



Phylogeography of rubella virus in Asia: Vaccination and demography shape synchronous outbreaks

Brooke A. Bozick*, Colin J. Worby, C. Jessica E. Metcalf

Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, United States

ARTICLE INFO

Keywords:

Rubella virus
Phylogeography
Metapopulation
Human movement
Epidemic dynamics

ABSTRACT

Rubella virus causes mild disease in children but for women in the early stages of pregnancy, it can cause spontaneous abortion, congenital rubella syndrome (CRS) and associated birth defects. Despite the availability of an effective vaccine, rubella virus continues to circulate endemically in several regions of the world. This is particularly true in East and Southeast (E/SE) Asia, where control efforts vary widely among countries that are well connected through travel and immigration. It is therefore important to understand how the regional persistence of rubella is affected both by dynamics occurring across countries and susceptibility within countries. Here, we use genetic and epidemiological data from countries in E/SE Asia to explore the phylogeography of rubella virus in this region. Our results underline that metapopulation dynamics are key for rubella persistence and highlight the source-sink population structure of the region. We identify countries that contribute to the regional metapopulation network and link epidemic dynamics to susceptibility profiles within each country. Our results indicate that human movement plays an important role in driving epidemic dynamics in E/SE Asia.

1. Introduction

Rubella virus generally causes a mild childhood infection but can lead to severe birth defects when contracted by women in the early stages of pregnancy. A safe and inexpensive vaccine that confers complete immunity has been available since the 1970s (Plotkin, 2006); it is currently part of the routine immunization program in 159 countries worldwide (World Health Organization, 2017) and has considerably restricted the previously global reach of this pathogen. However, although the virus has been eliminated from the Americas and burdens have been drastically reduced in many developed countries, the WHO still estimates that over 100,000 babies/year are born with congenital rubella syndrome worldwide (Vynnycky et al., 2016; World Health Organization, 2011). Rubella continues to circulate endemically in many countries across both Africa and Asia (Mirambo et al., 2015; Zhu et al., 2015; Pham et al., 2013).

Human movement from endemic countries continues to cause outbreaks of rubella, even in settings where local transmission has been interrupted (Figueiredo et al., 2012; Vauloup-Fellous et al., 2010; Icenogle et al., 2006; Martínez-Torres et al., 2016). As global connectedness grows, it is increasingly apparent that an understanding of how human mobility drives the spread of infectious diseases is necessary to predict and control outbreaks of novel or reintroduced

pathogens, especially in a context of growing but incomplete vaccination coverage. The identification of genotypes and/or serologically different viral strains with divergent geographic distributions can be used to distinguish new introductions from ongoing but unobserved circulation (Icenogle et al., 2006; Rivaller et al., 2017). To our knowledge, just a single study has addressed global diffusion of rubella (Cloete, 2014); findings indicated that infection tends to flow out from regions of low vaccination coverage, education index and income rather than necessarily between regions in close geographic proximity. While this result is expected at broad global scales, it is not clear that these patterns will hold at finer spatial resolutions, particularly in regions of Asia where geographically adjacent countries that are strongly connected through travel and immigration have highly variable levels of vaccination coverage.

While countries such as Japan introduced the rubella vaccine to adolescent school girls as early as 1977 (Terada, 2003), populous countries like China and Vietnam only included it in their routine childhood immunization schedule in 2008 and 2015, respectively (Zhu et al., 2012; Gavi TVA, 2019). Theory (Bauch and Earn, 2003) and observation indicate that, for rubella, peaks in incidence occurring annually (Metcalf et al., 2011a; Metcalf et al., 2011b; Metcalf et al., 2013; Wesolowski et al., 2016) or every five years (Terada, 2003; Anderson and May, 1991) are expected in the absence of vaccination,

* Corresponding author.

E-mail address: bbozick@princeton.edu (B.A. Bozick).

<https://doi.org/10.1016/j.epidem.2019.100346>

Received 19 June 2018; Received in revised form 16 January 2019; Accepted 27 May 2019

Available online 28 May 2019

1755-4365/© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

and that erratic transient outbreaks should occur following incomplete introduction of the vaccine (Lessler and Metcalf, 2013; Winter et al., 2018). Data from East and Southeast (E/SE) Asia are incomplete but suggest that a mixture of these patterns is observed (supplementary Fig. 1). From 2011–2014, multiple countries in E/SE Asia, including those reporting high vaccine coverage rates, were affected by a large outbreak, suggesting a degree of regional synchronization. Our understanding of rubella persistence will be strengthened by characterizing the extent to which dynamics within countries are mediated by epidemics occurring in neighboring countries as well as outside the region. Furthermore, the dynamics of susceptibility play an integral role, as outbreaks can only occur where sufficient numbers of susceptible individuals are present. Inferring susceptibility is complex for pathogens with waning immunity, or complex cross-immunity (e.g. influenza), but as rubella infection and vaccination are completely immunizing, estimation of the size of the susceptible population becomes relatively tractable (Bjørnstad et al., 2002) with knowledge on the size of unvaccinated birth cohorts. Characterizing susceptible dynamics will improve our understanding of rubella persistence and may contribute to optimizing public health intervention measures.

Rubella is a single-stranded, positive-sense RNA virus of the *Togaviridae* family. RNA viruses are often amenable to phylogenetic analysis, as their evolutionary dynamics occur on the same time scale as their ecological dynamics (Grenfell et al., 2004). Although only one serotype exists, the CDC has identified two groups of rubella viruses consisting of ten and three genotypes, respectively (World Health Organization, 2013). These genotypes are immunologically identical in terms of both natural infection and vaccination. While the genome comprises approximately 10 kb, genotyping is based on a 739 base pair window of the E1 structural protein, which contains antigenic sites and other important functional domains necessary for cellular invasion (Banatvala and Peckham, 2007). Intraspecific genotypic variation is approximately 5% across genotypes of group 1 and 8% for group 2; maximum variation between groups 1 and 2 is 8–10%. Such genotype classifications are an essential component of virological surveillance and are used primarily to determine whether outbreaks are due to endemic circulation or importation events. Of those genotypes that are currently active, 1E, 1G, and 2B have achieved the broadest global distributions (Rivailler et al., 2017; Abernathy et al., 2011; Mulders et al., 2016).

In this study, we aimed to assess the importance of metapopulation dynamics to rubella persistence in E/SE Asia. We used publicly available genetic sequences to map pathways of regional spread from 2000 to 2016 and identify countries that act as sources and sinks of viral diffusion. We compared the dynamics of the two most commonly sampled genotypes in Asia, 1E and 2B (Mulders et al., 2016), and found that, despite being immunologically identical, their circulation patterns differ. Our results suggest that there are identifiable sources and sinks for rubella, but that the directionality of spatial spread may be primarily governed by stochasticity and founder effects. We additionally estimated country-specific susceptibility profiles to link temporal trends in rubella incidence and susceptible population size with the timing of viral importation and exportation events. Finally, we discuss limitations arising from analyzing spatially and/or temporally biased samples, a common challenge with studies of this scope.

2. Methods

2.1. Sequence data

Rubella virus sequences were downloaded from NCBI Genbank in September 2016 (National Center for Biotechnology Information, 2019). We limited our search to sequences from 2000 to 2016 for which a country and date of collection were available. We did not limit our search by demography, so adult, child and CRS cases were all potentially represented in the analysis; not all sequences could be classified

demographically due to the limited information available from Genbank. Sequences were aligned using the MUSCLE program in Geneious v5.6.7 (Kearse, 2012), and those that did not map to the 739bp window used for rubella genotype identification (World Health Organization, 2005) were discarded. For the majority of sequences, Genbank provided additional information on either the province or city from which the sample was collected. Separate datasets were created for genotypes 1E and 2B, and each dataset was randomly down-sampled so that no city or province was represented by more than one sequence per year in countries with greater than 50 sequences available (supplementary Table 1). In total, we retained 491 sequences matching our criteria that mapped to the 739bp identification window for genotype 1E and 743 sequences for genotype 2B. In the subsampled datasets, we included 226 sequences from genotype 1E and 250 sequences from genotype 2B. We created and analyzed three versions of each dataset to ensure that our results were robust to differences introduced by the random sampling of sequences (supplementary Table 2). The phylogenies from each of the three subsampled datasets were qualitatively similar for each genotype (supplementary Figs. 2–3).

Sequences collected from within Asia were geo-referenced by country, and sequences from outside Asia were collectively categorized as “Other”. Due to its wide geographic span and cultural ties to countries in Europe, we included sequences from Russia in the “Other” categorization. Sequences from Eurasia and the Middle East were also included in “Other”, as very few sequences were available from these regions.

2.2. Phylogenetic analysis

The software package BEAST v1.8.4 (Drummond et al., 2012; Ayres et al., 2012) was used to perform a phylogeographic analysis, from which we obtained the rate and magnitude of viral movement between locations. We used the HKY + I + G substitution model (Hasegawa et al., 1985; Yang, 1994) with a strict molecular clock, the Bayesian skyride coalescent prior (Minin et al., 2008) and an asymmetric, non-reversible, continuous time transition matrix (Ferreira and Suchard, 2008; Gong et al., 2013) for spatial locations using the Bayesian stochastic search variable selection (BSSVS) procedure (Lemey et al., 2009) to identify significant connections. For the BSSVS analysis, a 50% prior probability of identifying no well-supported connections is assumed. The expected numbers of transitions between location states and the timing of these events were calculated using a robust counting procedure (Minin and Suchard, 2008, 2007). For each dataset, three MCMC chains were run for 100 million generations, subsampled every 10,000 steps, and the first 1000 trees were discarded as burn-in. Convergence was assessed using the program Tracer v1.6, to ensure that all estimated parameters had effective sample sizes (ESS) greater than 200. In the case of runs for which ESS values were between 150–200, multiple chains were combined using LogCombiner. Maximum clade credibility trees for each of the sample datasets were produced using TreeAnnotator v1.8.4.

In order to identify well-supported connections between countries, Bayes factors were calculated in the spread3 program (Bielejec et al., 2016). As suggested in previous literature (Faria et al., 2013), the following cut-offs were used to classify support for each transition: greater than 100 for strong support, less than 100 but greater than 30 for good support, less than 30 but greater than 10 for support, less than 10 but greater than 3 for weak support and less than 3 for no support.

To identify predictors of viral spread within Asia, we selected one representative dataset for each genotype (sample 1; supplementary Table 2, supplementary Figs. 2–3), removed sequences collected outside E/SE Asia and used a phylogenetic generalized linear model (GLM) as implemented in BEAST v1.8.4 (Lemey et al., 2014). The phylogenetic GLM assesses support for various ecological and sampling factors predicted to influence epidemic diffusion patterns across the phylogeny. We selected 16 different variables encompassing geographic,

ecological, demographic, environmental and epidemiological differences between countries to include in the model (see supplementary material). To determine whether the set of variables driving diffusion over the entire 16-year period differed from that responsible for viral spread during the most recent epidemics, we performed the analysis on the dataset of sequences from E/SE Asia collected from 2000 to 2016, as well as the subset of these sequences collected during the height of the regional epidemics, which, according to our dataset, occurred from 2011 to 2015.

2.3. Reconstructing the profile of susceptibility

Susceptible populations in countries from which sequences were available were reconstructed using data on vaccination rates and birth rates to predict the size of the unvaccinated birth cohorts (see supplementary material). These were further depleted based on reported incidence using a previously developed framework for modelling natural infection (Takahashi et al., 2015). From this, we estimated the proportion of susceptible individuals within each country; this eases comparison of countries of such different sizes, but is also likely to be epidemiologically relevant as transmission of directly-transmitted childhood infections often scales in a frequency dependent fashion (Bjørnstad et al., 2002). Estimated coverage achieved in rubella vaccination programs, timing and age range of supplementary immunization activities (SIAs; wide age-range campaigns that occur at intervals of 2–4 years), and time-series of rubella incidence are publicly available from the World Health Organization (World Health Organization, 2019). Demographic data on population size, birth cohort size and birth rates are publicly available from the World Bank databank (World Bank, 2019).

3. Results

3.1. Genotype 1E circulation

The most recent common ancestor (MRCA) of all lineages currently circulating in Asia can be traced back to a common ancestor that existed in China in the late 1990s (mean: 1999.02; highest posterior density interval (HPDI): 1997.77–2000.08) (Fig. 1a). Two major, well-supported lineages containing sequences from Asia are evident: one that became dominant in China (Fig. 1a, clade 1) and one that spread to

multiple countries after a period of prolonged circulation in Taiwan (Fig. 1a, clade 2). The remaining clades (Fig. 1a, clades 3–9) are composed exclusively of sequences from outside Asia and are geographically segregated (Russia/Eastern Europe, Africa/Middle East, France). Genotype 1E therefore appears to persist in Asia without continual introductions from outside the region; viral exchange between Asia and the outer world was limited to movement from Asia to other continents.

Clade 1, a large clade of primarily Chinese sequences, corresponds to the previously identified Chinese 1E lineage (Zhu et al., 2015, 2012; Zhu et al., 2016). In line with previous estimates, we identified a divergence time of 2000.4 (HPDI: 1999.53–2001.04). Sequences from Japan, Taiwan, and outside Asia are interspersed throughout the clade but do not cluster together, signifying frequent but sporadic exports from China. The reconstructed ancestral location of four multi-national clusters lies in Taiwan.

Clade 2 contains sequences from multiple countries in and outside of Asia and is characterized by internal clusters with high posterior support (> 0.9) that are hierarchically organized by date and country; sequences from similar epidemic periods tend to cluster together and, within these clusters, group by country. This pattern is consistent with intermittent introductions over time. In contrast, finer subdivisions within the Chinese lineage (clade 1) are not well-supported, indicating continuous evolution and persistence of the Chinese 1E lineage throughout China. Clade 2 appears to have descended from an ancestor that originally circulated in China prior to late 2002 before spreading to Taiwan (mean: 2002.73; HPDI: 2001.52–2003.77). The lineage was detected continuously in Taiwan until around mid-2009, when it began to appear elsewhere in Asia.

Phylogeographic analysis of the circulation patterns of genotype 1E reinforce the importance of China and Taiwan. Within the region, Bayes factors (BF) indicate decisive support (BF > 100) for movement of the virus from China to Japan and Taiwan, and from Taiwan to Malaysia (Fig. 2a). Total viral exportations from China far exceeded importations (Fig. 2c). Numerous imports and exports in and out of Taiwan are tied to the country's significant connections to both East and Southeast Asia (supplementary Fig. 13). Additional analyses of datasets containing only sequences collected from recent cases in Taiwan (occurring from 2010 to 2015) confirmed that older sequences from Taiwan were not biasing the ancestral state characterization (see supplementary information). In contrast, there was no support for viral migration out of

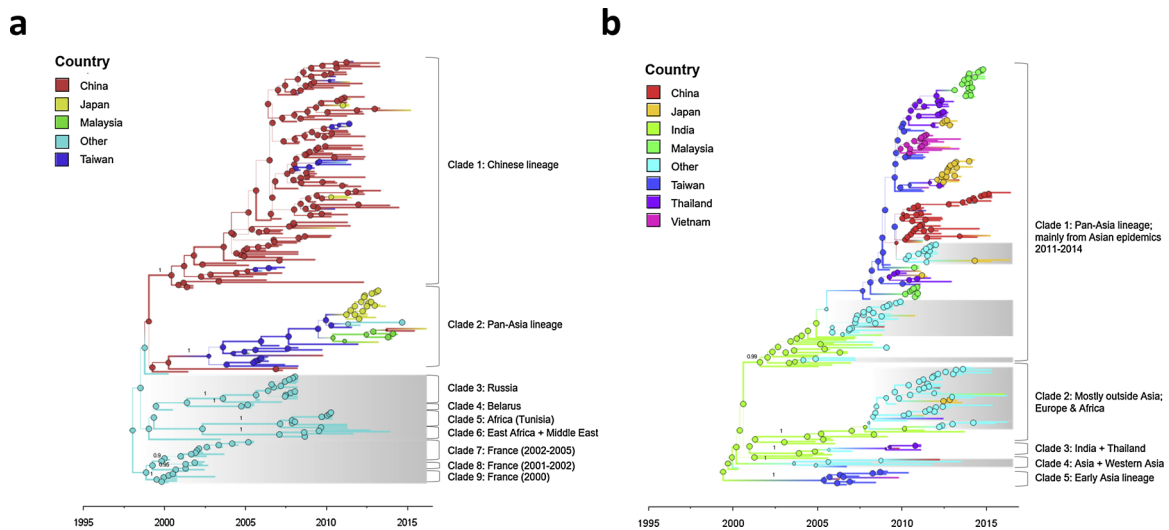


Fig. 1. Phylogenetic trees. Phylogenetic trees for a) genotype 1E and b) genotype 2B. Branch colors represent the most probable location state of the lineage at that time. Posterior support values indicated for each of the identified clades. X-axis represented in years. Branch widths sized according to posterior probability, where thicker branches indicate greater posterior support. Clades shaded in grey represent continuous transmission of a lineage outside of Asia. Interior nodes are colored based on their inferred geographic location and sized by the posterior support for this placement, where larger circles indicate greater posterior support.

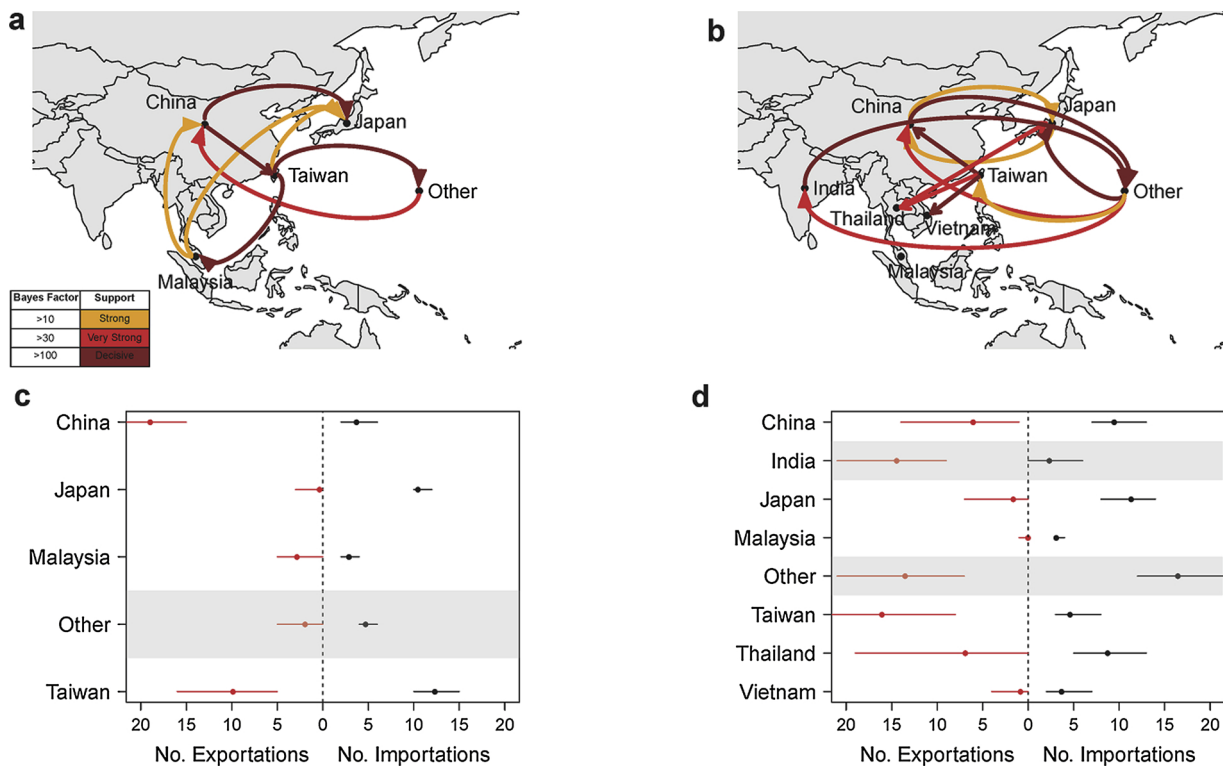


Fig. 2. Circulation patterns of rubella virus in Asia. Circulation maps of a) genotype 1E and b) genotype 2B. “Other” represents sequences collected from outside of Asia. Connections between countries are represented by lines colored by their Bayes factor (BF). Dark red (BF > 100) indicates decisive support for the connection, red indicates very strong support (100 > BF > 30), orange indicates strong support (30 > BF > 10). Connections with weak support (10 > BF > 3) are not mapped. Total estimated number of viral importations and exportations per country for c) genotype 1E and d) genotype 2B. Locations identified as being outside the metapopulation circulation network (i.e. no significant connections to the countries of E/SE Asia) are shaded in grey in panels c and d (For interpretation of the references to coloured in this figure legend, the reader is referred to the web version of this article).

Japan, where importations far exceeded expected exportations (Fig. 2c, supplementary Fig. 11). Of the 16 predictors tested in the phylogenetic GLM, there was weak support for geographic distance metrics and destination country sample size having an effect on viral diffusion when examining the entire time period (supplementary Fig. 6). However, none of our selected predictors were significant when considering only sequences collected during the recent epidemics from 2011 to 2015 (supplementary Fig. 7).

3.2. Genotype 2B circulation

The common ancestor of all sampled global lineages existed in India in the late 1990s (mean: 1999.3; HPDI 1998.33–2000.15) (Fig. 1b). However, this could be biased if few locations were sampled during this period; according to the World Health Organization, reporting of rubella cases in many countries did not start until 2005 (supplementary Fig. 1). Genotype 2B circulates widely throughout the region, as evidenced by the large multi-national clade containing many E/SE Asian sequences from the 2011–2014 epidemics (Fig. 1b, clade 1). Similar to the pan-Asia lineage of genotype 1E, sequences cluster by outbreak and country. Correspondingly, there is no evidence for long-term persistence of country-specific lineages. The existence of a distantly-related Asian clade predating the 2011–2014 epidemics (Fig. 1b, clade 5), and the probable placement outside of Asia of the ancestor of all E/SE Asian sequences in clade 1 (posterior: 0.65) suggests that the recent outbreaks were caused by a lineage newly introduced from outside the region, rather than by an E/SE Asian lineage that persisted across inter-epidemic periods. In total, we identified five well-supported clades, all of which contain sequences from within and outside of Asia. Further, there is strong support for the progenitors of each of these clades to have existed outside of E/SE Asia, indicating that historic circulation of

genotype 2B was maintained by introductions from outside the region.

Few direct transitions between India and E/SE Asia are observed, suggesting that viral exchange is minimal; with the exception of the early Asia lineage (clade 5), all groups of E/SE Asian sequences evolved from an ancestor predicted to have existed outside of Asia. However, sparse sampling from more recent outbreaks in India could potentially bias this conclusion as the majority of sequences from India were sampled from 2004–2008. As a populous country with low vaccination coverage (Dewan and Gupta, 2012), the existence of a sustained epidemic trough post-2008 is unlikely. However, the presence of two Indian sequences from 2011 and 2013 which cluster with others from outside of E/SE Asia supports the conclusion of limited direct viral exchange.

Phylogeographic analysis of the circulation patterns of genotype 2B again suggests an important regional role for Taiwan (Fig. 2b, supplementary Fig. 13). We find decisive support (BF > 100) for movement from Taiwan to both China and Southeast Asia. Additional analyses of datasets containing only sequences from recent cases in Taiwan (occurring from 2010 to 2015) show that exportations from Taiwan scale with the number of recent sequences included in the analysis, suggesting high genetic diversity among recently sampled Taiwan sequences rather than a collection of highly homogenous sequences sampled from a few related outbreaks (see supplementary material). Bidirectional exchange between China and Japan is also strongly supported (BF > 10) although few movements between these countries occurred. Again, far more importations than exportations are observed for Japan (Fig. 2d, supplementary Fig. 11), and significant movement from Thailand to Japan is apparent (Fig. 2b, supplementary Fig. 14). Viral exchange between India and countries outside of Asia is strongly supported (BF > 100), though we find no significant connections between India and E/SE Asia, suggesting that India does not play a role in

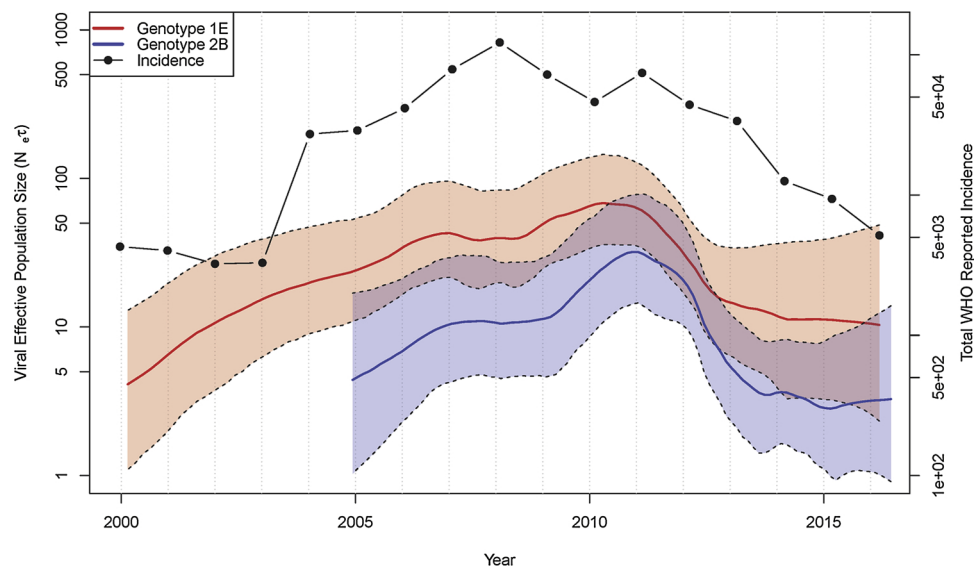


Fig. 3. Viral effective population sizes (N_e). Estimates of viral N_e in E/SE Asia for each genotype plotted with total reported cases across the six focal countries. Sparse sample collection precluded estimating dynamics for genotype 2B prior to 2005.

the E/SE Asia metapopulation. None of the 16 predictors included in the phylogenetic GLM were significantly correlated with viral diffusion when considering the entire period from 2000 to 2016 (supplementary Fig. 8). When considering only sequences collected during the recent epidemics from 2011 to 2015, geographic distance and destination country sample size were marginally significant (supplementary Fig. 9).

3.3. Viral demography in the context of epidemiological dynamics

WHO case data indicates that incidence increased in most of the focal countries between 2010 and 2015 (supplementary Fig. 1a). With the exception of Vietnam, all focal countries report coverage of rubella-containing vaccine (RCV1) to be above 90% by 2012 (supplementary Fig. 1b), suggesting a level of susceptibility well below the theoretical threshold necessary for elimination based on previous estimates of R_0 ranging from 3 to 8 (Metcalfe et al., 2011b) and a vaccination coverage threshold of 80% ($1-1/R_0$ for $R_0 \sim 5$). That rubella continues to circulate in many of these countries suggests uncertainty in reporting of coverage (Lessler et al., 2011) and/or heterogeneity underlying national coverage rates (Takahashi et al., 2017).

For each of the focal countries, accumulations of susceptible individuals precede each peak in reported cases (supplementary Fig. 1c). Apart from Thailand and Taiwan, heightened susceptibility between 2009 and 2014 was apparent. Although total proportion susceptible in Taiwan was consistently low, estimated accumulations of susceptible individuals occur around 1999 and 2006. In Thailand, the estimated proportion susceptible remained stable and relatively low throughout the entire period.

Rubella importation can only be successful if sufficient susceptible individuals are present. The timing of viral importation events into each country appear to precede epidemic peaks (supplementary Fig. 4) and tend to occur during local peaks in susceptible population size (supplementary Fig. 5). For example, imports of genotype 2B into China are observed around 2006 and 2009, the times during which susceptible population size peaked just prior to each major outbreak. As the epidemiological data suggests that the susceptible population size in China falls after 2005 (supplementary Fig. 1c), we might expect that there are fewer opportunities for the introduction of new strains. However, if erratic local dynamics cause extinctions in areas where not enough susceptibles are present, susceptibles throughout China will eventually accumulate enough to allow new introductions. In support, the phylogeny suggests that introductions continued to happen in China across

the time course of this study. Further, the small spike in introductions to China in 2009 coincides with an accumulation of susceptibles following the epidemic in 2008 even though the overall proportion susceptible is consistently decreasing (supplementary Fig. 5). As only 1–2 epidemics occurred during the time period of study in each of the focal countries, we were unable to statistically test this observation, but in the context of the previously described phylogenetic patterns it is suggestive of the importance of metapopulation processes in the epidemic dynamics of rubella in this region.

Demographic reconstruction of the viral effective population sizes (N_e) of each genotype within E/SE Asia reveals strikingly similar patterns: two periods of epidemic growth interspersed with an approximately two-year period of stasis (Fig. 3). Peaks are apparent for both genotypes around 2011, whereas growth phases are observed from 2005 to 2007 and from 2009–2011. While improved sampling could underlie this sharp increase, it also coincides with the period in which several Asian countries experienced rubella epidemics (supplementary Fig. 1a). Broad concordance between peaks in total incidence and N_e are observed, particularly for genotype 2B (Fig. 3).

4. Discussion

To understand whether metapopulation dynamics contribute to the persistence of rubella across neighboring countries that vary in their level of vaccination coverage, we combined both phylogenetic and epidemiological lines of evidence. Our analysis indicates that circulation patterns may differ between rubella virus genotypes 1E and 2B. As no evidence of a substantial competitive advantage for either subtype exists to date, these differences therefore suggest that the current distributions are likely a product of stochastic forces and founder effects stemming from the timing of their initial invasion. We confirmed the presence of a major lineage of genotype 1E evolving in China and found evidence for an additional 1E lineage circulating throughout Asia. Although some exchange with countries outside of the region was apparent, recent genotype 1E circulation in Asia appears isolated from the rest of the world. In contrast, we found evidence for multiple, recent introductions of genotype 2B from outside of Asia, suggesting continual global exchange. There was a clear increase in the effective population size of each genotype just prior to the 2011–2014 epidemics in Asia, suggesting that simultaneous outbreaks of both occurred. While China serves as a regional source population, particularly for genotype 1E, Taiwan likely acts as an important indicator for regional rubella activity

across East and Southeast Asia. In line with high reported levels of vaccination coverage, we find that Japan is a regional sink. Further, India does not have strong epidemiological links to E/SE Asia.

While other studies have looked at global rubella circulation, our study focuses on regional spread in an area of endemic transmission. Our results suggest that metapopulation dynamics are integral to the persistence of the virus, as lineages tend to circulate continuously across the region without a strong signal of spatial isolation within countries between epidemics. Although clear source and sink populations were evident for genotype 1E and, to some extent, genotype 2B, we also identified widely distributed lineages that were introduced from outside Asia. However, we were not able to identify any strong predictors of viral diffusion throughout the region for either of the genotypes we studied. Although geographic proximity appeared to be weakly correlated with viral diffusion, this relationship was erratic among the two genotypes. This finding was likely spurious as distance is probably a poor measure for the connections that are relevant to rubella spread. Our finding that viral persistence is maintained through a complex, lineage-specific circulation network is similar to that described for other childhood diseases in the E/SE Asia region, notably enterovirus A71 (Thao et al., 2017). Limited viral exchange occurred between India and E/SE Asia over the period of our study, an observation likely due to cultural differences across the Asian continent. Owing to historical events, India is more strongly connected through travel and immigration with countries in Europe, North America and, more recently, the Middle East (Naujoks, 2009; United Nations, 2016). As very few sequences are available from the Middle East and other countries in Western Asia and the Pacific where large populations of Indian immigrants reside, we cannot say at this time whether these connections are relevant to the spread of rubella. It presents an interesting avenue for future research, in particular because of observed differences in epidemic dynamics between India and the Asia-Pacific region for other infectious diseases (Pons-salort et al., 2015). Characterizing patterns such as the limited relevance of imports from India for re-introduction of rubella in E/SE Asia will become increasingly important as the region approaches rubella elimination.

Given the introduction of rubella-containing vaccine to all children in Japan in 1989, Japan's role as a sink for viral importation is unsurprising. The outbreaks represented in this study occurred primarily among adult men that were not vaccinated during childhood (Tanaka-Taya et al., 2013), as initial vaccination campaigns in Japan starting in 1977 only targeted adolescent schoolgirls (Terada, 2003). Though outbreaks in Japan continue to occur, the introduction of the MMR vaccine in 1989 and the policy switch towards the targeting of infants and young children has resulted in epidemics of decreased magnitude (Terada, 2003) with an increased average age of infection (Sugishita et al., 2015; Kinoshita and Nishiura, 2016). This, coupled with an increasing share of foreign residents immigrating from Southeast Asia (Green, 2017), make Japan a prime location for the importation of viral lineages, an observation which, paradoxically, is a metric indicating successful control. It remains to be seen how rubella dynamics throughout the region will affect Japan as it continues to near elimination.

While China's importance is easily explained by its large population size, relatively recent introduction of routine rubella vaccination, and prevalence of travel connections, our finding of Taiwan's role as a significant disseminator of infection (also suggested previously (Cloete, 2014)) is less obvious. Rubella is a notifiable disease in Taiwan and a subset of sequences collected from 2005 to 2013 have been well characterized epidemiologically. Many of these sequences were traced to importation events from Vietnam, and to a lesser extent Malaysia, Thailand and China, while other outbreaks that were untraceable in origin occurred within immigrant communities (Cheng et al., 2013). Although our phylogenetic analysis identifies Taiwan as an ancestral location, an alternative explanation is that the high intensity of surveillance could lead to outbreaks being identified earlier here than in

the countries from which they originated. Under this hypothesis and coupled with reported vaccination rates of greater than 90% since 2000, Taiwan is likely not an actual source of infection; instead, Taiwan probably acts as a sentinel for the rest of the region.

Reinforcing this, in addition to its well-established surveillance system which facilitates viral detection and characterization, there are multiple factors that predispose Taiwan to rubella importations. Past serological studies show that seronegativity is higher among immigrant women (Lin et al., 2010; Wang et al., 2007) and until 2002, Taiwan did not require proof of rubella vaccination upon immigration (Lin et al., 2010; Taiwan Centers for Disease Control, 2019). Further, an increasing proportion of immigrants now come from nations in Southeast Asia, where endemic rubella transmission is still widespread and vaccination has only recently been introduced (Lin et al., 2010; Wang et al., 2007; Tsay, 2015). In addition, from 2006 to 2010, 15–28% of Taiwanese men were married to foreign-born women, primarily from Southeast Asia (Lin et al., 2010). Finally, Taiwan has recently experienced an increase in international air travel. In 2014, 9.9 and 11.8 million people, respectively, travelled in to and out of the country and an estimated 84–88% of these travelers either visited or visited from countries in E/SE Asia (Shih et al., 2016). While these patterns are unlikely to outweigh high vaccination coverage and result in Taiwan contributing importantly to regional dynamics, such high rates of immigration and international travel, coupled with its strong geopolitical ties to China, a clear source of rubella infection, makes Taiwan an ideal sentinel for the identification of new lineages in the region.

The observed simultaneous increases in the effective population size of both genotypes is striking given that rubella virus immunity is completely cross-protective. Previous work has shown that the interpretation of viral effective population size estimates drawn from Bayesian coalescent skyline techniques should focus on changes over time, as observed increases in N_e correlate best with increases in incidence at the start of an epidemic (Frost and Volz, 2010). The synchronized phases of exponential viral population growth (2005–2007 and 2009–2011) that coincide with the two epidemic periods observed in the WHO case reports are evidence that both genotypes simultaneously circulated and contributed to epidemics in the region. This is surprising for a fully cross-immunizing infection, as it suggests that the entire E/SE Asia region is nearly in phase. While we would not expect two consecutive years of viral population growth during any single rubella epidemic in a given location, closely synchronized outbreaks across multiple countries could appear as consecutive years of growth when taken together. Two scenarios may be invoked as an explanation for epidemic synchronization across the region.

First, a country or city might act as a major pulse driver, akin to that of London for pre-vaccination measles epidemics in the United Kingdom (Grenfell et al., 2001). However, this scenario would require that only one major location play a key role and that sink populations are only weakly connected such that rescue effects after epidemic extinction are rare. These conditions are unlikely, given that many cities are above the critical community size for maintaining persistent rubella circulation (Metcalfe et al., 2013) and that high international connectivity in E/SE Asia has been cited as a possible explanation for the persistence of other viruses (Russell et al., 2008). However, we find strong evidence that China functions as a source population for genotype 1E. Given that we also did not find support for repeated introductions from outside Asia for this genotype, there is a possibility that epidemiological dynamics in China drive the regional spread of genotype 1E.

The second scenario we envision to explain synchronized phases of viral growth emerges from extremely high regional connectivity. If the existence of a connection, rather than the magnitude of movement across that connection, is the chief predictor of outbreak timing, we might expect that an epidemic arising in one location will immediately ignite outbreaks in all connected populations whose susceptible population size is above the threshold necessary for a sustained epidemic to

occur. Comparable seasonal travel patterns across settings (e.g. driven by similar timing of school holidays (Wesolowski et al., 2017)) will reinforce this alignment of timing, creating a positive feedback loop to further synchronize dynamics. Sporadic introductions from weakly connected locations that are not part of the core circulation network would then reseed epidemics once the pool of susceptibles has again reached an appropriate size. This scenario is consistent with our study, in that (1) the magnitude of connectivity measured by travel flows was not predictive of viral diffusion, (2) extensive travel and immigration ties exist between countries throughout the region and (3) repeated introductions from outside E/SE Asia were observed for genotype 2B.

An inherent limitation of using publicly available data is the inevitability of skewed and potentially biased sampling. Sequences sourced from Genbank originate from multiple studies with different methodologies and different motivations for collection. As a result, the quantity of data publicly available from any single location or time period varies. The number of available sequences from Japan, Taiwan and China, where surveillance is more extensive, far exceeds that from countries in Southeast and South Asia, where rubella is often not a notifiable disease. Further, even in the most well sampled locations, sampling biases exist over both space and time. Despite our efforts to correct for this, we cannot exclude the possibility that additional, unsampled lineages circulate, or have circulated, in Southeast Asia and India where sampling is sparse. Further, our analysis does not account for viral exchange with countries from which no samples are present. The possibility therefore remains that other countries may also act as sources. Broader sampling throughout Asia, both in terms of locations represented and the quantity of sequences available, as well across the globe, particularly from rubella-endemic countries that lack access to vaccination, will further refine our understanding of rubella circulation and persistence. Literature on the historical genotypic characterization of global rubella viruses indicates that many subtypes and lineages have gone extinct from 1960 through to the present (World Health Organization, 2013) and it remains to be seen whether, for example, the introduction of genotype 2B into China will lead to local extinction of genotype 1E. However, as sampling is not consistent or complete across this region, it is difficult to say with any certainty whether lineages have actually gone extinct, are still present at low levels, or are circulating in countries with minimal available data. Our results suggest that rubella lineages are not confined to specific countries and therefore that surveillance and elimination efforts must take into account the epidemiology of the region as a whole and population movements between countries in order to be successful.

Ethics

All genetic data were publicly available and anonymized; thus, this study did not require approval by an ethics committee or institutional review board.

Data, code and materials

Genbank accession numbers for all sequences used in this study can be found in supplementary table 2.

Competing interests

The authors declare no competing interests.

Authors' contributions

B.A.B and C.J.E.M. conceived the analysis. B.A.B., C.J.E.M and C.J.W. designed the analysis. B.A.B. performed the analysis. B.A.B., C.J.E.M. and C.J.W. wrote the paper.

Funding

This work was supported by a grant from the Bill and Melinda Gates Foundation (CJEM, BAB; OPP1094793).

Acknowledgements

We acknowledge B.T. Grenfell, S. Takahashi, A. Winter and L. Salvador for thoughtful comments and discussion of the manuscript.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:<https://doi.org/10.1016/j.epidem.2019.100346>.

References

- Abernathy, E.S., et al., 2011. Status of global virologic surveillance for rubella viruses. *J. Infect. Dis.* 204, S524–S532. <https://doi.org/10.1093/infdis/jir099>.
- Anderson, R.M., May, R.M., 1991. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.
- Ayres, D.L., et al., 2012. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* 61, 170–173. <https://doi.org/10.1093/sysbio/syr100>.
- Banatvala, J.E., Peckham, C.A., 2007. *Rubella Viruses*. Elsevier B.V.
- Bauch, C.T., Earn, D.J.D., 2003. Transients and attractors in epidemics. *Proc. R. Soc. B Biol. Sci.* 270, 1573–1578. <https://doi.org/10.1098/rspb.2003.2410>.
- Bielejec, F., Baele, G., Vrancken, B., Suchard, M.A., Rambaut, A., Lemey, P., 2016. Spread3: interactive visualization of spatiotemporal history and trait evolutionary processes. *Mol. Biol. Evol.* 33, 2167–2169. <https://doi.org/10.1093/molbev/msw082>.
- Bjørnstad, O.N., Finkenstädt, B.F., Grenfell, B.T., 2002. Dynamics of Measles epidemics: estimating scaling of transmission rates using a time series SIR model. *Ecol. Monogr.* 72, 169–184. [https://doi.org/10.1890/0012-9615\(2002\)072\[0169:DOMEES\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2002)072[0169:DOMEES]2.0.CO;2).
- Cheng, W.-Y., Wang, H.-C., Liu, M.-T., Wu, H.-S., 2013. Molecular surveillance of rubella viruses in Taiwan from 2005 to 2011. *J. Med. Virol.* 85, 745–753. <https://doi.org/10.1002/jmv.23451>.
- Cloete, L.J., 2014. *The Molecular Evolution and Epidemiology of Rubella Virus*. University of the Western Cape.
- Green, David, 2017. As Its Population Ages, Japan Quietly Turns to Immigration. <https://www.migrationpolicy.org/article/its-population-ages-japan-quietly-turns-immigration>.
- Dewan, P., Gupta, P., 2012. Burden of congenital rubella syndrome (CRS) in India: a systematic review. *Indian Pediatr.* 49, 377–399. <https://doi.org/10.1007/s13312-012-0087-4>.
- Drummond, A.J., Suchard, M.A., Xie, D., Rambaut, A., 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. <https://doi.org/10.1093/molbev/mss075>.
- Faria, N.R., Suchard, M.A., Rambaut, A., Streicker, D.G., Lemey, P., 2013. Simultaneously reconstructing viral cross-species transmission history and identifying the underlying constraints. *Philos. Trans. R. Soc. B Biol. Sci.* 368, 20120196. <https://doi.org/10.1098/rstb.2012.0196>.
- Ferreira, M.A.R., Suchard, M.A., 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. *Can. J. Stat.* 36, 355–368. <https://doi.org/10.1002/cjs.5550360302>.
- Figueiredo, C.A., Oliveira, M.L., Curti, S.P., Afonso, A.M.S., Yu AL, Frugi, Gualberto, F., Durigon, E.L., 2012. Phylogenetic analysis of rubella virus strains during an outbreak in São Paulo, 2007–2008. *J. Med. Virol.* 84, 1666–1671. <https://doi.org/10.1002/jmv.23382>.
- Frost, S.D.W., Volz, E.M., 2010. Viral phylodynamics and the search for an 'effective number of infections'. *Philos. Trans. R. Soc. B Biol. Sci.* 365, 1879–1890. <https://doi.org/10.1098/rstb.2010.0060>.
- Gavi TVA. Gavi country factsheet : Vietnam. <http://www.gavi.org/country/vietnam/> (accessed on 26 October 2017).
- Gong, L.L., Suchard, M.A., Bloom, J.D., 2013. Stability-mediated epistasis constrains the evolution of an influenza protein. *Elife* 2, e00631. <https://doi.org/10.7554/eLife.00631>.
- Grenfell, B.T., Bjørnstad, O.N., Kappey, J., 2001. Travelling waves and spatial hierarchies in measles epidemics. *Nature* 414, 716–723. <https://doi.org/10.1038/414716a>.
- Grenfell, B.T., Pybus, O.G., Gog, J.R., Wood, J.L.N., Daly, J.M., Mumford, J.A., Holmes, E.C., 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 303 <https://doi.org/10.1126/science.1090727>. 327 LP-332.
- Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* 22, 160–174.
- Icenogle, J.P., Frey, T.K., Abernathy, E., Reef, S.E., Schnurr, D., Stewart, J.A., 2006. Genetic analysis of rubella viruses found in the United States between 1966 and 2004: evidence that indigenous rubella viruses have been eliminated. *Clin. Infect. Dis.* 43, S133–S140. <https://doi.org/10.1086/505945>.
- Kearse, M., 2012. *Geneious*.

- Kinoshita, R., Nishiura, H., 2016. Assessing herd immunity against rubella in Japan: a retrospective seroepidemiological analysis of age-dependent transmission dynamics. *BMJ Open* 6, e009928. <https://doi.org/10.1136/bmjopen-2015-009928>.
- Lemey, P., Rambaut, A., Drummond, A.J., Suchard, M.A., 2009. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* 5, e1000520. <https://doi.org/10.1371/journal.pcbi.1000520>.
- Lemey, P., et al., 2014. Unifying viral genetics and human transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Pathog.* 10, e1003932. <https://doi.org/10.1371/journal.ppat.1003932>.
- Lessler, J., Metcalf, C.J.E., 2013. Balancing evidence and uncertainty when considering rubella vaccine introduction. *PLoS One* 8. <https://doi.org/10.1371/journal.pone.0067639>.
- Lessler, J., Metcalf, C.J.E., Grais, R.F., Luquero, F.J., Cummings, D.A.T., Grenfell, B.T., J.E., 2016. Genetic characterization of rubella virus strains detected in Spain, 1998–2014. *PLoS One* 11, e0162403. <https://doi.org/10.1371/journal.pone.0162403>.
- Metcalf, C.J.E., Bjørnstad, O.N., Ferrari, M.J., Klepac, P., Bharti, N., Lopez-Gatell, H., Grenfell, B.T., 2011a. The epidemiology of rubella in Mexico: seasonality, stochasticity and regional variation. *Epidemiol. Infect.* 139, 1029–1038. <https://doi.org/10.1017/S0950268810002165>.
- Metcalf, C.J.E., Munayco, C.V., Chowell, G., Grenfell, B.T., Bjørnstad, O.N., 2011b. Rubella metapopulation dynamics and importance of spatial coupling to the risk of congenital rubella syndrome in Peru. *J. R. Soc. Interface* 8, 369–376. <https://doi.org/10.1098/rsif.2010.0320>.
- Metcalf, C.J.E., et al., 2013. Implications of spatially heterogeneous vaccination coverage for the risk of congenital rubella syndrome in South Africa. *J. R. Soc. Interface* 10, 20120756. <https://doi.org/10.1098/rsif.2012.0756>.
- Minin, V.N., Suchard, M.A., 2007. Counting labeled transitions in continuous-time Markov models of evolution. *J. Math. Biol.* 56, 391–412. <https://doi.org/10.1007/s00285-007-0120-8>.
- Minin, V.N., Suchard, M.A., 2008. Fast, accurate and simulation-free stochastic mapping. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 3985–3995. <https://doi.org/10.1098/rstb.2008.0176>.
- Minin, V.N., Bloomquist, E.W., Suchard, M.A., 2008. Smooth skyride through a rough skyline: bayesian coalescent-based inference of population dynamics. *Mol. Biol. Evol.* 25, 1459–1471. <https://doi.org/10.1093/molbev/msn090>.
- Mirambo, M.M., Majigo, M., Aboud, S., Groß, U., Mshana, S.E., 2015. Serological markers of rubella infection in Africa in the pre vaccination era: a systematic review. *BMC Res. Notes* 8, 716. <https://doi.org/10.1186/s13104-015-1711-x>.
- Mulders, M.N., et al., 2016. Global measles and rubella laboratory network support for elimination goals, 2010–2015. *MMWR Morb. Mortal. Wkly. Rep.* 65, 438–442. <https://doi.org/10.15585/mmwr.mm6517a3>.
- National Center for Biotechnology Information, <https://www.ncbi.nlm.nih.gov/> (accessed on 21 September 2016).
- Naujoks, D., 2009. Emigration, immigration, and diaspora relations in India. *Migr. Inf. Source* 1–18.
- Pham, V.H., Van, Nguyen T., Nguyen, T.T.T., Dang, L.D., Hoang, N.H., Van, Nguyen T., Abe, K., 2013. Rubella epidemic in Vietnam: Characteristic of rubella virus genes from pregnant women and their fetuses/newborns with congenital rubella syndrome. *J. Clin. Virol.* 57, 152–156. <https://doi.org/10.1016/j.jcv.2013.02.008>.
- Plotkin, S.A., 2006. The history of rubella and rubella vaccination leading to elimination. *Clin. Infect. Dis.* 43, S164–S168. <https://doi.org/10.1086/505950>.
- Pons-salort, M., Parker, E.P.K., Grassly, N.C., 2015. The epidemiology of non-polio enteroviruses: recent advances and outstanding questions. *Curr. Opin. Infect. Dis.* 28, 479–487. <https://doi.org/10.1097/qco.0000000000000187>.
- Rivaille, P., Abernathy, E., Icenogle, J., 2017. Genetic diversity of currently circulating rubella viruses: a need to define more precise viral groups. *J. Gen. Virol.* <https://doi.org/10.1099/jgv.0.000680>.
- Russell, C.A., et al., 2008. The global circulation of seasonal influenza A (H3N2) viruses. *Science* (80-) 320, 340–346. <https://doi.org/10.1126/science.1154137>.
- Shih, C.T., Chang, Y.C., Wang, H.L., Lin, C.C., 2016. Comparing the rubella seronegativity in pregnant women who received one dose of rubella vaccine at different ages in Taiwan. *Vaccine* 34, 4787–4791. <https://doi.org/10.1016/j.vaccine.2016.08.024>.
- Sugishita, Y., Shimatani, N., Katow, S., Takahashi, T., Hori, N., 2015. Epidemiological characteristics of rubella and congenital rubella syndrome in the 2012–2013 epidemics in Tokyo, Japan. *Jpn. J. Infect. Dis.* 68, 159–165. <https://doi.org/10.7883/yoken.JJID.2014.195>.
- Taiwan Centers for Disease Control, <http://www.cdc.gov.tw/english/info.aspx?treeid=E79C7A9E1E9B1CDF&nowtreeid=E02C24F0DACDD729&tid=C07C33CE1061D072> (accessed on 25 October 2017).
- Takahashi, S., Metcalf, C.J.E., Ferrari, M.J., Moss, W.J., Truelove, S.A., Tatem, A.J., Grenfell, B.T., Lessler, J., 2015. Reduced vaccination and the risk of measles and other childhood infections post-Ebola. *Science* 347, 1240–1242. <https://doi.org/10.1126/science.aaa3438>.
- Takahashi, S., Metcalf, C.J.E., Ferrari, M.J., Tatem, A.J., Lessler, J., 2017. The geography of measles vaccination in the African Great Lakes region. *Nat. Commun.* 8, 15585. <https://doi.org/10.1038/ncomms15585>.
- Tanaka-Taya, Keiko, Satoh, Hiroshi, Arai, Satoru, Yamagishi, Takuya, Yahata, Yuichiro, Nakashima, Kazutoshi, Sugawara, Tamie, Ohkusa, Yasushi, Matsui, Tamano, Saito, Takehito, Kanou, Kazuhiko, Shimada, Tomoe, Kinoshita, Hitomi, Yamashita, Kazuyo, Yasui, Yoshinori, Yuki, K.O., 2013. Nationwide rubella epidemic - Japan, 2013. *MMWR Morb. Mortal. Wkly. Rep.* 62, 457–462.
- Terada, K., 2003. Rubella and congenital rubella syndrome in Japan: epidemiological problems. *Jpn. J. Infect. Dis.* 56, 81–87.
- Thao, N.T.T., et al., 2017. Evolution and spatiotemporal dynamics of enterovirus A71 subgenogroups in Vietnam. *J. Infect. Dis.* 216, 1371–1379. <https://doi.org/10.1093/infdis/jix500>.
- Tsay, C.-L., 2015. Migration between Southeast Asia and Taiwan: trends, characteristics and implications. *J. ASEAN Stud.* 3, 68–92.
- United Nations, 2016. *International Migration Report 2015* (doi:ST/ESA/SER.A/384).
- Vauloup-Fellous, C., Hübschen, J.M., Abernathy, E.S., Icenogle, J., Gaidot, N., Dubreuil, P., Parent-du-Châtelet, I., Grangeot-Keros, L., Müller, C.P., 2010. Phylogenetic analysis of rubella viruses involved in congenital rubella infections in France between 1995 and 2009. *J. Clin. Microbiol.* 48, 2530–2535. <https://doi.org/10.1128/JCM.00181-10>.
- Vynnycky, E., et al., 2016. Using seroprevalence and immunisation coverage data to estimate the global burden of congenital rubella syndrome, 1996–2010: a systematic review. *PLoS One* 11, e0149160. <https://doi.org/10.1371/journal.pone.0149160>.
- Wang, J.J., Huang, L.M., Chen, H.H., Hwang, K.C., Chen, C.J., 2007. Seroprevalence of rubella infection after national immunization program in Taiwan: vaccination status and immigration impact. *J. Med. Virol.* 79, 97–103. <https://doi.org/10.1002/jmv.20764>.
- Wesolowski, A., et al., 2016. Introduction of rubella-containing-vaccine to Madagascar: implications for roll-out and local elimination. *J. R. Soc. Interface* 13, 20151101. <https://doi.org/10.1098/rsif.2015.1101>.
- Wesolowski, A., et al., 2017. Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. *Nat. Commun.* 8, 2069. <https://doi.org/10.1038/s41467-017-02064-4>.
- Winter, A.K., Pramanik, S., Lessler, J., Ferrari, M., Grenfell, B.T., Metcalf, C.J.E., 2018. Rubella Vaccination in India: Identifying broad consequences of vaccine introduction and key knowledge gaps. *Epidemiol. Infect.* 146, 65–77.
- World Bank. World Bank Open Data. <https://data.worldbank.org/> (accessed on 25 October 2017).
- World Health Organization, 2005. Standardization of the nomenclature for genetic characteristics of wild-type rubella viruses. *Wkly. Epidemiol. Rec.* 80, 126–132.
- World Health Organization, 2011. Rubella vaccines: WHO position paper. *Wkly. Epidemiol. Rec.* 86, 301–316. <https://doi.org/10.1016/j.vaccine.2011.08.061>.
- World Health Organization, 2013. Rubella virus nomenclature update: 2013. *Wkly. Epidemiol. Rec.* 32, 337–348.
- World Health Organization, 2017. Data, Statistics and Graphics. http://www.who.int/immunization/monitoring_surveillance/data/en/ (accessed on 26 October 2017).
- World Health Organization, Rubella - Reported cases by country. <http://apps.who.int/gho/data/node.main.WHS3.57?lang=en> (accessed on 1 September 2017).
- Yang, Z., 1994. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. *J. Mol. Evol.* 39, 306–314.
- Zhu, Z., et al., 2012. Emergence and continuous evolution of genotype 1E rubella viruses in China. *J. Clin. Microbiol.* 50, 353–363. <https://doi.org/10.1128/JCM.01264-11>.
- Zhu, Z., et al., 2015. Evolutionary analysis of rubella viruses in mainland China during 2010–2012: endemic circulation of genotype 1E and introductions of genotype 2B. *Sci. Rep.* 5, 7999. <https://doi.org/10.1038/srep07999>.
- Zhu, Z., et al., 2016. Analysis of complete genomes of the rubella virus genotypes 1E and 2B which circulated in China, 2000–2013. *Nature*. <https://doi.org/10.1038/srep39025>.