

1 Article type: Short Communication (1,256)

2

3 **Soil pH influences the structure of virus communities at local and global scales**

4

5 Sungeun Lee¹, Jackson W. Sorensen², Robin L. Walker³, Joanne B. Emerson²,
6 Graeme W. Nicol¹, Christina Hazard¹

7

8 1. Environmental Microbial Genomics Group, Laboratoire Ampère, École Centrale de
9 Lyon, CNRS UMR 5005, Université de Lyon, Écully, 69134, France

10 2. Department of Plant Pathology, University of California, Davis, One Shields Avenue,
11 Davis, 95616, California, USA

12 3. School of Rural Land Use, Scotland's Rural College, Craibstone Estate, Aberdeen,
13 AB21 9YA, Scotland, United Kingdom

14

15 Corresponding author

16 Email: christina.hazard@ec-lyon.fr

17 Tel: +33 4 72 18 60 91

18

19

20

21

22

23

24

25

26

27 **Abstract**

28 The influence of biotic and abiotic factors on viral communities across environmental
29 gradients in soil is relatively unknown. While soil pH strongly influences microbial
30 community structure, it is unclear whether there is a similar influence on soil viruses.
31 In this study, prokaryotic and viral communities were characterized in soils from a long-
32 term pH-manipulated soil gradient (pH 4.5 and 7.5), and viral populations also
33 compared to those of other soils ranging in pH (4.0-7.5). Viral communities were
34 significantly influenced by pH at the local scale with 99% of viral operational taxonomic
35 units restricted to pH 4.5 or 7.5 soil only. Analysis of viromes from six other European
36 and North American soil systems demonstrated that a selection of viral clusters from
37 acidic and neutral pH soils were more associated with those from the local gradient
38 pH 4.5 or 7.5 soils, respectively. While direct pH effects on virion integrity and indirect
39 selection via host composition were not distinguished, the results reveal that soil pH
40 is a factor in structuring viral communities at local and global scales.

41

42 key words: viruses; prokaryotes; community structure; pH

43

44

45

46

47

48

49

50

51

52 **Main**

53 Viruses play a major role in controlling the abundance, structure and evolution of
54 microbial communities through cell lysis and the release of nutrients and modulation
55 of host cell metabolism (Pratama et al., 2018). In soils, viruses are diverse (Emerson
56 et al., 2018), abundant (Williamson et al., 2017), and have potential to impact nutrient
57 cycling and other ecosystem functions (Emerson et al., 2018; Kuzyakov and Mason-
58 Jones, 2018; Trubl et al., 2018). However, fundamental knowledge gaps in soil viral
59 ecology include a basic understanding of the biotic and abiotic drivers of viral
60 communities.

61 Soil bacterial and viral community dynamics co-vary, with the susceptibility of
62 hosts to infection from individual viruses varying over time (Vos et al., 2009) or viral
63 community shifts occurring as an indirect result of nutrient input altering host
64 community structure (Srinivasiah et al., 2015). While host communities are likely the
65 strongest factor for defining viral community structures, the host range of viruses may
66 also play a role in defining whether their community dynamics vary to the same extent
67 over physicochemical gradients (de Jonge et al., 2019). Closely related host
68 populations may be adapted to growth under different conditions, resulting in niche
69 differentiation and contrasting distribution across an ecological gradient. However, it
70 is unclear whether narrow or broad host ranges of their associated viruses reduce the
71 relative variation in virus community structure compared to prokaryotes. In addition,
72 changes in soil physicochemical characteristics may also directly impact the physical
73 integrity and dispersal of viruses (Trubl et al., 2016). As soil pH is a major determinant
74 of prokaryotic community composition at local and global scales (e.g. Bahram et al.,
75 2018; Bartram et al., 2014; Griffiths et al., 2011; Lammel et al., 2018; Lauber et al.,

76 2009), it is possible that virus communities may also exhibit pH-influenced community
77 structures.

78 Soil samples were taken from the extreme ends of an extensively characterized
79 contiguous soil pH gradient (pH 4.5 to 7.5) where total carbon, nitrogen and organic
80 matter do not change significantly across the gradient (Nicol et al., 2008; Bartram et
81 al., 2014). At this site, prokaryotic communities have been shown to vary in response
82 to pH (Nicol et al., 2008; Bartram et al., 2014), and distinct virus populations infecting
83 methylophilic communities associated with contrasting soil pH have recently been
84 observed (Lee et al., 2021). Both non-targeted total community metagenome and
85 virus-targeted virome libraries were prepared as previously described (Lee et al.,
86 2021), and the assembled predicted viral contigs obtained were clustered into 1,910
87 viral operational taxonomic units (vOTUs) (Table S1), and prokaryotic 16S rRNA gene
88 fragments were extracted and classified into 2,312 OTUs (Supplementary Methods).
89 Sequencing depth was sufficient to capture vOTU richness, although further
90 sequencing and sampling may have increased 16S rRNA OTU recovery (Fig. S1).

91 Of the metagenomic reads, 22.7% were taxonomically defined, with
92 Actinobacteria and Proteobacteria dominating in both pH soils (Fig. 1a). Similar to
93 other soil viral studies, only a small proportion of the prokaryotic viral community (7.0%
94 of vOTUs) was taxonomically defined. A total of 29.1% of vOTUs were linked to
95 predicted hosts, with the majority for both pH soils belonging to Actinobacteria and
96 Proteobacteria (Fig. 1a, Table S2).

97 Analysis of individual 16S rRNA OTUs and vOTUs demonstrated distinct
98 structures between pH 4.5 and 7.5 soils for both prokaryote and viral communities
99 (Fig. 1b). Specifically, 38.6% of OTUs (pH 4.5 OTUs, 263; pH 7.5 OTUs, 630) and
100 99.0% of the vOTUs (pH 4.5 vOTUs, 524; pH 7.5 vOTUs, 1,361) were found in only

101 one soil pH. Prokaryote and viral (virome) alpha-diversity were significantly greater in
102 pH 7.5 soil (Table S3). Although absorption of viruses to soil organic particles can
103 decrease with increasing pH (Sobsey et al., 1980), the effect of soil pH biasing
104 measurements of alpha-diversity may be small as organic matter content does not
105 change significantly across the gradient and a neutral pH buffer was used for
106 extracting viral particles from both soils, with no significant difference in the yields of
107 virome DNA.

108 While viromes produced 73x more assembled viral contigs than metagenomes,
109 read mapping of individual reads to all vOTUs demonstrated that both approaches
110 produced distinct viral community profiles between soils (Fig. 1b). Decreasing the
111 breadth (length of contig covered by mapped reads) threshold (<75%) for defining
112 vOTU detection disproportionately increased vOTU detection in metagenomes
113 compared to viromes (Fig. S2). This suggests that the appropriate breadth thresholds
114 for detection may be different in viromes compared to total metagenomes, with a cut-
115 off $\geq 75\%$ potentially too conservative for the total metagenomes. However, care
116 should be taken before reducing breadth threshold in other datasets that do not have
117 paired viromes to corroborate viral detection.

118 Soil viromes from the gradient soils were compared with those from six other
119 ecosystems varying in pH, soil type, land use and location and where viral contigs
120 were predicted using the same tools and standards (Table S4) (Emerson et al., 2018;
121 ter Horst et al., 2021; Santos-Medellin et al., 2021). Using gene-sharing network
122 analysis (Jang et al., 2019), the number of clusters containing vOTUs from these
123 additional soils and the local gradient (pH 4.5 only, pH 7.5 only or both) were
124 determined (Fig. 2, Table S5). On average, 31% of clusters were shared (range 1-
125 62%) in pairwise comparisons between all soils (Fig. S3). A clear trend was also

126 observed in comparison with the gradient soils (Fig. 2b) with shared viral clusters in
127 acidic and neutral pH soils more associated with those from the local gradient pH 4.5
128 or 7.5 soils, respectively. While this demonstrates that certain viral genes were
129 restricted to acidic vs neutral pH soils, no candidate genes were identified as
130 potentially conferring host adaptation to a particular pH range, with the majority
131 encoding uncharacterized hypothetical proteins. In addition, the relative proportion of
132 virus-encoded putative auxiliary metabolic genes for different metabolic functions was
133 similar between soils with no trends associated with pH (Fig. S4).

134 If individual viruses can infect multiple host populations at different soil pH,
135 there would be a potential for virus community structures to be less distinct over an
136 ecological gradient. However, analysis of samples from a continuous pH gradient
137 demonstrated that contrasting soil pH results in the selection for virus community
138 structures that are at least as distinct as prokaryote host community structures. While
139 a relatively limited number of sites for analysis of viromes in different soil systems and
140 that these analyses do not separate direct effects of pH on virion integrity from indirect
141 effects of host composition, they demonstrate that, as with prokaryote communities,
142 soil pH correlates with distinct patterns of virus community structures.

143

144 **Data availability**

145 Metagenome sequence reads are deposited in NCBI's GenBank under BioProject
146 accession nos. PRJNA621436–PRJNA621447. Metagenome draft assemblies are
147 accessible through the JGI Genome Portal (DOI: 10.25585/ 1487501). Assembled
148 metagenome-derived 16S rRNA gene sequences are available at [ftp://ftp-adn.ec-](ftp://ftp-adn.ec-lyon.fr/)
149 [lyon.fr/](ftp://ftp-adn.ec-lyon.fr/). Metagenome sequence reads from 'Agriculture 2' site are available through
150 NCBI BioProject PRJNA767554.

151 **Acknowledgments**

152 This work was funded by an AXA Research Chair awarded to GWN. The sequencing
153 data were generated under JGI Community Science Program proposal 503702
154 awarded to GWN and CH. The sequencing data from the soil pH gradient was
155 produced by the U.S. Department of Energy Joint Genome Institute, a DOE Office of
156 Science User Facility, which is supported by the Office of Science of the U.S.
157 Department of Energy under Contract No. DE-AC02-05CH11231. The pH gradient
158 experiment is funded through the Scottish Government RESAS 2016-2021 program.
159 The authors would like to thank Dr. Laurent Pouilloux for assistance with the Newton
160 high performance computing cluster at École Centrale de Lyon. 'Agriculture 2' site data
161 collection and analysis was predominantly funded by the UC Davis College of
162 Agricultural and Environmental Sciences and Department of Plant Pathology (new lab
163 start-up to JBE) and also in part by USDA National Institute of Food and Agriculture
164 (NIFA) Hatch project number CA-D-PPA-2464-H and USDA NIFA grant number 2021-
165 67013-34815-0 to JBE. The authors thank Dr. Laura Zinke for collecting and
166 processing the 'Agriculture 2' site samples.

167

168 **Competing Interests**

169 The authors declare no competing interests.

170

171 **References**

172 Al-Shayeb, B., Sachdeva, R., Chen, L.X., Ward, F., Munk, P., Devoto, A., Castelle,
173 C.J., Olm, M.R., Bouma-Gregson, Amano, Y., He, C., Méheust, R., Brooks, B.,
174 Thomsa, A., Lavy, A., Matheus-Carnevali, P., Sun, C., Goltsman, D.S.A., Borton, M.A.,
175 Sharrar, A., Jaffe, A.L., Nelson, T.C., Kantor, R., Keren, R., Lane, K.R., Farag, I.F.,

176 Lei, S., Finstad, K., Amundson, R., Anatharaman, K., Zhou, J., Probst, A.J., Power,
177 M.E., Tringe, S.G., Li, W-J., Wrighton, K., Harrison, S., Morowitz, M., Relman, D.A.,
178 Doudna, J.A., Lehours, A-C., Warren, L., Cate, J.H.D., Satini, J.M., Banfield, J.F.,
179 2020. Clades of huge phages from across Earth's ecosystems. *Nature* 578, 425-431.
180

181 Bahram, M., Hildebrand, F., Forslund, S.K., Anderson, J.L., Soudzilovskaia, H.A.,
182 Bodegom, P.M., Bengtsson-Palme, J., Anslan, S., Coelho, L.P., Harend, H, Huerta-
183 Cepas, J., Medema, M.H., Maltz, M.R., Mundra, S., Olsson, P.A., Pent, M., Pöhlme, S.,
184 Sunagawa, S., Ryberg, M., Tendersoo, L., Bork, P., 2018. Structure and function of
185 the global topsoil microbiome. *Nature* 560, 233-237.
186

187 Bartram, A.K., Jiang, X., Lynch, M.D., Masella, A.P., Nicol, G.W., Dushoff, J., Neufeld,
188 J.D., 2014. Exploring links between pH and bacterial community composition in soils
189 from the Craibstone Experimental Farm. *FEMS Microbiology Ecology* 87, 403-415.
190

191 Emerson, J.B., Roux, S., Brum, J.R., Bolduc, B., Woodcroft, B.J., Jang, H.B.,
192 Singleton, C.M., Solden, L.M., Naas, A.E., Boyd J.A., Hodgkins, S.B., Wilson, R.M.,
193 Trubl, G., Li, C., Frolking, S., Pope, P.B., Wrighton, K.C., Crill, P.M., Chanton, J.P.,
194 Saleska, S.R., Tyson G.W., Rich, V.I., Sullivan, M.B., 2018. Host-linked soil viral
195 ecology along a permafrost thaw gradient. *Nature Microbiology* 3, 870-880.
196

197 Griffiths, R.I., Thomson, B.C., James, P., Bell, T., Bailey, M., Whiteley, A.S., 2011. The
198 bacterial biogeography of British soils. *Environmental Microbiology* 13, 1642-1654.
199

200 ter Horst, A.M., Santos-Medellin, C., Sorensen, J.W., Zinke, L.A., Wilson, R.M.,
201 Johnston, E.R., Truble, G.G., Pett-Ridge, J., Blazewicz, S., Hanson, P.J., Chanton, P.,
202 Schadt, C.W., Kostka, J.E., Emerson, J.B., 2021. Minnesota peat viromes reveal
203 terrestrial and aquatic niche partitioning for local and global viral populations.
204 *Microbiome* 9, 233.

205

206 Jang, H.B., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister,
207 J.R., Kropinski, A.M., Krupovic, M., Lavigne, R., Turner, D., Sullivan, M.B., 2019.
208 Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-
209 sharing networks. *Nature Biotechnology* 37, 632-639.

210

211 de Jonge, P.A., Nobrega, F.L., Brouns, S.J.J., Dutilh, B.E, 2019. Molecular and
212 evolutionary determinants of bacteriophage host range. *Trends in Microbiology* 27, 51-
213 63.

214

215 Kuzyakov, Y., Mason-Jones, K., 2018. Viruses in soil: Nano-scale undead drivers of
216 microbial life, biogeochemical turnover and ecosystem functions. *Soil Biology and*
217 *Biochemistry* 127, 305-317.

218

219 Lammel, D.R., Barth, G., Ovaskainen, O., Cruz, L.M., Zanatta, J.A., Ryo, M., de
220 Souza, E.M., Pedrosa, F.O., 2018. Direct and indirect effects of a pH gradient bring
221 insights into the mechanisms driving prokaryotic community structures. *Microbiome* 6,
222 106.

223

224 Lauber, C.L., Hamady, M., Knight, R., Fierer, N., 2009. Pyrosequencing-based
225 assessment of soil pH as a predictor of soil bacterial community structure at the
226 continental scale. *Applied and Environmental Microbiology* 75, 5111-5120.
227

228 Lee, S., Sieradzki, E.T., Nicolas, A.M., Walker, R.L., Firestone, M.K., Hazard, C.,
229 Nicol, G.W., 2021. Methane-derived carbon flows into host-virus networks at different
230 trophic levels in soil. *Proceedings of the National Academy of Sciences of the United*
231 *States of America* 10, e2105124118.
232

233 Nicol, G.W., Leininger, S., Schleper, C., Prosser, J.I., 2008. The influence of soil pH
234 on the diversity, abundance and transcriptional activity of ammonia oxidizing archaea
235 and bacteria. *Environmental Microbiology* 10, 2966-2978.
236

237 Pratama, A.A., van Elsas, J.D., 2018. The 'neglected' soil virome – potential role and
238 impact. *Trends in Microbiology* 26, 649–662.
239

240 Santos-Medellin, C., Zinke, L.A., ter Horst, A.M., Gelardi, D.L., Parikh, S.J., Emerson,
241 J.B., 2021. Viromes outperform total metagenomes in revealing the spatiotemporal
242 patterns of agricultural soil viral communities. *ISME Journal* 15, 1956-1970.
243

244 Sobsey, M.D., Dean, C.H., Knuckles, M.E., Wagner, R.A., 1980. Interactions and
245 survival of enteric viruses in soil material. *Applied and Environmental Microbiology* 40,
246 92-101.
247

248 Srinivasiah, S., Lovett, J., Ghosh, D., Roy, K., Fuhrmann, J.J., Radosevich, M.,
249 Wommack, K.E., 2015. Dynamics of autochthonous soil viral communities parallels
250 dynamics of host communities under nutrient stimulation. *FEMS Microbiology Ecology*
251 91, fiv063.

252

253 Trubl, G., Solonenko, N., Chittick, L., Solonenko, S.A., Rich, VI, Sullivan, M.B., 2016.
254 Optimization of viral suspension methods for carbon-rich soils along a permafrost thaw
255 gradient. *PeerJ* 17, e1999.

256

257 Trubl, G., Jang, H.B., Roux, S., Emerson, J.B., Solonenko, N., Vik, D.R., Vik, D.R.,
258 Solden, L., Ellenbogen, J., Runyon, A.T., Bolduc, B., Woodcroft, B.J., Saleska, S.R.,
259 Tyson, G.W., Wrighton, K.C., Sullivan, M.B., Rich, V.I., 2018. Soil viruses are
260 underexplored players in ecosystem carbon processing. *mSystems* 3, e000076-18.

261

262 Vos, M., Birkett, P.J., Brich, E., Griffiths, R.J., Buckling, A., 2009. Local adaptation of
263 bacteriophages to their bacterial hosts in soil. *Science* 14, 833.

264

265 Williamson, K.E., Fuhrmann, J.J., Wommack, K.E., Radosevich, M., 2017. Viruses in
266 Soil Ecosystems: An Unknown Quantity Within an Unexplored Territory. *Annual*
267 *Review of Virology* 4, 201-219.

268

269

270

271

272

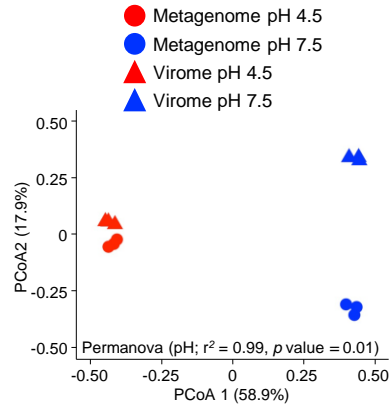
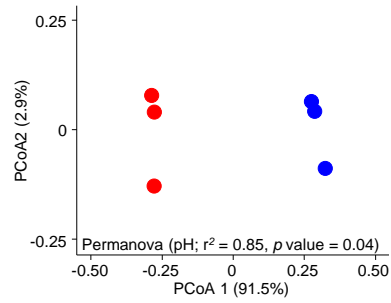
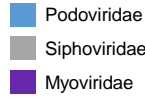
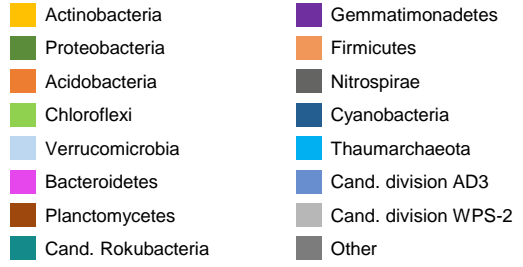
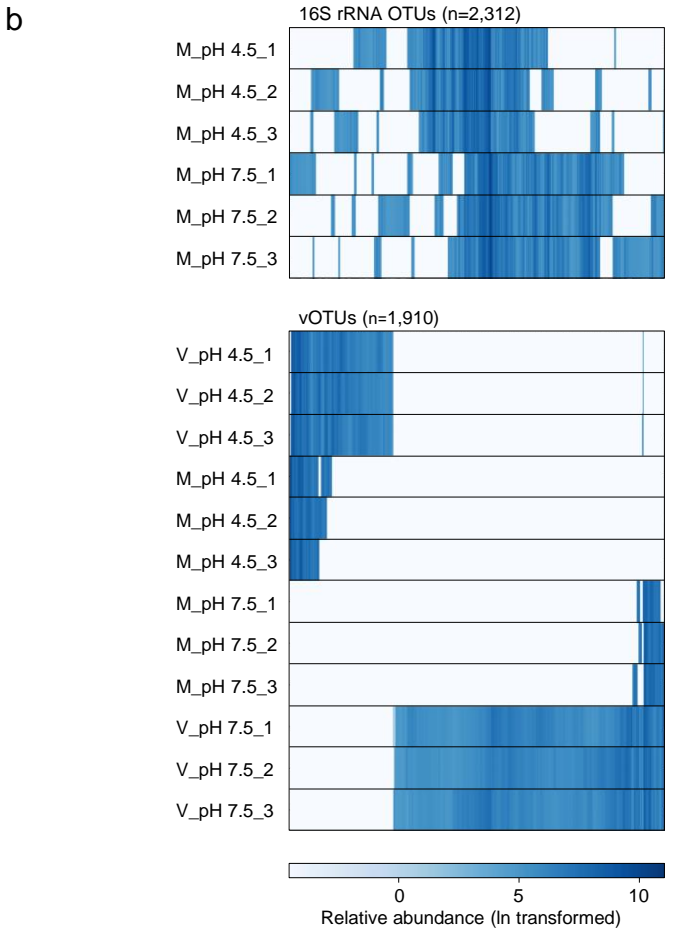
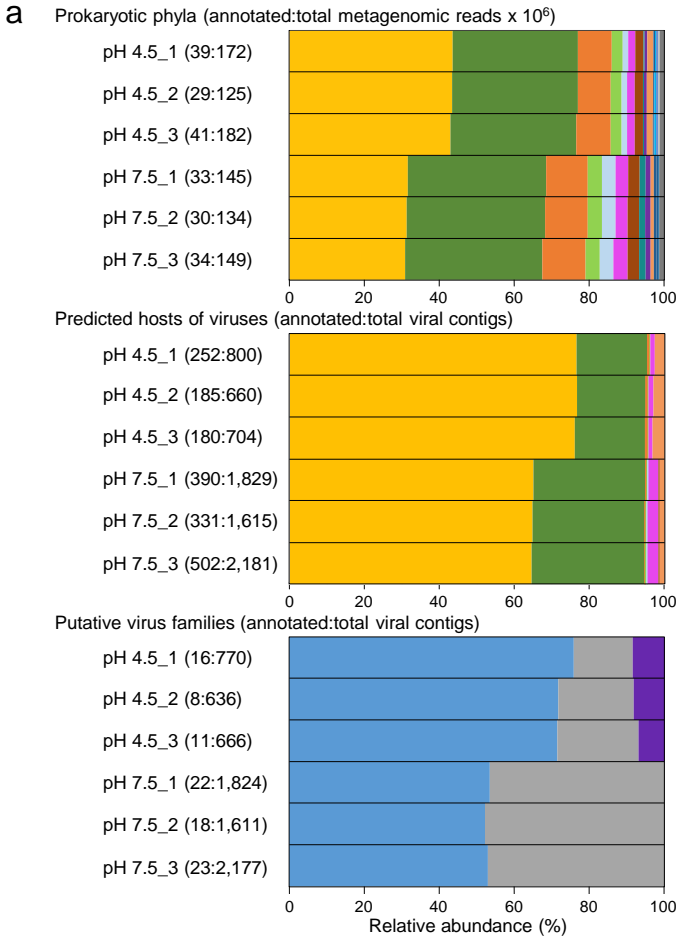
273 **Figure Legends**

274 **Fig. 1:** Taxonomic composition and community structure of prokaryotic and virus
275 communities in pH 4.5 and 7.5 replicate soil samples taken from the ends of a
276 contiguous pH gradient. a) Relative abundance of taxonomically-defined prokaryotes,
277 taxonomically-defined viruses and the predicted hosts of viruses. For prokaryotes,
278 reads from metagenomes were annotated at the phylum level using the NCBI nr
279 database. Numbers in parenthesis denote the number of mapped reads:total reads
280 analyzed. Viral contigs ≥ 10 kb were taxonomically defined at the family level based on
281 gene-sharing network analysis (Jang et al., 2019). Host prediction of viruses was
282 determined by using the gene-sharing network, and gene homology analysis (Al-
283 Shayeb et al., 2020). Numbers in parenthesis denote the number of annotated:total
284 reads or contigs for each sample, and plots display the relative proportion of annotated
285 reads only (i.e. annotated reads of prokaryotes or reads mapped to annotated viral
286 contigs). b) Normalized relative abundance of individual 16S rRNA OTUs and vOTUs
287 in soil samples determined by read-mapping. Only vOTUs where reads were mapped
288 with $\geq 1x$ coverage over 75% contig breadth were included. Ordinations show the
289 principal coordinate analysis of Bray-Curtis dissimilarities derived from relative
290 abundance tables. For virus communities, reads from both viromes (V) and
291 metagenomes (M) were analyzed for each sample. Details of all methods used are
292 provided in Supplementary Information.

293

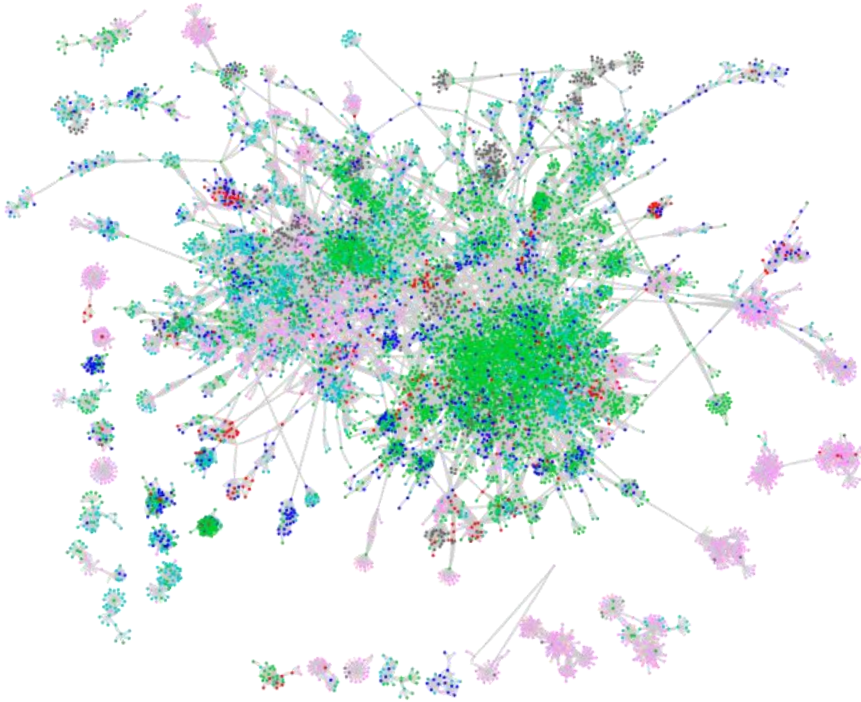
294 **Fig. 2:** Network analysis describing linkages of Gradient 4.5 and 7.5 vOTUs with six
295 sets of soils samples ranging in pH from 4.0 to 7.5 from Europe and North America
296 (Table S4). a) Gene sharing network of vOTUs showing viral clusters containing ≥ 25
297 vOTUs. b) Relative abundance of clusters from each soil that contain vOTUs shared

298 with those from gradient pH 4.5 (blue) and gradient pH 7.5 soil (red) or both (grey).
299 Numbers in parenthesis denote total number and percentage of shared viral clusters
300 with Gradient soils. Details of all methods used are provided in Supplementary
301 Information.



a

- Gradient 4.5 (this study)
- Gradient 7.5 (this study)
- Peat (ter Horst et al., 2021)
- Bog (Emerson et al., 2018)
- Palsa (Emerson et al., 2018)
- Fen (Emerson et al., 2018)
- Agriculture 1 (Santos-Medellin et al., 2021)
- Agriculture 2 (this study)
- RefSeq viruses



b

