UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of
Veterinary Science
Fakulteit Veeartsenykunde
Lefapha la Diseanse tša Bongakadiruiwa

Full genome sequencing and analysis of viral isolates of the 2017 HPAI H5N8 outbreaks in South Africa

Report 1, 28 September 2017

Prepared by: C Abolnik; Research Chair: Poultry Health and Production for the SAVA Poultry Group

Summary

- Eighteen HPAI H5N8 viruses from Mpumalanga, Gauteng, KwaZulu-Natal and the Western Cape Provinces were sequenced (period 20 June -17 August)
- Four local variants were detected based on reassortment analysis
- The Standerton virus had undergone reassortment with viruses known to be circulating in South African wild ducks and ostriches since 2012
- A dominant variant that circulated in Gauteng and Mpumalanga is evident. This variant spread to an ostrich farm in the Western Cape.
- A virus isolated from layers in the Abaqulisi district in KwaZulu-Natal on the 17th of August is distinct and could not be directly linked to any specific outbreak in the Gauteng or Mpumalanga provinces

Background

Illumina MiSeq or Ion Torrent Next Generation Sequencing (NGS) was used to obtain the full genome sequences for South African HPAI H5N8 viruses isolated in eggs since the start of the outbreak in late June 2017. Please note that only isolates from Deltamune, the University of Pretoria (UP) and Stellenbosch Provincial Veterinary Laboratory were analysed (n=18). We do not have access to isolates from samples submitted by veterinarians directly to the ARC-OVI, nor any sequence data generated by the ARC-OVI.

The Influenza A virus genome is split across eight separate genetic segments, which facilitates reassortment during co-infections. For the purpose of this and subsequent reports, the eight gene segments are designated as follows: PB2, PB1, PA, HA, NP, NA, NS.

The purpose of the study was to assess genetic drift and reassortment, and to trace the origins of strains where possible.

Results

Following assembly, phylogenetic trees were generated for each of the eight gene segments, including representative strains from the public sequence databases. Not all trees are included in this report, a representative tree for the PB2 gene is depicted in Fig. 1. Each phylogenetic tree was examined for topology (how branches are structured and ordered, and the statistical support for the branches), and clusters were assigned. A cluster is defined here as a group of closely related sequences that likely had a very recent common ancestor. Distance matrices and shared nucleotide residues were also considered in defining clusters. In most trees, the nucleotide sequence identity within a cluster was ≤99.4% (subject to revision as the epidemic progresses).

In Fig. 1, four distinct clusters of SA virus PB2 genes are evident: a, b, c and d. The same process was followed for all genes (data not shown). Table 1 provides a summary of the phylogenetic analysis for an overview of variants and reassortment between strains.

Table 1 shows that within 18 viruses analysed, there were four variants circulating in SA (five including the non-SA reference sequence). DM436893, DM440638, DM441587, 002, 17800046 and GDARD 11/08-004B Speckled dove share the same gene constellation. This cluster is the major variant. GDARD 10/08-009 and GDARD 13/7-001 share the same gene constellation. DM443397 has a unique gene constellation, as does the Standerton virus. The Villiers virus (index case) was not included, re-sequencing is underway.

Other interesting observations from the phylogenetic analyses (phylogenetic trees not shown)

- The Standerton virus provided the clearest example of HPAI H5N8 reassortment in the wild duck reservoir. The PB2 and PA genes are significantly different to those of other South African H5N8 viruses. The PA gene is related to those of H4N8 and H11N2 viruses isolated from Red-billed teals, Cape Shovellers and Yellow-billed ducks in South Africa in 2013, H7N7 viruses of ostriches in 2013 and H7N1 viruses of ostriches in 2012. Previous work hypothesized that influenza A virus genes can persist in regional duck populations for up to six years (Abolnik et al., Avian Diseases 60:286-295, 2016)
- Ostrich virus 17080046 was basal to the majority variant (viruses from the Gauteng and Mpumalanga outbreaks) across all eight genome segments. This implies that the ostrich strain was derived from this group, yet genetic drift had occurred. The genetic drift may have been caused by circulation in ostriches, or alternatively in wild birds.
- Two wild bird viruses, GDARD 13/7-001 (Egyptian goose) and GDARD 10/08-009 (Sacred ibis) were closely related across all gene segments. The Egyptian goose virus was collected in Centurion, Pretoria on 13 July, and the Sacred ibis was sampled at Bon Accord dam on the 15th of August. Thus, the same virus persisted in the Pretoria region for at least a month.
- DM436893 was isolated from commercial layers in the Ekurhuleni district on 6 July, and the Speckled dove isolate GDARD 11/08-004B was collected on the same farm a month later on 14 August. DM445187 was isolated in the middle of this period, on 4 August in the Ekuruhleni district. These three viruses cluster together on most trees (M, NA, NP, HA, PA, PB2), representing a localised outbreak strain.

Isolations and genome sequencing continues, the next report will be prepared as soon as more data becomes available.

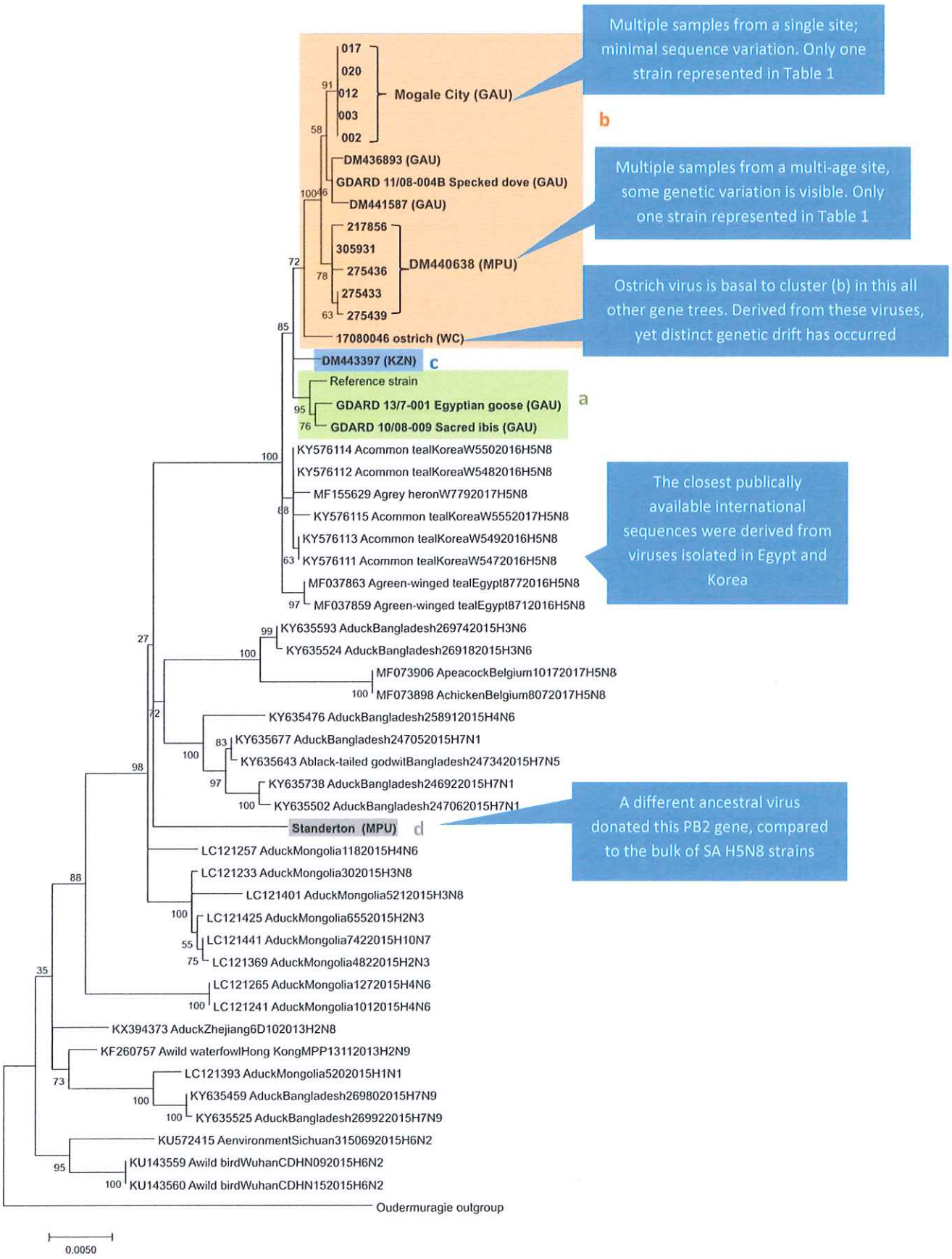

Figure 1. Maximum likelihood phylogenetic tree of South African and reference HPAI H5N8 sequences for the PB2 protein encoding gene segment. Clusters (a-d) are labelled and coloured. These clusters were confirmed using sequence analysis (distance matrices and shared mutations).

Table 1. Summary of phylogenetic results for South African HPAI H5N8 viruses

Date	Location (OIE data)	Specie	Lab number	PB2	PB1	PA	HA	NP	NA	M	NS
	Reference strain			a	a	a	a	a	a	a	a
20-Jun	Dipaleseng, MPU	chickens	Standerton	d	a	d	a	a	a	a	a
06-Jul	Ekuruhleni, GAU	chickens	DM436893	b	b	b	b	a	b	b	a
13-Jul	City of Tshwane, GAU	Egyptian goose	GDARD 13/7-001	a	a	a	a	b	a	a	a
01-Aug	Lekwa, MPU	chickens	DM440638*	b	b	b	b	a	b	b	a
04-Aug	Ekurhuleni, GAU	chickens	DM441587	b	b	b	b	a	b	b	a
04-Aug	Mogale City, GAU	chickens	002, 003, 012, 017, 020	b	b	b	b	a	b	b	a
07-Aug	Steve Tshwete, MPU	chickens	DM441839	b	b	b	b	a	b	b	a
09-Aug	Hessequa, WC	ostrich	17080046	b_1	b_1	b_1	b_1	a	b_1	b_1	a_1
14-Aug	Ekurhuleni, GAU	Speckled dove	GDARD 11/08-004B	b	b	b	b	a	b	b	a
15-Aug	City of Tshwane, GAU	Sacred ibis	GDARD 10/08-009	a	a	a	a	b	a	a	a
17-Aug	Abaqulisi, KZN	chickens	DM443397	c	c	c	b_2	a	c	c	b

*217856, 305931, 275436, 275433, 275439

Interpreting this table

Colour coded genes are useful for assessing genetic relationships between strains at a glance. Each viral genome is represented horizontally (NB. there is no relationship between a's, b's, c's or d's horizontally). Each strain is stacked for comparison with other viruses across the same gene segment, with letters and colours according to the assigned cluster. Please compare the PB2 column above to the corresponding tree in Fig. 1 as an example.

The ostrich virus 17080046 is labelled as b_1 in Table 1 (a_1 for the NS gene). Sequence homology placed it within the b cluster in each case, but it was also basal to the remainder of the b cluster viruses. Similarly, the HA gene of DM443397 was significantly unique to flag it as "b_2". Note that the NP and NS genes provide relatively poor discriminatory power (i.e. it's less easy to identify and assign clusters because the gene sequences are more conserved).