# From emergence to endemicity: highly pathogenic H5 avian influenza viruses in Taiwan

**Authors:**

Yao-Tsun Li[1*], Hui-Ying Ko[2], Joseph Hughes[3], Ming-Tsan Liu[4], Yi-Ling Lin[2,5], Katie Hampson[1], Kirstyn Brunker[1*]

**Affiliations:**

1. School of Biodiversity, One Health & Veterinary Medicine, University of Glasgow, Glasgow, UK
2. Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan
3. MRC-University of Glasgow Centre for Virus Research, Glasgow, UK
4. Center for Research, Diagnostic and Vaccine Development, Centers for Disease Control, Ministry of Health and Welfare, Taipei, Taiwan
5. Biomedical Translation Research Center, Academia Sinica, Taipei, Taiwan

***Corresponding authors:**

Yao-Tsun Li Yao-Tsun.Li@glasgow.ac.uk

Kirstyn Brunker Kirstyn.Brunker@glasgow.ac.uk

University of Glasgow, Glasgow G12 8QQ, UK

**Author contributions:** Conceptualisation: YTL, HYK. Formal analysis: YTL. Methodology: YTL, JH. Resources: HYK, YLL. Writing - original draft preparation: YTL. Writing - review & editing: YTL, HYK, JH, MTL, YLL, KH, KB. Supervision: JH, KH, KB. Funding: KH, KB.

**Data accessibility statement:** The custom codes used in analyses, and the XML files required for BEAST can be found at https://github.com/yaotli/endemic. The accession numbers of the Taiwan sequences analysed in our study are listed in Supplementary Table 2.

**Keywords:** avian influenza, endemic, zoonotic

## Abstract

A/goose/Guangdong/1/96-like (GsGd) highly pathogenic avian influenza (HPAI) H5 viruses cause severe outbreaks in poultry when introduced. Since emergence in 1996, control measures in most countries have suppressed local GsGd transmission following introductions, making persistent transmission in domestic birds rare. However, geographical expansion of clade 2.3.4.4 sublineages has raised concern about establishment of endemic circulation, while mechanistic drivers leading to endemicity remain unknown. We reconstructed the evolutionary history of GsGd sublineage, clade 2.3.4.4c, in Taiwan using a time-heterogeneous rate phylogeographic model. During Taiwan's initial epidemic wave (January 2015 - August 2016), we inferred that localised outbreaks had multiple origins from rapid spread between counties/cities nationwide. Subsequently, outbreaks predominantly originated from a single county, Yunlin, where persistent transmission harbours the trunk viruses of the sublineage. Endemic hotspots determined by phylogeographic reconstruction largely predicted the locations of re-emerging outbreaks in Yunlin. The transition to endemicity involved a shift to chicken-dominant circulation, following the initial bidirectional spread between chicken and domestic waterfowl. Our results suggest that following their emergence in Taiwan, source-sink dynamics from a single county have maintained GsGd endemicity, pointing to where control efforts should be targeted to eliminate the disease.

## Introduction

Preventing emerging zoonotic viruses from establishing endemic circulation is essential to public health, considering the great health burden caused by endemic zoonoses globally[1]. An example of particular concern is the A/Goose/Guangdong/96-like (GsGd) H5 viruses, a lineage of highly pathogenic avian influenza (HPAI) first identified in southern China. From 2003-2006 HPAI GsGd viruses spread globally, infecting poultry in countries across Asia and Europe[2]. Most countries have subsequently eliminated HPAI GsGd viruses[3,4], but, since 2010, a few countries, including Bangladesh, China, India, Indonesia, and Vietnam, have been classified as endemic, due to continued circulation in poultry and sporadic human cases[5]. Although there has been extensive characterisation of the genetic origins and the migration routes of HPAI GsGd viruses[6–8], the mechanisms facilitating persistence and the transition to endemicity remain unknown.

Since emergence GsGd viruses have undergone diversification by accumulating mutations in their surface protein hemagglutinin (HA), resulting in multiple antigenically distinct sublineages, termed clades[2,3]. These genetic changes can alter host preferences, determining the virus' reservoir, dissemination and persistence[9]. GsGd viruses of the clade 2.3.4.4, including clade 2.3.4.4a-2.3.4.4h, are distinguished from previous GsGd viruses by various neuraminidase (NA) subtypes (N2, N5, N6 and N8) besides N1, and have rapidly increased in the global population since 2014[10,11]. These "H5Nx" viruses have caused poultry and wild animal outbreaks in previously unaffected regions, such as the Americas[10,12], and have also demonstrated unexpected transmission from Europe back to China[13]. Moreover, the viruses have led to mammal-to-mammal transmission in minks and cows[14,15]. The increasingly sustained circulation in areas previously thought to be sinks raises concerns about expanding establishment of endemic circulation[16].

The clade 2.3.4.4c virus, one of the H5Nx sublineages, caused severe outbreaks in Korean poultry farms in early 2014[10,17]. Following migratory bird flyways, the virus spread to Japan, the USA and Taiwan by the end of that year[18–20]. In 2015, Taiwan experienced a devastating epidemic in domestic birds and culled over 5 million birds to curb localised outbreaks[21]. This marked the first significant transmission of HPAI GsGd viruses in the country[22,23]. While clade 2.3.4.4c viruses were eliminated in other countries before 2017[11,20], the virus continued to circulate as H5N2 subtype in Taiwan until at least 2019[20,24]. With multiple genotypes generated by reassortment with local low pathogenic avian influenza (LPAI) viruses[20,23], the HPAI clade 2.3.4.4c virus in Taiwan has been recognized as circulating endemically[21]. How this clade dispersed in Taiwan and the factors underlying its endemic establishment are not understood.

In this study, we quantified the persistence of GsGd outbreaks globally using virus HA genes, and identified the source of the endemic clade that established in Taiwan within this global context. We then curated the HA genes of the clade 2.3.4.4c virus in Taiwan, isolated from January 2015 to March 2019, representing approximately 20% of outbreaks that occurred within the country. Using this data set, we performed time-heterogeneous phylogeographic analyses to reconstruct the geographical and ecological dispersal dynamics of the clade as it transitioned from emergence to endemicity. Specifically, we elucidate the factors facilitating viral spread among counties/cities in Taiwan following the first epidemic wave in 2015.

## Results

### Global persistence of GsGd

To better understand local epidemics following GsGd introductions on a global scale, we identified descendent sublineages. We found about 400 sublineages from a uniformly downsampled dataset using the H5 genes of GsGd viruses up to September 2022 (Figure 1). These GsGd sublineages circulated across 78 countries, but their circulation was generally short-lived, with an average duration of 0.6 years based on the date of sample collection. Closer examination of sublineages that persisted for more than three years (based on a 95% quantile of 2.64), revealed 15 persistent sublineages in Southeast Asia, South Asia and the Middle East, including the clade 2.3.4.4c virus in Taiwan (Figure 1). Geographical reconstruction informed by outbreak reports supported this finding that most introductions had limited onward transmission (on average, 0.61 years) and identified the same countries with prolonged viral circulation (data not shown).

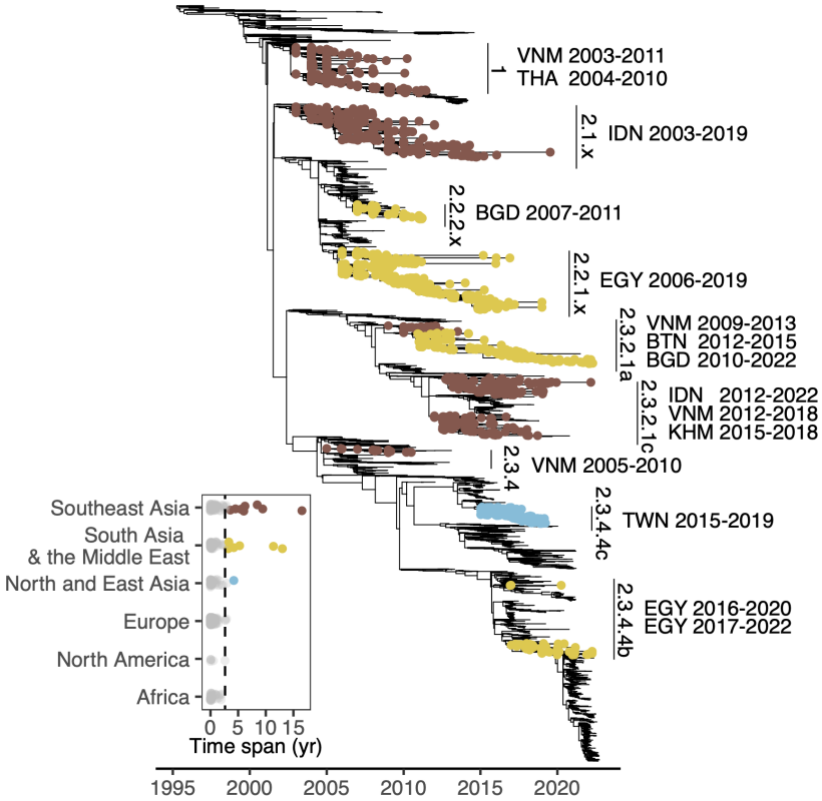


**Figure 1. The duration of GsGd outbreak lineages**. The time-scaled phylogenetic tree inferred with HA genes highlights fifteen lineages, which persistently circulated for at least three years. These are labelled with H5 nomenclature classification, country of isolation and duration. The inset shows the lineage duration classified by geographical area, with persistent sublineages highlighted as tips on the phylogeny. The dashed line denotes 95% quantile. VNM, Vietnam; THA, Thailand; IDN, Indonesia; BGD, Bangladesh; EGY, Egypt; BTN, Bhutan; KHM; Cambodia; TWN, Taiwan.

The emergence and spread of clade 2.3.4.4c in Taiwan

In January 2015, over 600 outbreaks of clade 2.3.4.4c were reported in Taiwan across 10 counties/cities (Figure 2A). Subsequently, reported HPAI outbreaks in Taiwan declined and became sporadic in previously affected areas, with only three cases reported during July-August 2016, before transmission resumed in winter 2016-2017 (Figure 2B). To understand the spread of the clade 2.3.4.4c virus in Taiwan, we curated a genetic data set of viral HA sequences. The sequences are distributed over time from January 2015 to March 2019, and their sample sizes per county/city correlate well with the numbers of reported outbreaks (Figure 2D).

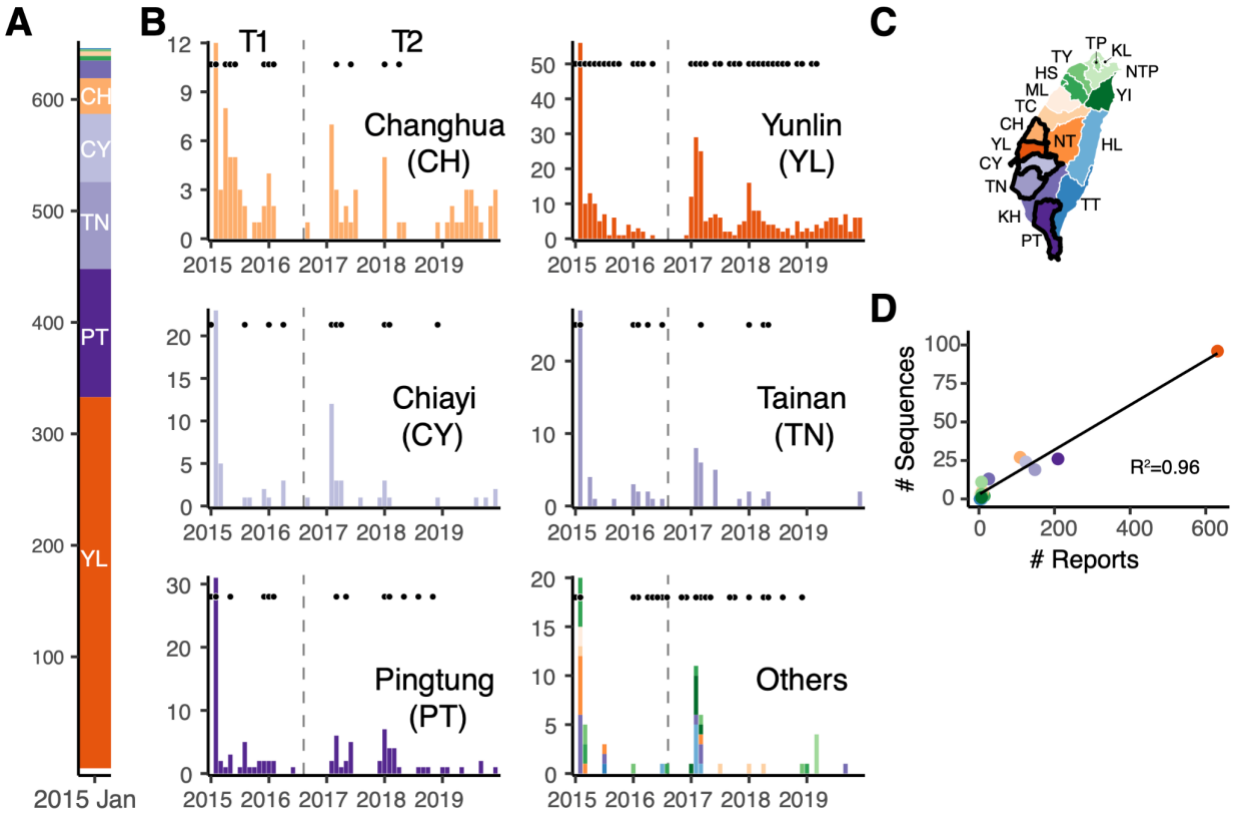



**Figure 2. The emergence of HPAI GsGd clade 2.3.4.4c in Taiwan and subsequent circulation from 2015-2019.** (A and B) The number of reports of infected poultry in farms or abandoned carcasses. The January 2015 numbers are pooled in panel A, while the monthly numbers of reports are presented by location in panel B. Black dots above the bars indicate the availability of genomic data in the corresponding months. The dashed vertical lines indicate the start of September 2016, the transition from the emergence (T1) to the endemic phase (T2). (C) The map highlights the five counties/cities with the most reports. (D) The correlation between reported outbreaks and available sequences for each of the 14 counties/cities. The points are coloured using the same colour scheme as the map and the other panels. NTP, New Taipei; TP, Taipei; KL, Keelung; TY, Taoyuan; HS, Hsinchu; YI, Yilan; ML, Miaoli; TC, Taichung; CH, Changhua; NT, Nantou; YL, Yunlin; CY, Chiayi; TN, Tainan; KH, Kaohsiung; PT, Pingtung; HL, Hualien; TT, Taitung.

To understand the dispersal dynamics of the 2.3.4.4c clade in Taiwan, we reconstructed the viral spread using continuous phylogeographical approaches which inferred the spatial locations of nodes. The Bayesian MCMC sampling process used priors informed by surveillance data to specify the locations of viruses (see Methods). The results, stratified by year, indicate that in 2015 there was intensive virus dispersal within and between western counties/cities (CH, YL, CY and TN), as well as more limited dispersal in southern counties/cities (KH and PT)(Figure 3). Long-distance dispersal events mostly originated from the main cluster in the west, and the northward dispersal events did not appear to seed sustained outbreaks. Since 2016, the number of dispersal events has significantly decreased and since 2017 Yunlin County (YL) has become the primary source of both short- and long-distance spread (Figure 3). The pattern is further illustrated by discretizing the inferred locations, with diffusion within YL dominating in all directions (Supplementary Figure 1).

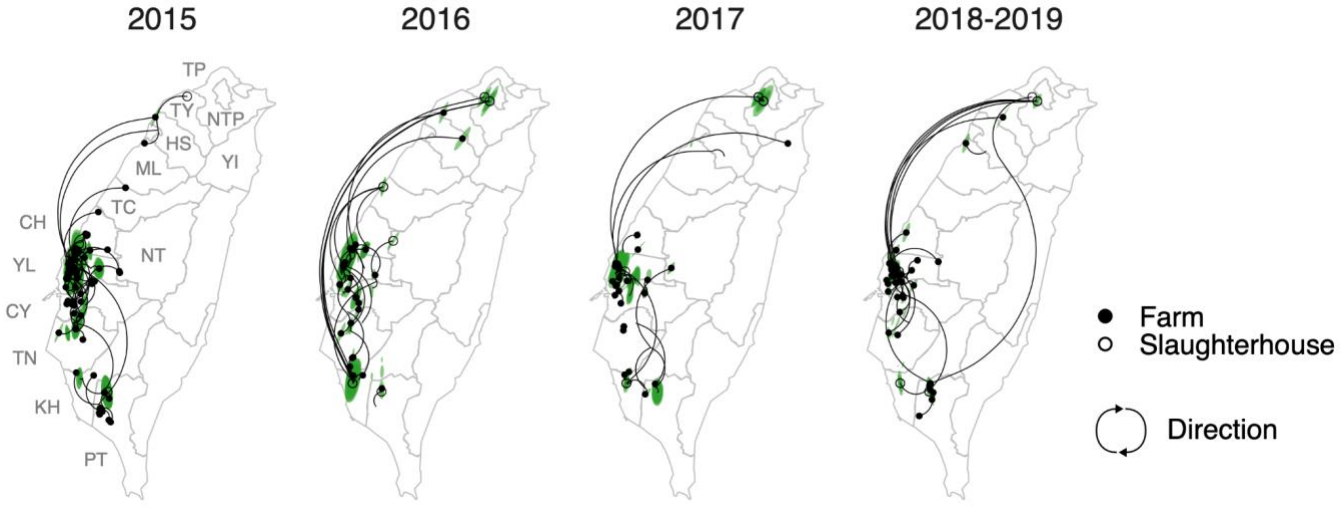


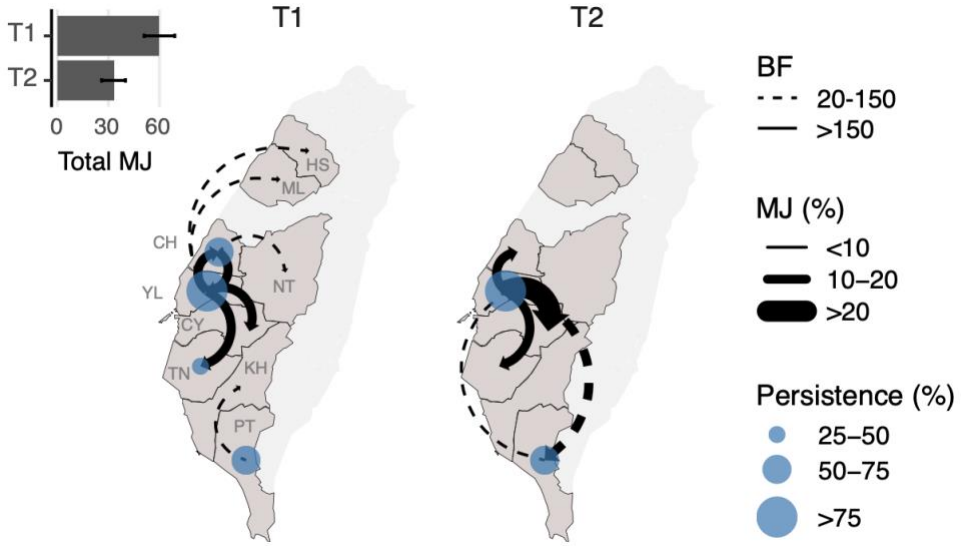

**Figure 3. Reconstruction of the dispersal of clade 2.3.4.4c in Taiwan, 2015-2019.** The points on the maps indicate the location of the virus samples, while the coloured areas represent the 80% HPD (highest posterior density) of the nodes inferred by the continuous phylogeography. Curves represent the branches

of the maximum clade credibility (MCC) tree and are categorised into four periods based on the timing of parental nodes. Dispersal direction (clockwise) of the viral lineage is indicated by the curvature.

Use of explicit phylogeographic methods can reduce sampling biases[25,26]. Here, the locations of samples from slaughterhouses or rendering factories introduced inaccuracies into our inference. This is because these locations were unlikely to be the origins of viral spread, particularly those from the northern cities (e.g. TP and NTP) where agricultural activity is minimal (Supplementary Figure 2). To reconstruct the dispersal history of the virus, we therefore used discrete diffusion methods to reassign the locations of slaughterhouse samples. Specifically, we employed priors informed by the surveillance data, which represented possible sources of animals transported to slaughterhouses or rendering factories (Supplementary Figure 3B). Additionally, we used a heterogeneous rate model to identify changes in dispersal patterns between two epidemiological phases defined from the surveillance data (Figure 2): from January 2015 to August 2016 (T1) and from September 2016 onwards (T2).

The statistically supported routes that we inferred indicate a shift in dispersal from outbreaks having multiple sources to predominantly originating from a single source (Figure 4 and Supplementary Figure 4A), and are generally consistent with inference from the continuous phylogeographic analyses. During the emergence phase (T1), substantial transition events were detected between two western counties (YL and CH), in addition to dispersal from YL to nearby counties, with less frequent transitions identified from CH and PT to neighbouring counties/cities. In contrast, during the transition to endemicity (T2), all supported routes originated from YL, except for a less frequent path from PT. To account for local transmission, we quantified the maximum time interval assembled by branches with the same inferred state (Figure 4 and Supplementary Figure 5A). YL appears to be the only site where viruses

197 persisted for more than 75% of both time periods (T1 and T2), suggesting its potential as a source of virus
198 spread. Furthermore, when considering the viruses leading to the most recent isolates, CH harboured the
199 "trunk" of the clade in 2015, but since 2016, YL has consistently been the major site harbouring viruses
200 leading to the most distant progeny (Supplementary Figure 6A). These results suggest that YL has become
201 the most prominent origin for nationwide dispersal following the initial epidemic wave. In contrast,
202 circulation in other locations, particularly CH, has been largely suppressed, preventing spread to other
203 counties/cities.
204




205
206

207 **Figure 4. Schematic illustration of the main diffusion routes of the Taiwan clade 2.3.4.4c virus**
208 **during the first epidemic wave and subsequent transition to endemicity.** The statistically supported
209 directions between counties/cities (Bayes factor>20) inferred by the discrete phylogenetic analyses are
210 indicated as curves with arrows. The different line types reflect the strength of support (BF). The thickness
211 of the curve reflects the frequency of transition events, presented as the proportion of total Markov jumps
212 in the time period (T1 or T2). Total Markov jumps are indicated on the upper left inset. The size of the

blue circles reflects the degree of persistence, calculated by unifying phylogeny branches with both nodes inferred as the same geographical area divided by the T1/T2 time period. Only areas where viruses circulated for more than 25% of the period are labelled. T1, January 2015 to August 2016; T2, September 2016 to March 2019.

To investigate the impact of sampling, we created two subsampled data sets by downsampling genetic sequences from county YL, resulting in YL no longer being the most frequently sampled site in each dataset. The subsampled data showed a shift in the origins of viral diffusion to CH or TN during the emergence phase (T1), but the inferred routes during the endemic phase (T2) were almost identical to those obtained using the full dataset (Supplementary Figure 4B and 4C). Similarly, persistence quantified from subsampling shows that CH replaced YL as the location with the most persistent circulation during the initial epidemic wave (T1). However, the estimates during the endemic phase (T2) were largely unaffected by the targeted subsampling (Supplementary Figure 5B and 5C).

To understand how the surveillance data may have biassed our inference, we also performed the same phylogeographic approaches based on the original collection county/city, with unknown samples assigned an uninformative prior containing all collection locations. This inference, which is independent of outbreak records, reveals a more widespread distribution of dispersal origins during the initial epidemic wave (T1). In addition to PT, more southern locations (i.e. TN and KH) were also identified as dispersal origins (Supplementary Figure 4D). During the endemic phase (T2), the pattern is consistent with the results obtained from considering the outbreak records. Notable differences were found in routes involving locations where samples were mainly from slaughterhouses or rendering factories (i.e. NTP and TP, Supplementary Figure 4D). Estimates of persistence based on record-independent priors (Supplementary

Figure 5D) and all trunk probability conditions tested (Supplementary Figure 6B-D) indicate similar results to the full data set. Taken together, these findings suggest that our observation of nationwide geographical spread is robust to both sampling and state assignment.

Location-specific factors drive the risk of viral dispersal

To identify the risk factors that facilitated the spread of clade 2.3.4.4c in Taiwan, we implemented a Generalised Linear Model (GLM) model accommodated by the time-heterogeneous discrete phylogeographic method. This allowed for estimates of predictor effect size and inclusion probability by time period. The results show that sample size is positively and significantly correlated with diffusion during the initial epidemic wave (T1, Figure 5A, RandEffect-). The number of poultry farms has a prominent inclusion probability during the endemic phase (T2), but this was not statistically significant. To account for location-specific factors, the GLM model was adjusted by adding random effects for each site, in addition to the existing predictors. The new model reduced the effects of sample size, and no further predictors were identified (Figure 5A, RandEffect+). Interestingly, the result shows that dispersal routes to high incidence counties/cities (CH, YL, CY and TN), and the direction from YL, have significant positive effects (Figure 5B). To avoid overparameterization, we conducted GLM estimations with only one of the predictors that had higher inclusion probabilities and sample size. Consistent with the full model, the reduced models that include predictors for either poultry farm, road distance, or cropland show no significant effects for these factors (Supplementary Figure 7). Furthermore, the random effects that specify directions to YL, CY, or TN are significant in all three reduced models, while the effect of the direction from YL is significant in one reduced model. The results remained consistent when the full model was examined in a time-homogeneous manner (Supplementary Figure 8). These findings suggest

that location-specific factors are more likely to explain the spatial spread of the virus in Taiwan than shared agricultural risk factors.

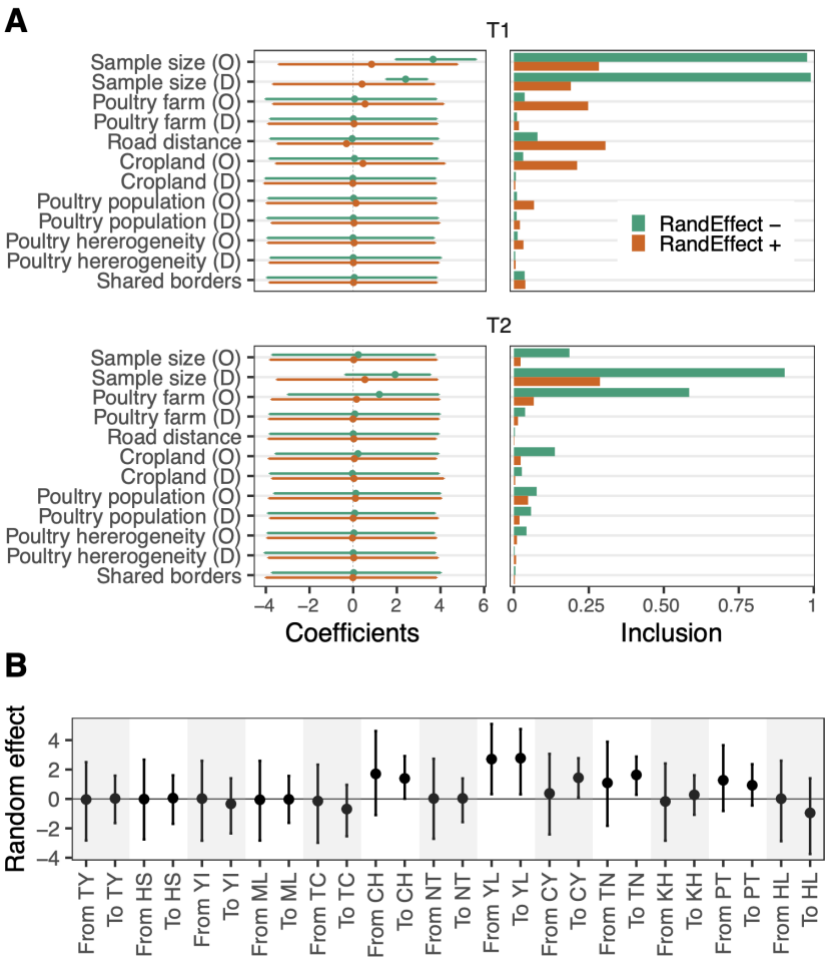



**Figure 5. Potential predictors of dispersal of clade 2.3.4.4c virus between counties/cities in Taiwan.** (A) The conditional effect sizes quantified by coefficients (left) and the inclusion probabilities of predictors (right) estimated by the time-heterogeneous phylogenetic generalised linear model (GLM). Results of the models with and without random effect variables are shown. The predictor names are denoted by O in parentheses for origin and D for destination. (B) Location-specific random effects in the GLM model. The effect sizes are in log space and are presented as mean with 95% HPD interval.

Most YL outbreaks re-emerge from endemic hotspots

The phylogeographic results suggest a pivotal role of YL county in driving outbreaks across the country. However, the determinants that distinguish YL country from other locations remain unclear. To investigate this, we evaluated re-emerging HPAI outbreaks in the persistent foci during the endemic phase (T2). Based on the practice of establishing surveillance zones after case identification[21,27], hotspots were defined by overlapping buffer areas within 1 km of node locations inferred by the continuous phylogeographic methods, or within the same distance of the node plus outbreak locations (Figure 6A). When mapping new outbreaks in YL from 2019 to 2022, over 50% of these occurred within the phylogenetically-defined hotspots. Moreover, over 70% of the outbreaks were found in the hotspots defined by the phylogenies supplemented with contemporary outbreak sites (Figure 6B). The areas defined by both methods account for less than 20% of the area of YL (Supplementary Figure 9). When the buffer radius was set to 0.5 km, over 30% and 60% of the new outbreaks were still identified within the hotspots defined by the two methods, accounting for less than 10% of the county's area (Supplementary Figure 9). When the approach is applied to four other high incidence areas in Taiwan (Figure 2B), YL shows significantly higher percentages of recurrence in the hotspots defined by the two buffer distances (Figure 6C). These findings suggest that the inability to interrupt transmission within the county has led to the development of an endemic focus of persistent viral circulation in YL.

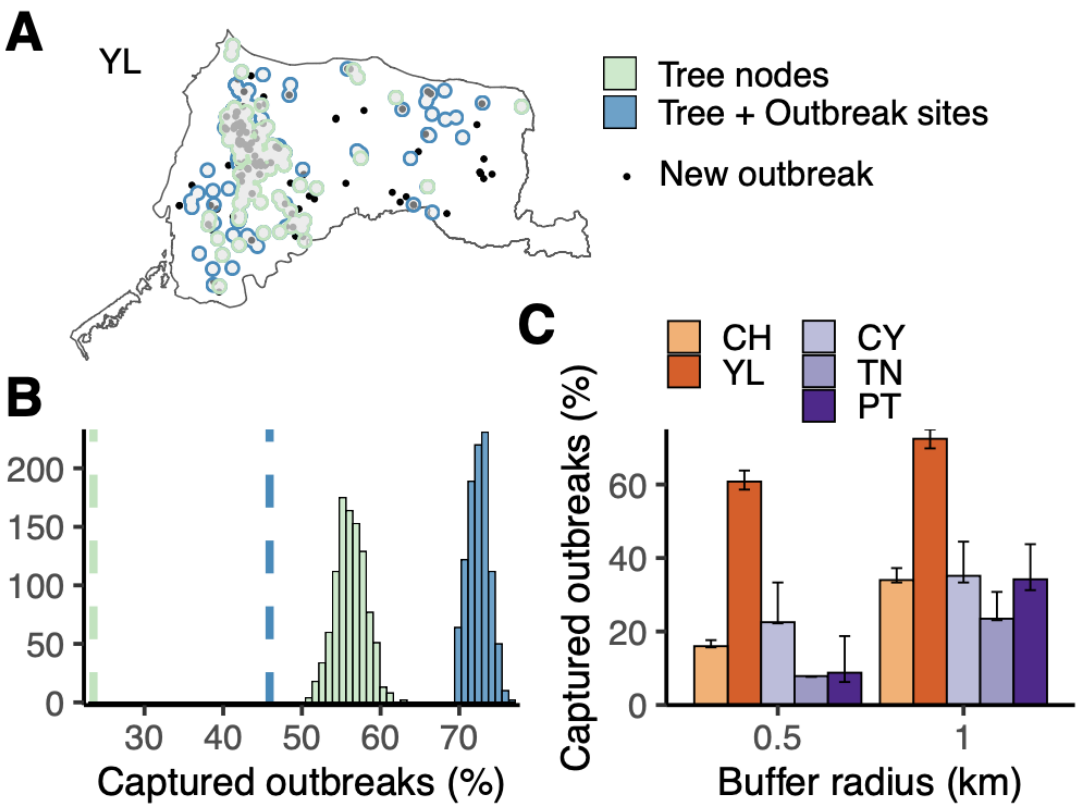



**Figure 6. Evaluating the spatial distribution of re-emergent H5 HPAI outbreaks in Yunlin county (YL).** (A) An example from a randomly selected posterior tree. The black points indicate reported outbreak sites that occurred between 2019 and 2022, after the latest available genetic data. The green/blue areas created by merged buffers indicate the inferred hotspots. (B) Distributions of captured new outbreaks by inferred epidemic hotspots based on 1 km-buffers with centres of nodes only (green) or nodes plus outbreak sites (blue). The vertical dashed lines indicate the mean of randomly generated circulating areas with identical buffer numbers as tree nodes (green) or tree nodes plus outbreak sites (blue). (C) Comparing the re-emergence of outbreaks within inferred hotspots in different locations. Proportions of captured outbreaks were calculated with buffers of two radius distances. Error bars indicate 95% credible intervals.

Development of chicken-dominant circulation

The initial 2015 epidemic wave in Taiwan was characterised by widespread transmission in farms rearing domestic waterfowl[28]. This corresponded to a higher proportion of outbreaks and viral samples attributed to ducks or geese compared to chickens in 2015 (Figure 7B and Supplementary Figure 3A). To assess transmission between poultry species, we applied similar phylogeographic methods to the same genetic data, classified into Anseriformes (duck and goose) or Galliformes (chicken, turkey and quail) species. The inferred diffusion indicates that transmission between the two host groups was bidirectional during the initial epidemic wave (T1), with 80% of transmission attributed to spread from waterfowl, while both routes were statistically supported (Figure 7A). During the endemic phase (T2), transmission shifted to gallinaceous birds, while the transmission route from waterfowl was not supported (Figure 7A). This finding was further confirmed by an NA data set containing cognate N2 genes of Taiwan clade 2.3.4.4c viruses (Supplementary Figure 10). The trunk analyses show that the terrestrial poultry has maintained the gene source of the lineage since its introduction (Figure 7C). These results suggest that terrestrial poultry, mainly chickens, served as reservoir for the clade 2.3.4.4c virus during the endemic phase (T2), following the first epidemic wave during which inter-host group transmission dominated and more likely originated from domestic waterfowl.

By detecting signals of positive selection, we assessed the evidence for host adaptation. The dN/dS (w) estimate of HA (H5) genes in Taiwan clade 2.3.4.4c viruses is similar to that of HA genes in a North American H5N2 lineage introduced to the Taiwanese chicken population as early as 2003[20,29], while the estimate of NA (N2) genes in the clade 2.3.4.4c viruses is higher than that of the North American lineage (Supplementary Table 1). In addition, there are less than five positively selected sites detected in both HA and NA of clade 2.3.4.4c in Taiwan, which is comparable to the results found in the North American H5N2 lineage. Note that all of the substitutions occurring at these sites were not fixed in the samples

collected in 2019 (Supplementary Table 1). These results do not support selection pressure from host adaptation during the transition to endemicity and establishment of chicken-dominant circulation.

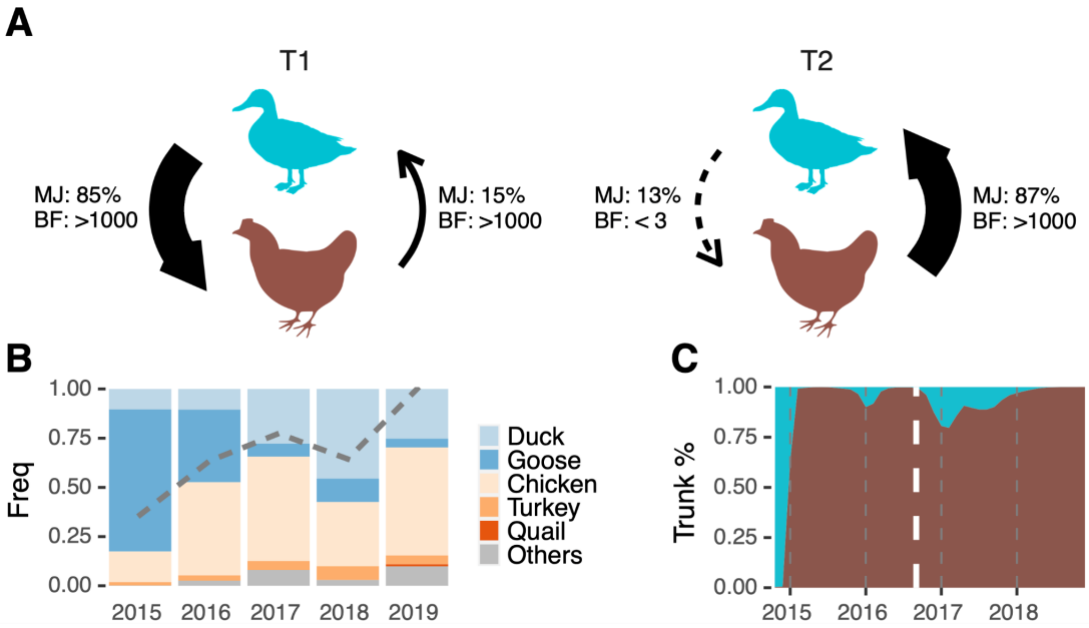



**Figure 7. Transmission shifted from predominantly domestic waterfowl to chicken during the transition from epidemic to endemic circulation of 2.3.4.4c in Taiwan.** (A) Shift in diffusion patterns between Anseriformes (duck and goose) and Galliformes (chicken, turkey and quail). Frequencies of transition events (Markov jump, MJ) and statistical support (Bayes factor, BF) were inferred using the discrete phylogenetic method. (B) The host distribution over time based on the surveillance data. The grey dashed line indicates the proportions of Galliformes in the genetic data. (C) Inferred host categories, i.e. Anseriformes or Galliformes species, of the trunk of clade 2.3.4.4c. The white dashed line divides T1 and T2, with areas coloured using the same scheme as panel (A).

## Discussion

The circulation of clade 2.3.4.4c in Taiwan presents an opportunity to study how the GsGd lineage established endemic transmission following its emergence. By comprehensively characterizing the dynamics of clade 2.3.4.4c in Taiwan, this study focuses on viral dispersal and mechanisms driving endemic transmission after the initial epidemic wave. While the introduction and the genetic origins of the virus have been reported previously[20,23], the dispersal dynamics and mechanisms driving persistence were previously unknown. Our time-heterogeneous phylogeographic analyses show that nearly all viral diffusion after mid-2016 originated from a single county, in contrast to the interconnected and multi-origin diffusion between counties/cities during emergence in 2015 (Figure 3 and 4). These endemic hotspots in Taiwan, where the disease was persistent, continued to seed outbreaks in other agricultural areas, while new outbreaks were mostly under control. Therefore, our geographical reconstruction provides insights into strategic control measures to eliminate GsGd viruses in Taiwan, emphasising the need to focus on the endemic hotspots. Note that the clade 2.3.4.4c virus was still being identified in 2023[24], although viral genetic data isolated after 2019 was unavailable.

Previous studies using phylogenetic GLMs have identified the poultry trade[30], geographic distance[31] and the wild bird migration network[13] as predictors of GsGd viral dispersal. These studies have investigated transmission patterns globally as well as single epidemic waves, but have not examined the risk of establishment of endemic transmission. Our GLMs did not identify any significant agricultural predictors for GsGd viruses in Taiwan. Instead, our analyses highlighted location-specific factors that were not explained by the tested predictors. These results indirectly support nationwide spread from a specific endemic county with other risk factors playing only a minor role. It is possible that overparameterization due to sample size could not be excluded, although different models were

360 extensively explored (Supplementary Figures 7-8). Of all the predictors tested in the GLMs, the number
361 of poultry farms had the highest inclusion probability (Figure 5 and Supplementary Figures 7-8), a finding
362 corroborated by a spatial case-control study focusing on clustering of H5 viruses in Taiwan from 2015-
363 2017[27].
364
365 Lineage-specific host preferences can be targeted to prevent viral spread[32]. Control measures for
366 HPAI have targeted domestic waterfowl due to their biological similarity to wild waterfowl, which are
367 capable of long-distance dissemination[8,33]. In addition, infected waterfowl exhibit mild clinical
368 manifestations compared with Galliformes species like chickens, increasing the risk of silent transmission
369 and co-infection with multiple strains[2]. Our diffusion analyses indicate that GsGd transmission in Taiwan
370 shifted from bidirectionality, with mixing between domestic waterfowl (Anseriformes species) and
371 terrestrial species (Galliformes species), to a predominantly chicken-origin circulation (Figure 7). The
372 greater transitions from waterfowl to terrestrial species during the first epidemic wave, aligned with
373 devastating outbreaks in geese[21,28], that did not lead to sustained descendant lineages (Figure 7C). Based
374 on the selection analyses of the surface proteins, it is unlikely that the change in dispersal patterns between
375 these epidemiological phases reflects virus adaptation from waterfowl to terrestrial birds (Supplementary
376 Table 1). Instead, the results suggest that unknown factors in the chicken farming networks supported
377 endemic circulation of GsGd in Taiwan, while viral lineages circulating in goose and duck populations
378 were mostly eliminated. During the transition to endemicity, more than half of the Anseriformes samples
379 were either isolated or inferred to be from Yunlin county (Supplementary Figure 3), indicating that the
380 diffusion pattern in host ecology was not confounded by the spatial pattern described earlier.
381

Although these data have provided invaluable information on H5 HPAI circulation in Taiwan, it is unlikely that the surveillance data from the Taiwanese agricultural authorities identified all of the GsGd or clade 2.3.4.4c outbreaks. This limitation is implied by the results of the discrete diffusion analysis that relied solely on genetic sequences without outbreak records (Supplementary Figure 4D). The analysis exclusively identified statistically supported transmission routes originating from locations in southern Taiwan, particularly city KH, during the first epidemic phase. It was also noted that some viral samples do not precisely correspond to inferred locations (Figure 2B and Methods). Our reconstruction schemes demonstrate that state assignment strategies have little effect on all tested characteristics during the endemic phase, T2 (Supplementary Figure 4-6). There may also be an issue with the lack of explicit viral classification for each entry in the surveillance data, leading to inaccurate priors for the locations of the 2.3.4.4c virus. However, there is no evidence of continued circulation of other GsGd lineages, including clade 2.3.4.4b, in Taiwan during 2015-2019, according to available sequences (see Methods) and outbreak reporting[24].

Our findings indicate that in Taiwan, HPAI control should prioritise the source region, specifically by strengthening measures to interrupt transmission and improve biosecurity in poultry farms in Yunlin county. Meanwhile, epidemics in other regions could be readily suppressed. On a global scale, attention should be paid to countries newly affected by GsGd viruses after 2020[12,16], as well as countries where viruses were previously seeded (Figure 1). Effective interventions can be implemented through continued local and global surveillance efforts to discern dispersal patterns supported by genomic data.

## Materials and Methods

Sequence data preparation

H5 hemagglutinin (HA) sequences of avian influenza A viruses were downloaded from NCBI (Influenza Virus Resource, https://www.ncbi.nlm.nih.gov/genomes/FLU/)[34] and GISAID (EpiFlu, https://www.gisaid.org)[35] on 17 October, 2022. N2 neuraminidase (NA) sequences of part of the associated H5 viruses in Taiwan were also downloaded from the same databases. Short sequences (<1500 for HA and <1200 for NA) or sequences containing notable (n=100) ambiguous nucleotides were removed from the download. Filtered genetic data were then aligned using Nextalign v2.3.0[36] or MAFFT v7.490[37] and trimmed to coding regions. Multiple basic amino acids at the H5 cleavage site were also removed. Sequences of Taiwan GsGd clade 2.3.4.4c viruses analysed in this study have all been described previously[20,21,23,28,38,39].

Quantifying the persistence of GsGd lineages globally

The above GsGd H5 sequences were first subjected to location-focused downsampling. That is, for each country where genetic data were isolated, five sequences were randomly sampled on different dates within a monthly interval. When countries had samples with incomplete temporal information, up to 50 sequences were allowed to be randomly included per year. The downsampled data set was then used to reconstruct a time-scaled tree using Treetime v0.11.1[40] and a maximum likelihood (ML) tree inferred by IQ-TREE v2.2.0[41], justified by a strong temporal signal ($R^2$=0.95). The *mugration* function in Treetime was performed to infer the country as ancestral states for each node. Post-introduction sublineages were defined by first identifying internal branches representing state transition and all the descending taxa of those branches. For each of these sublineages, nested transition branches were removed, resulting in a lineage each comprising taxa isolated in the same country and sharing a most common ancestor linked to viral movement. Considering the variation in sequencing effort across countries, ancestral reconstruction

based on transition rates informed by frequencies of HPAI H5 outbreaks during 2000-2022[42] was also performed.

Phylogenetic analyses and data set design

A ML tree was inferred for the HA alignment using IQ-TREE. Major H5 viral lineages including clades 2.3.4.4b and c, were classified according to the GsGd H5N1 nomenclature proposed by the World Health Organization[3,11]. All available H5 viruses isolated in Taiwan were identified on the tree. A total of 292 HA sequences isolated in Taiwan were classified as clade 2.3.4.4c, accounting for nearly all the available GsGd isolates in the country (292/295). The main working dataset was generated by removing duplicated sequences and viruses isolated in apparently the same outbreak (n=252). To verify results in the discrete phylogeographic analyses, more evenly distributed datasets were prepared by downsampling the sequences from Yunlin county (YL). The downsampled datasets, which include no more than six samples per year in YL, contain comparable sample sizes as the other locations where disease was prevalent (sample size in county YL, 26; CH, 27; CY, 24; PT, 26). In addition, to verify patterns in host diffusion, the N2 genes corresponding to the curated H5 datasets were collected (n=183); N2 is the predominant subtype among clade 2.3.4.4c viruses in Taiwan (Supplementary Figure 3). Temporal signal in each data set was examined by TempEst v1.5.3[43], taking an ML tree inferred by IQ-TREE and times of sample collection as input. All phylogenetic trees in this study were visualised with the ggtree package[44].

Epidemiological and spatial information of HPAI H5 viruses in Taiwan

The Taiwanese government has been actively conducting virological surveillance of HPAI H5 viruses in poultry farms and abandoned bird carcasses prior to 2015[28,39]. Information on the collection date, host species and subtypes of each identification report can be found on a publicly available website[24]. The

sampling sites' geographical coordinates for the reported events were obtained by parsing an interactive map embedded on the same website using the RSelenium package[45]. Metadata for the genetic data, including subtype, time, location at the county/city level and host, were acquired along with the sequences. Information on the sampling environment types, i.e., farms or slaughterhouses, were obtained from a recent publication by the Taiwanese government[21].

Continuous phylogeographic analyses

We adopted a recently developed method to account for uncertainty in the sampling location for evolutionary reconstructions in continuous space[46,47]. Specifically, for each sequence, a prior describing multiple polygons and their corresponding probabilities was defined in a Keyhole Markup Language (KML) file, and thereby incorporated into the Bayesian Markov chain Monte Carlo (MCMC) process performed by BEAST[48]. To construct the priors in this study, candidate locations for each genetic sequence were defined based on the sampling sites of the surveillance data. The candidate locations were gathered by matching the sequence's metadata, i.e. host, subtype, and county/city, to the reported events, along with a dynamic time-interval starting with (1) the collection date; if no hit was found with compatible information and in the time-interval, (2) the collection date +-7 was used. After conducting two searches, the centroids of all subdivisions in the county/city (i.e. town or area) of the sample were assigned to each unmatched sequence. The probabilities of a prior were uniform for sequences that were compatible with events in the surveillance data. However, for sequences that we were unable to link to reported events, the probabilities were assigned proportionally to the incidence of H5 HPAI during 2015-2019 in subdivisions. Each site with geographical coordinates was expanded to a minimum area of latitude +-0.0003 and longitude +-0.0005 to fit the polygon format in the original approach[46]. A sample collected in a slaughterhouse or rendering factory was assigned to the location of one arbitrarily selected

473 slaughterhouse for each county/city based on the registry data (www.baphiq.gov.tw, accessed 22 March
474 2023).
475
476 Continuous phylogeographic analyses were performed with the Cauchy Relaxed Random Walk
477 (RRW) model implemented in BEAST v1.10.4[25,49]. A general time reversible (GTR) substitution model,
478 a gamma-distributed rate and an uncorrelated lognormal relaxed molecular clock[50] were employed in the
479 Bayesian analyses, along with the Skygrid demographic model[51]. The MCMC was run for 550 million
480 steps, with samples taken every 10,000 steps after removing 50 million steps as burn-in. Parallel runs were
481 conducted to confirm the observed pattern. A starting tree was added to each run to facilitate the sampling
482 of tree topologies. The results of continuous phylogeographic analyses were visualised with the
483 SERAPHIM package[52].
484
485 Discrete phylogeographic analyses with an epoch model
486 To evaluate differences in spatial dispersal patterns throughout different epidemiological phases, from the
487 introduction to the development of endemicity, we employed a time-heterogeneous transition rate model
488 with the discrete phylogeographic method[53]. The method enables the Bayesian discrete state inference to
489 accommodate multiple rate matrices, each of which specified by one time interval (epoch). Therefore,
490 transition frequencies and statistically supported directions in different intervals can be simultaneously
491 estimated on the same phylogeny. We defined two time intervals, T1 and T2, divided by the date 1
492 September, 2016, based on sample sizes of the sequence data and the epidemiological trajectory (Figure
493 2). Transition parameters were estimated using an asymmetric substitution model with Bayesian stochastic
494 search variable selection (BSSVS)[54].
495

To better approximate the locations of the original farms, uncertain discrete state assignment was applied to the samples collected in slaughterhouses or rendering factories[55]. Similar to the strategy we used for the analyses in continuous space, we informed the priors of the uncertain discrete locations with surveillance data. We gathered candidate locations at the county/city level by matching the sequence's metadata to the reports within a 15-day interval (collection date +-7). If no matched events were found, we extended the interval to 90 days. The probabilities linked to the locations were proportional to the corresponding events in different counties and cities. The dispersal between hosts was also assessed by classifying sequences into the order of Anseriformes or Galliformes.

Additionally, we used a generalised linear model (GLM) incorporated into the discrete phylogenetic framework to investigate the potential predictors' contribution to the transition between locations[56]. The predictors included (1) road distance, (2) shared border, (3) number of poultry farms, (4) poultry population, (5) poultry heterogeneity, (6) area of cropland and (7) number of viral sequences. Poultry farm, poultry heterogeneity and cropland area were added to the model as covariates for both origin and destination. Poultry farm, poultry population, poultry heterogeneity and sample size were treated as time-heterogeneous by averaging the annual values in the two intervals (2015-2016 and 2016-2019)(Supplementary Figure 2), while all predictors were log-transformed and standardised before inclusion in the matrices. The GLM model was adapted using previously described methods to assume temporally heterogeneous effect sizes and inclusion probabilities[13,57]. To account for unexplained variability, models that included time-homogeneous random effect variables specifying the effects as both origin and destination of each location were also implemented[58,59]. To validate the results, we performed reduced GLM models with no more than two predictors and a time-homogeneous GLM model including all the aforementioned predictors.

519

520 BEAST v1.10.5 (pre-released v0.1.2) was used to implement time-heterogeneous rate models with
521 MCMC chains of 110 million steps[48]. Performance of the GLM estimations were improved by applying a
522 set of empirical trees. All BEAST runs were facilitated by the BEAGLE library[60]. Convergence was
523 examined using Tracer v1.7.2[61], ensuring that effective sampling sizes (ESS) were greater than 200 for all
524 continuous parameters.

525

526 Posterior analyses. The number of transitions between the two states was quantified by Markov
527 jumps[62]. Bayes factors (BF) were calculated based on the indicators of transition rates resulting from
528 BSSVS to determine statistically supported diffusion routes. Routes with BF>20 were considered
529 supported and were classified into two categories: 20-150 and >150, which are typically referred to as
530 'strong' and 'very strong'[63]. The duration of viral persistence was evaluated using maximised time intervals
531 that were unified by branch lengths, where both nodes were inferred to be in the same location. Trunk
532 probabilities over time were summarised by PACT v0.9.4 (https://github.com/trvrb/PACT). Trunks were
533 defined as ancestral branches shared by sequences isolated in 2019. Persistence and trunk probability were
534 calculated using 1000 MCMC trees sampled from the posterior phylogeny distributions.

535

536 Probing the spatial distribution of re-emergent outbreaks

537 Buffers representing disease hotspots on the map were created by parsing geographical coordinates
538 inferred at the nodes of the posterior trees in the continuous phylogeographic calibrations, including tips
539 and internal nodes. A buffer was defined as the area within 1 km of the inferred point location. The buffers
540 created by the phylogenetic inference were merged with or without the buffers created by the sampling
541 sites of contemporary outbreak events reported in the surveillance data for each tree. Only nodes estimated

no earlier than 1 September, 2016 (i.e. during the endemic phase, T2) and the outbreaks during the same period were taken as central points to create buffers. The nodes and the outbreaks were stratified by county/city before being transformed from points to polygons. Coverage ("captured") was determined by the sites of new outbreaks covered by the merged buffer areas. New outbreaks were defined as outbreaks occurring after the latest genetic sample, during 2019-2022. The captured outbreaks were summarised as proportions with 1000 posterior trees, whose node numbers were also used to generate randomly distributed null models. To evaluate the coverage performance, sensitivity analyses were conducted on buffer size considering multiple radius distances, including 0.5, 1 and 2 km. The spatial data manipulation was performed using the *sf* package[64]. The map data (shp files) were obtained from the government open data platform (data.gov.tw).

Data sources of potential predictors

Road distances between counties and cities in Taiwan were measured with Google Maps. Data on agricultural statistics, including the number of poultry farms, poultry population and cropland area during 2015-2019 were obtained from the Agricultural Statistics Database of the Ministry of Agriculture (https://agrstat.moa.gov.tw/sdweb). Poultry heterogeneity was calculated as Simpson's Index of Diversity[65].

Selection analyses

SLAC[66] and MEME[67] in the Datamonkey platform[68] were used to detect pervasive and episodic positive selection sites, respectively, with a default statistical significance level of $p<0.1$. HyPhy v.2.3.14[69] was used to calculate non-synonymous/synonymous rate ratio, dN/dS (ω) estimates. The results of the surface

proteins in the clade 2.3.4.4c viruses were compared with those in an enzootic, non-GsGd H5N2 lineage previously described in Taiwan[20,29].

Data and code availability

The custom codes used in anaylses, and the XML files required for BEAST can be found at https://github.com/yaotli/endemic. The accession numbers of the Taiwan sequences analysed in our study are listed in Supplementary Table 2.

**Acknowledgements**

The work was supported by the UK Research and Innovation Global effort on COVID-19 (MR/V035444/1), Wellcome (207569/Z/17/Z, 224520/Z/21/Z), , the UK Medical Research Council New Investigator Research Grant (MR/X002047/1) and a University of Glasgow Lord Kelvin/Adam Smith Fellowship to KB. We are grateful for the continued surveiilnace efforts of the Animal and Plant Health Inspection Agency, MOA, Taiwan. We acknowledge the authors, originating laboratories, and submitting entities responsible for providing the sequences obtained from GISAID.

## References

1. Karesh, W. B. *et al.* Ecology of zoonoses: natural and unnatural histories. *The Lancet* **380**, 1936–1945 (2012).
2. Guan, Y. & Smith, G. J. D. The emergence and diversification of panzootic H5N1 influenza viruses. *Virus Research* **178**, 35–43 (2013).
3. Smith, G. J. D., Donis, R. O., & World Health Organization/World Organisation for Animal Health/Food and Agriculture Organization (WHO/OIE/FAO) H5 Evolution Working Group. Nomenclature updates resulting from the evolution of avian influenza A(H5) virus clades 2.1.3.2a, 2.2.1, and 2.3.4 during 2013–2014. *Influenza and Other Respiratory Viruses* **9**, 271–276 (2015).
4. Li, Y.-T., Su, Y. C. F. & Smith, G. J. D. H5Nx Viruses Emerged during the Suppression of H5N1 Virus Populations in Poultry. *Microbiology Spectrum* **9**, e01309-21 (2021).
5. *Approaches to Controlling, Preventing and Eliminating H5N1 Highly Pathogenic Avian Influenza in Endemic Countries*. (FAO, Rome, 2011).
6. Vijaykrishna, D. *et al.* Evolutionary Dynamics and Emergence of Panzootic H5N1 Influenza Viruses. *PLOS Pathogens* **4**, e1000161 (2008).
7. Tian, H. *et al.* Avian influenza H5N1 viral and bird migration networks in Asia. *Proceedings of the National Academy of Sciences* **112**, 172–177 (2015).
8. Trovão, N. S., Suchard, M. A., Baele, G., Gilbert, M. & Lemey, P. Bayesian Inference Reveals Host-Specific Contributions to the Epidemic Expansion of Influenza A H5N1. *Molecular Biology and Evolution* **32**, 3264–3275 (2015).
9. Medina, R. A. & García-Sastre, A. Influenza A viruses: new research developments. *Nat Rev Microbiol* **9**, 590–603 (2011).
10. Saito, T. *et al.* Intracontinental and intercontinental dissemination of Asian H5 highly pathogenic

avian influenza virus (clade 2.3.4.4) in the winter of 2014–2015. *Reviews in Medical Virology* **25**, 388–405 (2015).

11. WHO. Antigenic and genetic characteristics of zoonotic influenza A viruses and development of candidate vaccine viruses for pandemic preparedness (February, 2020). (2020).
12. Cruz, C. D. *et al.* Highly Pathogenic Avian Influenza A(H5N1) from Wild Birds, Poultry, and Mammals, Peru. *Emerg Infect Dis* **29**, 2572–2576 (2023).
13. Zhang, G. *et al.* Bidirectional Movement of Emerging H5N8 Avian Influenza Viruses Between Europe and Asia via Migratory Birds Since Early 2020. *Molecular Biology and Evolution* **40**, msad019 (2023).
14. Agüero, M. *et al.* Highly pathogenic avian influenza A(H5N1) virus infection in farmed minks, Spain, October 2022. *Eurosurveillance* **28**, 2300001 (2023).
15. Nguyen, T.-Q. *et al.* Emergence and interstate spread of highly pathogenic avian influenza A(H5N1) in dairy cattle. 2024.05.01.591751 Preprint at https://doi.org/10.1101/2024.05.01.591751 (2024).
16. Pohlmann, A. *et al.* Has Epizootic Become Enzootic? Evidence for a Fundamental Change in the Infection Dynamics of Highly Pathogenic Avian Influenza in Europe, 2021. *mBio* **13**, e00609-22 (2022).
17. Lee, Y.-J. *et al.* Novel Reassortant Influenza A(H5N8) Viruses, South Korea, 2014. *Emerg Infect Dis* **20**, 1087–1089 (2014).
18. Ip, H. S. *et al.* Novel Eurasian highly pathogenic avian influenza A H5 viruses in wild birds, Washington, USA, 2014. *Emerg Infect Dis* **21**, 886–890 (2015).
19. Kanehira, K. *et al.* Characterization of an H5N8 influenza A virus isolated from chickens during an outbreak of severe avian influenza in Japan in April 2014. *Arch Virol* **160**, 1629–1643 (2015).
20. Li, Y.-T., Chen, C.-C., Chang, A.-M., Chao, D.-Y. & Smith, G. J. D. Co-circulation of both low and

highly pathogenic avian influenza H5 viruses in current poultry epidemics in Taiwan. *Virus Evolution* **6**, veaa037 (2020).

21. Huang, C.-W. *et al.* Evolutionary history of H5 highly pathogenic avian influenza viruses (clade 2.3.4.4c) circulating in Taiwan during 2015–2018. *Infection, Genetics and Evolution* **92**, 104885 (2021).
22. Chang, C.-F. *et al.* Lessons from the Largest Epidemic of Avian Influenza Viruses in Taiwan, 2015. *Avian Diseases* **60**, 156–171 (2016).
23. Huang, P.-Y. *et al.* Genetic characterization of highly pathogenic H5 influenza viruses from poultry in Taiwan, 2015. *Infection, Genetics and Evolution* **38**, 96–100 (2016).
24. BAPHIQ. Epidemic Situation of Avian Influenza. https://twai.baphiq.gov.tw/AI/ accessed 22 Feb 2023 (2023).
25. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography Takes a Relaxed Random Walk in Continuous Space and Time. *Molecular Biology and Evolution* **27**, 1877–1885 (2010).
26. Dellicour, S. *et al.* Relax, Keep Walking — A Practical Guide to Continuous Phylogeographic Inference with BEAST. *Molecular Biology and Evolution* **38**, 3486–3493 (2021).
27. Liang, W.-S. *et al.* Ecological factors associated with persistent circulation of multiple highly pathogenic avian influenza viruses among poultry farms in Taiwan during 2015-17. *PLOS ONE* **15**, e0236581 (2020).
28. Lee, M.-S. *et al.* Highly pathogenic avian influenza viruses H5N2, H5N3, and H5N8 in Taiwan in 2015. *Veterinary Microbiology* **187**, 50–57 (2016).
29. Lee, C.-C. D. *et al.* Emergence and Evolution of Avian H5N2 Influenza Viruses in Chickens in Taiwan. *Journal of Virology* **88**, 5677–5686 (2014).
30. Yang, Q. *et al.* Assessing the role of live poultry trade in community-structured transmission of

avian influenza in China. *Proceedings of the National Academy of Sciences* **117**, 5949–5954 (2020).

31. Hicks, J. T. *et al.* Agricultural and geographic factors shaped the North American 2015 highly pathogenic avian influenza H5N2 outbreak. *PLOS Pathogens* **16**, e1007857 (2020).
32. Haydon, D. T., Cleaveland, S., Taylor, L. H. & Laurenson, M. K. Identifying reservoirs of infection: a conceptual and practical challenge. *Emerging infectious diseases* **8**, 1468–1473 (2002).
33. Bahl, J. *et al.* Ecosystem Interactions Underlie the Spread of Avian Influenza A Viruses with Pandemic Potential. *PLOS Pathogens* **12**, e1005620 (2016).
34. Bao, Y. *et al.* The influenza virus resource at the National Center for Biotechnology Information. *J Virol* **82**, 596–601 (2008).
35. Shu, Y. & McCauley, J. GISAID: Global initiative on sharing all influenza data – from vision to reality. *Eurosurveillance* **22**, 30494 (2017).
36. Aksamentov, I., Roemer, C., Hodcroft, E. B. & Neher, R. A. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *Journal of Open Source Software* **6**, 3773 (2021).
37. Kuraku, S., Zmasek, C. M., Nishimura, O. & Katoh, K. aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Research* **41**, W22–W28 (2013).
38. The Global Consortium for H5N8 and Related Influenza Viruses. Role for migratory wild birds in the global spread of avian influenza H5N8. *Science* **354**, 213–217 (2016).
39. Chen, L.-H. *et al.* Molecular Epidemiology in H5Nx Subtypes of Avian Influenza Virus in Taiwan in 2015. *Experimental report of national institute for animal health* **51**, 53–68 (2017).
40. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evolution* **4**, vex042 (2018).
41. Minh, B. Q. *et al.* IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the

Genomic Era. *Molecular Biology and Evolution* **37**, 1530–1534 (2020).

42. FAO. Empres-i. https://empres-i.apps.fao.org/ accessed 19 Aug 2023 (2023).
43. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* **2**, vew007 (2016).
44. 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 in Ecology and Evolution* **8**, 28–36 (2017).
45. Harrison, J., Kim, J. Y. & Völkle, J. RSelenium: R Bindings for 'Selenium WebDriver'. (2023).
46. Dellicour, S. *et al.* Incorporating heterogeneous sampling probabilities in continuous phylogeographic inference — Application to H5N1 spread in the Mekong region. *Bioinformatics* **36**, 2098–2104 (2020).
47. Dellicour, S., Lemey, P., Suchard, M. A., Gilbert, M. & Baele, G. Accommodating sampling location uncertainty in continuous phylogeography. *Virus Evolution* **8**, veac041 (2022).
48. Suchard, M. A. *et al.* Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution* **4**, vey016 (2018).
49. Pybus, O. G. *et al.* Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proceedings of the National Academy of Sciences* **109**, 15066–15071 (2012).
50. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed Phylogenetics and Dating with Confidence. *PLOS Biology* **4**, e88 (2006).
51. Gill, M. S. *et al.* Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol Biol Evol* **30**, 713–724 (2013).
52. Dellicour, S., Rose, R., Faria, N. R., Lemey, P. & Pybus, O. G. SERAPHIM: studying environmental rasters and phylogenetically informed movements. *Bioinformatics* **32**, 3204–3206

(2016).

53. Bielejec, F., Lemey, P., Baele, G., Rambaut, A. & Suchard, M. A. Inferring Heterogeneous Evolutionary Processes Through Time: from Sequence Substitution to Phylogeography. *Systematic Biology* **63**, 493–504 (2014).

54. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian Phylogeography Finds Its Roots. *PLOS Computational Biology* **5**, e1000520 (2009).

55. Scotch, M. *et al.* Incorporating sampling uncertainty in the geospatial assignment of taxa for virus phylogeography. *Virus Evolution* **5**, vey043 (2019).

56. Lemey, P. *et al.* Unifying Viral Genetics and Human Transportation Data to Predict the Global Transmission Dynamics of Human Influenza H3N2. *PLOS Pathogens* **10**, e1003932 (2014).

57. Lemey, P. *et al.* Untangling introductions and persistence in COVID-19 resurgence in Europe. *Nature* **595**, 713–717 (2021).

58. Dudas, G. *et al.* Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309–315 (2017).

59. Hong, S. L. *et al.* In Search of Covariates of HIV-1 Subtype B Spread in the United States—A Cautionary Tale of Large-Scale Bayesian Phylogeography. *Viruses* **12**, 182 (2020).

60. Ayres, D. L. *et al.* BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics. *Systematic Biology* **68**, 1052–1061 (2019).

61. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology* **67**, 901–904 (2018).

62. Minin, V. N. & Suchard, M. A. Counting labeled transitions in continuous-time Markov models of evolution. *J Math Biol* **56**, 391–412 (2008).

63. Kass, R. E. & Raftery, A. E. Bayes Factors. *Journal of the American Statistical Association* **90**, 773–795 (1995).

64. Pebesma, E. Simple Features for R: Standardized Support for Spatial Vector Data. *The R Journal* **10**, 439–446 (2018).

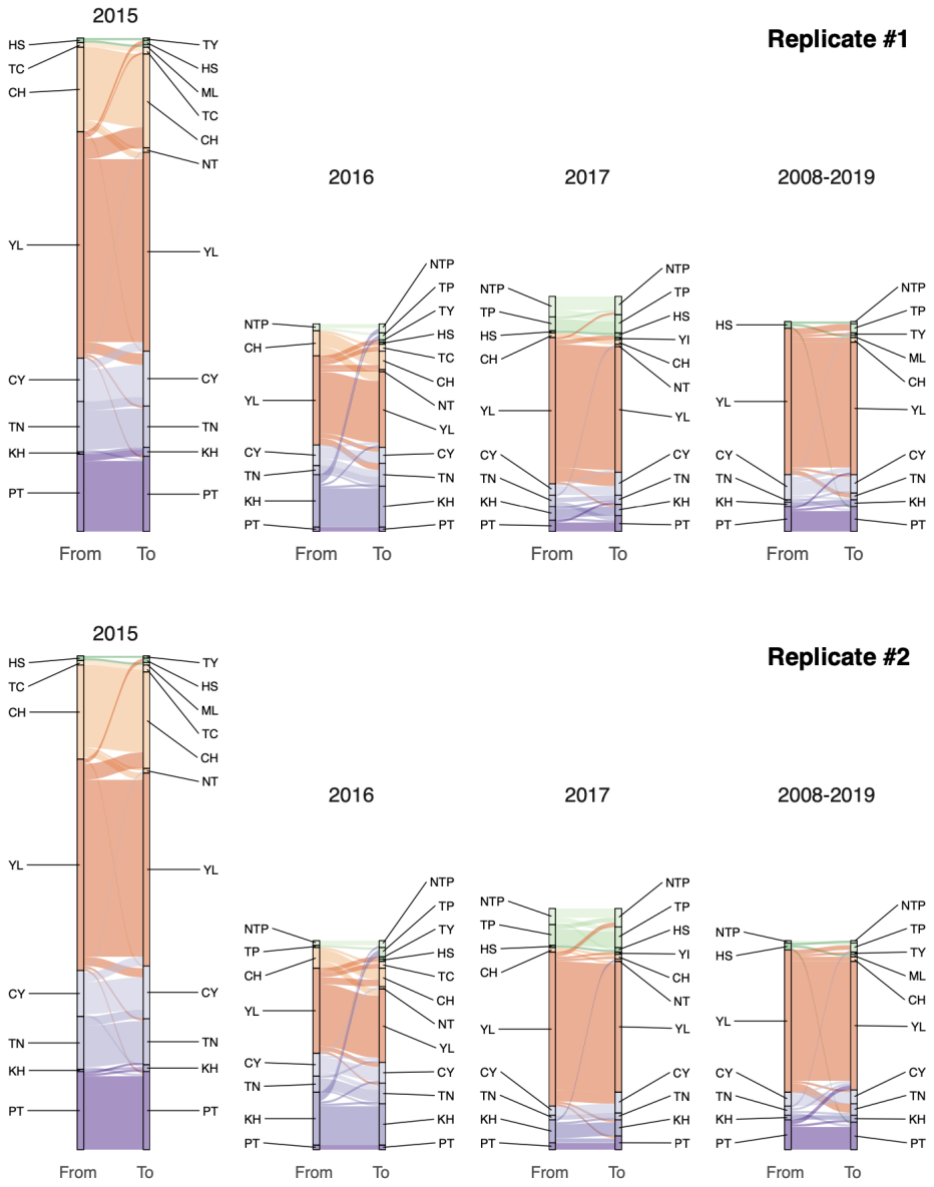
65. Simpson, E. H. Measurement of Diversity. *Nature* **163**, 688–688 (1949).

66. Kosakovsky Pond, S. L. & Frost, S. D. W. Not So Different After All: A Comparison of Methods for Detecting Amino Acid Sites Under Selection. *Molecular Biology and Evolution* **22**, 1208–1222 (2005).

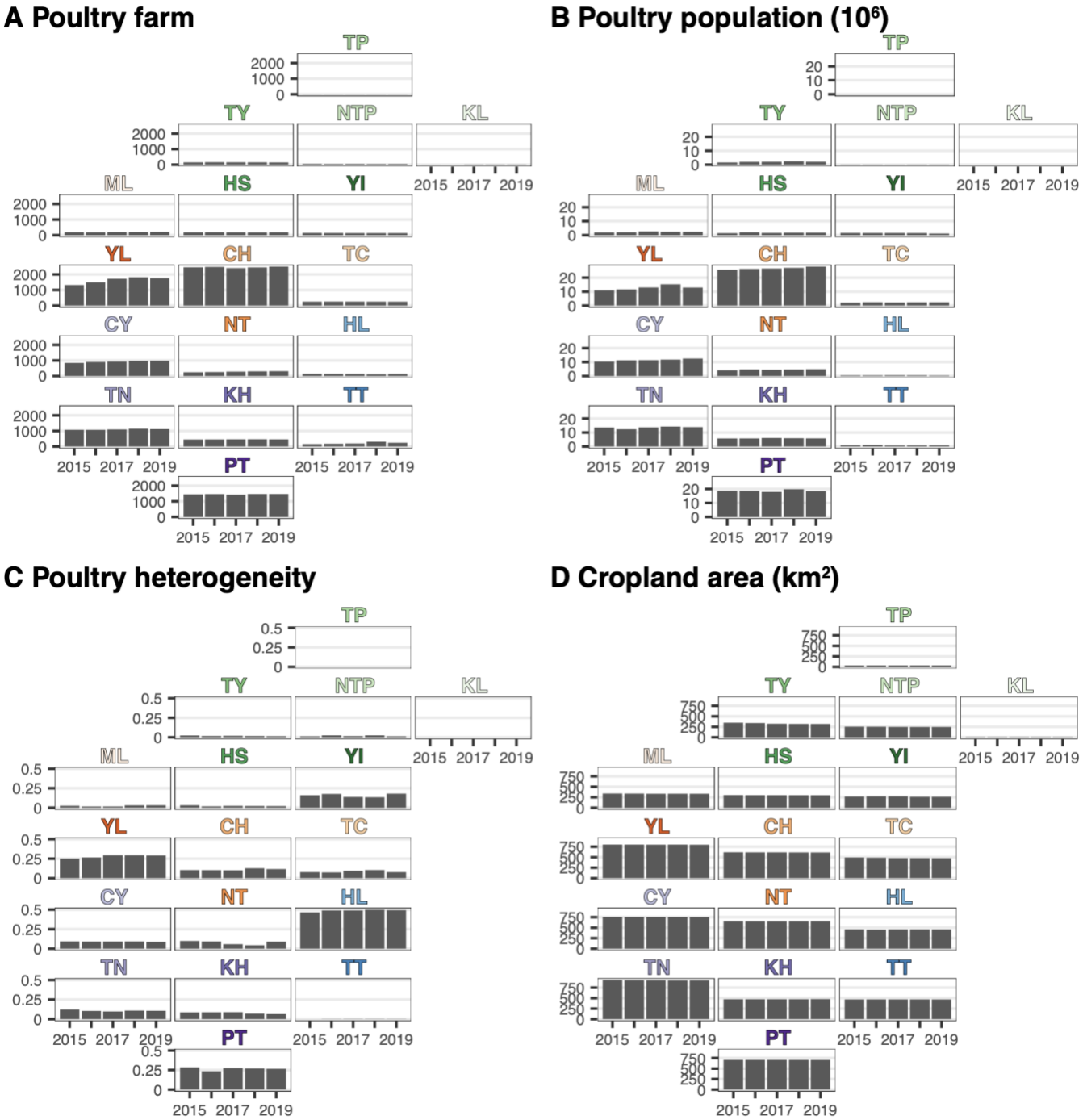
67. Murrell, B. *et al.* Detecting Individual Sites Subject to Episodic Diversifying Selection. *PLOS Genetics* **8**, e1002764 (2012).

68. Weaver, S. *et al.* Datamonkey 2.0: A Modern Web Application for Characterizing Selective and Other Evolutionary Processes. *Molecular Biology and Evolution* **35**, 773–777 (2018).

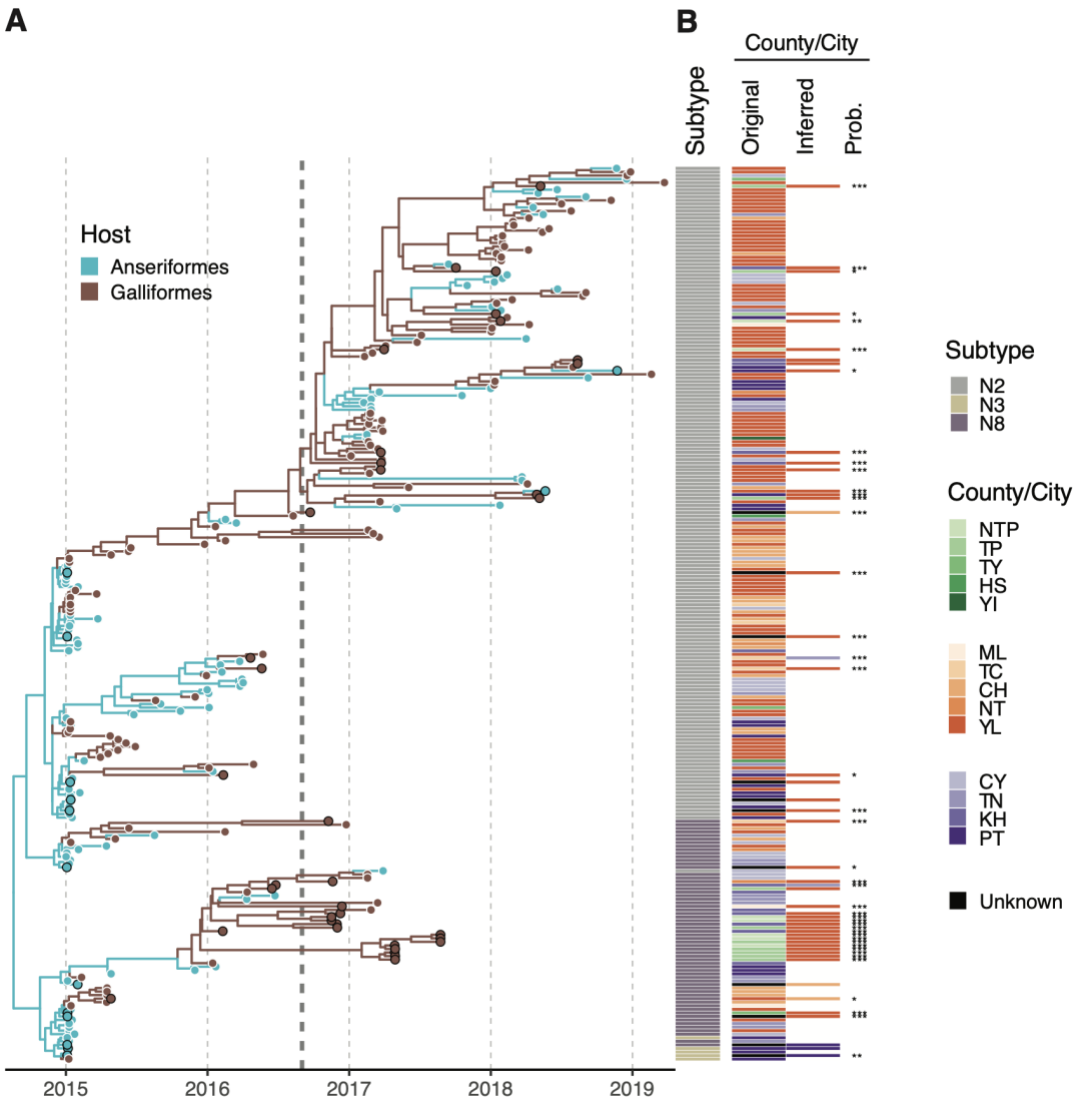
69. Pond, S. L. K., Frost, S. D. W. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).




**Supplementary figure 1.** Visualisation of the dispersal pattern of the clade 2.3.4.4c virus in discrete geographical divisions. The connected lines represent branches inferred by the continuous phylogeographic approach (see Figure 3), with origin and destination locations determined as county or city-level divisions. The lines were coloured based on the origin. Replicated results from parallel runs are shown.

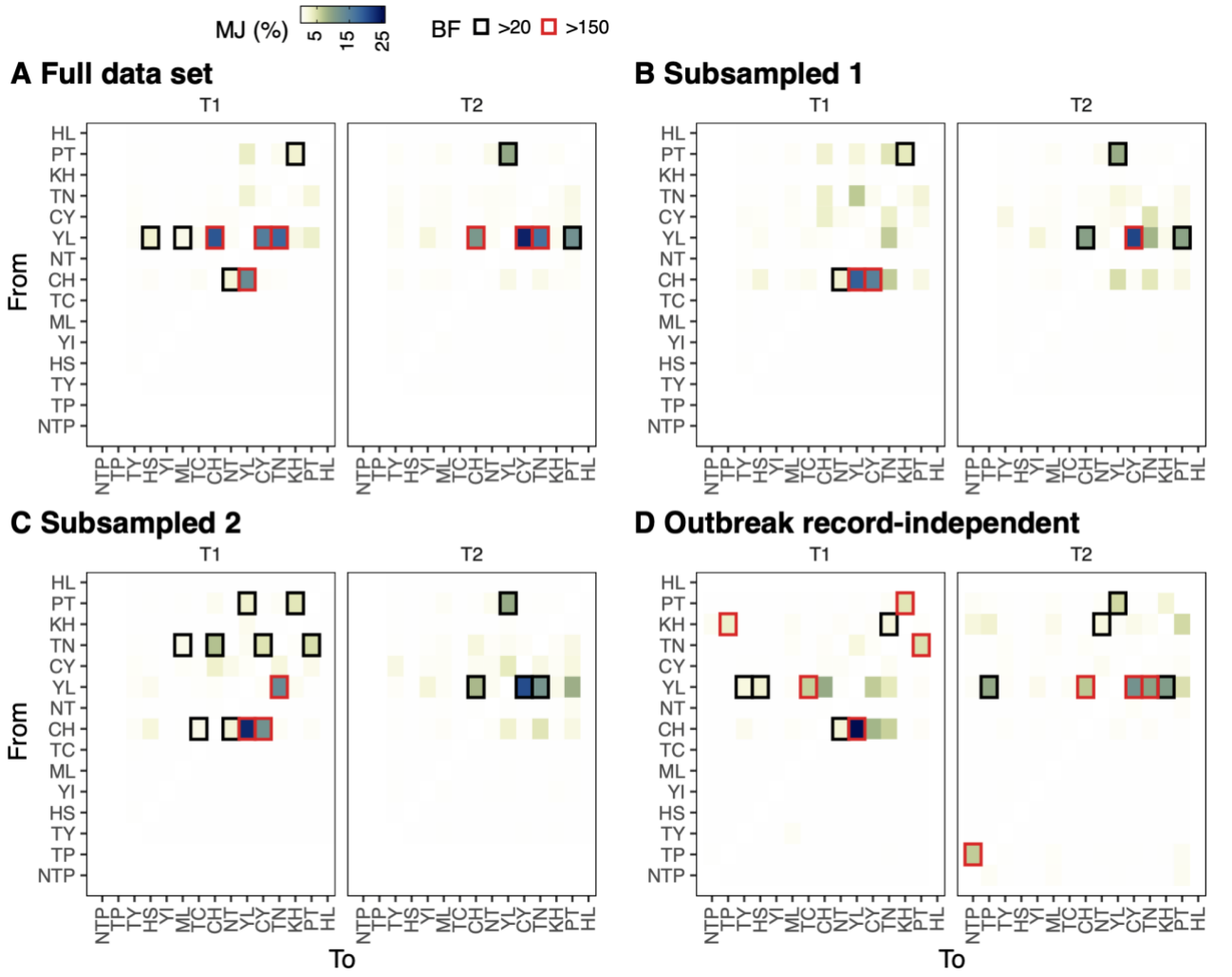



**Supplementary figure 2.** Agricultural characteristics of county or city-level administrative areas in Taiwan.




**Supplementary figure 3.** Summary of the host and geographical information associated with the genomic data used in this study. (A) The time-scaled phylogeny was reconstructed using the HA genes of the clade 2.3.4.4c virus in Taiwan. The tips on the phylogeny are colored according to the host genera, Galliformes (chicken, turkey and quail) or Anseriformes (duck and goose), logged in the sequence metadata. The branches are colored according to the states inferred on the summarized maximum clade credibility (MCC) tree. The tips with black borders indicate taxa that have uncertain state assignments in geographical analyses. The vertical dashed line denotes September 1st, 2016. (B) The adjacent heatmap shows the NA

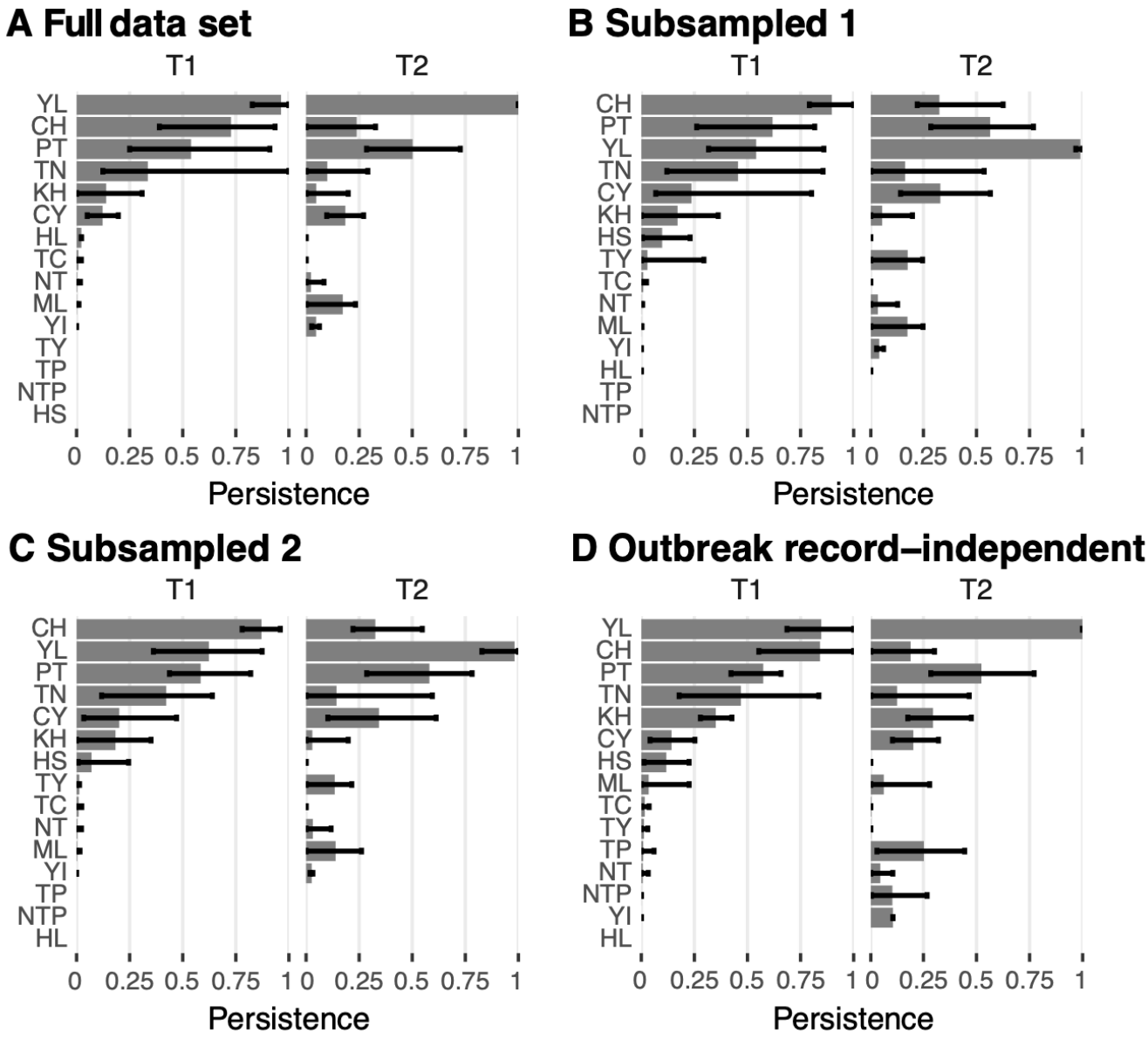
760 subtypes and the corresponding collection locations of the tree taxa. The discrete diffusion model was
761 used to estimate the uncertain geographical locations and posterior probabilities. Only the results of
762 samples collected in slaughterhouses/rendering factories and samples lacking geospatial information,
763 which were implemented with state uncertainty, are shown in the 'Inferred' column. The symbols *, **
764 and *** denote posterior probability>0.7, 0.8 and 0.9, respectively.
765
766
767
768
769
770
771



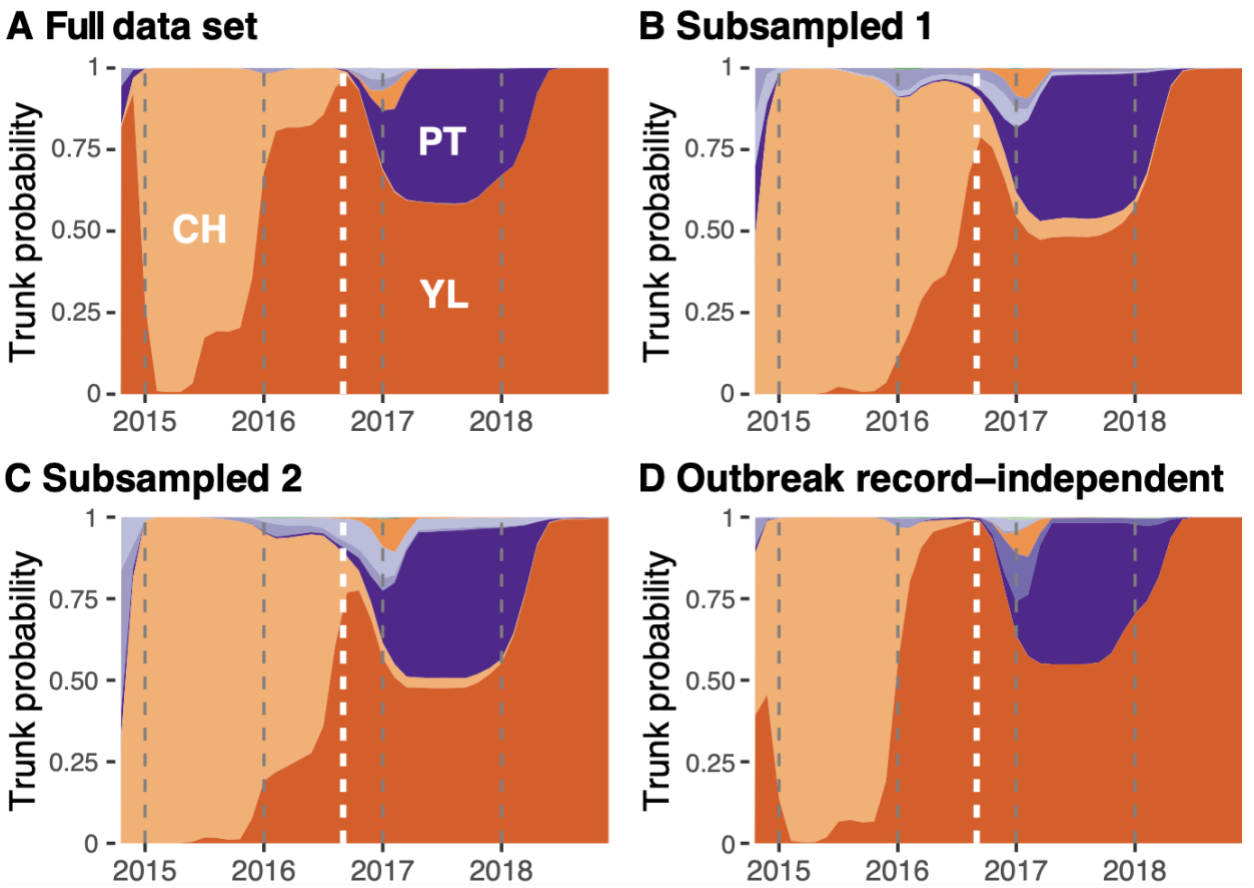

772

773 **Supplementary figure 4.** Estimated number of transition events between counties/cities in Taiwan
774 inferred by the discrete phylogeographic method. The color of the heatmap reflects the proportions of
775 between-locations Markov jumps to the total jumps in the distinct time period. The directions supported
776 by Bayesian factor (BF) are highlighted by black (>20) or red (>150) edges. Results of a full genomic data
777 set (A), along with two data sets with reduced samples from county YL are shown (B and C). Using the
778 same genomic data as panel (A), Bayesian inference was performed for panel (D) with geographical states
779 assigned independently of outbreak records. The locations were determined by the sequence metadata if
780 available, or using a uniform prior over all isolated counties/cities if the collection site was unknown.

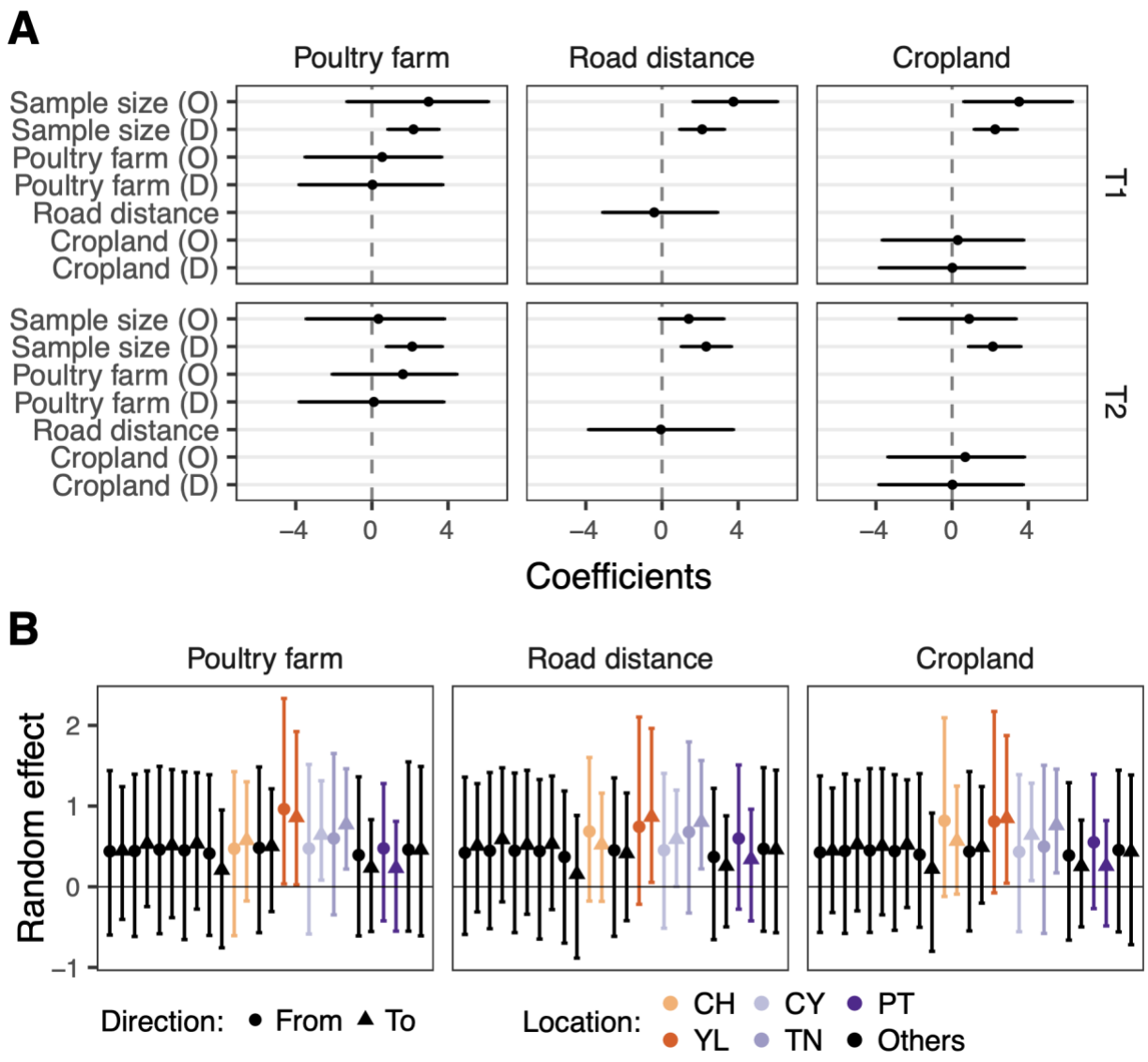
781



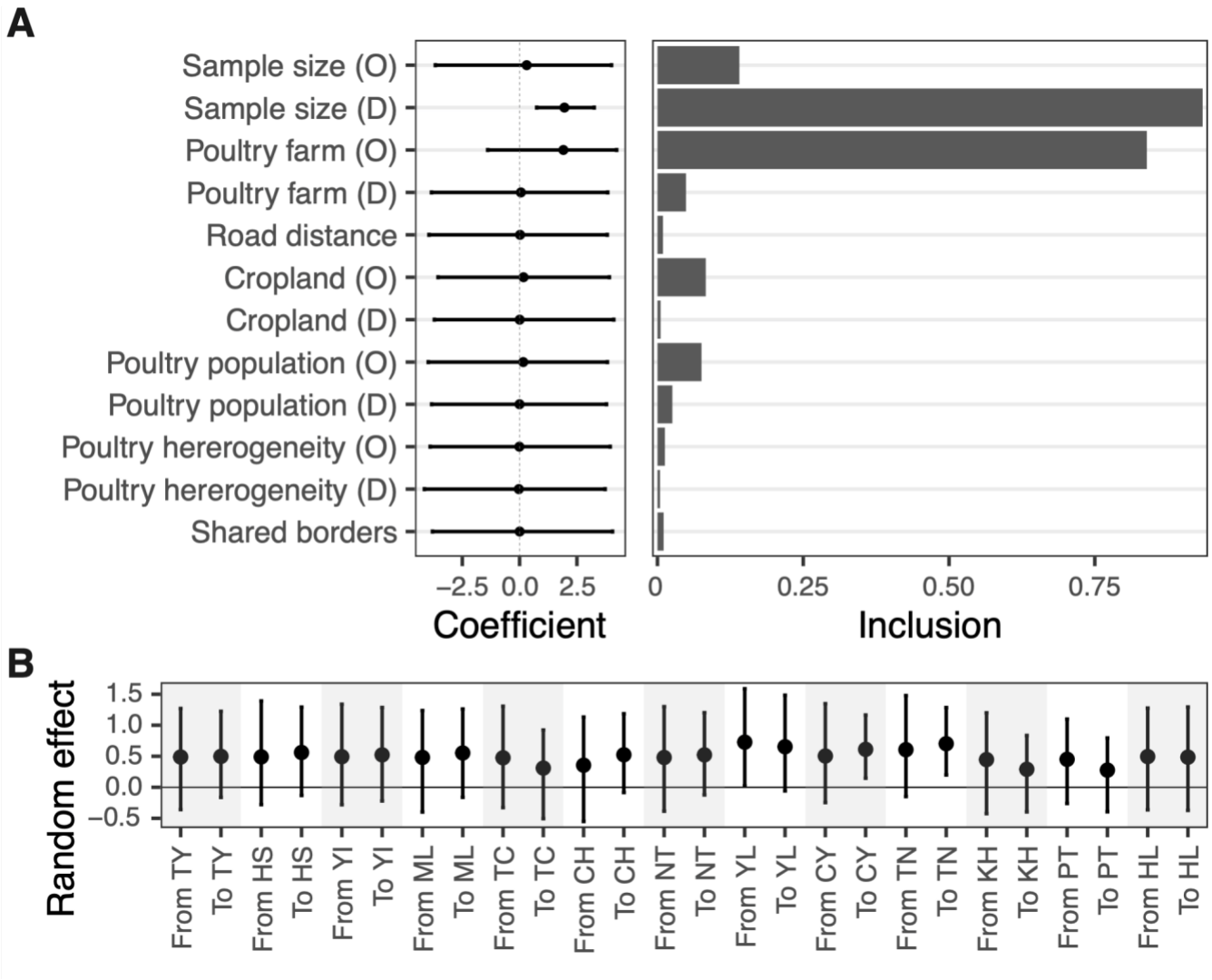

**Supplementary figure 5.** Persistence of the virus assessed by the discrete phylogeographic method. To calculate the value for each location, the branch length was unified from branches where both nodes were estimated to share the same state. The unified time interval was presented as a proportion divided by the time span of T1/T2. The mean and 95% HPD of the proportions summarized by 1000 posterior trees are indicated by the bars. The results of the four panels were obtained from the same MCMC runs as the corresponding panels in Supplementary Figure 4.



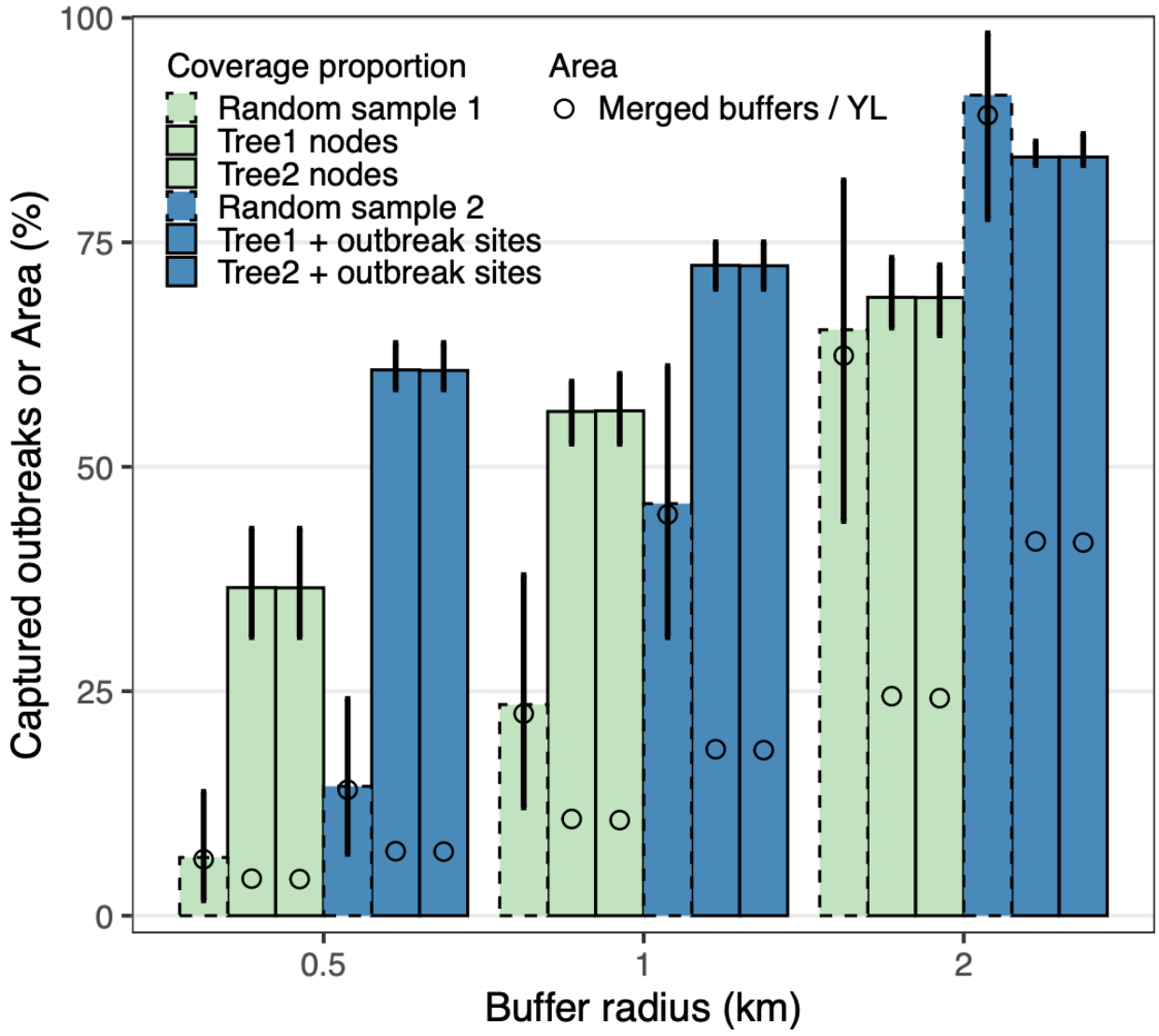

**Supplementary figure 6.** Inferred trunk locations of Taiwan clade 2.3.4.4c phylogenies through time in different genetic data sets (A-C) or with a simple geographic state assignment scheme (D). The proportions at each time point indicate the posterior support for viruses circulating in a particular county/city occupying the trunk of the tree. The areas are colored using the same scheme as Figure 2. The white dashed lines indicate the boundary of T1 and T2, September 1st, 2016.




**Supplementary figure 7.** Evaluating predictors in reduced models of time-heterogeneous phylogenetic GLM. (A) The conditional effect sizes in models containing respective predictors. The predictor names are denoted by O in parentheses for origin and D for destination. (B) Location-specific random effects in the three models, with high incidence areas colored. The estimates are in log space and presented as mean with 95% HPD interval.

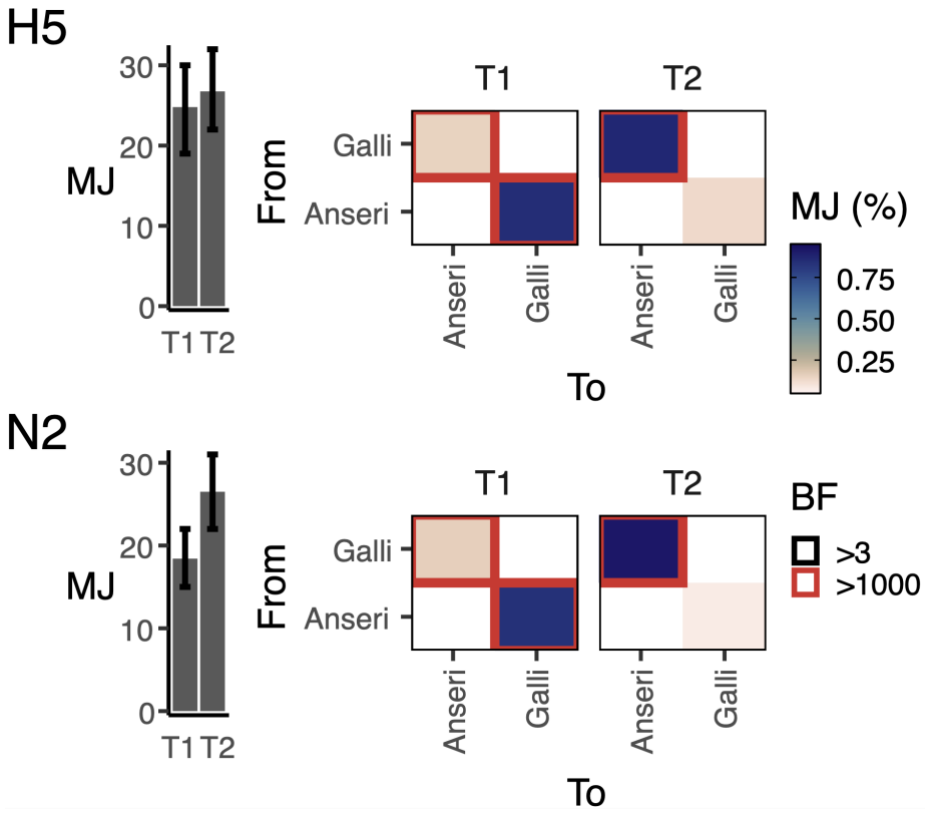



**Supplementary figure 8.** Evaluating predictors in a time-homogeneous model. (A) The conditional effect sizes and the inclusion probabilities of predictors estimated by the conventional phylogenetic GLM models. The predictor names are denoted by O in parentheses for origin and D for destination. (B) Location-specific random effects in the GLM model. The estimates are in log space and presented as mean with 95% HPD interval.




**Supplementary figure 9.** Impact of buffer radius on evaluating the re-emergence of new outbreaks in county YL. The estimated locations of tree nodes using the continuous phylogenetic method during T2, combined with or without contemporary outbreak sites, served as central points to create buffers with different radius distances on the map. The proportion of new outbreak sites covered by the buffer areas was calculated for each posterior tree. These outbreak sites were reported between 2019 and 2022, after the latest available genetic data. Null models were created by randomly distributing sites in YL with the same number of tree nodes (Random sample 1, green) or tree nodes plus outbreak sites (Random sample

2, blue). The bars show the mean values of two parallel continuous phylogeographic analyses each with 1000 posterior trees, and null models. Error bars indicate 95% credible intervals. The open circle in each bar represents the ratio of the area of merged buffers to the area of YL.




841

842

843 **Supplementary figure 10.** Diffusion between host groups. Markov jumps (MJ) and Bayes factors (BF)
844 were estimated by the discrete phylogenetic method. Total MJ counts in different time spans are shown
845 on the left. The heatmaps present the jumps as proportions to the total counts. Anseri refers to the
846 Anseriformes genus and Galli refers to the Galliformes genus. Results based on both HA (H5) and NA
847 (N2) are shown.

848

849

850

851

852

853

**Supplementary Table1.** Selection pressure for the surface proteins of H5 avian influenza lineages in Taiwan.

| | H5 | | N2 | |
|---|---|---|---|---|
| | GsGd Clade 2.3.4.4 | North American | GsGd Clade 2.3.4.4 | North American |
| dN/dS (ω) (95% CI) | 0.204 (0.180-0.230) | 0.207 (0.174-0.245) | 0.355 (0.316-0.397) | 0.233 (0.192-0.279) |
| Pervasive positively selected sites | 131 | - | 77 | - |
| Episodic positively selected sites | 131, 387, 403, 545 | 17, 102, 137, 155, 317 | 77, 313, 416, 439 | 5, 6, 261, 307 |
| No. of fixed sites among selected sites | 0 | 0 | 0 | 0 |

871 **Supplementary Table 2.** List of database accession numbers from GenBank and GISAID.
872

| | accession no. | strain name |
|---|---|---|
| 1 | MW333030 | duck/Nantou/15A03445-11-20T/2015 |
| 2 | MW333038 | chicken/Changhua/15010120-1/2015 |
| 3 | MW333046 | chicken/Changhua/15010288/2015 |
| 4 | MW333054 | chicken/Changhua/15010324-2/2015 |
| 5 | MW333062 | chicken/Changhua/15010361-1/2015 |
| 6 | MW333070 | chicken/Changhua/15010362-2/2015 |
| 7 | MW333078 | chicken/Changhua/15010363-2/2015 |
| 8 | MW333086 | chicken/Changhua/15040036-1/2015 |
| 9 | MW333094 | chicken/Changhua/15040043-1/2015 |
| 10 | MW333102 | chicken/Changhua/15040037-2/2015 |
| 11 | MW333110 | chicken/Changhua/15040042/2015 |
| 12 | MW333118 | chicken/Changhua/15050013-1/2015 |
| 13 | MW333126 | chicken/Changhua/15050038-1/2015 |
| 14 | MW333134 | chicken/Changhua/15060027-2/2015 |
| 15 | MW333142 | chicken/Changhua/15060029/2015 |
| 16 | MW333150 | chicken/Changhua/15060037-1/2015 |
| 17 | MW333158 | chicken/Changhua/15120008/2015 |
| 18 | MW333166 | chicken/Changhua/16010060-2/2016 |
| 19 | MW333174 | chicken/Changhua/16020042-2/2016 |
| 20 | MW333182 | chicken/Changhua/17030063-1/2017 |
| 21 | MW333190 | chicken/Changhua/17060006/2017 |
| 22 | MW333198 | chicken/Changhua/18010011-1/2018 |
| 23 | MW333206 | chicken/Changhua/18040008-1/2018 |
| 24 | MW333214 | chicken/Chiayi/15010287-2/2015 |
| 25 | MW333222 | chicken/Chiayi/15010409-2/2015 |
| 26 | MW333230 | chicken/Chiayi/15010454-1/2015 |
| 27 | MW333238 | chicken/Chiayi/17030059-2/2017 |
| 28 | MW333246 | chicken/Chiayi/17040009-2/2017 |
| 29 | MW333254 | chicken/Chiayi/18120003-1/2018 |
| 30 | MW333262 | chicken/Hsinchu/16080004-2/2016 |
| 31 | MW333270 | chicken/Kaohsiung/16010047/2016 |
| 32 | MW333278 | chicken/Kaohsiung/16020017-2/2016 |
| 33 | MW333294 | chicken/Kaohsiung/16070003/2016 |
| 34 | MW333302 | chicken/Kaohsiung/16120012/2016 |
| 35 | MW333310 | chicken/Kaohsiung/16120034/2016 |
| 36 | MW333326 | chicken/Kaohsiung/17030023/2017 |
| 37 | MW333334 | chicken/Kaohsiung/17040006-2/2017 |
| 38 | MW333342 | chicken/Kaohsiung/17040007/2017 |
| 39 | MW333350 | chicken/Kaohsiung/17100002/2017 |
| 40 | MW333358 | chicken/Kaohsiung/18080004-1/2018 |
| 41 | MW333366 | chicken/Kaohsiung/18080004-2/2018 |
| 42 | MW333374 | chicken/Miaoli/15010421-1/2015 |
| 43 | MW333382 | chicken/Miaoli/18040011-1/2018 |
| 44 | MW333390 | chicken/Nantou/16110023-1/2016 |
| 45 | MW333398 | chicken/New_Taipei_City/16110019-1/2016 |
| 46 | MW333406 | chicken/New_Taipei_City/16110020-2/2016 |
| 47 | MW333414 | chicken/New_Taipei_City/17040002-2/2017 |
| 48 | MW333422 | chicken/New_Taipei_City/17050004-2/2017 |
| 49 | MW333430 | chicken/New_Taipei_City/17090001/2017 |
| 50 | MW333454 | chicken/New_Taipei_City/18010021-3/2018 |
| 51 | MW333462 | chicken/Pingtung/15010402/2015 |
| 52 | MW333470 | chicken/Pingtung/15050006-1/2015 |
| 53 | MW333478 | chicken/Pingtung/16020009-3/2016 |
| 54 | MW333486 | chicken/Pingtung/18010017-1/2018 |
| 55 | MW333494 | chicken/Pingtung/18010019-2/2018 |
| 56 | MW333502 | chicken/Pingtung/18020015-1/2018 |
| 57 | MW333510 | chicken/Pingtung/18050006-1/2018 |
| 58 | MW333518 | chicken/Pingtung/18080001-1/2018 |
| 59 | MW333526 | chicken/Taichung/15010399/2015 |
| 60 | MW333534 | chicken/Taichung/16040032-3/2016 |
| 61 | MW333542 | chicken/Taichung/16050044-2/2016 |
| 62 | MW333550 | chicken/Taichung/16110017-3/2016 |
| 63 | MW333558 | chicken/Tainan/15020214/2015 |
| 64 | MW333566 | chicken/Tainan/16040035-1/2016 |
| 65 | MW333574 | chicken/Tainan/17030058-1/2017 |
| 66 | MW333582 | chicken/Tainan/18040014-2/2018 |

| | | |
|---|---|---|
| 67 | MW333590 | chicken/Taipei_City/17090004/2017 |
| 68 | MW333598 | chicken/Taipei_City/18010007-3/2018 |
| 69 | MW333606 | chicken/Taipei_City/18010008-3/2018 |
| 70 | MW333614 | chicken/Taipei_City/18050007-1/2018 |
| 71 | MW333622 | chicken/Taipei_City/18050010-8/2018 |
| 72 | MW333646 | chicken/Taipei/16120021/2016 |
| 73 | MW333654 | chicken/Taipei/17050005-1/2017 |
| 74 | MW333662 | chicken/Taipei/17050006-2/2017 |
| 75 | MW333686 | chicken/Taipei/17050009-3/2017 |
| 76 | MW333694 | chicken/Yunlin/15020048-1/2015 |
| 77 | MW333702 | chicken/Yunlin/15040002-2/2015 |
| 78 | MW333710 | chicken/Yunlin/15040027/2015 |
| 79 | MW333718 | chicken/Yunlin/15050012/2015 |
| 80 | MW333726 | chicken/Yunlin/15050037-2/2015 |
| 81 | MW333734 | chicken/Yunlin/15050039-2/2015 |
| 82 | MW333742 | chicken/Yunlin/15060014-1/2015 |
| 83 | MW333750 | chicken/Yunlin/15080024/2015 |
| 84 | MW333758 | chicken/Yunlin/16010001-1/2016 |
| 85 | MW333766 | chicken/Yunlin/16010008-1/2016 |
| 86 | MW333774 | chicken/Yunlin/16020038/2016 |
| 87 | MW333782 | chicken/Yunlin/16050053/2016 |
| 88 | MW333790 | chicken/Yunlin/17010001/2017 |
| 89 | MW333798 | chicken/Yunlin/17010014/2017 |
| 90 | MW333806 | chicken/Yunlin/17020033-2/2017 |
| 91 | MW333814 | chicken/Yunlin/17020094-1/2017 |
| 92 | MW333822 | chicken/Yunlin/17020095-2/2017 |
| 93 | MW333830 | chicken/Yunlin/17030001-1/2017 |
| 94 | MW333838 | chicken/Yunlin/17030007-1/2017 |
| 95 | MW333846 | chicken/Yunlin/17030010-1/2017 |
| 96 | MW333854 | chicken/Yunlin/17030014-1/2017 |
| 97 | MW333862 | chicken/Yunlin/17030020-2/2017 |
| 98 | MW333870 | chicken/Yunlin/17030021-1/2017 |
| 99 | MW333878 | chicken/Yunlin/17030032-2/2017 |
| 100 | MW333886 | chicken/Yunlin/17030065-2/2017 |
| 101 | MW333894 | chicken/Yunlin/17040003-1/2017 |
| 102 | MW333902 | chicken/Yunlin/17040004-2/2017 |
| 103 | MW333910 | chicken/Yunlin/17040005-1/2017 |
| 104 | MW333918 | chicken/Yunlin/17040010-1/2017 |
| 105 | MW333926 | chicken/Yunlin/17040011-1/2017 |
| 106 | MW333934 | chicken/Yunlin/17040012-1/2017 |
| 107 | MW333942 | chicken/Yunlin/17060019/2017 |
| 108 | MW333950 | chicken/Yunlin/17070010/2017 |
| 109 | MW333958 | chicken/Yunlin/18010001-2/2018 |
| 110 | MW333966 | chicken/Yunlin/18010003-3/2018 |
| 111 | MW333974 | chicken/Yunlin/18010023-1/2018 |
| 112 | MW333982 | chicken/Yunlin/18010024-2/2018 |
| 113 | MW333990 | chicken/Yunlin/18020003-2/2018 |
| 114 | MW333998 | chicken/Yunlin/18020004-1/2018 |
| 115 | MW334006 | chicken/Yunlin/18020005-2/2018 |
| 116 | MW334014 | chicken/Yunlin/18020008-2/2018 |
| 117 | MW334022 | chicken/Yunlin/18020021-1/2018 |
| 118 | MW334030 | chicken/Yunlin/18030002-1/2018 |
| 119 | MW334038 | chicken/Yunlin/18040006-3/2018 |
| 120 | MW334046 | chicken/Yunlin/18050013-1/2018 |
| 121 | MW334054 | chicken/Yunlin/18060001-2/2018 |
| 122 | MW334062 | chicken/Yunlin/18070003-3/2018 |
| 123 | MW334070 | chicken/Yunlin/18080006-1/2018 |
| 124 | MW334078 | duck/Changhua/15A03431-1-20T/2015 |
| 125 | MW334086 | duck/Changhua/15A03473-1-20T/2015 |
| 126 | MW334094 | duck/Chiayi/15080023/2015 |
| 127 | MW334102 | duck/Hsinchu/15020216/2015 |
| 128 | MW334110 | duck/Kaohsiung/15010213/2015 |
| 129 | MW334118 | duck/Nantou/15A3437/2015 |
| 130 | MW334126 | duck/Nantou/17A0243/2017 |
| 131 | MW334134 | duck/Pingtung/15010145/2015 |
| 132 | MW334142 | duck/Pingtung/15010157-1/2015 |
| 133 | MW334150 | duck/Pingtung/15010180/2015 |
| 134 | MW334158 | duck/Pingtung/15010444/2015 |
| 135 | MW334166 | duck/Pingtung/15010461-1/2015 |
| 136 | MW334174 | duck/Pingtung/15120006/2015 |

| | | |
|---|---|---|
| 137 | MW334182 | duck/Pingtung/15A03399-1-10T/2015 |
| 138 | MW334190 | duck/Pingtung/17A00408-1-10T/2017 |
| 139 | MW334198 | duck/Pingtung/17A0160/2017 |
| 140 | MW334206 | duck/Pingtung/18A00003-11-20T/2018 |
| 141 | MW334214 | duck/Pingtung/18A00076-11-20T/2018 |
| 142 | MW334222 | duck/Pingtung/18X00077/2018 |
| 143 | MW334230 | duck/Taichung/15010400-1/2015 |
| 144 | MW334238 | duck/Tainan/15010403/2015 |
| 145 | MW334262 | duck/Taoyuan/16010023-1/2016 |
| 146 | MW334270 | duck/Yilan/17A0092/2017 |
| 147 | MW334278 | duck/Yunlin/15010115-2/2015 |
| 148 | MW334286 | duck/Yunlin/15010244-1/2015 |
| 149 | MW334294 | duck/Yunlin/15010277/2015 |
| 150 | MW334302 | duck/Yunlin/15010310-4/2015 |
| 151 | MW334310 | duck/Yunlin/15010318-2/2015 |
| 152 | MW334318 | duck/Yunlin/15010369-1/2015 |
| 153 | MW334326 | duck/Yunlin/15010458-1/2015 |
| 154 | MW334334 | duck/Yunlin/15010538/2015 |
| 155 | MW334342 | duck/Yunlin/15070002-2/2015 |
| 156 | MW334350 | duck/Yunlin/15A03403-1-20T/2015 |
| 157 | MW334358 | duck/Yunlin/15A04883-1-10T/2015 |
| 158 | MW334366 | duck/Yunlin/15A5019/2015 |
| 159 | MW334374 | duck/Yunlin/17110001/2017 |
| 160 | MW334382 | duck/Yunlin/17A0086/2017 |
| 161 | MW334390 | duck/Yunlin/17A0686/2017 |
| 162 | MW334398 | duck/Yunlin/18010004-2/2018 |
| 163 | MW334406 | duck/Yunlin/18060006-1/2018 |
| 164 | MW334414 | duck/Yunlin/18060007-3/2018 |
| 165 | MW334422 | duck/Yunlin/18X00007-3/2018 |
| 166 | MW334446 | goose/Changhua/15040008/2015 |
| 167 | MW334454 | goose/Chiayi/15010008-3/2015 |
| 168 | MW334462 | goose/Chiayi/15010016-1/2015 |
| 169 | MW334470 | goose/Chiayi/15020022/2015 |
| 170 | MW334478 | goose/Chiayi/16010017-2/2016 |
| 171 | MW334486 | goose/Chiayi/16040003-1/2016 |
| 172 | MW334494 | goose/Chiayi/16040005/2016 |
| 173 | MW334502 | goose/Chiayi/16040007-1/2016 |
| 174 | MW334510 | goose/Chiayi/17020027-1/2017 |
| 175 | MW334518 | goose/Chiayi/17040014/2017 |
| 176 | MW334526 | goose/Chiayi/18010006/2018 |
| 177 | MW334534 | goose/Chiayi/18020002-2/2018 |
| 178 | MW334566 | goose/Pingtung/15010007-4/2015 |
| 179 | MW334574 | goose/Pingtung/15010037-1/2015 |
| 180 | MW334582 | goose/Pingtung/15050014/2015 |
| 181 | MW334590 | goose/Pingtung/16010061/2016 |
| 182 | MW334598 | goose/Tainan/16010031-2/2016 |
| 183 | MW334606 | goose/Tainan/16020041-2/2016 |
| 184 | MW334614 | goose/Tainan/16040018-1/2016 |
| 185 | MW334622 | goose/Tainan/16070002/2016 |
| 186 | MW334630 | goose/Tainan/17030022-1/2017 |
| 187 | MW334638 | goose/Tainan/17030024-1/2017 |
| 188 | MW334646 | goose/Tainan/18010022-1/2018 |
| 189 | MW334654 | goose/Tainan/18050014-1/2018 |
| 190 | MW334662 | goose/Taoyuan/15010041/2015 |
| 191 | MW334670 | goose/Taoyuan/18120002-2/2018 |
| 192 | MW334678 | goose/Yunlin/15010005/2015 |
| 193 | MW334686 | goose/Yunlin/15010009-1/2015 |
| 194 | MW334694 | goose/Yunlin/15010010-2/2015 |
| 195 | MW334702 | goose/Yunlin/15010011-1/2015 |
| 196 | MW334710 | goose/Yunlin/15010029-1/2015 |
| 197 | MW334718 | goose/Yunlin/15040040/2015 |
| 198 | MW334726 | goose/Yunlin/16020008-4/2016 |
| 199 | MW334734 | goose/Yunlin/16020011-1/2016 |
| 200 | MW334742 | goose/Yunlin/16030042-2/2016 |
| 201 | MW334750 | goose/Yunlin/16040002-2/2016 |
| 202 | MW334758 | goose/Yunlin/17100005/2017 |
| 203 | MW334782 | goose/Yunlin/18030020-3/2018 |
| 204 | MW334790 | goose/Yunlin/18040002-3/2018 |
| 205 | MW334798 | goose/Yunlin/18040018-2/2018 |
| 206 | MW334806 | goose/Yunlin/18050003-2/2018 |

873

| | | |
|---|---|---|
| 207 | MW334814 | goose/Yunlin/18090002-3/2018 |
| 208 | MW334822 | goose/Yunlin/18090006-1/2018 |
| 209 | MW334830 | goose/Yunlin/18110002-3/2018 |
| 210 | MW334854 | turkey/Changhua/16010011-1/2016 |
| 211 | MW334862 | turkey/Chiayi/17020086-2/2017 |
| 212 | MW334870 | turkey/Chiayi/17020092-2/2017 |
| 213 | MW334878 | turkey/Chiayi/18010013-3/2018 |
| 214 | MW334886 | turkey/Yunlin/15030069-1/2015 |
| 215 | MW334894 | turkey/Yunlin/15060049-1/2015 |
| 216 | MW334902 | turkey/Yunlin/18020020-2/2018 |
| 217 | MW334910 | turkey/Yunlin/18070001-2/2018 |
| 218 | MW334918 | turkey/Yunlin/18090003-3/2018 |
| 219 | MW334926 | turkey/Yunlin/18110001-3/2018 |
| 220 | MW334934 | turkey/Yunlin/18120005-2/2018 |
| 221 | EPI588952 | goose/Taiwan/a015/2015 |
| 222 | EPI588960 | duck/Taiwan/a043/2015 |
| 223 | EPI588976 | duck/Taiwan/a068/2015 |
| 224 | KT388444 | goose/Taiwan/TNC1/2015 |
| 225 | KT388572 | goose/Taiwan/TNO3/2015 |
| 226 | KU646901 | goose/Taiwan/01038/2015 |
| 227 | KU646917 | duck/Taiwan/A3400/2015 |
| 228 | KU646925 | goose/Taiwan/01019/2015 |
| 229 | KU646933 | goose/Taiwan/01026/2015 |
| 230 | KU646941 | goose/Taiwan/01039/2015 |
| 231 | KU646861 | duck/Taiwan/01006/2015 |
| 232 | KU646869 | goose/Taiwan/01022/2015 |
| 233 | KU646877 | goose/Taiwan/01023/2015 |
| 234 | KU646885 | goose/Taiwan/01031/2015 |
| 235 | KU646893 | goose/Taiwan/01040/2015 |
| 236 | EPI961922 | chicken/Taiwan/u7/2016 |
| 237 | EPI961933 | chicken/Taiwan/x37/2016 |
| 238 | KP714479 | goose/Taiwan/01-003/2015 |
| 239 | KP714480 | goose/Taiwan/01-004/2015 |
| 240 | KP714481 | goose/Taiwan/01-042/2015 |
| 241 | MW334254 | duck/Tainan/16010024-1/2016 |
| 242 | MW333318 | chicken/Kaohsiung/16MB0889/2016 |
| 243 | MW333638 | chicken/Taipei/16060023-1/2016 |
| 244 | MW333670 | chicken/Taipei/17050007-2/2017 |
| 245 | MW333438 | chicken/New_Taipei_City/17090002/2017 |
| 246 | KU646909 | chicken/Taiwan/01174/2015 |
| 247 | MW334558 | goose/Chiayi/18020014-3/2018 |
| 248 | MW334766 | goose/Yunlin/18030020-1/2018 |
| 249 | MN988773 | chicken/Taiwan/A3/2019 |
| 250 | MN988805 | chicken/Taiwan/D9/2019 |
| 251 | KT388452 | goose/Taiwan/TNC2/2015 |
| 252 | KT388476 | goose/Taiwan/TNC5/2015 |

874
875