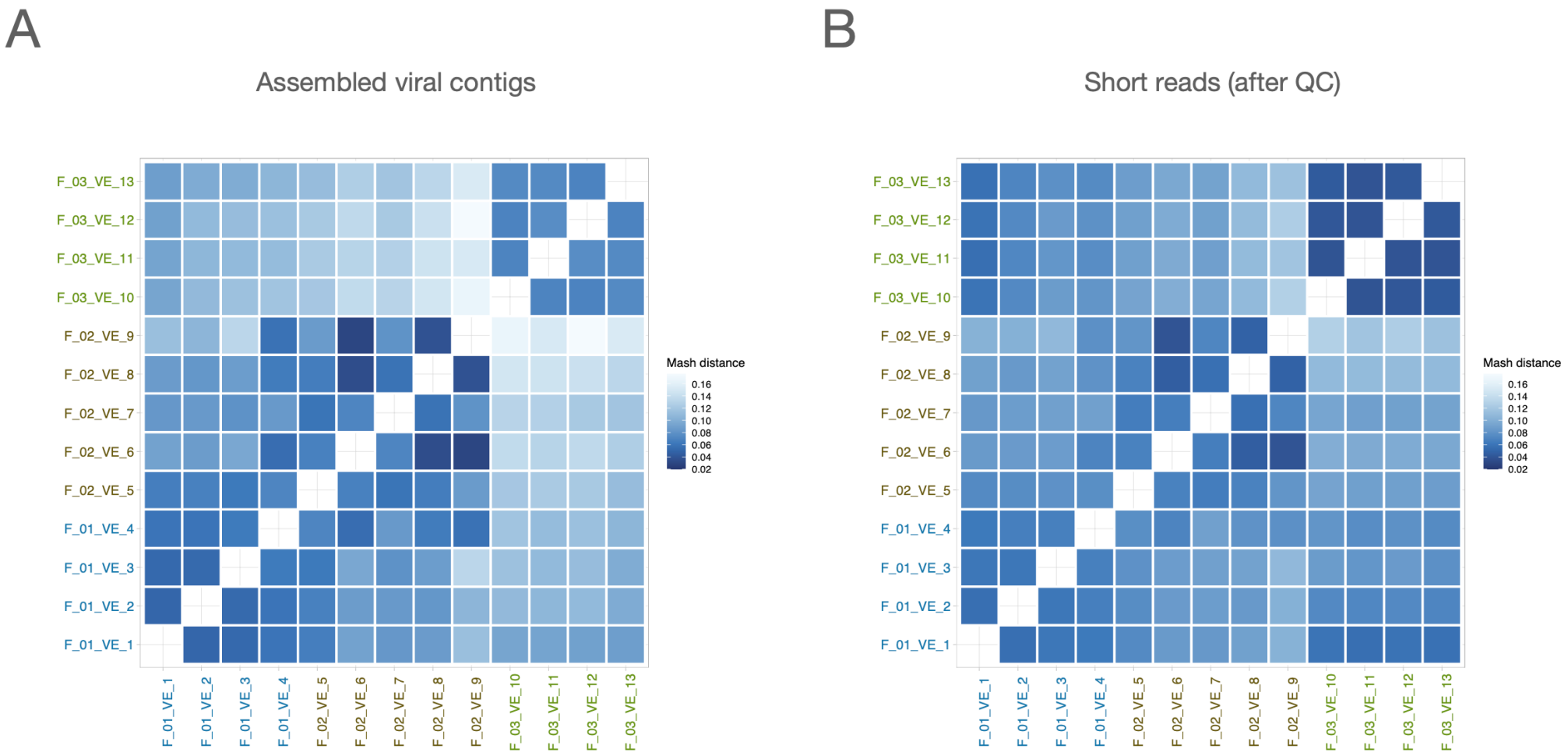
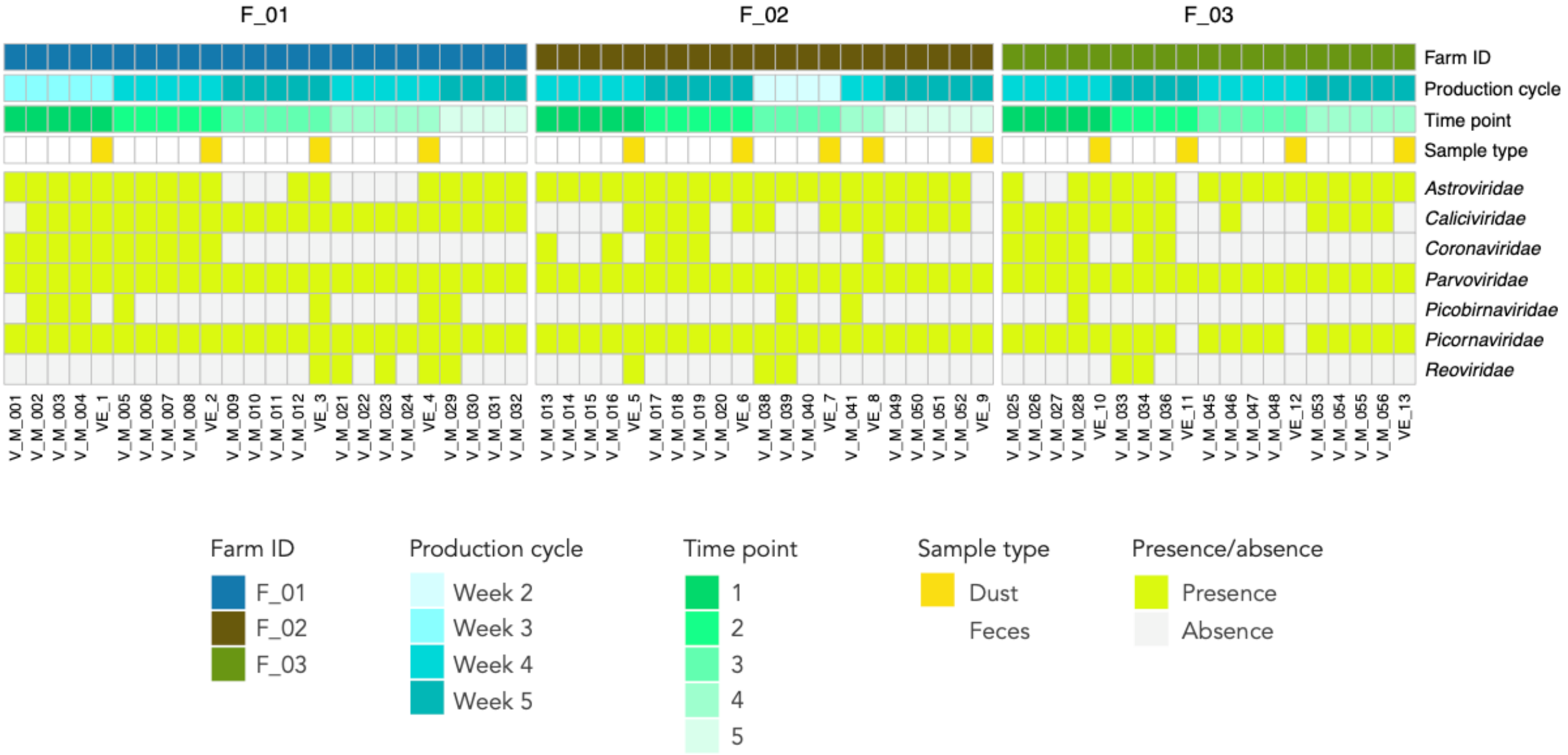

Farm ID
F_01
F_02
F_03
Chicken_V_M_052_calici_7
Dust_VE_8_calici_66
Chicken_V_M_041_calici_11
Chicken_V_M_050_calici_2
Chicken_V_M_051_calici_5
Chicken_V_M_019_calici_3
Chicken_V_M_030_calici_8
Chicken_V_M_056_calici_1
Chicken_V_M_053_calici_1
Chicken_V_M_054_calici_1
Chicken_V_M_036_calici_2
Chicken_JQ347523_F10026n_DEU_2010
Chicken_JQ347527_F10034n_DEU_2010
C1
AmericanBlackDuck_MN175556_B76_CAN_2015
Mallard_MN175552_B6_CAN_2015
Avian_MT138028_xftoti59cal1_CHN_2018
PinkEaredDuck_MK204416_MW23_AUS2017
Penguin_MT025075_ATA_Antarctic11_2016
Turkey_JQ347522_L11043_DEU_2011
Chicken_V_M_021_calici_5
Chicken_V_M_029_calici_1
Chicken_V_M_032_calici_1
Chicken_HQ010042_Bavaria04V0021_DEU_2004
Chicken_MN810875_GA1479_USA_2004
Chicken_KM254170_CALICID62_KOR_2013
Chicken_KM254171_CaliciQ45_KOR_2013
Chicken_MG846434_RS–BR–15–1R–1_BRA_2015
Chicken_MG846433_RS–BR–15–6R_BRA_2015
C2
Human_JX993277_BE–HPI01_DEU_2012
Human_MF431582_GI.2_IRL_2016
0.5 aa subs/site

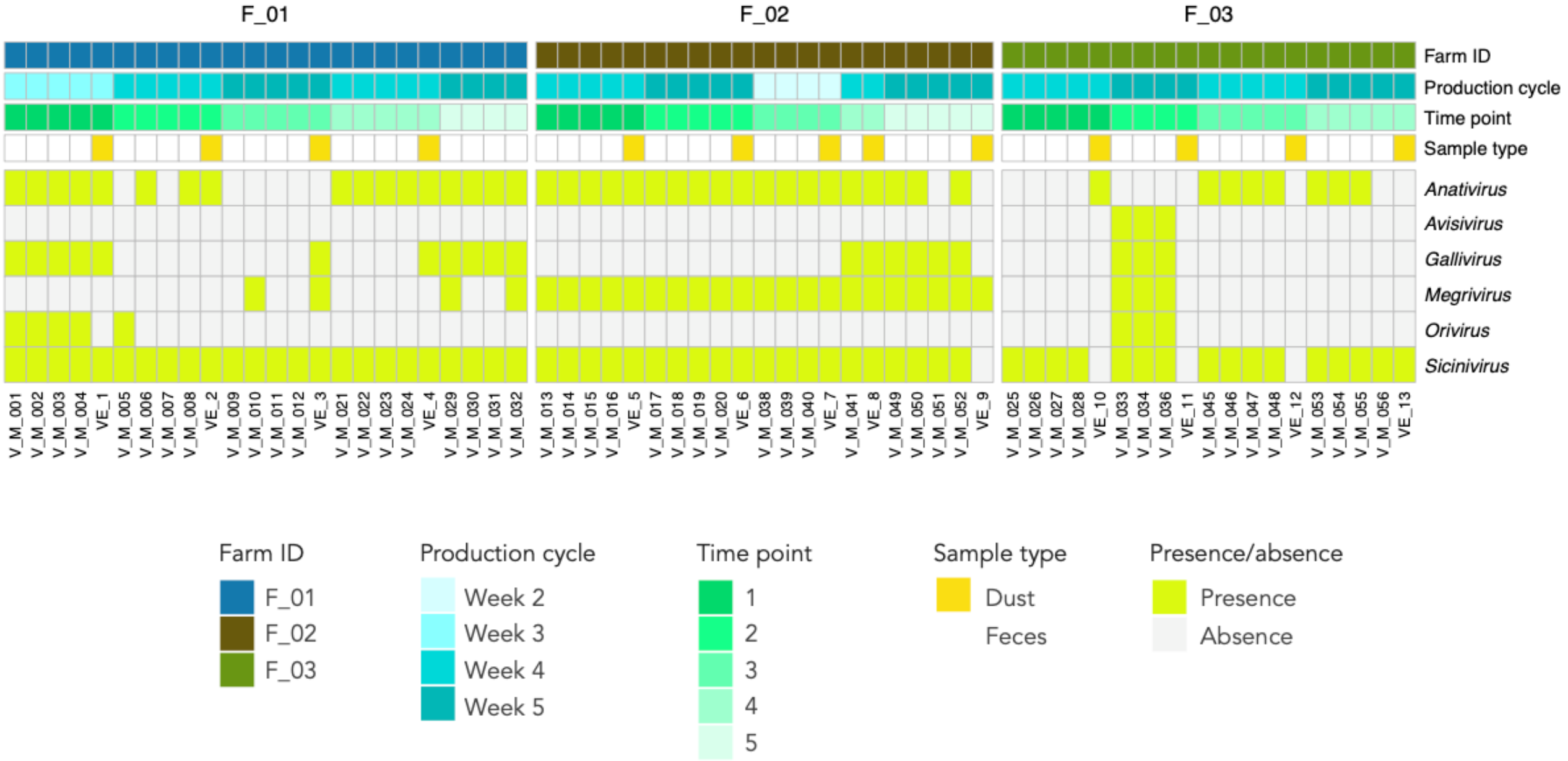
Figure 7


A
Assembled viral contigs
B
Short reads (after QC)
F_03_VE_13
F_03_VE_12
F_03_VE_11
F_03_VE_10
F_02_VE_9
F_02_VE_8
F_02_VE_7
F_02_VE_6
F_02_VE_5
F_01_VE_4
F_01_VE_3
F_01_VE_2
F_01_VE_1
Mash distance
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02

Supplementary Figure 1



Supplementary Figure 2



Supplementary Materials

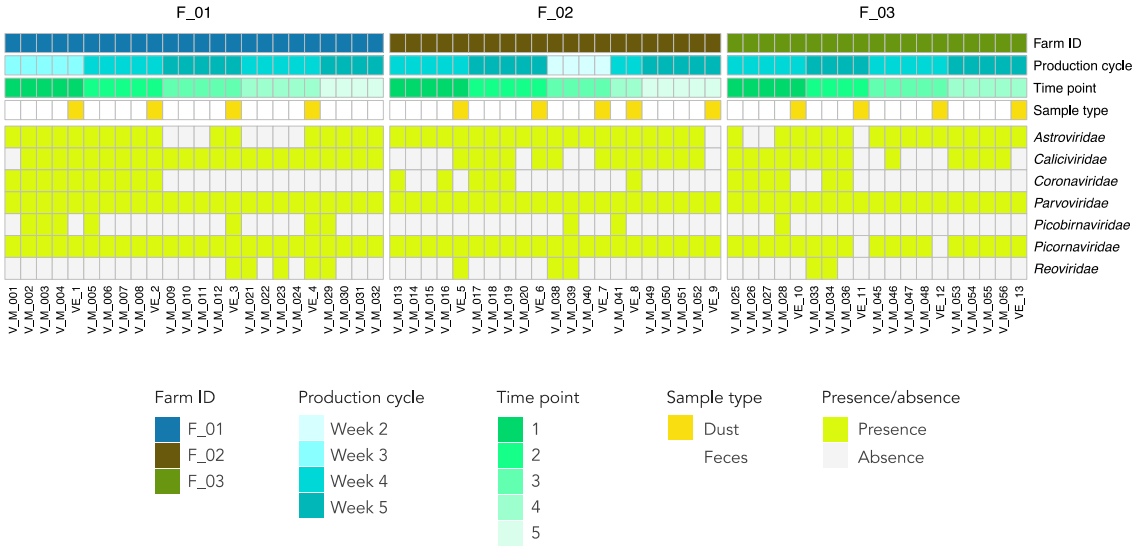


Figure S1. An overview of virus family prevalence in chicken feces and farm dust samples. Associated metadata is shown at the top panel. The color intensity of the heatmap (bottom panel) is determined by number of contigs with minimum length of 300 nt, at least 70% identity and an e-value threshold of $1x10^{-10}$ when comparing with closest reference in our database. Sample order is assorted by farm ID, sampling time point. Time point is an arbitrary number. Samples were collected at the same time and place when both Farm ID and Time point match.

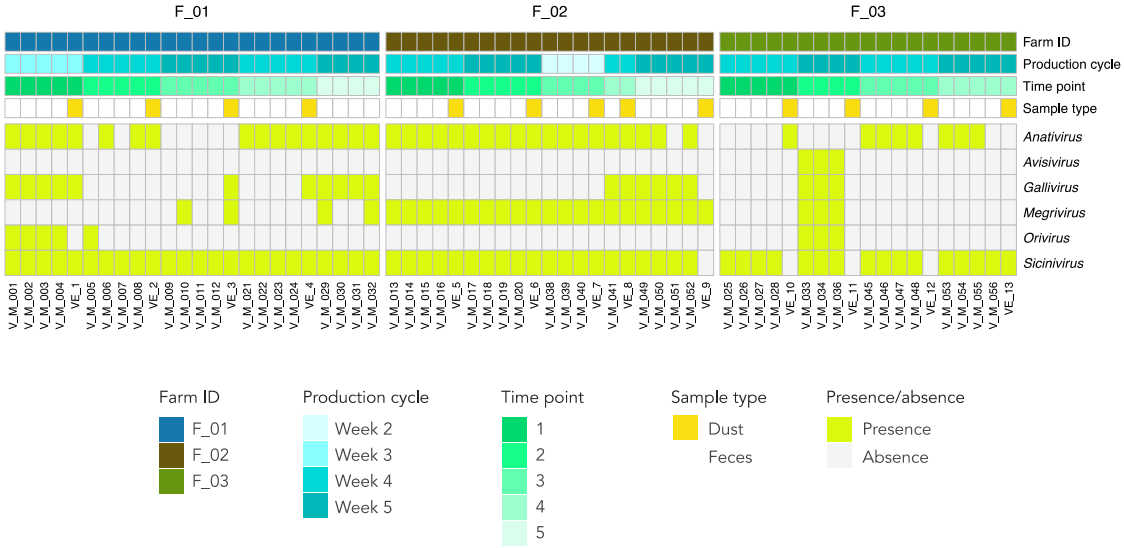


Figure S2. An overview of picornavirus genera prevalence in chicken feces and farm dust samples. Associated metadata is shown at the top panel. The color intensity of the heatmap (bottom panel) is determined by number of contigs with minimum length of 300 nt, at least 70% identity and an e-value threshold of $1x10^{-10}$ when comparing with closest reference in our database. Sample order is assorted by farm ID, sampling time point. Time point is an arbitrary number. Samples were collected at the same time and place when both Farm ID and Time point match.