

Supplementary Information for

Glycoblotting-Based Ovo-Sulphoglycomics Reveals Phosphorylated *N*-Glycans as a Possible Host-Factor of AIV Prevalence in Waterfowls

Bryan M. Montalban[†] and Hiroshi Hinou^{*,†,‡}

[†] Laboratory of Advanced Chemical Biology, Graduate School of Life Science, Hokkaido University, Sapporo 001-0021, Japan

[‡] Frontier Research Center for Advanced Material and Life Science, Faculty of Advanced Life Science, Hokkaido University, Sapporo 001-0021, Japan

*Corresponding author: Hiroshi Hinou, Email: hinou@sci.hokudai.ac.jp

Hiroshi Hinou

Frontier Research Center for Advanced Material and Life Science, Faculty of Advanced Life Science, Hokkaido University, Sapporo 001-0021, Japan

Email: hinou@sci.hokudai.ac.jp

Abbreviation:

BOA	benzyloxyamine
DABP	3,4-diaminobenzophenone
MTT	3-methyl-1- <i>p</i> -tolyltriazene
LVP	low viral prevalence
HVP	high viral prevalence
VIP	variable importance to projection

Keywords: glycoblotting, ovo-sulphoglycomics, on-bead esterification, MTT, virus prevalence.

Table S1. The list of egg whites from various species of Order Anseriformes (4 families, 27 genera, 72 species) used in this study. Classification was based on Sibley's DNA-DNA hybridization.

Sample ID	Scientific name	Family	Sub-Family	Common Name
D1	<i>Aix galericulata</i>	Anatidae	Anatinae	Mandarin Duck
D2	<i>Anas platyrhynchos</i>	Anatidae	Anatinae	Mallard Duck
D3	<i>Lophodytes cucullatus</i>	Anatidae	Anatinae	Hooded Merganser
D4	<i>Aythya americana</i>	Anatidae	Anatinae	Red head
D5	<i>Anas versicolor</i>	Anatidae	Anatinae	Silver Teal
D6	<i>Anser anser</i>	Anatidae	Anserinae	Graylag Goose
D7	<i>Anser indicus</i>	Anatidae	Anserinae	Bar Headed Goose
D8	<i>Dendrocygna eytoni</i>	Dendrocygnidae	Dendrocygnae	Plumed Whistling Duck
D9	<i>Tadorna radjah</i>	Anatidae	Anserinae	White Headed Shelduck
D10	<i>Sarkidiornis melanotos</i>	Anatidae	Anserinae	Knob-billed Goose
D11	<i>Anseranas semipalmata</i>	Anseranatidae	Anseranatidae	Maggie Goose
D12	<i>Aix sponsa</i>	Anatidae	Anatinae	Wood Duck
D13	<i>Alopochen aegyptiaca</i>	Anatidae	Anserinae	Egyptian Goose
D14	<i>Anas platyrhynchos domesticus</i>	Anatidae	Anatinae	White Call Duck
D15	<i>Anser anser domesticus (America)</i>	Anatidae	Anserinae	Buff Goose
D16	<i>Anser anser domesticus (France)</i>	Anatidae	Anserinae	Dewlap Toulouse Goose
D17	<i>Anser anser domesticus (Germany)</i>	Anatidae	Anserinae	Emdens Goose
D18	<i>Anser canagicus</i>	Anatidae	Anserinae	Emperor Goose
D19	<i>Anser cygnoides domesticus</i>	Anatidae	Anserinae	African Goose
D20	<i>Anser cygnoides domesticus</i>	Anatidae	Anserinae	White China Goose
D21	<i>Anser cygnoides domesticus</i>	Anatidae	Anserinae	Brown China Goose
D22	<i>Anser erythropus</i>	Anatidae	Anserinae	Lesser white-fronted Goose
D23	<i>Callonetta leucophrys</i>	Anatidae	Anatinae	Ringed Teal
D24	<i>Dendrocygna arborea</i>	Dendrocygnidae	Dendrocygnae	West Indian Whistling Duck
D25	<i>Lophonetta specularioides</i>	Anatidae	Anatinae	Crested Duck
D26	<i>Netta rufina</i>	Anatidae	Anatinae	Red Crested Pochard
D27	<i>Oxyura jamaicensis</i>	Anatidae	Oxyurinae	Ruddy duck
D28	<i>Chloephaga picta picta</i>	Anatidae	Anserinae	Magellan Goose
D29	<i>Branta leucopsis</i>	Anatidae	Anserinae	Barnacle Goose
D30	<i>Branta sandvicensis</i>	Anatidae	Anserinae	Hawaiian Goose
D31	<i>Anas gibberifrons</i>	Anatidae	Anatinae	Indonesian Teal
D32	<i>Anas laysanensis</i>	Anatidae	Anatinae	Laysan Duck
D33	<i>Anas luzonica</i>	Anatidae	Anatinae	Philippine Duck
D34	<i>Anas rubripes</i>	Anatidae	Anatinae	American Black Duck
D35	<i>Anas clypeata</i>	Anatidae	Anatinae	Northern Shoveler
D36	<i>Oxyura vittata</i>	Anatidae	Oxyurinae	Lake Duck
D37	<i>Anas melleri</i>	Anatidae	Anatinae	Meller's Duck
D38	<i>Oxyura australis</i>	Anatidae	Oxyurinae	Blue-billed Duck
D39	<i>Branta canadensis maxima</i>	Anatidae	Anserinae	Canada Goose
D40	<i>Dendrocygna viduata</i>	Dendrocygnidae	Dendrocygnae	White-faced Whistling Duck
D41	<i>Anser brachyrhynchus</i>	Anatidae	Anserinae	Pink-footed Goose
D42	<i>Chauna torquata</i>	Anhimidae	Anhimidae	Southern Screamer
D43	<i>Thalassornis leucotos</i>	Dendrocygnidae	Dendrocygnae	White-backed Duck
D44	<i>Tadorna tadornoides</i>	Anatidae	Anserinae	Australian Shelduck
D45	<i>Chenonetta jubata</i>	Anatidae	Anatinae	Australian Wood Duck
D46	<i>Somateria mollissima</i>	Anatidae	Anatinae	Common Eider
D47	<i>Anser albifrons</i>	Anatidae	Anserinae	Greater whitefronted Goose
D48	<i>Mergus serrator</i>	Anatidae	Anatinae	Red-breasted Merganser
D49	<i>Aythya affinis</i>	Anatidae	Anatinae	Lesser Scaup
D50	<i>Dendrocygna autumnalis</i>	Dendrocygnidae	Dendrocygnae	Black-bellied whistling Duck
D51	<i>Clangula hyemalis</i>	Anatidae	Anatinae	Longtailed Duck
D52	<i>Anas discors</i>	Anatidae	Anatinae	Blue-winged Teal
D53	<i>Oxyura punctata</i>	Anatidae	Oxyurinae	
D54	<i>Anas strepera</i>	Anatidae	Anatinae	Gadwall
D55	<i>Heteronetta atricapilla</i>	Anatidae	Anatinae	Black-headed Duck
D56	<i>Anas superciliosa</i>	Anatidae	Anatinae	Pacific black Duck
D57	<i>Anser caerulescens</i>	Anatidae	Anserinae	Snow Goose
D58	<i>Aythya fuligula</i>	Anatidae	Anatinae	Tufted Duck
D59	<i>Aythya ferina</i>	Anatidae	Anatinae	Common Pochard
D60	<i>Anas formosa</i>	Anatidae	Anatinae	Baikal Teal

D61	<i>Dendrocygna bicolor</i>	Dendrocygnidae	Dendrocygninae	Fulvous Whistling Duck
D62	<i>Anas querquedula</i>	Anatidae	Anatinae	Garganey
D63	<i>Malacorhynchus membranaceus</i>	Anatidae	Anatinae	Pink-eared Duck
D64	<i>Anas hottentota</i>	Anatidae	Anatinae	Hottentot Teal
D65	<i>Anas georgica</i>	Anatidae	Anatinae	Yellow-billed Pintail
D66	<i>Biziura lobata</i>	Anatidae	Oxyurinae	Musk Duck
D67	<i>Cygnus atratus</i>	Anatidae	Cygninae	Black swan
D68	<i>Aythya australis</i>	Anatidae	Anatinae	Hardhead
D69	<i>Amazonetta brasiliensis</i>	Anatidae	Anatinae	Brazilian Teal
D70	<i>Lophonetta cristata</i>	Anatidae	Anatinae	
D71	<i>Anas crecca</i>	Anatidae	Anatinae	Common Teal
D72	<i>Tadorna tadorna</i>	Anatidae	Anserinae	Common Shelduck

Table S2. List of 89 acidic *N*-glycans identified from the egg whites of 72 waterfowl species. Glycan structures were inferred from glycan composition based on the observed monoisotopic masses.

Glycan ID	Observed mass, m/z [M - H] ⁻	Calculated Glycoform Mass [M - BOA]	Mass Error, ppm	Monosaccharide composition	Glyconnect Database Links
A1	1094.339	972.274	-10.8	Hex3 HexNAc2 Su1	
A2	1108.327	972.284	-30.5	Hex3 HexNAc2 Pho1	Glyconnect
A3	1135.373	1013.301	-4.2	Hex2 HexNAc3 Su1	
A4	1149.392	1013.310	4.8	Hex2 HexNAc3 Pho1	
A5	1256.320	1134.327	-66.7	Hex4 HexNAc2 Su1	
A6	1270.431	1134.336	14.1	Hex4 HexNAc2 Pho1	
A7	1295.303	1159.368	-109.5	Hex2 HexNAc3 dHex1 Pho1	
A8	1297.420	1175.353	-7.5	HexNAc1 Su1 + Man3 GlcNAc2	Glyconnect
A9	1311.445	1175.363	3.7	HexNAc1 Pho1 + Man3 GlcNAc2	
A10	1338.419	1216.380	-28.2	Hex2 HexNAc4 Su1	
A11	1352.412	1216.390	-40.5	Hex2 HexNAc4 Pho1	
A12	1418.430	1296.380	-18.9	Hex2 Su1 + Man3 GlcNAc2	
A13	1432.469	1293.389	2.4	Hex2 Pho1 + Man3 GlcNAc2	Glyconnect
A14	1441.498	1305.426	-3.3	Unknown structure	
A15	1443.485	1321.411	-1.9	HexNAc1 dHex1 Su1 + Man3 GlcNAc2	
A16	1459.475	1337.406	-5.3	Hex1 HexNAc1 Su1 + Man3 GlcNAc2	
A17	1473.505	1337.416	8.1	Hex1 HexNAc1 Pho1 + Man3 GlcNAc2	
A18	1500.497	1378.433	-8.5	HexNAc2 Su1 + Man3 GlcNAc2	Glyconnect
A19	1514.530	1378.442	7.5	HexNAc2 Pho1 + Man3 GlcNAc2	
A20	1580.496	1458.432	-8.1	Hex3 Su1 + Man3 GlcNAc2	Glyconnect
A21	1589.525	1467.469	-13.1	HexNAc1 dHex2 Su1 + Man3 GlcNAc2	
A22	1594.528	1458.442	5.5	Hex3 Pho1 + Man3 GlcNAc2	Glyconnect
A23	1603.574	1467.479	11.4	HexNAc1 dHex2 Pho1 + Man3 GlcNAc2	
A24	1605.533	1483.464	-4.8	Hex1 HexNAc1 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A25	1621.533	1499.459	-1.7	Hex2 HexNAc1 Su1 + Man3 GlcNAc2	
A26	1635.542	1499.469	-2.2	Hex2 HexNAc1 Pho1 + Man3 GlcNAc2	
A27	1646.558	1524.491	-5.9	HexNAc2 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A28	1662.556	1540.486	-4.1	Hex1 HexNAc2 Su1 + Man3 GlcNAc2	Glyconnect
A29	1676.605	1540.495	19.8	Hex1 HexNAc2 Pho1 + Man3 GlcNAc2	
A30	1703.584	1581.512	-2.8	HexNAc3 Su1 + Man3 GlcNAc2	Glyconnect
A31	1717.608	1581.522	5.6	HexNAc3 Pho1 + Man3 GlcNAc2	
A32	1749.602	1613.537	-6.8	Unknown structure	
A33	1751.599	1629.522	0.1	Hex1 HexNAc1 dHex2 Su1 + Man3 GlcNAc2	
A34	1764.667	1628.502	50.0	Hex1 HexNAc1 NeuAc1 Su1 + Man3 GlcNAc2	
A35	1765.618	1629.532	5.3	Hex1 HexNAc1 dHex2 Pho1 + Man3 GlcNAc2	
A36	1767.620	1645.517	14.8	Hex2 HexNAc1 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A37	1783.585	1661.512	-2.1	Hex3 HexNAc1 Su1 + Man3 GlcNAc2	Glyconnect
A38	1792.814	1670.549	105.0	HexNAc2 dHex2 Su1 + Man3 GlcNAc2	
A39	1797.610	1661.521	6.6	Hex3 HexNAc1 Pho1 + Man3 GlcNAc2	Glyconnect
A40	1806.657	1670.558	12.5	HexNAc2 dHex2 Pho1 + Man3 GlcNAc2	
A41	1808.607	1686.544	-7.6	Hex1 HexNAc2 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A42	1824.608	1702.538	-3.7	Hex2 HexNAc2 Su1 + Man3 GlcNAc2	Glyconnect
A43	1838.661	1702.548	19.5	Hex2 HexNAc2 Pho1 + Man3 GlcNAc2	
A44	1865.620	1743.565	-11.7	Hex1 HexNAc3 Su1 + Man3 GlcNAc2	Glyconnect
A45	1879.661	1743.575	4.8	Hex1 HexNAc3 Pho1 + Man3 GlcNAc2	
A46	1888.924	1766.543	161.1	Hex4 dHex1 Su1 + Man3 GlcNAc2	
A47	1906.655	1784.592	-7.2	HexNAc4 Su1 + Man3 GlcNAc2	Glyconnect
A48	1920.682	1784.601	2.3	HexNAc4 Pho1 + Man3 GlcNAc2	
A49	1945.612	1823.565	-15.3	Hex4 HexNAc1 Su1 + Man3 GlcNAc2	
A50	1952.677	1816.616	-7.9	HexNAc2 dHex3 Pho1 + Man3 GlcNAc2	
A51	1954.866	1832.601	96.3	Hex1 HexNAc2 dHex2 Su1 + Man3 GlcNAc2	Glyconnect
A52	1959.599	1823.574	-26.5	Hex4 HexNAc1 Pho1 + Man3 GlcNAc2	Glyconnect
A53	1970.726	1848.596	26.9	Hex2 HexNAc2 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A54	1986.667	1864.591	-0.5	Hex3 HexNAc2 Su1 + Man3 GlcNAc2	
A55	2000.577	1864.601	-50.2	Hex3 HexNAc2 Pho1 + Man3 GlcNAc2	
A56	2009.653	1873.638	-30.7	HexNAc3 dHex2 Pho1 + Man3 GlcNAc2	
A57	2011.663	1889.623	-18.1	Hex1 HexNAc3 dHex1 Su1 + Man3 GlcNAc2	Glyconnect
A58	2027.682	1905.618	-6.1	Hex2 HexNAc3 Su1 + Man3 GlcNAc2	

A59	2041.713	1905.627	4.3	Hex2 HexNAc3 Pho1 + Man3 GlcNAc2	
A60	2050.704	1928.596	15.2	Hex5 dHex1 Su1 + Man3 GlcNAc2	
A61	2068.719	1946.644	-1.0	Hex1 HexNAc4 Su1 + Man3 GlcNAc2	
A62	2082.738	1946.654	3.3	Hex1 HexNAc4 Pho1 + Man3 GlcNAc2	
A63	2109.758	1987.671	5.0	HexNAc5 Su1 + Man3 GlcNAc2	Glyconnect
A64	2114.753	1978.669	3.4	Hex1 HexNAc2 dHex3 Pho1 + Man3 GlcNAc2	
A65	2129.999	1993.634	135.3	Hex2 HexNAc2 NeuAc1 Su1 + Man3 GlcNAc2	Glyconnect
A66	2146.364	2010.659	-173.2	Hex3 HexNAc2 dHex1 Pho1 + Man3 GlcNAc2	
A67	2148.660	2026.644	-28.3	Hex4 HexNAc2 Su1 + Man3 GlcNAc2	
A68	2158.155	2035.681	184.1	Hex1 HexNAc3 dHex2 Su1 + Man3 GlcNAc2	Glyconnect
A69	2171.782	2035.690	6.9	Hex1 HexNAc3 dHex2 Pho1 + Man3 GlcNAc2	
A70	2189.759	2067.671	5.1	Hex3 HexNAc3 Su1 + Man3 GlcNAc2	
A71	2203.771	2067.680	6.5	Hex3 HexNAc3 Pho1 + Man3 GlcNAc2	
A72	2212.764	2090.649	17.2	Hex6 dHex1 Su1 + Man3 GlcNAc2	
A73	2230.794	2108.697	8.8	Hex2 HexNAc4 Su1 + Man3 GlcNAc2	
A74	2244.834	2108.707	22.5	Hex2 HexNAc4 Pho1 + Man3 GlcNAc2	
A75	2271.838	2149.724	16.2	Hex1 HexNAc5 Su1 + Man3 GlcNAc2	
A76	2278.900	2156.707	51.0	Hex3 HexNAc2 dHex2 Su1 + Man3 GlcNAc2	
A77	2287.809	2165.744	-5.2	HexNAc3 dHex4 Su1 + Man3 GlcNAc2	
A78	2320.103	2197.734	126.0	Hex2 HexNAc3 dHex2 Su1 + Man3 GlcNAc2	
A79	2374.818	2252.702	16.6	Hex7 dHex1 Su1 + Man3 GlcNAc2	
A80	2392.895	2270.750	28.3	Hex3 HexNAc4 Su1 + Man3 GlcNAc2	
A81	2406.937	2270.760	41.8	Hex3 HexNAc4 Pho1 + Man3 GlcNAc2	
A82	2433.956	2311.777	42.1	Hex2 HexNAc5 Su1 + Man3 GlcNAc2	
A83	2449.933	2327.797	24.4	Hex1 HexNAc3 dHex4 Su1 + Man3 GlcNAc2	
A84	2479.957	2343.801	31.9	Hex2 HexNAc3 dHex3 Pho1 + Man3 GlcNAc2	
A85	2555.030	2432.803	58.9	Hex4 HexNAc4 Su1 + Man3 GlcNAc2	
A86	2569.031	2432.812	55.4	Hex4 HexNAc4 Pho1 + Man3 GlcNAc2	
A87	2596.066	2473.829	61.6	Hex3 HexNAc5 Su1 + Man3 GlcNAc2	
A88	2758.164	2635.882	74.5	Hex4 HexNAc5 Su1 + Man3 GlcNAc2	
A89	2920.151	2797.935	47.7	Hex5 HexNAc5 Su1 + Man3 GlcNAc2	

*Monosaccharide nomenclatures are based on the SNFG: Hexose (Hex), *N*-acetyl hexosamine (HexNAc), Mannose (Man), *N*-acetyl glucosamine (GlcNAc), Fucose (dHex), Sulfate (Su), and Phosphate (Pho). The number of units corresponding to each monosaccharide are indicated after each abbreviation.

*The links to the Glyconnect database of the Swiss Institute of Bioinformatics are provided for selected monoisotopic peaks found in the database.

*From the 89 monoisotopic masses, 55 sulfated and 34 phosphorylated *N*-glycans were identified based on their glycan composition and MS/MS analysis. Fucosylated acidic *N*-glycan structures were also found in trace abundance relative to un-fucosylated acidic *N*-glycans.

*Glycoform mass is the mass of unlabeled *N*-glycan structure denoted as [M-BOA], BOA is benzyloxyamine with a molecular mass of 123.0684 Da.

Table S3. List of acidic *N*-glycans that gave the variation between groups in PCA.

PCA Group No.	Glycan ID	Observed mass, m/z [M-H] ⁻	¹ Glycan composition	Probable structures	² trans-Gal structure type
1	A6	1270.431	Hex1 PMe1 + Core		trans-Gal(-)
	A22	1594.528	Hex3 PMe1 + Core		trans-Gal(-)
	A37	1783.585	Hex3 HexNAc1 Su1 + Core		trans-Gal(+)
	A39	1797.610	Hex3 HexNAc1 PMe1 + Core		trans-Gal(+)
2	A3	1135.373	Hex2 HexNAc3 Su1		trans-Gal(-)
	A8	1297.420	HexNAc1 Su1 + Core		trans-Gal(-)
	A16	1459.475	Hex1 HexNAc1 Su1 + Core		trans-Gal(-)
	A18	1500.497	HexNAc2 Su1 + Core		trans-Gal(-)
	A25	1621.533	Hex2 HexNAc1 Su1 + Core		trans-Gal(-)
	A28	1662.556	Hex1 HexNAc2 Su1 + Core		trans-Gal(+)
	A30	1703.584	HexNAc3 Su1 + Core		trans-Gal(-)
	A44	1865.620	Hex1 HexNAc3 Su1 + Core		trans-Gal(+)
3	A42	1824.608	Hex2 HexNAc2 Su1 + Core		trans-Gal(+)
	A47	1906.655	HexNAc4 Su1 + Core		trans-Gal(-)
	A58	2027.682	Hex2 HexNAc3 Su1 + Core		trans-Gal(+)
	A61	2068.719	Hex1 HexNAc4 Su1 + Core		trans-Gal(+)
	A70	2189.759	Hex3 HexNAc3 Su1 + Core		trans-Gal(+)
	A73	2230.794	Hex2 HexNAc4 Su1 + Core		trans-Gal(+)
	A80	2392.895	Hex3 HexNAc4 Su1 + Core		trans-Gal(+)
	A85	2555.030	Hex4 HexNAc4 Su1 + Core		trans-Gal(+)

¹Monosaccharide nomenclatures are based on the SNFG: Hexose (Hex), *N*-acetyl hexosamine (HexNAc), Sulfate (Su), methylated Phosphate (PMe) and *N*-glycan core Man3GlcNAc2 (Core). The number of units corresponding to each monosaccharide are indicated after each abbreviation.

²The definition of trans-Gal(+/-) classifications were based on Hirose, *et al.*[1]

Table S4. GenBank accession numbers for various genes of the 72 Anseriformes species in this study.

Sample ID	Scientific name	CO1	Cty <i>b</i>	ND2	Complete mtDNA
D01	<i>Aix galericulata</i>	JN703260	EU585604	EU585667	KF437906
D02	<i>Anas platyrhynchos</i>	Mk262361	EU585609	EU585672	MN720361
D03	<i>Lophodytes cucullatus</i>		EU585650	EU585713	
D04	<i>Aythya americana</i>	DQ434316	NC 000877	NC 000877	NC 000877
D05	<i>Anas versicolor</i>	FJ027121	AF059094	AF059154	
D06	<i>Anser anser</i>	GU571243	EU585613	EU585676	NC 011196
D07	<i>Anser indicus</i>	GU571246	EU585619	EU585682	NC 025654
D08	<i>Dendrocygna eytoni</i>	MZ153330	EU585647	EU585710	
D09	<i>Tadorna radjah</i>		EU585665	EU585728	
D10	<i>Sarkidiornis melanotos</i>	FJ028237	EU585660	EU585723	
D11	<i>Anseranas semipalmata</i>	MN356217	NC 005933		MN356217
D12	<i>Aix sponsa</i>	AY666569	EU585605	EU585668	
D13	<i>Alopochen aegyptiaca</i>	Mf580159	EU585606	EU585669	
D14	<i>Anas platyrhynchos domesticus</i>				
D15	<i>Anser anser domesticus (America)</i>				
D16	<i>Anser anser domesticus (France)</i>				
D17	<i>Anser anser domesticus (Germany)</i>				
D18	<i>Anser canagica</i>	DQ432849	EU585615	EU585678	
D19	<i>Anser cygnoides domesticus</i>	LC145060	EU585616	EU585679	NC 023832
D20	<i>Anser cygnoides domesticus</i>				
D21	<i>Anser cygnoides domesticus</i>				
D22	<i>Anser erythropus</i>	GU571729	EU161871	EU585680	
D23	<i>Callonetta leucophrys</i>	FJ027277	EU914157	AF059157	
D24	<i>Dendrocygna arborea</i>				
D25	<i>Lophonetta specularioides</i>	JN801488	AF059102	AF059162	
D26	<i>Netta rufina</i>	GQ482234	EU585657	EU585720	NC 024922
D27	<i>Oxyura jamaicensis</i>	AY666448	EU585658	EU585721	MW574354
D28	<i>Chloephaga picta</i>	FJ027353	AF515262	AF515266	
D29	<i>Branta leucopsis</i>	GU571283	EU585630	EU585693	
D30	<i>Branta sandvicensis</i>	JF498832	EU585632	EU585695	
D31	<i>Anas gibberifrons</i>	JQ174015	AF059076	AF059136	
D32	<i>Anas laysanensis</i>	JF498830	AF059078	AF059138	
D33	<i>Anas luzonica</i>	KT151721	AF059079	AF059139	
D34	<i>Anas rubripes</i>	AY666211	AF059088	AF059148	
D35	<i>Anas clypeata</i>	GU571236	AF059062	AF059122	NC 028346
D36	<i>Oxyura vittata</i>	JQ175648	EU585659	EU585722	
D37	<i>Anas melleri</i>		AF059080	AF059140	
D38	<i>Oxyura australis</i>		AF119167	AY747867	
D39	<i>Branta canadensis</i>	GU571280	EU585629	EU585692	NC 007011
D40	<i>Dendrocygna viduata</i>	FJ027502	EU585649	EU585712	
D41	<i>Anser brachyrhynchus</i>	GU571244	EU585614	EU585677	
D42	<i>Chauna torquata</i>	AY140730	AY274030	AY274053	NC 052807
D43	<i>Thalassornis leuconotos</i>	U97738			
D44	<i>Tadorna tadornoides</i>		EU585666	EU585729	
D45	<i>Chenonetta jubata</i>	JN801436	AF059100	AF059160	
D46	<i>Somateria mollissima</i>	GU571620	EU585661	EU585724	MW849292
D47	<i>Anser albifrons</i>	DQ433314	EU585612	EU585675	NC 004539
D48	<i>Mergus serrator</i>	GU571482	EU585655	EU585718	MZ365040
D49	<i>Aythya affinis</i>	DQ434308	EU585621	EU585684	
D50	<i>Dendrocygna autumnalis</i>	FJ027495			
D51	<i>Clangula hyemalis</i>	GU571339	EU585638	EU585701	MW849278
D52	<i>Anas discors</i>	AY666325	EU914146	AF059128	
D53	<i>Oxyura punctata</i>				
D54	<i>Anas strepera (Mareca strepera)</i>	GQ481327	EU574791	AF059169	NC 045373
D55	<i>Heteronetta atricapilla</i>	FJ027649			

D56	<i>Anas superciliosa</i>	JN801396	AF059092	AF059152	
D57	<i>Anser caerulescens</i>	DQ434537	FJ423758		
D58	<i>Aythya fuligula</i>	JF499099	KU697802	EU585687	NC 024595
D59	<i>Aythya ferina</i>	JF499098	EU585623	EU585686	NC 024602
D60	<i>Anas formosa</i>	JN703250	AF059073	AF059133	NC 015482
D61	<i>Dendrocygna bicolor</i>		EU585646	EU585709	
D62	<i>Anas querquedula</i>	GQ481326	EU585610	EU585673	
D63	<i>Malacorhynchus membranaceus</i>		EU585651	EU585714	
D64	<i>Anas hottentota (Anas punctata)</i>		EU585608	EU585671	
D65	<i>Anas georgica</i>	FJ027096	AF059075	AF059135	
D66	<i>Biziura lobata</i>		EU585627	EU585690	
D68	<i>Cygnus atratus</i>	NC 012843	EU585641	EU585704	NC 012843
D67	<i>Aythya australis</i>	MW151626	EU585622	EU585685	
D69	<i>Amazonetta brasiliensis</i>	FJ027059	AF059054	AF059115	
D70	<i>Lophonetta cristata</i>				
D71	<i>Anas crecca</i>	KC771255	AF059064	EU585670	NC 022452
D72	<i>Tadorna tadorna</i>	KU140668	AF059113	AF059173	NC 024750
OG1	<i>Gallus gallus</i>		AF195631		NC 040902
OG2	<i>Struthio camelus</i>	LC145063	MZ545713		NC 002785

*Accession numbers of the mitochondrial gene sequences of the cytochrome *b* (Cty *b*), cytochrome oxidase subunit 1 (CO1), NADH dehydrogenase subunit 2 (ND2) and the complete mitochondrial DNA (mtDNA) was taken from GenBank.

Table S5. Waterfowl classification based on their virus prevalence (VP).

Sample ID	Species	Sub-family Classification	PCA Group	VP Values	VP Classification
D2	<i>Anas platyrhynchos</i>	Anatinae	1	12.9	HVP
D56	<i>Anas superciliosa</i>	Anatinae	2	5.7	HVP
D34	<i>Anas rubripes</i>	Anatinae	2	18.1	HVP
D71	<i>Anas creeca</i>	Anatinae	2	4.0	LVP
D31	<i>Anas gibberifrons</i>	Anatinae	2	5.8	HVP
D54	<i>Anas strepera</i>	Anatinae	2	1.5	LVP
D52	<i>Anas discors</i>	Anatinae	1	11.2	HVP
D45	<i>Chenonetta jubata</i>	Anatinae	2	2.0	LVP
D12	<i>Aix sponsa</i>	Anatinae	2	2.2	LVP
D68	<i>Aythya australis</i>	Anatinae	2	2.8	LVP
D72	<i>Tadorna tadorna</i>	Anatinae	2	6.5	HVP
D44	<i>Tadorna tadornoides</i>	Anatinae	2	5.0	LVP
D63	<i>M. Membranaceus</i>	Anatinae	1	6.3	HVP
D39	<i>Branta canadensis</i>	Anserinae	3	0.8	LVP
D47	<i>Anser albifrons</i>	Anserinae	2	2.2	LVP
D06	<i>Anser anser</i>	Anserinae	3	1.1	LVP

*Each waterfowl species was classified either as a high virus prevalence (HVP) or low virus prevalence (LVP). Classification was based on the average virus prevalence (5.5%). LVP < 5.50% < HVP.

*Virus prevalence data of the 16 species shown on the table was taken from the work of Wille, M. *et al.* [2] and Olsen, B. *et al.* [3].

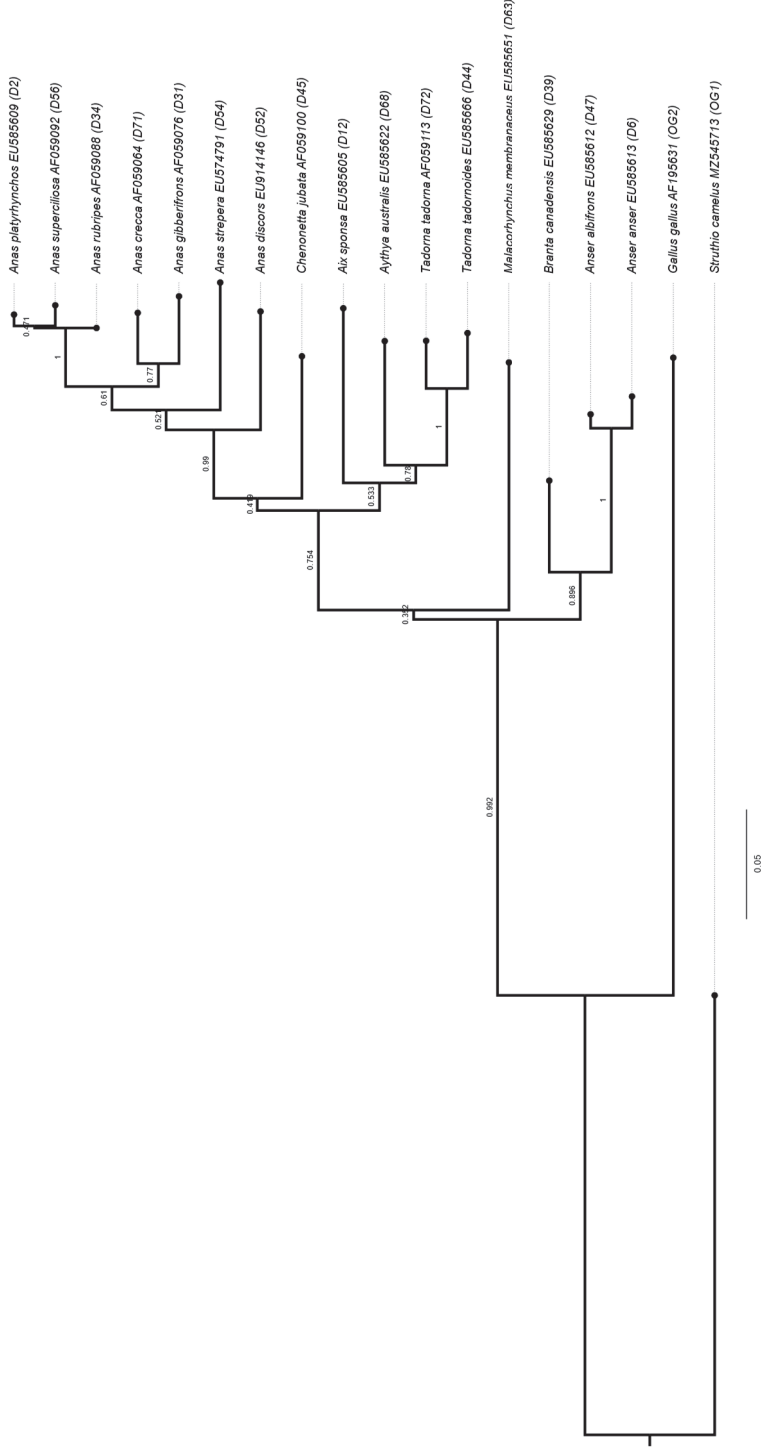


Figure S2. The evolutionary history was inferred by using the Maximum Likelihood method and Tamura-Nei model. The tree with the highest log likelihood (-5786.80) is shown. The percentage of trees in which the associated taxa clustered together is shown below the branches. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Tamura-Nei model, and then selecting the topology with superior log likelihood value. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.3642)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. This analysis involved 18 nucleotide sequences. There were a total of 1082 positions in the final dataset. Evolutionary analyses were conducted in MEGA11.

References

- [1] Hirose, K., Amano, M., Hashimoto, R., Lee, Y.C., Nishimura, S.I., Insight into glycan diversity and evolutionary lineage based on comparative avio-N-glycomics and sialic acid analysis of 88 egg whites of galloanserae. *Biochemistry* 2011, 50, 4757–4774.
- [2] Wille, M., Lisovski, S., Roshier, D., Ferenczi, M., et al., Strong host phylogenetic and ecological effects on host competency for avian influenza in Australian wild birds. *Proceedings of the Royal Society B: Biological Sciences* 2023, 290, 1–9.
- [3] Olsen, B., Munster, V.J., Wallensten, A., Waldenström, J., et al., Global Patterns of Influenza A Virus in Wild Birds. *Science (1979)* 2006, 312, 384–388.