Supplementary Information for

Ice nucleation by viruses and their potential for cloud glaciation

Michael P. Adams[1, #], Nina S. Atanasova[2, 3, #], Svetlana Sofieva[2, 3], Janne Ravantti[3], Aino Heikkinen[2 +], Zoé Brasseur[4], Jonathan Duplissy[4,5], Dennis H. Bamford[3,*] & Benjamin J. Murray[1*]

[1] - Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds, UK
[2] - Finnish Meteorological Institute, Helsinki, Finland
[3] - Molecular and Integrative Biosciences Research Programme, Faculty of Biological and Environmental Sciences, University of Helsinki, Finland
[4] - Institute for Atmospheric and Earth System Research / Physics, Faculty of Science, University of Helsinki, Finland
[5] - Helsinki Institute of Physics, University of Helsinki, Helsinki, Finland
[#] Authors contributed equally to the work
+ Now at the Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Helsinki, Finland

Correspondence to: Dennis.Bamford@helsinki.fi and B.J.Murray@leeds.ac.uk

This file includes:

Viral INA vs buffer INA

Each virus was compared to the buffer solution it was suspended in (listed in table S1) to determine if the virus INA was distinguishable from the buffer INA, shown in Figure S2. Multiple experiments were performed for each buffer solution, and have been complied into one freezing spectra as described in Polen et al. (2018).

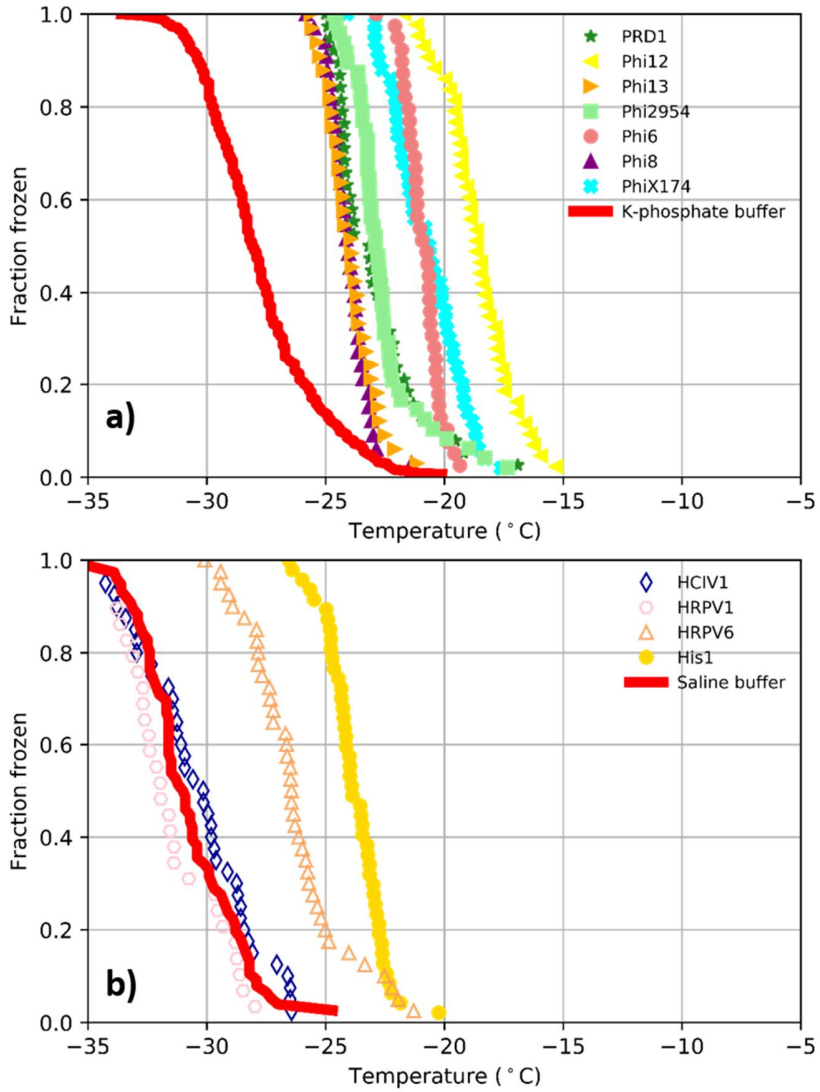

Figure S1. Fraction from curves for virus samples compared to the buffer solutions they were suspended in. A. Viruses suspended in K-phosphate buffer. B. Viruses suspended in Saline buffer. These fraction frozen curves are not adjusted for salt concentrations in the buffer solution, but both samples were suspended in the same buffer solution and would have experienced the same freezing point depression due to NaCl.

Protein profiles

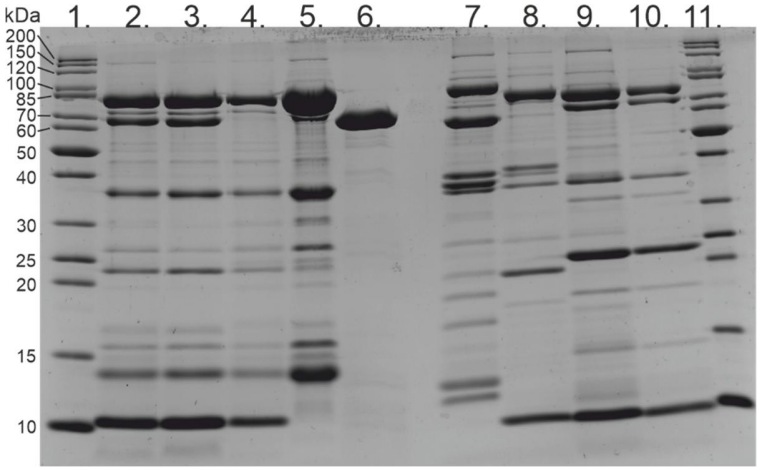

Figure S2. Subviral particles of Phi6 and protein profiles of Cystoviruses. **1.** Protein standard, sizes marked on the left side in kilodaltons (kDa). **2.** 1× purified Phi6 virus. **3.** 2× purified Phi6 virus. **4.** BHT treated Phi6. **5.** NC of Phi6. **6.** P3 protein of Phi6. **7.** Phi8 virus. **8.** Phi12 virus. **9.** Phi13 virus. **10.** Phi2954 virus. **11.** Protein standard.

Phi6 1X vs 2X purification

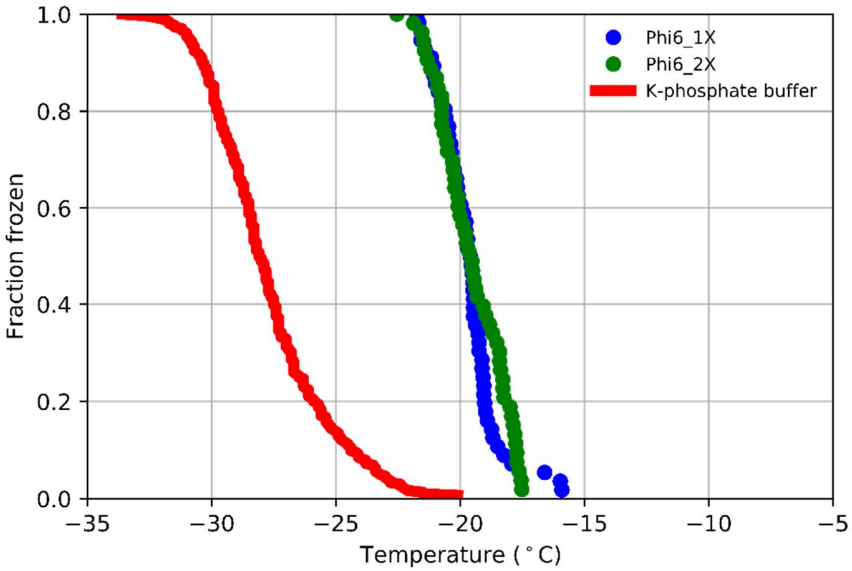

Figure S3. Fraction frozen curves showing the INA of Phi6 when purified using the 1X and 2X methods. The lack of reduction in INA when the sample is purified further implies that the INA is driven by the virus particles, not any contaminants.

Viral INA vs host INA

Each of the hosts for viruses that showed INA activity distinct from that of the buffer solution they were also tested for their INA activity. The virus and host activity of these virus are shown in Figures S4-7.

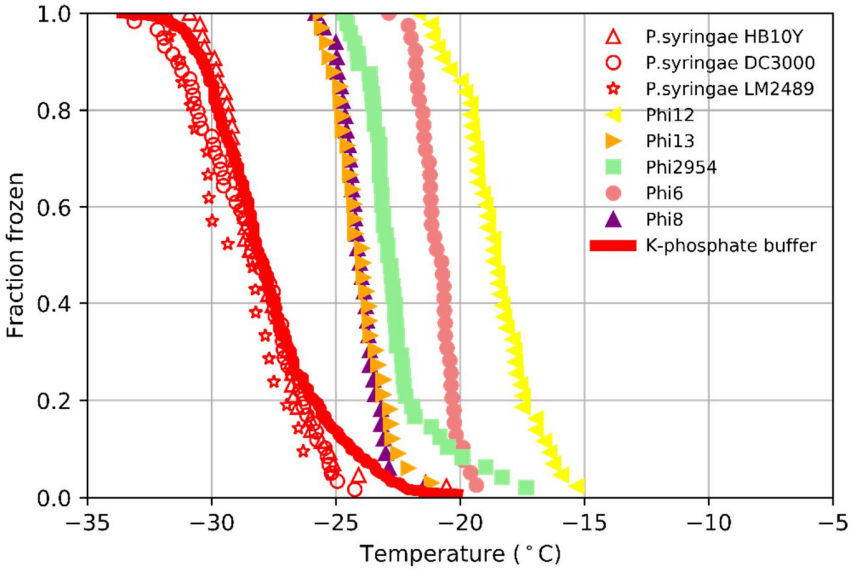

Figure S4. Fraction frozen curves of the bacterial viruses of P.syringae, the P.syringae strains used as hosts, and the K-phosphate buffer they were suspended in. The host bacteria did not give an INA signal distinguishable from the K-phosphate buffer. These fraction frozen curves are not adjusted for salt concentrations in the buffer solution, but both samples were suspended in the same buffer solution and would have experienced the same freezing point depression due to NaCl.

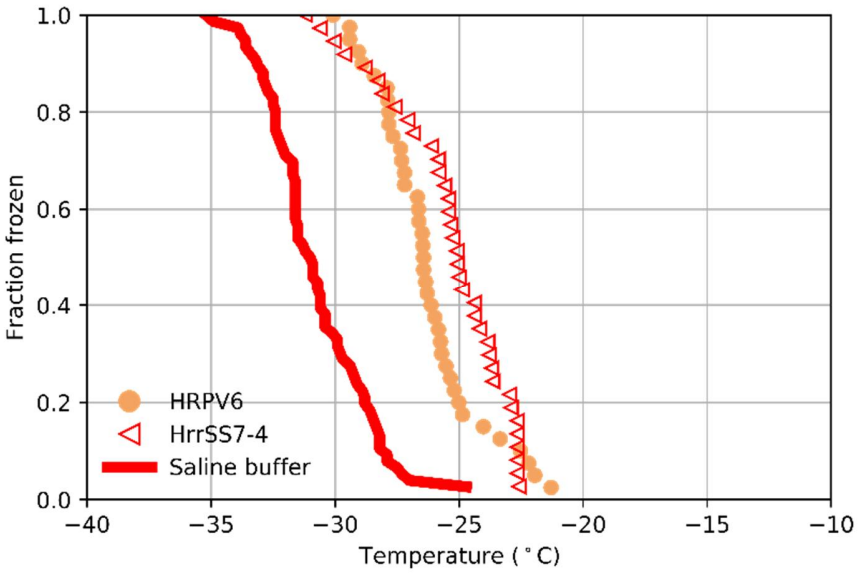

Figure S5. Fraction frozen curves of the archael virus HRPV6, its host, HrrSS7-4, and the saline buffer they were suspended in. HRPV6 did not give an INA signal distinguishable from its host. These fraction frozen curves are not adjusted for salt ***concentrations in the buffer solution, but both samples were suspended in the same buffer solution and would have experienced the same freezing point depression due to NaCl.***

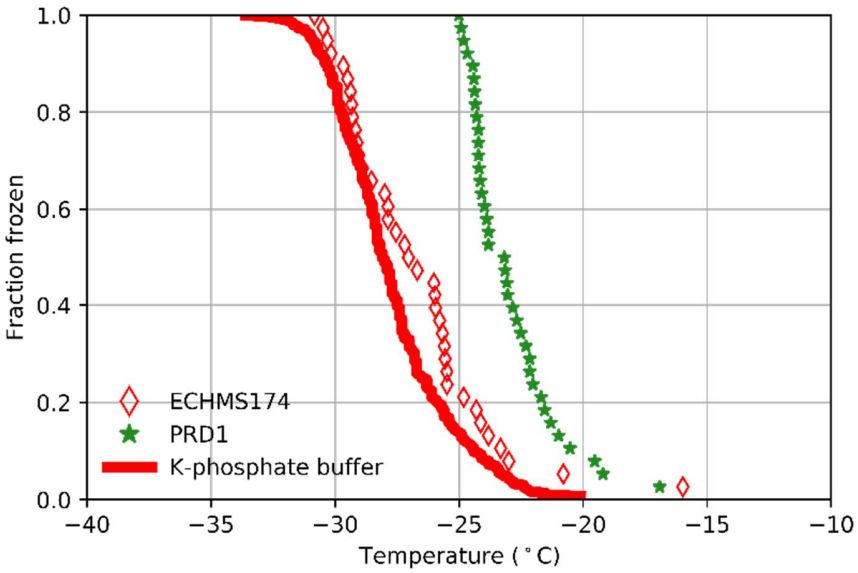

Figure S6. Fraction frozen curves of the bacterial virus PRD1, its host, ECHMS174, and the K-phosphate buffer in which they were suspended. PRD1 gave an INA distinguishable from both the K-phosphate buffer and its host. These fraction frozen curves are not adjusted for salt concentrations in the buffer solution, but both samples were suspended in the same buffer solution and would have experienced the same freezing point depression due to NaCl.

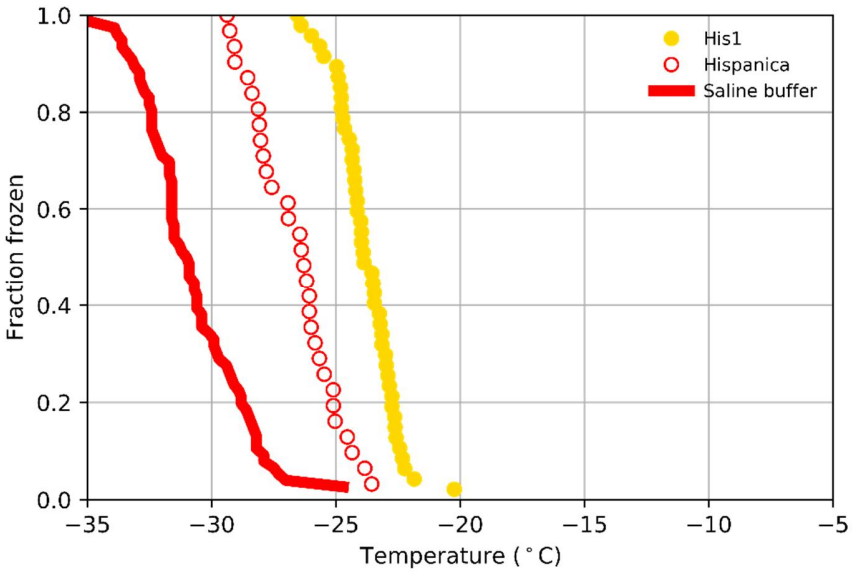

Figure S7. Fraction frozen curves of the bacterial virus His1, its host, Haloarcula hispanica, and the K-phosphate buffer they were suspended in. PRD1 gave an INA distinguishable from both the Saline buffer and its host. These fraction frozen curves are not adjusted for salt concentrations in the buffer solution, but both samples were suspended in the same buffer solution and would have experienced the same freezing point depression due to NaCl.

Genetic sequencing

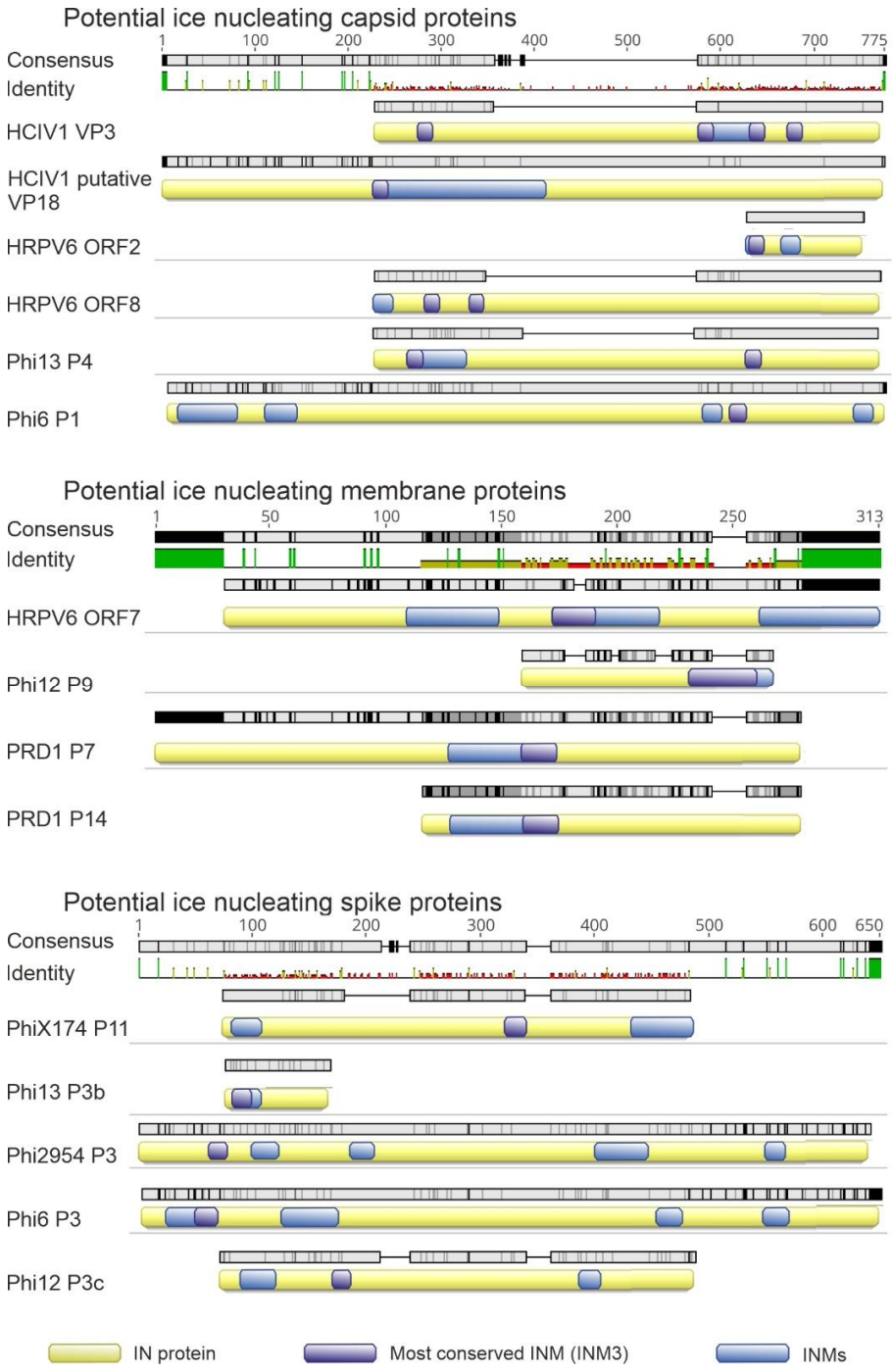

Figure S8. Multiple sequence alignment of potential IN proteins in the tested INA viruses. The predicted INMs shown in blue and the most conserved motif marked in violet. Multiple sequence alignment was done using Muscle program in Geneious Prime.

Viruses studied in this paper information

Table S1. Viruses and virus hosts used in this study.

Virus	Virus morphology	Virus host	Host domain	Virus origin	Virus buffer	Reference
Phi6	Enveloped; icosahedral	*Pseudomonas syringae*[d] HB10Y	*Bacteria*	Bacteria-infected bean, USA	K-phosphate buffer[a]	Vidaver et al., 1973
Phi6		*P. syringae* DC3000	*Bacteria*			
Phi8	Enveloped; icosahedral	*P. syringae*[d] LM2489	*Bacteria*	Bacteria-infected pea, USA	K-phosphate buffer	Mindich et al., 1999
Phi12	Enveloped; icosahedral	*P. syringae*[d] LM2489	*Bacteria*	Bacteria-infected bacil, USA	K-phosphate buffer	Mindich et al., 1999
Phi13	Enveloped; icosahedral	*P. syringae*[d] LM2489	*Bacteria*	Bacteria-infected radish, USA	K-phosphate buffer	Mindich et al., 1999
Phi2954	Enveloped; icosahedral	*P. syringae*[d] HB10Y	*Bacteria*	Bacteria-infected radish, USA	K-phosphate buffer	Qiao et al., 2010.
PRD1	Icosahedral with inner membrane	*Salmonella enterica* DS88	*Bacteria*	Sewage water, USA	K-phosphate buffer	Caldentey et al., 1990.
PRD1		*Escherichia coli* HMS174	*Bacteria*			
PhiX174	Icosahedral	*E. coli* C122	*Bacteria*	Human samples, Paris, France	Tris-HCl buffer[b]	McKenna et al., 1992
His1	Lemon-shaped	*Haloarcula hispanica*	*Archaea*	Salt lake water, Australia	Saline buffer[c]	Bath et al., 2006; Pietilä et al., 2013
HRPV-1	Pleomorphic	*Halorubrum* sp. PV6	*Archaea*	Saltern water, Italy	Saline buffer	Pietilä et al., 2009
HRPV-6	Pleomorphic	*Halorubrum* sp. SS7-4	*Archaea*	Salt crystals, Thailand	Saline buffer	Pietilä et al., 2012
HCIV-1	Icosahedral with inner membrane	*Haloarcula californiae*	*Archaea*	Saltern water, Italy	Saline buffer	Demina et al., 2016

[a] K-phosphate buffer contains: 20 mM K-phosphate pH 7.2, 1 mM $MgCl_2$

[b] Tris-HCl buffer contains: 50 mM Tris-HCl pH 7.2, 100 mM NaCl

[c] Saline buffer contains: 20 mM K-phosphate pH 7.2, 500 mM NaCl, 1 mM $MgCl_2$

[d] The classification was updated, also known as *Pseudomonas savastanoi*

Ice nucleation motifs and the MEME database

Table S2. A list of known INMs as generalized nucleotide sequences from SPRINT database in IUPAC codes used for MEME searches.

INM	Sequence
Motif 1	GAYCAYKGNGGNHTNRTHTGGCCNNYNNBNGGNHYNGTNGARWSNMRNTWYTGG
Motif 2	YTNWSNNYNMAYGCNGAYGCNMRNHGNRWNGTNKSNGARGTNRMNRYNGVNGANHKNHTN
Motif 3	YTNACNRCNGGNTAYGGNWSNACNHSNACNGCNGGNGCNGAYWSN
Motif 4	TAYYTNACNGCNGGNGAYMGNWSNAARYTNACNGCNGGNVAYGAYWSNRYNYTNATGGCNGGNGAY
Motif 5	YTNATHTTYMGNYKNTGGGAYGGNRARMGNTAYMSNMANBTNGTNGYNMRNACNGGN
Motif 6	GRNRTHGARDSNGAYRTNCCNTAYYANRTNRAYGANGANDVNRAYNTNBTNRWNAARSCN

Genbank IDs of viral genomes

Table S3. Genbank IDs of the analyzed viral genomes. For Cystoviruses, DNA fragments are specified in brackets (S/M/L).

Virus	Genbank ID
PhiX174	NC_001422.1
HRPV6	NC_017089.1
HRPV1	NC_012558.1
His1	NC_013758.1
PRD1	NC_001421.2
Phi2954	NC_012091.2 (L), NC_012092.1 (M), NC_012093.1 (S)
Phi13	NC_004172.1 (L), NC_004171.1 (M), NC_004170.1 (S)
Phi12	NC_004173.1 (L), NC_004175.1 (M), NC_004174.1 (S)
Phi8	NC_003299.1 (L), NC_003300.1 (M), NC_003301.1 (S)
Phi6	NC_003715.1 (L), NC_003716.1 (M), NC_003714.1 (S)
HCIV1	NC_030848.1

References for data shown in Figure 4

Table S4. A list of studies from which data used in the creation of the field measurement envelopes for Figure 4 was obtained.

Study	Environment	Location
Price et al., 2018	Terrestrial influenced	Eastern tropical Atlantic in African dust plumes
O'Sullivan et al., 2018	Terrestrial	UK (rural site)
Ardon-dyer, Levin., 2014	Terrestrial	Israel
Petters and Wright, 2015	Terrestrial	Terrestrial mid-latitudes
McCluskey et al., 2018	Marine	Ireland (Macehead)
McCluskey et al., 2018	Marine	Southern Ocean

Ice nucleation protein information

Table S5. Ice nucleation proteins, coding genes and references.

Protein name	Gene name	Organism	Length (aa)	Reference
Ice nucleation protein	*inaV*	*Pseudomonas syringae*	1196	Schmid, Daniel, Pridmore, David, Capitani, Guido, Battistutta, Roberto, Neeser, Jean-Richard and Jann, Alfred(1997), Molecular organisation of the ice nucleation protein InaV from Pseudomonas syringae , FEBS Letters, 414, doi: 10.1016/S0014-5793(97)01079-X
Ice nucleation protein	*inaZ*	*Pseudomonas syringae* pv. syringae	1200	Green, R., Warren, G. Physical and functional repetition in a bacterial ice nucleation gene. Nature 317, 645–648 (1985). https://doi.org/10.1038/317645a0
Ice nucleation protein InaU	*inaU*	*Pantoea ananas (Erwinia uredovora)*	1034	Yasuyuki Michigami, Satoshi Watabe, Keiko Abe, Hitoshi Obata & Soichi Arai (1994) Cloning and Sequencing of an Ice Nucleation Active Gene of Erwinia uredovora, Bioscience, Biotechnology, and Biochemistry, 58:4, 762-764, DOI: 10.1271/bbb.58.762
Ice nucleation protein	*iceE*	*Enterobacter agglomerans (Erwinia herbicola) (Pantoea agglomerans)*	1258	Gareth Warren, Loren Corotto. The consensus sequence of ice nucleation proteins from Erwinia herbicola, Pseudomonas fluorescens and Pseudomonas syringae, Gene, Volume 85, Issue 1, 1989, Pages 239-242, ISSN 0378-1119, https://doi.org/10.1016/0378-1119(89)90488-5.
Ice nucleation protein	*inaK*	*Pseudomonas syringae*	1148	Directly submitted to the EMBL/GenBank/DDBJ databases July 1997; UniProtKB: locus ICEK_PSESX, accession O30611.
Ice nucleation protein	*inaW*	*Pseudomonas fluorescens*	1210	Warren G, Corotto L, Wolber P. Conserved repeats in diverged ice nucleation structural genes from two species of Pseudomonas. Nucleic Acids Res. 1986 Oct 24;14(20):8047-60. doi: 10.1093/nar/14.20.8047. PMID: 3774551; PMCID: PMC311833.
Ice nucleation protein InaA	*inaA*	*Pantoea ananas (Erwinia uredovora)*	1322	Abe, Keiko, Watabe, Satoshi, Emori, Yasufumi, Watanabe, Michiko and Arai, Soichi (1989), An ice nucleation active gene of Erwinia ananas , FEBS Letters, 258, doi: 10.1016/0014-5793(89)81678-3
Ice nucleation protein	*inaX*	*Xanthomonas campestris* pv. translucens	1567	Zhao, J., Orser, C.S. Conserved repetition in the ice nucleation gene inaX from Xanthomonas campestris pv. translucens . Mol Gen Genet 223, 163–166 (1990). https://doi.org/10.1007/BF00315811

SI References

Polen, M., Brubaker, T., Somers, J., & Sullivan, R. C. (2018). Cleaning up our water: reducing interferences from nonhomogeneous freezing of “pure” water in droplet freezing assays of ice-nucleating particles. *Atmos. Meas. Tech*, *11*, 5315–5334. https://doi.org/10.5194/amt-11-5315-2018